

Bachelorarbeit
im Bachelorstudiengang
Information Management Automotive
an der Hochschule für angewandte Wissenschaften Neu-Ulm

Thema

Mobility Process Mining
Eine Untersuchung der Auswirkungen der Anwendung des Process Mining Ansatzes auf Mobilitätsprozesse

Erstkorrektor: Prof. Dr. Joerg-Oliver Vogt

Verfasserin: Christine Kuck (Matrikel-Nr.: 259241)

Thema erhalten: 01.06.2022

Arbeit abgegeben: 26.09.2022

LOSE BEILAGEN

Digitaler Anhang

Inhaltsverzeichnis

Abbildungsverzeichnis.....	1
Tabellenverzeichnis	7
Abkürzungsverzeichnis.....	8
1. Einleitung.....	9
1.1 Problemstellung	10
1.2 Ziel der Arbeit.....	10
1.3 Vorgehensweise	11
1.4 Aufbau der Arbeit	12
2. Theoretische Grundlagen.....	14
2.1 Process Mining	14
2.1.1 Was ist Process Mining?	14
2.1.2 IEEE Task Force on Process Mining	15
2.1.3 Begriffsklärung	15
2.1.3.1 Event-Log	15
2.1.3.2 CSV-Format.....	17
2.1.3.3 XES Standard.....	17
2.1.3.4 Prozessmodell	17
2.1.4 Arten des Process Mining.....	18
2.1.4.1 Process Discovery	18
2.1.4.2 Conformance Checking	19
2.1.4.3 Model Enhancement	20
2.1.4.4 Operational Support.....	21
2.1.5 Tools	22
2.1.6 Einsatzgebiete von Process Mining in der Theorie und der Realität.....	23
2.1.6.1 Erkennung realer Prozesse.....	23

2.1.6.2 Prozessoptimierung.....	23
2.1.6.3 Prozessmanagement.....	24
2.1.6.4 Unterstützung entstehender und etablierter Unternehmen.....	24
2.1.6.5 reale Anwendungsfälle von Process Mining	25
2.2 Mobilität.....	27
2.2.1 Räumliche Mobilität	27
2.2.2 Was sind Mobilitätsprozesse?.....	28
2.2.3 Mobilitätsprozesse als Bestandteil von ‚Mobility as a Process‘	30
2.3 Anwendung von Process Mining auf Mobilitätsprozesse.....	30
2.3.1 Anwendung von Process Mining auf Mobilitätsprozesse in der Literatur	31
2.3.1.1 Untersuchung der relevanten Literatur	32
2.3.1.2 Schlussfolgerung der Literaturrecherche	33
2.3.2 Aufzeichnung von Mobilitätsdaten durch Google Maps.....	33
2.3.3 Definition von Kennzahlen zur Bewertung von Mobilitätsprozessen	34
2.3.4 Hypothesen zur Bewertung von Mobilitätsprozessen	35
3. Methodik.....	37
3.1 Erhebung der Mobilitätsdaten.....	37
3.2 Import der Event-Logs in das Tool ‚Disco‘	42
3.3 Vorgehen bei der Auswertung der Event-Logs.....	44
3.4 Erfüllung der Gütekriterien der Forschung.....	47
4. Ergebnisse.....	48
4.1 Überblick über das Process Mining Tool ‚Disco‘	48
4.2 Benötigte Informationen	55
4.3 Fall 1	56
4.3.1 Allgemeine Informationen	56
4.3.2 Prozessmodell	56
4.3.3 Dauer und zurückgelegte Entfernung der Cases.....	57

4.3.4 Durchschnittliche Dauer und Ereignisse pro Variante	59
4.3.5 Mobilitätsprozesse	59
4.3.6 Berechnung der Prozesskennzahl V	60
4.3.7 Berechnung der Prozesskennzahl G	61
4.3.8 Berechnung der Prozesskennzahl D	62
4.3.9 Filter: Kürzeste Cases	63
4.3.10 Filter: Kürzer als Durchschnitt	64
4.3.11 Filter: Cases mit Wartezeiten.....	64
4.4 Fall 2	64
4.4.1 Allgemeine Informationen	64
4.4.2 Prozessmodell	64
4.4.3 Dauer und zurückgelegte Entfernung der Cases.....	65
4.4.4 Durchschnittliche Dauer und Ereignisse pro Variante	67
4.4.5 Mobilitätsprozesse	67
4.4.6 Berechnung der Prozesskennzahl V	68
4.4.7 Berechnung der Prozesskennzahl G	69
4.4.8 Berechnung der Prozesskennzahl D	70
4.4.9 Filter: Kürzeste Cases	71
4.4.10 Filter: Kürzer als Durchschnitt	72
4.4.11 Filter: Cases mit Wartezeiten.....	72
4.5 Fall 3	72
4.5.1 Allgemeine Informationen	72
4.5.2 Prozessmodell	72
4.5.3 Dauer und zurückgelegte Entfernung der Cases.....	73
4.5.4 Durchschnittliche Dauer und Ereignisse pro Variante	74
4.5.5 Mobilitätsprozesse	75
4.5.6 Berechnung der Prozesskennzahl V	76

4.5.7 Berechnung der Prozesskennzahl G	76
4.5.8 Berechnung der Prozesskennzahl D	77
4.5.9 Filter: Kürzeste Cases	78
4.5.10 Filter: Kürzer als Durchschnitt	79
4.5.11 Filter: Cases mit Wartezeiten.....	79
4.6 Fall 4	79
4.6.1 Allgemeine Informationen	79
4.6.2 Prozessmodell	79
4.6.3 Dauer und zurückgelegte Entfernung der Cases	80
4.6.4 Durchschnittliche Dauer und Ereignisse pro Variante	81
4.6.5 Mobilitätsprozesse	82
4.6.6 Berechnung der Prozesskennzahl V	83
4.6.7 Berechnung der Prozesskennzahl G	84
4.6.8 Berechnung der Prozesskennzahl D	84
4.6.9 Filter: Kürzeste Cases	85
4.6.10 Filter: Kürzer als Durchschnitt	85
4.6.11 Filter: Cases mit Wartezeiten.....	86
4.7 Prüfung der aufgestellten Hypothesen	86
4.7.1 Hypothese H1	86
4.7.2 Hypothese H2	86
4.7.3 Hypothese H3	86
5. Diskussion.....	87
5.1 Zusammenfassung und Interpretation der Ergebnisse	87
5.1.1 Möglichkeiten des Process Mining-Tools ‚Disco‘	87
5.1.2 Betrachtung der einzelnen Fälle und Prüfung der Hypothesen	89
5.1.2.1 Prozesskennzahl V	90
5.1.2.2 Prozesskennzahl G	92

5.1.2.3 Prozesskennzahl D	94
5.1.2.4 Anwendung von Filtern	95
5.1.2.5 Hypothesen	97
5.1.2.6 Möglichkeiten zur Nutzung der Informationen aus dem Tool ‚Disco‘	102
5.1.2.7 Weitere Erkenntnisse	102
5.2 Beschränkungen der Forschung	102
5.3 Empfehlung für weiterführende Forschung	103
6. Fazit	105
6.1 Erkenntnisse der Forschung	105
6.2 Beantwortung der Forschungsfrage	108
6.3 Kritische Würdigung und Lösungsansätze	108
6.4 Ausblick	109
Literaturverzeichnis	110
Anhang	113
Eidesstattliche Erklärung	164

Abbildungsverzeichnis

Abbildung 1: Input/Output Modell der Process Mining Ausprägung Discovery. Eigene Darstellung (2022) nach W. van der Aalst et al. (2012, p. 175)	19
Abbildung 2: Input/Output Modell der Process Mining-Ausprägung Conformance Checking. Eigene Darstellung (2022) nach W. van der Aalst et al. (2012, p. 175).....	20
Abbildung 3: Input/Output Modell der Process Mining-Ausprägung Enhancement. Eigene Darstellung (2022) nach W. van der Aalst et al. (2012, p. 175)	21
Abbildung 4: Verteilung der Anzahl der Anwendungsfälle von Process Mining von 2005-2020, Quelle: Cotroneo et al. (2021, p. 245)	25
Abbildung 5: Mobility as a Process. Eigene Darstellung (2022) nach Vogt, Mobility Services and Fleet Management, Vorlesungsskript 2, Folie 11	30
Abbildung 6: unbearbeitete CSV-Datei. Eigene Darstellung (2022)	38
Abbildung 7: In Spalten sortierte Google Maps Timeline-Daten. Eigene Darstellung (2022)	39
Abbildung 8: betrachteter Mobilitätsprozess in Excel. Eigene Darstellung (2022)	39
Abbildung 9: Anonymisierung und Vereinheitlichung der Orte. Eigene Darstellung (2022)	39
Abbildung 10: Unterscheidung der Fortbewegungsarten (Spalte E). Eigene Darstellung (2022).....	40
Abbildung 11: aktualisierter Start-Zeitpunkt für ‚Start‘. Eigene Darstellung (2022).....	40
Abbildung 12: Hinzufügen einer Case ID. Eigene Darstellung (2022)	40
Abbildung 13: auswertbarer Event-Log. Eigene Darstellung (2022)	41
Abbildung 14: Endgültiger Event-Log. Eigene Darstellung (2022).....	42
Abbildung 15: Startseite des Process Mining-Tools ‚Disco‘. Eigene Darstellung (2022) ..	42
Abbildung 16: Deklaration der Spalten in ‚Disco‘. Eigene Darstellung (2022).....	43
Abbildung 17: Darstellung des Mobilitätsprozesses direkt nach dem Import. Eigene Darstellung (2022)	44
Abbildung 18: Prozess der Erhebung und Auswertung der Mobilitätsdaten inkl. genutzter Tools. Eigene Darstellung (2022)	46
Abbildung 19: CRISP-DM-Modell, Quelle: Wuttke (2020)	46
Abbildung 20: Darstellung des Mobilitätsprozesses mit 100%iger Genauigkeit. Eigene Darstellung (2022)	48
Abbildung 21: Darstellung des Mobilitätsprozesses mit 0% bei Activities und Paths. Eigene Darstellung (2022)	49

Abbildung 22: Überblick über alle Werte einer Aktivität/eines Pfades. Eigene Darstellung (2022).....	50
Abbildung 23: Ausschnitt aus einer Animation. Eigene Darstellung (2022)	50
Abbildung 24: Statistiken zum Event-Log in ‚Disco‘. Eigene Darstellung (2022).....	51
Abbildung 25: Tabelle ‚Variants‘. Eigene Darstellung (2022)	52
Abbildung 26: Übersicht über den Punkt ‚Activity‘. Eigene Darstellung (2022).....	52
Abbildung 27: Überblick sonstige Attribute. Eigene Darstellung (2022)	53
Abbildung 28: Überblick über die einzelnen Cases in ‚Disco‘ mit Graph. Eigene Darstellung (2022).....	54
Abbildung 29: Überblick über die einzelnen Cases in ‚Disco‘ mit Tabelle. Eigene Darstellung (2022)	54
Abbildung 30: Fenster zur Anwendung von Filtern & Filtermöglichkeiten in ‚Disco‘. Eigene Darstellung (2022)	55
Abbildung 31: Auswertung Fall 1 - Prozessmodell. Eigene Darstellung (2022)	57
Abbildung 32: Auswertung Fall 1 - Ergebnisse KPI V. Eigene Darstellung (2022).....	61
Abbildung 33: Auswertung Fall 1 - Ergebnisse KPI G. Eigene Darstellung (2022).....	62
Abbildung 34: Auswertung Fall 1 - Ergebnisse KPI D. Eigene Darstellung (2022).....	63
Abbildung 35: Auswertung Fall 2 - Prozessmodell. Eigene Darstellung (2022)	65
Abbildung 36: Auswertung Fall 2 - Ergebnisse KPI V. Eigene Darstellung (2022).....	69
Abbildung 37: Auswertung Fall 2 - Ergebnisse KPI G. Eigene Darstellung (2022).....	70
Abbildung 38: Auswertung Fall 2 - Ergebnisse KPI D. Eigene Darstellung (2022).....	71
Abbildung 39: Auswertung Fall 3 - Prozessmodell. Eigene Darstellung (2022)	73
Abbildung 40: Auswertung Fall 3 - Ergebnisse KPI V. Eigene Darstellung (2022).....	76
Abbildung 41: Auswertung Fall 3 - Ergebnisse KPI G. Eigene Darstellung (2022).....	77
Abbildung 42: Auswertung Fall 3 - Ergebnisse KPI D. Eigene Darstellung (2022).....	78
Abbildung 43: Auswertung Fall 4 - Prozessmodell. Eigene Darstellung (2022)	80
Abbildung 44: Auswertung Fall 4 - Ergebnisse KPI V. Eigene Darstellung (2022).....	83
Abbildung 45: Auswertung Fall 4 - Ergebnisse KPI G. Eigene Darstellung (2022).....	84
Abbildung 46: Auswertung Fall 4 - Ergebnisse KPI D. Eigene Darstellung (2022).....	85
Abbildung 47: Prozessmodelle und Anzahl der Cases. Eigene Darstellung (2022).....	98
Abbildung 48: Anleitung Text in Spalten, Teil 1. Eigene Darstellung (2022).....	123
Abbildung 49: Anleitung Text in Spalten, Teil 2. Eigene Darstellung (2022).....	123
Abbildung 50: Unterschied Prozessmodell mit Schlusszeit als Timestamp (links) und ohne Schlusszeit als Timestamp (rechts). Eigene Darstellung (2022)	123

Abbildung 51: Auswahlmöglichkeiten unter dem Reiter 'Frequency'. Eigene Darstellung (2022).....	124
Abbildung 52: Auswahlmöglichkeiten unter dem Reiter 'Performance'. Eigene Darstellung (2022).....	124
Abbildung 53: Einstellung von Geschäfts- und Urlaubszeiten. Eigene Darstellung (2022)	125
Abbildung 54: Auswertung Fall 1 - Allgemeine Informationen. Eigene Darstellung (2022)	125
Abbildung 55: Auswertung Fall 1 - Varianten in Prozent. Eigene Darstellung (2022).....	126
Abbildung 56: Auswertung Fall 1 - Tabelle: 'Variants'. Eigene Darstellung (2022).....	126
Abbildung 57: Auswertung Fall 1 - Mobilitätsprozess Variante 1. Eigene Darstellung (2022)	126
Abbildung 58: Auswertung Fall 1 - Mobilitätsprozess Variante 2. Eigene Darstellung (2022)	127
Abbildung 59: Auswertung Fall 1 - Mobilitätsprozess Variante 3. Eigene Darstellung (2022)	127
Abbildung 60: Auswertung Fall 1 - Mobilitätsprozess Variante 4. Eigene Darstellung (2022)	128
Abbildung 61: Auswertung Fall 1 - Mobilitätsprozess Variante 5. Eigene Darstellung (2022)	128
Abbildung 62: Auswertung Fall 1 - Mobilitätsprozess Variante 6. Eigene Darstellung (2022)	129
Abbildung 63: Auswertung Fall 1 - Filter: Kürzeste Cases. Eigene Darstellung (2022) ..	131
Abbildung 64: Auswertung Fall 1 - Filter: Kürzer als Durchschnitt. Eigene Darstellung (2022).....	132
Abbildung 65: Auswertung Fall 1 - Filter: Cases mit Wartezeiten. Eigene Darstellung (2022)	132
Abbildung 66: Auswertung Fall 2 - allgemeine Informationen. Eigene Darstellung (2022)	133
Abbildung 67: Auswertung Fall 2 - Varianten in Prozent. Eigene Darstellung (2022).....	133
Abbildung 68: Auswertung Fall 2 - Dauer und Entfernung der Cases. Eigene Darstellung (2022).....	134
Abbildung 69: Auswertung Fall 2 - Tabelle: ‚Variants‘. Eigene Darstellung (2022)	134
Abbildung 70: Auswertung Fall 2 - Prozessablauf Variante 1. Eigene Darstellung (2022)	134

Abbildung 71: Auswertung Fall 2 - Prozessablauf Variante 2. Eigene Darstellung (2022)	135
Abbildung 72: Auswertung Fall 2 - Prozessablauf Variante 3. Eigene Darstellung (2022)	135
Abbildung 73: Auswertung Fall 2 - Prozessablauf Variante 4. Eigene Darstellung (2022)	136
Abbildung 74: Auswertung Fall 2 - Prozessablauf Variante 5. Eigene Darstellung (2022)	136
Abbildung 75: Auswertung Fall 2 - Prozessablauf Variante 6. Eigene Darstellung (2022)	137
Abbildung 76: Auswertung Fall 2 - Prozessablauf Variante 7. Eigene Darstellung (2022)	137
Abbildung 77: Auswertung Fall 2 - Prozessablauf Variante 8. Eigene Darstellung (2022)	138
Abbildung 78: Auswertung Fall 2 - Prozessablauf Variante 9. Eigene Darstellung (2022)	138
Abbildung 79: Auswertung Fall 2 - Prozessablauf Variante 10. Eigene Darstellung (2022)	139
Abbildung 80: Auswertung Fall 2 - Prozessablauf Variante 11. Eigene Darstellung (2022)	139
Abbildung 81: Auswertung Fall 2 - Prozessablauf Variante 12. Eigene Darstellung (2022)	140
Abbildung 82: Auswertung Fall 2 - Prozessablauf Variante 13. Eigene Darstellung (2022)	140
Abbildung 83: Auswertung Fall 2 - Prozessablauf Variante 14. Eigene Darstellung (2022)	141
Abbildung 84: Auswertung Fall 2 - Prozessablauf Variante 15. Eigene Darstellung (2022)	141
Abbildung 85: Auswertung Fall 2 - Filter: Kürzeste Cases. Eigene Darstellung (2022) ..	143
Abbildung 86: Auswertung Fall 2 - Filter: Kürzer als Durchschnitt. Eigene Darstellung (2022).....	144
Abbildung 87: Auswertung Fall 2 - Filter: Cases mit Wartezeit. Eigene Darstellung (2022)	144
Abbildung 88: Auswertung Fall 3 - Allgemeine Informationen. Eigene Darstellung (2022)	145

Abbildung 89: Auswertung Fall 3 - Varianten in Prozent. Eigene Darstellung (2022).....	145
Abbildung 90: Auswertung Fall 3 - Dauer und Entfernung der Cases. Eigene Darstellung (2022).....	146
Abbildung 91: Auswertung Fall 3 - Tabelle: ‚Variants‘. Eigene Darstellung (2022)	146
Abbildung 92: Auswertung Fall 3 - Prozessablauf Variante 1. Eigene Darstellung (2022)	146
Abbildung 93: Auswertung Fall 3 - Prozessablauf Variante 2. Eigene Darstellung (2022)	147
Abbildung 94: Auswertung Fall 3 - Prozessablauf Variante 3. Eigene Darstellung (2022)	147
Abbildung 95: Auswertung Fall 3 - Prozessablauf Variante 4. Eigene Darstellung (2022)	148
Abbildung 96: Auswertung Fall 3 - Prozessablauf Variante 5. Eigene Darstellung (2022)	148
Abbildung 97: Auswertung Fall 3 - Prozessablauf Variante 6. Eigene Darstellung (2022)	149
Abbildung 98: Auswertung Fall 3 - Prozessablauf Variante 7. Eigene Darstellung (2022)	149
Abbildung 99: Auswertung Fall 3 - Prozessablauf Variante 8. Eigene Darstellung (2022)	150
Abbildung 100: Auswertung Fall 3 - Prozessablauf Variante 9. Eigene Darstellung (2022)	150
Abbildung 101: Auswertung Fall 3 - Prozessablauf Variante 10. Eigene Darstellung (2022)	151
Abbildung 102: Auswertung Fall 3 - Filter: Kürzeste Cases. Eigene Darstellung (2022)	152
Abbildung 103: Auswertung Fall 3 - Filter: Kürzer als Durchschnitt. Eigene Darstellung (2022).....	153
Abbildung 104: Auswertung Fall 3 - Filter: Cases mit Wartezeiten. Eigene Darstellung (2022).....	153
Abbildung 105: Auswertung Fall 4 - allgemeine Informationen. Eigene Darstellung (2022)	154
Abbildung 106: Auswertung Fall 4 - Varianten in Prozent. Eigene Darstellung (2022)...	154
Abbildung 107: Auswertung Fall 4 - Dauer und Entfernung der Cases. Eigene Darstellung (2022).....	155
Abbildung 108: Auswertung Fall 4 - Tabelle: ‚Variants‘. Eigene Darstellung (2022)	155

Abbildung 109: Auswertung Fall 4 - Prozessablauf Variante 1. Eigene Darstellung (2022)	155
Abbildung 110: Auswertung Fall 4 - Prozessablauf Variante 2. Eigene Darstellung (2022)	156
Abbildung 111: Auswertung Fall 4 - Prozessablauf Variante 3. Eigene Darstellung (2022)	156
Abbildung 112: Auswertung Fall 4 - Prozessablauf Variante 4. Eigene Darstellung (2022)	157
Abbildung 113: Auswertung Fall 4 - Prozessablauf Variante 5. Eigene Darstellung (2022)	157
Abbildung 114: Auswertung Fall 4 - Prozessablauf Variante 6. Eigene Darstellung (2022)	158
Abbildung 115: Auswertung Fall 4 - Prozessablauf Variante 7. Eigene Darstellung (2022)	158
Abbildung 116: Auswertung Fall 4 - Prozessablauf Variante 8. Eigene Darstellung (2022)	159
Abbildung 117: Auswertung Fall 4 - Filter: Kürzeste Cases. Eigene Darstellung (2022)	161
Abbildung 118: Auswertung Fall 4 - Filter: Kürzer als Durchschnitt. Eigene Darstellung (2022)	161
Abbildung 119: Auswertung Fall 4 - Filter: Cases mit Wartezeiten. Eigene Darstellung (2022)	162
Abbildung 120: Fall 4 - Variante 7: fehlerhafte Aufzeichnung. Eigene Darstellung (2022)	162
Abbildung 121: Auswertung Fall 1 - Berechnung des Mittelwertes der KPI G in Excel. Eigene Darstellung (2022)	162
Abbildung 122: Auswertung Fall 2 - Berechnung des Mittelwertes der KPI G in Excel. Eigene Darstellung (2022)	162
Abbildung 123: Auswertung Fall 3 - Berechnung des Mittelwertes der KPI G in Excel. Eigene Darstellung (2022)	163
Abbildung 124: Auswertung Fall 4 - Berechnung des Mittelwertes der KPI G in Excel. Eigene Darstellung (2022)	163
Abbildung 125: Auswertung Fall 4 - Ermittlung der Durchschnittsdauer ohne Cases 1 und 28. Eigene Darstellung (2022)	163

Tabellenverzeichnis

Tabelle 1: Benötigte Informationen. Eigene Darstellung (2022)	55
Tabelle 2: Auswertung Fall 1 - Dauer und Entfernung der Cases. Eigene Darstellung (2022)	57
Tabelle 3: Auswertung Fall 1 - durchschnittliche Dauer und Anzahl der Ereignisse pro Variante. Eigene Darstellung (2022)	59
Tabelle 4: Auswertung Fall 1 - Mobilitätsprozesse der Varianten. Eigene Darstellung (2022)	60
Tabelle 5: Auswertung Fall 1 - Ergebnisse des Filters 'Case duration max. 21,5min'. Eigene Darstellung (2022)	63
Tabelle 6: Auswertung Fall 2 - Dauer und Entfernung der Cases. Eigene Darstellung (2022)	65
Tabelle 7: Auswertung Fall 2 - durchschnittliche Dauer und Anzahl der Ereignisse pro Variante. Eigene Darstellung (2022)	67
Tabelle 8: Auswertung Fall 2 - Mobilitätsprozesse der Varianten. Eigene Darstellung (2022)	68
Tabelle 9: Auswertung Fall 2 - Ergebnisse des Filters 'Case duration max. 20,5min'. Eigene Darstellung (2022)	71
Tabelle 10: Auswertung Fall 3 - Dauer und Entfernung der Cases. Eigene Darstellung (2022)	73
Tabelle 11: Auswertung Fall 3 - durchschnittliche Dauer und Anzahl der Ereignisse pro Variante. Eigene Darstellung (2022)	74
Tabelle 12: Auswertung Fall 3 - Mobilitätsprozesse der Varianten. Eigene Darstellung (2022)	75
Tabelle 13: Auswertung Fall 3 - Ergebnisse des Filters 'Case duration max. 93min'. Eigene Darstellung (2022)	78
Tabelle 14: Auswertung Fall 4 - Dauer und Entfernung der Cases. Eigene Darstellung (2022)	80
Tabelle 15: Auswertung Fall 4 - durchschnittliche Dauer und Anzahl der Ereignisse pro Variante. Eigene Darstellung (2022)	82
Tabelle 16: Auswertung Fall 4 - Mobilitätsprozesse der Varianten. Eigene Darstellung (2022)	82
Tabelle 17: Auswertung Fall 4 - Ergebnisse des Filters 'Case duration max. 47min'. Eigene Darstellung (2022)	85

Tabelle 18: Anzahl der Ereignisse und Durchschnittsdauer der Varianten aller Fälle. Eigene Darstellung (2022)	99
Tabelle 19: Berechnung des Anteil der Varianten, die nur einmalig aufgetreten sind. Eigene Darstellung (2022)	101
Tabelle 20: Wurde Process Mining bereits auf Mobilitätsprozesse angewendet? - Übersicht Literaturrecherche. Eigene Darstellung (2022)	113
Tabelle 21: Auswertung Fall 1 - Ergebnisse KPI V. Eigene Darstellung (2022)	129
Tabelle 22: Auswertung Fall 1 - Ergebnisse KPIs G & D. Eigene Darstellung (2022)	129
Tabelle 23: Auswertung Fall 2 - Ergebnisse KPI V. Eigene Darstellung (2022)	141
Tabelle 24: Auswertung Fall 2 - Ergebnisse KPIs G & D. Eigene Darstellung (2022)	142
Tabelle 25: Auswertung Fall 3 - Ergebnisse KPI V. Eigene Darstellung (2022)	151
Tabelle 26: Auswertung Fall 3 - Ergebnisse KPIs G & D. Eigene Darstellung (2022)	151
Tabelle 27: Auswertung Fall 4 - Ergebnisse KPI V. Eigene Darstellung (2022)	159
Tabelle 28: Auswertung Fall 4 - Ergebnisse KPIs G & D. Eigene Darstellung (2022)	159

Abkürzungsverzeichnis

BPI	<i>Business Process Improvement</i>
BPM	<i>Business Process Management</i>
BPMN	<i>Business Process Modeling Notation</i>
CPM	<i>Corporate Performance Management</i>
CRISP-DM	<i>Cross Industry Standard Process for Data Mining</i>
CSV	<i>Comma-Separated-Value</i>
EPK	<i>Ereignisgesteuerte Prozesskette</i>
GPM (s. BPM)	<i>Geschäftsprozessmanagement</i>
GPS	<i>Global Positioning System</i>
IT	<i>Informationstechnik</i>
KML	<i>Keyhole Markup Language</i>
KPI	<i>Key Performance Indicator</i>
UML	<i>Unified Modeling Language</i>
WFM	<i>Workforce Management</i>
XES	<i>eXtensible Event Stream</i>

1. Einleitung

Moderne Technologien ermöglichen es der Menschheit immer mehr, das menschliche Verhalten zu verstehen, indem zunehmend Daten darüber gesammelt werden – seien dies nun Informationen über das individuelle Kaufverhalten, persönliche Interessen oder die menschliche Mobilität.

Eine wichtige Rolle dabei spielen mobile Geräte wie zum Beispiel Smartwatches oder Smartphones. Die Omnipräsenz letzterer führt unter anderem dazu, dass Unternehmen wie beispielsweise Google Informationen über personenbezogene Verhaltensmuster anhand von Cookies in Web-Browsern oder durch die Appnutzung der Individuen sammeln und so unter anderem zielgerichtete, personalisierte Werbung schalten können (Ullah et al., 2020, p. 143818).

Durch integrierte Global Positioning System (GPS)-Sensoren ist es außerdem mehr denn je möglich, menschliche Mobilitätsmuster zu verfolgen. Google Maps zum Beispiel bietet die Funktion ‚*Google Maps Timeline*‘ an. Ist diese aktiviert, werden die GPS-Daten des Smartphones getrackt und im Google-Konto des Nutzers gespeichert. Dies ermöglicht es dem Nutzer, die eigenen zurückgelegten Wege und besuchten Orte in einer Karte zu sehen und zu bearbeiten, an welchen Orten man sich aufgehalten hat oder welche Verkehrsmittel man genutzt hat. Was viele jedoch nicht wissen, ist, dass man die aufgezeichneten Mobilitätsdaten exportieren und so einen detaillierten Datensatz mit den Rohdaten der Mobilitätsvorgänge erhalten kann.

Doch wie lassen sich diese Aufzeichnungen nun überhaupt nutzen, um die eigenen Mobilitätsprozesse aufzuschlüsseln und diese zu analysieren? Ein möglicher Ansatzpunkt wäre hier die Anwendung von Process Mining-Techniken auf die aufgezeichneten Mobilitätsdaten.

Bei Process Mining handelt es sich um einen Ansatz zur Analyse von Prozessen, welcher in den späten 1990er Jahren entwickelt wurde (W. M. P. van der Aalst, 2022, p. 4). Besonders im vergangenen Jahrzehnt hat Process Mining in der Geschäftswelt zunehmend an Bedeutung gewonnen (Cotroneo, Carbone, Boggini, & Cerini, 2021, p. 245).

In der Vergangenheit wurde Process Mining bisher auf Geschäftsprozesse in unterschiedlichen Branchen, wie z.B. der Finanz-, Dienstleistungs- oder der Telekommunikationsbranche, angewandt (Cotroneo et al., 2021, p. 3). Besonders häufig

konnte der Einsatz von Process Mining-Techniken auf Prozesse im Gesundheitswesen festgestellt werden (Dakic, Stefanovic, Cosic, Lolic, & Medojevic, 2018, p. 0871).

Um Process Mining anwenden zu können, gibt es ein großes Angebot an Software-Programmen. Neben einigen umfassenden, wissenschaftlichen Process Mining-Programmen, wie unter anderem ‚*ProM*‘, gibt es aktuell bereits über 40 kommerzielle Process Mining-Tools (W. M. P. van der Aalst, 2022, p. 4).

Die Forschungsliteratur beschäftigt sich seit nunmehr über 20 Jahren mit der Thematik Process Mining. Über einen Bezug zwischen Process Mining und Mobilitätsprozessdaten ist bislang jedoch nur wenig bekannt. Die Forschung dazu steht noch am Anfang.

1.1 Problemstellung

Die vorliegende Arbeit behandelt die Problemstellung, wie Process Mining-Techniken auf Mobilitätsdaten angewendet werden können. Da es bisher kaum Forschung über eine Verbindung zwischen diesen beiden Aspekten gibt, muss zunächst eine Grundlage geschaffen werden. Im Rahmen dieser Bachelorarbeit soll deshalb zunächst der Frage nachgegangen werden, welche Erkenntnisse sich über die alltäglichen Mobilitätsprozessdaten von Personen mithilfe der in dieser Arbeit genutzten Process Mining-Software gewinnen und direkt aus dieser ablesen lassen.

Weiterhin ist von Interesse, wie die Auswertungsergebnisse des Tools weiterverarbeitet werden können, um die betrachteten Mobilitätsprozesse zu analysieren und zu bewerten. Aufbauend auf den Ergebnissen der ersten Teilfrage soll deshalb die Frage, wie sich die aus dem Tool gewonnenen Informationen nutzen lassen, um eine Bewertungsgrundlage für die Mobilitätsprozesse zu erhalten, beantwortet werden.

Anhand der Antworten zu den beiden gestellten Teilfragen und der Erkenntnisse der hier durchgeführten Forschung soll abschließend die Forschungsfrage dieser Arbeit beantwortet werden:

„Welche Möglichkeiten bietet die Anwendung eines Process Mining-Tools, wie z.B. ‚*Disco*‘, auf die alltäglichen Mobilitätsprozesse von Personen?“

1.2 Ziel der Arbeit

Das Ziel dieser Arbeit ist es, zunächst den aktuellen Forschungsstand zu den beiden Thematiken ‚*Process Mining*‘ und ‚*Mobilitätsprozesse*‘ zusammenzufassen. Zudem soll anhand einer Fachliteraturrecherche geklärt werden, ob das Process Mining in der

Vergangenheit bereits Anwendung auf Mobilitätsprozessdaten in der wissenschaftlichen Literatur gefunden hat.

Anschließend soll das Process Mining auf reale Mobilitätsprozesse angewendet werden. Dabei soll ein Überblick darüber geschaffen werden, welche Möglichkeiten Process Mining bietet, um menschliche Mobilitätsprozesse zu analysieren und welche Erkenntnisse und Informationen aus einem Process Mining-Tool gewonnen werden können.

Weiterhin sollen Beispiele dafür, wie man die aus der Software gewonnenen Informationen weiterverarbeiten und nutzen kann, um die untersuchten Prozesse zu bewerten, erarbeitet und demonstriert werden.

Abschließendes Ziel dieser Arbeit ist es, die Wirkung des Einsatzes von Process Mining-Techniken auf Mobilitätsprozessdaten zu erforschen und herauszuarbeiten, welche Möglichkeiten dadurch für die Analyse der persönlichen Mobilitätsprozesse entstehen.

Es wird erwartet, dass die unterschiedlichen Mobilitätsprozesse vom Tool zuverlässig erkannt und ausgewertet werden.

Außerdem wird erwartet, dass auf Grundlage der Informationen, welche durch das genutzte Process Mining-Tool zu Verfügung gestellt werden, bereits zahlreiche Entscheidungen über die Mobilitätsprozesse getroffen werden können, wie z.B. über die zeitliche Effizienz dieser.

1.3 Vorgehensweise

Die Forschung dieser Arbeit erfolgt anhand eines Mixed-Methods-Ansatzes. Begonnen wird mit der Sammlung von Mobilitätsdaten im Rahmen einer qualitativen Fallstudie, welche gewählt wurde, um die zu einem spezifischen Fall passenden Mobilitätsprozessdaten zu erheben und somit unterschiedliche Perspektiven und Ausprägungen eines ausgewählten Falles zu beleuchten. Untersucht wird dabei der Fall *„Die Mobilitätsprozesse einer Person von oder zu ihrer Arbeitsstätte bzw. von oder zum Arbeitsort“*. Die gesammelten Daten werden anschließend so weiterverarbeitet, dass diese von der Process Mining-Software genutzt werden können.

Nachdem die Datensätze in das Tool importiert worden sind, erfolgt eine qualitative Inhaltsanalyse der Funktionen des ausgewählten Process Mining-Programms sowie der bereitgestellten Informationen über die Mobilitätsprozesse.

Weiterhin werden ausgewählte Informationen aller untersuchten Fälle gesammelt. Im Anschluss sollen diese dafür genutzt werden, die entdeckten Mobilitätsprozesse anhand quantitativer Methoden, wie beispielsweise der Berechnung von Prozesskennzahlen, zu

bewerten. Gewählt wurde die quantitative Methodik, um messbare Werte als Bewertungsgrundlage für die Prozesse zu erhalten.

1.4 Aufbau der Arbeit

Die vorliegende Bachelorarbeit beginnt mit der Schaffung eines theoretischen Rahmens in Kapitel 2. Zunächst wird der Process Mining Ansatz definiert und dessen zentralen Aufgaben erklärt sowie die bedeutendste Forschungsinstanz des Process Mining, die „*IEEE Task Force on Process Mining*“, vorgestellt. Anschließend werden wichtige Begriffe im Zusammenhang des Process Mining erläutert, um ein umfassendes Grundverständnis über das Themenfeld zu schaffen. Weiterhin werden die unterschiedlichen Ausprägungen dieser Technik, welche alle unterschiedliche Ziele verfolgen, aufgezeigt sowie ein Überblick über die verfügbare Software geschaffen. Beendet wird das Thema Process Mining mit einer Übersicht von theoretischen Anwendungsgebieten und Möglichkeiten sowie realen Anwendungsfällen aus der Vergangenheit. Als Nächstes folgen theoretische Hintergründe zum Begriff ‚*Mobilität*‘, welchem unterschiedliche Bedeutungen zugrunde liegen. Aus diesem Grund wird zunächst die Mobilitätsform herausgearbeitet, welche für diese Arbeit von Bedeutung ist. Auf dieser Basis wird eine Definition des Terminus ‚*Mobilitätsprozess*‘ formuliert. Zudem wird der Mobilitätsprozess in den Überbegriff ‚*Mobility as a Process*‘ eingeordnet und dessen Bestandteile erläutert. Der abschließende Teil des zweiten Kapitels handelt von der Anwendung von Process Mining auf Mobilitätsprozesse. Hier erfolgt zunächst eine Fachliteraturrecherche zur Klärung des aktuellen Forschungsstandes. Im Anschluss wird Google Maps als mögliche Datenquelle von Mobilitätsprozessdaten vorgestellt. Abschließend werden Prozesskennzahlen, anhand welcher Mobilitätsprozesse bewertet werden können, etabliert und Hypothesen aufgestellt, welche im Laufe dieser Forschungsarbeit überprüft werden sollen.

In Kapitel 3 wird der empirische Teil dieser Arbeit in Form eines Mixed-Methods-Ansatzes zur Untersuchung der Auswirkungen und der Möglichkeiten des Einsatzes von Process Mining-Techniken auf Mobilitätsprozesse durchgeführt. Zunächst erfolgt eine qualitative Fallstudie, bei der der Fall ‚*Die Mobilitätsprozesse einer Person von oder zu ihrer Arbeitsstätte bzw. von oder zum Arbeitsort*‘ betrachtet wird. Es wird dargelegt, wie die Mobilitätsprozessdaten erhoben und weiterverarbeitet wurden, um diese für Process Mining-Algorithmen nutzbar zu machen. Im gleichen Zuge wird aufgezeigt, welche Informationen die erhobenen Mobilitätsdaten über die Prozesse preisgeben. Anschließend wird der Import der Datensätze in das Process Mining-Tool detailliert beschrieben und aufgezeigt, welchen Informationen welche Attribute zugeordnet werden, damit das Tool mit den Daten arbeiten

kann. Weiterhin wird geschildert, wie bei der Auswertung der entstandenen Informationen und Werte vorgegangen wurde.

Die Ergebnisse der Untersuchung werden in Kapitel 4 zusammengefasst. Zu Beginn wird ein Überblick über die Funktionen des verwendeten Tools sowie über die bereitgestellten Daten geschaffen. Außerdem werden die Werte, welche für jeden betrachteten Fall benötigt werden, und die Mobilitätsprozesse bewerten zu können, herausgearbeitet. Anschließend werden diese für jeden Fall einzeln zusammengefasst sowie genutzt, um die Prozesskennzahlen zu berechnen und Filter im Process Mining-Tool anzuwenden. Weiterhin werden die in Kapitel 2 aufgestellten Hypothesen auf ihre Richtigkeit hin untersucht.

In Kapitel 5 werden die Ergebnisse zusammengefasst und interpretiert. Unter anderem werden die Möglichkeiten des Process Mining-Tools diskutiert, die Ergebnisse der Berechnungen der Kennzahlen sowie der Anwendung von Filtern gedeutet und die Erreichung der Resultate der Prüfung der Hypothesen erläutert. Anhand dessen können die ersten beiden Teilfragen dieser Arbeit beantwortet werden. Das Kapitel wird abgerundet mit der Nennung der Beschränkungen dieser Forschung sowie mit Empfehlungen für weiterführende Forschungsprojekte in der Zukunft.

2. Theoretische Grundlagen

Das folgende Kapitel befasst sich mit den theoretischen Hintergründen zu Process Mining sowie zu Mobilitätsprozessen. Es wird weiterhin geklärt, ob die beiden Bereiche in der Fachliteratur bereits miteinander in Verbindung gebracht worden sind. Zudem werden Prozesskennzahlen definiert und Hypothesen zur weiteren Prüfung aufgestellt.

2.1 Process Mining

Der Begriff ‚*Process Mining*‘ beschreibt eine relativ neue Technik zur Prozessdatenverarbeitung, welche um die Jahrtausendwende herum entwickelt worden ist (W. M. P. van der Aalst, 2022, pp. V, 4). Nachfolgend wird sich mit dieser Technik befasst und unter anderem der Begriff Process Mining definiert, wichtige Begriffe in dessen Zusammenhang erklärt, die unterschiedlichen Arten des Process Mining erläutert, Software-Tools vorgestellt und theoretische sowie reale Einsatzgebiete dieser Technik genannt.

2.1.1 Was ist Process Mining?

Der Begriff ‚*Process Mining*‘ beschreibt eine prozessorientierte und faktenbasierte Technik zur systematischen Analyse von End-to-End-Prozessen, welche es Unternehmen ermöglichen soll, die Funktionsweise ihrer Geschäftsprozesse besser zu verstehen und diese managen zu können (Laue, Koschmider, & Fahland, 2021, p. 165; Peters & Nauroth, 2019, p. 3; W. van der Aalst, 2012, pp. 1-2). Konkret besteht das Ziel von Process Mining darin, **reale Prozesse** zu **entdecken**, zu **kontrollieren** und zu **optimieren**, indem Wissen aus Event-Logs (dt. Ereignisprotokolle) gewonnen wird (W. van der Aalst, 2012, p. 1; W. van der Aalst et al., 2012, pp. 170, 172).

Dies beinhaltet Aufgaben wie unter anderem die Bereitstellung von Einblicken in den jeweiligen Prozess, die Identifizierung von Engpässen, die Vorhersage von zukünftig auftretenden Schwierigkeiten und die Empfehlung von Gegenmaßnahmen (W. van der Aalst et al., 2012, p. 174).

Hierbei verbindet Process Mining Ansätze des prozessorientierten Geschäftsprozessmanagements, des datenzentrierten Data-Mining und des Workflow-Managements miteinander (Peters & Nauroth, 2019, p. 3; W. van der Aalst, 2012, p. 2). Zudem ist Process Mining eine Enabling-Technologie für Managementansätze wie Corporate Performance Management (CPM), Business Process Improvement (BPI), Six Sigma und viele weitere (W. van der Aalst et al., 2012, p. 172).

In der heutigen Zeit gewinnt Process Mining immer mehr an Bedeutung. Das liegt zum einen daran, dass zunehmend Aktivitäten aufgezeichnet werden, welche detaillierte Informationen zu Prozessen preisgeben (W. van der Aalst et al., 2012, p. 170). Die allgegenwärtige Verfügbarkeit solcher Daten erlaubt es, Prozesse anhand der aufgezeichneten historischen Fakten zu analysieren, anstatt sich an manuell erstellten Soll-Prozessmodellen zu orientieren (W. van der Aalst, 2012, p. 3). Zum anderen gewinnt Process Mining an Bedeutung, da es notwendig ist, Geschäftsprozesse in sich schnell ändernden Wettbewerbsumgebungen unterstützen und optimieren zu können (W. van der Aalst et al., 2012, p. 70).

2.1.2 IEEE Task Force on Process Mining

Im Jahr 2009 wurde die „*IEEE Task Force on Process Mining*“ gegründet, zu welcher zahlreiche Softwareunternehmen, Beratungsfirmen, Endnutzer und Forschungsinstitute gehören (W. van der Aalst, 2012, p. 9; W. van der Aalst et al., 2012, p. 173).

Ziele der Task Force sind, die Forschung, Entwicklung, Lehre und das Verständnis von Process Mining voranzubringen (W. van der Aalst, 2012, p. 9). Es soll über den aktuellen Stand der Technik informiert, die Nutzung von Process Mining-Techniken vorangetrieben und die Formate für Eventdaten standardisiert werden. Außerdem werden Sondertagungen, Workshops und Gremien organisiert sowie Artikel, Bücher, Videos und Journalbeiträge veröffentlicht (W. van der Aalst, 2012, p. 9; W. van der Aalst et al., 2012, p. 173).

Eine bedeutende Veröffentlichung der Task Force ist das „*Process Mining Manifesto*“, welches Leitsätze und Herausforderungen des Process Mining beschreibt (W. van der Aalst, 2012, p. 9).

2.1.3 Begriffsklärung

Im Kontext des Process Mining gibt es einige Begriffe, die der Klärung bedürfen. Diese werden im folgenden Kapitel definiert und genauer erläutert.

2.1.3.1 Event-Log

Event-Logs sind eine „Sammlung von Ereignisdaten“ (Peters & Nauroth, 2019, p. 15). und bilden die Ausgangsbasis für die Nutzung von Process Mining-Techniken, welche Prozesse automatisiert identifizieren, kontrollieren und optimieren (Laue et al., 2021, p. 166; W. van der Aalst, 2012, p. 1; W. van der Aalst et al., 2012, p. 174).

Die Event-Logs speichern Prozessaktivitäten, die von Menschen, Maschinen und Software-Programmen ausgeführt werden (W. van der Aalst, 2012, p. 1). Wird also ein Arbeitsschritt

des betrachteten Prozesses von einer Ressource durchgeführt, kann dies von einem IT-System im Event-Log protokolliert werden (Laue et al., 2021, p. 166).

Ein Event-Log besteht aus den folgenden Elementen:

- **Aktivität (engl. activity):** ein klar definierter Schritt in einem Prozess (W. van der Aalst et al., 2012, p. 193).
- **Ereignis (engl. event):** eine im Event-Log aufgezeichnete Aktion, welche die Durchführung - z.B. das Starten, das Beenden oder das Abbrechen - einer *Aktivität* darstellt (Laue et al., 2021, p. 166; W. van der Aalst et al., 2012, p. 193).
- **Fall bzw. Prozessinstanz (engl. case/process instance):** die Instanz, die von dem zu analysierenden Prozess bearbeitet wird (W. van der Aalst et al., 2012, p. 194).

Die genannten Bestandteile eines Event-Logs stehen in folgenden Beziehungen zueinander:

- Jedes **Ereignis** bezieht sich auf eine **Aktivität** (W. van der Aalst et al., 2012, p. 174).
- **Ereignisse** werden immer einem **Fall** zugeordnet (W. van der Aalst et al., 2012, p. 194).
- **Aktivitäten** werden immer innerhalb eines **Falles** durchgeführt (Laue et al., 2021, p. 167; Peters & Nauroth, 2019, p. 15).
- Ein **Event-Log** speichert zahlreiche **Ereignisse** aus unterschiedlichen **Fällen** (Laue et al., 2021, p. 167).

Aus diesen Beziehungen heraus lassen sich weitere Elemente von Event-Logs definieren und erklären:

- **Variante bzw. Ablauf (engl. variant/trace):** alle chronologisch geordneten *Ereignisse* eines *Falles* (Laue et al., 2021, p. 167).
- **Attribut (engl. attribute):** Attribute beinhalten Informationen zu einem spezifischen *Ereignis*. Standard-Attribute sind der *Aktivitätsname* (engl. *activity name*; welche Aktivität wurde beim Auftreten des jeweiligen Ereignisses ausgeführt?), der *Zeitstempel* (engl. *timestamp*; wann wurde das Ereignis aufgezeichnet?) und der *Fallidentifikator* (engl. *case identifier* – kurz: *Case ID*; in welchem Fall wurde die Aktivität durchgeführt?) (Laue et al., 2021, pp. 167-169).

Peters und Nauroth (2019, p. 15) geben an, dass ein Ereignisprotokoll mindestens Informationen zu dem konkreten Vorgang (Fall/Case) und den jeweiligen Aktivitäten sowie zu den Zeitstempeln und den Ressourcen enthalten sollte, um daraus Erkenntnisse mithilfe von Process Mining gewinnen zu können.

Ebenso ist es möglich, den Event-Logs, ergänzend zu den obligatorischen Informationen, zusätzliche Attribute hinzuzufügen, welche anschließend dazu genutzt werden können, die Analyse nach unterschiedlichen Attributen und Merkmalen zu filtern (Peters & Nauroth, 2019, p. 15).

Ferner ist es zwingend notwendig, dass die in den Event-Logs gesammelten Ereignisdaten die Prozessaktivitäten in chronologischer, aufsteigender Reihenfolge darstellen, „um kausale Zusammenhänge im Prozess-Modell“ (Peters & Nauroth, 2019, p. 15) erkennen zu können (Peters & Nauroth, 2019, p. 15; W. van der Aalst et al., 2012, p. 174).

2.1.3.2 CSV-Format

Das Comma-Separated-Value (CSV)-Format ist ein text-basiertes Format zur Speicherung von Daten. Es wird von den meisten Process Mining-Programmen unterstützt (Laue et al., 2021, p. 170).

2.1.3.3 XES Standard

Das eXtensible Event Stream (XES)-Format stellt ein strukturiertes Ereignisprotokoll dar (Laue et al., 2021, p. 176). Es wurde von der *IEEE Task Force on Process Mining* als das Standardformat für Event-Logs festgelegt (W. van der Aalst et al., 2012, p. 194).

Im Gegensatz zum CSV-Format definiert der XES-Standard ein Datenmodell sowie Kriterien für die Konsistenz der Daten in Event-Logs. Außerdem ermöglicht es das XES-Format, jedes Ereignis nach dem Fall, in dem es aufgetreten ist, zu gruppieren und chronologisch einem spezifischen Ablauf zuzuordnen (Laue et al., 2021, pp. 170, 174).

Das CSV-Format lässt sich in das XES-Format überführen. Dafür müssen jeweils ein Aktivitäts- und ein Fall-Attribut gewählt werden, welche für jedes Ereignis definiert sind (Laue et al., 2021, pp. 176-177).

Das wissenschaftliche Process Mining Tool ‚*ProM*‘, welches unter Punkt 2.1.5 vorgestellt wird, bietet zu dem Zwecke der Überführung des CSV-Formates in den XES-Standard ein geeignetes Plugin.

2.1.3.4 Prozessmodell

Prozessmodelle sind eine formale Beschreibung aller zentralen Elemente und Aspekte eines (Geschäfts)Prozesses, die zur Ausführung von einzelnen Prozessaufgaben erforderlich sind, und stellen diese unter Verwendung von Modellierungssprachen verständlich dar (Borrmann, König, Koch, & Beetz, 2021, p. 77; Laue et al., 2021, p. 2).

Das Ziel der Prozessmodellierung ist es, reale Prozesse und Vorgänge darzustellen. Process Mining spielt dabei deshalb eine bedeutende Rolle, da Prozessmodelle vorwiegend manuell erstellt werden und somit das Problem besteht, dass sich die Prozessabläufe im Modell häufig von den Abläufen in der Realität unterscheiden (Laue et al., 2021, p. 165). Durch die Anwendung von Process Mining-Techniken können also Prozessmodelle auf Grundlage realer, historischer Ereignisprotokoll-Daten erstellt werden, wodurch die Gefahr, dass sich die Modelle von der Realität differenzieren, minimiert wird. Infolgedessen wird der eigentliche Zweck eines Prozessmodells, nämlich die Darstellung realer Prozessvorgänge, erfüllt.

2.1.4 Arten des Process Mining

Es kann grundsätzlich zwischen drei Ausprägungen des Process Mining unterschieden werden: **Process Discovery** (*Entdeckung realer Prozesse*), **Conformance Checking** (*Abweichungsanalyse*) und **Model Enhancement** (*Erweiterung*). Des Weiteren entwickelte sich in den vergangenen Jahren eine weitere Art des Process Mining: der **Operational Support** (dt. *operative Unterstützung*) (Laue et al., 2021, pp. 196-197; Peters & Nauroth, 2019, p. 6). Die Ausprägungen erfüllen jeweils unterschiedliche Funktionen und können einzeln oder auch ergänzend zueinander angewendet werden. Im Folgenden werden die Disziplinen des Process Mining näher erläutert.

2.1.4.1 Process Discovery

Process Discovery ist die bekannteste Ausprägung des Process Mining (W. van der Aalst et al., 2012, p. 175).

Bei dieser Art des Process Mining werden Ereignisprotokolle genutzt, anhand derer automatisch Ist-Prozessmodelle mit einer erkennbaren Kontrollflusslogik erstellt werden (Laue et al., 2021, pp. 196, 236; Peters & Nauroth, 2019, p. 6; W. van der Aalst, 2012, p. 3). Diese sollen die realen Prozessabläufe, die in den Event-Logs gespeichert sind, möglichst exakt wiedergeben und somit Ungenauigkeiten, die während einer manuellen Prozessmodellierung entstehen können, vermeiden (Laue et al., 2021, pp. 196-197).

In Abbildung 1 wird der Process Discovery Ablauf in Form eines Input/Output Modells veranschaulicht. Als Input dient hier ein Event-Log, welcher mithilfe von Process Discovery Techniken analysiert wird. Der Output ist ein Modell des Ist-Prozesses.

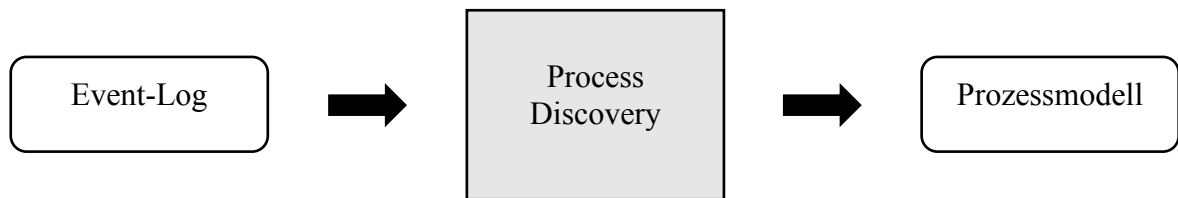


Abbildung 1: Input/Output Modell der Process Mining Ausprägung Discovery. Eigene Darstellung (2022) nach W. van der Aalst et al. (2012, p. 175)

Um die entstandenen Prozessmodelle darzustellen, eignen sich unterschiedliche Notationen wie z.B. Petri-Netze, Business Process Modeling Notation (BPMN), Unified Modeling Language (UML) Aktivitätsdiagramme, Ereignisgesteuerte Prozessketten (EPK) und weitere (W. van der Aalst, 2012, p. 5).

Laut W. van der Aalst (2012) können die beim Process Discovery entdeckten Prozessmodelle genutzt werden, um z.B.:

- Vorschläge zur Prozessoptimierung zu erarbeiten,
- Modelle zu verbessern und erweitern, und
- Workforce Management (WFM)/Business Process Management (BPM) Systeme zu konfigurieren (W. van der Aalst, 2012, p. 4).

2.1.4.2 Conformance Checking

Mithilfe der Process Mining-Ausprägung Conformance Checking wird überprüft, inwiefern der reale Prozess mit einem Soll-Prozess übereinstimmt (W. van der Aalst et al., 2012, p. 175).

Hierbei wird ein vorgegebenes Soll-Prozessmodell mit dem Ist-Modell bzw. mit dem Event-Log des entsprechenden Prozesses verglichen und Abweichungen vom erwünschten Verhalten aufgezeigt (Laue et al., 2021, p. 197; Peters & Nauroth, 2019, p. 6). Besonders bei nicht vollautomatisch ausgeführten Prozessen treten vermehrt Differenzen auf (Laue et al., 2021, p. 199).

Abbildung 2 stellt ein Input/Output Modell der Process Mining-Ausprägung Conformance Checking dar. Diese nutzt einen Event-Log sowie ein Prozessmodell als Input. Das Prozessmodell kann entweder manuell erstellt oder durch Process Discovery entdeckt worden sein (W. van der Aalst, 2012, p. 6). Der Output nach der Analyse durch einen Conformance Checking-Algorithmus ist ein Vergleich zwischen dem Soll- und Ist-Prozess, welcher Übereinstimmungen und Abweichungen aufzeigt.

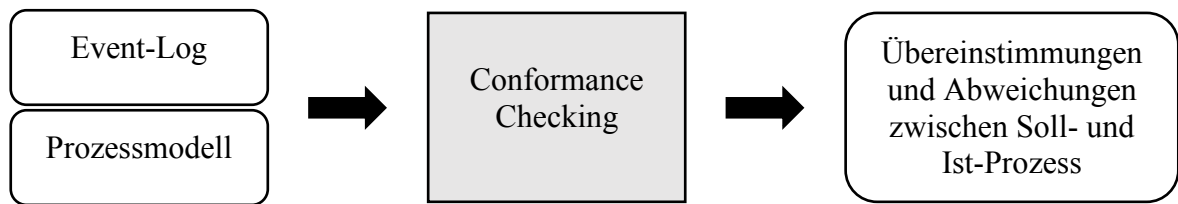


Abbildung 2: Input/Output Modell der Process Mining-Ausprägung Conformance Checking. Eigene Darstellung (2022) nach W. van der Aalst et al. (2012, p. 175)

Laue et al. (2021, p. 203) unterscheiden zwischen zwei Arten von Abweichungen. Zum einen kann Verhalten, welches im Event-Log enthalten ist, jedoch nicht vom Soll-Prozessmodell erlaubt wird, auftreten. Zum anderen gibt es Verhalten, das zwar im Soll-Modell enthalten ist, jedoch nicht im Event-Log.

Conformance Checking kann unter anderem genutzt werden:

- um die Qualität der aufgezeichneten Prozesse zu überprüfen,
- um abweichende Vorgänge zu entdecken,
- um Stellen im Prozess zu identifizieren, an denen häufig Abweichungen auftreten,
- zur Standardisierung,
- für aktives Geschäftsprozessmanagement in der Analyse- und Optimierungsphase,
- zu Prüfzwecken,
- zur internen Revision,
- um die Qualität eines entdeckten Prozessmodells zu bewerten,
- zur Steuerung von Algorithmen zur Entdeckung evolutionärer Prozesse, und
- als Ausgangsbasis für die dritte Process Mining-Ausprägung Enhancement (Peters & Nauroth, 2019, p. 6; W. van der Aalst, 2012, p. 7).

2.1.4.3 Model Enhancement

Die Process Mining-Ausprägung Enhancement hat zum Ziel, ein bestehendes Prozessmodell mithilfe von Informationen aus dem entsprechenden Event-Log, sprich anhand des realen Prozesses, zu erweitern oder zu optimieren (W. van der Aalst, 2012, p. 8; W. van der Aalst et al., 2012, p. 175).

Abbildung 3 stellt diesen Vorgang in Form eines Input/Output Modells dar. Der Input besteht aus einem Event-Log und einem Prozessmodell. Durch Process Enhancement wird ein erweitertes oder verbessertes Modell entwickelt.



Abbildung 3: Input/Output Modell der Process Mining-Ausprägung Enhancement. Eigene Darstellung (2022) nach W. van der Aalst et al. (2012, p. 175)

Process Enhancement kann unter anderem eingesetzt werden, um ein Prozessmodell mit weiteren Attributen aus dem Event-Log zu ergänzen. Diese sind z.B. der Zeitstempel, die Ressource, das Alter einer beteiligten Person (z.B. Kunde) etc. (W. van der Aalst, 2012, pp. 8-9).

Enhancement wird genutzt, um beispielsweise:

- aktives Geschäftsprozessmanagement zu betreiben,
- Wartezeiten zwischen Aktivitäten zu analysieren,
- Engpässe zu identifizieren,
- die Rollen der Ressourcen zu entdecken,
- auf Grundlage des Arbeitsflusses Netzwerke zu konstruieren, und
- die Leistung der Ressourcen zu untersuchen (Peters & Nauroth, 2019, p. 6; W. van der Aalst, 2012, p. 9).

Weitere Anwendungsgebiete des Process Enhancement sind ebenso die Standardisierung und das Blueprinting für IT-Implementierungen (Peters & Nauroth, 2019, p. 6).

2.1.4.4 Operational Support

Im Laufe der Zeit hat sich im fortgeschrittenen ökonomischen Gebrauch eine weitere Process Mining-Disziplin entwickelt. Diese wird als ‚Operational Support‘ bezeichnet, zu Deutsch bedeutet das sinngemäß in etwa „operative Unterstützung IT-basierter Systeme“ (Laue et al., 2021, p. 197; Peters & Nauroth, 2019, p. 6).

Die Anwendung dieser Process Mining-Art ermöglicht es dem Nutzer, bei der Bearbeitung von Prozessen Echtzeithilfe sowie Vorhersagen und Empfehlungen zu erhalten (Laue et al., 2021, p. 197; Peters & Nauroth, 2019, p. 6).

Laue et al. (2021) nennen zwei mögliche Anwendungsfälle für den Operational Support. Es lässt sich z.B. vorhersagen, wann ein Fall voraussichtlich abgeschlossen sein wird. Dadurch wird es möglich, Fälle zu priorisieren und somit „lange Liegezeiten [in der Produktion] zu vermeiden“ (Laue et al., 2021, p. 197). Eine weitere Anwendungsmöglichkeit besteht darin,

Entscheidungsempfehlungen vorzuschlagen, „auf deren Grundlage Entscheidungen konsistent zu anderen Fällen getroffen werden können.“ (Laue et al., 2021, p. 197).

2.1.5 Tools

Mit der Zeit wurden eine Vielzahl an wissenschaftlichen und kommerziellen Process Mining Tools entwickelt (Peters & Nauroth, 2019, p. 41).

Wissenschaftliche Tools ermöglichen flexibleres Arbeiten und tragen dazu bei, dass der Process Mining Ansatz aus wissenschaftlicher Perspektive besser erforscht werden kann (Peters & Nauroth, 2019, p. 41). Das Open-Source-Tool ‚*ProM*‘, welches im Jahre 2004 veröffentlicht wurde, stellt den Standard unter den Process Mining-Tools dar, da es eine einzigartige Auswahl an unterschiedlichen Process Mining-Algorithmen bietet und bereits über 1500 Plug-ins enthält (W. van der Aalst, 2012, p. 11; W. M. P. van der Aalst, 2022, p. 4).

Um wissenschaftliche Software-Tools anzuwenden, bedarf es fundierter Fähigkeiten in der Informationstechnik (IT) und im Process Mining (Peters & Nauroth, 2019, p. 41; W. van der Aalst, 2012, p. 11).

Aus unternehmerischer Sicht eignen sich **kommerzielle Process Mining-Tools** besser, da diese ein benutzerfreundliches User-Interface bieten. Infolgedessen ermöglicht kommerzielle Software aktives Geschäftsprozessmanagement (GPM) mit möglichst geringem Aufwand (Peters & Nauroth, 2019, p. 41).

Auf dem Markt gibt es bereits eine große Vielfalt an kommerziellen Process Mining-Tools. Einige davon sind unter anderem *Celonis* Process-Mining, *myInvenio*, *Disco*, *LANA* Process Mining, *Comprehend*, *Futura Reflect*, *ARIS* Process Mining und *Signavio Process Intelligence* (Peters & Nauroth, 2019, pp. 43-57; W. van der Aalst, 2012, p. 11; W. van der Aalst et al., 2012, p. 172).

Stand Juli 2012 stellten alle kommerziellen Tools lediglich Process Discovery-Techniken zur Verfügung. Kommerzielle Software legte zu Beginn ihren Fokus also auf die Entdeckung von Prozessen und die Messung der Performance (W. van der Aalst, 2012, pp. 11-12).

Peters und Nauroth (2019, pp. 42-57) untersuchten einige Jahre später insgesamt 20 kommerzielle Process Mining-Tools. Dabei stellten sie fest, dass im Vergleich zu 2012 bereits einige Programme mehrere Process Mining-Arten anboten. Jedoch unterschieden sich die Möglichkeiten und die unterstützten Process Mining-Ausprägungen von Tool zu

Tool. Inzwischen gibt es über 40 kommerzielle Process Mining-Tools auf dem Markt (W. M. P. van der Aalst, 2022, p. 4).

Daraus lässt sich ableiten, dass nicht jedes Process Mining-Tool für alle individuellen Ziele einsetzbar ist. Nutzer sollten ihre Ziele im Vorfeld klar formulieren und anhand dessen ein geeignetes Tool, welches die nötigen Eigenschaften und Fähigkeiten besitzt, auswählen.

2.1.6 Einsatzgebiete von Process Mining in der Theorie und der Realität

Durch die Nutzung von Process Mining-Techniken entsteht eine Vielzahl an Möglichkeiten die gewonnenen Erkenntnisse zu nutzen. Theoretische und reale Anwendungsmöglichkeiten werden auf den nachfolgenden Seiten erläutert.

2.1.6.1 Erkennung realer Prozesse

Process Mining, genauer die Ausprägung Process Discovery, ermöglicht es, ein vollständiges Bild der tatsächlichen Vorgänge eines Prozesses zu erstellen. Dadurch wird eine 100-prozentige Transparenz der Prozessabläufe erzielt, während Soll-Prozessmodelle lediglich den idealen Ablauf darstellen (Peters & Nauroth, 2019, p. 25).

Dadurch entsteht eine Vielzahl an weiteren möglichen Anwendungsbereichen, z.B.:

- die Identifikation von Arbeitsschritten, welche der Nacharbeit dienen,
- die Erkennung von Metriken, Relationen und Mustern sowie deren Ursachen,
- die Ermittlung zeitlicher Faktoren wie Durchlaufzeiten, Engpässen und Verzögerungen,
- die Bestimmung von Automatisierungsgrad und Prozesskosten,
- die Erkennung von Prozessbrüchen,
- etc. (Peters & Nauroth, 2019, pp. 25-26)

2.1.6.2 Prozessoptimierung

Durch den Informationsreichtum der Ereignisprotokolle ist es möglich, aus dem entdeckten Ist-Prozessmodell nach unterschiedlichen Attributen zu filtern und so „Zugriff auf die tatsächlichen Stellschrauben der Geschäftsprozesse“ (Peters & Nauroth, 2019, p. 26) zu erhalten.

Dies ermöglicht in der Theorie:

- die Behebung von Engpässen und Verzögerungen,
- die Gestaltung von durchgängigen Prozessen,
- die Automatisierung bzw. die Optimierung des vorhandenen Automatisierungsgrades, und

- die Reduktion von Durchlaufzeiten und Prozesskosten (Peters & Nauroth, 2019, p. 26).

2.1.6.3 Prozessmanagement

Neben der Entdeckung und der Optimierung von Prozessen ermöglicht es Process Mining den Unternehmen, die es anwenden, durch das gewonnene Wissen ihre Geschäftsprozesse zu managen.

Folgende Möglichkeiten entstehen durch den Einsatz von Process Mining unter anderem für das Prozessmanagement:

- agiles Geschäftsprozessmanagement,
- Benchmarking (Vergleich von Performance und Strukturen des Unternehmens mit internen und/oder externen Referenzen (Rosemann, Schwegmann, & Delfmann, 2012, p. 55)) der Geschäftsprozesse zwischen den unterschiedlichen Unternehmensbereichen, Werken und Organisationseinheiten,
- Anregung des Re-Engineerings der Geschäftsprozesse,
- Unterstützung von Echtzeit-Geschäftsprozessen anhand von historischen Prozessdaten,
- Umsetzung schneller, iterativer Prozessänderungen, und
- effiziente Gestaltung und Realisierung von Prozessänderungen sowie Messung der Folgen dieser Änderungen (Peters & Nauroth, 2019, p. 27).

2.1.6.4 Unterstützung entstehender und etablierter Unternehmen

Peters und Nauroth (2019) zufolge kann Process Mining „in verschiedenen Phasen der Unternehmensführung helfen.“ (Peters & Nauroth, 2019, p. 28)

Ist ein Unternehmen erst kürzlich gegründet worden, kann Process Mining dabei behilflich sein, zu überprüfen, ob das Geschäftsmodell funktioniert. Dabei werden die tatsächlichen Geschäftsprozesse mithilfe von Process Mining analysiert (Peters & Nauroth, 2019, p. 29).

Ein etabliertes Unternehmen kann Process Mining nutzen, um zu erfahren, ob die Geschäftsprozesse effizient und effektiv sind. Nachdem der Ist-Zustand der Prozesse entdeckt und Einblicke in den Ressourceneinsatz, Automatisierungsgrad, usw. gewonnen wurden, können die realen Prozesse mit den eigenen Ansprüchen des Unternehmens sowie mit dem Wettbewerb verglichen werden (Peters & Nauroth, 2019, p. 29).

Auch im Falle einer Einführung, Ablösung oder Integration von IT-Systemen ist der Einsatz von Process Mining von Vorteil. Bevor dies geschehen kann, sollten die manifestierten

Prozesse überprüft und gegebenenfalls optimiert werden. Process Mining kann dabei eingesetzt werden, um die Status-Quo-Prozesse zu analysieren und künftige Prozesse zu designen. Durch aktives Prozessdesign lässt sich so ein langfristiger Wettbewerbsvorteil erarbeiten (Peters & Nauroth, 2019, pp. 29-30).

Des Weiteren bietet der Einsatz von Process Mining im IT-Projektmanagement zahlreiche Vorteile. Diese sind unter anderem eine Verbesserung der Produktivität, Termintreue, Investitionssicherheit, Wettbewerbsfähigkeit und der Mitarbeiterzufriedenheit (Peters & Nauroth, 2019, p. 30).

2.1.6.5 reale Anwendungsfälle von Process Mining

Die Studie „*Process Mining: A Database of Applications*“ der HSPI Management Consulting untersucht seit 2016 Einsatzgebiete und Anwendungsfälle des Process Mining in der Realität mit dem Ziel, eine aktuelle und möglichst vollständige Datenbank praktischer Anwendungen des Process Mining Ansatzes zu erstellen (Cotroneo et al., 2021, p. 2). Die Datenbank wird seither regelmäßig erweitert und aktualisiert.

In der Datenbanksammlung befinden sich (Stand 2021) 677 reale Anwendungsfälle des Process Mining (Cotroneo et al., 2021, pp. 7-244). Sortiert man die Fälle zeitlich, wird ersichtlich, dass Process Mining mit den Jahren zunehmend an Bedeutung gewonnen hat, wie aus Abbildung 4 hervorgeht (Cotroneo et al., 2021, p. 245).

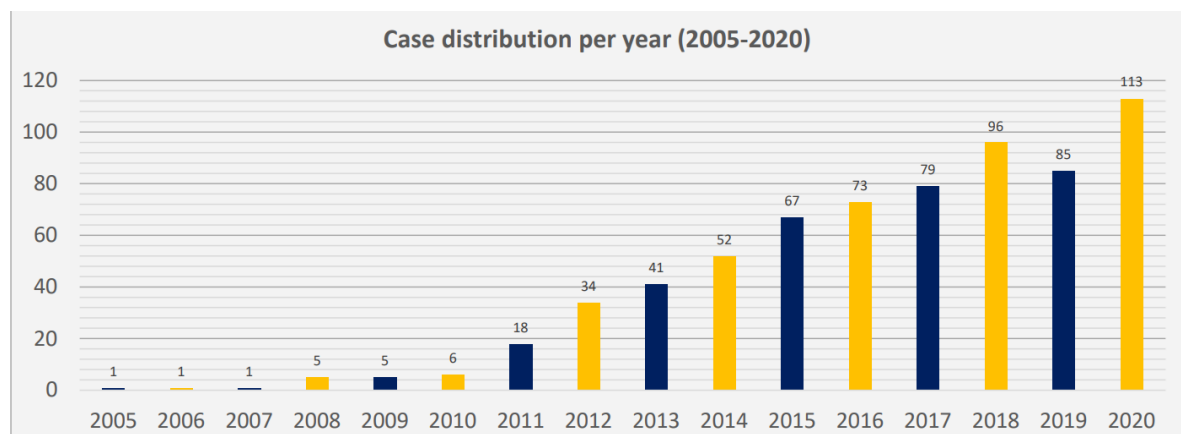


Abbildung 4: Verteilung der Anzahl der Anwendungsfälle von Process Mining von 2005-2020, Quelle: Cotroneo et al. (2021, p. 245)

Process Mining-Techniken wurden bereits in den unterschiedlichsten Branchen eingesetzt. Diese sind der Studie zufolge Finanzen, Gesundheit, Öffentlichkeit, Dienstleistungen, Technologie, Telekommunikation, Versorgung, Industrie, Konsumgüter, Verbraucherdienste, Grundstoffe und E-Learning (Cotroneo et al., 2021, p. 3).

22% der Fälle lassen sich der Industrie zuordnen, gefolgt von der Finanzbranche mit 17%, dem Gesundheitswesen mit 15% und der Telekommunikationsbranche mit 10% (Cotroneo et al., 2021, p. 248).

Weiterhin untersuchten Dakic et al. (2018, p. 0866) anhand einer Fachliteraturrecherche die Einsatzgebiete des Process Mining in der Wirtschaft in einem Zeitraum von zehn Jahren. Es konnten zwölf Branchen, in denen Process Mining angewandt wurde, festgestellt werden: Gesundheitswesen (in 28% der Arbeiten), IT-Branche (14%), Finanzbranche (11%), Produktion, Bildung, Regierung, Energie (jeweils 8%), Landwirtschaft (6%), Logistik, Telekommunikation, Informationsmedien und Tourismus (jeweils 3%) (Dakic et al., 2018, p. 0871).

Anhand der Studie von Cotroneo et al. (2021) sowie der Literaturrecherche von Dakic et al. (2018) fällt auf, dass sich die Anwendung des Process Mining Ansatzes bereits in einigen Branchen etabliert hat. Darunter zählen vor allem die Industrie/Produktion, die Gesundheitsbranche, die Finanzbranche und die Telekommunikationsbranche.

Dakic et al. (2018, p. 0871) nennen in ihrer Arbeit außerdem konkrete Anwendungsfälle des Process Mining in der Gesundheits-, IT- und Finanzbranche. Im Gesundheitswesen lag der Fokus der Anwendung z.B. auf Behandlungspfaden, ambulanten Prozessen, der Nutzung medizinischer Apps und der Erstellung personalisierter Gesundheitsversorgungsprozesse. In der IT-Branche wurde Process Mining unter anderem bereits zur Prüfung von Softwareprozessen, zur Entdeckung von Geschäftsprozessen in vorhandenen Systemen und zur Erkennung ungewöhnlicher Nutzeraktivitäten in sozialen Netzwerken angewandt. Im Finanzwesen lag der Fokus beispielsweise auf der Analyse von Dokumentenmanagement-Systemen, von Kreditgenehmigungsverfahren und von Prozessen zur Regulierung von Versicherungsansprüchen.

Rudnitckaia (2016, p. 2) nennt darüber hinaus fünf weitere konkrete, reale Anwendungsbeispiele für das Process Mining. Diese sind:

- die Analyse von Behandlungsprozessen in Krankenhäusern,
- die Verbesserung von Prozessen im Kundenservice eines international agierenden Unternehmens,
- das Verstehen des Such- und Navigationsverhaltens der Kunden auf einer Buchungs-Website,
- die Analyse von Fehlfunktionen und Störungen in einem Gepäckabfertigungssystem, und

- die Optimierung der Nutzeroberfläche eines Röntgengerätes.

2.2 Mobilität

Der Begriff Mobilität beschreibt „die Beweglichkeit [von Personen] in mehrfacher Hinsicht“ (Jarass, 2012, p. 19) und kann unterschiedlich gedeutet werden. Es kann unter anderem zwischen der sozialen Mobilität und der räumlichen Mobilität differenziert werden (Jarass, 2012, p. 19; Spode, 2017, p. 24).

Während man unter sozialer Mobilität den „Auf- oder Abstieg innerhalb bzw. entlang gesellschaftlicher Schichten“ (Jarass, 2012, p. 20) (*vertikale* Mobilität (Spode, 2017, p. 24)) sowie einen Funktionswechsel (Jarass, 2012, p. 19; Spode, 2017, p. 24) (*horizontale* Mobilität (Spode, 2017, p. 24)), z.B. im beruflichen Kontext, versteht, spricht die räumliche Mobilität von „der Überwindung von Distanzen im Raum“ (Jarass, 2012, p. 20), also der physischen Fortbewegung von Personen.

Von Bedeutung für diese Arbeit ist die **räumliche Mobilität**, da diese den Positionswechsel von Individuen im geografischen Raum zum Gegenstand hat. Daher beschäftigt sich das folgende Kapitel mit der räumlichen Mobilität. Weiterhin wird eine Definition für den Begriff ‚Mobilitätsprozess‘ vorgeschlagen. Abgeschlossen wird die Thematik ‚*Mobilität*‘ mit einer Gegenüberstellung der Begriffe ‚*Mobilitätsprozess*‘ und ‚*Mobility as a Process*‘.

2.2.1 Räumliche Mobilität

Die räumliche Mobilität, welche synonym zu horizontaler Mobilität (Spode, 2017, pp. 24-25) verwendet werden kann, ist ein Grundbedürfnis (Pfahler, Rubey, & Kördel, 2018, p. 33) der heutigen Gesellschaft und hat einen hohen Stellenwert im Leben eines Menschen, welcher „nach einem idealen Maß an Mobilität“ (Pfahler et al., 2018, p. 33) strebt.

Durch das Hinzuziehen von zeitlichen Dimensionen lässt sich die räumliche Mobilität weiterhin in die punktuelle und die zirkuläre Mobilität unterteilen (Kley, 2016, p. 482).

Die *punktuelle* Mobilität beschreibt räumliche Bewegungen, die in der Regel selten auftreten und zeitlich eng eingegrenzt werden können. Beispiele dafür sind ein Umzug oder auch der Wechsel des Arbeitsplatzes (Kley, 2016, p. 482).

Die *zirkuläre* Mobilität hingegen erfüllt die Voraussetzungen, dass die Bewegungen dieselben Start- und Zielpunkte haben, meistens wiederkehrend auftreten und innerhalb eines überschaubaren, relativ kurzen Zeitrahmens stattfinden. Charakteristisch für die zirkuläre Mobilität ist also, dass es eine Intention des Zurückkehrens an den Anfangspunkt gibt (Kley, 2016, pp. 485-486; Spode, 2017, p. 27). Dieser Definition zufolge kann man also

zeitlich absehbare Reisen sowie das Pendeln zum Arbeitsort mit Übernachtung zu dieser Art der Mobilität zählen (Kley, 2016, p. 485).

Daneben ist auch die *Alltagsmobilität* Bestandteil der zirkulären Mobilität (Kley, 2016, p. 485). Laakso (2017, p. 137) zufolge lassen sich fünf Gründe für diese Form der Mobilität identifizieren: die geschäftliche Mobilität (Weg zur Arbeit), der Weg zur Bildungsstätte oder anderen Betreuungseinrichtungen (z.B. Schule, Kindergarten), der Besuch von Familie und Freunden, das Einkaufen und die Mobilität in der Freizeit.

Für den weiteren Verlauf dieser Arbeit liegt der Fokus auf der **Alltagsmobilität**.

2.2.2 Was sind Mobilitätsprozesse?

Pfahler et al. (2018, pp. 38-41) beleuchteten den Begriff Mobilitätsprozess am Beispiel von körperlich beeinträchtigten Personen, die die Deutsche Bahn nutzen. Demnach beinhaltet ein Mobilitätsprozess nicht nur den reinen Transport von Personen durch den Mobilitätsdienstleister (hier: Deutsche Bahn), sondern auch Schritte, die sowohl davor als auch danach stattfinden. Ein Mobilitätsprozess beschreibt also nach der Definition von Pfahler et al. (2018) alle Etappen vom Reisewunsch und der Planung über das Verlassen des Startpunktes (z.B. der Haustüre) und das Erreichen des Zieles bis hin zu Erfahrungsberichten und möglicher Kontaktaufnahme (in diesem Falle mit der Deutschen Bahn) seitens des Reisenden nach Abschluss der Fahrt.

Hierbei ist jedoch zu beachten, dass der Mobilitätsprozess von körperlich behinderten Personen einen Sonderfall darstellt und sich von einem alltäglichen Mobilitätsprozess unbeeinträchtigter Personen unterscheidet. Dies liegt daran, dass die Planung einer Reise oder einer Fahrt für Menschen mit Behinderung einen im Vergleich größeren Stellenwert hat, da vor allem auf die Barrierefreiheit geachtet werden muss (Pfahler et al., 2018, p. 38). In der alltäglichen Mobilität hingegen werden Wege wiederholt zurückgelegt (Manderscheid, 2021, p. 418), weshalb davon ausgegangen werden kann, dass es sich hierbei um eine Gewohnheit handelt, bei der keine besondere Planung notwendig ist. Aus diesen Gründen wird angenommen, dass der Aspekt der Planung für die Alltagsmobilität von nicht körperlich eingeschränkten Personen zu vernachlässigen ist. Des Weiteren werden auch die Aspekte der Kontaktaufnahme und der Erfahrungsberichte als vernachlässigbar betrachtet, da diese in der alltäglichen Mobilität, bei der gewohnte Strecken zurückgelegt werden, eine untergeordnete Rolle spielen.

Somit erstreckt sich der alltägliche Mobilitätsprozess einer körperlich unbeeinträchtigten Person vom Verlassen des Ausgangspunktes (z.B. Wohnort) bis zum Erreichen des Zielpunktes (z.B. Büro).

Mithilfe der Definition des Begriffes ‚*Weg*‘ lässt sich der Mobilitätsprozess genauer erläutern. Von einem Weg ist die Rede, wenn eine Person eine Positionsveränderungen außerhalb des eigenen Wohnortes durchführt, indem sie eine oder mehrere Fortbewegungsmöglichkeiten (z.B. zu Fuß, Bus, Fahrrad, Auto, etc.) nutzt (Ahrens, Wittwer, Hubrich, Wittig, & Ließke, 2015, p. 16; Nobis & Kuhnimhof, 2018, p. 20). Die Zurücklegung eines Weges dient immer einem Zweck, z.B. in die Arbeit zu kommen (Ahrens et al., 2015, p. 16). Das bedeutet, dass es immer einen Start- und einen Zielpunkt eines Weges gibt. Demnach sind die Rahmenbedingungen eines Weges und eines Mobilitätsprozesses identisch – nämlich die Ortsveränderung einer Person von einem Start- zu einem Zielpunkt.

Um die beiden Begriffe voneinander zu differenzieren, wird für den Mobilitätsprozess folgende Definition vorgeschlagen:

Der Begriff ‚**Mobilitätsprozess**‘ beschreibt die Positionsveränderung einer Person im geografischen Raum von einem Startpunkt A zu einem Zielpunkt B unter Nutzung einer oder mehrerer Fortbewegungsmöglichkeiten.

Demzufolge kann ein Mobilitätsprozess aus einer oder mehreren Etappen bestehen (z.B. vom Wohnort zum Bahnhof laufen, auf den Zug warten, mit dem Zug fahren, vom Bahnhof zur Bushaltestelle laufen, usw.). Die einzelnen Etappen eines Mobilitätsprozesses lassen sich zeitlich klar voneinander abgrenzen.

Ein spezifischer Mobilitätsprozess beschreibt eine immergleiche Abfolge von Etappen.

Zu beachten ist, dass ein Weg durch unterschiedliche Mobilitätsprozesse beschrieben werden kann, da es, je nach Verfügbarkeit der Fortbewegungsmittel, in der Regel mehrere Möglichkeiten gibt, ein und denselben Weg zurückzulegen.

Die Begriffe Weg und Mobilitätsprozess stehen in folgender Beziehung zueinander: Einem Weg können mehrere Mobilitätsprozesse zugeordnet werden, einem Mobilitätsprozess kann jedoch nur ein spezifischer Weg zugeordnet werden.

2.2.3 Mobilitätsprozesse als Bestandteil von ‚Mobility as a Process‘

Wie im vorherigen Kapitel festgestellt wurde, entspricht ein Mobilitätsprozess der reinen Ausführung einer Positionsveränderung im Raum. Diesen Vorgang kann man als Bestandteil eines größeren Konstruktes sehen.

‚*Mobility as a Process*‘ (kurz: MaaP) beschreibt die Abfolge dreier Prozesse, die in der räumlichen Mobilität miteinander auftreten können (s. Abbildung 5). Begonnen wird hierbei mit dem *Vorbereitungsprozess*, bei dem beispielsweise geklärt wird, welche Route wie zurückgelegt werden soll, welche Kosten dadurch voraussichtlich entstehen und ob es Einschränkungen gibt, gefolgt von der Buchung von z.B. Mobilitätsdienstleistungen. Auf die Vorbereitung folgt der Positionswechsel, sprich die Ausführung des *Mobilitätsprozesses*. Nach Abschluss des Mobilitätsprozesses folgen die *After-Mobility-Services*, also alles, was nach einer Reise oder Fahrt passiert. So kann man hier beispielsweise die Schlüssel des Mietwagens zurückgeben, die tatsächlich entstandenen Kosten kalkulieren und offene Rechnungen begleichen.

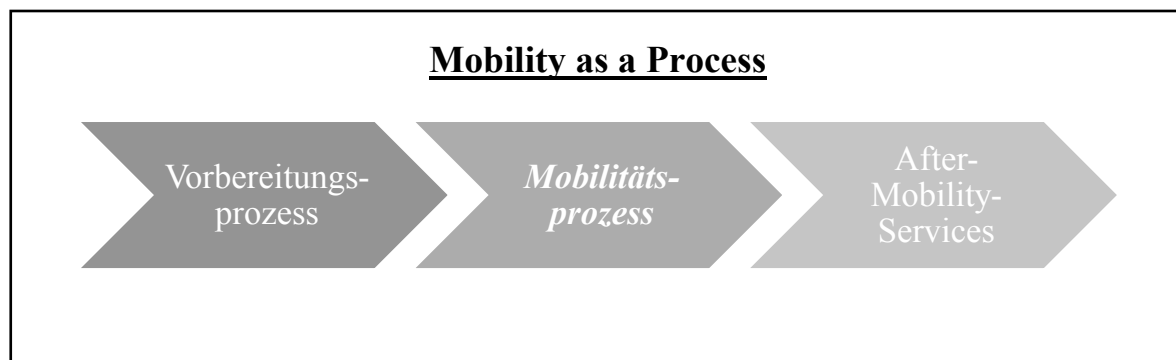


Abbildung 5: *Mobility as a Process*. Eigene Darstellung (2022) nach Vogt, *Mobility Services and Fleet Management*, Vorlesungsskript 2, Folie 11

2.3 Anwendung von Process Mining auf Mobilitätsprozesse

Moderne Systeme - seien dies nun GPS, Funk oder WLAN – ermöglichen es, eine Vielzahl an Mobilitäts- und Standortdaten zu sammeln. Die Omnipräsenz solcher Daten erlaubt es, Routen und Mobilitätsprozesse von Personen oder Fortbewegungsmitteln zu verstehen. An dieser Stelle könnte Process Mining theoretisch dafür eingesetzt werden, Mobilitätsprozesse zu analysieren und optimieren.

Zunächst stellt sich also die Frage, ob sich in der Literatur bereits mit diesem Thema befasst wurde. Aus diesem Grund soll nachfolgend anhand einer Fachliteraturrecherche geklärt werden, ob der Process Mining Ansatz in wissenschaftlichen Veröffentlichungen bereits auf Mobilitätsprozessdaten angewandt wurde.

Weiterhin wird Google Maps als mögliche Quelle von Mobilitätsprozessdaten für die Nutzung von Process Mining-Techniken vorgestellt.

Abschließend werden Kennzahlen zur Bewertung der erfassten Mobilitätsprozesse definiert und Hypothesen über die Anwendung von Process Mining-Techniken auf Mobilitätsprozesse aufgestellt.

2.3.1 Anwendung von Process Mining auf Mobilitätsprozesse in der Literatur

Um die Fragestellung, ob Process Mining in der Vergangenheit in wissenschaftlichen Schriften auf Mobilitätsprozesse angewandt wurde, zu beantworten, wurde eine Literaturrecherche durchgeführt. Durchsucht wurden dabei die wissenschaftlichen Datenbanken ProQuest, Oxford Academic, IEEE Xplore und ScienceDirect.

Die Suchanfragen enthielten immer den Begriff Process Mining in unterschiedlichen Schreibweisen (Process Mining, Process-Mining und „Process Mining“) in Verbindung mit den Begriffen Mobilitätsprozess, Alltagsmobilität, Transportprozess, räumliche Mobilität, ÖPNV, everyday mobility, public transport, travel und trip. Die Begriffe, die aus zwei Wörtern bestanden, wurden dabei ebenfalls anhand unterschiedlicher Schreibweisen (mit Bindestrich und mit Anführungszeichen) durchsucht.

Nachdem die Suche ohne zusätzliche Filter abgeschlossen war, wurden weiterhin die Filter „Abstract“/„Kurzfassung“ bzw. „Title, Abstract, Keywords“ sowie „Index Terms“ verwendet, um diejenigen Suchanfragen, die mehr als 200 Ergebnisse hervorbrachten, weiter einzugrenzen. Da unter Verwendung des Filters „Index Terms“ mehrere Tausend Ergebnisse vorgeschlagen wurden, wurde in den restlichen Suchen auf diesen Filter verzichtet. Trotz des Einsatzes von Filtern konnte keine neue relevante Literatur gefunden werden.

In Tabelle 20 (s. Anhang) wurden alle Suchanfragen sowie die Gesamtergebnisse und die relevanten Ergebnisse dokumentiert.

Um die relevanten Ergebnisse aus den Gesamtergebnissen herauszufiltern, wurden zunächst die Titel und die Kurzfassungen (Abstract) der Ergebnisse im Gesamtkontext betrachtet und auf den Begriff ‚*Process Mining*‘ hin untersucht. Dadurch ergaben sich insgesamt sieben unterschiedliche Ergebnisse, die als relevant erachtet wurden. Davon stammt eine Publikation aus der Datenbank ProQuest und jeweils drei aus den Datenbanken IEEE Xplore und ScienceDirect.

Im weiteren Verlauf wurden die relevanten Schriften auf die Frage untersucht, ob diese tatsächlich von der Anwendung des Process Mining auf Mobilitätsprozesse handeln.

2.3.1.1 Untersuchung der relevanten Literatur

Ribeiro, Fontes, Soares, und Borges (2020) behandelten in ihrer Publikation „*Process discovery on geolocation data*“ den Einsatz von Process Discovery-Techniken auf geografische Daten am Beispiel von Busrouten eines ÖPNV-Unternehmens mit dem Ziel, Erkenntnisse über Transportprozesse, Geschäftsabläufe und Servicequalität zu gewinnen (Ribeiro et al., 2020, pp. 139-146). Dabei wurde zwischen stationären (Standort ändert sich nicht) und nicht-stationären (Standort ändert sich) Events unterschieden. Im Falle des Busunternehmens beschrieben stationäre Events z.B. die Wartung der Fahrzeuge oder das Tanken (Ribeiro et al., 2020, pp. 140-141).

Die Arbeit kam unter anderem zu dem Entschluss, dass sich die Anwendung von Process Mining auf geografische Daten besonders bei gut strukturierten Abläufen eignet. Außerdem sollte die Aufzeichnung der Daten in einer hohen zeitlichen Frequenz erfolgen, da durch zu niedrige Frequenzen die Genauigkeit der Event-Daten nicht gegeben sei. So könnten z.B. der Beginn bzw. das Ende eines Events durch eine zu niedrige Frequenz nicht akkurat aufgezeichnet oder einzelne Events gänzlich ignoriert werden (Ribeiro et al., 2020, p. 145)

Aufbauend darauf hatte die Arbeit „*Detection of vehicle-based operations from geolocation data*“ von Tavares, Ribeiro, und Fontes (2022) zum Ziel, fahrzeugbasierte, stationäre Vorgänge wie das Tanken, das Anhalten (z.B. an einer Bushaltestelle), Pannen oder Staus anhand von Positionsdaten zu entdecken (Tavares et al., 2022, p. 342). Hierfür wurde eine Fallstudie am Beispiel eines öffentlichen Verkehrsnetzes in Rio de Janeiro durchgeführt (Tavares et al., 2022, p. 346).

Yamaguchi, Terada, Manaskasemsak, Rungsawang, und Leelaprute (2017) verwendeten Process Mining in ihrer Arbeit „*Tour Miner: Mining System of Tour Plans from SNS – Smelting Function from Travel Records to Tour Routes*“, um vorher aus sozialen Netzwerken gesammelte Routeninformationen in Prozessmodelle umzuwandeln. Diese konnten dann von einem weiteren Algorithmus weiterverarbeitet werden, um Tour-Routen zu erstellen (Yamaguchi et al., 2017, pp. 239-240).

Auch die Publikation „*An Interest-based Tour Planning Tool by Process Mining from Twitter*“ von Gou et al. (2016) nutzte Process Mining, um aus Routen von Nutzern der Social Media Plattform ‚Twitter‘, die anhand von Schlüsselbegriffen bestimmten Interessensgebieten zugeordnet wurden, interessenbasierte Tour-Routen zu erstellen (Gou et al., 2016, pp. 1-4).

Blank, Maurer, Siebenhofer, Rogge-Solti, und Schonig (2016) untersuchten in ihrer Arbeit „*Location-Aware Path Alignment in Process Mining*“ die Wege von Kunden, die sich in Gruppen in einem Einkaufsgeschäft bewegten (Blank et al., 2016, pp. 1-8).

Die Forschungsarbeit „*Detecting train reroutings with process mining*“ von Janssenswillen, Depaire, und Verboven (2018) handelte unter anderem davon, Informationen aus Zugregistriereinrichtungen des belgischen Bahnnetzes mithilfe von Process Mining zu analysieren, um Abweichungen der tatsächlichen Route der Züge von der geplanten Route zu entdecken (Janssenswillen et al., 2018, pp. 2,4).

Mit der Entdeckung von alltäglichen Fortbewegungsmustern einzelner oder mehrerer Nutzer durch Process Mining beschäftigte sich die Arbeit „*Discovering commute patterns via process mining*“ von Yousfi und Weske (2019, pp. 691-693, 703). Dadurch konnten Einblicke darüber gewonnen werden, wo sich die Individuen am häufigsten aufhielten, welche Routen sie mit welchen Fortbewegungsmitteln zurücklegten und wie lange sie für einzelnen Routen unterwegs waren (Yousfi & Weske, 2019, p. 692). Die Autoren kamen zu dem Schluss, dass Process Mining die wichtigen Orte und die Fortbewegungsmittel mit hoher Genauigkeit erkannt hat, während die Routen und die Dauer mit geringerer Exaktheit erkannt wurden (Yousfi & Weske, 2019, pp. 710-711).

2.3.1.2 Schlussfolgerung der Literaturrecherche

Es lässt sich schlussfolgern, dass alle der sieben relevanten Schriften den Process Mining Ansatz auf standortbezogene Daten angewandt haben. Allerdings legte allein die Arbeit von Yousfi und Weske (2019) den Fokus auf die Entdeckung von alltäglichen Fortbewegungsprozessen von Individuen. Hierbei untersuchten sie jedoch nicht einzelne Mobilitätsprozesse wie sie in dieser Arbeit definiert wurden, sondern die ganzheitlichen Fortbewegungsmuster – also alle Wege, die zurückgelegt wurden – der Teilnehmer.

Man kann ableiten, dass die Anwendung von Process Mining auf standortbezogene Daten in der Literatur bisher nur wenig erforscht wurde und die Forschung zu diesem Thema somit erst am Anfang steht. Zu der Anwendung von Process Mining auf Mobilitätsprozesse gemäß der Definition dieser Arbeit konnten keine wissenschaftlichen Publikationen gefunden werden.

2.3.2 Aufzeichnung von Mobilitätsdaten durch Google Maps

In der heutigen Zeit erfreuen sich Smartphones einer großen Beliebtheit und werden tagtäglich von einer Vielzahl von Menschen genutzt (Shi et al., 2017, p. 1). Die mobilen Geräte eignen sich unter anderem dafür, menschliche Mobilitätsmuster aufzuzeichnen. Das

hat den Grund, dass Smartphones bei vielen Nutzern stark in den Alltag integriert sind und vor allem die Fähigkeit besitzen, durch einen eingebauten GPS-Sensor den genauen Standpunkt des Gerätes zu ermitteln (Chon, Talipov, Shin, & Cha, 2014, p. 512).

Die bekannteste Applikation, die sich dieser Technik bedient, ist *Google Maps* (Goudarzi, 2018, p. 1404; Sonkin, Alpert, & Jaffe, 2020, p. 1398). Google Maps kam bereits im Jahre 2005 auf den Markt (He, Chow, & Zhang, 2019, p. 2612). Dabei handelt es sich um einen webbasierten Karten-Service, welcher unter anderem Landkarten, Stadtpläne, Satellitenbilder und Echtzeit-Informationen zur Verkehrslage bereitstellt und als Navigationssystem genutzt werden kann, indem es Routen für Fußgänger, Fahrradfahrer, Autofahrer, die Nutzung öffentlicher Verkehrsmittel usw. vorschlägt (Baji, Salini, Sandra, Thomas, & Ravikumar, 2017, p. 52; Goudarzi, 2018, p. 1404).

Anhand von Google Maps ist es dem Nutzer folglich möglich, Routen zu finden und die erwartete Dauer für die Zurücklegung dieser zu erhalten. Um die ungefähre Dauer bestimmter Routen zu berechnen, bedient sich Google Maps sowohl historischer als auch Echtzeit-Daten. Letztere kommen zustande, indem Echtzeit-GPS-Informationen von Smartphones an Google übermittelt werden (He et al., 2019, p. 2612).

Als Nutzer hat man zudem die Möglichkeit, manuell die Google Maps-Funktion ‚*Google Maps Timeline*‘ zu aktivieren. Ist diese aktiviert, zeichnet Google Maps den Standort des Gerätes kontinuierlich auf und erkennt dabei, auf welche Weise sich der Nutzer fortbewegt (z.B. zu Fuß, mit dem Auto, mit der Bahn, etc.) (Sonkin et al., 2020, p. 1398). So entstehen detaillierte Informationen darüber, zu welchen Uhrzeiten sich eine Person an den von ihr besuchten Orten aufgehalten hat und welche Routen sie zurückgelegt hat. Wie bei Google üblich, werden die aufgezeichneten Informationen im Google-Konto des Nutzers gespeichert (Abdulrazaq et al., 2020, p. 2301).

Die Daten aus der Google Maps Timeline können als Keyhole Markup Language (KML)-Dateien exportiert werden (Sardianos, Varlamis, & Bouras, 2018, p. 693; Sonkin et al., 2020, p. 1398). Diese enthalten unter anderem die Koordinaten der besuchten Orte (Sonkin et al., 2020, p. 1398). Die KML-Dateien können wiederum in Browserprogrammen wie Google Maps oder Google Earth geöffnet werden (Sardianos et al., 2018, p. 693).

2.3.3 Definition von Kennzahlen zur Bewertung von Mobilitätsprozessen

Bergsmann (2011, p. 227) beschreibt Prozesskennzahlen (engl. Key Performance Indicator – kurz: KPI) als Mittel zum Zweck, um die Erreichung der Ziele, die für einen Prozess verfolgt werden, sicherzustellen. Die KPIs werden in der Regel dazu genutzt, Prozesse zu

bewerten, Optimierungspotenziale aufzudecken, Prozesse zu verbessern und interne und externe Prozesse miteinander zu vergleichen (Prozessbenchmarking) (Kahl & Zimmer, 2017, p. 77).

Wesentlich ist, dass KPIs speziell auf den jeweiligen Prozess abgestimmt sein sollten, um die konkreten Ziele des Prozesses erreichen und deren Fortschritt überprüfen zu können (Bergsmann, 2011, p. 228; Krstev & Krneta, 2022, p. 1; Nanz, 2012, p. 211). Demnach ist die Nutzung gängiger sowie zu vieler Prozesskennzahlen zu vermeiden. Besser ist es, nur wenige, jedoch für den Prozess wichtige KPIs zu definieren (Bergsmann, 2011, p. 228).

Google Maps sammelt und stellt nur messbare Informationen zur Dauer und zur zurückgelegten Entfernung eines Mobilitätsprozesses bereit (genaue Beschreibung der bereitgestellten Daten in Kapitel 3.1). Daher lassen sich weder Aussagen zu den Kosten oder der Nachhaltigkeit der Mobilitätsprozesse noch zu der Termintreue von Mobilitätsdienstleistern treffen.

In Anbetracht dessen lassen sich die folgenden Kennzahlen identifizieren:

- Verhältnis der tatsächlich gebrauchten Zeit zur kürzesten Dauer für dieselbe Route
$$(V) = \frac{\text{tatsächliche Dauer}}{\text{kürzeste Dauer für diese Route}}, \text{ mit } V=1 \text{ als exakte Übereinstimmung der tatsächlichen und der kürzesten Dauer, } V<1 \text{ tatsächliche Dauer schneller als die kürzeste Dauer und } V>1 \text{ tatsächliche Dauer länger als die kürzeste Dauer.}$$
- Durchschnittsgeschwindigkeit in km/h (G) = $\frac{\text{zurückgelegte Entfernung in m}}{\text{Dauer in s}} \times 3,6.$
- Verhältnis zwischen Durchlaufzeit eines Cases und durchschnittlicher Durchlaufzeit
$$(D) = \frac{\text{Durchlaufzeit Case}}{\text{durchschnittliche Durchlaufzeit}}, \text{ mit } D=1 \text{ als Übereinstimmung der tatsächlichen und der durchschnittlichen Durchlaufzeit, } D<1 \text{ Dauer des Cases kürzer als die durchschnittliche Durchlaufzeit und } D>1 \text{ Dauer des Cases länger als die Durchschnittsdauer.}$$

2.3.4 Hypothesen zur Bewertung von Mobilitätsprozessen

Wendet man Process Mining-Techniken auf Mobilitätsprozessdaten an, werden die Mobilitätsprozesse, die bei der Zurücklegung eines bestimmten Weges aufgetreten sind, von der Software entdeckt und ausgewertet. Je nachdem, über welchen Zeitraum - also über wie viele Tage (Anzahl der Tage = Anzahl der Cases) - die Datensammlung stattgefunden hat, unterscheidet sich die Menge an auswertbaren Daten. Daher liegt die Vermutung nahe, dass die Ergebnisse der Auswertung durch Process Mining umso genauer und umfangreicher werden, je mehr Cases in die Analyse mit einfließen. Dies würde sich folglich im durch

Process Mining-Techniken entdeckten Prozessmodell bemerkbar machen. Dadurch ergibt sich die folgende Behauptung:

H1. Je mehr Cases ein Event-Log enthält, desto umfangreicher ist das entdeckte Prozessmodell.

Durch die Anwendung von Process Mining auf Mobilitätsdaten können außerdem die unterschiedlichen Mobilitätsprozesse (Varianten), welche in einem Event-Log aufgetreten sind, entdeckt werden. Diese können sich durch die Anzahl ihrer Prozessschritte (Ereignisse) unterscheiden. Dies liegt in der Regel an der Auswahl der Fortbewegungsmöglichkeiten, welche für die Zurücklegung einer Strecke genutzt werden. Da es sich bei den verschiedenen Varianten immer um denselben Weg, jedoch mit unter Umständen unterschiedlich vielen Ereignissen, handelt, liegt die Vermutung nahe, dass sich die Varianten anhand ihrer Dauer unterscheiden. Dies führt zu der folgenden Annahme:

H2. Je mehr Ereignisse eine Variante enthält, desto länger dauert diese durchschnittlich im Vergleich zu Varianten mit einer geringeren Anzahl an Ereignissen.

Ebenso kann es vorkommen, dass einige Varianten nur ein einziges Mal in dem gesamten Event-Log aufgetreten sind und somit eher eine Ausnahme darstellen. Geht man von einem proportionalen Verlauf aus, kann folgende Hypothese angenommen werden:

H3. Je mehr Varianten in einem Event-Log auftreten, desto höher ist der Anteil der Varianten, welche nur einmal aufgetreten sind.

3. Methodik

Um die Forschungsfrage der vorliegenden Arbeit zu beantworten, wurde mit dem Mixed-Methods-Ansatz gearbeitet und dabei induktiv vorgegangen. Zunächst erfolgte eine qualitativ durchgeführte Fallstudie zum Fall *„Die Mobilitätsprozesse einer Person von oder zu ihrer Arbeitsstätte bzw. von oder zum Arbeitsort“*. Untersucht wurden vier unterschiedliche Fälle.

Die im Rahmen der Fallstudie erhobenen Mobilitätsdaten wurden aus Google Maps gewonnen und anschließend in Excel so bearbeitet, dass diese einen Event-Log gebildet haben, welcher durch das Process Mining-Tool *„Disco“* genutzt werden konnte. Dieses wurde gewählt, um zu veranschaulichen, welche Informationen Personen ohne spezielle Vorkenntnisse aus der Analyse ihrer Mobilitätsprozessdaten mithilfe eines kommerziellen, benutzerfreundlichen Tools gewinnen können.

Zur Beantwortung der ersten Teilfrage wurde anhand einer qualitativen Inhaltsanalyse analysiert, welche Informationen das Tool *„Disco“* über die untersuchten Mobilitätsprozesse bereitstellt und welche Möglichkeiten das Tool bietet, um mit diesen Informationen zu arbeiten.

Die aus dem Tool gewonnenen Informationen waren sowohl qualitativer als auch quantitativer Natur. Letztere wurden genutzt, um Prozesskennzahlen zu berechnen, Filter anzuwenden und die aufgestellten Hypothesen zu überprüfen. Die Durchführung dieser Schritte diente der Beantwortung der Fragestellung, wie die Informationen aus dem Tool genutzt werden können, um die Mobilitätsprozesse zu bewerten.

3.1 Erhebung der Mobilitätsdaten

Es wurden insgesamt vier konkrete Fälle untersucht. Bei dreien dieser Fälle handelte es sich um den täglichen Weg zum Arbeitsplatz. Ein Fall stellte die wöchentliche Pendelstrecke einer Person zwischen zwei Orten dar. Erhoben wurden bereits vorhandene Mobilitätsdaten, die mittels der Funktion *„Google Maps Timeline“* aufgezeichnet wurden. Anhand dieser kann man auf einer Landkarte all seine zurückgelegten Strecken eines Tages sehen. Es wurden die Tage, an denen der zu betrachtende Weg zurückgelegt wurde, ausgewählt und die Aufzeichnungen tageweise als KML-Datei exportiert. Dabei variierte der Zeitraum der exportierten Daten der Teilnehmer je nach Verfügbarkeit, welche beispielsweise von zeitlich begrenzten Praktika oder einem Wohnortwechsel beeinflusst wurde. Außerdem konnten nicht immer die Mobilitätsprozesse aller Tage exportiert werden, da Google Maps an einigen Tagen ungenaue Aufzeichnungen bereitstellte und die Daten somit nicht nutzbar gewesen

wären. Dies machte sich dadurch bemerkbar, dass anstelle einer Route lediglich eine gerade Linie vom Start- zum Zielpunkt führte. Schlussendlich variierte die Menge der Cases der einzelnen Fälle zwischen einer Anzahl von 20 bis 55 Cases.

Da die Daten im KML-Format nicht von Process Mining-Tools verarbeitet werden können, mussten diese zuallererst in das CSV-Format konvertiert werden. Hierfür wurde sich der Webseite www.convertcsv.com/kml-to-csv.htm bedient. Abbildung 6 zeigt die in Excel geöffnete Datei des siebten Tages von Fall 1 nach dem Konvertieren in das entsprechende Format. Sensible Informationen wurden zur Wahrung der Anonymität in dieser sowie in den folgenden Abbildungen zensiert.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	name,address,ExtendedData/Data/0/value,ExtendedData/Data/0/_name,ExtendedData/Data/1/value,ExtendedData/Data/1/_name,ExtendedData/Data/2/value,ExtendedData/Data/2/_name,descript														
2															
3	Driving,		Email,Driving,Category,5690,Distance,Driving from 2021-03-25T06:11:43.999Z to 2021-03-25T06:25:39.999Z. Distance 5690m,2021-03-25T06:11:43.999Z,2021-03-25T												
4	Walking,		Email,Walking,Category,739,Distance,Walking from 2021-03-25T06:25:39.999Z to 2021-03-25T06:36:43.010Z. Distance 739m,2021-03-25T06:25:39.999Z,2021-03-25T												
5															
6	Walking,		Email,Walking,Category,691,Distance,Walking from 2021-03-25T14:45:02.091Z to 2021-03-25T14:56:14.963Z. Distance 691m,2021-03-25T14:45:02.091Z,2021-03-25T												
7	Driving,		Email,Driving,Category,5467,Distance,Driving from 2021-03-25T14:56:14.963Z to 2021-03-25T15:12:08.999Z. Distance 5467m,2021-03-25T14:56:14.963Z,2021-03-25T												
8															
9	Driving,		Email,Driving,Category,2694,Distance,Driving from 2021-03-25T16:41:44.000Z to 2021-03-25T16:49:54.999Z. Distance 2694m,2021-03-25T16:41:44.000Z,2021-03-25T												
10	Aral,														
11	Driving,		Email,Driving,Category,148414,Distance,Driving from 2021-03-25T16:56:01.000Z to 2021-03-25T18:45:11.000Z. Distance 148414m,2021-03-25T16:56:01.000Z,2021-03-												
12	Ulm,														
13	Driving,		Email,Driving,Category,14379,Distance,Driving from 2021-03-25T19:13:58.907Z to 2021-03-25T19:35:35.999Z. Distance 14379m,2021-03-25T19:13:58.907Z,2021-03-2												
14	Zuhause														
15															

Abbildung 6: unbearbeitete CSV-Datei. Eigene Darstellung (2022)

Um die Daten für das Process Mining nutzbar zu machen, mussten diese in dem Microsoft-Programm Excel weiterverarbeitet werden. Hierfür wurde wie folgt vorgegangen:

1. Ausgefüllte Zellen aus Spalte A markieren,
2. Unter dem Reiter ‚Daten‘ das Feld ‚Text in Spalten‘ auswählen,
3. Dateityp ‚Getrennt‘ festlegen und auf ‚Weiter‘ (s. Anhang, Abbildung 48),
4. Trennzeichen ‚Komma‘ festlegen und ‚Fertig stellen‘ (s. Anhang, Abbildung 49).

Daraufhin wurden die Informationen übersichtlich in Spalten sortiert (s. Abbildung 7). Betrachtet man diese, wird ersichtlich, welche Informationen aus Google Maps exportiert wurden. Diese sind unter anderem die Bezeichnung des Ortes oder der Aktivität (z.B. Gehen, Fahren, etc.), die Adresse der besuchten Orte, die E-Mail-Adresse des Nutzers, die Kategorie des Ortes oder der Fortbewegungsart (z.B. Gebäude, Parkplatz, etc.), die zurückgelegte Entfernung in Metern, eine Beschreibung, die Start- und Schlusszeit, der Längen- und Breitengrad, die Höhe und die Koordinaten.

	A	B	C	D	E	F	G	H	
1	name	address	ExtendedData/Data/0/value	ExtendedData/Data/0/_name	ExtendedData/Data/1/value	ExtendedData/Data/1/_name	ExtendedData/Data/2/value	ExtendedData/Data/2/_name	dr
2				Email	Gebäude	Category	0 Distance		G
3	Driving			Email	Driving	Category	5690 Distance		D
4	Walking			Email	Walking	Category	739 Distance		W
5				Email	Automobilwerk	Category	0 Distance		A
6	Walking			Email	Walking	Category	691 Distance		W
7	Driving			Email	Driving	Category	5467 Distance		D
8				Email	Gebäude	Category	0 Distance		G
9	Driving			Email	Driving	Category	2694 Distance		D
10	Aral			Email	Tankstelle	Category	0 Distance		T
11	Driving			Email	Driving	Category	148414 Distance		D
12	Ulm			Email		Category	0 Distance		f
13	Driving			Email	Driving	Category	14379 Distance		D
14	Zuhause			Email	Gebäude	Category	0 Distance		G
15									

Abbildung 7: In Spalten sortierte Google Maps Timeline-Daten. Eigene Darstellung (2022)

Da nur ein bestimmter Abschnitt des Tages, nämlich der Weg zur Arbeitsstätte, betrachtet werden sollte, wurden an dieser Stelle alle Zeilen, die nicht zu diesem Abschnitt gehörten, gelöscht. Übrig blieben alle Aktionen und Orte zwischen einschließlich dem Start- und dem Zielort (s. Abbildung 8).

	A	B	C	D	E	F	G	H	
1	name	address	ExtendedData/Data/0/value	ExtendedData/Data/0/_name	ExtendedData/Data/1/value	ExtendedData/Data/1/_name	ExtendedData/Data/2/value	ExtendedData/Data/2/_name	dr
2				Email	Gebäude	Category	0 Distance		G
3	Driving			Email	Driving	Category	5690 Distance		D
4	Walking			Email	Walking	Category	739 Distance		W
5				Email	Automobilwerk	Category	0 Distance		A
6									
7									

Abbildung 8: betrachteter Mobilitätsprozess in Excel. Eigene Darstellung (2022)

Weiterhin wurden die besuchten Orte vereinheitlicht und anonymisiert. Das bedeutet, man änderte die Bezeichnung (s. Abbildung 9, Spalte E) des jeweiligen Ortes so, dass dieser für jeden durchlaufenen Mobilitätsprozess identisch war. Der Startort, in der Regel der Wohnort des Nutzers, erhielt in diesem Fall die Bezeichnung ‚Start‘ und der Zielort/Arbeitsort die Bezeichnung ‚Ziel‘. Wurden auf dem Weg weitere Orte, wie z.B. ein Parkhaus oder eine Bäckerei, besucht, erhielten diese die Bezeichnung ‚Gebäude C‘, ‚Gebäude D‘ usw. Wichtig zu beachten war hierbei, dass jedem spezifischen Ort auch nur eine spezifische Bezeichnung zugeordnet wird.

	A	B	C	D	E	F	G	H	
1	name	address	ExtendedData/Data/0/value	ExtendedData/Data/0/_name	ExtendedData/Data/1/value	ExtendedData/Data/1/_name	ExtendedData/Data/2/value	ExtendedData/Data/2/_name	desc
2				Email	Start	Category	0 Distance		Geb
3	Driving			Email	Driving	Category	5690 Distance		Driv
4	Walking			Email	Walking	Category	739 Distance		Wali
5				Email	Ziel	Category	0 Distance		Auto
6									

Abbildung 9: Anonymisierung und Vereinheitlichung der Orte. Eigene Darstellung (2022)

Außerdem wurden die Fortbewegungsarten (z.B. Driving, Walking) im Falle von unterschiedlich zurückgelegten Strecken voneinander differenziert (s. Abbildung 10). Dies wurde überprüft, indem die KML-Dateien der jeweiligen Tage in Google Maps geöffnet und die genauen Routen betrachtet wurden. Traten also beispielsweise insgesamt zwei unterschiedliche Mobilitätsprozesse auf, um mit dem Auto von einem Start- zu einem Endpunkt zu gelangen, so bezeichnete man die erste Strecke als ‚Driving A‘ und die zweite Strecke als ‚Driving B‘. Dasselbe galt für alle anderen Fortbewegungsarten.

Dies hatte den Vorteil, dass später im Process Mining-Tool unterschieden werden konnte, wie viele unterschiedliche Mobilitätsprozesse (Varianten) genutzt wurden, um den Weg zurückzulegen.

	A	B	C	D	E	F	G	H	
1	name	address	ExtendedData/Data/0/value	ExtendedData/Data/0/_name	ExtendedData/Data/1/value	ExtendedData/Data/1/_name	ExtendedData/Data/2/value	ExtendedData/Data/2/_name	desc
2				Email	Start	Category	0 Distance		Gebäude
3	Driving			Email	Driving A	Category	5690 Distance		Driving
4	Walking			Email	Walking	Category	739 Distance		Walking
5				Email	Ziel	Category	0 Distance		Auto

Abbildung 10: Unterscheidung der Fortbewegungsarten (Spalte E). Eigene Darstellung (2022)

Im nächsten Schritt wurde die Startzeit des Startortes gelöscht und durch die Schlusszeit ersetzt, sodass sowohl im Startzeit-Feld als auch im Schlusszeit-Feld derselbe Wert eingetragen war (s. Abbildung 11, wichtig: das Feld durfte nicht leer bleiben, da es sonst im Tool eine Fehlermeldung gegeben hätte). Analog dazu wurde die Schlusszeit des Zielortes gelöscht und durch die Startzeit ersetzt. Andernfalls würde das Process Mining-Programm an diesen Stellen einen zeitlichen Engpass identifizieren, da in der Regel mehrere Stunden an diesen Orten verbracht wurden. Für den Mobilitätsprozess waren jedoch nur die Zeiten von Bedeutung, zu welchen man den Startort verlassen bzw. den Zielort erreicht hat.

	I	J	K	L	M	N	O	P
1	description	TimeSpan/begin	TimeSpan/end	longitude	latitude	altitude	LineString/altitudeMode	LineString/extension
2	Gebäude von 2021-03-24T18:20:14.439Z to 2021-03-25T06:11:43.999Z. Distance 0m	2021-03-25T06:11:43.999Z	2021-03-25T06:11:43.999Z		114.486.472	48.769.095.199.999.900	0	
3	Driving from 2021-03-25T06:11:43.999Z to 2021-03-25T06:25:39.999Z. Distance 5690m	2021-03-25T06:11:43.999Z	2021-03-25T06:25:39.999Z				clampToGround	
4	Walking from 2021-03-25T06:25:39.999Z to 2021-03-25T06:36:43.010Z. Distance 739m	2021-03-25T06:25:39.999Z	2021-03-25T06:36:43.010Z				clampToGround	
5	Automobilwerk von 2021-03-25T06:36:43.010Z to 2021-03-25T14:45:02.091Z. Distance 0m	2021-03-25T06:36:43.010Z	2021-03-25T14:45:02.091Z	11.410.447.699.999.900	487.919.966	0		

Abbildung 11: aktualisierter Start-Zeitpunkt für ‚Start‘. Eigene Darstellung (2022)

Der zu diesem Zeitpunkt aktuelle Stand der Datei war allerdings noch immer nicht für die Anwendung von Process Mining geeignet. Da ein Event-Log immer eine Case ID benötigt, musste diese manuell hinzugefügt werden. Hierfür wurde eine neue Spalte mit dem Titel ‚ID‘ eröffnet (s. Abbildung 12, Spalte A). Die Werte, die in die Spalte eingetragen wurden, waren pro Datei für alle Aktivitäten identisch. Handelte es sich z.B. um den ersten Tag des betrachteten Zeitraumes, erhielten alle Zeilen die ID-Nummer 1. So ging es chronologisch weiter. Der 15. Tag bekam dementsprechend die ID-Nummer 15.

	A	B	C	D	E	F	G	H	I
1	ID	name	address	ExtendedData/Data/0/value	ExtendedData/Data/0/_name	ExtendedData/Data/1/value	ExtendedData/Data/1/_name	ExtendedData/Data/2/value	ExtendedData/Data/2/_name
2	1				Email	Start	Category	0 Distance	Gebäude
3	1	Driving			Email	Driving A	Category	5690 Distance	Driving
4	1	Walking			Email	Walking	Category	739 Distance	Walking
5	1				Email	Ziel	Category	0 Distance	Auto

Abbildung 12: Hinzufügen einer Case ID. Eigene Darstellung (2022)

Hatte man die CSV-Dateien aller Tage so bearbeitet, konnten die Informationen in chronologischer Reihenfolge in einer neuen Excel-Tabelle zusammengefügt werden. Dadurch erhielt man ein für Process Mining nutzbares Ereignisprotokoll (s. Abbildung 13).

ID	genauer Ort oder Aktivität	Adresse	ExtendedData/Data/0/value	ExtendedData/Data/0/_name	Gebäude/Aktivität	ExtendedData/Data/1/_name	Entfernung in m
1	1			Email	Start	Category	0 Distance
2	1			Email	Driving A	Category	5658 Distance
3	1			Email	Gebäude C	Category	0 Distance
4	1			Email	Walking	Category	474 Distance
5	1			Email	Ziel	Category	0 Distance
6	2			Email	Start	Category	0 Distance
7	2			Email	Driving A	Category	5725 Distance
8	2			Email	Walking	Category	659 Distance
9	2			Email	Ziel	Category	0 Distance
10	3			Email	Start	Category	0 Distance
11	3			Email	Driving A	Category	5734 Distance
12	3			Email	Gebäude C	Category	0 Distance
13	3			Email	Walking	Category	490 Distance
14	3			Email	Ziel	Category	0 Distance
15	4			Email	Start	Category	0 Distance
16	4			Email	Driving A	Category	5610 Distance
17	4			Email	Gebäude C	Category	0 Distance
18	4			Email	Walking	Category	613 Distance
19	4			Email	Ziel	Category	0 Distance
20	5			Email	Start	Category	0 Distance
21	5			Email	Driving A	Category	5362 Distance
22	5			Email	Walking	Category	613 Distance
23	5			Email	Ziel	Category	0 Distance

Abbildung 13: auswertbarer Event-Log. Eigene Darstellung (2022)

Jedoch waren nicht alle Spalten für die Anwendung mit Process Mining-Techniken von Bedeutung. Der Übersichtlichkeit und der Anonymität der Teilnehmer wegen wurden unwichtige Spalten gelöscht.

Abschließend enthielten die Daten folgende Attribute (s. Abbildung 14):

- Eine **Case ID**, welche chronologisch den jeweiligen einzelnen Vorgängen zugeordnet wurde,
- Eine **Aktivität** bzw. ein besuchter Ort, wobei jede/r Aktivität/Ort eine spezifische Bezeichnung erhielt,
- Die **Entfernung in Metern**, die angab, wie weit die zurückgelegte Strecke der einzelnen Aktivitäten war,
- Ein **Beginn-Zeitstempel**, welcher das genaue Datum und die Uhrzeit enthielt, zu welcher eine Aktivität begonnen hat, und
- Ein **Schluss-Zeitstempel**, welcher das Datum und die Uhrzeit enthielt, zu welcher eine Aktivität abgeschlossen wurde.

	A	B	C	D	E
1	Case ID	Gebäude/Aktivität	Entfernung in m	Beginn	Ende
2	1	Start	0	2021-03-16T06:20:03.000Z	2021-03-16T06:20:03.000Z
3	1	Driving A	5658	2021-03-16T06:20:03.999Z	2021-03-16T06:32:48.000Z
4	1	Gebäude C	0	2021-03-16T06:32:48.000Z	2021-03-16T06:33:50.000Z
5	1	Walking A	474	2021-03-16T06:33:50.000Z	2021-03-16T06:44:43.416Z
6	1	Ziel	0	2021-03-16T06:44:43.416Z	2021-03-16T06:44:43.416Z
7	2	Start	0	2021-03-17T06:19:16.999Z	2021-03-17T06:19:16.999Z
8	2	Driving A	5725	2021-03-17T06:19:16.999Z	2021-03-17T06:36:03.999Z
9	2	Walking A	659	2021-03-17T06:36:03.999Z	2021-03-17T06:45:46.355Z
10	2	Ziel	0	2021-03-17T06:45:46.355Z	2021-03-17T06:45:46.355Z
11	3	Start	0	2021-03-18T06:17:38.000Z	2021-03-18T06:17:38.000Z
12	3	Driving A	5734	2021-03-18T06:17:38.000Z	2021-03-18T06:32:54.000Z

Abbildung 14: Endgültiger Event-Log. Eigene Darstellung (2022)

Das hier beschriebene Vorgehen wurde für jeden einzelnen Case aller vier Fälle wiederholt.

3.2 Import der Event-Logs in das Tool ‚Disco‘

Das vorher erstellte Ereignisprotokoll konnte anschließend vom Process Mining-Tool ‚Disco‘ ausgewertet werden. Das Tool akzeptierte die Dateien sowohl im CSV- als auch im Excel-Format.

Um die Event-Log-Datei zu importieren, öffnete man zunächst ‚Disco‘ und wählte dann unter dem Punkt ‚Load your own data‘ die Schaltfläche ‚Open file...‘ aus (s. Abbildung 15). Anschließend wählte man die benötigte Datei aus und öffnete diese.

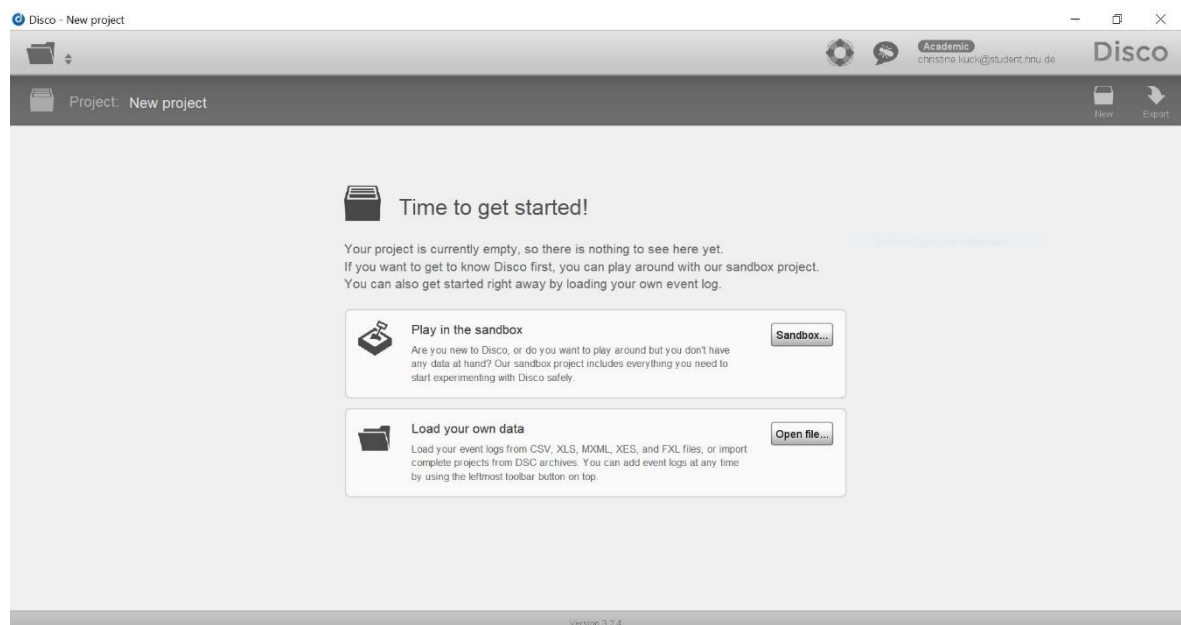


Abbildung 15: Startseite des Process Mining-Tools ‚Disco‘. Eigene Darstellung (2022)

Es öffnete sich ein Fenster, in dem der komplette Datensatz zu sehen war. Dort konnte man auswählen, welche Spalten in der Auswertung berücksichtigt werden sollen und diese deklarieren, sodass das Tool wusste, wie es die jeweiligen Spalten zu behandeln hat.

Die Spalten wurden folgendermaßen deklariert: die Spalte ‚Case ID‘ als *Case*, ‚Gebäude/Aktivität‘ als *Activity*, ‚Beginn‘ und ‚Ende‘ als *Timestamps* und die Spalte ‚Entfernung in m‘ als *Other* (s. Abbildung 16).

Case ID	Gebäude/Aktivität	Entfernung in m	Beginn	Ende
1	Gebäude A	0	2021-03-16T06:20:03.000Z	2021-03-16T06:20:03.000Z
2	Driving A	5658	2021-03-16T06:20:03.999Z	2021-03-16T06:32:48.000Z
3	Gebäude C	0	2021-03-16T06:32:48.000Z	2021-03-16T06:33:50.000Z
4	Walking A	474	2021-03-16T06:33:50.000Z	2021-03-16T06:44:43.416Z
5	Gebäude B	0	2021-03-16T06:44:43.416Z	2021-03-16T15:41:05.017Z
6	Gebäude A	0	2021-03-17T06:19:16.999Z	2021-03-17T06:19:16.999Z
7	Driving A	5725	2021-03-17T06:19:16.999Z	2021-03-17T06:36:03.999Z
8	Walking A	659	2021-03-17T06:36:03.999Z	2021-03-17T06:45:46.355Z
9	Gebäude B	0	2021-03-17T06:45:46.355Z	2021-03-17T15:29:19.510Z
10	Gebäude A	0	2021-03-18T06:17:38.000Z	2021-03-18T06:17:38.000Z
11	Driving A	5734	2021-03-18T06:17:38.000Z	2021-03-18T06:32:54.000Z
12	Gebäude C	0	2021-03-18T06:32:54.000Z	2021-03-18T06:37:45.336Z
13	Walking A	490	2021-03-18T06:37:45.336Z	2021-03-18T06:44:25.507Z
14	Gebäude B	0	2021-03-18T06:44:25.507Z	2021-03-18T15:39:23.504Z
15	Gebäude A	0	2021-03-19T06:15:06.000Z	2021-03-19T06:15:06.000Z
16	Driving A	5610	2021-03-19T06:15:06.000Z	2021-03-19T06:28:27.999Z
17	Gebäude C	0	2021-03-19T06:28:27.999Z	2021-03-19T06:31:39.999Z
18	Walking A	613	2021-03-19T06:31:39.999Z	2021-03-19T06:39:36.000Z
19	Gebäude B	0	2021-03-19T06:39:36.000Z	2021-03-19T13:07:43.000Z
20	Gebäude A	0	2021-03-23T06:16:10.999Z	2021-03-23T06:16:10.999Z
21	Driving A	5362	2021-03-23T06:16:10.999Z	2021-03-23T06:29:50.000Z
22	Walking A	839	2021-03-23T06:29:50.000Z	2021-03-23T06:41:21.279Z
23	Gebäude B	0	2021-03-23T06:41:21.279Z	2021-03-23T15:50:36.229Z
24	Gebäude A	0	2021-03-24T06:18:38.999Z	2021-03-24T06:18:38.999Z

Abbildung 16: Deklaration der Spalten in ‚Disco‘. Eigene Darstellung (2022)

Im Anschluss wählte man die Schaltfläche ‚Start import‘ in der unteren rechten Ecke des Fensters aus.

Daraufhin erschien ein automatisch erzeugtes Prozessmodell, welches die tatsächlichen Mobilitätsprozesse anhand der historischen Informationen aus dem Event-Log visualisierte (s. Abbildung 17).

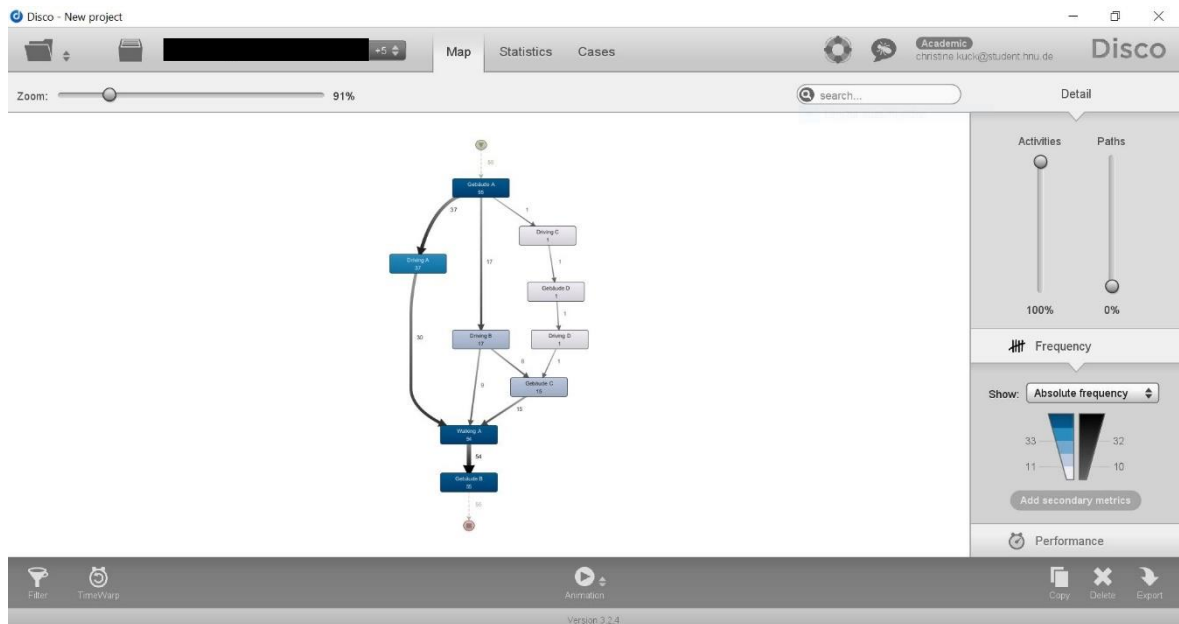


Abbildung 17: Darstellung des Mobilitätsprozesses direkt nach dem Import. Eigene Darstellung (2022)

Weiterhin wurde derselbe Event-Log erneut importiert, allerdings mit der Unterscheidung, dass die Spalte mit der Schlusszeit nicht als Timestamp deklariert, sondern ignoriert wurde. Dies ermöglichte ergänzende Informationen zu der Performance des Prozesses, welche in dem Prozessmodell, bei dem die Schlusszeit berücksichtigt wurde, nicht ersichtlich geworden sind (vgl. Anhang, Abbildung 50).

Um das vollständige Prozessmodell betrachten zu können, welches unter anderem für die Prüfung der ersten Hypothese von Bedeutung war, wurden zuallererst die Schieberegler für ‚Activities‘ und ‚Paths‘ auf 100% eingestellt. So ließen sich ohne weiteren Aufwand bereits das vollständige Prozessmodell sowie zahlreiche Informationen über die betrachteten Mobilitätsprozesse direkt aus dem Tool ablesen.

3.3 Vorgehen bei der Auswertung der Event-Logs

Nachdem die jeweiligen Ereignisprotokolle, wie in den vorherigen Kapiteln beschrieben, erstellt und in das Process Mining-Tool ‚Disco‘ importiert wurden, konnte mit der Auswertung begonnen werden.

Zur Beantwortung der ersten Teilfrage wurde sich zunächst ein Überblick über alle Möglichkeiten des Tools und über die Informationen, welche das Tool über die Event-Logs preisgibt, verschafft.

Anhand des gewonnenen Überblickes konnte festgestellt werden, welche Informationen für die Berechnung der in Kapitel 2.3.3 eingeführten Prozesskennzahlen sowie für die

Überprüfung der aufgestellten Hypothesen von Bedeutung sind. Diese Daten wurden dann für jeden der betrachteten vier Fälle gesammelt und zusammengefasst.

Daraufhin wurden die KPIs berechnet. Die Prozesskennzahl V steht für das Verhältnis der tatsächlichen Dauer für die Zurücklegung eines Weges zur kürzesten Dauer für die Zurücklegung desselben Weges. Hier wurden die Mobilitätsprozesse (Varianten), die in dem jeweiligen Event-Log aufgetreten sind, betrachtet. Es wurde die Durchschnittsdauer der jeweiligen Variante als tatsächlich benötigte Zeit verwendet. Die kürzeste Dauer wurde ermittelt, indem dieselbe Route mit denselben Fortbewegungsarten in Google Maps eingetragen wurde. Abschließend wurde der Wert für die tatsächlich benötigte Zeit durch den Wert für die kürzeste Dauer dividiert.

Die KPI G gibt die durchschnittliche Geschwindigkeit in km/h an. Hierbei wurde jeder Case einzeln betrachtet. Zuerst wurde die zurückgelegte Strecke in Metern ermittelt. Diese ließ sich aus dem Tool für jede Aktivität ablesen und wurde anschließend für jeden Case addiert. Daraufhin wurde die Dauer der jeweiligen Cases in Sekunden umgerechnet. Abschließend wurden die Werte in die Gleichung übernommen und die Durchschnittsgeschwindigkeit ermittelt.

Weiterhin beschreibt die Kennzahl D das Verhältnis der benötigten Dauer eines Cases zur Durchschnittsdauer aller Cases. Letztere wurde in Minuten angegeben und ließ sich aus dem Process Mining-Tool entnehmen. Auch die Durchlaufzeit der einzelnen Cases wurde hier in Minuten umgerechnet. Um ein Ergebnis für D zu erhalten, wurde der Wert für die Durchlaufzeit des Cases durch den Durchschnittswert dividiert.

Im nächsten Schritt wurden für jeden Event-Log Filter innerhalb des Process Mining-Tools angewandt, um die folgenden Fragen zu klären:

- „Welche waren die kürzesten Cases?“,
- „Wie viele Cases waren kürzer als die durchschnittliche Durchlaufzeit?“, und
- „Bei wie vielen Fällen traten Wartezeiten auf?“.

Abschließend wurde überprüft, ob die aufgestellten Hypothesen zutreffend sind.

Die Berechnung der Prozesskennzahlen sowie die Anwendung von Filtern und die Prüfung der Hypothesen dienten der Beantwortung der zweiten Teilfrage.

In Abbildung 18 wurde der in dieser Arbeit angewandte Datenerhebungs- und Auswertungsprozess, welcher in den vorangegangenen Kapiteln beschrieben wurde, grafisch dargestellt.

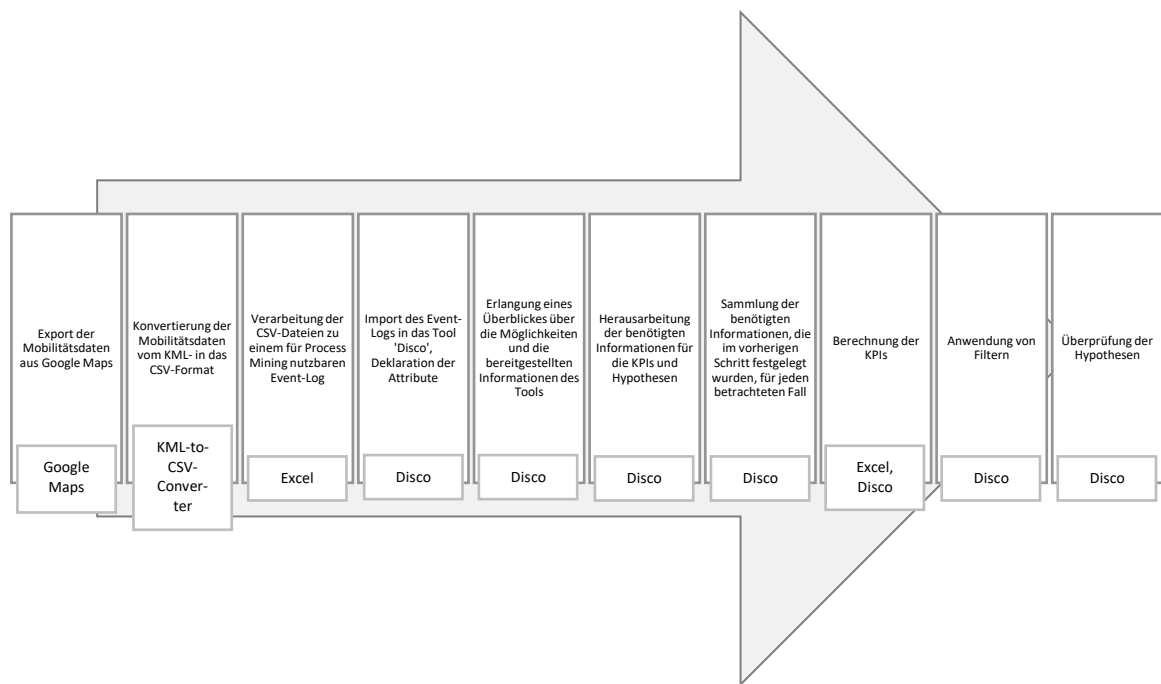


Abbildung 18: Prozess der Erhebung und Auswertung der Mobilitätsdaten inkl. genutzter Tools. Eigene Darstellung (2022)

Die methodische Vorgehensweise dieser Arbeit ähnelt dem Cross Industry Standard Process for Data Mining (CRISP-DM)-Modell (Mertens, 2019, p. 27). Dieses beschreibt eine seit den 1990er Jahren (Saltz, 2021, p. 2337) verwendete Methode für Data-Mining Projekte. Das CRISP-DM-Modell besteht aus den folgenden sechs iterativen Phasen:

1. Geschäftsmodell verstehen (Business Understanding),
2. Daten verstehen (Data Understanding),
3. Daten aufbereiten (Data Preparation),
4. Modellierung (Modeling),
5. Evaluation, und
6. Einsatz (Deployment) (Mertens, 2019, p. 27; Saltz, 2021, p. 2337).

In Abbildung 19 wird das CRISP-DM-Modell grafisch veranschaulicht.



Abbildung 19: CRISP-DM-Modell, Quelle: Wuttke (2020)

Die ersten beiden Prozessphasen wurden in dieser Arbeit durch die Schaffung eines theoretischen Rahmens sowie die Erhebung der Rohdaten repräsentiert. Die Phase ‚Data Preparation‘ beschreibt den hier stattgefundenen Prozess der Verarbeitung der Daten zu einem von Process Mining nutzbaren Ereignisprotokoll. Die vierte Phase stellt die Anwendung der Process Mining-Techniken auf die Event-Logs dar. Unter den Prozessschritt ‚Evaluation‘ fallen die Sammlung und die Weiterverarbeitung der Informationen aus dem Tool sowie die abschließende Interpretation der Ergebnisse. Einzig die Phase ‚Deployment‘ fand hier keine Anwendung.

3.4 Erfüllung der Gütekriterien der Forschung

Die Gütekriterien der qualitativen Forschung sind erfüllt, da die einzelnen Schritte der Forschung transparent dargelegt wurden. Außerdem ist die Reichweite gegeben, da die verwendeten Mobilitätsdaten aus der Google Maps Timeline exportiert wurden und aus diesem Grund unabhängig von der Person immer dieselben Attribute enthalten. Die erneute Anwendung des hier beschriebenen Verfahrens würde deshalb ähnliche Ergebnisse erzielen. Ebenso ist das Gütekriterium der Intersubjektivität erfüllt, da die Ergebnisse offen diskutiert wurden.

Weiterhin sind auch die Gütekriterien der quantitativen Forschung gegeben. Die Forschung ist valide, da alle Berechnungen mithilfe vorgegebener Formeln durchgeführt wurden, deren Ergebnisse dem Forschungsinteresse gerecht wurden. Die Messungen wurden zudem immer mit denselben Formeln und Attributen durchgeführt, wodurch die Reliabilität gegeben ist. Außerdem erfüllte die quantitative Forschung das Gütekriterium der Objektivität, da sich an die errechneten Ergebnisse gehalten wurde.

4. Ergebnisse

Bei der Durchführung der in Kapitel 3 genannten Schritte konnten zahlreiche Ergebnisse gesammelt werden. Dieses Kapitel beschreibt zunächst das Tool ‚Disco‘ und nennt, welche Informationen aus dem Tool für die Berechnung der Prozesskennzahlen, die Prüfung der Hypothesen sowie die Anwendung von Filtern notwendig waren. Anschließend werden die Ergebnisse aller Fälle gesammelt, die Prozesskennzahlen berechnet, Filter angewandt und die Hypothesen überprüft.

4.1 Überblick über das Process Mining Tool ‚Disco‘

Nach dem Import des Event-Logs erschien das von der Software entdeckte Prozessmodell. Dieses ließ sich nach Wunsch feiner oder gröber darstellen, indem man die Schieberegler für ‚Activities‘ (dt. Aktivitäten) und ‚Paths‘ (dt. Pfade) auf der rechten Seite des Fensters eingestellt hat. Zu Beginn waren die Aktivitäten bei 100% und die Pfade bei 0% eingestellt. Stellte man zweiteres ebenfalls auf 100%, wurden (unter Umständen) weitere Pfade sichtbar. Bei dieser Einstellung wurden alle Aktivitäten und Pfade, die in dem Event-Log aufgetreten sind, dargestellt, sodass man ein vollständiges Prozessmodell mit allen Informationen erhalten hat (s. Abbildung 20).

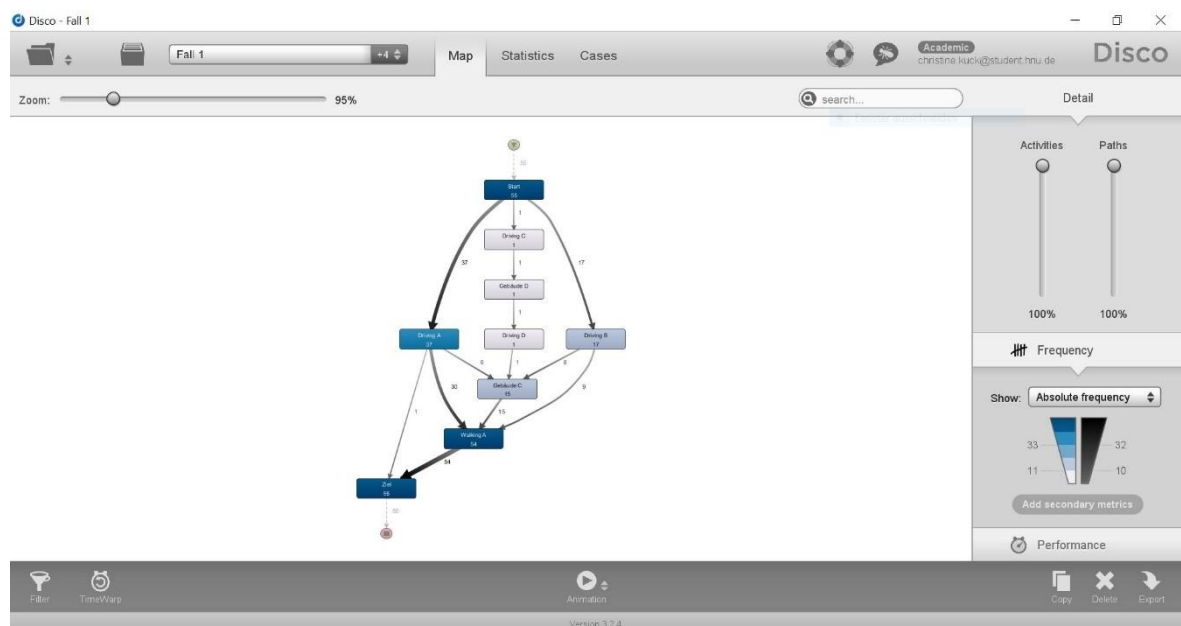


Abbildung 20: Darstellung des Mobilitätsprozesses mit 100%iger Genauigkeit. Eigene Darstellung (2022)

Stellte man hingegen beide Regler auf 0%, wurden nur der direkte Pfad und nur die Aktivitäten angezeigt, die in der am häufigsten aufgetretenen Variante vorgekommen sind (s. Abbildung 21). Es wurde also die Abfolge von Prozessschritten angezeigt, welche am häufigsten aufgetreten ist.

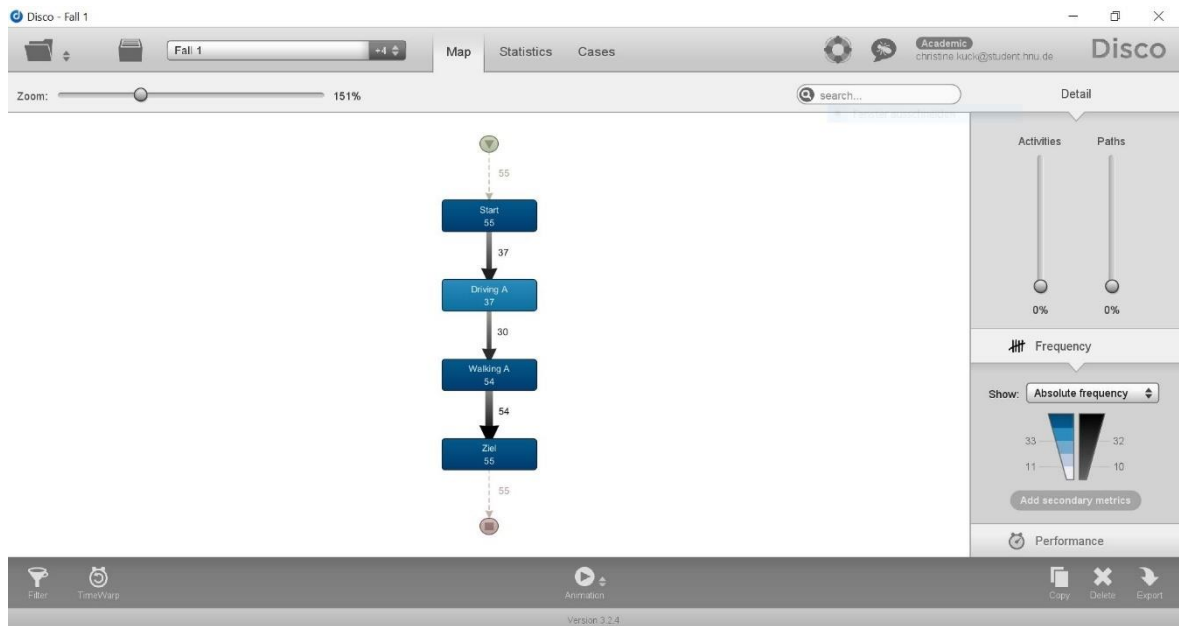


Abbildung 21: Darstellung des Mobilitätsprozesses mit 0% bei Activities und Paths. Eigene Darstellung (2022)

Die Zahlen in den Aktivitätsfeldern bzw. neben den Pfaden zeigten an, wie oft die jeweilige Aktivität oder der jeweilige Pfad insgesamt eingetreten sind (*Absolute frequency*). Unter dem Reiter ‚Frequency‘ ließen sich jedoch weitere Informationen auswählen, die anstelle der oder zusätzlich zu den Informationen in den Feldern bzw. neben den Pfeilen angezeigt werden sollten. Neben der ‚Absolute frequency‘ konnte man aus der ‚Case frequency‘ (in wie vielen Cases/Fällen ist diese Aktivität/dieser Pfad aufgetreten?), den ‚Max. repetitions‘ (wie häufig kamen die jeweiligen Aktivitäten/Pfade maximal pro Case vor?) und der ‚Case coverage‘ (in wie viel Prozent der Cases kamen die jeweiligen Aktivitäten/Pfade vor?) wählen (s. Anhang, Abbildung 51).

Des Weiteren ließen sich unter dem Reiter ‚Performance‘ Angaben zu der Dauer der einzelnen Aktivitäten (wenn Start- und Schlusszeit als Timestamps deklariert wurden) oder der Pfade (wenn nur Beginn als Timestamp deklariert wurde) anzeigen, welche die Dauer zwischen zwei Aktivitäten beschrieben. Hier konnte man zwischen der ‚Total duration‘ (insgesamte Dauer für alle Cases), ‚Median duration‘ (mediane Dauer), ‚Mean duration‘ (durchschnittliche Dauer), ‚Max. duration‘ (maximale Dauer) und der ‚Min. duration‘ (minimale Dauer) wählen (s. Anhang, Abbildung 52).

Sowohl unter ‚Frequency‘ als auch unter ‚Performance‘ war es möglich, eine weitere Metrik auszuwählen, sodass im Prozessmodell zwei Werte gleichzeitig betrachtet werden konnten.

Die Farben der Aktivitäten sowie die Breite der Pfade in den Prozessmodellen deuteten zudem darauf hin, wie häufig die jeweiligen Aktivitäten bzw. Pfade aufgetreten sind oder wie lange diese gedauert haben.

Durch Auswählen einer Aktivität oder eines Pfades im Prozessmodell erhielt man außerdem einen Überblick über alle Performance- und Frequency-Werte der jeweiligen Aktivität bzw. des jeweiligen Pfades (s. Abbildung 22).

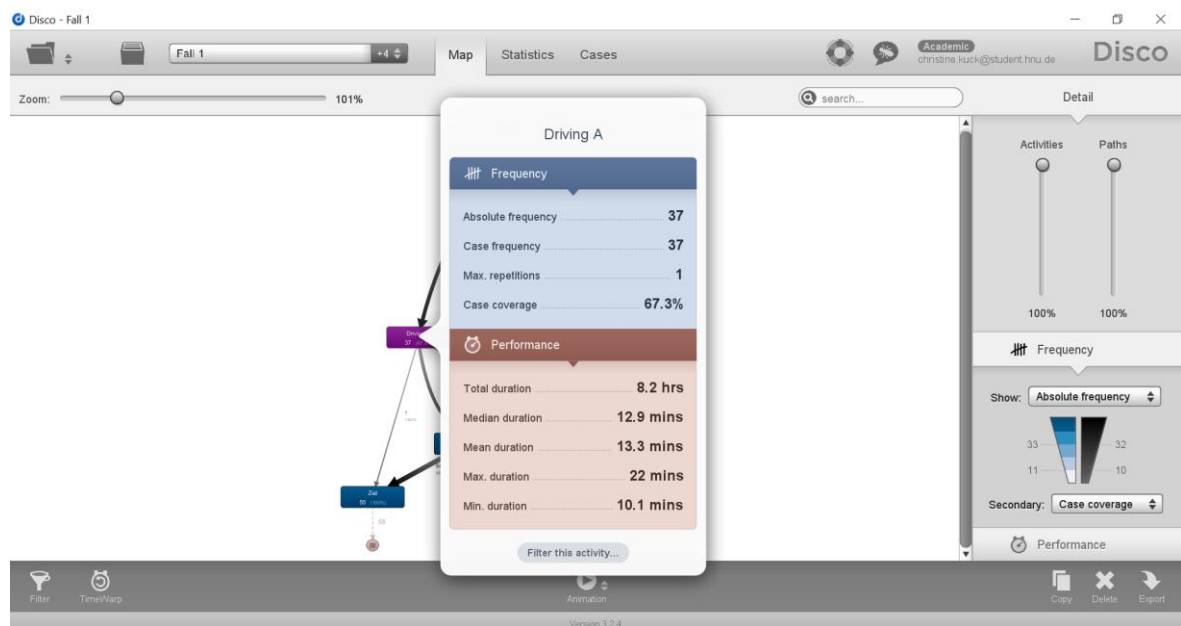


Abbildung 22: Überblick über alle Werte einer Aktivität/eines Pfades. Eigene Darstellung (2022)

Weiterhin bestand die Möglichkeit, Geschäftszeiten sowie Urlaubszeiten einzustellen (s. Anhang, Abbildung 53).

Ein weiteres Feature des Tools war, dass man sich eine Animation des Event-Logs anzeigen lassen konnte. Abbildung 23 stellt einen Ausschnitt aus einer solchen Animation des ersten Falles dar.

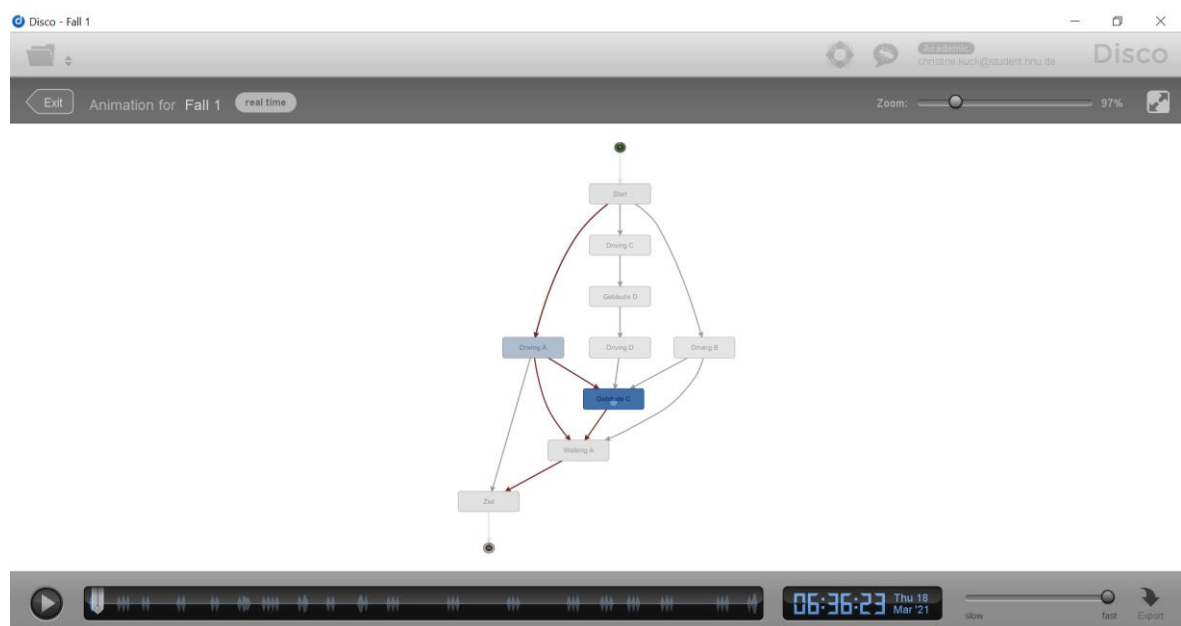


Abbildung 23: Ausschnitt aus einer Animation. Eigene Darstellung (2022)

Unter dem Reiter ‚Statistics‘ konnte man sich unter dem Punkt ‚Overview‘ einen Überblick über diverse Statistiken verschaffen. Diese waren: die Ereignisse verteilt über die Zeitspanne des Event-Logs (‚Events over time‘), die aktiven Cases während der Gesamtdauer des Event-Logs (‚Active cases over time‘), die Anzahl sowie der kumulierte Anteil der Varianten (‚Case variants‘), die Anzahl der Ereignisse pro Case (‚Events per case‘), die Dauer der einzelnen Cases (‚Case duration‘), die Anzahl der Cases, welche die Zeit zu einem gewissen Prozentsatz aktiv genutzt haben (‚Case utilization‘), die durchschnittliche Dauer der Aktivitäten (‚Mean activity duration‘) und die Anzahl der Cases, bei denen eine gewisse Wartezeit eingetreten ist (‚Mean waiting time‘).

Zudem erhielt man auf einen Blick wichtige Informationen über die Anzahl der Ereignisse, Cases und Aktivitäten, die mediane und die durchschnittliche Dauer der Cases, sowie das Start- und das Enddatum inklusive Uhrzeiten (s. Abbildung 24).

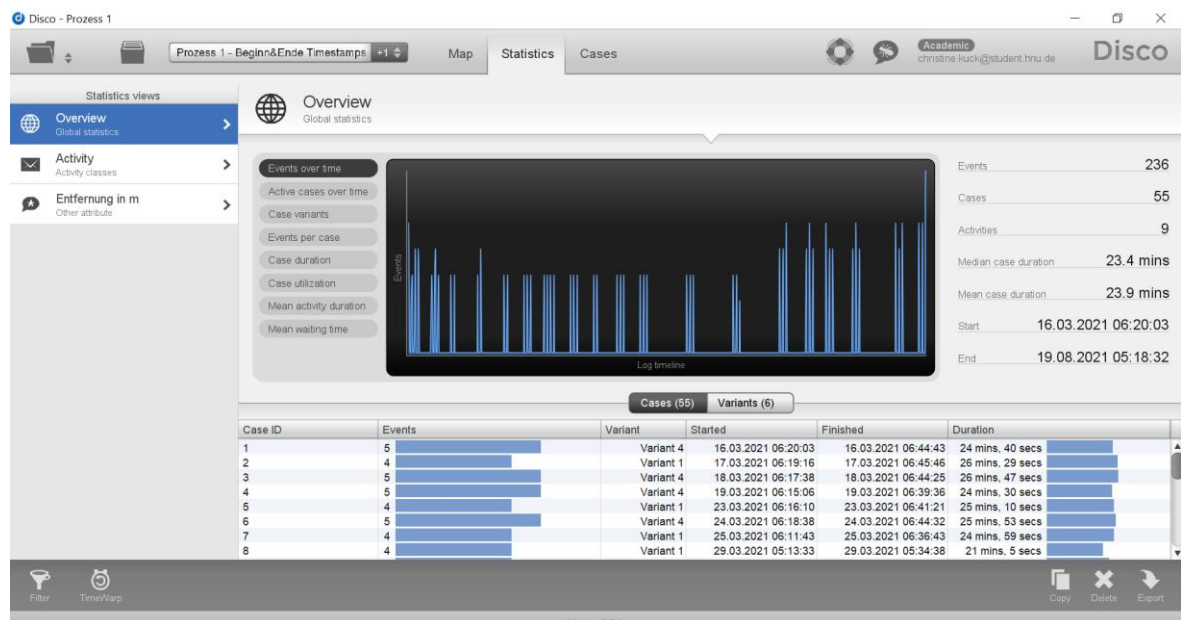


Abbildung 24: Statistiken zum Event-Log in ‚Disco‘. Eigene Darstellung (2022)

Weiterhin ließen sich in der unteren Hälfte des Fensters die beiden Tabellen ‚Cases‘ (s. Abbildung 24) und ‚Variants‘ (s. Abbildung 25) betrachten. Erstere informierte den Nutzer über die Anzahl der Ereignisse, die zugehörige Variante, den Start- und Schluss-Zeitstempel sowie die Dauer der einzelnen Cases. Die Tabelle ‚Variants‘ hingegen enthielt die Anzahl der Cases und Ereignisse sowie die mediane und die durchschnittliche Dauer der einzelnen Varianten.

Cases (55) Variants (6)					
Variant	▲ Cases	Events	Median duration		Mean duration
Variant 1	30	4	22 mins, 27 secs		22 mins, 40 secs
Variant 2	9	4	24 mins, 12 secs		24 mins, 54 secs
Variant 3	8	5	23 mins, 56 secs		24 mins, 22 secs
Variant 4	6	5	24 mins, 35 secs		24 mins, 50 secs
Variant 5	1	3	22 mins, 1 sec		22 mins, 1 sec
Variant 6	1	7	45 mins, 40 secs		45 mins, 40 secs

Abbildung 25: Tabelle ‚Variants‘. Eigene Darstellung (2022)

Unter dem Punkt ‚Activity‘ fand man ebenfalls zahlreiche Statistiken, Informationen und Tabellen in Bezug auf die Aktivitäten (s. Abbildung 26). Auf den ersten Blick konnte man erfahren, wie viele unterschiedliche Aktivitäten aufgetreten sind, welche die minimale, mediane, durchschnittliche und maximale Häufigkeit der Aktivitäten in dem Event-Log waren, sowie die Standardabweichung der Häufigkeit der Aktivitäten.

Statistiken, welche man sich unter ‚Activity‘ anzeigen lassen konnte, waren die Häufigkeit und der kumulierte Anteil der Aktivitäten (‚Frequency‘), die mediane Dauer der Aktivitäten (‚Median duration‘), die durchschnittliche Dauer (‚Mean duration‘), der zeitliche Spielraum, der bei den einzelnen Aktivitäten aufgetreten ist (‚Duration range‘) und die Gesamtdauer der einzelnen Aktivitäten (‚Aggregate duration‘).

Die Tabelle ‚All activities‘ enthielt die absolute sowie die relative Häufigkeit der Aktivitäten, die mediane und die durchschnittliche Dauer, genauso wie den zeitlichen Spielraum der Aktivitäten. Die anderen beiden Tabellen ‚First in case‘ und ‚Last in case‘ zeigten nur die Aktivität(en) an, welche zu Beginn bzw. zum Schluss der Cases aufgetreten ist/sind.

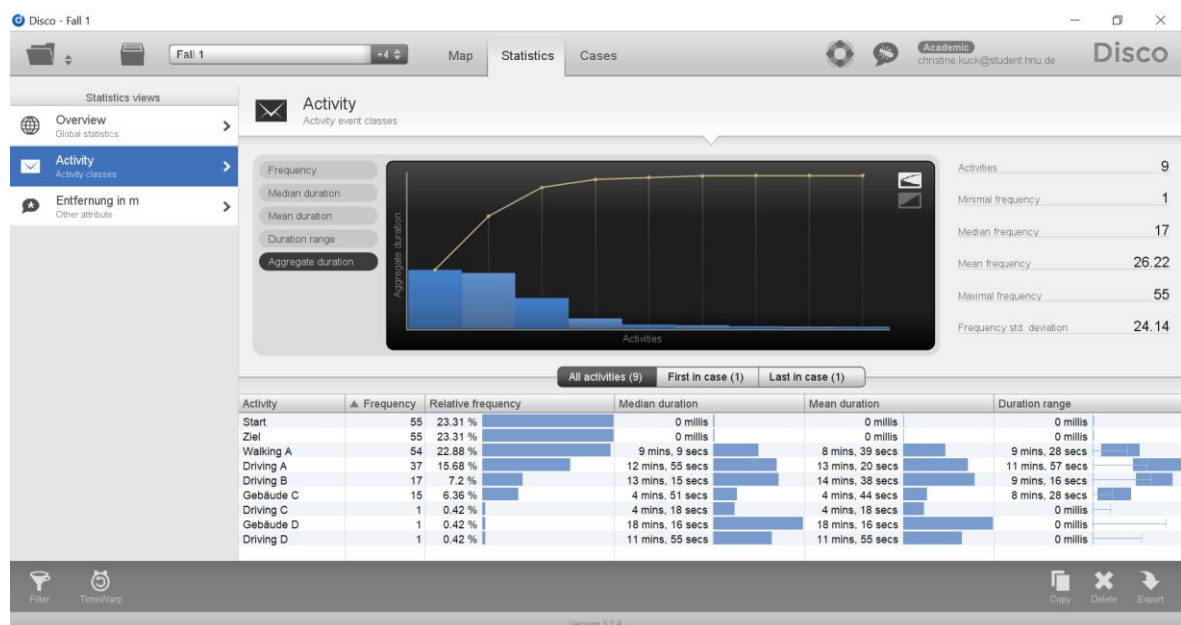


Abbildung 26: Übersicht über den Punkt ‚Activity‘. Eigene Darstellung (2022)

Hat man bei der Deklaration des Event-Logs außerdem sonstige Attribute bestimmt, konnte man sich ebenfalls unter dem Reiter ‚Statistics‘ Informationen über das jeweilige Attribut anzeigen lassen. Dort konnte man sehen, wie schon bei den Aktivitäten, wie viele unterschiedliche Werte des Attributes aufgetreten sind, was dessen minimale, durchschnittliche und maximale Häufigkeit war, sowie die Standardabweichung. Daneben konnte man einen Graphen betrachten, welcher anzeigte, wie häufig die jeweiligen Werte vorgekommen sind. Auch hier gab es eine Tabelle ‚All values‘ welche die absolute und die relative Häufigkeit der aufgetretenen Werte enthielt. Die Tabellen ‚First in case‘ und ‚Last in case‘ zeigten zudem an, welche Werte zu Beginn bzw. zum Schluss der Cases vorgekommen sind (s. Abbildung 27).

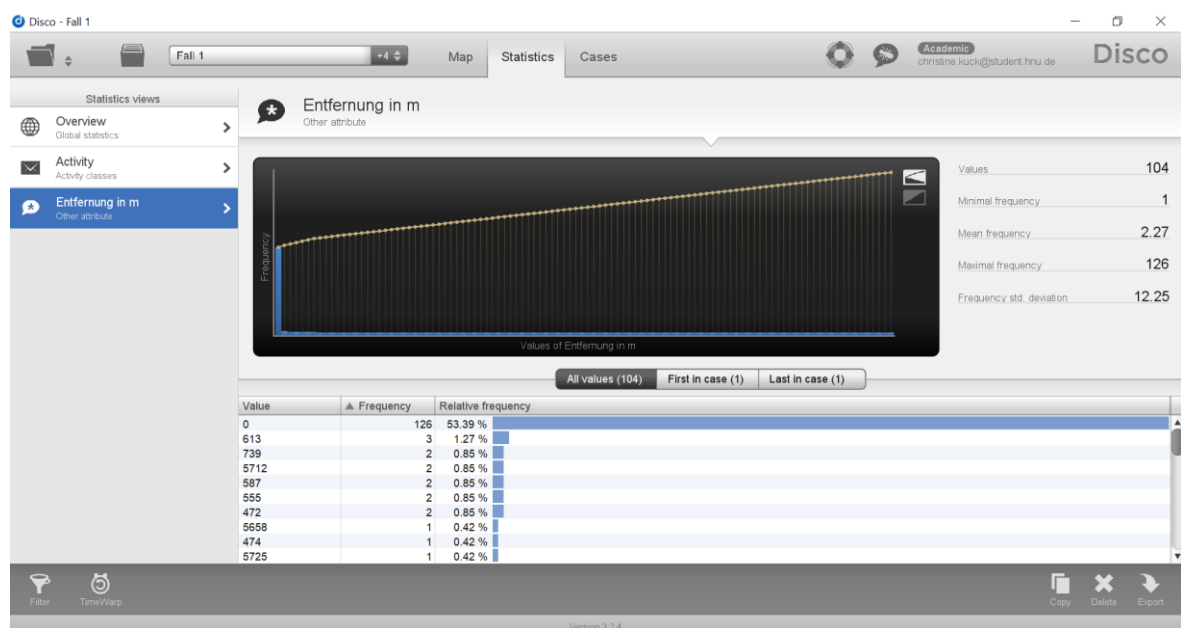


Abbildung 27: Überblick sonstige Attribute. Eigene Darstellung (2022)

Abschließend fand man unter dem Reiter ‚Cases‘ auf der linken Seite des Fensters unter der Spalte ‚Variants‘ den kompletten Event-Log sowie alle Varianten mit der jeweiligen Anzahl an Cases und einer Prozentangabe, welche angab, wie viel Prozent aller Cases zu der Variante gehörten. Je nachdem, ob man den gesamten Event-Log oder eine spezielle Variante ausgewählt hat, erschienen in der Spalte ‚Cases‘ alle zugehörigen Cases (s. Abbildung 28).

Wählte man einen Case aus, erschienen Informationen dazu, wie viele Ereignisse dieser Case enthält, an welchem Datum und zu welcher Uhrzeit dieser begonnen hat, wie lange der Case gedauert hat und zu wie viel Prozent die Zeit aktiv genutzt wurde. Außerdem wurde in einem Graphen in rot markiert, an welcher Stelle des Event-Logs sich der jeweilige Case befunden hat. Weiterhin konnte man sich den Case als Graph (s. Abbildung 28) oder als Tabelle (s.

Abbildung 29) anzeigen lassen. Der Graph enthielt Informationen zum Startdatum und der Start-Uhrzeit der Aktivitäten sowie deren Dauer und etwaige Wartezeiten. In der Tabelle wurden anstelle der Wartezeiten zusätzlich die Entfernungen der Aktivitäten aufgelistet.

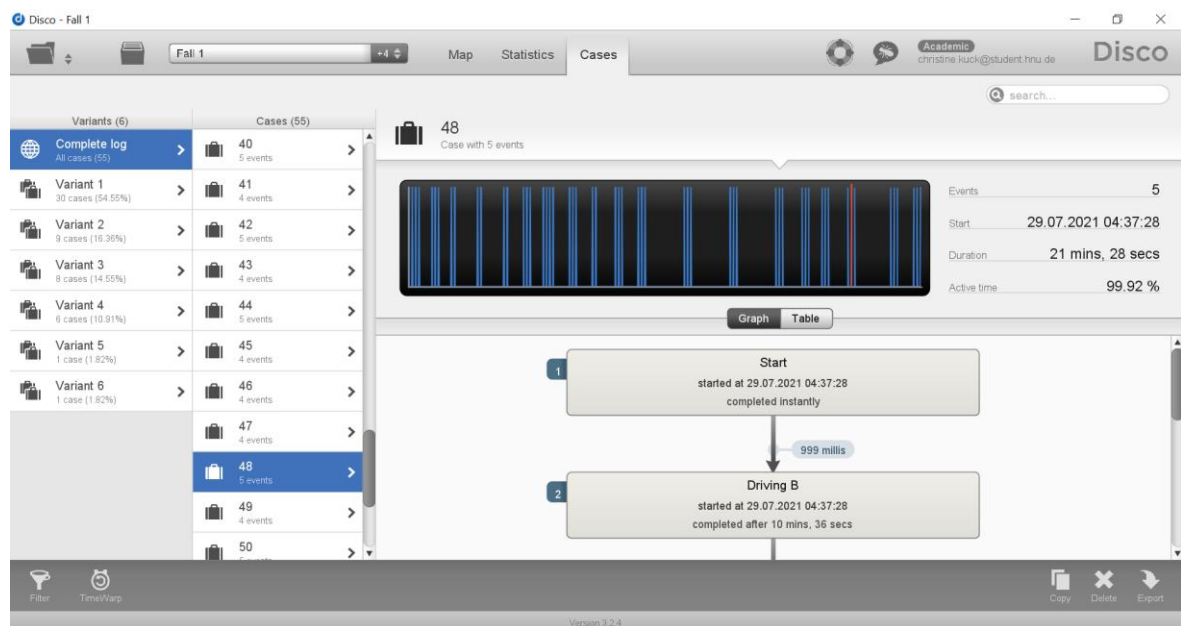


Abbildung 28: Überblick über die einzelnen Cases in ‚Disco‘ mit Graph. Eigene Darstellung (2022)

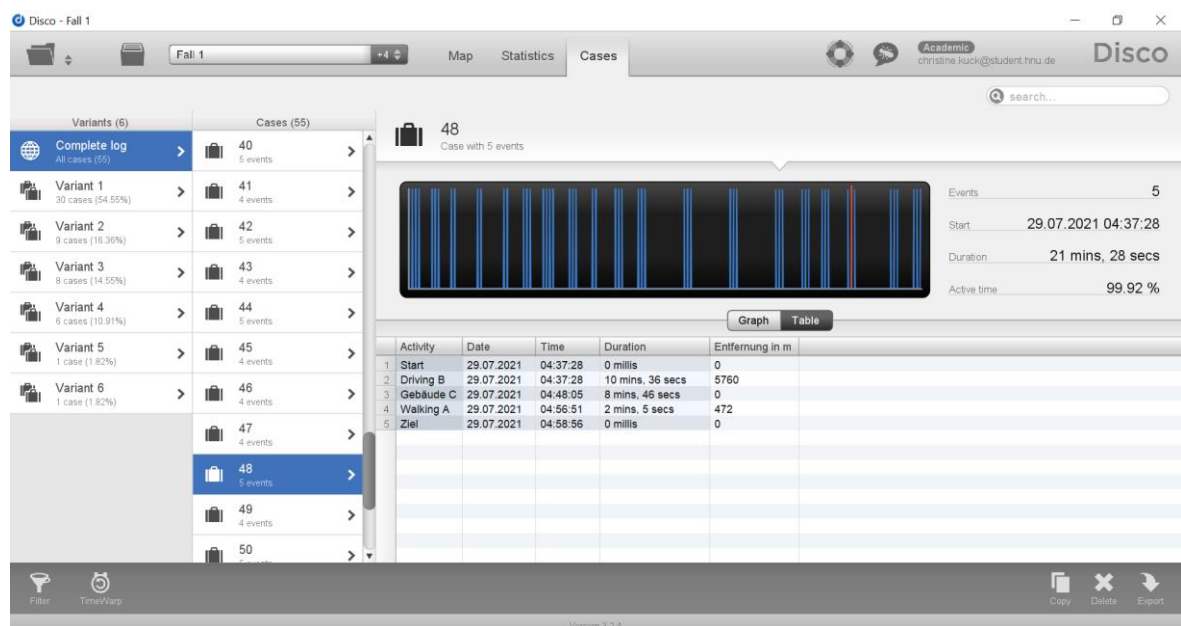


Abbildung 29: Überblick über die einzelnen Cases in ‚Disco‘ mit Tabelle. Eigene Darstellung (2022)

‚Disco‘ bot außerdem die Möglichkeit, die Ergebnisse nach einer Vielzahl von Aspekten zu filtern. Dabei konnte man zwischen Zeiträumen, Varianten, der Performance, unvollständigen Cases, Attributen oder aufeinanderfolgenden Ereignissen wählen oder auch mehrere Filter auf einmal anwenden (s. Abbildung 30).

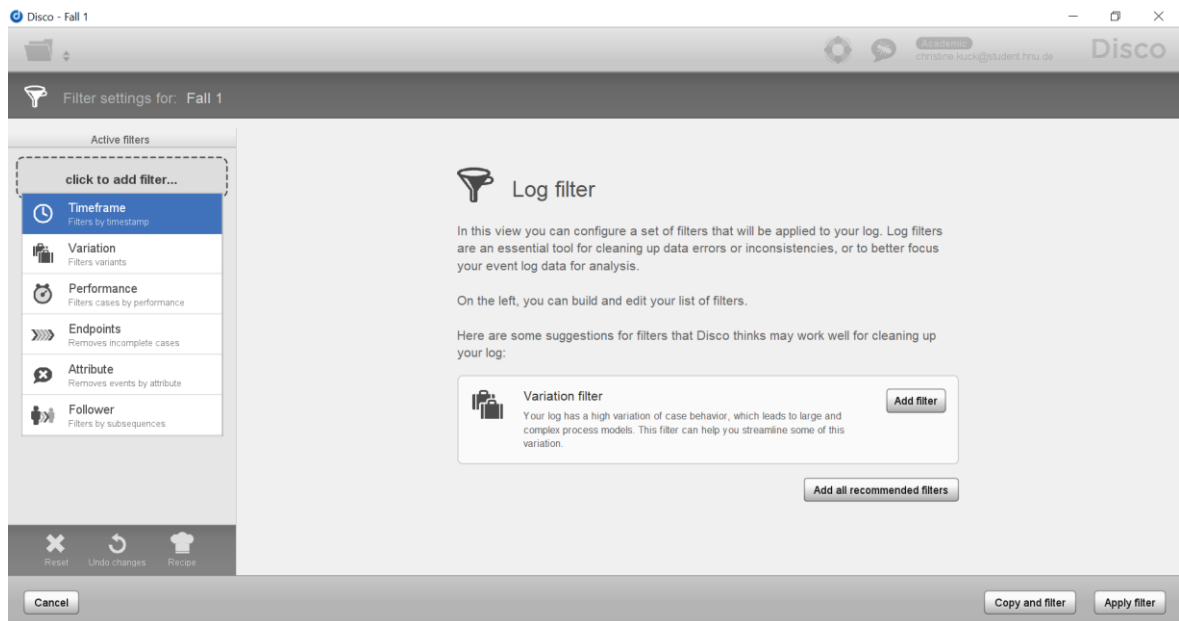


Abbildung 30: Fenster zur Anwendung von Filtern & Filtermöglichkeiten in ‚Disco‘. Eigene Darstellung (2022)

4.2 Benötigte Informationen

Nachdem sich ein Überblick über die Informationen, welche man aus dem Process Mining-Tool ‚Disco‘ entnehmen kann, verschafft wurde, konnte festgestellt werden, welche dieser Informationen für die Prüfung der Hypothesen sowie für die Berechnung der Prozesskennzahlen von Bedeutung sind.

In Tabelle 1 wurde zusammengefasst, welche Daten gebraucht wurden, wo diese zu finden waren und für welche Hypothesen oder Kennzahlen sie eingesetzt werden sollten.

Tabelle 1: Benötigte Informationen. Eigene Darstellung (2022)

Benötigte Information	Zu finden unter	Hypothese/KPI
<i>Anzahl Cases pro Event-Log</i>	Statistics – Overview; Cases	H1
<i>Anzahl Varianten pro Event-Log</i>	Statistics – Tabelle: Variants; Statistics – Graph: Case Variants; Cases - Variants	H3
<i>Anzahl Varianten, die nur 1x aufgetreten sind</i>	Statistics – Tabelle: Variants; Statistics – Graph: Case variants; Cases - Variants	H3
<i>Anzahl Ereignisse pro Variante</i>	Statistics – Tabelle: Variants; Cases - Variants	H2
<i>Ø Dauer der Varianten</i>	Statistics – Tabelle: Variants	KPI V, H2

<i>Prozessablauf der Varianten</i>	Cases – Variants – Graph/Table	KPI <i>V</i>
<i>Ø Dauer Gesamt</i>	Statistics – Overview	KPI <i>D</i>
<i>Dauer einzelner Cases</i>	Statistics – Tabelle: Cases; Cases - Cases	KPI <i>G</i> , KPI <i>D</i>
<i>Entfernung in Metern von jedem Case</i>	Cases - Cases	KPI <i>G</i>
<i>Prozessmodell</i>	Map	H1

4.3 Fall 1

Bei diesem Fall handelte es sich um einen innerstädtischen Arbeitsweg, welcher etwa sechs Kilometer lang war. Die Strecke wurde zum größten Teil mit dem Auto zurückgelegt. Dieses wurde in einem Parkhaus abgestellt und die restliche Strecke bis zum Arbeitsort zu Fuß gegangen.

4.3.1 Allgemeine Informationen

Der Event-Log des ersten Falles beinhaltete **55 Cases**, welche aus insgesamt neun unterschiedlichen Aktivitäten und 236 Events bestanden. Die durchschnittliche Dauer der Cases betrug **23,9 Minuten** (s. Anhang, Abbildung 54).

Bei der Zurücklegung des Weges des ersten Falles sind **sechs Varianten**, bzw. Mobilitätsprozesse, welche beschreiben, auf welche Art und Weise ein Weg zurückgelegt wurde, aufgetreten. **Zwei** davon, die Varianten 5 und 6, beinhalteten jeweils nur **einen Case** (s. Anhang, Abbildung 55).

4.3.2 Prozessmodell

Das vom Process Mining-Tool entdeckte Prozessmodell, welches in Abbildung 31 zu sehen ist, wurde für die Prüfung der ersten Hypothese benötigt.

Anhand der Farben sowie der Zahlen in den Aktivitätsfeldern lässt sich erkennen, dass die Aktivitäten *Start*, *Ziel* und *Walking A* am häufigsten aufgetreten sind. Von den Driving-Aktivitäten ist *Driving A* am häufigsten genutzt worden. Außerdem fällt auf, dass die Aktivitäten *Driving C*, *Driving D* sowie *Gebäude D* jeweils nur einmal vorgekommen sind.

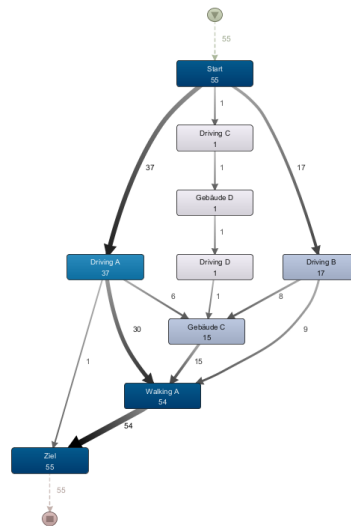


Abbildung 31: Auswertung Fall 1 - Prozessmodell. Eigene Darstellung (2022)

4.3.3 Dauer und zurückgelegte Entfernung der Cases

Für die beiden Prozesskennzahlen G und D war die Dauer jeden einzelnen Cases von Bedeutung. Diese konnte unter dem Reiter ‚Cases‘ für jeden Case abgelesen werden. Im gleichen Zuge konnte die zurückgelegte Entfernung eines jeden Cases, welche ebenfalls für die KPI G benötigt wurde, in der Tabelle gefunden werden (vgl. Abbildung 29).

Die Angaben zur Dauer und zur Entfernung wurden in Tabelle 2 zusammengefasst. Die Dauer der Cases lag in einem Bereich zwischen 20 Minuten 42 Sekunden und 45 Minuten 40 Sekunden, während die Entfernung zwischen 6128 Meter und 9003 Meter betrug.

Tabelle 2: Auswertung Fall 1 - Dauer und Entfernung der Cases. Eigene Darstellung (2022)

Case	Dauer	Zurückgelegte Entfernung (in Metern)
1	24 mins, 40 secs	$5658 + 474 = 6132$
2	26 mins, 29 secs	$5725 + 659 = 6384$
3	26 mins, 47 secs	$5734 + 490 = 6224$
4	24 mins, 30 secs	$5610 + 613 = 6223$
5	25 mins, 10 secs	$5362 + 839 = 6201$
6	25 mins, 53 secs	$5653 + 475 = 6128$
7	24 mins, 59 secs	$5690 + 739 = 6429$
8	21 mins, 5 secs	$5730 + 696 = 6426$
9	23 mins, 30 secs	$5757 + 554 = 6311$
10	23 mins, 21 secs	$5754 + 567 = 6321$
11	23 mins, 49 secs	$5599 + 642 = 6241$
12	22 mins, 29 secs	$5712 + 1070 = 6782$

13	22 mins, 43 secs	$5689 + 688 = 6377$
14	23 mins, 10 secs	$5620 + 1122 = 6742$
15	22 mins, 15 secs	$5704 + 1029 = 6733$
16	22 mins, 52 secs	$5761 + 589 = 6350$
17	21 mins, 36 secs	$5802 + 582 = 6384$
18	20 mins, 42 secs	$5768 + 1058 = 6826$
19	22 mins, 59 secs	$5672 + 882 = 6554$
20	22 mins, 4 secs	$5564 + 647 = 6211$
21	22 mins, 24 secs	$5731 + 613 = 6344$
22	21 mins, 32 secs	$5685 + 629 = 6314$
23	23 mins, 16 secs	$5753 + 827 = 6580$
24	23 mins, 14 secs	$5644 + 592 = 6236$
25	21 mins, 10 secs	$5627 + 1068 = 6695$
26	20 mins, 42 secs	$5547 + 739 = 6286$
27	22 mins, 1 sec	$5617 + 735 = 6352$
28	21 mins, 32 secs	$5707 + 613 = 6320$
29	22 mins, 14 secs	$5727 + 535 = 6262$
30	21 mins, 43 secs	$5726 + 604 = 6330$
31	22 mins, 27 secs	$5712 + 645 = 6357$
32	23 mins, 45 secs	$5720 + 587 = 6307$
33	24 mins, 39 secs	$5628 + 788 = 6416$
34	23 mins, 38 secs	$5701 + 555 = 6256$
35	22 mins, 27 secs	$5600 + 545 = 6145$
36	21 mins, 12 secs	$5591 + 678 = 6269$
37	22 mins, 1 sec	6436
38	26 mins, 47 secs	$6735 + 623 = 7358$
39	22 mins, 58 secs	$5813 + 572 = 6385$
40	28 mins, 25 secs	$6580 + 472 = 7052$
41	23 mins, 2 secs	$5832 + 597 = 6429$
42	23 mins, 42 secs	$5667 + 489 = 6156$
43	24 mins, 5 secs	$5873 + 556 = 6429$
44	25 mins, 45 secs	$5733 + 521 = 6254$
45	24 mins, 12 secs	$6125 + 215 = 6340$
46	23 mins, 52 secs	$5903 + 3100 = 9003$

47	26 mins, 52 secs	$5983 + 459 = 6442$
48	21 mins, 28 secs	$5760 + 472 = 6232$
49	26 mins, 51 secs	$5786 + 587 = 6373$
50	24 mins, 19 secs	$5962 + 484 = 6446$
51	25 mins, 26 secs	$6328 + 541 = 6869$
52	23 mins, 37 secs	$6088 + 485 = 6573$
53	24 mins, 10 secs	$5993 + 690 = 6683$
54	23 mins, 26 secs	$6237 + 555 = 6792$
55	45 mins, 40 secs	$1879 + 5246 + 576 = 7701$

4.3.4 Durchschnittliche Dauer und Ereignisse pro Variante

Um die zweite Hypothese prüfen zu können, wurde die Anzahl der Ereignisse pro Variante sowie die durchschnittliche Dauer der Varianten benötigt. Letzteres war ebenso für die KPI *V* von Bedeutung. Diese beiden Informationen wurden unter dem Reiter ‚Statistics‘ aus der Tabelle ‚Variants‘ (s. Anhang, Abbildung 56) entnommen. Die Ergebnisse wurden in Tabelle 3 zusammengetragen.

Alle Varianten bestanden aus drei bis sieben Ereignissen. Zwei der Varianten beinhalteten jeweils vier Ereignisse, weitere zwei beinhalteten fünf Ereignisse, und die verbliebenen zwei Varianten bestanden aus drei und sieben Ereignissen. Die minimale Durchschnittsdauer lag bei 22 Minuten und einer Sekunde. Die maximale Durchschnittsdauer hingegen betrug 45 Minuten 40 Sekunden.

Tabelle 3: Auswertung Fall 1 - durchschnittliche Dauer und Anzahl der Ereignisse pro Variante. Eigene Darstellung (2022)

Variante	Anzahl der Ereignisse	Ø Dauer
Variante 1	4	22 mins, 40 secs
Variante 2	4	24 mins, 54 secs
Variante 3	5	24 mins, 22 secs
Variante 4	5	24 mins, 50 secs
Variante 5	3	22 mins, 1 sec
Variante 6	7	45 mins, 40 secs

4.3.5 Mobilitätsprozesse

Unter dem Reiter ‚Cases‘ ließ sich ablesen, welcher Mobilitätsprozess hinter den einzelnen Varianten steckte (s. Anhang, Abbildungen 57-62). Die spezifischen Prozesse wurden für

die Berechnung der Prozesskennzahl V benötigt. Tabelle 4 fasst die aufgetretenen Mobilitätsprozesse zusammen.

Anhand der Mobilitätsprozesse fällt auf, dass nach dem Startpunkt immer eine Fahrt mit dem Auto folgte. In zwei Fällen folgte auf die Autofahrt eine Walking-Aktivität zum Ziel, in zwei weiteren Fällen befand sich zwischen der Fahrt und der Walking-Aktivität ein Gebäude. Bei der fünften Variante folgte auf die Fahrt direkt das Ziel. Bei der sechsten Variante lassen sich auf dem Weg zur Arbeit mehrere Zwischenstopps erkennen.

Tabelle 4: Auswertung Fall 1 - Mobilitätsprozesse der Varianten. Eigene Darstellung (2022)

Variante	Mobilitätsprozess
<i>Variante 1</i>	Start → Driving A → Walking A → Ziel
<i>Variante 2</i>	Start → Driving B → Walking A → Ziel
<i>Variante 3</i>	Start → Driving B → Gebäude C → Walking A → Ziel
<i>Variante 4</i>	Start → Driving A → Gebäude C → Walking A → Ziel
<i>Variante 5</i>	Start → Driving A → Ziel
<i>Variante 6</i>	Start → Driving C → Gebäude D → Driving D → Gebäude C → Walking A → Ziel

4.3.6 Berechnung der Prozesskennzahl V

Für die Prozesskennzahl V ergaben sich unter Verwendung der Mobilitätsprozesse der Varianten (s. Kapitel 4.3.5) und der durchschnittlichen Dauer ebendieser (s. Kapitel 4.3.4) Werte, welche in Tabelle 21 (s. Anhang) aufgelistet sind. In Abbildung 32 wurden die errechneten Werte außerdem in Form eines Graphen visualisiert.

Anhand des Graphen lässt sich erkennen, dass alle Werte über dem Wert ,1‘ liegen. Während der Wert bei den ersten drei Varianten relativ konstant ist, ist ab der vierten Variante ein Anstieg bemerkbar, welcher bei der sechsten Variante bei einem Wert von über zwei endet.

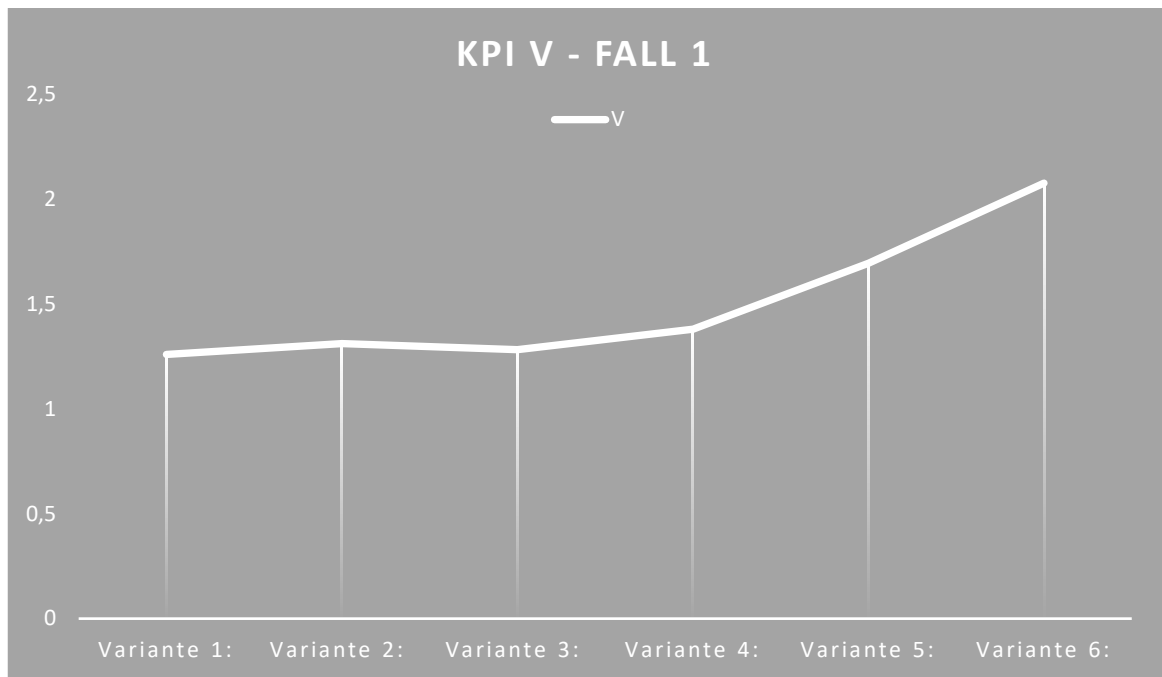


Abbildung 32: Auswertung Fall 1 - Ergebnisse KPI V. Eigene Darstellung (2022)

4.3.7 Berechnung der Prozesskennzahl G

Um die Prozesskennzahl G zu berechnen, wurden die zurückgelegte Entfernung der einzelnen Cases sowie die Dauer der Cases (s. Kapitel 4.3.3) genutzt. Tabelle 22 (s. Anhang) beinhaltet die Ergebnisse der Berechnungen der KPIs G und D .

Abbildung 33 stellt die errechneten Ergebnisse in Form eines Graphen dar. Anhand dessen ist zu erkennen, dass Case 46 der Case mit der mit Abstand schnellsten Durchschnittsgeschwindigkeit war. Auch bei Case 18 kann man eine Spitze erkennen. Ebenfalls schnell unterwegs waren die Cases 25, 8, 26, 15 und 12.

Insgesamt waren 30 der 55 Cases schneller als der Durchschnitt. Folglich waren also 25 Cases langsamer als der Durchschnitt. Der langsamste war mit großem Abstand Case 55.

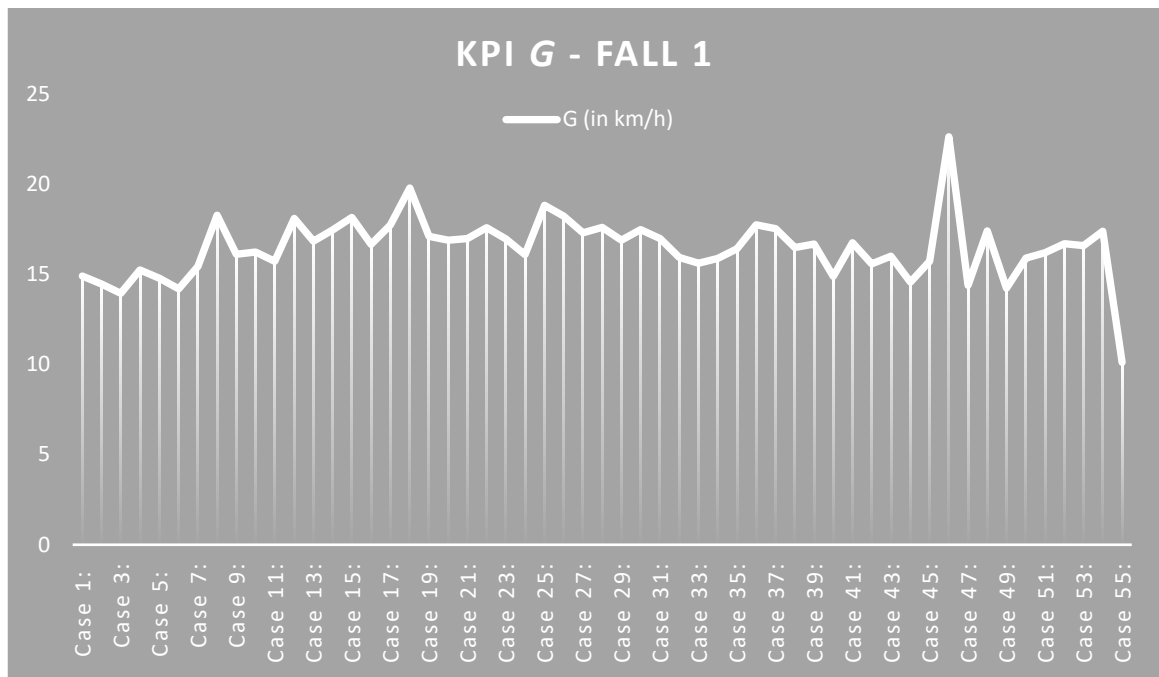


Abbildung 33: Auswertung Fall 1 - Ergebnisse KPI G. Eigene Darstellung (2022)

4.3.8 Berechnung der Prozesskennzahl D

Die Dauer der einzelnen Cases (s. Kapitel 4.3.3) wurde gemeinsam mit der durchschnittlichen Dauer aller Cases (s. Kapitel 4.3.1) zur Berechnung der KPI D verwendet. Die genauen Ergebnisse finden sich in Tabelle 22 (s. Anhang).

Die Berechnungen haben gezeigt, dass insgesamt 36 der 55 Cases eine kürzere Dauer als der Durchschnitt hatten. Wie man sowohl in der Tabelle als auch in Abbildung 34, welche die Ergebnisse grafisch darstellt, sehen kann, waren die Cases 18 und 26 mit einer exakt gleichen Dauer am kürzesten. Darauf folgten die Cases 8, 36, 25 und. Die Cases 43 und 46 lagen beide sehr nah an der Durchschnittsdauer. Weit über den durchschnittlichen 23,9 Minuten war hingegen Case 55.

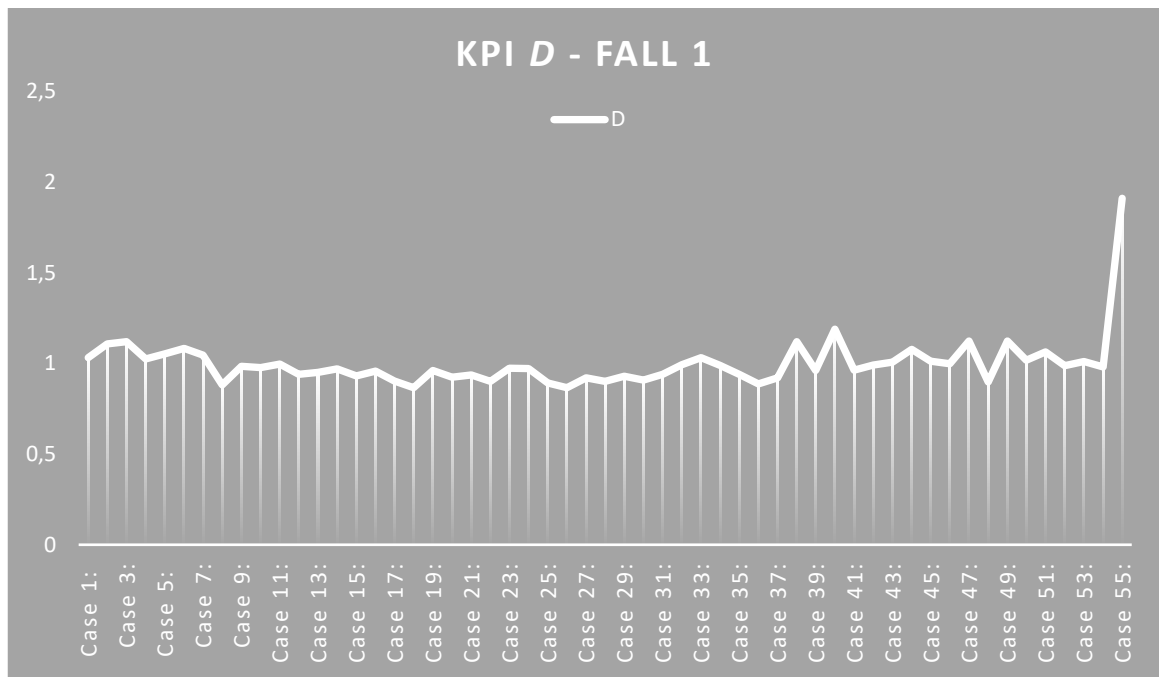


Abbildung 34: Auswertung Fall 1 - Ergebnisse KPI D. Eigene Darstellung (2022)

4.3.9 Filter: Kürzeste Cases

Der erste Filter, der angewandt wurde, bezog sich auf die Dauer der Cases und diente der Fragestellung „Welche waren die kürzesten Cases?“. Da der kürzeste Case 20 Minuten und 42 Sekunden gedauert hat, wurde nach allen Fällen gefiltert, welche maximal 21 Minuten und 30 Sekunden dauerten.

Insgesamt sechs Cases trafen auf dieses Kriterium zu. Dies entspricht zehn Prozent aller Fälle (s. Anhang, Abbildung 63). In Tabelle 5 wurden die Case IDs, die Dauer und die Varianten, zu denen die jeweiligen Cases gehören, aufgelistet.

Tabelle 5: Auswertung Fall 1 - Ergebnisse des Filters 'Case duration max. 21,5min'. Eigene Darstellung (2022)

<i>Case ID</i>	<i>Dauer</i>	<i>Variante</i>
8	21min 5sek	1
18	20min 42sek	1
25	21min 10sek	1
26	20min 42sek	1
36	21min 12sek	1
48	21min 28sek	2

Nach Anwendung des Filters betrug die Durchschnittszeit der ersten Variante 20 Minuten und 58 Sekunden. Bei der zweiten Variante betrug diese 21 Minuten und 28 Sekunden und war somit identisch zu der Dauer von Case 48.

4.3.10 Filter: Kürzer als Durchschnitt

Auch der zweite Filter bezog sich auf die Dauer der Fälle und sollte die Frage „Wie viele Cases waren schneller als der Durchschnitt?“ beantworten. Als Maximalwert wurde hierfür der Durchschnittswert, also 23,9 Minuten (= 23 Minuten 54 Sekunden), eingesetzt.

Auf den Filter trafen insgesamt 36 Fälle, entsprechend 65% aller Cases, zu. 26 (72,22%) davon ließen sich Variante 1 zuordnen, vier (11,11%) der zweiten Variante, drei (8,33%) der dritten Variante, zwei (5,56%) der vierten Variante und einer (2,78%) der fünften Variante (s. Anhang, Abbildung 64).

4.3.11 Filter: Cases mit Wartezeiten

Um die Frage, bei wie vielen Cases Wartezeiten auftraten, zu beantworten, wurde abschließend ein Filter angewandt, welcher sich auf die durchschnittliche Wartezeit bezog. Um alle Cases mit einer Wartezeit zu erhalten, wurde nach einer Mindest-Wartedauer von einer Millisekunde gefiltert.

Als Ergebnis kamen 17 Fälle heraus, was 30% aller Cases entspricht. Davon konnten acht (47,06%) der ersten Variante zugeordnet werden, sechs (35,29%) gehörten zur zweiten Variante, zwei (11,76%) zur dritten Variante und einer (5,88%) zur vierten Variante (s. Anhang, Abbildung 65).

4.4 Fall 2

Dieser Case beschrieb ebenfalls einen Arbeitsweg, für welchen das Auto genutzt wurde. Da sich der Wohnort 26 Kilometer entfernt vom Arbeitsort befand, führte die Strecke in den meisten Fällen über Landstraßen und eine Autobahn.

4.4.1 Allgemeine Informationen

Der zweite Fall bestand aus **41 Cases**. Es traten insgesamt 20 Aktivitäten und 158 Ereignisse auf. Durchschnittlich dauerte ein Case **30,1 Minuten** (s. Anhang, Abbildung 66).

Im gesamten Event-Log sind **15 Mobilitätsprozesse** (Varianten) aufgetreten. **Sechs** dieser Varianten (9 bis 15) traten nur **einmalig** auf (s. Anhang, Abbildung 67).

4.4.2 Prozessmodell

Abbildung 35 zeigt das durch die Anwendung von Process Mining auf den Event-Log des zweiten Falles entdeckte Prozessmodell, welches für die Prüfung von Hypothese H1 von Bedeutung war.

Auf den ersten Blick ist zu erkennen, dass das Modell zahlreiche Aktivitätsfelder enthält. Acht davon sind lediglich ein Mal aufgetreten. Neben *Start* und *Ziel*, welche in jedem Case

aufgetreten sind, ist die Aktivität *Walking* am häufigsten vorgekommen. Die häufigste Driving-Aktivität war *Driving D*, gefolgt von *Driving F*.

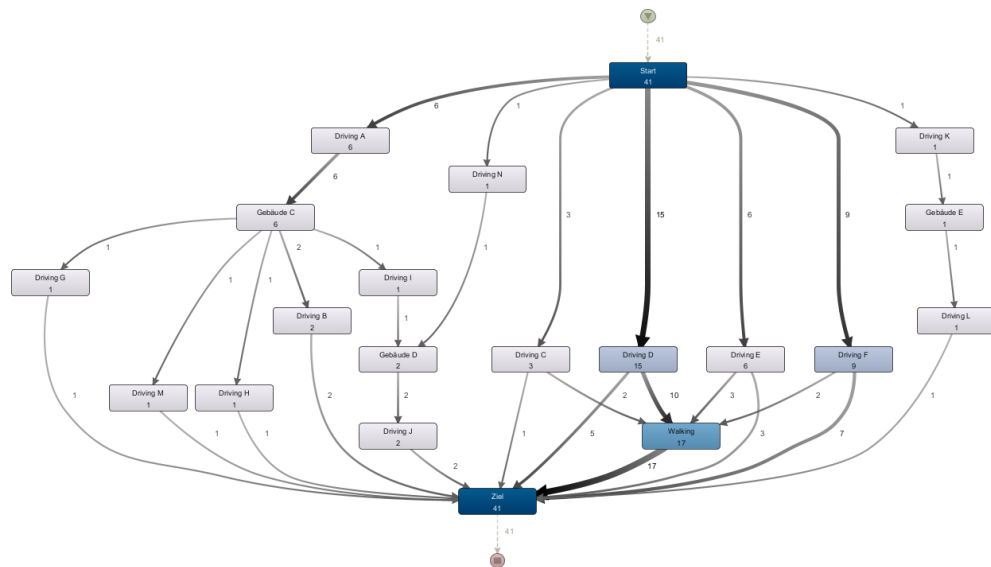


Abbildung 35: Auswertung Fall 2 - Prozessmodell. Eigene Darstellung (2022)

4.4.3 Dauer und zurückgelegte Entfernung der Cases

Die Dauer der einzelnen Cases wurde benötigt, um die KPIs *G* und *D* berechnen zu können. Weiterhin wurde ebenfalls für die KPI *G* die zurückgelegte Entfernung eines jeden Cases benötigt. Diese beiden Informationen konnten unter dem Reiter ‚Cases‘ im Tool ‚Disco‘ herausgelesen werden (s. Anhang, Abbildung 68).

Die Angaben zur Dauer und zur Entfernung wurden in Tabelle 6 zusammengefasst. Die Entfernung der Cases bewegte sich in einem Bereich zwischen 25875 und 33639 Metern. Auch die Dauer lag zwischen 19 Minuten 41 Sekunden und 48 Minuten 27 Sekunden.

Tabelle 6: Auswertung Fall 2 - Dauer und Entfernung der Cases. Eigene Darstellung (2022)

Case	Dauer	Zurückgelegte Entfernung (in Metern)
1	33 mins, 53 secs	$8700 + 24100 = 32800$
2	29 mins, 53 secs	$6904 + 20097 = 27001$
3	33 mins, 33 secs	$29962 + 184 = 30146$
4	29 mins, 18 secs	28915
5	25 mins, 18 secs	$27094 + 219 = 27313$
6	22 mins, 29 secs	26168
7	33 mins, 34 secs	$29925 + 265 = 30190$
8	31 mins, 28 secs	30082
9	22 mins, 30 secs	25875

10	39 mins, 50 secs	$8719 + 24920 = 33639$
11	19 mins, 55 secs	25894
12	20 mins, 53 secs	27000
13	19 mins, 54 secs	26809
14	47 mins, 42 secs	$8561 + 24101 = 32662$
15	26 mins, 12 secs	$26965 + 237 = 27202$
16	19 mins, 41 secs	26848
17	25 mins, 32 secs	$26937 + 316 = 27253$
18	27 mins, 13 secs	26417
19	32 mins, 1 sec	$28796 + 174 = 28970$
20	30 mins, 49 secs	$28773 + 176 = 28949$
21	33 mins, 33 secs	$28804 + 184 = 28988$
22	27 mins, 56 secs	$29024 + 221 = 29245$
23	48 mins, 27 secs	$8623 + 22586 + 932 = 32141$
24	28 mins, 40 secs	29235
25	24 mins, 8 secs	$25778 + 291 = 26069$
26	25 mins, 25 secs	26122
27	31 mins, 12 secs	28997
28	25 mins, 25 secs	27117
29	23 mins, 20 secs	25993
30	29 mins, 14 secs	$28735 + 197 = 28932$
31	33 mins, 53 secs	$28863 + 235 = 29098$
32	29 mins, 51 secs	$28683 + 236 = 28919$
33	29 mins, 6 secs	29004
34	43 mins, 54 secs	$8671 + 19480 = 28151$
35	43 mins, 49 secs	$6879 + 23410 = 30289$
36	41 mins, 22 secs	$28217 + 508 = 28725$
37	31 mins, 26 secs	$28826 + 248 = 29074$
38	28 mins, 45 secs	28884
39	33 mins, 44 secs	$28892 + 227 = 29119$
40	29 mins, 19 secs	$28692 + 218 = 28910$
41	19 mins, 59 secs	$25749 + 315 = 26064$

4.4.4 Durchschnittliche Dauer und Ereignisse pro Variante

Zur Prüfung von Hypothese H2 waren die Anzahl der Ereignisse pro Variante sowie die Durchschnittsdauer der einzelnen Varianten von Bedeutung. Letzteres wurde ebenso für die Berechnung der KPI V benötigt. Aus der Tabelle ‚Variants‘ (s. Anhang, Abbildung 69) im Process Mining-Tool ‚Disco‘ konnten diese Informationen entnommen werden. In Tabelle 7 wurden die Ergebnisse zusammengefasst.

Die Anzahl der Ereignisse betrug zwischen drei und sieben pro Variante. Vier der insgesamt 15 Varianten bestanden aus drei Ereignissen, vier Varianten bestanden aus jeweils vier Ereignissen, sechs bestanden aus fünf Ereignissen, und eine Variante beinhaltete sieben Ereignisse. Die niedrigste Durchschnittsdauer betrug 21 Minuten 59 Sekunden, während die höchste bei 48 Minuten 27 lag.

Tabelle 7: Auswertung Fall 2 - durchschnittliche Dauer und Anzahl der Ereignisse pro Variante. Eigene Darstellung (2022)

Variante	Anzahl der Ereignisse	Ø Dauer
Variante 1	4	31 mins, 11 secs
Variante 2	3	22 mins, 58 secs
Variante 3	3	29 mins, 24 secs
Variante 4	4	25 mins, 41 secs
Variante 5	3	21 mins, 59 secs
Variante 6	5	31 mins, 53 secs
Variante 7	4	33 mins, 34 secs
Variante 8	4	22 mins, 3 secs
Variante 9	3	31 mins, 28 secs
Variante 10	5	39 mins, 50 secs
Variante 11	5	47 mins, 42 secs
Variante 12	7	48 mins, 27 secs
Variante 13	5	43 mins, 54 secs
Variante 14	5	43 mins, 49 secs
Variante 15	5	41 mins, 22 secs

4.4.5 Mobilitätsprozesse

Zur Berechnung der Prozesskennzahl V wurden die genauen Prozessabläufe der einzelnen Varianten (s. Anhang, Abbildungen 70-84) benötigt. Diese konnten unter dem Reiter ‚Cases‘ ausfindig gemacht werden. In Tabelle 8 werden die Mobilitätsprozesse der Varianten übersichtlich aufgezeigt.

Die Mobilitätsprozesse zeigen auf, dass nach dem Start immer eine Autofahrt folgte. In vier Fällen wurde daraufhin bereits das Ziel erreicht. In weiteren vier Fällen folgte auf die Autofahrt ein Fußweg bis zum Zielort. Bei weiteren sieben Varianten wurde die Fahrt zum Zielort von einem Zwischenstopp unterbrochen. Bei einer Variante traten auf der Fahrt zwei Stopps auf.

Tabelle 8: Auswertung Fall 2 - Mobilitätsprozesse der Varianten. Eigene Darstellung (2022)

Variante	Mobilitätsprozess
<i>Variante 1</i>	Start → Driving D → Walking → Ziel
<i>Variante 2</i>	Start → Driving F → Ziel
<i>Variante 3</i>	Start → Driving D → Ziel
<i>Variante 4</i>	Start → Driving E → Walking → Ziel
<i>Variante 5</i>	Start → Driving E → Ziel
<i>Variante 6</i>	Start → Driving A → Gebäude C → Driving B → Ziel
<i>Variante 7</i>	Start → Driving C → Walking → Ziel
<i>Variante 8</i>	Start → Driving F → Walking → Ziel
<i>Variante 9</i>	Start → Driving C → Ziel
<i>Variante 10</i>	Start → Driving A → Gebäude C → Driving G → Ziel
<i>Variante 11</i>	Start → Driving A → Gebäude C → Driving H → Ziel
<i>Variante 12</i>	Start → Driving A → Gebäude C → Driving I → Gebäude D → Driving J → Ziel
<i>Variante 13</i>	Start → Driving K → Gebäude E → Driving L → Ziel
<i>Variante 14</i>	Start → Driving A → Gebäude C → Driving M → Ziel
<i>Variante 15</i>	Start → Driving N → Gebäude D → Driving J → Ziel

4.4.6 Berechnung der Prozesskennzahl V

Unter Verwendung der durchschnittlichen Dauer (s. Kapitel 4.4.4) und der Prozessabläufe der Varianten (s. Kapitel 4.4.5) ergaben sich die in Tabelle 23 (s. Anhang) aufgelisteten Werte für die KPI *V*. Zudem wurden die errechneten Werte in der folgenden Abbildung 36 visualisiert.

Den niedrigsten Wert erreichte, wie in dem Graphen zu sehen ist, die siebte Variante. Darauf folgten die erste, die neunte und die dritte Variante. Bei der achten Variante kann man behaupten, dass diese nahezu exakt der schnellsten Dauer entsprochen hat. Die Varianten 12 und 14 wiesen die höchsten Werte auf.

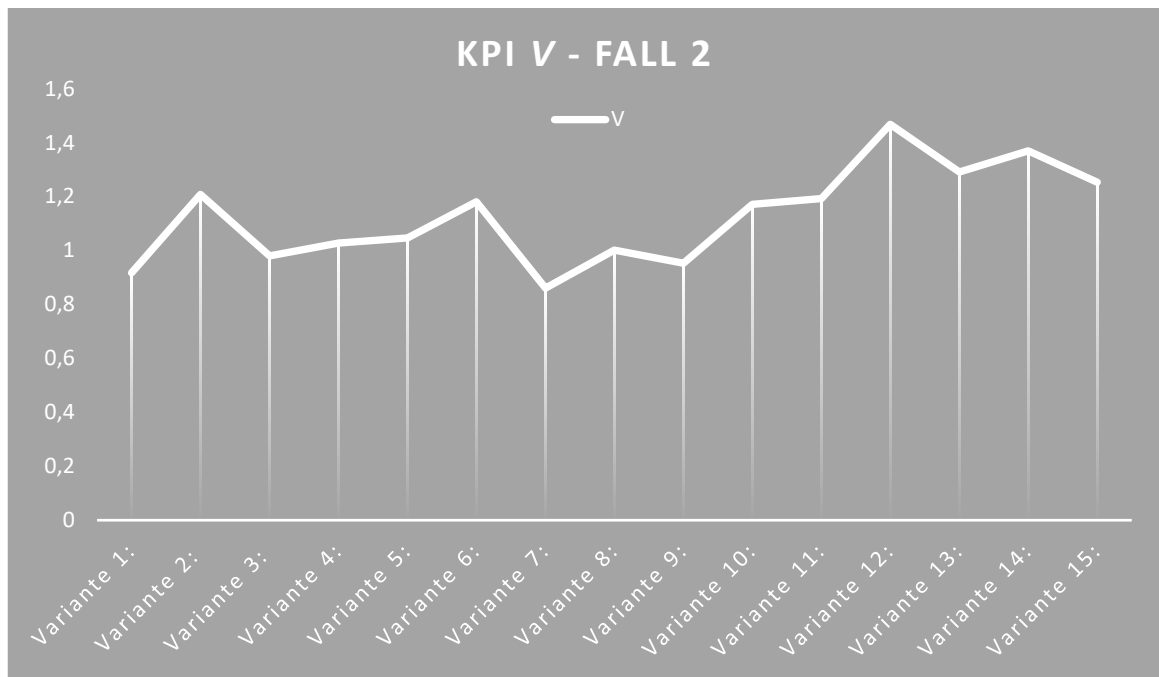


Abbildung 36: Auswertung Fall 2 - Ergebnisse KPI V. Eigene Darstellung (2022)

4.4.7 Berechnung der Prozesskennzahl G

Die zurückgelegte Entfernung der einzelnen Cases sowie die Dauer der Cases (s. Kapitel 4.4.3) wurden genutzt, um die Prozesskennzahl G zu berechnen.

Die Ergebnisse der Berechnungen wurden in Tabelle 24 (s. Anhang) notiert. Aus dieser lässt sich ablesen, dass 19 der insgesamt 41 Cases über und 22 Cases unter dem Durchschnittswert lagen. In Abbildung 37 lassen sich eindeutige Spitzen bei den Cases 16 und 13 erkennen. Darauf folgten die Cases 41, 11 und 12. Der langsamste aller Cases war Case 34. Ebenfalls zu den langsameren Cases gehörten Case 23 und Case 14, bei denen deutliche Einbrüche im Graphen zu erkennen sind.

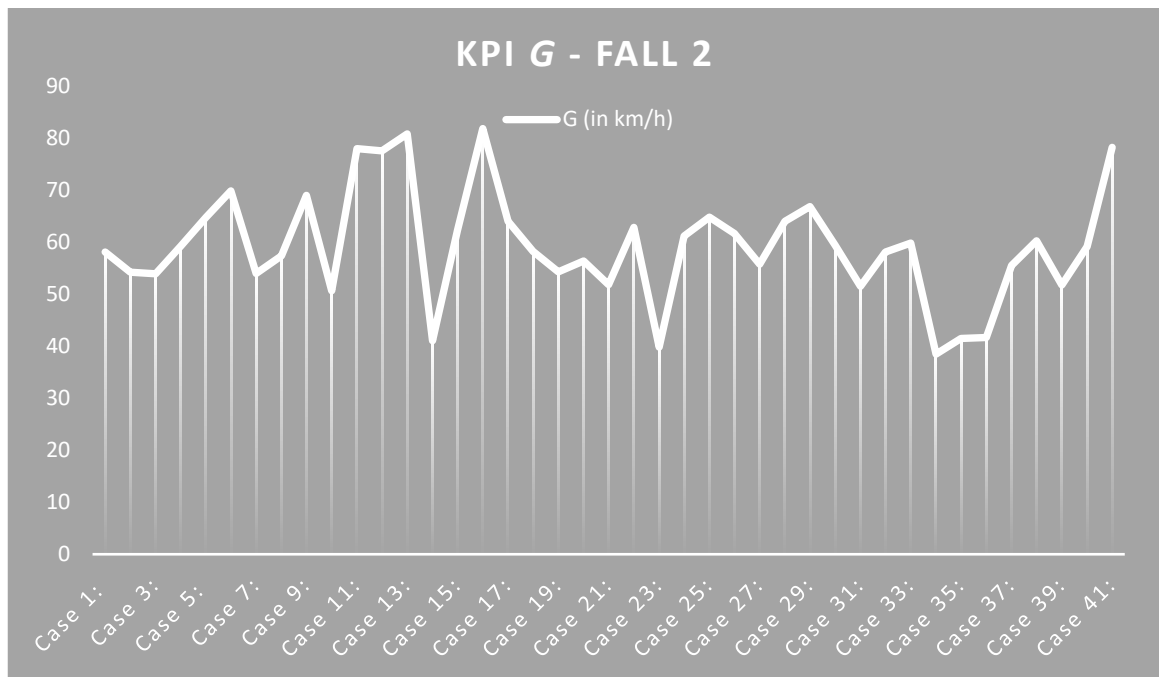


Abbildung 37: Auswertung Fall 2 - Ergebnisse KPI G. Eigene Darstellung (2022)

4.4.8 Berechnung der Prozesskennzahl D

Zur Berechnung der KPI D kamen die Dauer der Cases (s. Kapitel 4.4.3) sowie die durchschnittliche Dauer aller Cases (s. Kapitel 4.4.1) zum Einsatz. Genau wie bei der KPI G wurden die Ergebnisse in Tabelle 24 (s. Anhang) zusammengefasst.

Anhand der Berechnungen konnte festgestellt werden, dass 24 Cases unter der Durchschnittszeit von 30,1 Minuten und 17 Cases darüber lagen.

Zu den schnellsten gehörten die Cases 16, 13, 11, 41 und 12, was in Abbildung 38 an den Tälern zu erkennen ist. Am nächsten an der Durchschnittszeit waren die Cases 32 und 2, welche beide einen Wert von nahezu ,1‘ aufwiesen. Deutlich zu sehen ist auch, dass die Cases 14, 23, 34 und 35 besonders weit von der Durchschnittszeit entfernt waren.

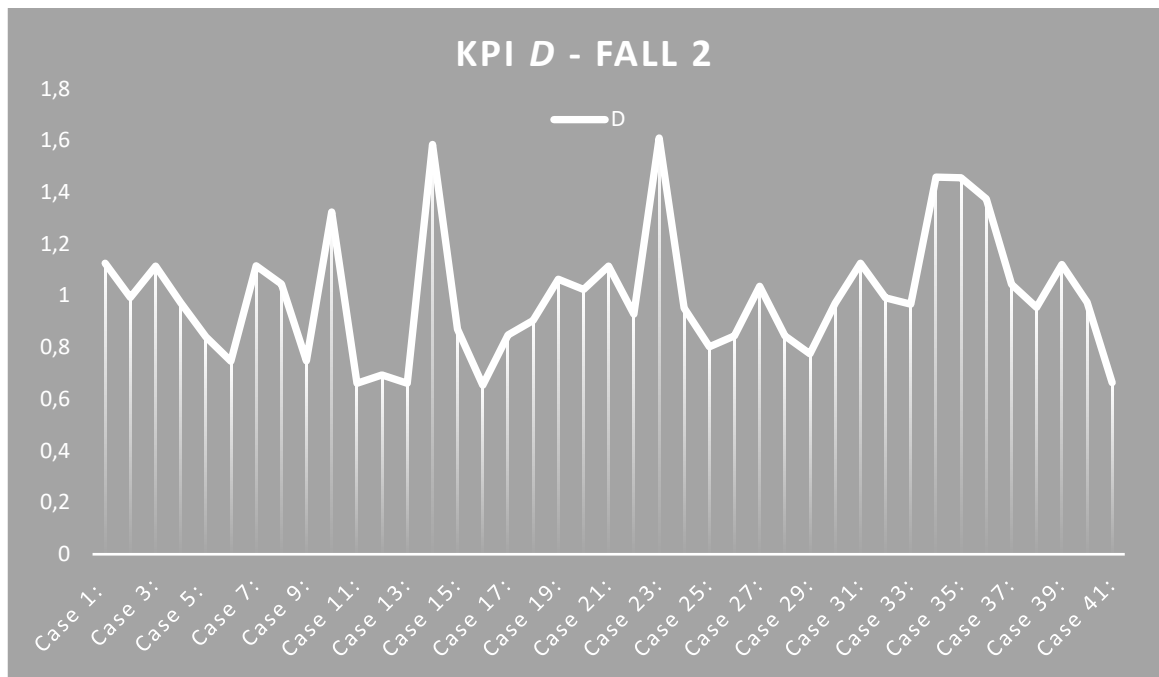


Abbildung 38: Auswertung Fall 2 - Ergebnisse KPI D. Eigene Darstellung (2022)

4.4.9 Filter: Kürzeste Cases

Da der kürzeste Case 19 Minuten und 41 Sekunden gedauert hat, wurde nach den Fällen gesucht, die maximal 20 Minuten und 30 Sekunden gedauert haben, um die Frage nach den kürzesten Cases zu beantworten.

Vier Fälle trafen auf den Filter zu. Dies entspricht neun Prozent aller Cases (s. Anhang, Abbildung 85). In Tabelle 9 wurden die Case IDs, die Dauer und die Varianten, zu denen die jeweiligen Cases gehören, aufgelistet.

Tabelle 9: Auswertung Fall 2 - Ergebnisse des Filters 'Case duration max. 20,5min'. Eigene Darstellung (2022)

Case ID	Dauer	Variante
11	19min 55sek	1
13	19min 54sek	1
16	19min 41sek	2
41	19min 59sek	3

Die durchschnittliche Durchlaufzeit der Varianten betrug nach Anwendung des Filters 19 Minuten und 54 Sekunden für die erste Variante, 19 Minuten und 41 Sekunden für die zweite Variante und 19 Minuten und 59 Sekunden für die dritte Variante. Da die zweite und die dritte Variante beide nur einmalig vorgekommen sind, entspricht deren Durchschnittsdauer der Gesamtdauer der Cases.

4.4.10 Filter: Kürzer als Durchschnitt

Um die Frage, wie viele Cases schneller als der Durchschnitt waren, zu beantworten, wurde nach der Maximalzeit von 30,1 Minuten (= 30 Minuten 6 Sekunden) gefiltert, was der Durchschnittszeit aller Cases des Event-Logs entsprach.

Insgesamt 24 Cases, also 58% aller Fälle, trafen auf den Filter zu. Sieben davon (29,17%) wurden der ersten Variante zugeordnet, jeweils vier (je 16,67%) den Varianten 2 und 3, jeweils drei (je 12,5%) den Varianten 4 und 5, zwei (8,33%) der Variante 6 und eines (4,17%) der siebten Variante (s. Anhang, Abbildung 86).

4.4.11 Filter: Cases mit Wartezeiten

Hier sollte die Frage beantwortet werden, bei wie vielen Cases Wartezeiten vorgekommen sind. Nachdem nach einer Mindest-Wartezeit von einer Millisekunde gefiltert wurde, konnten keine Ergebnisse gefunden werden (s. Anhang, Abbildung 87).

4.5 Fall 3

Der dritte Fall beschrieb im Gegensatz zu den beiden vorherigen Cases keinen täglichen Arbeitsweg, sondern eine Pendelstrecke. Der Nutzer pendelte zwischen zwei etwa 140 Kilometer weit entfernten Orten. Betrachtet wurde hier der Rückweg vom Zweitwohnsitz zum Hauptwohnsitz gegen Ende einer Arbeitswoche. Genutzt wurde auch hier das Auto. Die Strecke führte hauptsächlich über Autobahnen und Bundesstraßen.

4.5.1 Allgemeine Informationen

Der Event-Log des dritten Falles beinhaltete **20 Cases**, 22 Aktivitäten und 110 Events. Die Dauer der Cases lag durchschnittlich bei **111,6 Minuten** (s. Anhang, Abbildung 88).

Es sind **zehn** unterschiedliche Varianten aufgetreten. **Sieben** davon (Varianten 4 bis 10) enthielten nur **einen Case** (s. Anhang, Abbildung 89).

4.5.2 Prozessmodell

Wie bei den vorherigen Fällen war auch beim dritten Fall das Prozessmodell von Bedeutung, um die erste Hypothese beantworten zu können. Abbildung 39 zeigt das Modell, welches durch die Anwendung des Process Mining Ansatzes auf den Event-Log entdeckt wurde.

Das Modell zeigt, dass die Aktivitäten *Driving A* sowie *Gebäude C* in 17 von 20 Fällen aufgetreten sind. Nur drei Mal wurde vom Startpunkt aus eine andere Route gewählt. Außerdem gab es nach *Gebäude C* insgesamt acht unterschiedliche Driving-Aktivitäten. Mit Abstand am häufigsten auf *Gebäude C* folgte *Driving G*, während die anderen sieben

Driving Aktivitäten nur ein bzw. zwei Mal aufgetreten sind. Es ist ebenso zu erkennen, dass jeder End-to-End-Prozess aus mindestens fünf Aktivitäten bestand.

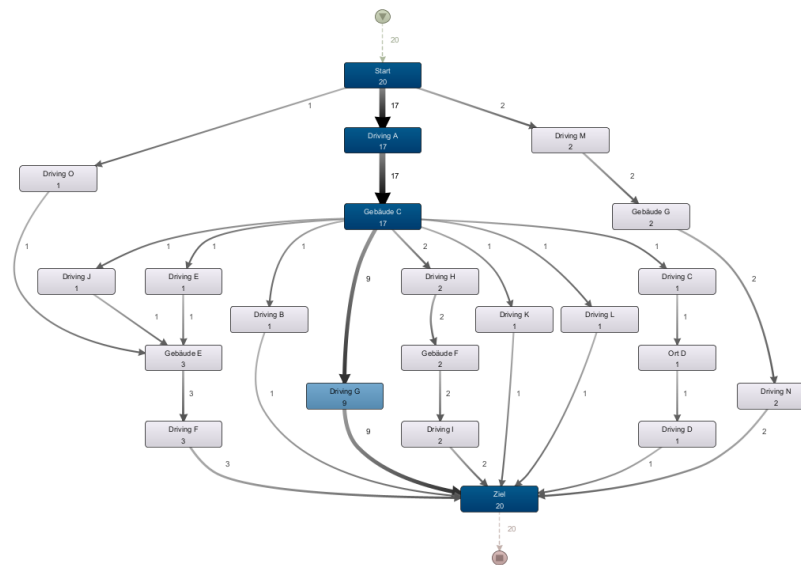


Abbildung 39: Auswertung Fall 3 - Prozessmodell. Eigene Darstellung (2022)

4.5.3 Dauer und zurückgelegte Entfernung der Cases

Unter dem Reiter ‚Cases‘ konnten die Dauer der einzelnen Cases sowie deren zurückgelegte Entfernung in Metern entnommen werden (s. Anhang, Abbildung 90). Letzteres war für die Berechnung der Prozesskennzahl G von Bedeutung, während ersteres sowohl für die KPI G als auch für die KPI D benötigt wurde.

Die Angaben zur Dauer und zur Entfernung wurden in Tabelle 10 zusammengefasst. Die Minstdauer betrug eine Stunde 31 Minuten, während die maximale Dauer bei zwei Stunden 54 Minuten lag. Die Entfernung bewegte sich in einem Bereich zwischen 137393 und 165487 Metern.

Tabelle 10: Auswertung Fall 3 - Dauer und Entfernung der Cases. Eigene Darstellung (2022)

Case	Dauer	Zurückgelegte Entfernung (in Metern)
1	2 hours, 54 mins	$2724 + 144396 = 147120$
2	2 hours, 53 mins	$2694 + 148414 + 14379 = 165487$
3	1 hour, 53 mins	$2736 + 125045 + 12296 = 140077$
4	1 hour, 32 mins	$2789 + 134868 = 137657$
5	1 hour, 38 mins	$2715 + 134788 = 137503$
6	2 hours, 17 mins	$2669 + 143127 + 9416 = 155212$
7	1 hour, 31 mins	$2713 + 134831 = 137544$
8	2 hours, 4 mins	$2707 + 143321 + 8574 = 154602$

9	1 hour, 32 mins	$2753 + 134640 = 137393$
10	1 hour, 33 mins	$2753 + 134770 = 137523$
11	1 hour, 50 mins	$2715 + 126404 + 12267 = 141386$
12	1 hour, 33 mins	$2750 + 134748 = 137498$
13	1 hour, 52 mins	$3364 + 139022 = 142386$
14	2 hours, 6 mins	$2712 + 140636 = 143348$
15	1 hour, 35 mins	$2762 + 134789 = 137551$
16	1 hour, 31 mins	$112930 + 25567 = 138497$
17	1 hour, 35 mins	$113032 + 25692 = 138724$
18	1 hour, 35 mins	$2701 + 135023 = 137724$
19	1 hour, 41 mins	$125320 + 12339 = 137659$
20	1 hour, 56 mins	$2686 + 134833 = 137519$

4.5.4 Durchschnittliche Dauer und Ereignisse pro Variante

Aus der Tabelle ‚Variants‘ (s. Anhang, Abbildung 91), welche unter dem Reiter ‚Statistics‘ zu finden war, konnten Informationen über die Anzahl der Ereignisse pro Variante sowie über die durchschnittliche Dauer der einzelnen Varianten gefunden werden. Beide Informationen waren für die Prüfung der zweiten Hypothese von Bedeutung. Außerdem wurde die Durchschnittsdauer ebenfalls zur Berechnung der Prozesskennzahl V benötigt. In Tabelle 11 wurden die Ergebnisse zusammengefasst.

Sechs der insgesamt zehn Varianten bestanden aus jeweils fünf Ereignissen, während die übrigen vier Varianten aus jeweils sieben Ereignissen bestanden. Außerdem bewegte sich die durchschnittliche Dauer der Varianten zwischen einer Stunde 33 Minuten und zwei Stunden 54 Minuten.

Tabelle 11: Auswertung Fall 3 - durchschnittliche Dauer und Anzahl der Ereignisse pro Variante. Eigene Darstellung (2022)

Variante	Anzahl der Ereignisse	Ø Dauer
Variante 1	5	1 hour, 36 mins
Variante 2	7	2 hours, 11 mins
Variante 3	5	1 hour, 33 mins
Variante 4	5	2 hours, 54 mins
Variante 5	7	2 hours, 53 mins
Variante 6	7	1 hour, 53 mins
Variante 7	7	1 hour, 50 mins

<i>Variante 8</i>	5	1 hour, 52 mins
<i>Variante 9</i>	5	2 hours, 6 mins
<i>Variante 10</i>	5	1 hour, 41 mins

4.5.5 Mobilitätsprozesse

Unter dem Reiter ‚Cases‘ konnten die Mobilitätsprozesse der Varianten (s. Anhang, Abbildungen 92-101) betrachtet werden. Diese wurden verwendet, um die Prozesskennzahl V berechnen zu können. In Tabelle 12 wurden die Mobilitätsprozesse der Varianten gesammelt.

Alle zehn Varianten begannen mit einer Autofahrt, einem anschließenden Zwischenstopp sowie einer weiteren Fahrt mit dem Auto. Bei sechs Varianten endete dieser Prozess mit der Erreichung des Zielortes. Bei den anderen vier Mobilitätsprozessen folgte auf die zweite Fahrt ein weiterer Zwischenstopp, bevor die Fahrt zum Zielort fortgeführt wurde.

Tabelle 12: Auswertung Fall 3 - Mobilitätsprozesse der Varianten. Eigene Darstellung (2022)

<i>Variante</i>	<i>Mobilitätsprozess</i>
<i>Variante 1</i>	Start → Driving A → Gebäude C → Driving G → Ziel
<i>Variante 2</i>	Start → Driving A → Gebäude C → Driving H → Gebäude F → Driving I → Ziel
<i>Variante 3</i>	Start → Driving M → Gebäude G → Driving N → Ziel
<i>Variante 4</i>	Start → Driving A → Gebäude C → Driving B → Ziel
<i>Variante 5</i>	Start → Driving A → Gebäude C → Driving C → Ort D → Driving D → Ziel
<i>Variante 6</i>	Start → Driving A → Gebäude C → Driving E → Gebäude E → Driving F → Ziel
<i>Variante 7</i>	Start → Driving A → Gebäude C → Driving J → Gebäude E → Driving F → Ziel
<i>Variante 8</i>	Start → Driving A → Gebäude C → Driving K → Ziel
<i>Variante 9</i>	Start → Driving A → Gebäude C → Driving L → Ziel
<i>Variante 10</i>	Start → Driving O → Gebäude E → Driving F → Ziel

4.5.6 Berechnung der Prozesskennzahl V

In Tabelle 25 (s. Anhang) wurden die Ergebnisse der Berechnungen der KPI V aufgelistet. Diese wurden berechnet, indem die durchschnittliche Dauer (s. Kapitel 4.5.4) und die Mobilitätsprozesse (s. Kapitel 4.5.5) der Varianten verwendet wurden.

In Abbildung 40 wurden die Werte, welche berechnet wurden, visualisiert. Dort ist zu sehen, dass die zehnte Variante in etwa genau so lang wie die kürzeste, von Google Maps vorgeschlagene Dauer war. Mit einem kleinen Unterschied folgten die dritte und die erste Variante. Bei Variante 4 ist ein deutlicher Anstieg zu erkennen.

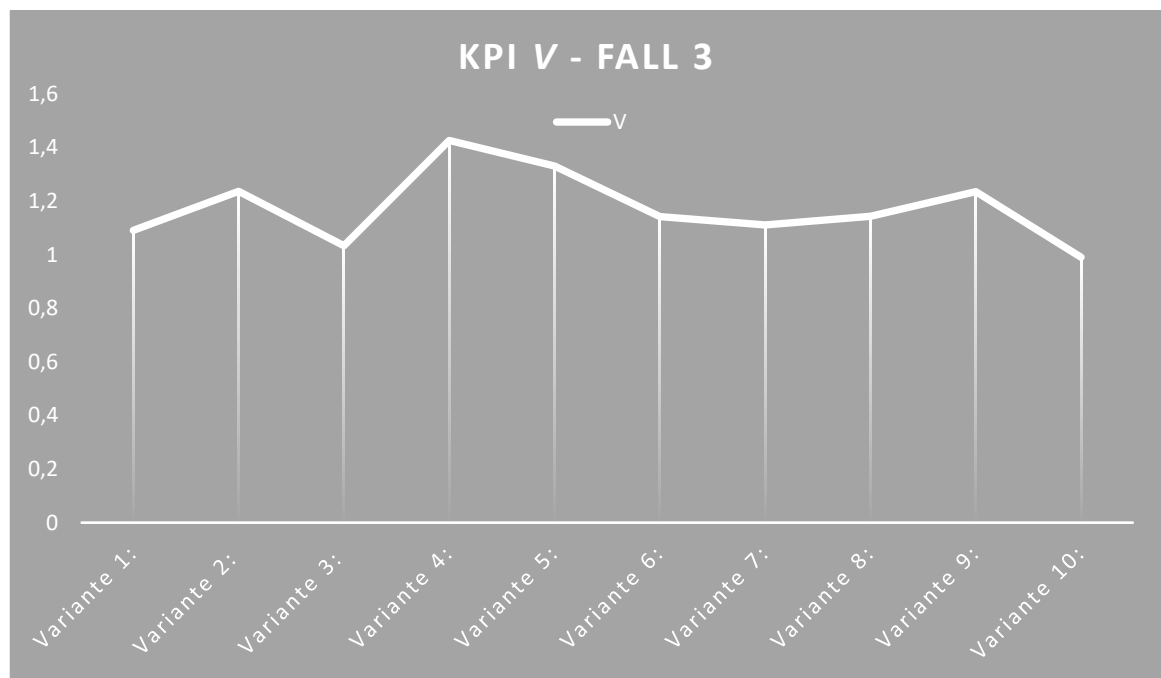


Abbildung 40: Auswertung Fall 3 - Ergebnisse KPI V . Eigene Darstellung (2022)

4.5.7 Berechnung der Prozesskennzahl G

In Tabelle 26 (s. Anhang) wurden die Berechnungsergebnisse der Prozesskennzahlen G und D aufgelistet. Um G zu berechnen, kamen die Entfernung und die Dauer der jeweiligen Cases (s. Kapitel 4.5.3) zum Einsatz.

Die Ergebnisse haben gezeigt, dass die Cases 16, 7, 4 und 9 die höchsten Durchschnittsgeschwindigkeiten hatten (s. Abbildung 41). Die langsamsten Cases hingegen waren eindeutig die Cases 1 und 2, deren Werte mit Abstand unter den restlichen Werten lagen.

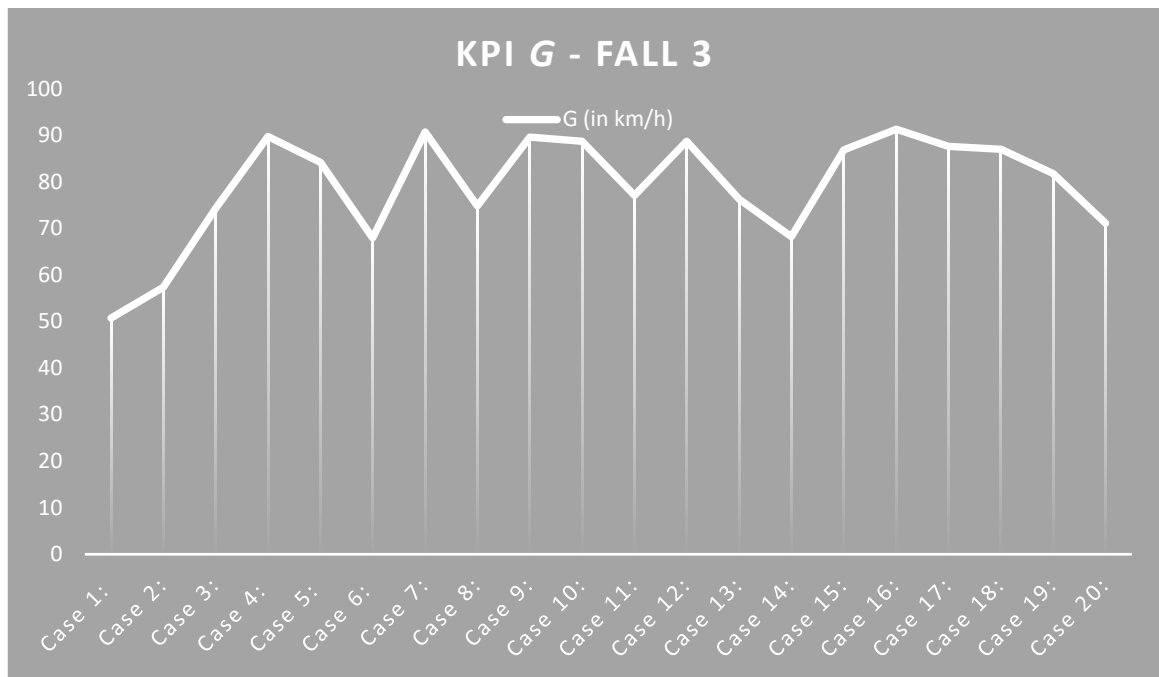


Abbildung 41: Auswertung Fall 3 - Ergebnisse KPI G. Eigene Darstellung (2022)

4.5.8 Berechnung der Prozesskennzahl D

Zur Berechnung von D wurden die Dauer der einzelnen Cases (s. Kapitel 4.5.3) sowie die Durchschnittsdauer aller Cases (s. Kapitel 4.5.1) verwendet und die Ergebnisse anschließend in Tabelle 26 (s. Anhang) zusammengefasst.

Es stellte sich heraus, dass von insgesamt 20 Cases zwölf Cases unter und acht Cases über der Durchschnittszeit von 111,6 Minuten lagen.

Die kürzesten Cases waren der siebte und der 16. Case, gefolgt von den Cases 4, 9, 10, 12, 15, 17 und 18. An all diesen Stellen sind im Graphen in Abbildung 42 die Tiefpunkte zu erkennen. Ebenfalls unter dem Durchschnitt lagen die Cases 5, 19 und 11. Am längsten gedauert haben die Cases 1 und 2, an deren Stellen eindeutige Höchstwerte zu erkennen sind. Diese haben mehr als 1,5-mal so lange wie der Durchschnitt gebraucht.

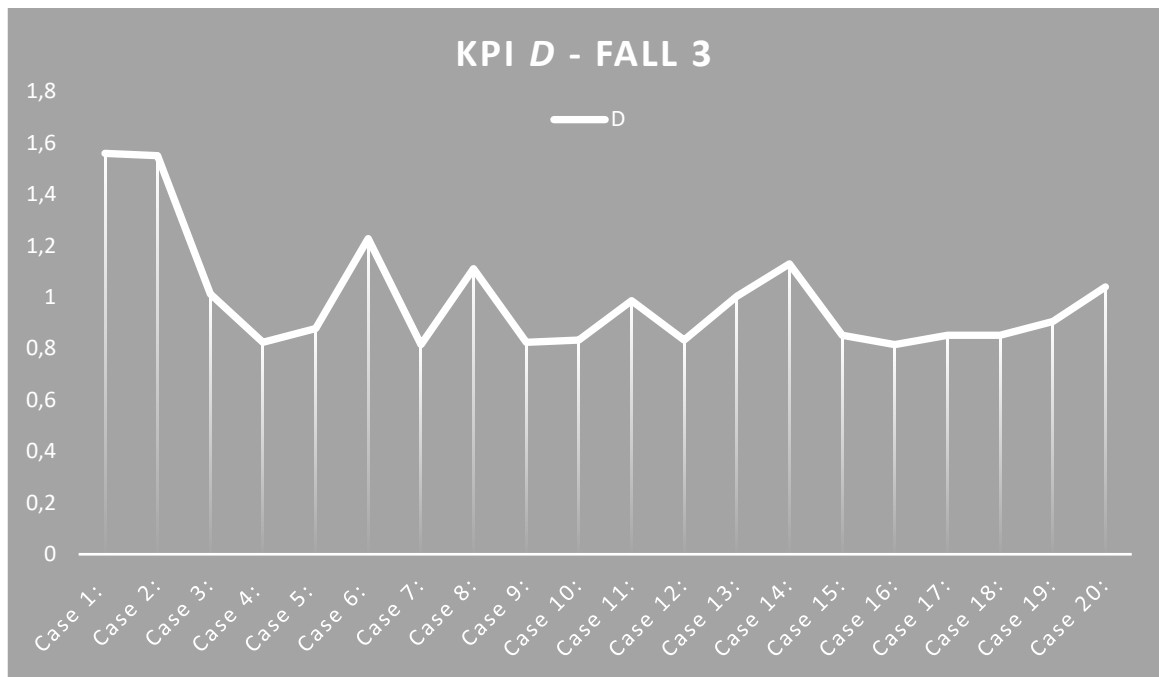


Abbildung 42: Auswertung Fall 3 - Ergebnisse KPI D. Eigene Darstellung (2022)

4.5.9 Filter: Kürzeste Cases

Der kürzeste Case dieses Event-Logs dauerte eine Stunde und 31 Minuten. Aus diesem Grund wurde nach allen Fällen gesucht, die maximal eine Stunde und 33 Minuten gedauert haben.

Auf den Filter trafen vier Cases zu, was 20% aller Cases entspricht (s. Anhang, Abbildung 102). Tabelle 13 fasst die Case IDs, die Dauer und die Varianten der Cases zusammen.

Tabelle 13: Auswertung Fall 3 - Ergebnisse des Filters 'Case duration max. 93min'. Eigene Darstellung (2022)

<i>Case ID</i>	<i>Dauer</i>	<i>Variante</i>
4	1h 32min	1
7	1h 31min	1
9	1h 32min	1
16	1h 31min	2

Nach Anwendung des Filters hatte die erste Variante eine durchschnittliche Durchlaufzeit von einer Stunde und 32 Minuten. Die Durchschnittszeit der zweiten Variante lag bei einer Stunde und 31 Minuten, also identisch zur Gesamtdauer von Case 16.

4.5.10 Filter: Kürzer als Durchschnitt

Die Durchschnittszeit aller Cases lag bei 111,6 Minuten. Deshalb wurde hier nach einer maximalen Zeit von 111,6 Minuten (= eine Stunde 51 Minuten) gefiltert, um die Cases zu erhalten, welche kürzer als der Durchschnitt waren.

12 Cases erfüllten die Bedingungen des Filters. Dies entspricht 60% aller Fälle. Davon wurden acht Cases (66,67%) der ersten Variante zugeschrieben, zwei Cases (16,67%) der zweiten Variante, und jeweils ein Case (je 8,33%) der dritten und vierten Variante (s. Anhang, Abbildung 103).

4.5.11 Filter: Cases mit Wartezeiten

Um alle Fälle, bei denen Wartezeiten aufgetreten sind, herauszufiltern, wurde nach einer durchschnittlichen Wartezeit von mindestens einer Millisekunde gesucht.

Gefunden wurden vier Cases. Es traten also bei 20% aller Fälle Wartezeiten auf. Jeder dieser Cases (jeweils 25%) wurde einer anderen Variante (Varianten 1-4) zugeordnet (s. Anhang, Abbildung 104).

4.6 Fall 4

Der folgende Case handelte von einem täglichen Arbeitsweg in einen etwa 28 Kilometer entfernten Ort. Die Strecke wurde hauptsächlich mit dem Zug zurückgelegt. Nach Ankunft am Bahnhof der Zielstadt wurde die verbleibende Strecke zu Fuß gegangen.

4.6.1 Allgemeine Informationen

Der vierte Fall bestand aus **32 Cases**, 17 Aktivitäten und 195 Ereignissen. Die Durchschnittsdauer der Fälle lag bei **61,2 Minuten** (s. Anhang, Abbildung 105).

Insgesamt sind **acht** Varianten aufgetreten, wovon **fünf** (Varianten 4 bis 8) nur **einmalig** vorgekommen sind. (s. Anhang, Abbildung 106).

4.6.2 Prozessmodell

Abbildung 43 zeigt das Prozessmodell des vierten Falles. Dieses wurde benötigt, um die Hypothese H1 zu prüfen.

Anhand der Farben zeichnet sich auf den ersten Blick ein bestimmter Mobilitätsprozess ab, welcher am häufigsten aufgetreten ist. Dieser ist *Start* → *Walking A* → *Haltestelle A* → *Zugfahrt A* → *Walking G* → *Ziel*. Alle verbliebenen Aktivitäten traten maximal drei Mal auf.

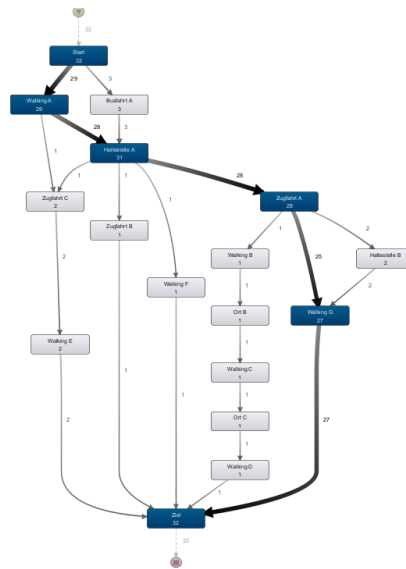


Abbildung 43: Auswertung Fall 4 - Prozessmodell. Eigene Darstellung (2022)

4.6.3 Dauer und zurückgelegte Entfernung der Cases

Zur Berechnung der Prozesskennzahlen G und D wurden unter anderem die Dauer (KPI G , KPI D) sowie die zurückgelegte Entfernung (KPI G) der einzelnen Cases (s. Anhang, Abbildung 107) benötigt. Beides konnte unter dem Reiter ‚Cases‘ entnommen werden und wurde in Tabelle 14 zusammengefasst. Die Entfernung der Cases lag zwischen 23664 und 26919 Metern. Die Dauer bewegte sich zwischen 43 Minuten 37 Sekunden und drei Stunden 23 Minuten.

Tabelle 14: Auswertung Fall 4 - Dauer und Entfernung der Cases. Eigene Darstellung (2022)

Case	Dauer	Zurückgelegte Entfernung (in Metern)
1	2 hours, 57 mins	$792 + 22035 + 1693 + 753 + 1117 = 26390$
2	43 mins, 37 secs	$2677 + 23002 = 25679$
3	47 mins, 6 secs	$1106 + 17019 + 5751 = 23876$
4	54 mins	$812 + 22852 = 23664$
5	1 hour, 6 mins	$778 + 21725 + 1374 = 23877$
6	59 mins	$1387 + 21906 + 1412 = 24705$
7	50 mins, 36 secs	$2272 + 21779 + 1577 = 25628$
8	48 mins, 45 secs	$825 + 21714 + 1310 = 23849$
9	56 mins, 17 secs	$827 + 21712 + 1327 = 23866$
10	58 mins, 15 secs	$792 + 22251 + 1253 = 24296$
11	55 mins, 37 secs	$772 + 21630 + 1421 = 23823$
12	54 mins, 3 secs	$795 + 21803 + 1455 = 24053$

13	50 mins, 8 secs	$793 + 21977 + 1815 = 24585$
14	46 mins, 3 secs	$816 + 21649 + 1308 = 23773$
15	54 mins, 50 secs	$780 + 21882 + 1027 = 23689$
16	53 mins, 41 secs	$949 + 17757 + 5253 = 23959$
17	48 mins, 31 secs	$860 + 21665 + 1439 = 23964$
18	50 mins, 9 secs	$747 + 21723 + 1378 = 23848$
19	49 mins, 23 secs	$779 + 21760 + 1214 = 23753$
20	51 mins, 21 secs	$753 + 21820 + 1204 = 23777$
21	49 mins, 38 secs	$845 + 21853 + 1316 = 24014$
22	49 mins, 48 secs	$819 + 21870 + 1478 = 24167$
23	51 mins, 49 secs	$799 + 22100 + 1335 = 24234$
24	51 mins, 43 secs	$815 + 22413 + 1043 = 24271$
25	47 mins, 24 secs	$751 + 21879 + 1417 = 24047$
26	1 hour, 45 secs	$787 + 24342 + 1227 = 26356$
27	54 mins, 49 secs	$794 + 21679 + 1481 = 23954$
28	3 hours, 23 mins	$1355 + 22390 + 475 = 24220$
29	51 mins, 48 secs	$769 + 22161 + 1424 = 24354$
30	56 mins, 28 secs	$956 + 21833 + 1373 = 24162$
31	51 mins, 52 secs	$945 + 21804 + 1493 = 24242$
32	54 mins, 32 secs	$3684 + 21663 + 1572 = 26919$

4.6.4 Durchschnittliche Dauer und Ereignisse pro Variante

Aus der Tabelle ‚Variants‘ (s. Anhang, Abbildung 108) konnten Informationen über die Anzahl der Ereignisse pro Variante sowie über die durchschnittliche Dauer der einzelnen Varianten entnommen werden, welche für die Prüfung der zweiten Hypothese von Bedeutung waren. Die Durchschnittsdauer wurde außerdem zur Berechnung der Prozesskennzahl V verwendet. Die Anzahl der Ereignisse sowie die durchschnittliche Dauer der jeweiligen Varianten wurden in Tabelle 15 zusammengefasst.

Alle Varianten beinhalteten zwischen fünf und maximal zehn Ereignisse. Von den insgesamt acht Mobilitätsprozessen bestanden drei aus jeweils fünf Ereignissen, wieder drei aus jeweils sechs Ereignissen, sowie jeweils eine Variante aus sieben und zehn Varianten. Die Durchschnittsdauer bewegte sich zwischen 43 Minuten 37 Sekunden und maximal zwei Stunden 57 Minuten.

Tabelle 15: Auswertung Fall 4 - durchschnittliche Dauer und Anzahl der Ereignisse pro Variante. Eigene Darstellung (2022)

Variante	Anzahl der Ereignisse	Ø Dauer
Variante 1	6	53 mins, 10 secs
Variante 2	6	1 hour, 44 mins
Variante 3	7	49 mins, 46 secs
Variante 4	10	2 hours, 57 mins
Variante 5	5	43 mins, 37 secs
Variante 6	6	47 mins, 6 secs
Variante 7	5	54 mins
Variante 8	5	53 mins, 41 secs

4.6.5 Mobilitätsprozesse

Um die Prozesskennzahl V berechnen zu können, waren unter anderem die Mobilitätsprozesse der einzelnen Varianten (s. Anhang, Abbildungen 109-116) von Bedeutung. In Tabelle 16 wurden diese gesammelt.

In sieben von acht Fällen begann der Prozess mit einem Fußweg. Die verbliebene Variante hingegen startete mit einer Busfahrt. Ebenfalls in sieben Fällen folgte auf den Fußweg bzw. die Busfahrt ein Aufenthalt an einer Haltestelle, worauf in sechs Fällen eine Zugfahrt folgte. Bei Variante 5 endete an dieser Stelle der Prozess mit der Erreichung des Ziels. Bei der siebten Variante folgte auf die Haltestelle eine Walking-Aktivität zum Zielort, während bei der achten Variante auf den ersten Fußweg direkt eine Zugfahrt und ein weiterer Fußweg bis zum Zielort folgten. Bei den Varianten 1 und 2 wurde die Zugfahrt mit einer Walking-Aktivität fortgeführt, welche zum Ziel führte. Bei der dritten Variante hingegen gab es zwischen der Zugfahrt und dem Fußweg zum Ziel einen Aufenthalt an der Zielhaltestelle. Abschließend folgten bei der vierten Variante auf die Zugfahrt Walking Aktivitäten, welche durch zwei Zwischenstopps unterbrochen wurden.

Tabelle 16: Auswertung Fall 4 - Mobilitätsprozesse der Varianten. Eigene Darstellung (2022)

Variante	Mobilitätsprozess
Variante 1	Start → Walking A → Haltestelle A → Zugfahrt A → Walking G → Ziel
Variante 2	Start → Busfahrt A → Haltestelle A → Zugfahrt A → Walking G → Ziel
Variante 3	Start → Walking A → Haltestelle A → Zugfahrt A → Haltestelle B → Walking G → Ziel

<i>Variante 4</i>	Start → Walking A → Haltestelle A → Zugfahrt A → Walking B → Ort B → Walking C → Ort C → Walking D → Ziel
<i>Variante 5</i>	Start → Walking A → Haltestelle A → Zugfahrt B → Ziel
<i>Variante 6</i>	Start → Walking A → Haltestelle A → Zugfahrt C → Walking E → Ziel
<i>Variante 7</i>	Start → Walking A → Haltestelle A → Walking F → Ziel
<i>Variante 8</i>	Start → Walking A → Zugfahrt C → Walking E → Ziel

4.6.6 Berechnung der Prozesskennzahl V

Zur Berechnung der KPI V wurden die Mobilitätsprozesse der Varianten (s. Kapitel 4.6.5) sowie die Durchschnittsdauer ebendieser (s. Kapitel 4.6.4) genutzt. Tabelle 27 (s. Anhang) enthält die Ergebnisse der Berechnungen. Die folgende Abbildung 44 stellt die Werte außerdem in graphischer Form dar.

Der mit Abstand niedrigste Wert konnte bei der siebten Variante aufgewiesen werden. Auch bei den Varianten 6 und 8 betrug der errechnete Wert weniger als 0,5. Die fünfte Variante lag ebenfalls bei einem Wert unter der ,1'. Bei der dritten Variante stimmten die beiden Werte für die tatsächliche und die schnellste Dauer nahezu überein. Variante 4 erreichte den mit Abstand höchsten Wert.

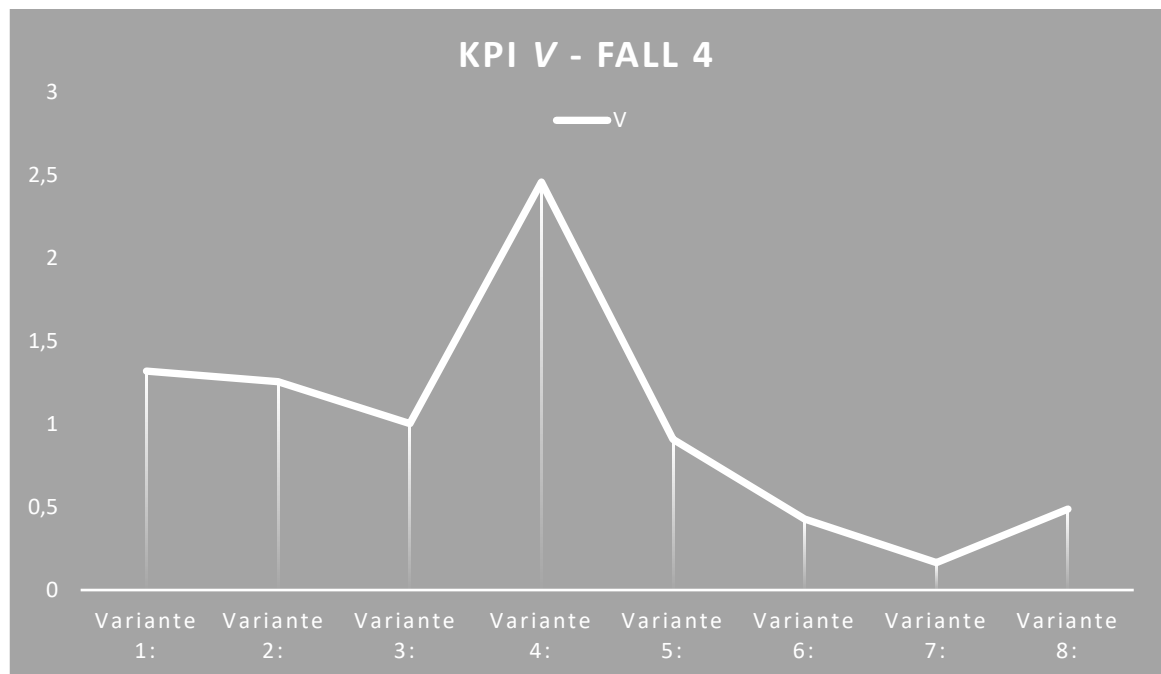


Abbildung 44: Auswertung Fall 4 - Ergebnisse KPI V. Eigene Darstellung (2022)

4.6.7 Berechnung der Prozesskennzahl G

Um die Prozesskennzahl G zu berechnen, wurde die Dauer der einzelnen Cases verwendet. Außerdem war die Entfernung der Cases (s. Kapitel 4.6.3) von Bedeutung. Die Ergebnisse wurden in die Tabelle 28 (s. Anhang) eingetragen.

In Abbildung 45 wurden die Ergebnisse außerdem grafisch dargestellt. Der schnellste Case war mit Abstand der zweite, wie man anhand der Spitze im Graphen erahnen kann. Darauf folgte mit über vier km/h Abstand Case 14. Knapp dahinter liegen die Cases 25, 3 und 7. Bei Case 1 und Case 28 kann man stattdessen die eindeutig niedrigsten Werte ablesen.

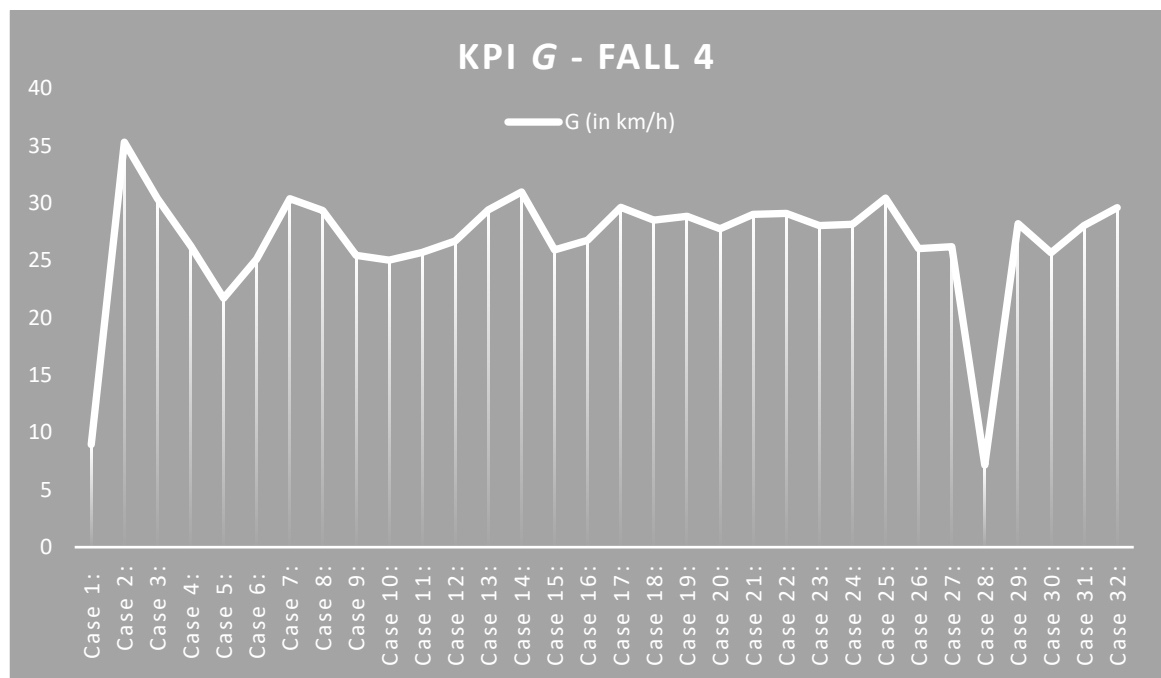


Abbildung 45: Auswertung Fall 4 - Ergebnisse KPI G. Eigene Darstellung (2022)

4.6.8 Berechnung der Prozesskennzahl D

Neben der Dauer aller Cases (s. Kapitel 4.6.3) war für die Berechnung der KPI D die Durchschnittsdauer aller Cases (s. Kapitel 4.6.1) von Bedeutung. Die Berechnungsergebnisse wurden in Tabelle 28 (s. Anhang) zusammengefasst.

Die Berechnungen von D haben gezeigt, dass von insgesamt 32 Cases 29 Stück unter der Durchschnittsdauer von 61,2 Minuten liegen. Nur drei Cases haben die Durchschnittszeit überschritten.

Case 2 war der kürzeste Case des Event-Logs, gefolgt von Case 14, Case 3 und Case 25. Die Cases 1 und 28 dauerten hingegen ein Vielfaches der durchschnittlichen Dauer, was in Abbildung 46 deutlich zu sehen ist. Ebenfalls die Durchschnittsdauer überschritten hat Case 5, bei dem jedoch nur ein leichter Anstieg zu erkennen ist.

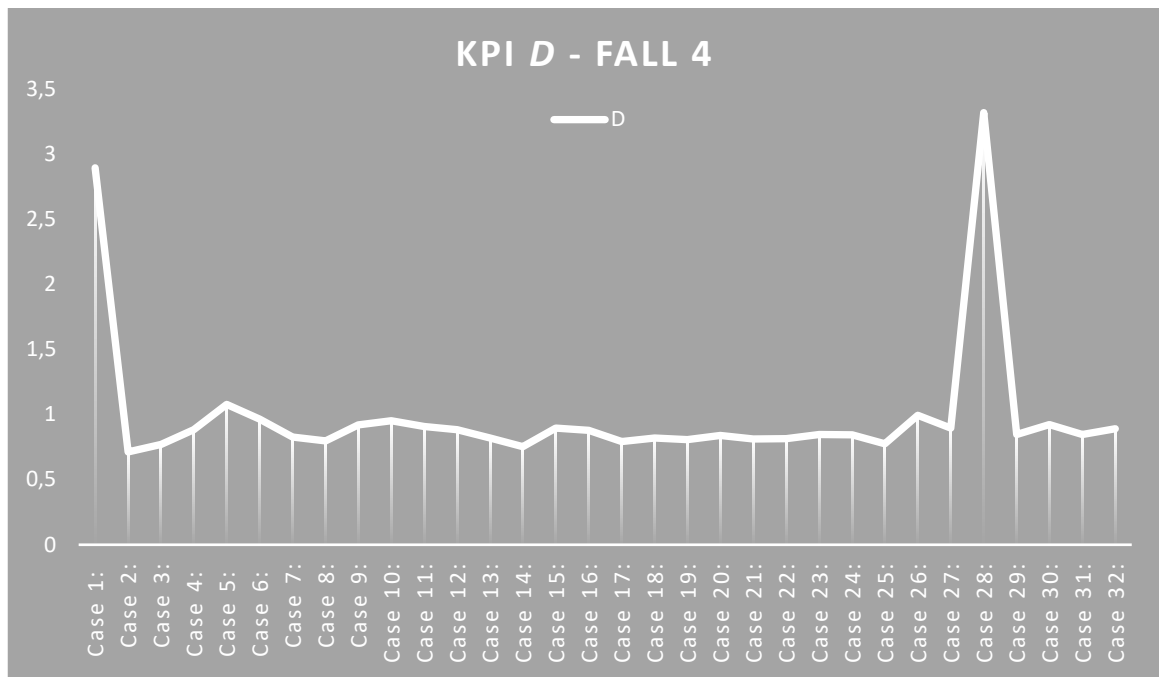


Abbildung 46: Auswertung Fall 4 - Ergebnisse KPI D. Eigene Darstellung (2022)

4.6.9 Filter: Kürzeste Cases

Hier wurde nach Fällen gesucht, welche maximal 47 Minuten gedauert haben, da der kürzeste Case eine Länge von 43 Minuten 37 Sekunden hatte.

Leidlich zwei Cases trafen auf das Suchkriterium zu. Dies sind sechs Prozent aller Cases (s. Anhang, Abbildung 117). In Tabelle 17 sind die Case ID, die Dauer und die Variante der beiden Cases zu finden.

Tabelle 17: Auswertung Fall 4 - Ergebnisse des Filters 'Case duration max. 47min'. Eigene Darstellung (2022)

Case ID	Dauer	Variante
2	43min 37sek	1
14	46min 3sek	2

Die durchschnittliche Durchlaufzeit der zwei Varianten entsprach hier der Dauer der beiden Cases.

4.6.10 Filter: Kürzer als Durchschnitt

Um alle Cases, die kürzer als der Durchschnitt waren, zu erhalten, wurde nach einer maximalen Dauer von 61,2 Minuten gefiltert, da dies der Durchschnittszeit aller Cases entsprach.

Insgesamt 29 Cases, also 90% aller Fälle, trafen auf den Filter zu. 21 Fälle (72,41%) gehörten zur ersten Variante, jeweils zwei Cases (je 6,9%) zu den Varianten 2 und 3, und jeweils ein Case (je 3,45%) zu den Varianten 4 bis 7 (s. Anhang, Abbildung 118).

4.6.11 Filter: Cases mit Wartezeiten

Hier wurde nach den Cases gesucht, bei denen Wartezeiten auftraten. Es wurde nach einer durchschnittlichen Wartezeit von mindestens einer Millisekunde gefiltert. Jedoch konnten keine Ergebnisse gefunden werden (s. Anhang, Abbildung 119).

4.7 Prüfung der aufgestellten Hypothesen

Abschließend konnten auf Grundlage der über die betrachteten Fälle gesammelten Informationen die drei in Kapitel 2.3.4 aufgestellten Hypothesen verifiziert bzw. falsifiziert werden.

4.7.1 Hypothese H1

Anhand der Anzahl der Cases pro Event-Log (s. Kapitel 4.3.1, 4.4.1, 4.5.1 und 4.6.1) sowie der Prozessmodelle der jeweiligen Fälle (s. Kapitel 4.3.2, 4.4.2, 4.5.2 und 4.6.2) konnte die Hypothese H1 „Je mehr Cases ein Event-Log enthält, desto umfangreicher ist das entdeckte Prozessmodell“ überprüft werden.

Es konnte festgestellt werden, dass diese Hypothese nicht zutreffend war.

4.7.2 Hypothese H2

Die Hypothese H2 „Je mehr Ereignisse eine Variante enthält, desto länger dauert diese durchschnittlich im Vergleich zu Varianten mit einer geringeren Anzahl an Ereignissen“ wurde anhand der Ergebnisse zur Anzahl der Ereignisse pro Variante sowie der durchschnittlichen Dauer der Varianten (s. Kapitel 4.3.4, 4.4.4, 4.5.4 und 4.6.4) geprüft.

Auch diese Hypothese erwies sich als unwahr.

4.7.3 Hypothese H3

Abschließend wurde die dritte Hypothese „Je mehr Varianten in einem Event-Log auftreten, desto höher ist der Anteil der Varianten, welche nur einmal aufgetreten sind“ unter Verwendung der Gesamtanzahl der Varianten pro Event-Log sowie der Anzahl derer Varianten, welche nur einmalig aufgetreten sind (s. Kapitel 4.3.1, 4.4.1, 4.5.1 und 4.6.1), geprüft.

Die Hypothese H3 konnte ebenfalls falsifiziert werden.

5. Diskussion

Um die Auswirkungen der Anwendung des Process Mining Ansatzes auf Mobilitätsprozesse zu untersuchen, wurden anhand einer Mixed-Methods Forschung die alltäglichen Mobilitätsprozesse von Personen mithilfe des Process Mining-Tools *Disco* analysiert. Hierfür wurde eine qualitative Fallstudie zum Fall *Der Mobilitätsprozess einer Person von oder zu ihrer Arbeitsstätte bzw. von oder zu ihrem Arbeitsort* durchgeführt. Um die Event-Logs, welche essenziell für die Anwendung von Process Mining-Techniken sind, erstellen zu können, wurden die bereits verfügbaren Mobilitätsdaten von vier berufstätigen Personen aus ihrer Google Maps Timeline exportiert und so bearbeitet, dass die Datensätze für Process Mining nutzbar waren. Dabei unterschied sich der betrachtete Zeitraum je nach Person. Anschließend wurden die Ereignisprotokolle in *Disco* importiert und einige der Informationen, welche durch das Process Mining-Tool zur Verfügung standen, anhand quantitativer Methoden ausgewertet.

5.1 Zusammenfassung und Interpretation der Ergebnisse

Die Ergebnisse der Forschung haben gezeigt, dass die Anwendung von Process Mining-Techniken auf Mobilitätsprozessdaten einen detaillierten Einblick auf die sich dahinter verbergenden Mobilitätsprozesse bieten kann. Die zur Verfügung gestellten Informationen über die Prozesse konnten des Weiteren dazu genutzt werden, diese zu bewerten.

In den folgenden Kapiteln werden die Ergebnisse zusammengefasst und interpretiert.

5.1.1 Möglichkeiten des Process Mining-Tools *Disco*

Die erste Teilfrage dieser Arbeit verfolgte das Interesse herauszufinden, welche Erkenntnisse sich über die alltäglichen Mobilitätsdaten von Personen mithilfe des Process Mining-Tools *Disco* gewinnen und ohne weiteren Aufwand daraus ablesen lassen. In Kapitel 4.1 wurde deshalb ein Überblick über die Möglichkeiten des Tools *Disco* geschaffen.

Die Ergebnisse haben gezeigt, dass das Tool es dem Nutzer ermöglichte, ein Prozessmodell des betrachteten Falles zu sehen, es je nachgewünschter Granularität anzupassen und durch Auswählen von den Aktivitäten oder Pfaden detaillierte Informationen über die Häufigkeit und die Performance dieser zu erhalten. Des Weiteren werden die vom Nutzer ausgewählten Werte über die Häufigkeit (*Frequency*) oder die Performance im Prozessmodell farblich dargestellt, sodass bereits, ohne einen Blick auf die Zahlen zu werfen, intuitiv ein Überblick gewonnen werden kann.

An dieser Stelle fiel auf, dass das Tool vermeintlich zu lang dauernde Aktivitäten und Pfade in einer dunkelroten Farbe im Prozessmodell markierte. Ob diese tatsächlich einen Engpass darstellten, lag jedoch im Ermessen des Betrachters, da das Tool keine Kenntnis über die optimale Dauer der Aktivitäten/Pfade verfügte.

Das Tool entdeckte außerdem die unterschiedlichen Mobilitätsprozesse des untersuchten Event-Logs. Weiterhin zählte „Disco“ die absoluten und berechnete die relativen Häufigkeiten der einzelnen Aktivitäten und berechnete unter anderem die minimale, durchschnittliche, mediane und maximale Dauer der Aktivitäten, Cases und Varianten sowie die Anteile der Ereignisse und der Varianten an der Gesamtheit aller Cases. Diese Informationen wurden in diversen Statistiken, Graphen und Tabellen zusammengefasst. Es ist aufgefallen, dass diese entweder den Aspekt Häufigkeit oder den Aspekt Zeit behandelten. Dies ist dadurch zu erklären, dass die aufgezeichneten Daten, welche im Event-Log genutzt wurden, lediglich die Attribute Case ID, Aktivität sowie die Beginn- und Schluss-Zeitstempel enthielten.

In der Einleitung dieser Arbeit wurde die Erwartung geäußert, dass man auf Grundlage der Informationen aus dem Tool in der Lage ist, Entscheidungen über die Mobilitätsprozesse zu treffen. Durch die Verfügbarkeit von Auswertungen in Form von Statistiken und Tabellen ist dies gegeben. Es ist beispielsweise möglich, eine Entscheidung darüber zu treffen, welche der aufgetretenen Varianten am zeitlich effizientesten war, da das Tool die Durchschnittsdauern der Varianten bereitstellt und diese so miteinander verglichen werden können.

Zudem ließen sich Statistiken über jene Attribute anzeigen, welche zu Beginn als „Other“ deklariert wurden. Im Falle der hier untersuchten Mobilitätsdaten handelte es sich um das Attribut „Entfernung in m“. Die Entfernungsdaten, welche von Google Maps aufgezeichnet worden sind, enthielten bei allen Fällen so viele unterschiedliche Werte, dass diese in der Regel nur ein bis zwei Mal vorgekommen sind. Dies führte dazu, dass die Statistiken und Graphen dieses Attributes nicht sinnvoll nutzbar waren. Zudem konnte das Tool die Entfernung nicht mit anderen Attributen in Verbindung bringen, da es sich um ein sonstiges Attribut handelte.

Weiterhin hatte man die Möglichkeit, jeden der aufgetretenen Cases einzeln zu betrachten. Dabei konnte man unter anderem ein Prozessmodell und eine Tabelle des spezifischen Cases sehen, welche Informationen wie die Aktivitäten, deren Startdatum sowie -uhrzeit, die Dauer, die aufgetretenen Wartezeiten und die zurückgelegte Entfernung zusammenfassten.

Außerdem ließen sich die Cases hier den aufgetretenen Mobilitätsprozessen bzw. Varianten zuordnen.

An dieser Stelle ist zu erwähnen, dass zu Beginn davon ausgegangen wurde, dass die Mobilitätsprozesse, welche in den Ereignisprotokollen aufgetreten sind, vom Tool zuverlässig erkannt und ausgewertet werden können. Diese Erwartung konnte durch die inhaltliche Analyse des Tools bestätigt werden, da die Mobilitätsprozesse sowohl erkannt als auch analysiert worden sind. Das Tool stellte etwa zu jedem aufgetretenen Mobilitätsprozess ein passendes Prozessmodell sowie Informationen, wie z.B. die Anzahl der Cases und Events und die durchschnittliche Dauer, bereit.

Ferner ließen sich die Ergebnisse im Tool filtern, sodass nur ausgewählte Informationen und Prozessmodelle angezeigt werden konnten, welche einem gewissen Interesse entsprachen.

Bei den Analyseergebnissen des Tools handelte es sich eher um eine objektive Zusammenfassung und übersichtliche Darstellung und Visualisierung von Daten sowie zahlreichen Berechnungen in Form von Modellen, Diagrammen und Tabellen. Eine Wertung bzw. Interpretation der Ergebnisse fand durch das Tool nicht statt und blieb dem Nutzer überlassen, da das Tool nicht die nötigen Möglichkeiten sowie keine Kenntnis über beispielsweise die optimale Dauer besaß. Dennoch waren die verfügbaren Informationen sehr umfangreich und boten dem Nutzer zahlreiche Möglichkeiten, diese innerhalb oder auch außerhalb des Tools weiter zu verwenden.

Weiterhin ist aufgefallen, dass die in dieser Arbeit analysierten Mobilitätsprozessdaten alle historischer Natur waren. Dies hat zur Folge, dass alle Bewertungen und Entscheidungen, welche auf Grundlage der Process Mining-Auswertungen getroffen werden können, auf historischen Daten basieren. Im Gegensatz dazu ist es Google Maps möglich, Routenvorschläge und deren erwartete Dauer auf Basis von Echtzeit-Daten bereitzustellen, wie aus der Publikation von He et al. (2019) (s. Kapitel 2.3.2) hervorging. Dies führt zu der Erkenntnis, dass die prädiktive Routenplanung von Google Maps vermutlich besonders geeignet ist, wenn unbekannte Strecken zurückgelegt werden sollen. Auf der anderen Seite eignet sich die historische Analyse durch Process Mining-Techniken, um bekannte Strecken zu bewerten und die für einen persönlich effizienteste Route auf Grundlage der historischen Daten zu ermitteln.

5.1.2 Betrachtung der einzelnen Fälle und Prüfung der Hypothesen

Wie bereits erwähnt wurde, konnten die aus dem Tool entnehmbaren Informationen weiter genutzt werden, um den Event-Log bzw. die darin aufgetretenen Mobilitätsprozesse zu

bewerten und somit neue Erkenntnisse zu sammeln. Das Tool selbst bot z.B. die Möglichkeit, die Ergebnisse nach ausgewählten Attributen zu filtern und sich dadurch nur die Informationen anzeigen zu lassen, welche von Interesse für die Beantwortung gewisser Fragestellungen waren.

Zum Zwecke der Beantwortung der zweiten Teilfrage „Wie lassen sich die aus dem Tool gewonnenen Informationen nutzen, um den/die Mobilitätsprozess/e zu bewerten?“ wurden deshalb anhand der Informationen über die Mobilitätsprozesse aus dem Tool jeweils drei Prozesskennzahlen berechnet sowie Hypothesen überprüft, welche zur Bewertung von Mobilitätsprozessen herangezogen werden können. Zusätzlich wurden drei Filter in ‚Disco‘ angewandt.

5.1.2.1 Prozesskennzahl V

Die KPI *V* wurde berechnet, um das Verhältnis zwischen der tatsächlich gebrauchten Zeit für die Zurücklegung einer Strecke und der laut Google Maps schnellsten Dauer für die Zurücklegung derselben Strecke bewerten zu können.

Ein Wert von ‚1‘ bedeutete, dass die tatsächliche und die schnellste Zeit übereinstimmten. Ein Wert von ‚>1‘ zeigte, dass es in Realität länger gedauert hat, diese Strecke zurückzulegen. Demzufolge zeigte ein Wert von ‚<1‘, dass die Strecke in Realität schneller zurückgelegt wurde, als es laut Google Maps der Fall sein sollte. Ein niedriger Wert könnte erstrebenswert sein, da dies indiziert, dass man in Realität bei Nutzung einer spezifischen Route mit bestimmten Fortbewegungsmitteln schneller vorankommt, als dies von Google Maps berechnet wurde. Sucht man auf der anderen Seite nach Varianten, welche zeitlich in etwa der berechneten Dauer von Google Maps entsprechen, ist ein Wert möglichst nahe an der ‚1‘ erstrebenswert, da dies zeigt, dass die entsprechende Variante ‚termintreu‘ gewesen ist.

5.1.2.1.1 Bewertung der Mobilitätsprozesse

Die Ergebnisse der Berechnung der KPI *V* für den ersten Fall haben gezeigt, dass alle aufgetretenen Varianten im Durchschnitt länger gedauert haben als von Google Maps für dieselbe Route mit denselben Fortbewegungsmitteln vorgeschlagen. Es konnte festgestellt werden, dass die Varianten 1 und 3 am nächsten an der „optimalen“ Dauer waren. Wie im vorangegangenen Absatz bereits erwähnt wurde, ist ein möglichst niedriger oder nahe an der ‚1‘ liegender Wert erstrebenswert. Aus diesem Grund können für den ersten Fall also die erste und dritte Variante empfohlen werden, da diese am ehesten der laut Google Maps

optimalen Dauer für die variantenspezifischen Routen und Fortbewegungsmöglichkeiten entsprochen haben.

Bei der Berechnung für den zweiten Fall hat sich herausgestellt, dass vier der insgesamt 15 Varianten in Realität schneller waren als die von Google Maps berechnete Dauer, allen voran die Varianten 7 und 1. Dies deutet darauf hin, dass diese beiden Varianten empfehlenswert sein könnten, da sie in Realität weniger gedauert haben, als Google Maps für diesen Prozess berechnet hat. Sucht man allerdings Empfehlungen für Varianten, welche der von Google berechneten Zeit am ehesten entsprechen, kann auf Grundlage der Ergebnisse hier die achte Variante empfohlen werden.

Beim dritten Fall konnte nur bei Variante 10 eine in Realität schnellere Zeit festgestellt werden. Bei den übrigen Varianten dauerte die Zurücklegung der Strecke im Durchschnitt länger, als von Google Maps berechnet wurde. Doch auch die Varianten 3 und 1 waren nur minimal über der durch Google vorgegebenen Zeit. Somit sind diese, neben der zehnten Variante, zu empfehlen.

Die Ergebnisse der Berechnungen haben beim vierten Fall gezeigt, dass die Hälfte aller Varianten in Realität eine kürzere Dauer aufgewiesen haben, als von Google Maps berechnet wurde. Eine der Varianten hingegen dauerte mehr als doppelt so lange. Die mit großem Abstand niedrigsten Werte wiesen die Varianten 6, 7 und 8 auf. Anhand der Definition der KPI V wären diese Varianten dementsprechend zu empfehlen. Allerdings fällt bei genauerem Betrachten der genannten Varianten auf, dass diese von Google Maps fehlerhaft aufgezeichnet worden sind. So wurde z.B. bei der siebten Variante ein Fußweg aufgezeichnet, welcher über 22 Kilometer lang war und in einer Zeit von 32 Minuten zurückgelegt wurde (s. Anhang, Abbildung 120). Dies ist selbstverständlich unmöglich. Abgesehen von diesen drei Varianten erreichten die Variante 5 einen Wert von unter 1, sowie die Variante 3 einen Wert von knapp über 1, weshalb sich diese empfehlen lassen.

5.1.2.1.2 Schlussfolgerungen KPI V

Anhand der Ergebnisse der KPI V konnte bewertet werden, welche Mobilitätsprozesse sich eignen, um eine Strecke möglichst schneller oder in der Zeit, die Google Maps für dieselbe Strecke berechnet, zurückzulegen. Bei den Berechnungen ist aufgefallen, dass die Ergebnisse stark davon abhängig sind, ob die Mobilitätsdaten korrekt aufgezeichnet worden sind. Dies hat sich z.B. bei dem vierten Fall gezeigt. Bei der siebten Variante wurde aufgezeichnet, dass eine Strecke von über 22 Kilometern in einer halben Stunde zu Fuß gegangen wurde. Eigentlich wurde der Großteil dieser Strecke mit dem Zug zurückgelegt.

Der Unterschied zwischen der Dauer der Aufzeichnung und der von Google Maps berechneten Dauer war deshalb sehr groß. An dieser Stelle ist also eine fehlerhafte Aufzeichnung entstanden, wodurch sich das Ergebnis der Berechnung der KPI stark verfälscht hat. Dies könnte möglicherweise z.B. an einer zu niedrigen Aufzeichnungsfrequenz liegen, welche hier bewirkte, dass das genutzte Verkehrsmittel nicht erkannt werden konnte. Auch Ribeiro et al. (2020) wiesen in ihrer Arbeit darauf hin, dass durch zu niedrige Frequenzen fehlerhafte Aufzeichnungen entstehen können.

Bei korrekter Aufzeichnung ist die KPI V jedoch ein zuverlässiges Mittel, um zu bewerten, ob die Dauer der Mobilitätsprozesse in Realität kürzer, gleich oder länger war im Vergleich zu der von Google Maps berechneten Dauer. Zu beachten ist allerdings, dass Varianten mit einem niedrigen Wert nicht automatisch auch eine kürzere Gesamtdauer haben. Es wäre denkbar, dass eine Variante, welche einen höheren Wert bei der KPI V erzielt hat, dennoch kürzer in der Dauer ist als eine Variante, welche einen niedrigeren Wert für V hat.

5.1.2.2 Prozesskennzahl G

Anhand der zurückgelegten Entfernung sowie der Dauer der einzelnen Cases konnte die Durchschnittsgeschwindigkeit aller Cases eines Falles berechnet werden. Eine höhere durchschnittliche Geschwindigkeit deutet darauf hin, dass man bei dem jeweiligen Case schnell und mit mutmaßlich wenigen Unterbrechungen vorangekommen ist.

5.1.2.2.1 Bewertung der Cases

Die Ergebnisse haben beim ersten Fall gezeigt, dass sich die Durchschnittsgeschwindigkeit der Cases in einem Bereich von rund zehn bis 23 km/h befand. Berechnet man den Mittelwert aus allen Geschwindigkeiten, ergibt sich ein Wert von rund 16,48 km/h als Durchschnittsgeschwindigkeit für den gesamten Fall (s. Anhang, Abbildung 121). Die höchste Durchschnittsgeschwindigkeit ist bei Case 46 aufgetreten. Auch die Cases 18, 25, 8, 26, 15 und 12 wiesen hohe Werte auf. Betrachtet man weiterhin im Tool, zu welchen Varianten die schnellsten Cases gehören, fällt auf, dass alle Cases (Cases 8, 12, 15, 18, 25, 26) bis auf einen (Case 46 – Variante 2) der ersten Variante zuzuordnen sind. Dies könnte ein Indikator dafür sein, dass man mit dem Mobilitätsprozess der ersten Variante besonders schnell im Vergleich zu anderen Mobilitätsprozessen vorankommt.

Die durchschnittlichen Geschwindigkeiten der Cases des zweiten Falles schwankten zwischen rund 38,48 km/h und 81,84 km/h. Der Durchschnitt aller Cases liegt bei etwa 59,26 km/h (s. Anhang, Abbildung 122). Die Ergebnisse haben gezeigt, dass man bei den Cases 16, 13, 41, 11 und 12 am schnellsten vorangekommen ist. Aus dem Tool kann man daraufhin

erfahren, dass zwei der Cases zu Variante 2 gehören, weitere zwei Cases werden der fünften Variante zugeordnet und der verbleibende Case gehört zu Variante 8. Diese drei Varianten scheinen sich also anzubieten, wenn man möglichst schnell – auf die Geschwindigkeit bezogen – vorankommen möchte.

Beim dritten Fall lag die Mindest-Durchschnittsgeschwindigkeit bei etwa 50,73 km/h, während die höchste Durchschnittsgeschwindigkeit bei etwa 91,32 km/h lag. Der Mittelwert aller Ergebnisse liegt bei rund 79,22 km/h (s. Anhang, Abbildung 123). Die höchsten durchschnittlichen Geschwindigkeiten konnten bei den Cases 16, 7, 4 und 9 aufgewiesen werden. Die letzten drei der genannten Cases lassen sich alle der ersten Variante zuschreiben. Case 16 hingegen gehörte zu der dritten Variante. Hier könnten sich also besonders die Variante 1 anbieten, möchte man mit einer vergleichsweise hohen Geschwindigkeit vorankommen.

Die Durchschnittsgeschwindigkeiten des vierten Falles bewegten sich im Bereich zwischen etwa 7,16 km/h und 35,32 km/h. Im Mittel betrug die Durchschnittsgeschwindigkeit rund 26,69 km/h (s. Anhang, Abbildung 124). Zu den schnellsten Cases gehörten die Cases 2, 14, 25, 3 und 7. Drei dieser Cases ließen sich der ersten Variante zuordnen und jeweils ein Case gehörte zu den Varianten 5 und 6. So lassen sich auch hier drei Mobilitätsprozesse, jedoch besonders die Variante 1, empfehlen, um mit schnellerer Geschwindigkeit unterwegs zu sein.

5.1.2.2.2 Schlussfolgerung KPI G

Mit der Prozesskennzahl G wurde für jeden einzelnen Case berechnet, welche Durchschnittsgeschwindigkeit dieser hatte. Legt man Wert darauf, mit einer möglichst hohen durchschnittlichen Geschwindigkeit voranzukommen, ist diese KPI als Bewertungsmittel geeignet, da dadurch indiziert wird, dass weniger Wartezeiten bzw. Streckenabschnitte, an denen man langsamer fahren muss, auftreten. Es wird also nicht die Gesamtdauer der Cases bewertet, sondern die Geschwindigkeit. Man kann annehmen, dass je höher die Durchschnittsgeschwindigkeit ist, desto weniger Unterbrechungen durch z.B. Ortschaften, Ampeln oder andere Faktoren, die die Geschwindigkeit beeinflussen, treten auf. Es ist also möglich, dass ein Case mit einer höheren durchschnittlichen Geschwindigkeit auch eine längere Strecke mit sich bringt und man dadurch nicht zwingend früher an sein Ziel kommt.

Auffällig ist außerdem, dass die einzelnen Cases, wie sie hier betrachtet wurden, nicht aussagekräftig genug waren. Ein einzelner Case könnte nämlich immer einen Ausnahmefall

darstellen. Da allerdings die Entfernung für die Berechnung von Nöten war, welche für jeden Case unterschiedlich war, war es nicht anders möglich. Aus diesem Grund empfiehlt sich hier, die ausgewählten Cases den Varianten zuzuordnen, um Muster erkennen zu können.

5.1.2.3 Prozesskennzahl D

Die Berechnung der Prozesskennzahl D hatte zum Ziel, das Verhältnis zwischen der Dauer eines Cases und der Durchschnittsdauer aller Cases bewerten und so herausfinden zu können, welche Cases unter oder überdurchschnittlich lang gedauert haben.

Ähnlich wie bei der Prozesskennzahl V bedeutete ein Wert von genau $,1'$, dass die Dauer des Cases und die Durchschnittsdauer identisch waren. Ein Wert von $,<1'$ indizierte, dass der betrachtete Case schneller durchlaufen wurde als durchschnittlich. Folglich zeigte ein Wert von $,>1'$, dass der betrachtete Case langsamer als der Durchschnitt war. Je niedriger der errechnete Wert also ist, desto zeitlich effizienter war der jeweilige Case.

5.1.2.3.1 Bewertung der Cases

Die Durchschnittsdauer aller Cases des ersten Falles betrug 23,9 Minuten. Anhand der Berechnungen ließ sich erkennen, dass 36 der insgesamt 55 Cases unter diesem Wert lagen. Bei den Cases 18, 26, 8, 36 und 25 handelte es sich um die kürzesten Cases. Mit Blick auf das Tool kann man herausfinden, dass alle der genannten Cases der ersten Variante zuzuordnen sind. Dies deutet darauf hin, dass die erste Variante somit die zeitlich effizienteste war.

Beim zweiten Fall betrug die Durchschnittsdauer 30,1 Minuten. Insgesamt 24 Cases waren kürzer als der Durchschnitt. Am kürzesten waren die Cases 16, 13, 11, 41 und 12. Jeweils zwei der Cases gehörten zu den Varianten 2 und 5, und ein weiterer zur achten Variante. Die Varianten 2, 5 und 8 scheinen dementsprechend zeitlich effizient gewesen zu sein.

Die durchschnittliche Dauer aller Cases aus dem Event-Log des dritten Falles betrug 111,6 Minuten. Zwölf der 20 Cases waren kürzer. Die kürzesten Cases waren diejenigen mit den IDs 7, 16, 4, 9, 10, 12, 15, 17 und 18. Sieben dieser Cases gehörten zur ersten Variante. Die anderen zwei Cases gehörten dagegen zur dritten Variante. Es lässt sich also die Variante 1 in besonderem Maße empfehlen, wenn man möglichst kurz unterwegs sein möchte. Doch auch die dritte Variante kann empfohlen werden.

Beim vierten Fall betrug die durchschnittliche Dauer der Cases 61,2 Minuten. 29 von insgesamt 32 Cases waren kürzer als der Durchschnittswert. Am kürzesten waren die Cases 2, 14, 3 und 25. Zwei davon ließen sich im Tool der ersten Variante zuordnen. Jeweils ein

Case gehörte zu den Varianten 5 und 6, welche beide nur einmal aufgetreten sind. Auch hier ist jedoch bei genauerem Hinsehen aufgefallen, dass es sich bei den Varianten 5 und 6 um fehlerhafte Aufzeichnungen seitens Google Maps handelt. Aus diesem Grund lässt sich hier lediglich die erste Variante empfehlen, um möglichst schnell an das Ziel zu gelangen. Zudem ist in den Ergebnissen bereits aufgefallen, dass die Cases 1 und 28 um ein Vielfaches länger brauchten als die restlichen Cases. Dies ist der Grund dafür, dass die Durchschnittsdauer des vierten Falles so hoch ausgefallen ist. Die beiden Cases zogen die durchschnittliche Dauer nämlich um rund 8,6 Minuten in die Höhe (s. Anhang, Abbildung 125).

5.1.2.3.2 Schlussfolgerung KPI D

Mithilfe der Prozesskennzahl D konnte berechnet werden, inwieweit die Dauer der einzelnen Cases der Durchschnittsdauer entsprochen haben. Je näher das Ergebnis an dem Wert ,1‘ war, desto eher entsprach die Dauer des betrachteten Cases der durchschnittlichen Dauer aller Cases des jeweiligen Falles. An einem Wert von ,<1‘ ließ sich außerdem erkennen, welche Cases unter der Durchschnittsdauer lagen.

Die KPI D ist also ein geeignetes Mittel, um zu erfahren, welche Cases unter oder über dem Durchschnitt waren, sowie welche Cases der Durchschnittsdauer am ehesten entsprochen haben.

5.1.2.4 Anwendung von Filtern

Anhand der Anwendung von Filtern konnte man sich nur diejenigen Cases anzeigen lassen, welche einer bestimmten Fragestellung gedient haben. Als Konsequenz wurden im Tool die Informationen, wie beispielsweise das Prozessmodell oder die Varianten, aktualisiert. Im Folgenden findet sich eine kurze Zusammenfassung aller angewandten Filter sowie eine Einschätzung des Nutzens der Anwendung solcher Filter.

5.1.2.4.1 Filter: Kürzeste Cases

Der Filter ,Kürzeste Cases‘ wurde bei jedem betrachteten Fall angewandt, um die Frage „Welche waren die kürzesten Cases?“ zu beantworten. Um die gesuchten Cases herauszufiltern, wurde sich an der Mindest-Dauer der Cases orientiert und diese um einige Sekunden bzw. Minuten erhöht, sodass zwischen zwei und sechs Ergebnisse gefunden werden konnten.

Bei dem ersten Fall wurde nach maximal 21 Minuten und 30 Sekunden gefiltert. Sechs Cases, also zehn Prozent aller Cases, trafen auf dieses Kriterium zu. Die ersten fünf Cases gehörten alle zu der ersten Variante. Nur das sechste Ergebnis ließ sich der zweiten Variante zuordnen.

Der Filter für den zweiten Fall suchte nach den Cases, welche maximal 20 Minuten und 30 Sekunden gedauert haben. Es konnten vier Cases, entsprechend neun Prozent aller Cases, gefunden werden. Zwei Cases gehörten zu Variante 1, jeweils ein Case zu Variante 2 und Variante 3.

Nach Cases, welche maximal eine Stunde 33 Minuten gedauert haben, wurde bei dem dritten Fall gefiltert. Bei der Suche kamen vier Cases heraus, was 20% aller Cases des Event-Logs sind. Die ersten drei Cases waren Variante 1 zugehörig, der vierte Case der zweiten Variante.

Bei dem vierten Fall suchte der Filter nach Cases, welche eine maximale Dauer von 47 Minuten hatten. Sechs Prozent aller Fälle, sprich zwei Cases, trafen darauf zu. Beide Cases gehörten einer anderen Variante an.

5.1.2.4.2 Filter: Kürzer als Durchschnitt

Das Ziel des Filters ‚Kürzer als Durchschnitt‘ war es herauszufiltern, wie viele Cases unterdurchschnittlich lange gedauert haben. Dementsprechend wurde nach einer maximalen Dauer, welche identisch mit der Durchschnittsdauer des jeweiligen Falles ist, gefiltert.

Da die durchschnittliche Dauer der Cases des ersten Falles 23,9 Minuten betrug, wurde nach denjenigen Cases gesucht, welche maximal 23,9 Minuten gedauert haben. Dabei kamen 36 Cases, was einem Anteil von 65% entspricht, heraus.

Bei dem zweiten Fall wurde nach maximal 30,1 Minuten gefiltert. 24 Cases, also 58% aller Fälle des zweiten Event-Logs, trafen darauf zu.

Die maximale Dauer, nach der beim dritten Fall gesucht wurde, war 111,6 Minuten. 60% aller Cases, sprich zwölf Cases, konnten herausgefiltert werden.

Bei dem vierten Fall wurde nach einer Höchstdauer von 61,2 Minuten gefiltert. Als Ergebnis erhielt man 29 Cases. Dies entspricht 90% aller Cases, welche in dem Event-Log aufgetreten sind.

5.1.2.4.3 Filter: Cases mit Wartezeiten

Mithilfe dieses Filters ‚Cases mit Wartezeiten‘ wurde nach der Anzahl der Cases gesucht, bei denen Wartezeiten auftraten. Wartezeit bedeutet in diesem Fall allerdings nicht, dass für die Person, deren Mobilitätsdaten aufgezeichnet wurden, Wartezeiten aufgetreten sind. Vielmehr geht es hier um Momente, an denen Google Maps keine Daten aufgezeichnet hat, wodurch Lücken in den aufgezeichneten Mobilitätsdaten entstanden sein könnten. Um alle Cases, bei denen solche Lücken aufgetreten sind, zu finden, wurde bei allen vier Fällen nach einer durchschnittlichen Wartezeit von mindestens einer Millisekunde gefiltert.

Bei dem ersten Fall konnten 17 Cases mit Wartezeiten gefunden werden. Dies entspricht einem Anteil von 30% aller Cases des Event-Logs. Bei dem dritten Fall wurden vier Cases, also 20% aller Fälle, gefunden. Bei dem zweiten und vierten Fall konnten hingegen keine Cases mit Wartezeiten herausgefiltert werden.

5.1.2.4.4 Nutzen der Anwendung von Filtern

Bei der Anwendung der Filter hat sich gezeigt, dass diese zwar nur noch diejenigen Cases im Tool anzeigten, welche auf den Filter zutrafen, jedoch neben einem aktualisierten Prozessmodell und angepassten Statistiken keine neuen nennenswerten Ergebnisse lieferten. Bis auf den Anteil der Cases und der Ereignisse des gesamten Event-Logs, welche auf den Filter zutrafen, wurde außerdem kein weiterer Vergleich zum ursprünglichen Ereignisprotokoll gezogen.

Aufgefallen ist zudem, dass die Varianten nach Anwendung eines Filters neu zugeordnet wurden. Das heißt, dass ein Ergebnis, welches nach dem Filtern zu der ersten Variante gehörte, im ursprünglichen Event-Log nicht zwingend der ersten Variante zugeordnet wurde. Dennoch kann man gut erkennen, welche Cases zu derselben Variante gehören, da sich die Prozessabläufe durch das Filtern nicht verändert haben.

Die Ergebnisse des Filters ‚Cases mit Wartezeiten‘ deuten außerdem darauf hin, dass das Auftreten von Wartezeiten weniger ein Indikator für die Qualität eines Mobilitätsprozesses, sondern vielmehr abhängig von dem genutzten Gerät, mit welchem die Daten aufgezeichnet wurden, ist. Es wäre denkbar, dass qualitativ weniger hochwertige GPS-Sensoren zu Lücken in den Aufzeichnungen führen können. Auch bestimmte Orte, an denen es eventuell keinen GPS-Empfang gibt, wie z.B. unterirdische Parkhäuser, könnten ein Grund für das Auftreten von Wartezeiten sein.

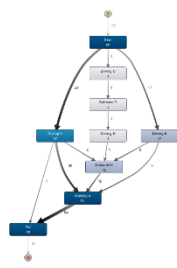
5.1.2.5 Hypothesen

Es wurde bereits festgestellt, dass die drei aufgestellten Hypothesen allesamt unzutreffend waren. Nachfolgend wird erläutert, wie diese Schlussfolgerungen zustande gekommen sind.

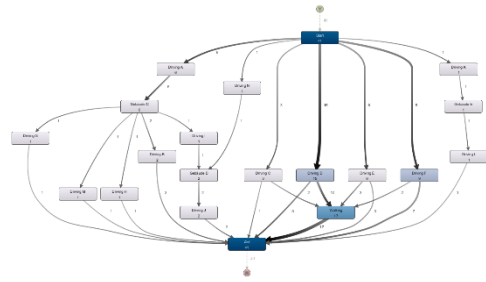
5.1.2.5.1 Hypothese 1

Die erste Hypothese lautete: „Je mehr Cases ein Event-Log enthält, desto umfangreicher ist das Prozessmodell“. Die Informationen, die man zur Überprüfung dieser Annahme benötigte, waren zum einen die Anzahl der Cases pro Event-Log sowie zum anderen das Prozessmodell. In Abbildung 47 wurden die Informationen für alle vier Fälle zusammengefasst.

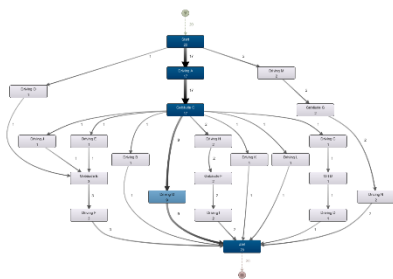
Prozessmodelle



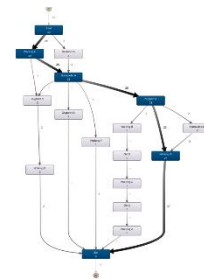
Fall 1: **55 Cases**



Fall 2: **41 Cases**



Fall 3: **20 Cases**



Fall 4: **32 Cases**

Abbildung 47: Prozessmodelle und Anzahl der Cases. Eigene Darstellung (2022)

Bereits beim ersten Blick auf die Modelle und die Case-Anzahl fällt auf, dass die Hypothese H1 nicht zutrifft. Der erste Fall hat zwar die meisten Cases, das Prozessmodell ist jedoch das kleinste. Der dritte Fall hat nur 20 Cases, das Prozessmodell ist hingegen eines der umfangreichsten. Auch der zweite Fall hat mit 41 Cases ein umfangreicheres Prozessmodell als der erste Fall. Gleiches gilt für Fall 4.

Dies ist damit zu erklären, dass sich der Umfang des Prozessmodells nicht an der Anzahl der Cases, sondern an der Anzahl der Aktivitäten orientiert. Da in einem Prozessmodell die Aktivitäten sowie die Pfade zwischen den Aktivitäten dargestellt werden, bestimmt die Anzahl ersterer den Umfang des Modells. Die Menge der Cases ist folglich völlig unerheblich für die Größe eines Prozessmodells. Um dies zu beweisen, wird ein Blick auf die Anzahl der Aktivitäten geworfen: der erste Fall besteht aus neun unterschiedlichen Aktivitäten, der zweite Fall aus 20 Aktivitäten, der dritte Fall aus 22 Aktivitäten und der vierte Fall aus 17 Aktivitäten (s. Kapitel 4.3.1, 4.4.1, 4.5.1, 4.6.1). Zählt man die Aktivitätsfelder in den Modellen nach, kommt man auf dieselben Werte. Die Annahme hätte also lauten müssen: „Je mehr unterschiedliche Aktivitäten in einem Event-Log vorkommen, desto umfangreicher ist das Prozessmodell“.

5.1.2.5.2 Hypothese 2

Hypothese H2 behauptete: „Je mehr Ereignisse eine Variante enthält, desto länger dauert diese durchschnittlich in Vergleich zu Varianten mit einer geringeren Anzahl an Ereignissen“. Hierfür wurden die Varianten der einzelnen Fälle betrachtet. Für die Prüfung der Annahme wurden die Anzahl der Ereignisse sowie die durchschnittliche Dauer der einzelnen Varianten benötigt. In Tabelle 18 wurden die Werte deshalb aufgelistet.

Tabelle 18: Anzahl der Ereignisse und Durchschnittsdauer der Varianten aller Fälle. Eigene Darstellung (2022)

Fall 1		
Variante 1	4 Ereignisse	22 mins, 40 secs
Variante 2	4 Ereignisse	24 mins, 54 secs
Variante 3	5 Ereignisse	24 mins, 22 secs
Variante 4	5 Ereignisse	24 mins, 50 secs
Variante 5	3 Ereignisse	22 mins, 1 sec
Variante 6	7 Ereignisse	45 mins, 40 secs
Fall 2		
Variante 1	4 Ereignisse	31 mins, 11 secs
Variante 2	3 Ereignisse	22 mins, 58 secs
Variante 3	3 Ereignisse	29 mins, 24 secs
Variante 4	4 Ereignisse	25 mins, 41 secs
Variante 5	3 Ereignisse	21 mins, 59 secs
Variante 6	5 Ereignisse	31 mins, 53 secs
Variante 7	4 Ereignisse	33 mins, 34 secs
Variante 8	4 Ereignisse	22 mins, 3 secs
Variante 9	3 Ereignisse	31 mins, 28 secs
Variante 10	5 Ereignisse	39 mins, 50 secs
Variante 11	5 Ereignisse	47 mins, 42 secs
Variante 12	7 Ereignisse	48 mins, 27 secs
Variante 13	5 Ereignisse	43 mins, 54 secs
Variante 14	5 Ereignisse	43 mins, 49 secs
Variante 15	5 Ereignisse	41 mins, 22 secs
Fall 3		
Variante 1	5 Ereignisse	1 hour, 36 mins
Variante 2	7 Ereignisse	2 hours, 11 mins
Variante 3	5 Ereignisse	1 hour, 33 mins

<i>Variante 4</i>	5 Ereignisse	2 hours, 54 mins
<i>Variante 5</i>	7 Ereignisse	2 hours, 53 mins
<i>Variante 6</i>	7 Ereignisse	1 hour, 53 mins
<i>Variante 7</i>	7 Ereignisse	1 hour, 50 mins
<i>Variante 8</i>	5 Ereignisse	1 hour, 52 mins
<i>Variante 9</i>	5 Ereignisse	2 hours, 6 mins
<i>Variante 10</i>	5 Ereignisse	1 hour, 41 mins
<i>Fall 4</i>		
<i>Variante 1</i>	6 Ereignisse	53 mins, 10 secs
<i>Variante 2</i>	6 Ereignisse	1 hour, 44 mins
<i>Variante 3</i>	7 Ereignisse	49 mins, 46 secs
<i>Variante 4</i>	10 Ereignisse	2 hours, 57 mins
<i>Variante 5</i>	5 Ereignisse	43 mins, 37 secs
<i>Variante 6</i>	6 Ereignisse	47 mins, 6 secs
<i>Variante 7</i>	5 Ereignisse	54 mins
<i>Variante 8</i>	5 Ereignisse	53 mins, 41 secs

Direkt bei Betrachtung des ersten Falles kann die Hypothese widerlegt werden. Denn die Varianten 3 und 4, welche beide aus jeweils fünf Ereignissen bestehen, sind im Durchschnitt kürzer als die zweite Variante, welche aus vier Ereignissen besteht. Auch beim zweiten Fall trat dies auf. Die vierte Variante (vier Ereignisse) ist nämlich um mehrere Minuten kürzer als die dritte Variante (drei Ereignisse). Auch die Variante 6 mit fünf Ereignissen ist kürzer als die siebte Variante mit vier Ereignissen. Abgesehen von diesen beiden Beispielen ist dieses Phänomen beim zweiten Fall noch häufiger aufgetreten. Weiterhin konnte es auch bei dem dritten und bei dem vierten Fall beobachtet werden.

Folglich gibt es also keine Korrelation zwischen der Anzahl der Ereignisse einer Variante und der Durchschnittsdauer derselbigen. Besonders bei Mobilitätsprozessen gibt es zahlreiche Faktoren, welche Einfluss auf die Dauer haben können. Diese sind unter anderem die Streckenlänge, die Straßenart, die Uhrzeit (z.B. Stoßzeiten), das Wetter, Verkehrsbehinderungen (Stau, Unfall etc.) und die gewählte(n) Fortbewegungsmethode(n). Abhängig davon kann es sich folglich durchaus lohnen, einen Mobilitätsprozess mit mehreren Schritten zu wählen.

5.1.2.5.3 Hypothese 3

Die dritte Behauptung lautete wie folgt: „Je mehr Varianten in einem Event-Log auftreten, desto höher ist der Anteil der Varianten, welche nur einmal aufgetreten sind“. Benötigt wurden für die Prüfung der Annahme die Anzahl der Varianten eines Event-Logs sowie die Anzahl derer Varianten, welche einmalig aufgetreten sind. In Tabelle 19 wurden die Anteile, welche berechnet werden konnten, indem die Anzahl der einmalig aufgetretenen Varianten durch die Gesamtanzahl der Varianten dividiert wurde, zusammengefasst.

Tabelle 19: Berechnung des Anteil der Varianten, die nur einmalig aufgetreten sind. Eigene Darstellung (2022)

Fall	Anzahl Varianten	1x aufgetretene Varianten	Anteil
<i>Fall 1</i>	6	2	33,33%
<i>Fall 2</i>	15	7	46,67%
<i>Fall 3</i>	10	7	70%
<i>Fall 4</i>	8	5	62,5%

Anschließend lassen sich die vier Fälle miteinander vergleichen. Der Hypothese zufolge müsste der zweite Fall den höchsten Anteil an einmalig aufgetretenen Varianten aufweisen, da dieser Fall die meisten Varianten hat. Jedoch weisen die Fälle 3 und 4 einen höheren Anteil auf. Außerdem fällt auf, dass die Behauptung für die Fälle 1, 3 und 4 zutreffen würde. Nur durch den zweiten Fall konnte die Annahme widerlegt werden.

Ein Erklärungsansatz dafür könnte sein, dass der Anteil bzw. die Anzahl der Varianten, die nur einmal aufgetreten sind, von der jeweiligen Person, deren Angewohnheiten, dem betrachteten Zeitraum sowie äußeren Einflüssen abhängig sind. Denn solche Varianten treten normalerweise auf, wenn sich auf dem Weg Ausnahmen ereignet haben. Solche Ausnahmen könnten z.B. sein, dass man einen Umweg fahren muss oder dass man auf dem Weg einen Zwischenstopp einlegt. Außerdem kann man annehmen, dass es Personen gibt, welche mehrere Mobilitätsprozesse routiniert und ohne Ausnahmen nutzen. Ebenso kann angenommen werden, dass es Personen gibt, welche noch keine Routine für ihren Arbeitsweg gefunden haben und unterschiedliche Varianten ausprobieren, wodurch der Anteil der einmaligen Varianten steigt. Nicht zuletzt ist auch der betrachtete Zeitraum von Bedeutung. Denn werden z.B. lediglich zwei Tage betrachtet, gibt es nur die beiden Möglichkeiten, dass der Anteil der einmalig aufgetretenen Varianten bei 0% oder bei 100% liegt, da die beiden Cases entweder zu der gleichen Variante oder zu zwei unterschiedlichen Varianten gehören können. Es kann folglich nicht pauschal behauptet werden, dass es eine

Verbindung zwischen der Gesamtanzahl aller Varianten und der Anzahl der einmalig aufgetretenen Varianten gibt.

5.1.2.6 Möglichkeiten zur Nutzung der Informationen aus dem Tool ‚Disco‘

Die Berechnung von Prozesskennzahlen, die Anwendung von Filtern sowie die Überprüfung von Hypothesen anhand von Werten aus dem Tool waren alles Beispiele dafür, wie man diese nutzen kann, um eine Bewertungsgrundlage für Mobilitätsprozesse zu erhalten. Die Vorstellung und Durchführung dieser Möglichkeiten diente der Beantwortung der zweiten Teilfrage: „Wie lassen sich die aus dem Tool gewonnenen Informationen nutzen, um den/die Mobilitätsprozess/e zu bewerten?“

Es hat sich gezeigt, dass es unterschiedliche Möglichkeiten zur manuellen Bewertung der Prozesse gibt, welche nicht durch das Tool selbst bereitgestellt werden. Dies liegt möglicherweise daran, dass Process Mining bisher kaum Anwendung auf Mobilitätsprozesse gefunden hat (s. Kapitel 2.3.1) und es deshalb, zumindest im hier verwendeten Tool, keine auf Mobilitätsprozessdaten angepassten Bewertungsalgorithmen gibt.

5.1.2.7 Weitere Erkenntnisse

Vergleicht man, welchen Zwecken die Prozesskennzahl D sowie der Filter ‚schneller als Durchschnitt‘ dienen, fällt auf, dass diese dasselbe Interesse verfolgen. Möchte man also durch die Berechnung der Kennzahl D herausfinden, welche Cases unter der Durchschnittszeit waren, kann man dasselbe Ergebnis durch die Anwendung eines Filters (hier: Kürzer als Durchschnitt) mit deutlich geringerem Aufwand erzielen. Doch auch für das Herausfiltern von Cases, welche über dem Durchschnitt liegen, wäre vermutlich die Anwendung eines Filters im Process Mining-Tool eine geeignetere Variante.

Die Prüfung der Hypothesen hat überdies zu der Erkenntnis geführt, dass man nicht pauschal von Korrelationen zwischen einzelnen Aspekten ausgehen kann. Besonders Mobilitätsprozesse sind so individuell und unvorhersehbar, dass sich die Verhältnisse zwischen Werten und Aspekten immer unterschiedlich verhalten können.

5.2 Beschränkungen der Forschung

Es ist nicht das übergeordnete Ziel dieser Arbeit gewesen, die untersuchten Fälle zu bewerten. Vielmehr waren die Bewertungen der Mobilitätsprozesse anhand von Prozesskennzahlen, Filtern und Hypothesen Teil der Beantwortung der Teil- und Forschungsfragen. Aus diesem Grund sind die Ergebnisse auf die Informationen beschränkt, welche für die Berechnung der Prozesskennzahlen, die Anwendung von Filtern und die Prüfung der Hypothesen benötigt wurden.

Außerdem konnten die Prozesse lediglich anhand ihrer Häufigkeiten und ihrer zeitlichen Performance bewertet werden, da die Aufzeichnungen der Mobilitätsprozessdaten aus Google Maps keine weiteren messbaren Informationen über die Mobilitätsprozesse preisgaben.

Es muss zudem berücksichtigt werden, dass in dieser Arbeit ausschließlich Mobilitätsdaten aus der Google Maps Timeline genutzt worden sind. Diese wurden in dem Process Mining-Tool *Disco* verarbeitet, welches lediglich Process Discovery-Techniken nutzte. Außerdem boten die Daten keine Auskunft darüber, wie ein optimaler Prozess aussehen sollte. Wie bereits in Kapitel 2.1.4 erläutert wurde, benötigen die Process Mining-Ausprägungen Conformance Checking und Model Enhancement neben einem Event-Log außerdem ein Soll-Prozessmodell. Durch das Fehlen eines solchen Modells sowie fehlender Möglichkeiten des Tools war es nicht möglich, andere Process Mining-Techniken als Process Discovery auf die Event-Logs anzuwenden. Auf anderem Wege erhobene Mobilitätsdaten, das Vorhandensein von Soll-Prozessmodellen sowie die Nutzung anderer Process Mining-Tools hätten möglicherweise zu abweichenden Ergebnissen geführt.

In Bezug auf die Qualität der aufgezeichneten Mobilitätsdaten hat sich ebenfalls eine Beschränkung ergeben. Beim vierten Fall z.B. sind Unterschiede von mehreren Stunden bei der Dauer aufgetreten. Mit Blick auf die aufgezeichneten Routen ließ sich feststellen, dass der Arbeitsort aus mehreren Gebäuden in einem größeren Gelände bestand. Es besteht folglich die Wahrscheinlichkeit, dass Google Maps erst später den „richtigen“ Arbeitsort erkannt hat, wodurch die Dauer der Prozesse stark variierte. Außerdem wurde bereits erwähnt, dass auch die genutzten Fortbewegungsmittel nicht immer korrekt erkannt wurden. Die Erkenntnis von Yousfi und Weske (2019), wonach die Orte und Fortbewegungsmittel durch die Anwendung von Process Mining-Techniken mit hoher Genauigkeit erkannt wurden, konnte in dieser Arbeit daher nicht bestätigt werden.

5.3 Empfehlung für weiterführende Forschung

Diese Bachelorarbeit hat erste Erkenntnisse zu der Anwendung des Process Mining Ansatzes auf den alltäglichen Mobilitätsprozess *Weg zur Arbeitsstelle* geliefert. Es empfiehlt sich, die Forschung weiterzuführen, indem umfangreichere Mobilitätsdatensätze (z.B. mit Wochentagen als Attribut) in einem umfassenderen Tool, wie beispielsweise *ProM*, ausgewertet werden. An dieser Stelle könnte erforscht werden, wie die Informationen, welche aus dem genutzten Tool zu entnehmen sind, für die Bewertung der Mobilitätsprozesse genutzt werden können und wie sich diese von den Informationen aus

dem Tool ‚Disco‘ unterscheiden. Außerdem wäre eine Befragung der Personen, deren Daten genutzt werden, von Vorteil, um Hintergründe über die Mobilitätsprozesse zu erfahren und zu erfragen, wie ein optimaler Mobilitätsprozess aussehen sollte. Hier sollte allerdings darauf geachtet werden, dass die Bewertung eines optimalen Mobilitätsprozesses von unterschiedlichen Faktoren, wie z.B. der Schnelligkeit, der Energieeffizienz, der Umweltbelastung oder der Streckenlänge abhängig sein kann.

Bei der Diskussion der ersten Hypothese ergab sich zudem eine neue Hypothese: „Je mehr unterschiedliche Aktivitäten in einem Event-Log vorkommen, desto umfangreicher ist das Prozessmodell“. In einem weiterführenden Forschungsprojekt könnte man daran anknüpfen und diese Annahme überprüfen.

Ein weiterer Ansatzpunkt für die Erforschung der Auswirkungen der Anwendung des Process Mining auf Mobilitätsprozesse wäre z.B. die Untersuchung von Bus- oder Zugrouten anstelle von personenbezogenen Mobilitätsprozessen.

6. Fazit

Die vorliegende Arbeit ging der Frage nach ,Welche Möglichkeiten bietet die Anwendung eines Process Mining-Tools, wie z.B. ,Disco‘, auf die alltäglichen Mobilitätsprozesse von Personen?“. Vorbereitend darauf wurden zunächst die Teilfragen ,Welche Erkenntnisse lassen sich über die alltäglichen Mobilitätsdaten von Personen mithilfe der in dieser Arbeit genutzten Process Mining-Software gewinnen und direkt aus dieser ablesen?“ und ,Wie lassen sich die aus dem Tool gewonnenen Informationen nutzen, um den/die Mobilitätsprozess/e zu bewerten?“ beantwortet. Für die Beantwortung der Fragen wurde mit einem Mixed-Methods-Ansatz gearbeitet.

Zunächst wurden reale Mobilitätsprozessdaten im Rahmen einer qualitativen Fallstudie erhoben, welche den Fall ,*Die Mobilitätsprozesse einer Person von oder zu ihrer Arbeitsstätte bzw. von oder zum Arbeitsort*“ untersuchte. Insgesamt wurden die Mobilitätsprozesse von vier berufstätigen Personen analysiert. Gewonnen wurden die Mobilitätsdaten aus der Google Maps Timeline der Teilnehmer.

Anschließend folgte eine qualitative Inhaltsanalyse der Möglichkeiten des kommerziellen Process Mining-Tools ,Disco‘, welches für die Auswertung der Daten genutzt wurde, sowie der bereitgestellten Informationen und Werte über die untersuchten Mobilitätsprozesse. Dies diente der Beantwortung der ersten Teilfrage.

Weiterhin wurde aufgezeigt, wie die Mobilitätsprozesse anhand der Informationen aus dem Tool bewertet werden können, um die zweite Teilfrage zu beantworten. An dieser Stelle wurden quantitative Methoden, wie die Berechnung von Prozesskennzahlen, angewandt. Außerdem wurden die Werte aus dem Tool herangezogen, um Hypothesen über die Mobilitätsprozesse zu prüfen und Filter innerhalb von ,Disco‘ anzuwenden.

6.1 Erkenntnisse der Forschung

Aus den Ergebnissen lässt sich schließen, dass mithilfe von Process Mining-Techniken ein detaillierter Einblick auf Mobilitätsprozesse möglich ist. Das Process Mining-Tool ,Disco‘ entdeckte zuverlässig die im Ereignisprotokoll aufgetretenen Mobilitätsprozesse und erstellte ein umfassendes Prozessmodell, welches durch farbliche Markierungen bereits einen intuitiven Überblick bietet. Zudem wurden die verfügbaren Informationen und die vom Tool berechneten Werte in Form von Prozessmodellen, Diagrammen und Tabellen übersichtlich zusammengefasst und visualisiert.

Bei der Analyse des Tools ist aufgefallen, dass die in den Tabellen und Statistiken dargestellten Informationen lediglich die Aspekte Häufigkeit und Zeit behandelten. Dies ist allerdings der Tatsache geschuldet, dass die in den Ereignisprotokollen enthaltenen Informationen nicht mehr preisgaben.

Weiterhin wurden zu Beginn dieser Arbeit zwei Erwartungen geäußert. Es wurde erwartet, dass das genutzte Process Mining-Tool alle aufgetretenen Mobilitätsprozesse erkennt und diese analysiert. Es hat sich gezeigt, dass das Tool alle Mobilitätsprozesse (Varianten) erkannt hat, für jeden dieser Prozesse ein passendes Prozessmodell erstellt sowie wichtige Informationen über z.B. die Anzahl der Cases und der Ereignisse pro Variante und die jeweiligen Durchschnittsdauern zusammengefasst hat. Somit wurde diese Erwartung erfüllt.

Außerdem wurde erwartet, dass sich anhand der Informationen, welche das Tool über die Ereignisprotokolle preisgibt, Entscheidungen über beispielsweise die zeitliche Effizienz der Mobilitätsprozesse treffen lassen. Das Tool stellte die Durchschnittsdauern der Varianten und auch der einzelnen Cases in Tabellen gegenüber, sodass man die Möglichkeit hatte, diese direkt miteinander zu vergleichen und zu erkennen, welche Prozesse oder Cases die beste zeitliche Effizienz hatten. Folglich konnte auch diese Erwartung erfüllt werden.

Des Weiteren konnte aus den Ergebnissen geschlossen werden, dass durch das Tool selbst keine Bewertung der Mobilitätsprozesse stattgefunden hat. Es handelte sich lediglich um eine objektive Zusammenfassung von Informationen und berechneten Werten. Dies liegt unter anderem daran, dass *Disco* keine Kenntnis über einen „optimalen“ Mobilitätsprozess verfügte. Darüber hinaus könnte eine mögliche Erklärung für die fehlende Bewertung sein, dass der Process Mining Ansatz bisher kaum Anwendung auf Mobilitätsprozesse gefunden hat, weshalb es noch keine angepassten Bewertungsalgorithmen für Mobilitätsprozesse gibt. Dennoch können auf Grundlage der bereitgestellten, umfangreichen Informationen Bewertungen manuell getroffen werden.

Aus diesem Grund wurden nachfolgend drei unterschiedliche Möglichkeiten, wie man die Daten aus dem Tool weiterverarbeiten kann, demonstriert. Begonnen wurde dabei mit der Berechnung von Prozesskennzahlen, welche im theoretischen Teil dieser Arbeit definiert worden sind.

Es hat sich gezeigt, dass die Ergebnisse der Berechnungen stark von der Korrektheit der Aufzeichnungen abhängig sind, da sonst verfälschte und irreführende Ergebnisse erreicht werden. Ein möglicher Grund für fehlerhafte Aufzeichnungen der Mobilitätsprozesse könnte

sein, dass die Aufzeichnungsfrequenz zu niedrig ist. Dies könnte unter anderem zur Folge haben, dass einzelne Ereignisse sowie die genutzten Verkehrsmittel nicht erkannt werden.

Außerdem hat die Berechnung der Prozesskennzahlen zu der Erkenntnis geführt, dass die Betrachtung einzelner Cases nicht aussagekräftig genug ist. Besser wäre es daher, die Cases sowie deren errechneten Werte der KPIs den jeweiligen Varianten zuzuordnen, um die einzelnen Mobilitätsprozesse bewerten zu können.

Darüber hinaus hat sich gezeigt, dass von den Ergebnissen der Prozesskennzahlen nicht automatisch auf z.B. eine kürzere Gesamtdauer geschlossen werden kann. Bei der KPI V zum Beispiel, welche das Verhältnis der Durchschnittsdauer der einzelnen Varianten zu der von Google Maps berechneten Dauer für die jeweilige Route bewertete, deutete ein niedriger Wert für V nicht automatisch darauf hin, dass die Durchschnittsdauer dieser Variante auch kürzer ist als die einer Variante, welche einen höheren Wert für V hatte. Auch bei der KPI G bedeutete eine höhere Durchschnittsgeschwindigkeit nicht gleich, dass man im Vergleich zu einem Case mit niedrigerer Durchschnittsgeschwindigkeit früher an das Ziel kam.

Weiterhin wurden Filter im Tool ‚Disco‘ angewandt. Die Ergebnisse der Anwendung der Filter haben gezeigt, dass kaum neue, sondern hauptsächlich nur ausgewählte Informationen angezeigt wurden. Lediglich das Prozessmodell wurde auf die herausgefilterten Mobilitätsprozesse angepasst sowie der Anteil der Cases und der Ereignisse in Prozent im Vergleich zum ursprünglichen Event-Log dargestellt.

Zudem konnte festgestellt werden, dass die Prozesskennzahl D sowie der Filter ‚Kürzer als Durchschnitt‘ dieselben Ergebnisse lieferten. In diesem Fall lohnt sich die Anwendung des Filters mehr, da dies weniger aufwändig und die Ergebnisse aussagekräftiger sind.

Anschließend wurden die Hypothesen, welche im theoretischen Teil dieser Bachelorarbeit aufgestellt worden sind, überprüft. Die Ergebnisse ließen darauf schließen, dass keine der drei Behauptungen zutreffend war. Dies führte zu der Erkenntnis, dass bei Mobilitätsprozessen aufgrund ihrer hohen Individualität nicht pauschal von Korrelationen ausgegangen werden kann.

Eine weitere Erkenntnis dieser Arbeit ist, dass sich die Anwendung des Process Mining eignet, um bekannte Strecken und Mobilitätsprozesse zu bewerten und die für einen persönlich effizienteste Variante zu ermitteln. Dies liegt daran, dass die hier genutzten Mobilitätsprozessdaten historischer Natur waren. Gegensätzlich dazu steht Google Maps, welches eine prädiktive Routenplanung auf Grundlage von Echtzeitdaten ermöglicht.

6.2 Beantwortung der Forschungsfrage

Durch die hier durchgeführte Mixed-Methods-Forschung wurde somit gezeigt, dass sich die Anwendung von Process Mining-Techniken auf die alltäglichen Mobilitätsprozesse mithilfe eines Tools eignet, wenn man seine persönlichen Mobilitätsprozesse aufschlüsseln, bewerten, und den für sich persönlich effizientesten Mobilitätsprozess ermitteln möchte.

Um dies zu erreichen, bietet das Tool *Disco* die Möglichkeit, ein umfassendes und intuitives Prozessmodell aller im Event-Log aufgetretenen Mobilitätsprozesse zu betrachten und mit diesem zu interagieren, indem mittels eines Mausklicks auf die jeweiligen Aktivitäten und Pfade zahlreiche Informationen über die Häufigkeiten und die Performance dieser angezeigt werden.

Weiterhin ermöglicht es das Tool, zahlreiche Informationen über die einzelnen Cases, Varianten und Aktivitäten, wie z.B. die (Durchschnitts)Dauer, die Anzahl der Ereignisse, die absolute und die relative Häufigkeit, den zeitlichen Spielraum, den Anteil der Cases, welche zu einer Variante gehören, und viele mehr, zu sehen. All diese Informationen werden übersichtlich in Form von Modellen, Diagrammen und Tabellen zusammengefasst.

Dies macht es wiederum möglich, Entscheidungen auf Grundlage der im Tool dargestellten Informationen zu treffen oder diese weiterzuverarbeiten, indem beispielsweise Prozesskennzahlen berechnet oder Filter angewandt werden, um weitere Bewertungsgrundlagen für die Mobilitätsprozesse zu erhalten.

Neben den visualisierenden und zusammenfassenden Funktionen, die das Tool enthält, können die gewonnenen Werte also auch genutzt werden, um gezielte, individuelle Fragestellungen und Interessen zu beantworten.

6.3 Kritische Würdigung und Lösungsansätze

Aufgrund fehlender Möglichkeiten des genutzten Process Mining-Tools sowie nicht vorhandener Soll-Prozessmodelle konnte in dieser Arbeit lediglich die Wirkung der Process Mining-Ausprägung *Discovery* auf Mobilitätsprozessdaten untersucht werden. Für zukünftige Forschungsprojekte empfiehlt sich daher, ein Tool zu wählen, welches neben Process Discovery auch Conformance Checking und/oder Model Enhancement unterstützt. Um ein Soll-Prozessmodell zu erhalten, sollte außerdem ein Interview mit den Teilnehmern stattfinden, in dem erfragt wird, wie der effizienteste Mobilitätsprozess für den betrachteten Weg aussehen sollte. Hierbei sollte unbedingt darauf geachtet werden, zu definieren, von welcher Art von Effizienz gesprochen wird. Zudem bietet sich ein Interview an, um Hintergründe zu den tatsächlichen Mobilitätsprozessen zu erfahren.

Weiterhin hat die Auswahl der Mobilitätsprozessdaten aus Google Maps gezeigt, dass die Qualität der Aufzeichnungen in einigen Fällen zu wünschen übrigließ, da Lücken aufgetreten sind oder die Fortbewegungsart nicht erkannt wurde. An dieser Stelle hätten die Aufzeichnungen genauer überprüft werden sollen, um nur die Dateien zu exportieren, welche fehlerfrei waren.

Dennoch konnten trotz der aufgetretenen Schwierigkeiten, besonders in Anbetracht des kaum vorhandenen Forschungsstandes, einige grundlegende Erkenntnisse über die Anwendung des Process Mining Ansatzes auf Mobilitätsprozessdaten gesammelt werden.

6.4 Ausblick

Diese Forschungsarbeit lieferte grundlegende Erkenntnisse über die Anwendung von Process Mining auf Mobilitätsprozesse. Aufbauend darauf bieten sich einige Forschungsansätze für die Zukunft an.

Im Allgemeinen empfiehlt es sich, informationsreichere Mobilitätsdaten zu erheben und diese mithilfe eines wissenschaftlichen Process Mining-Tools, wie *ProM*, zu analysieren, da diese umfassendere Möglichkeiten und Algorithmen anbieten.

Außerdem steht die Anwendung der Process Mining-Ausprägungen Conformance Checking und Model Enhancement noch aus. In Zukunft sollten diese auf Mobilitätsprozessdaten angewendet werden, um die Grundlagenforschung zu dieser Thematik voranzutreiben.

Zudem bietet es sich für weiterführende Forschungen an, nicht nur personenbezogene Mobilitätsprozessdaten, sondern auch Mobilitätsprozesse von beispielsweise öffentlichen Verkehrsmitteln auszuwerten. Dies könnte es Mobilitätsdienstleistern ermöglichen, ihre Prozesse zu optimieren.

Bei weiterer intensiver Forschung sowie idealerweise der Entwicklung angepasster Process Mining-Algorithmen hat Process Mining durchaus das Potenzial, sich in Zukunft zu einer gefragten Technik für die Auswertung und Bewertung von Mobilitätsprozessen zu entwickeln.

Literaturverzeichnis

- Abdulrazaq, A. P. D. M., Zuhriyah, H., Al-Zubaidi, S., Karim, S., Ramli, R., & Yusuf, E. (2020). NOVEL COVID-19 DETECTION AND DIAGNOSIS SYSTEM USING IOT BASED SMART HELMET. *International Journal of Psychosocial Rehabilitation*, 24, 2296-2303. doi:10.37200/IJPR/V24I7/PR270221
- Ahrens, G.-A., Wittwer, R., Hubrich, S., Wittig, S., & Ließke, F. (2015). Sonderauswertung zum Forschungsprojekt „Mobilität in Städten–SrV 2013 “Städtevergleich. *Technische Universität Dresden, Verkehrs-und Infrastrukturplanung, Dresden*.
- Baji, A. M., Salini, M., Sandra, S., Thomas, S. R., & Ravikumar, A. (2017, 20-22 July 2017). *Track&Go: A location prediction web application*. Paper presented at the 2017 International Conference on Networks & Advances in Computational Technologies (NetACT).
- Bergsmann, S. (2011). Prozessmanagement als Managementansatz. In S. Bergsmann (Ed.), *End-to-End-Geschäftsprozessmanagement: Organisationselement — Integrationsinstrument — Managementansatz* (pp. 197-254). Vienna: Springer Vienna.
- Blank, P., Maurer, M., Siebenhofer, M., Rogge-Solti, A., & Schonig, S. (2016, 5-9 Sept. 2016). *Location-Aware Path Alignment in Process Mining*. Paper presented at the 2016 IEEE 20th International Enterprise Distributed Object Computing Workshop (EDOCW).
- Borrmann, A., König, M., Koch, C., & Beetz, J. (2021). *Building Information Modeling - Technologische Grundlagen und industrielle Praxis* (Vol. 2): Springer Vieweg Wiesbaden.
- Chon, Y., Talipov, E., Shin, H., & Cha, H. (2014). SmartDC: Mobility Prediction-Based Adaptive Duty Cycling for Everyday Location Monitoring. *IEEE Transactions on Mobile Computing*, 13(3), 512-525. doi:10.1109/TMC.2013.14
- Cotroneo, G., Carbone, R., Boggini, S., & Cerini, M. (2021). Process Mining: A Database of Applications. from HSPI SpA - Management Consulting
- Dakic, D., Stefanovic, D., Cosic, I., Lolic, T., & Medojevic, M. (2018). Business Process Mining Application: A Literature Review. In *Proceedings of the 29th International DAAAM Symposium 2018* (pp. 0866-0875): DAAAM International.
- Gou, Z., Tsugawa, S., Korkaew, A., Yamaguchi, S., Khamket, T., Manaskasemsak, B., & Rungsawang, A. (2016, 11-14 Oct. 2016). *An interest-based tour planning tool by process mining from Twitter*. Paper presented at the 2016 IEEE 5th Global Conference on Consumer Electronics.
- Goudarzi, F. (2018, 28-30 June 2018). *Travel Time Prediction: Comparison of Machine Learning Algorithms in a Case Study*. Paper presented at the 2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS).
- He, Z., Chow, C. Y., & Zhang, J. D. (2019, 10-12 Aug. 2019). *A Comparative Analysis of Journey Time from Google Maps and Intelligent Transport System in Hong Kong*. Paper presented at the 2019 IEEE 21st International Conference on High Performance Computing and Communications; IEEE 17th International Conference on Smart City; IEEE 5th International Conference on Data Science and Systems (HPCC/SmartCity/DSS).
- Janssenswillen, G., Depaire, B., & Verboven, S. (2018). Detecting train reroutings with process mining: A Belgian application. *EURO Journal on Transportation and Logistics*, 7(1), 1-24. doi:<https://doi.org/10.1007/s13676-017-0105-8>

- Jarass, J. (2012). Zum Begriff der Mobilität. In J. Jarass (Ed.), *Wohnstandortpräferenzen und Mobilitätsverhalten: Verkehrsmittelwahl im Raum Köln* (pp. 19-23). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kahl, T., & Zimmer, F. (2017). Kontextspezifische Visualisierung von Prozesskennzahlen. In T. Barton, C. Müller, & C. Seel (Eds.), *Geschäftsprozesse: Von der Modellierung zur Implementierung* (pp. 75-97). Wiesbaden: Springer Fachmedien Wiesbaden.
- Kley, S. (2016). Regionale Mobilität in der Bevölkerungssoziologie. In Y. Niephaus, M. Kreyenfeld, & R. Sackmann (Eds.), *Handbuch Bevölkerungssoziologie* (pp. 481-500). Wiesbaden: Springer Fachmedien Wiesbaden.
- Krstev, S., & Krneta, D. (2022, 16-18 March 2022). *Data Warehouse to Support Creation of Key Performance Indicators (KPIs) in Electricity Supply*. Paper presented at the 2022 21st International Symposium INFOTEH-JAHORINA (INFOTEH).
- Laakso, S. (2017). Giving up cars – The impact of a mobility experiment on carbon emissions and everyday routines. *Journal of Cleaner Production*, 169, 135-142. doi:<https://doi.org/10.1016/j.jclepro.2017.03.035>
- Laue, R., Koschmider, A., & Fahland, D. (2021). *Prozessmanagement und Process-Mining*. Berlin/Boston: De Gruyter Oldenbourg.
- Manderscheid, K. (2021). Nachhaltige Mobilität: Eine soziologische Dimensionalisierung. *Soziologie der Nachhaltigkeit*. Bielefeld: transcript-Verlag, 417-434.
- Mertens, K. (2019). Herausforderungen beim Einsatz von Datenanalytik für eine ressourceneffiziente Produktion. *BHM Berg- und Hüttenmännische Monatshefte*, 164(1), 26-30. doi:10.1007/s00501-018-0803-z
- Nanz, G. (2012). zfo-Toolkit-Prozess-Alignment-Von der Unternehmensstrategie zum Geschäftsprozess. *Zeitschrift Führung und Organisation*, 81(3), 210.
- Nobis, C., & Kuhnimhof, T. (2018). *Mobilität in Deutschland - MiD: Ergebnisbericht*. Retrieved from Berlin-Adlershof:
- Peters, R., & Nauroth, M. (2019). *Process-Mining - Geschäftsprozesse: smart, schnell und einfach* (1 ed.). Wiesbaden: Springer Gabler.
- Pfahler, F., Rubey, T., & Kördel, H. (2018). Mobilitätslösungen für körperlich eingeschränkte Personen in einem gesamtheitlichen Kontext am Beispiel der Deutsche Bahn AG. In H. Wagner & S. Kabel (Eds.), *Mobilität 4.0 – neue Geschäftsmodelle für Produkt- und Dienstleistungsinnovationen* (pp. 31-71). Wiesbaden: Springer Fachmedien Wiesbaden.
- Ribeiro, J., Fontes, T., Soares, C., & Borges, J. L. (2020). Process discovery on geolocation data. *Transportation Research Procedia*, 47, 139-146. doi:<https://doi.org/10.1016/j.trpro.2020.03.086>
- Rosemann, M., Schwegmann, A., & Delfmann, P. (2012). Vorbereitung der Prozessmodellierung. In *Prozessmanagement* (pp. 47-111).
- Rudnitckaia, J. (2016). *Process Mining. Data science In action*. University of Technology, Faculty of Information Technology, Brno.
- Saltz, J. S. (2021, 15-18 Dec. 2021). *CRISP-DM for Data Science: Strengths, Weaknesses and Potential Next Steps*. Paper presented at the 2021 IEEE International Conference on Big Data (Big Data).
- Sardianos, C., Varlamis, I., & Bouras, G. (2018, 28-31 Aug. 2018). *Extracting User Habits from Google Maps History Logs*. Paper presented at the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM).
- Shi, L., Wang, W., Cai, W., Wang, Z., Zhang, S., & Zhou, W. (2017, 11-13 Oct. 2017). *Mobility patterns analysis of Beijing residents based on call detail records*. Paper presented at the 2017 9th International Conference on Wireless Communications and Signal Processing (WCSP).

- Sonkin, R., Alpert, E. A., & Jaffe, E. (2020). Epidemic investigations within an arm's reach – role of google maps during an epidemic outbreak. *Health and Technology*, 10(6), 1397-1402. doi:10.1007/s12553-020-00463-0
- Spode, H. (2017). Mobilität, Reisen, Tourismus. Transformationen der Terminologie zwischen Fremdenverkehrslehre und Mobility Turn. In H. Pechlaner & M. Volgger (Eds.), *Die Gesellschaft auf Reisen – Eine Reise in die Gesellschaft* (pp. 23-46). Wiesbaden: Springer Fachmedien Wiesbaden.
- Tavares, J., Ribeiro, J., & Fontes, T. (2022). Detection of vehicle-based operations from geolocation data. *Transportation Research Procedia*, 62, 341-349. doi:<https://doi.org/10.1016/j.trpro.2022.02.043>
- Ullah, I., Boreli, R., Kanhere, S. S., Chawla, S., Ahanger, T. A., & Tariq, U. (2020). Protecting Private Attributes in App Based Mobile User Profiling. *IEEE Access*, 8, 143818-143836. doi:10.1109/ACCESS.2020.3014424
- van der Aalst, W. (2012). Process Mining: Overview and Opportunities. *ACM Trans. Manage. Inf. Syst.*, 3(2), Article 7. doi:10.1145/2229156.2229157
- van der Aalst, W., Adriansyah, A., de Medeiros, A. K. A., Arcieri, F., Baier, T., Blickle, T., . . . Wynn, M. (2012). *Process Mining Manifesto*. Paper presented at the Business Process Management Workshops, Berlin, Heidelberg.
- van der Aalst, W. M. P. (2022). Process Mining: A 360 Degree Overview. In W. M. P. van der Aalst & J. Carmona (Eds.), *Process Mining Handbook* (pp. 3-34). Cham: Springer International Publishing.
- Wuttke, L. (2020). CRISP-DM: Grundlagen, Ziele und die 6 Phasen des Data Mining Prozess. Retrieved from <https://datasolut.com/crisp-dm-standard/>
- Yamaguchi, S., Terada, T., Manaskasemsak, B., Rungsawang, A., & Leelaprute, P. (2017, 12-14 June 2017). *Tour Miner: Mining system of tour plans from SNS: — Smelting function from travel records to tour routes —*. Paper presented at the 2017 IEEE International Conference on Consumer Electronics - Taiwan (ICCE-TW).
- Yousfi, A., & Weske, M. (2019). Discovering commute patterns via process mining. *Knowledge and Information Systems*, 60(2), 691-713. doi:<https://doi.org/10.1007/s10115-018-1255-1>

Anhang

Tabelle 20: Wurde Process Mining bereits auf Mobilitätsprozesse angewendet? - Übersicht Literaturrecherche. Eigene Darstellung (2022)

Suchanfrage	Filter	Datenbank	Ergebnisse gesamt	Davon relevant	Titel relevanter Ergebnisse	Notizen
Process Mining AND Mobilitätsprozess		ProQuest	/	/		
		Oxford Academic	/	/		
		IEEE Xplore	/	/		
		ScienceDirect	/	/		
Process-Mining AND Mobilitätsprozess		ProQuest	/	/		
		Oxford Academic	/	/		
		IEEE Xplore	/	/		
		ScienceDirect	/	/		
„Process Mining“ AND Mobilitätsprozess		ProQuest	/	/		
		Oxford Academic	/	/		
		IEEE Xplore	/	/		
		ScienceDirect	/	/		
Process Mining AND Mobilitätsprozess OR Alltagsmobilität		ProQuest	33	0		
		Oxford Academic	/	/		
		IEEE Xplore	/	/		
		ScienceDirect	20	0		
Process Mining AND Alltagsmobilität		ProQuest	1	0		
		Oxford Academic	/	/		
		IEEE Xplore	/	/		
		ScienceDirect	/	/		
Process-Mining AND Alltagsmobilität		ProQuest	/	/		
		Oxford Academic	/	/		
		IEEE Xplore	/	/		
		ScienceDirect	/	/		
„Process Mining“ AND Alltagsmobilität		ProQuest	/	/		
		Oxford Academic	/	/		
		IEEE Xplore	/	/		
		ScienceDirect	/	/		
Process Mining AND everyday mobility		ProQuest	8506	- s. Notizen		Anfrage zu ungenau – Filter anwen- den
		Oxford Academic	3330	- s. Notizen		Anfrage zu ungenau
		IEEE Xplore	3	0		
		ScienceDirect	3127	- s. Notizen		Anfrage zu ungenau
Process-Mining AND everyday mobility		ProQuest	31	1	„Discovering commute patterns via	

					Process Mining“	
		Oxford Academic	3330	- s. Notizen		Anfrage zu ungenau
		IEEE Xplore	/	/		
		ScienceDirect	3127	- s. Notizen		Anfrage zu ungenau
„Process Mining“ AND everyday mobility		ProQuest	31	1	„Discovering commute patterns via Process Mining“	
		Oxford Academic	5	0		
		IEEE Xplore	/	/		
		ScienceDirect	13	0		
Process Mining AND everyday-mobility		ProQuest	35	0		
		Oxford Academic	3330	- s. Notizen		Anfrage zu ungenau
		IEEE Xplore	/	/		
		ScienceDirect	3129	- s. Notizen		Anfrage zu ungenau
Process-Mining AND everyday-mobility		ProQuest	/	/		
		Oxford Academic	3330	- s. Notizen		Anfrage zu ungenau
		IEEE Xplore	/	/		
		ScienceDirect	3129	- s. Notizen		Anfrage zu ungenau
„Process Mining“ AND everyday-mobility		ProQuest	/	/		
		Oxford Academic	5	0		
		IEEE Xplore	/	/		
		ScienceDirect	13	0		
Process Mining AND „everyday mobility“		ProQuest	35	0		
		Oxford Academic	10	0		
		IEEE Xplore	/	/		
		ScienceDirect	27	0		
Process-Mining AND „everyday mobility“		ProQuest	/	/		
		Oxford Academic	10	0		
		IEEE Xplore	/	/		
		ScienceDirect	27	0		
„Process Mining“ AND „everyday mobility“		ProQuest	/	/		
		Oxford Academic	/	/		
		IEEE Xplore	/	/		
		ScienceDirect	/	/		
Process Mining AND Transportprozess		ProQuest	1	0		
		Oxford Academic	1	0		
		IEEE Xplore	/	/		
		ScienceDirect	1	0		
		ProQuest	1	0		

Process-Mining AND Transportprozess		Oxford Academic	1	0		
		IEEE Xplore	/	/		
		ScienceDirect	1	0		
„Process Mining“ AND Transportprozess		ProQuest	1	0		
		Oxford Academic	/	/		
		IEEE Xplore	/	/		
		ScienceDirect	/	/		
Process Mining AND ÖPNV		ProQuest	7	0		
		Oxford Academic	/	/		
		IEEE Xplore	/	/		
		ScienceDirect	1	0		
Process-Mining AND ÖPNV		ProQuest	/	/		
		Oxford Academic	/	/		
		IEEE Xplore	/	/		
		ScienceDirect	1	0		
„Process Mining“ AND ÖPNV		ProQuest	/	/		
		Oxford Academic	/	/		
		IEEE Xplore	/	/		
		ScienceDirect	/	/		
Process Mining AND public transport		ProQuest	86344	- s. Notizen		Anfrage zu ungenau
		Oxford Academic	25521	- s. Notizen		Anfrage zu ungenau
		IEEE Xplore	92	0		
		ScienceDirect	42153	- s. Notizen		Anfrage zu ungenau
Process-Mining AND public transport		ProQuest	240	- s. Notizen		Anfrage zu ungenau
		Oxford Academic	25522	- s. Notizen		Anfrage zu ungenau
		IEEE Xplore	/	/		
		ScienceDirect	42153	- s. Notizen		Anfrage zu ungenau
„Process Mining“ AND public transport		ProQuest	240	- s. Notizen		Anfrage zu ungenau
		Oxford Academic	32	0		
		IEEE Xplore	/	/		
		ScienceDirect	243	- s. Notizen		Anfrage zu ungenau
Process Mining AND public- transport		ProQuest	6076	- s. Notizen		Anfrage zu ungenau
		Oxford Academic	25522	- s. Notizen		Anfrage zu ungenau
		IEEE Xplore	66	0		

		ScienceDirect	42153	- s. Notizen		Anfrage zu ungenau
Process-Mining AND public- transport		ProQuest	20	0		
		Oxford Academic	25521	- s. Notizen		Anfrage zu ungenau
		IEEE Xplore	/	/		
		ScienceDirect	42153	- s. Notizen		Anfrage zu ungenau
„Process Mining“ AND public-transport		ProQuest	20	0		
		Oxford Academic	32	0		
		IEEE Xplore	/	/		
		ScienceDirect	243	- s. Notizen		Anfrage zu ungenau
Process Mining AND „public transport“		ProQuest	6076	- s. Notizen		Anfrage zu ungenau
		Oxford Academic	672	- s. Notizen		Anfrage zu ungenau
		IEEE Xplore	66	0		
		ScienceDirect	3055	- s. Notizen		Anfrage zu ungenau
Process-Mining AND „public transport“		ProQuest	20	0		
		Oxford Academic	672	- s. Notizen		Anfrage zu ungenau
		IEEE Xplore	/	/		
		ScienceDirect	3055	- s. Notizen		Anfrage zu ungenau
„Process Mining“ AND „public transport“		ProQuest	20	0		
		Oxford Academic	4	0		
		IEEE Xplore	/	/		
		ScienceDirect	14	1	„Detection of vehicle-based operations from geolocation data“	
Process Mining AND räumliche Mobilität		ProQuest	12	0		
		Oxford Academic	/	/		
		IEEE Xplore	/	/		
		ScienceDirect	1	0		
Process-Mining AND räumliche Mobilität		ProQuest	2	0		
		Oxford Academic	/	/		
		IEEE Xplore	/	/		
		ScienceDirect	1	0		
„Process Mining“ AND räumliche Mobilität		ProQuest	2	0		
		Oxford Academic	/	/		
		IEEE Xplore	/	/		
		ScienceDirect	/	/		

Process Mining AND räumliche- Mobilität		ProQuest	3	0		
		Oxford Academic	/	/		
		IEEE Xplore	/	/		
		ScienceDirect	1	0		
Process-Mining AND räumliche- Mobilität		ProQuest	/	/		
		Oxford Academic	/	/		
		IEEE Xplore	/	/		
		ScienceDirect	1	0		
„Process Mining“ AND räumliche- Mobilität		ProQuest	/	/		
		Oxford Academic	/	/		
		IEEE Xplore	/	/		
		ScienceDirect	/	/		
Process Mining AND „räumliche Mobilität“		ProQuest	3	0		
		Oxford Academic	/	/		
		IEEE Xplore	/	/		
		ScienceDirect	/	/		
Process-Mining AND „räumliche Mobilität“		ProQuest	/	/		
		Oxford Academic	/	/		
		IEEE Xplore	/	/		
		ScienceDirect	/	/		
„Process Mining“ AND „räumliche Mobilität“		ProQuest	/	/		
		Oxford Academic	/	/		
		IEEE Xplore	/	/		
		ScienceDirect	/	/		
Process Mining AND travel		ProQuest	79103	- s. Notizen		Anfrage zu ungenau
		Oxford Academic	19539	- s. Notizen		Anfrage zu ungenau
		IEEE Xplore	541	- s. Notizen		Anfrage zu ungenau (Ergeb- nisse ent- halten z.B. nur Begriff „Process“ ohne „Mining“ usw.)
		ScienceDirect	35094	- s. Notizen		Anfrage zu ungenau
Process-Mining AND travel		ProQuest	250	- s. Notizen		Anfrage zu ungenau
		Oxford Academic	19539	- s. Notizen		Anfrage zu ungenau
		IEEE Xplore	6	3	„An interest- based tour planning tool by Process	

					Mining from Twitter“ „Tour Miner: Mining system of tour plans from SNS - Smelting Function from Travel Records to Tour Routes“ „Location-aware path alignment in Process Mining“	
		ScienceDirect	35094	- s. Notizen		Anfrage zu ungenau
„Process Mining“ AND travel		ProQuest	250	- s. Notizen		Anfrage zu ungenau
		Oxford Academic	11	0		
		IEEE Xplore	5	3	„An interest-based tour planning tool by Process Mining from Twitter“ „Tour Miner: Mining system of tour plans from SNS - Smelting Function from Travel Records to Tour Routes“ „Location-aware path alignment in Process Mining“	
		ScienceDirect	190	2	„Process discovery on geolocation data“ „Detection of vehicle-based operations from geolocation data“	

Process Mining AND trip		ProQuest	31930	- s. Notizen		Anfrage zu ungenau
		Oxford Academic	6597	- s. Notizen		Anfrage zu ungenau
		IEEE Xplore	118	0		
		ScienceDirect	12134	- s. Notizen		Anfrage zu ungenau
Process-Mining AND trip		ProQuest	88	0		
		Oxford Academic	6597	- s. Notizen		Anfrage zu ungenau
		IEEE Xplore	1	0		
		ScienceDirect	12134	- s. Notizen		Anfrage zu ungenau
„Process Mining“ AND trip		ProQuest	88	0		
		Oxford Academic	4	0		
		IEEE Xplore	1	0		
		ScienceDirect	60	3	„Detecting train reroutings with process mining: A Belgian application“ „Process discovery on geolocation data“ „Detection of vehicle-based operations from geolocation data“	
Process Mining AND everyday mobility	Kurz- fassung	ProQuest	2	0		
	Abstract	Oxford Academic	/	/		
	Title, Abstract, Keywords	ScienceDirect	/	/		
Process-Mining AND everyday mobility	Abstract	Oxford Academic	/	/		
	Title, Abstract, Keywords	ScienceDirect	/	/		
Process Mining AND everyday- mobility	Abstract	Oxford Academic	/	/		
	Title, Abstract, Keywords	ScienceDirect	/	/		
	Abstract	Oxford Academic	/	/		

Process-Mining AND everyday- mobility	Title, Abstract, Keywords	ScienceDirect	/	/		
Process Mining AND public transport	Kurz- fassung	ProQuest	20	0		
	Abstract	Oxford Academic	1	0		
	Title, Abstract, Keywords	ScienceDirect	41	1	„Detection of vehicle-based operations from geolocation data“	
Process-Mining AND public transport	Kurz- fassung	ProQuest	/	/		
	Abstract	Oxford Academic	1	0		
	Title, Abstract, Keywords	ScienceDirect	41	1	„Detection of vehicle-based operations from geolocation data“	
„Process Mining“ AND public transport	Kurz- fassung	ProQuest	/	/		
	Title, Abstract, Keywords	ScienceDirect	1	1	„Detection of vehicle-based operations from geolocation data“	
Process Mining AND public- transport	Kurz- fassung	ProQuest	2	0		
	Abstract	Oxford Academic	1	0		
	Title, Abstract, Keywords	ScienceDirect	41	1	„Detection of vehicle-based operations from geolocation data“	
Process-Mining AND public- transport	Abstract	Oxford Academic	1	0		
	Title, Abstract, Keywords	ScienceDirect	41	1	„Detection of vehicle-based operations from geolocation data“	
„Process Mining“ AND public-transport	Title, Abstract, Keywords	ScienceDirect	1	1	„Detection of vehicle-based operations from geolocation data“	
Process Mining AND „public transport“	Kurz- fassung	ProQuest	2	0		
	Abstract	Oxford Academic	/	/		

	Title, Abstract, Keywords	ScienceDirect	14	1	„Detection of vehicle-based operations from geolocation data“	
Process-Mining AND „public transport“	Abstract	Oxford Academic	/	/		
	Title, Abstract, Keywords	ScienceDirect	14	1	„Detection of vehicle-based operations from geolocation data“	
Process Mining AND travel	Kurz- fassung	ProQuest	77	1	„Discovering commute patterns via Process Mining“	
	Abstract	Oxford Academic	/	/		
	Abstract, Index Terms	IEEE Xplore	42262	- s. Notizen		Anfrage zu ungenau (Index Terms)
	Abstract	IEEE Xplore	104	3	„Tour Miner: Mining system of tour plans from SNS - Smelting Function from Travel Records to Tour Routes“ „Location- aware path alignment in Process Mining“ „An interest- based tour planning tool by Process Mining from Twitter“	
	Title, Abstract, Keywords	ScienceDirect	88	0		
„Process Mining“ AND travel	Kurz- fassung	ProQuest	1	1	„Discovering commute patterns via Process Mining“	
	Abstract	Oxford Academic	/	/		
	Abstract, Index Terms	IEEE Xplore	25955	- s. Notizen		Anfrage zu ungenau

						(Index Terms)
	Abstract	IEEE Xplore	5	3	„Tour Miner: Mining system of tour plans from SNS - Smelting Function from Travel Records to Tour Routes“ „An interest-based tour planning tool by Process Mining from Twitter“ „Location-aware path alignment in Process Mining“	
	Title, Abstract, Keywords	ScienceDirect	1	0		
Process-Mining AND travel	Kurz-fassung	ProQuest	1	1	„Discovering commute patterns via Process Mining“	
	Abstract	Oxford Academic	/	/		
	Title, Abstract, Keywords	ScienceDirect	88	0		
Process Mining AND trip	Kurz-fassung	ProQuest	25	0		
	Abstract	Oxford Academic	2	0		
	Title, Abstract, Keywords	ScienceDirect	36	0		
Process-Mining AND trip	Abstract	Oxford Academic	2	0		
	Title, Abstract, Keywords	ScienceDirect	36	0		

Textkonvertierungs-Assistent - Schritt 1 von 3

Der Textkonvertierungs-Assistent hat erkannt, dass Ihre Daten eine feste Breite haben.

Wenn alle Angaben korrekt sind, klicken Sie auf 'Weiter', oder wählen Sie den korrekten Datentyp.

Ursprünglicher Datentyp

Wählen Sie den Datentyp, der Ihre Daten am besten beschreibt:

☒ Getrennt - Zeichen wie z.B. Kommas oder Tabstopps trennen Felder (Excel 4.0-Standard).

☐ Feste Breite - Felder sind in Spalten ausgerichtet, mit Leerzeichen zwischen jedem Feld.

Vorschau der markierten Daten:

1	name	address	ExtendedData/Data/0/value	ExtendedData/Data/0/_name	ExtendedData/Data/0/_value
2	Driving		Email, Driving	Category	5690, Distance
3	Walking		Email, Walking	Category	739, Distance
4	Walking		Email, Walking	Category	691, Distance
5	Driving		Email, Driving	Category	5467, Distance

Abbrechen < Zurück Weiter > Fertig stellen

Abbildung 48: Anleitung Text in Spalten, Teil 1. Eigene Darstellung (2022)

Textkonvertierungs-Assistent - Schritt 2 von 3

Dieses Dialogfeld ermöglicht es Ihnen, Trennzeichen festzulegen. Sie können in der Vorschau der markierten Daten sehen, wie Ihr Text erscheinen wird.

Trennzeichen

☐ Tabstopp

☐ Semikolon

☒ Komma

☐ Leerzeichen

☐ Andere:

☐ Aufeinanderfolgende Trennzeichen als ein Zeichen behandeln

Textqualifizierer: "

Datenvorschau

name	address	ExtendedData/Data/0/value
Driving		Email, Driving
Walking		Email, Walking
Walking		Email, Walking
Driving		Email, Driving

Abbrechen < Zurück Weiter > Fertig stellen

Abbildung 49: Anleitung Text in Spalten, Teil 2. Eigene Darstellung (2022)

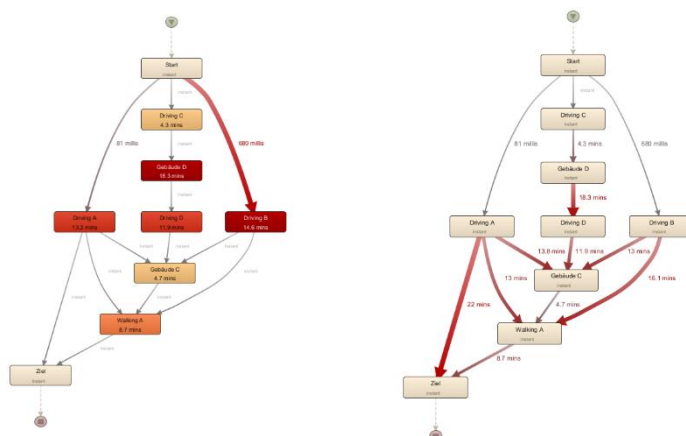


Abbildung 50: Unterschied Prozessmodell mit Schlusszeit als Timestamp (links) und ohne Schlusszeit als Timestamp (rechts). Eigene Darstellung (2022)

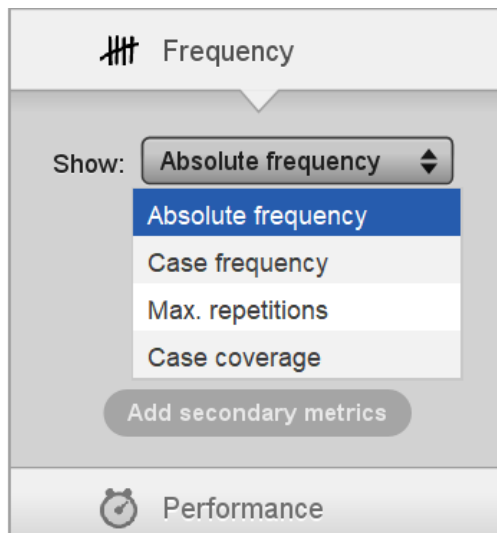


Abbildung 51: Auswahlmöglichkeiten unter dem Reiter 'Frequency'. Eigene Darstellung (2022)

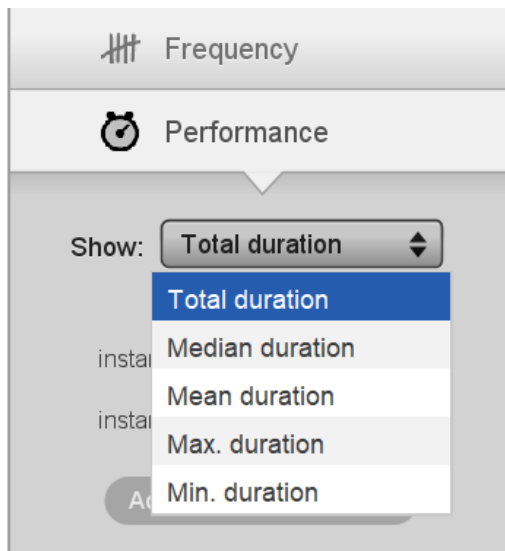


Abbildung 52: Auswahlmöglichkeiten unter dem Reiter 'Performance'. Eigene Darstellung (2022)

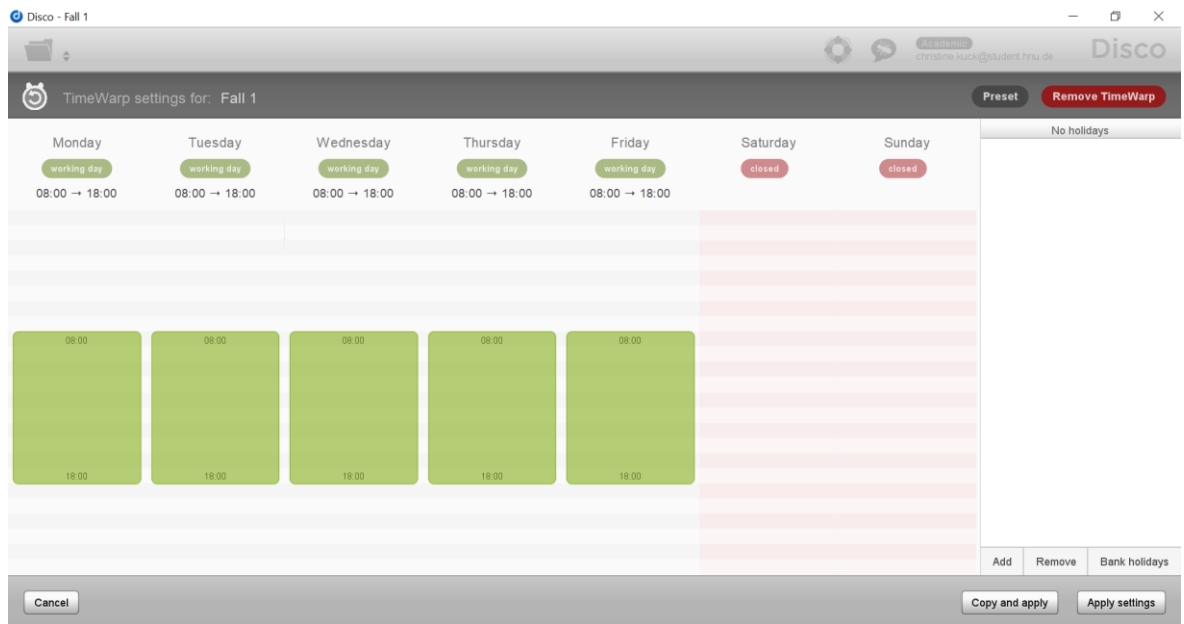


Abbildung 53: Einstellung von Geschäfts- und Urlaubszeiten. Eigene Darstellung (2022)

Events	236
Cases	55
Activities	9
Median case duration	23.4 mins
Mean case duration	23.9 mins
Start	16.03.2021 06:20:03
End	19.08.2021 05:18:32

Abbildung 54: Auswertung Fall 1 - Allgemeine Informationen. Eigene Darstellung (2022)



Abbildung 55: Auswertung Fall 1 - Varianten in Prozent. Eigene Darstellung (2022)

Cases (55) Variants (6)				
Variant	Cases	Events	Median duration	Mean duration
Variant 1	30	4	22 mins, 27 secs	22 mins, 40 secs
Variant 2	9	4	24 mins, 12 secs	24 mins, 54 secs
Variant 3	8	5	23 mins, 56 secs	24 mins, 22 secs
Variant 4	6	5	24 mins, 35 secs	24 mins, 50 secs
Variant 5	1	3	22 mins, 1 sec	22 mins, 1 sec
Variant 6	1	7	45 mins, 40 secs	45 mins, 40 secs

Abbildung 56: Auswertung Fall 1 - Tabelle: 'Variants'. Eigene Darstellung (2022)

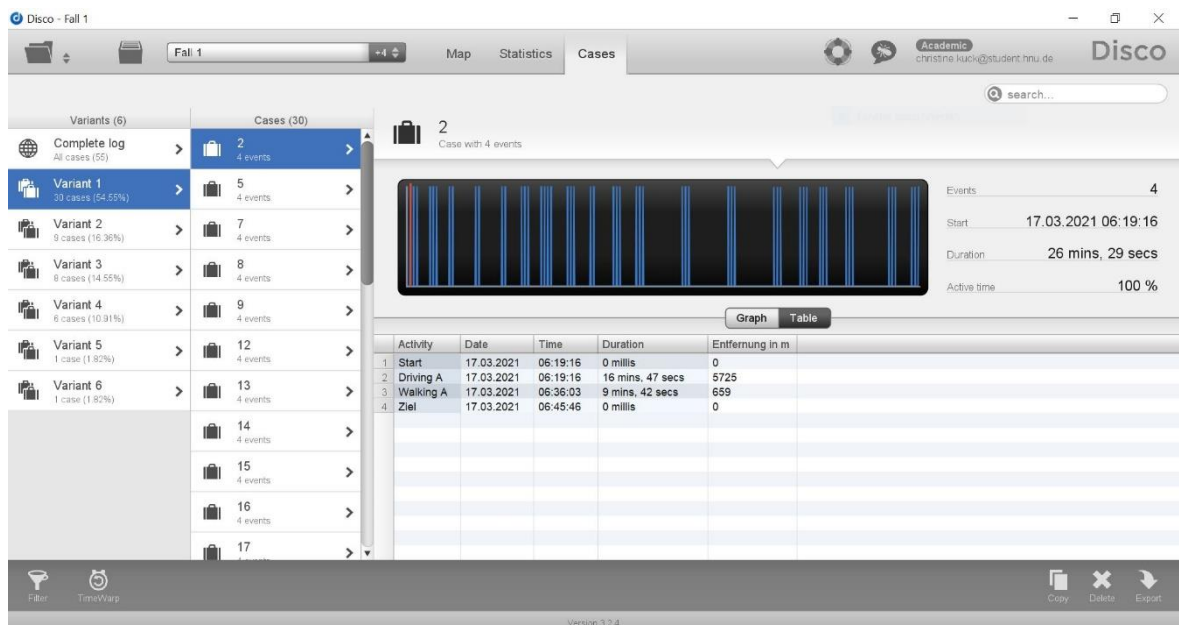


Abbildung 57: Auswertung Fall 1 - Mobilitätsprozess Variante 1. Eigene Darstellung (2022)

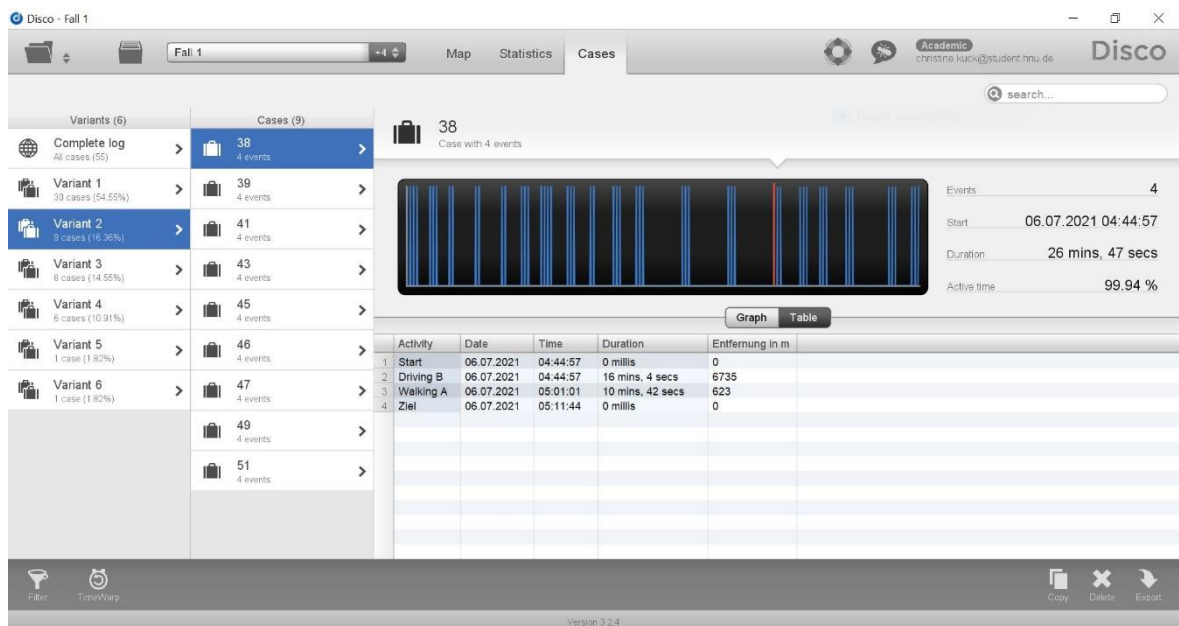


Abbildung 58: Auswertung Fall 1 - Mobilitätsprozess Variante 2. Eigene Darstellung (2022)

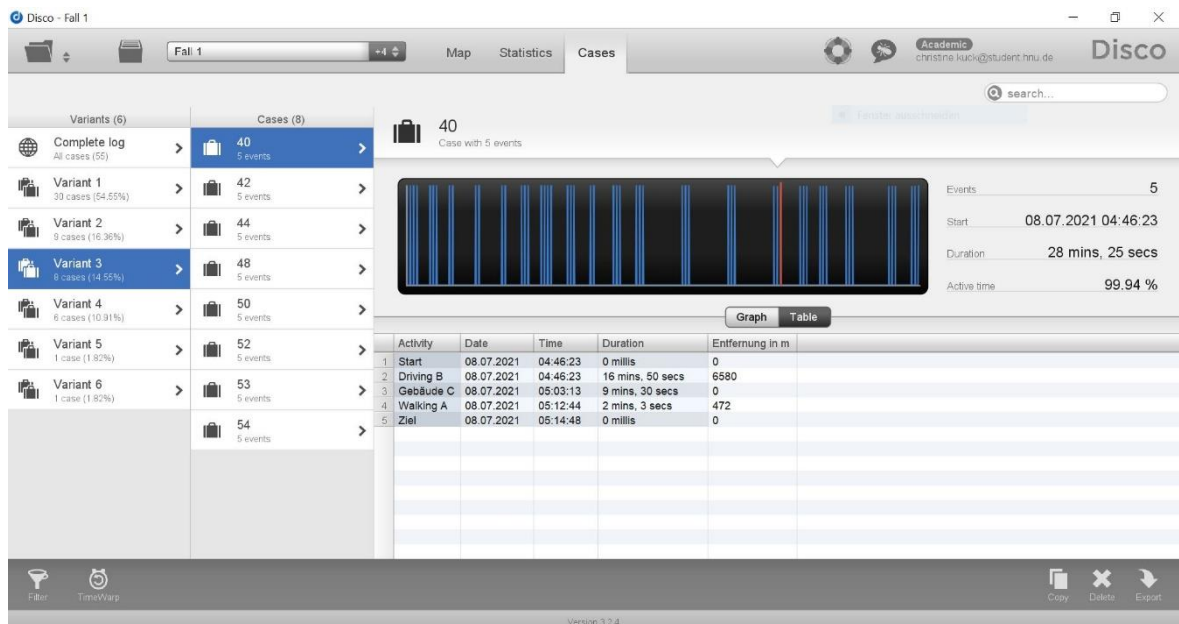


Abbildung 59: Auswertung Fall 1 - Mobilitätsprozess Variante 3. Eigene Darstellung (2022)

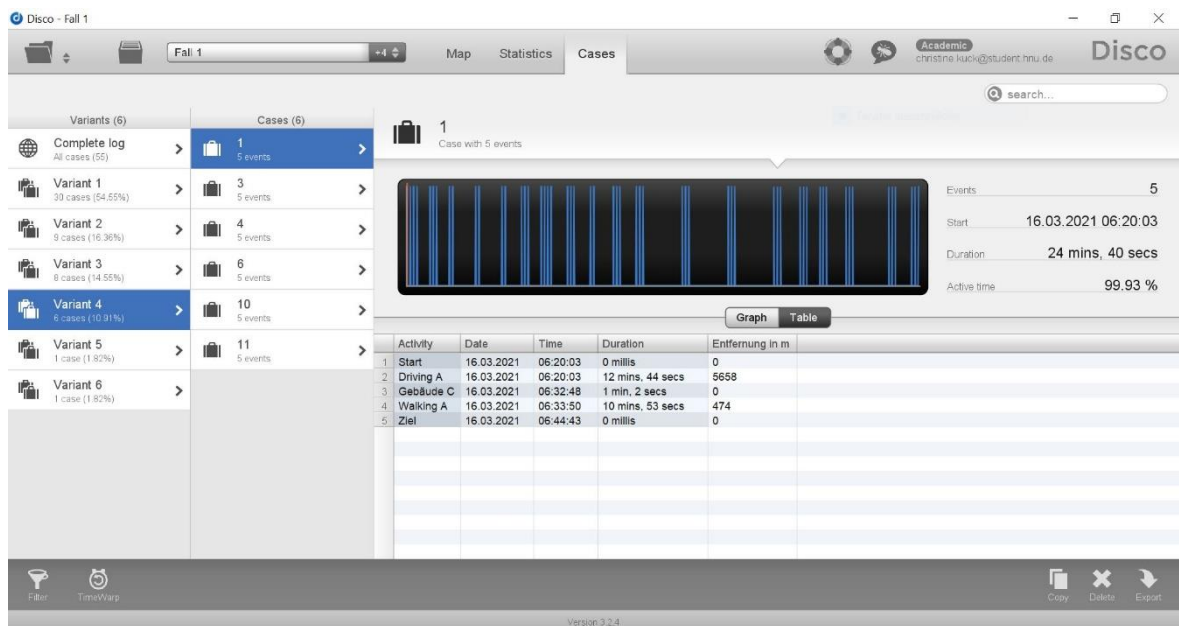


Abbildung 60: Auswertung Fall 1 - Mobilitätsprozess Variante 4. Eigene Darstellung (2022)

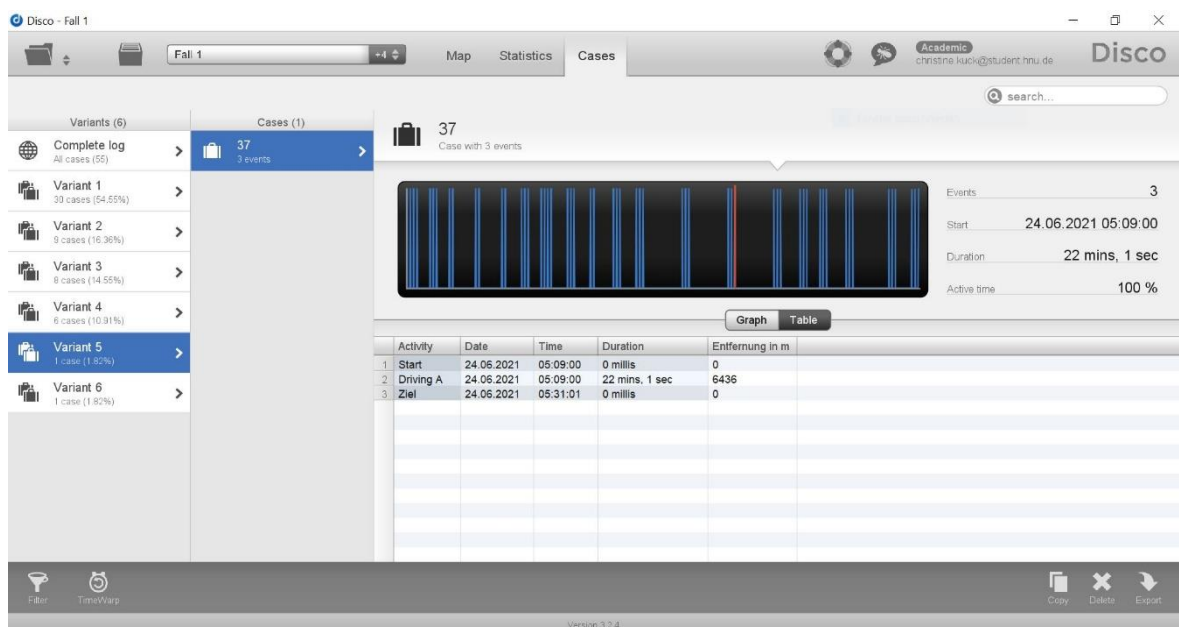


Abbildung 61: Auswertung Fall 1 - Mobilitätsprozess Variante 5. Eigene Darstellung (2022)

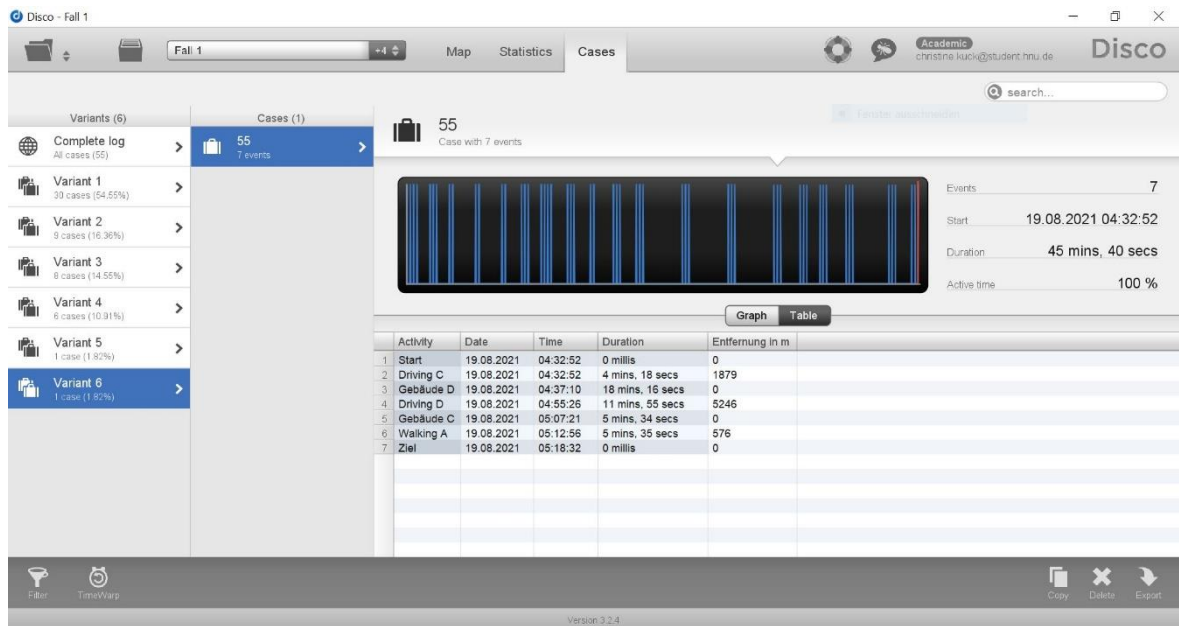


Abbildung 62: Auswertung Fall 1 - Mobilitätsprozess Variante 6. Eigene Darstellung (2022)

Tabelle 21: Auswertung Fall 1 - Ergebnisse KPI V. Eigene Darstellung (2022)

Variante	V
Variante 1	1,259259259
Variante 2	1,310526316
Variante 3	1,28245614
Variante 4	1,37962963
Variante 5	1,693589744
Variante 6	2,075757576

Tabelle 22: Auswertung Fall 1 - Ergebnisse KPIs G & D. Eigene Darstellung (2022)

Case	G (in km/h)	D
1	14,91567568	1,032078103
2	14,46343612	1,108089261
3	13,94299938	1,120641562
4	15,24	1,025104603
5	14,78384106	1,052998605
6	14,2052801	1,082984658
7	15,43989326	1,045327755
8	18,28743083	0,882147838
9	16,11319149	0,983263598
10	16,24239829	0,976987448

11	15,72260322	0,99651325
12	18,09873981	0,940725244
13	16,84314013	0,950488145
14	17,46129496	0,969316597
15	18,15640449	0,930962343
16	16,66180758	0,956764296
17	17,73333333	0,90376569
18	19,78550725	0,866108787
19	17,1097897	0,961645746
20	16,88791541	0,923291492
21	16,99285714	0,937238494
22	17,59318885	0,90097629
23	16,96848138	0,973500697
24	16,10444763	0,972105997
25	18,8296875	0,892608089
26	18,22028986	0,866108787
27	17,31052233	0,921199442
28	17,60990712	0,90097629
29	16,89895052	0,930264993
30	17,48887183	0,908647141
31	16,98975501	0,939330544
32	15,93347368	0,993723849
33	15,61703854	1,031380753
34	15,88265162	0,988842399
35	16,42316258	0,939330544
36	17,74245283	0,887029289
37	17,53943982	0,921199442
38	16,48338519	1,120641562
39	16,68069666	0,960948396
40	14,88985337	1,188981869
41	16,74703329	0,963737796
42	15,58481013	0,991631799
43	16,01688581	1,007670851
44	14,57242718	1,077405858

45	15,71900826	1,012552301
46	22,63324022	0,9986053
47	14,3866005	1,124128312
48	17,41863354	0,89818689
49	14,24134078	1,123430962
50	15,90514051	1,017433752
51	16,20471822	1,064156206
52	16,69922371	0,988145049
53	16,59227586	1,011157601
54	17,39061166	0,980474198
55	10,11810219	1,910739191

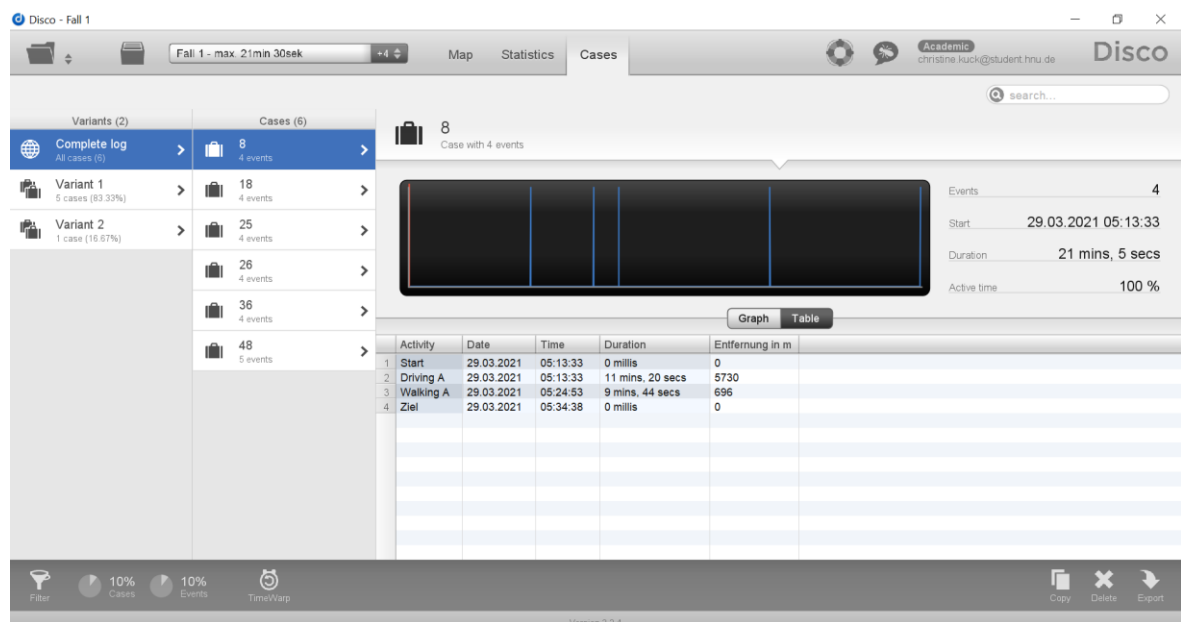


Abbildung 63: Auswertung Fall 1 - Filter: Kürzeste Cases. Eigene Darstellung (2022)

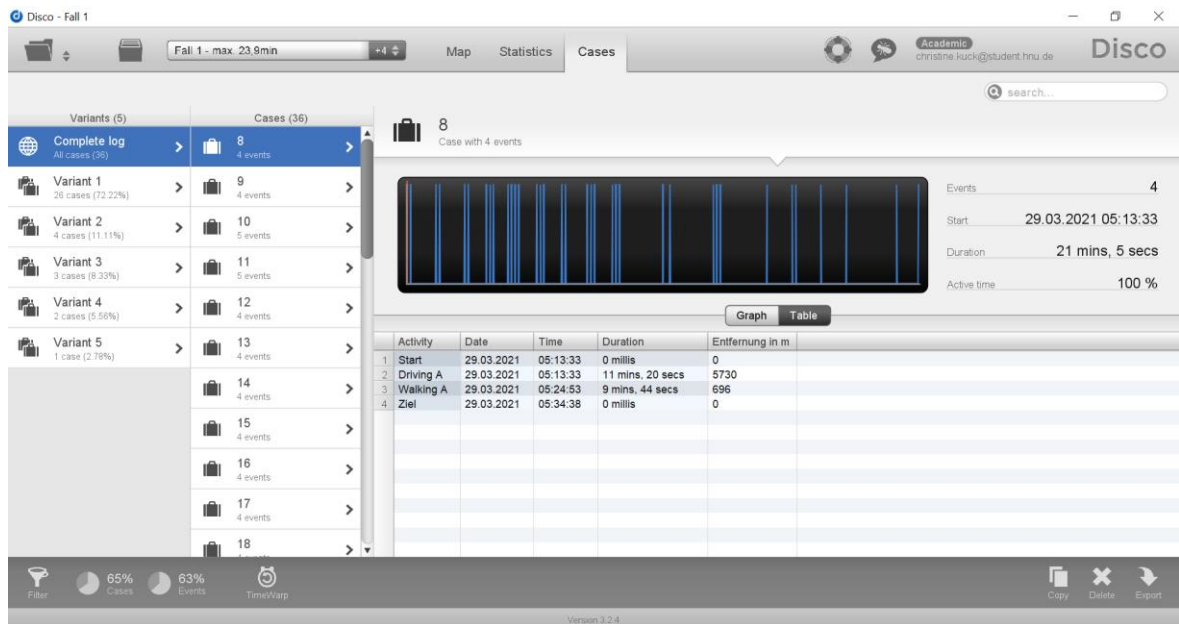


Abbildung 64: Auswertung Fall 1 - Filter: Kürzer als Durchschnitt. Eigene Darstellung (2022)

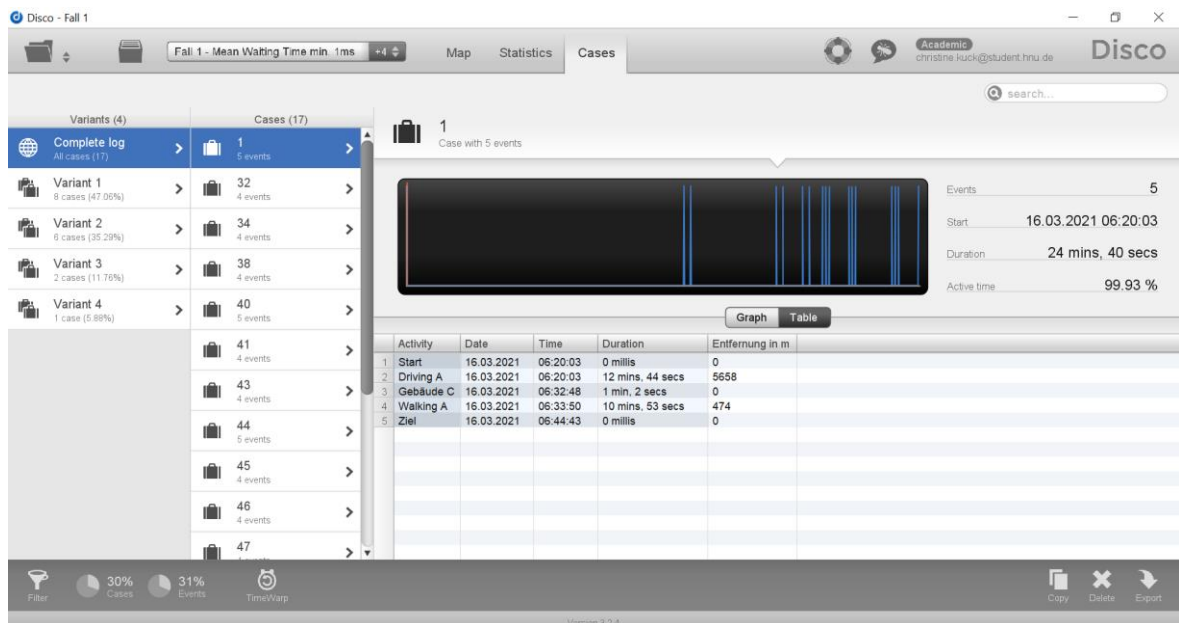


Abbildung 65: Auswertung Fall 1 - Filter: Cases mit Wartezeiten. Eigene Darstellung (2022)

Events	158
Cases	41
Activities	20
Median case duration	29.3 mins
Mean case duration	30.1 mins
Start	14.04.2020 03:20:46
End	26.06.2020 04:09:14

Abbildung 66: Auswertung Fall 2 - allgemeine Informationen. Eigene Darstellung (2022)

Variants (15)	
 Complete log	>
All cases (41)	
 Variant 1	>
10 cases (24.39%)	
 Variant 2	>
7 cases (17.07%)	
 Variant 3	>
5 cases (12.2%)	
 Variant 4	>
3 cases (7.32%)	
 Variant 5	>
3 cases (7.32%)	
 Variant 6	>
2 cases (4.88%)	
 Variant 7	>
2 cases (4.88%)	
 Variant 8	>
2 cases (4.88%)	
 Variant 9	>
1 case (2.44%)	
 Variant 10	>
1 case (2.44%)	
 Variant 11	>
1 case (2.44%)	
 Variant 12	>
1 case (2.44%)	
 Variant 13	>
1 case (2.44%)	
 Variant 14	>
1 case (2.44%)	
 Variant 15	>
1 case (2.44%)	

Abbildung 67: Auswertung Fall 2 - Varianten in Prozent. Eigene Darstellung (2022)

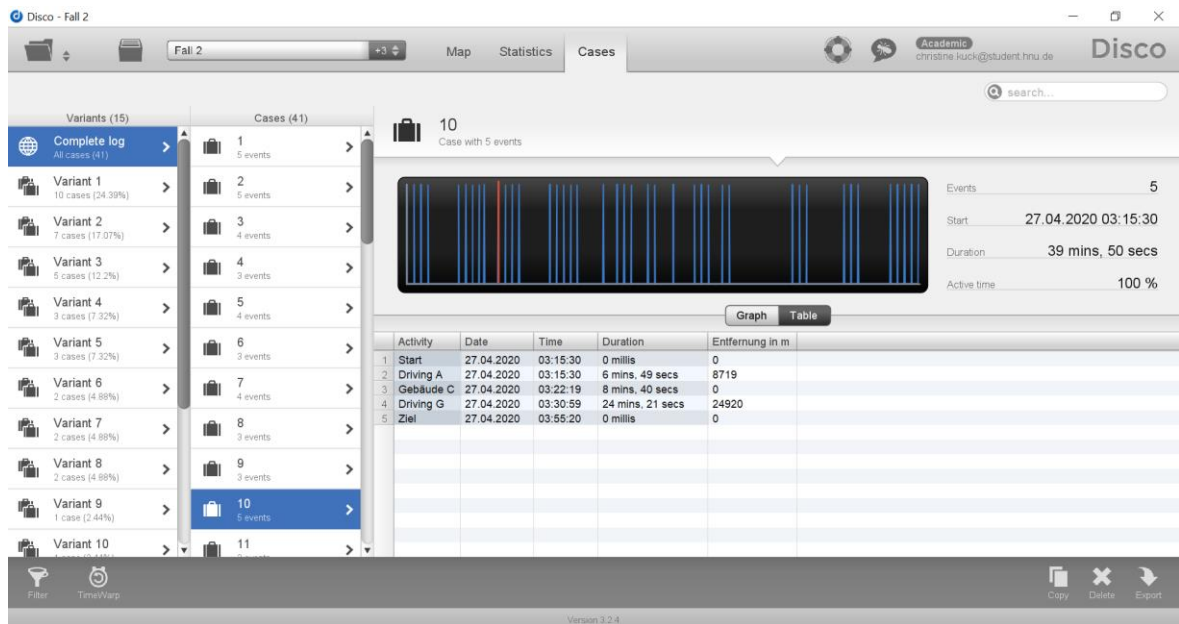


Abbildung 68: Auswertung Fall 2 - Dauer und Entfernung der Cases. Eigene Darstellung (2022)

		Cases (41)	Variants (15)
Variant	Cases	Events	Median duration
Variant 1	10	4	31 mins, 8 secs
Variant 2	7	3	22 mins, 30 secs
Variant 3	5	3	29 mins, 6 secs
Variant 4	3	4	25 mins, 32 secs
Variant 5	3	3	20 mins, 53 secs
Variant 6	2	5	31 mins, 53 secs
Variant 7	2	4	33 mins, 34 secs
Variant 8	2	4	22 mins, 3 secs
Variant 9	1	3	31 mins, 28 secs
Variant 10	1	5	39 mins, 50 secs
Variant 11	1	5	47 mins, 42 secs
Variant 12	1	7	48 mins, 27 secs
Variant 13	1	5	43 mins, 54 secs
Variant 14	1	5	43 mins, 49 secs
Variant 15	1	5	41 mins, 22 secs

Abbildung 69: Auswertung Fall 2 - Tabelle: 'Variants'. Eigene Darstellung (2022)

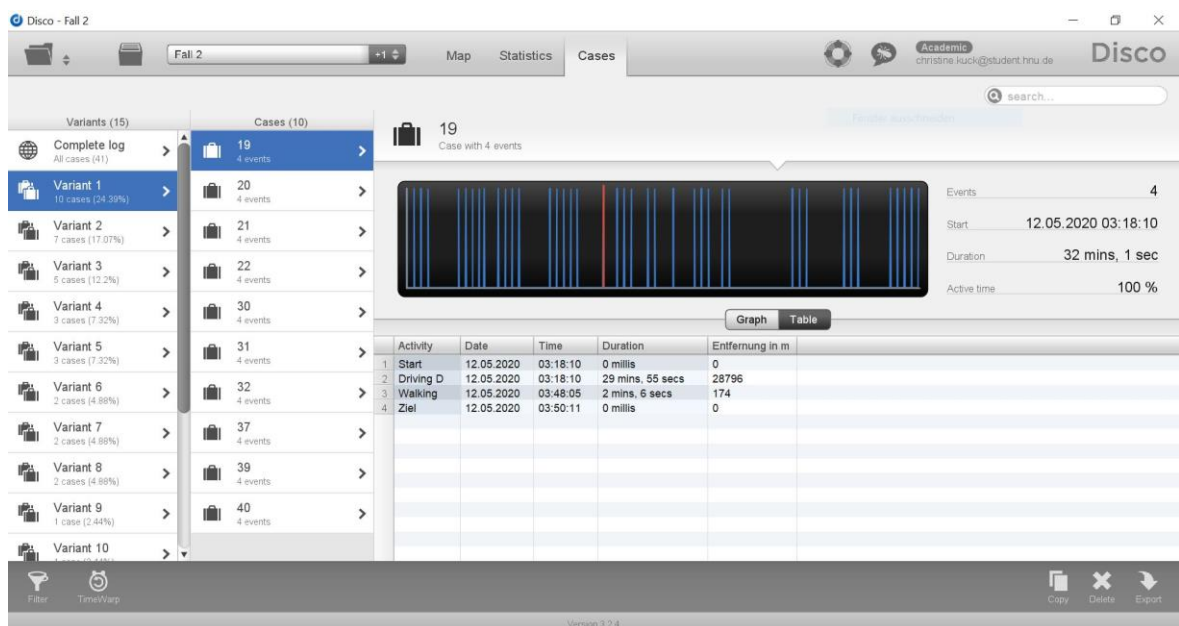


Abbildung 70: Auswertung Fall 2 - Prozessablauf Variante 1. Eigene Darstellung (2022)

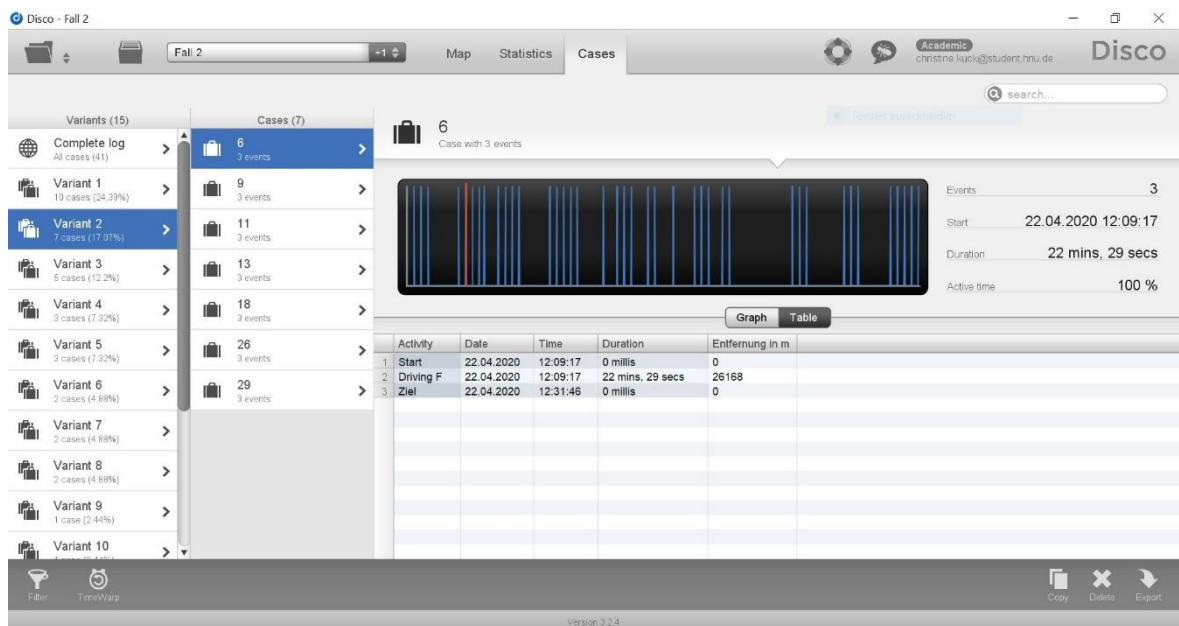


Abbildung 71: Auswertung Fall 2 - Prozessablauf Variante 2. Eigene Darstellung (2022)

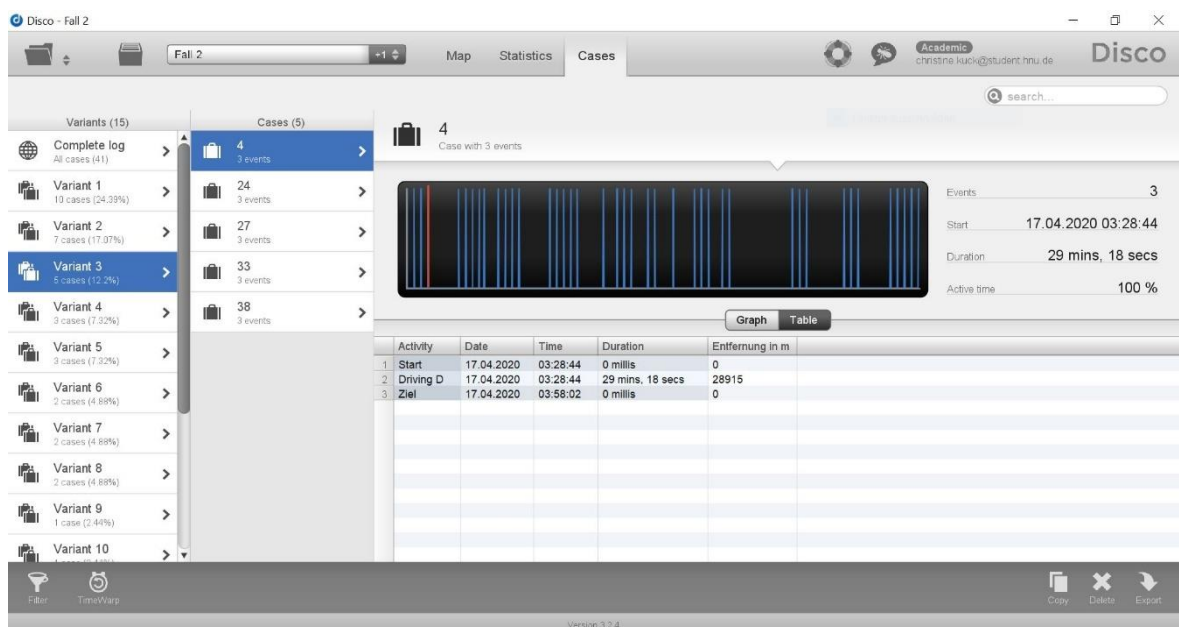


Abbildung 72: Auswertung Fall 2 - Prozessablauf Variante 3. Eigene Darstellung (2022)

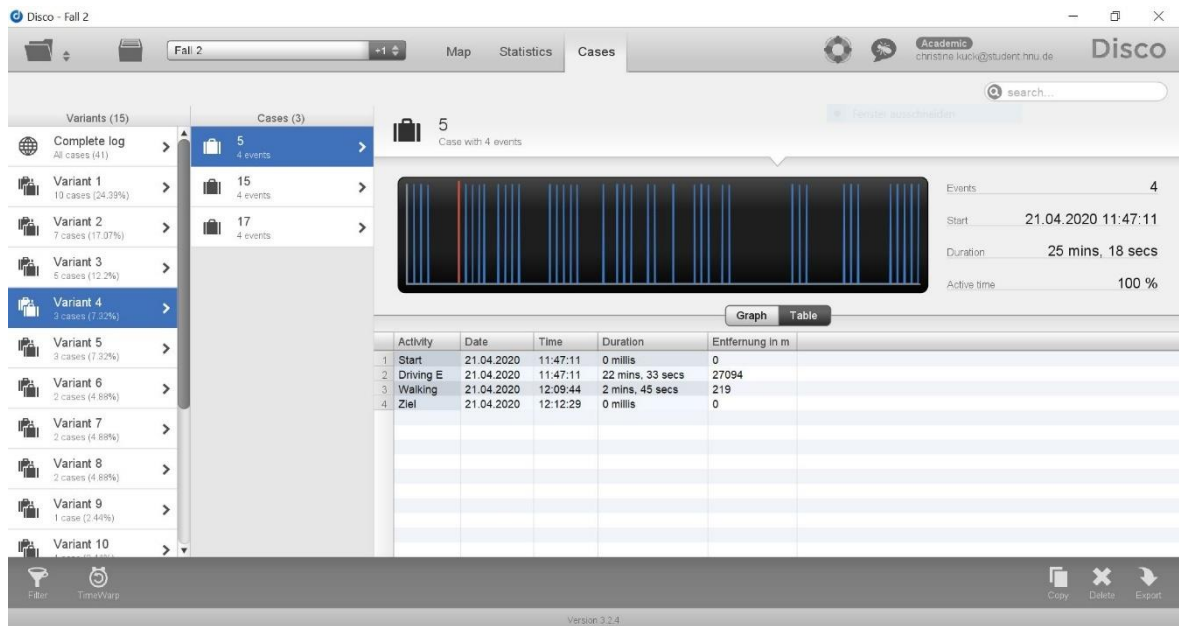


Abbildung 73: Auswertung Fall 2 - Prozessablauf Variante 4. Eigene Darstellung (2022)

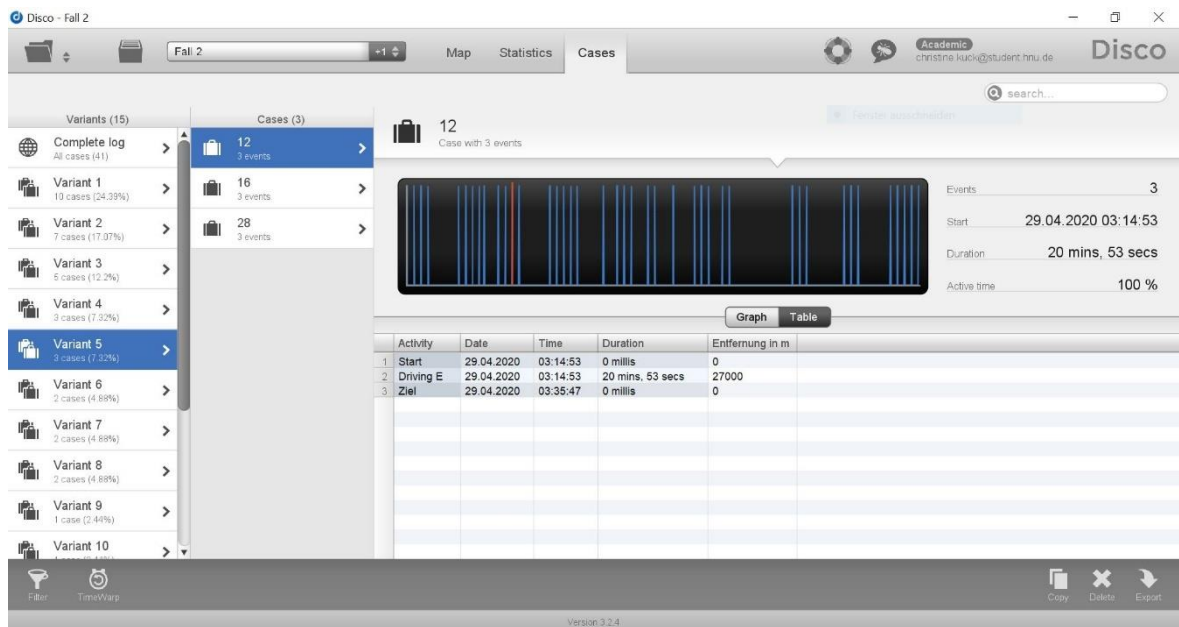


Abbildung 74: Auswertung Fall 2 - Prozessablauf Variante 5. Eigene Darstellung (2022)

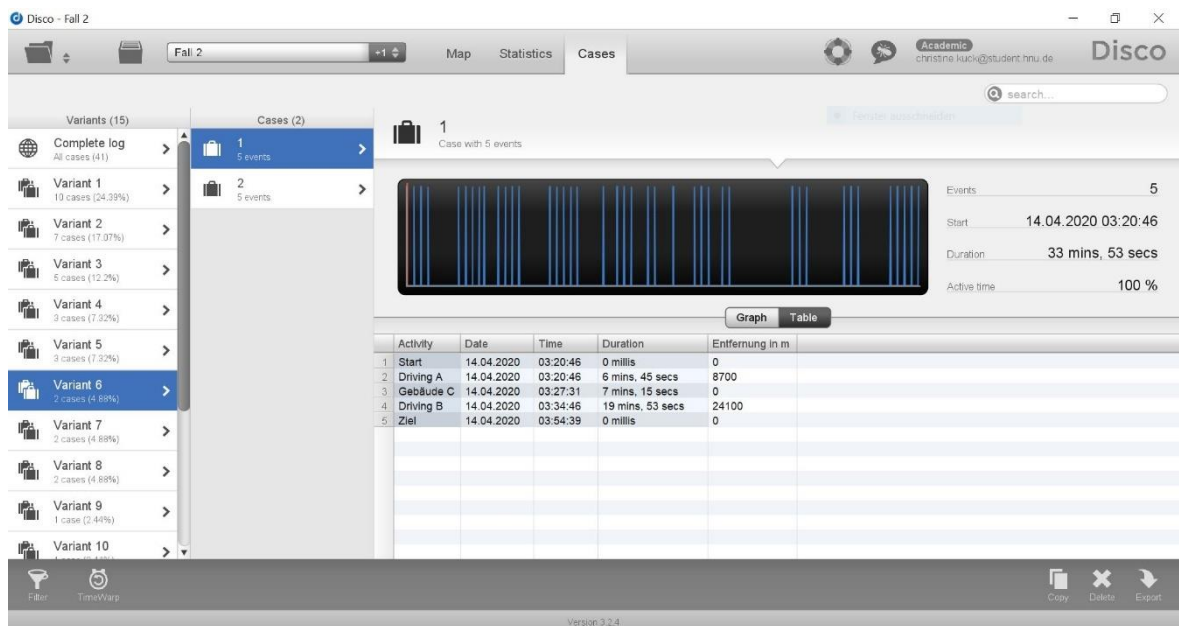


Abbildung 75: Auswertung Fall 2 - Prozessablauf Variante 6. Eigene Darstellung (2022)

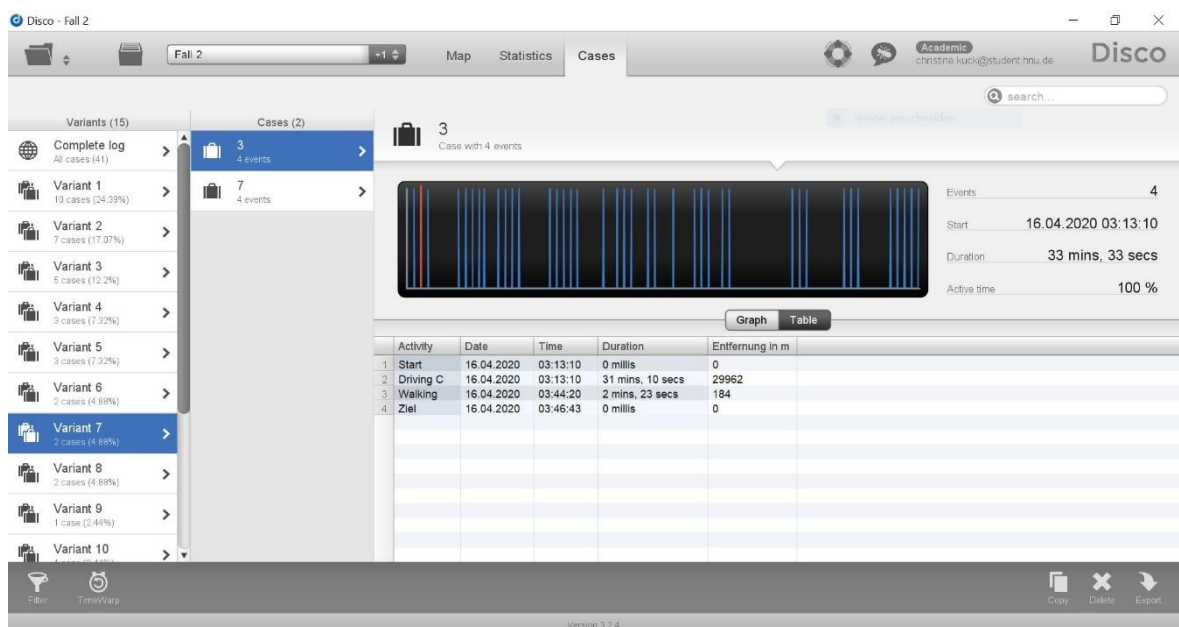


Abbildung 76: Auswertung Fall 2 - Prozessablauf Variante 7. Eigene Darstellung (2022)

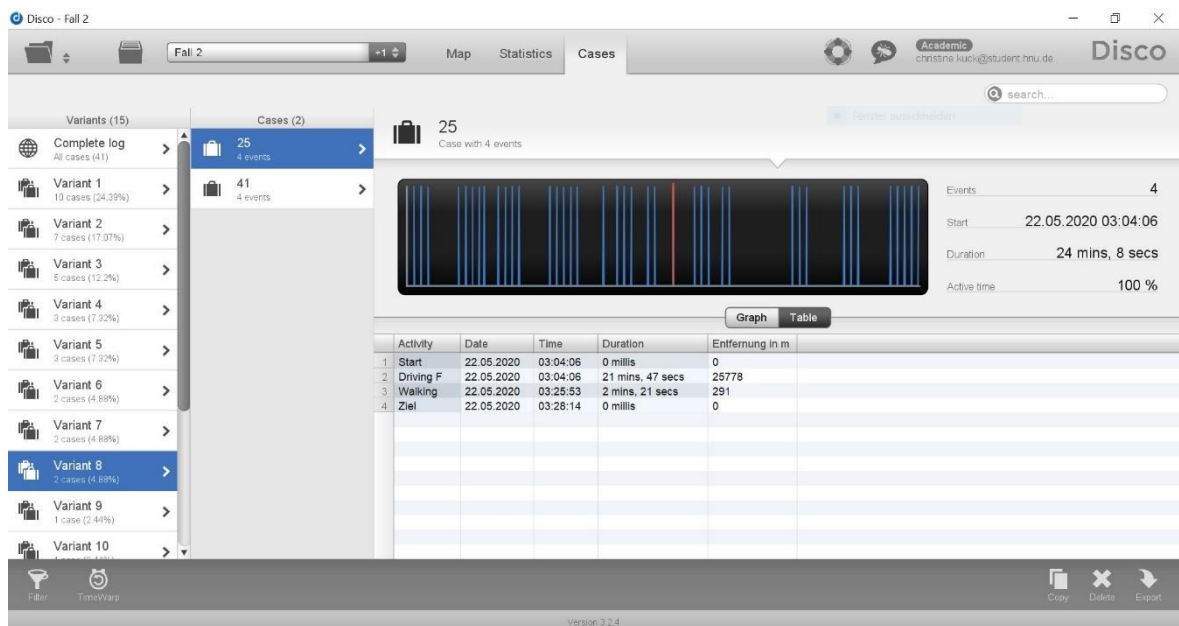


Abbildung 77: Auswertung Fall 2 - Prozessablauf Variante 8. Eigene Darstellung (2022)

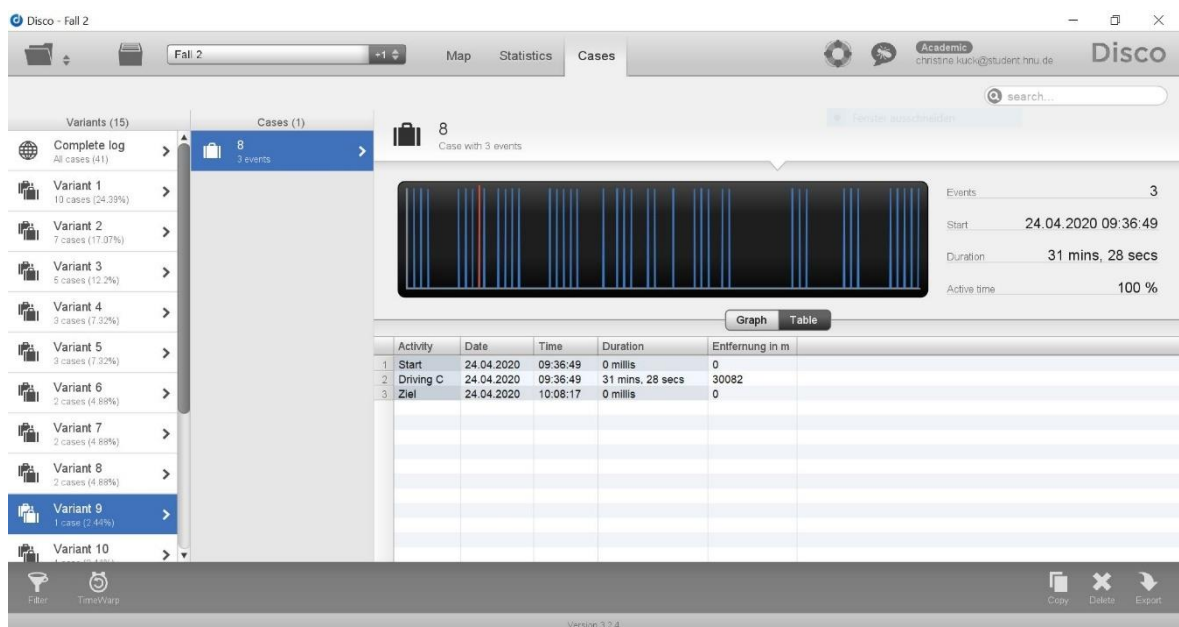


Abbildung 78: Auswertung Fall 2 - Prozessablauf Variante 9. Eigene Darstellung (2022)

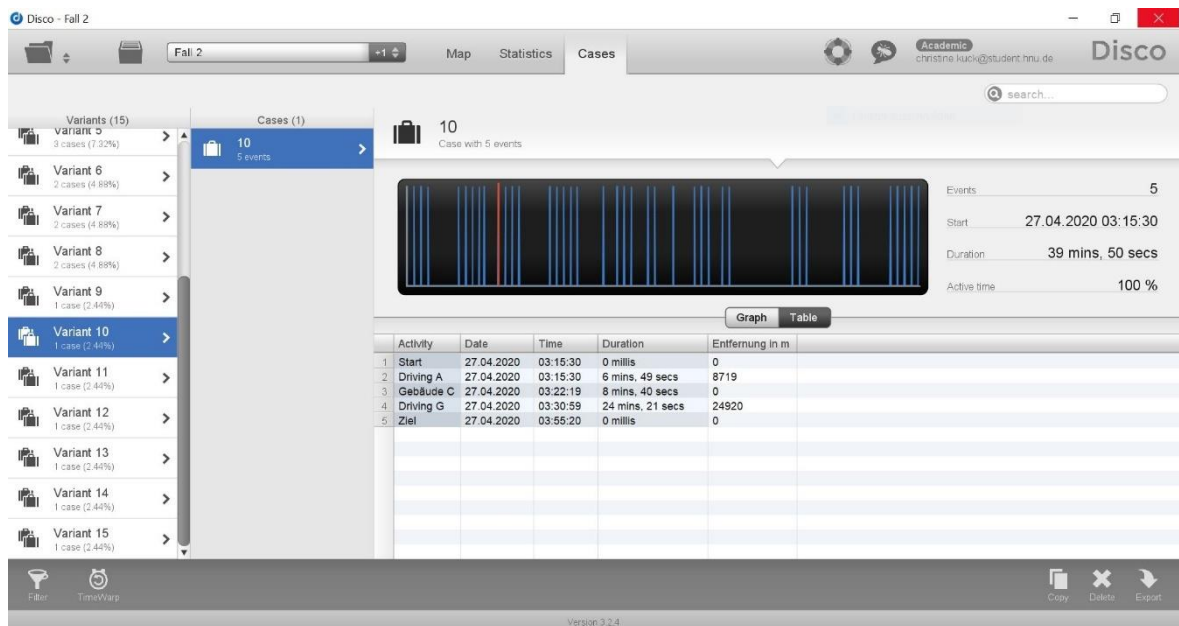


Abbildung 79: Auswertung Fall 2 - Prozessablauf Variante 10. Eigene Darstellung (2022)

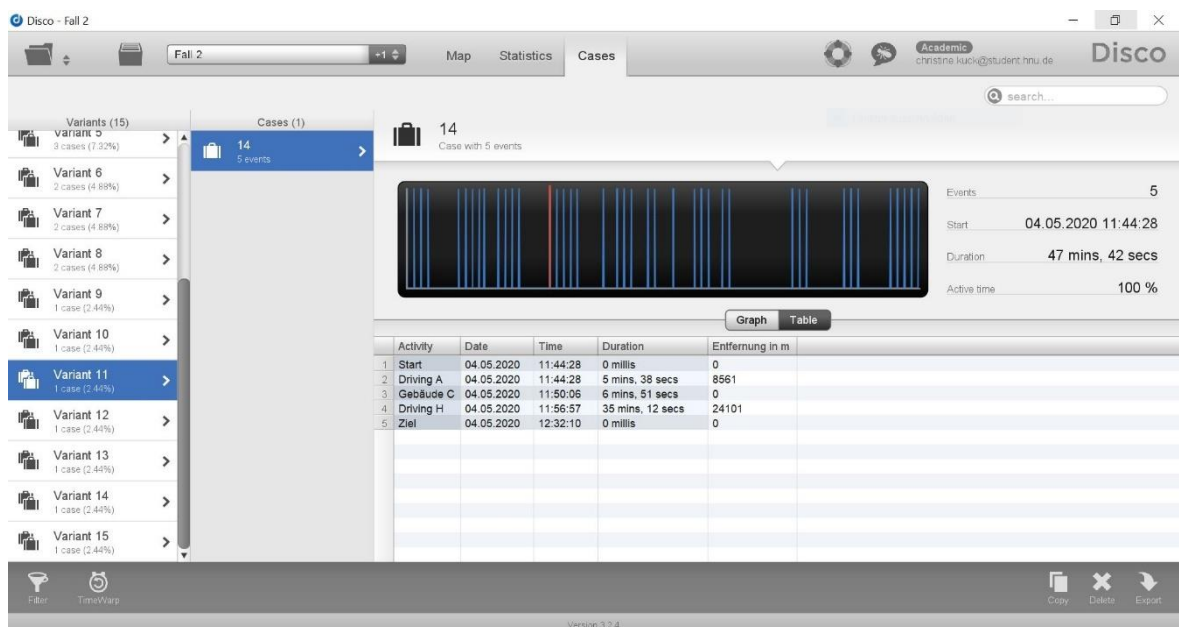


Abbildung 80: Auswertung Fall 2 - Prozessablauf Variante 11. Eigene Darstellung (2022)

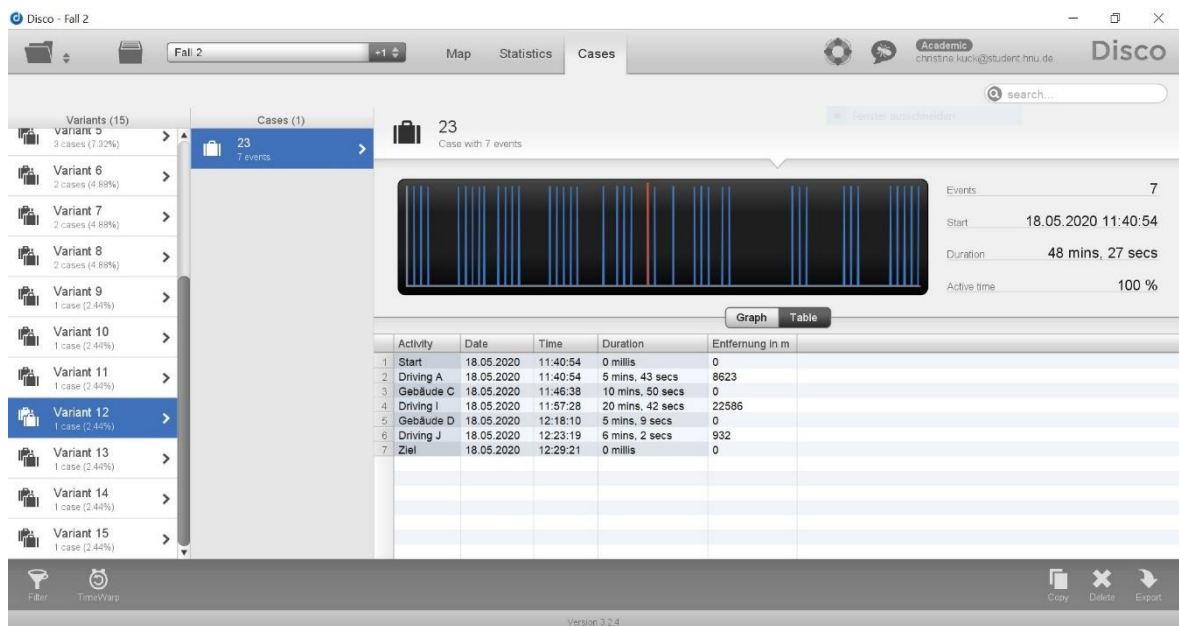


Abbildung 81: Auswertung Fall 2 - Prozessablauf Variante 12. Eigene Darstellung (2022)

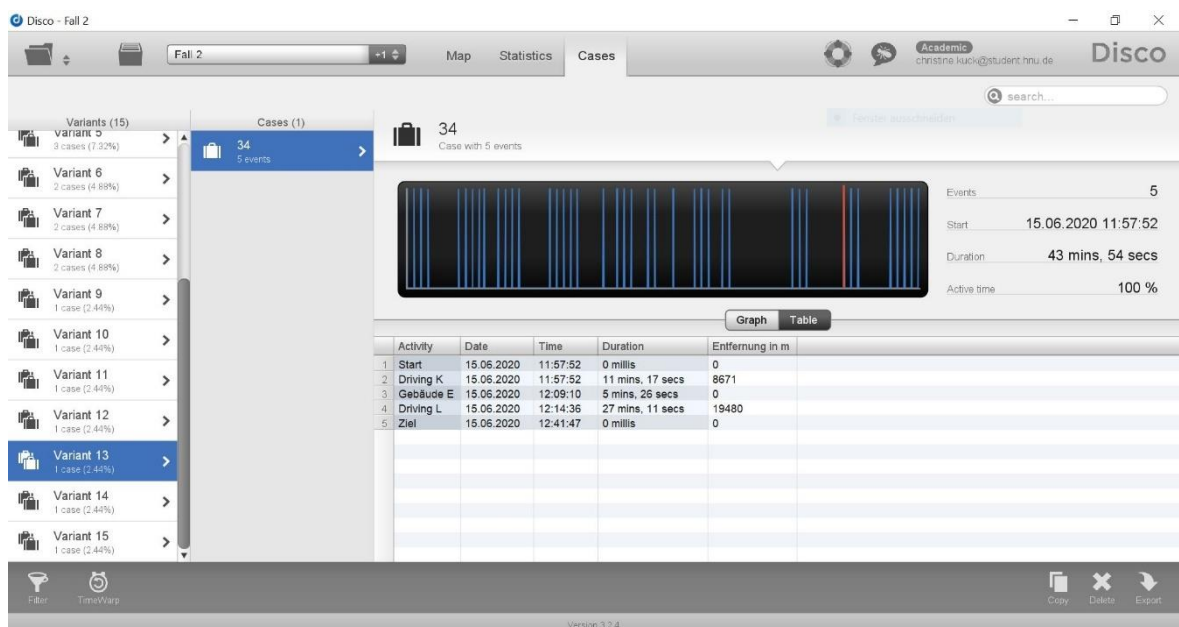


Abbildung 82: Auswertung Fall 2 - Prozessablauf Variante 13. Eigene Darstellung (2022)

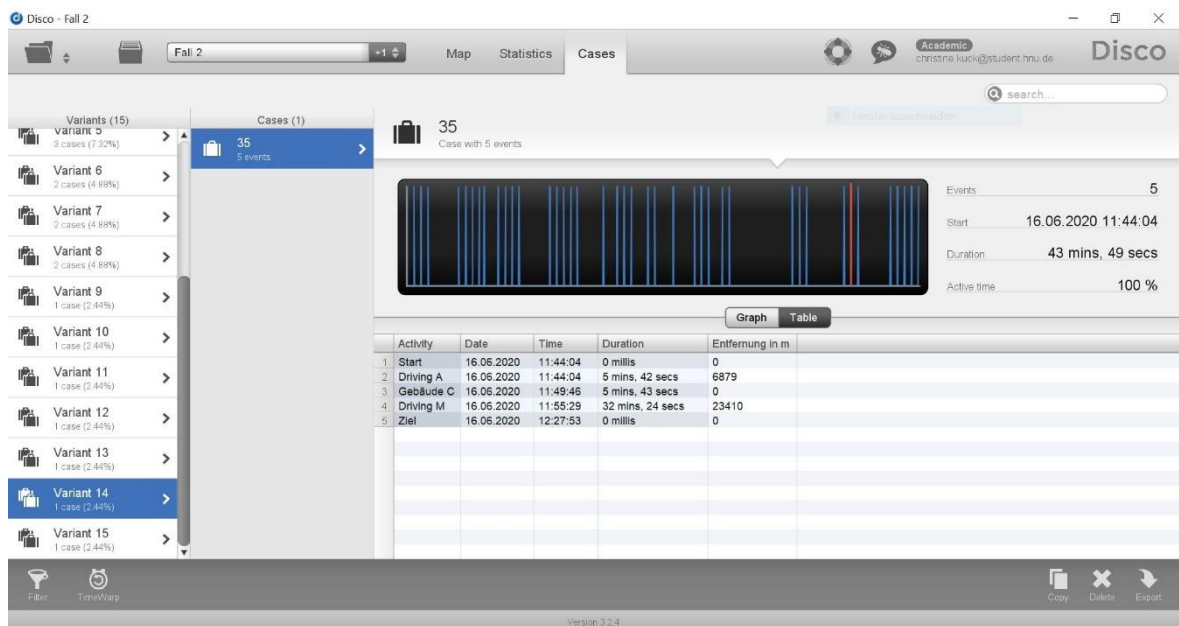


Abbildung 83: Auswertung Fall 2 - Prozessablauf Variante 14. Eigene Darstellung (2022)

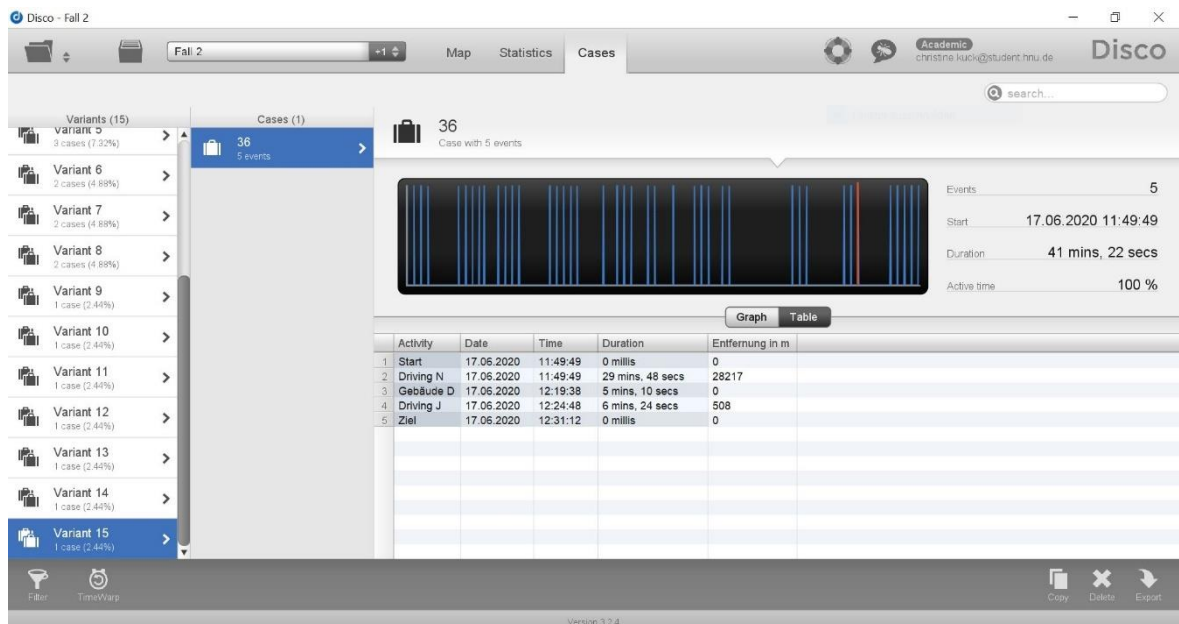


Abbildung 84: Auswertung Fall 2 - Prozessablauf Variante 15. Eigene Darstellung (2022)

Tabelle 23: Auswertung Fall 2 - Ergebnisse KPI V. Eigene Darstellung (2022)

Variante	V
Variante 1	0,917156863
Variante 2	1,20877193
Variante 3	0,98
Variante 4	1,027333333
Variante 5	1,046825397
Variante 6	1,180864198

<i>Variante 7</i>	0,860683761
<i>Variante 8</i>	1,002272727
<i>Variante 9</i>	0,953535354
<i>Variante 10</i>	1,171568627
<i>Variante 11</i>	1,1925
<i>Variante 12</i>	1,468181818
<i>Variante 13</i>	1,291176471
<i>Variante 14</i>	1,369270833
<i>Variante 15</i>	1,253535354

Tabelle 24: Auswertung Fall 2 - Ergebnisse KPIs G & D. Eigene Darstellung (2022)

Case	G (in km/h)	D
<i>1</i>	58,08165273	1,125692137
<i>2</i>	54,21282766	0,992801772
<i>3</i>	53,9123696	1,11461794
<i>4</i>	59,2116041	0,973421927
<i>5</i>	64,77391304	0,840531561
<i>6</i>	69,83306153	0,746954596
<i>7</i>	53,96425025	1,11517165
<i>8</i>	57,35974576	1,045404208
<i>9</i>	69	0,747508306
<i>10</i>	50,66962343	1,323366556
<i>11</i>	78,00702929	0,661683278
<i>12</i>	77,57382283	0,69379845
<i>13</i>	80,83115578	0,661129568
<i>14</i>	41,08427673	1,584717608
<i>15</i>	62,29465649	0,870431894
<i>16</i>	81,83979678	0,65393134
<i>17</i>	64,04099217	0,848283499
<i>18</i>	58,23710961	0,904208195
<i>19</i>	54,29047371	1,063676633
<i>20</i>	56,36365603	1,023809524
<i>21</i>	51,8414307	1,11461794
<i>22</i>	62,81742243	0,928017719

23	39,80309598	1,609634551
24	61,18953488	0,952380952
25	64,81243094	0,801771872
26	61,66504918	0,84440753
27	55,76346154	1,03654485
28	64,01390164	0,84440753
29	66,83914286	0,775193798
30	59,38152794	0,971207087
31	51,52621741	1,125692137
32	58,12864322	0,991694352
33	59,80206186	0,966777409
34	38,47517084	1,458471761
35	41,47599848	1,455703212
36	41,66398066	1,374307863
37	55,49650053	1,044296788
38	60,27965217	0,955149502
39	51,79268775	1,120708749
40	59,16770893	0,973975637
41	78,25721435	0,663898117

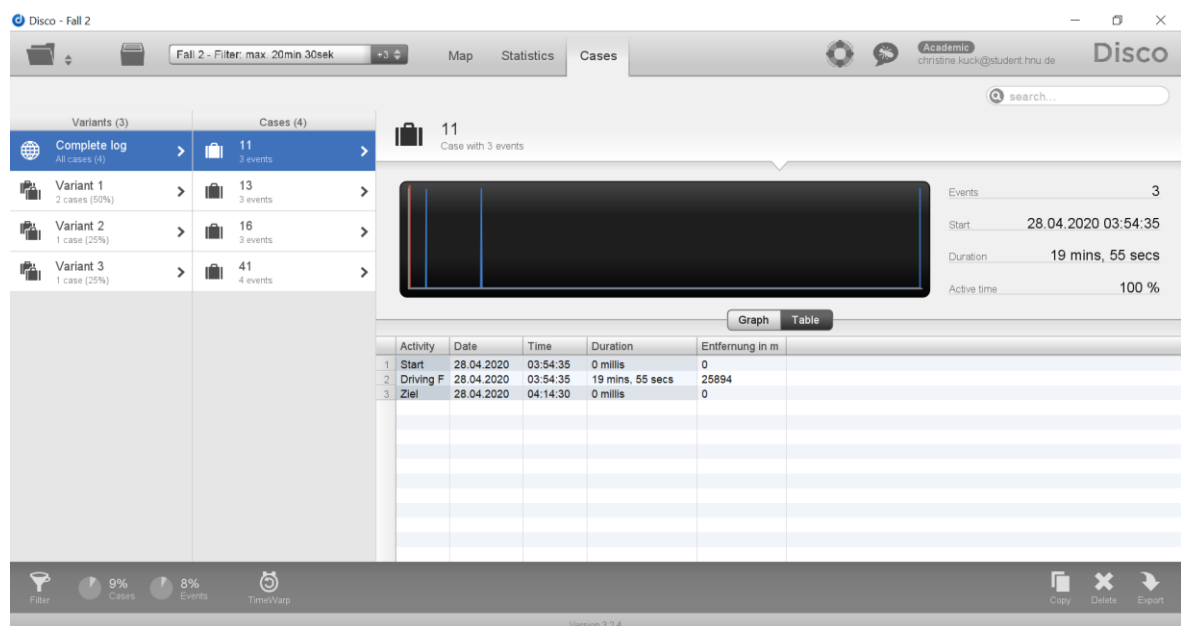


Abbildung 85: Auswertung Fall 2 - Filter: Kürzeste Cases. Eigene Darstellung (2022)

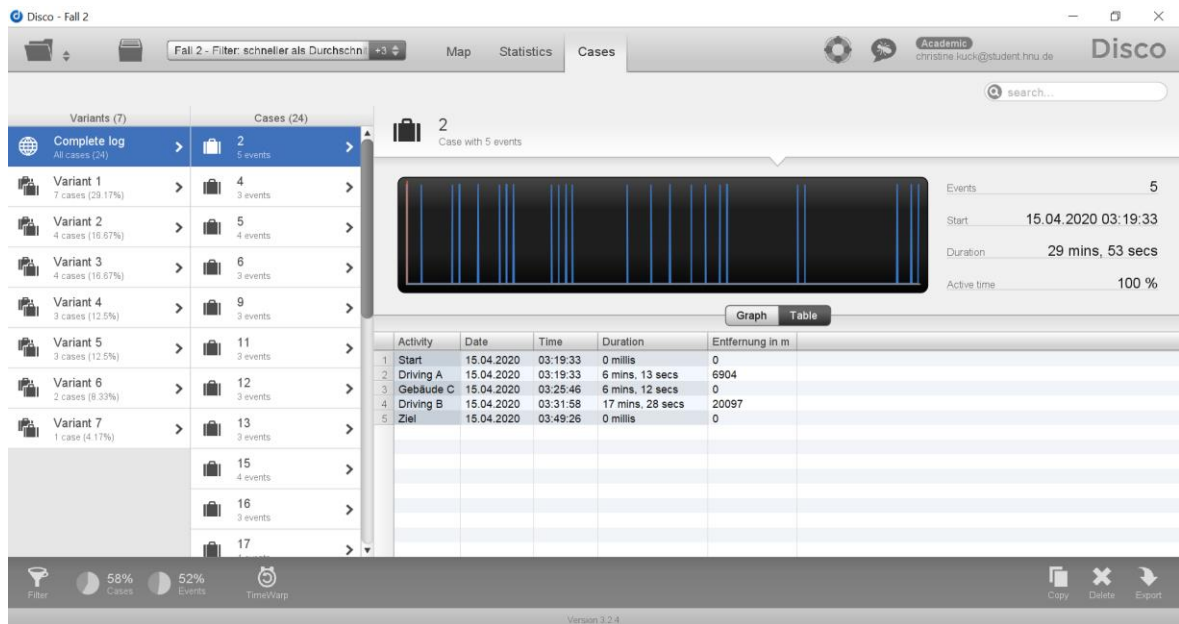


Abbildung 86: Auswertung Fall 2 - Filter: Kürzer als Durchschnitt. Eigene Darstellung (2022)

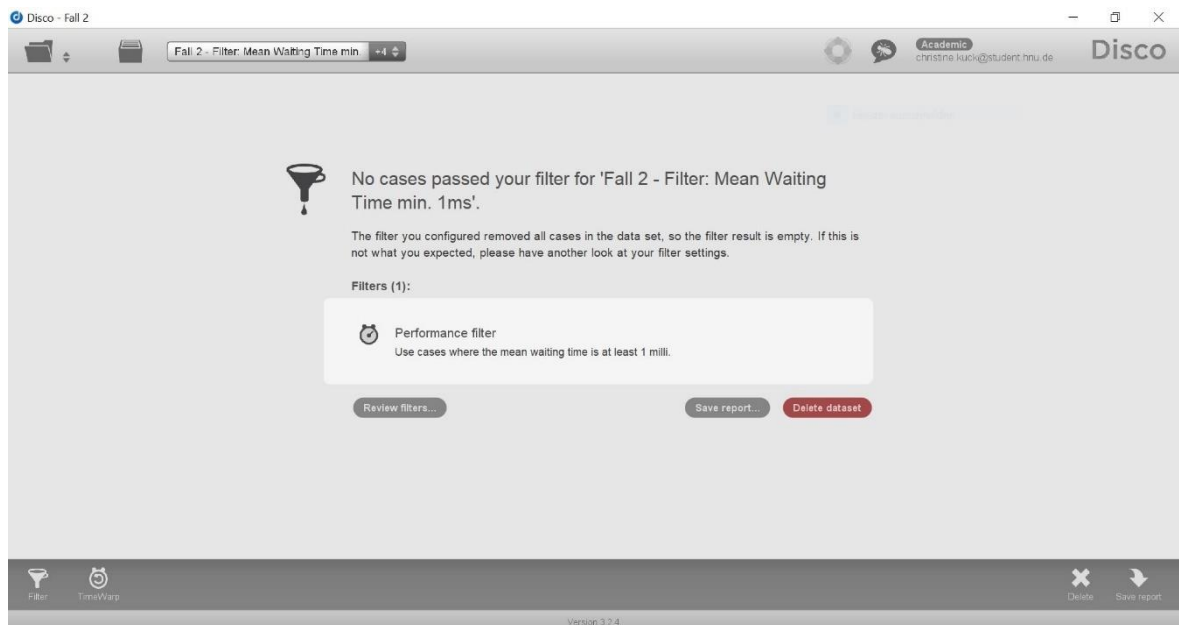


Abbildung 87: Auswertung Fall 2 - Filter: Cases mit Wartezeit. Eigene Darstellung (2022)

Events	110
Cases	20
Activities	22
Median case duration	100 mins
Mean case duration	111.6 mins
Start	19.03.2021 15:23:27
End	23.08.2021 13:50:38

Abbildung 88: Auswertung Fall 3 - Allgemeine Informationen. Eigene Darstellung (2022)

Variants (10)	
 Complete log	>
All cases (20)	
 Variant 1	>
9 cases (45%)	
 Variant 2	>
2 cases (10%)	
 Variant 3	>
2 cases (10%)	
 Variant 4	>
1 case (5%)	
 Variant 5	>
1 case (5%)	
 Variant 6	>
1 case (5%)	
 Variant 7	>
1 case (5%)	
 Variant 8	>
1 case (5%)	
 Variant 9	>
1 case (5%)	
 Variant 10	>
1 case (5%)	

Abbildung 89: Auswertung Fall 3 - Varianten in Prozent. Eigene Darstellung (2022)

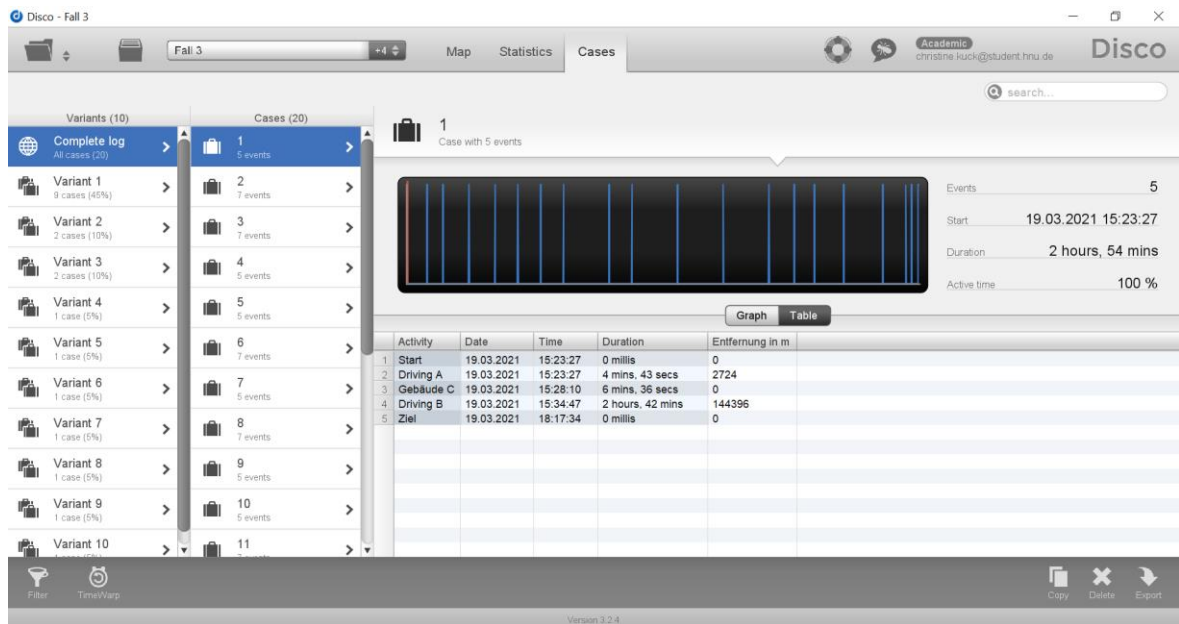


Abbildung 90: Auswertung Fall 3 - Dauer und Entfernung der Cases. Eigene Darstellung (2022)

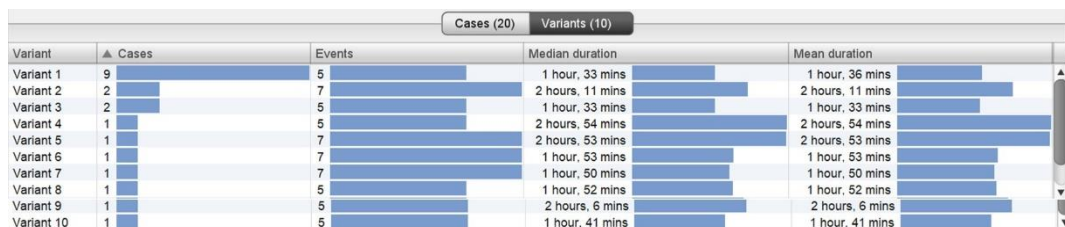


Abbildung 91: Auswertung Fall 3 - Tabelle: 'Variants'. Eigene Darstellung (2022)

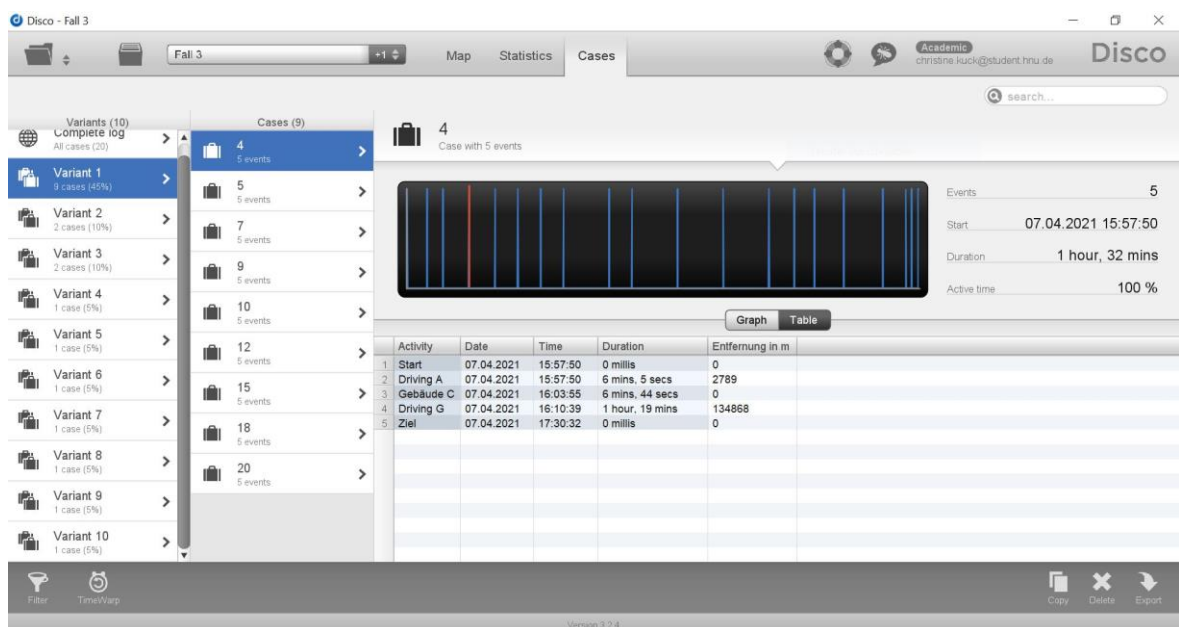


Abbildung 92: Auswertung Fall 3 - Prozessablauf Variante 1. Eigene Darstellung (2022)

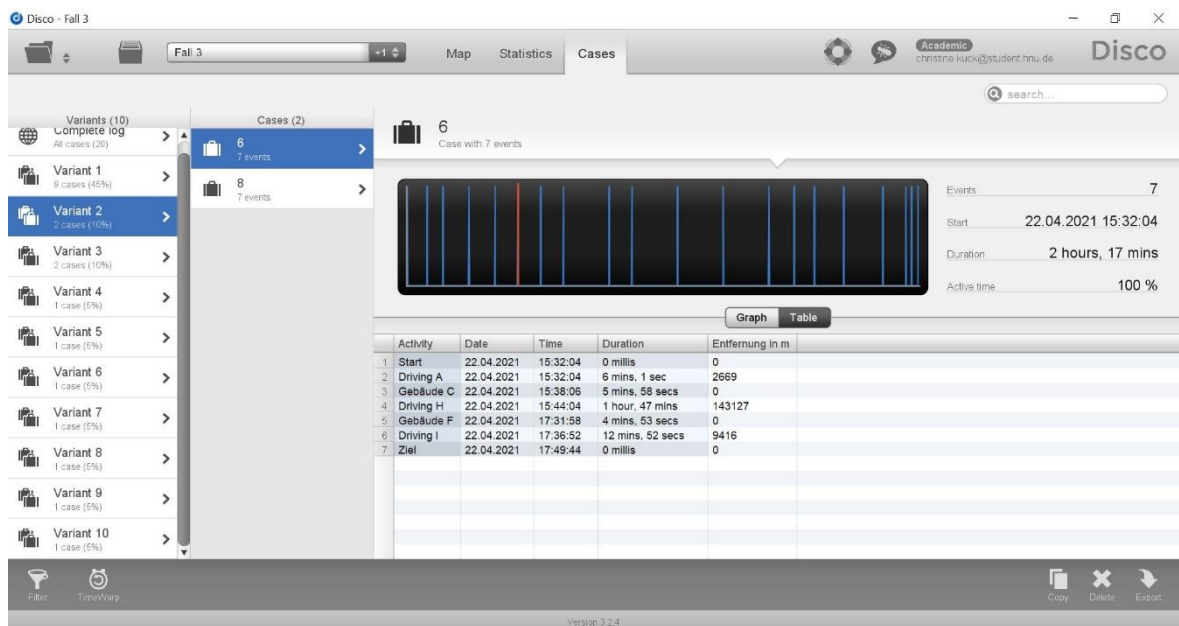


Abbildung 93: Auswertung Fall 3 - Prozessablauf Variante 2. Eigene Darstellung (2022)

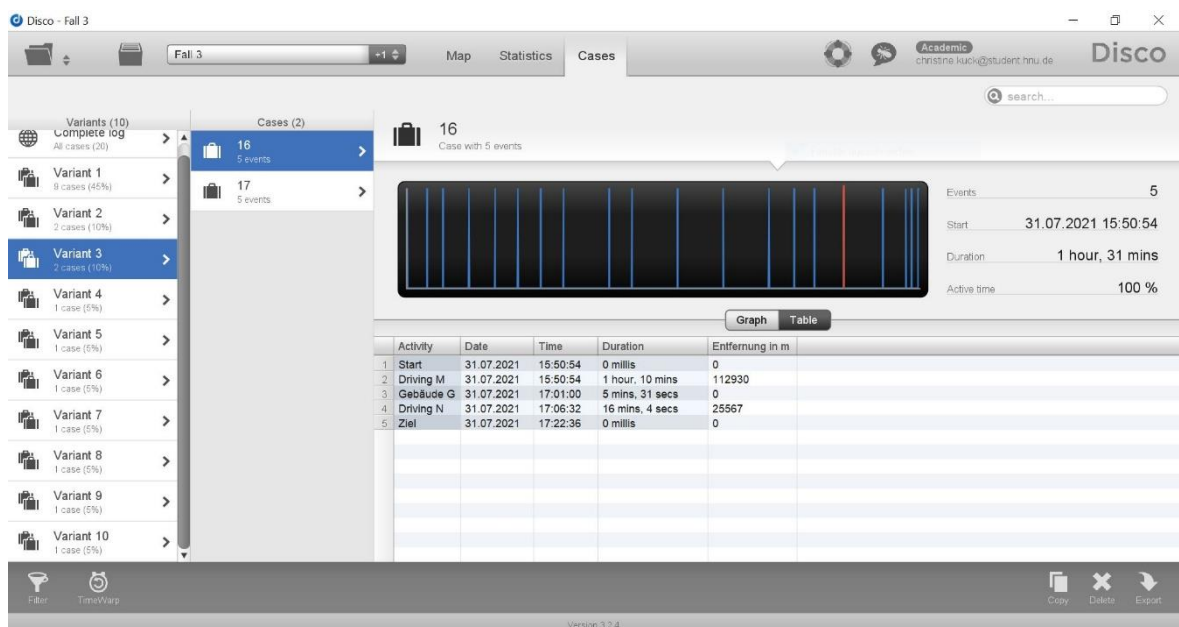


Abbildung 94: Auswertung Fall 3 - Prozessablauf Variante 3. Eigene Darstellung (2022)

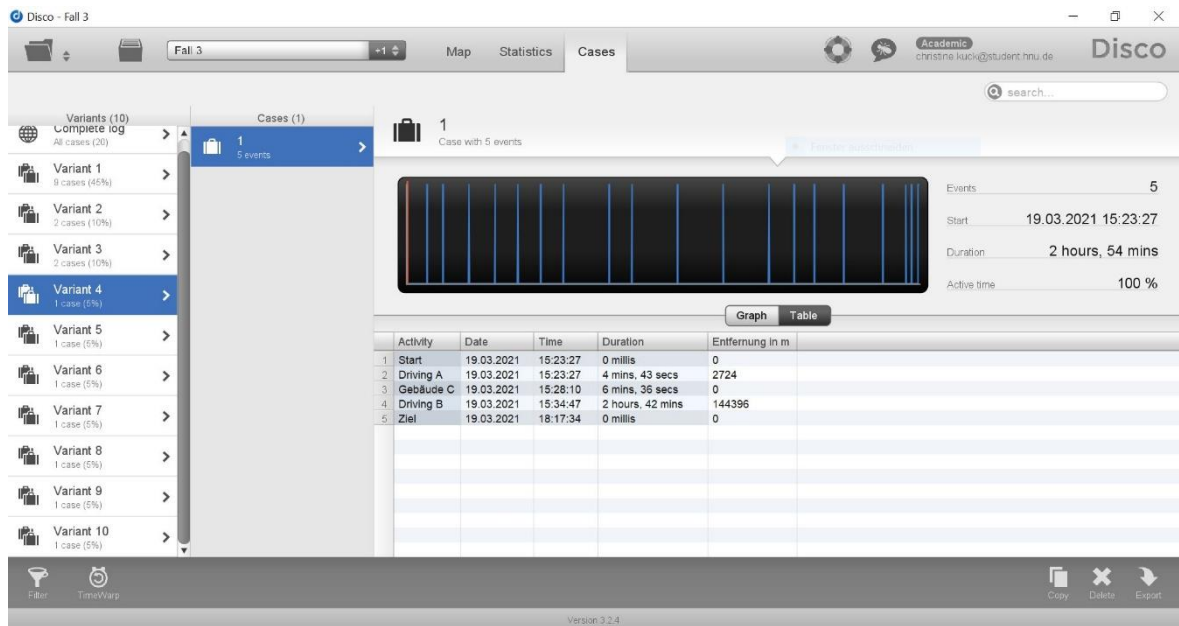


Abbildung 95: Auswertung Fall 3 - Prozessablauf Variante 4. Eigene Darstellung (2022)

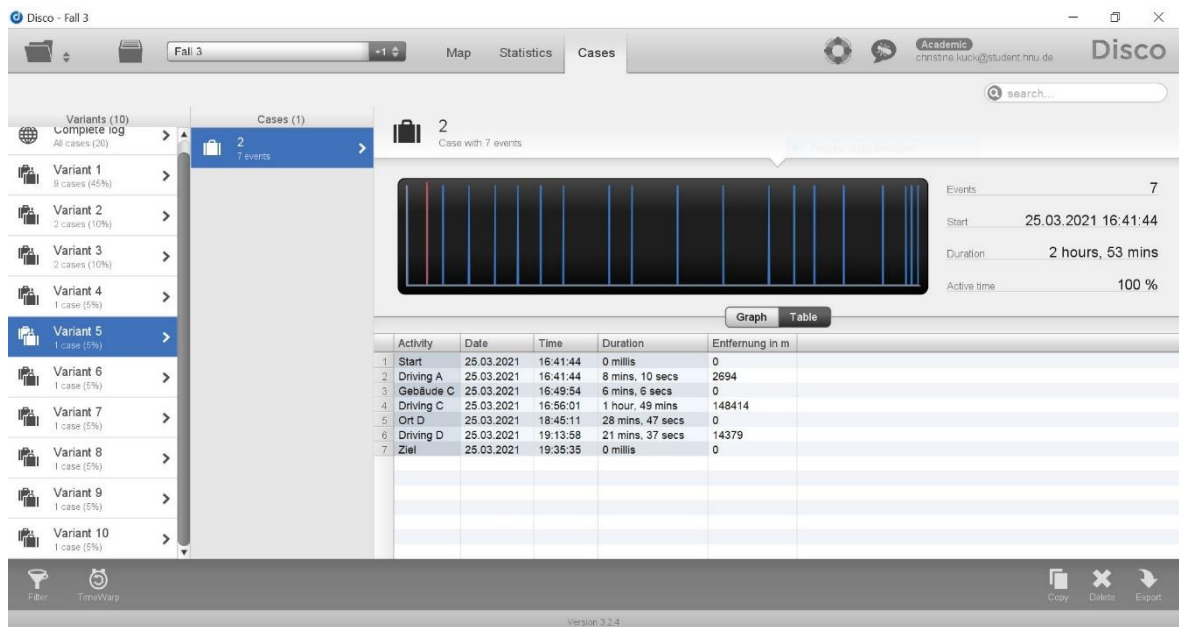


Abbildung 96: Auswertung Fall 3 - Prozessablauf Variante 5. Eigene Darstellung (2022)

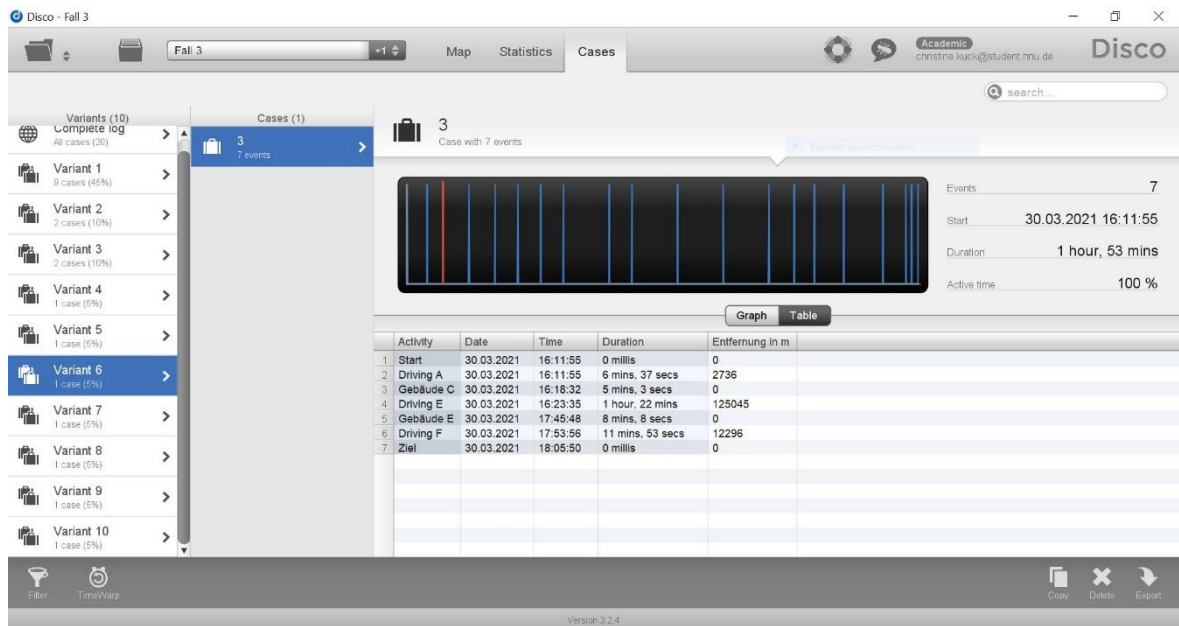


Abbildung 97: Auswertung Fall 3 - Prozessablauf Variante 6. Eigene Darstellung (2022)

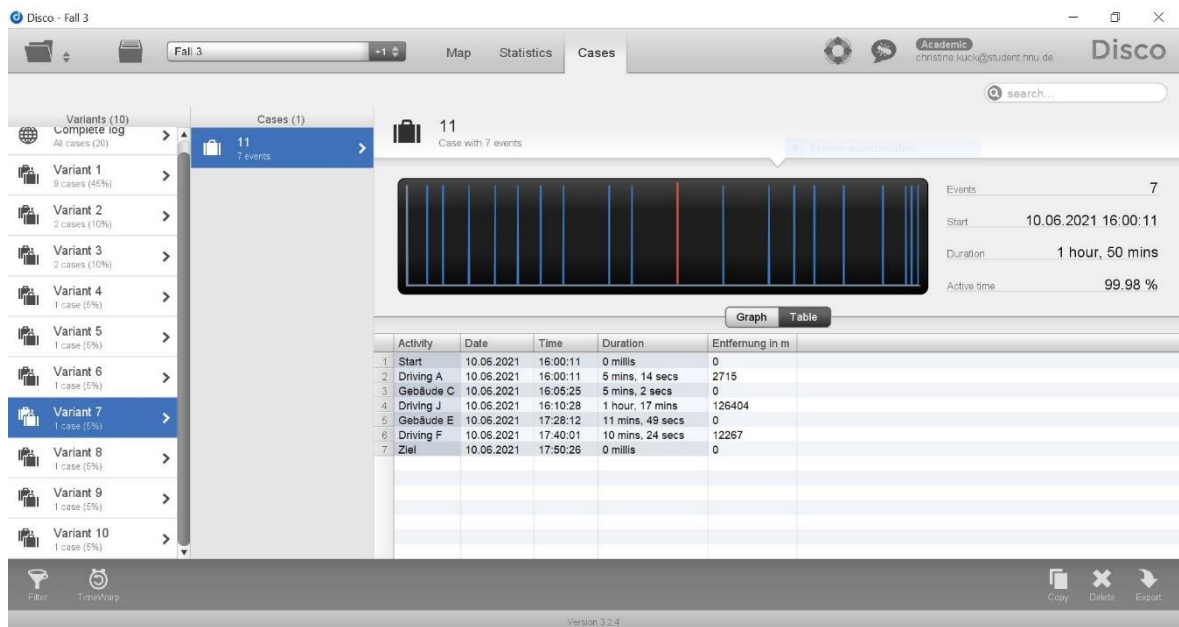


Abbildung 98: Auswertung Fall 3 - Prozessablauf Variante 7. Eigene Darstellung (2022)

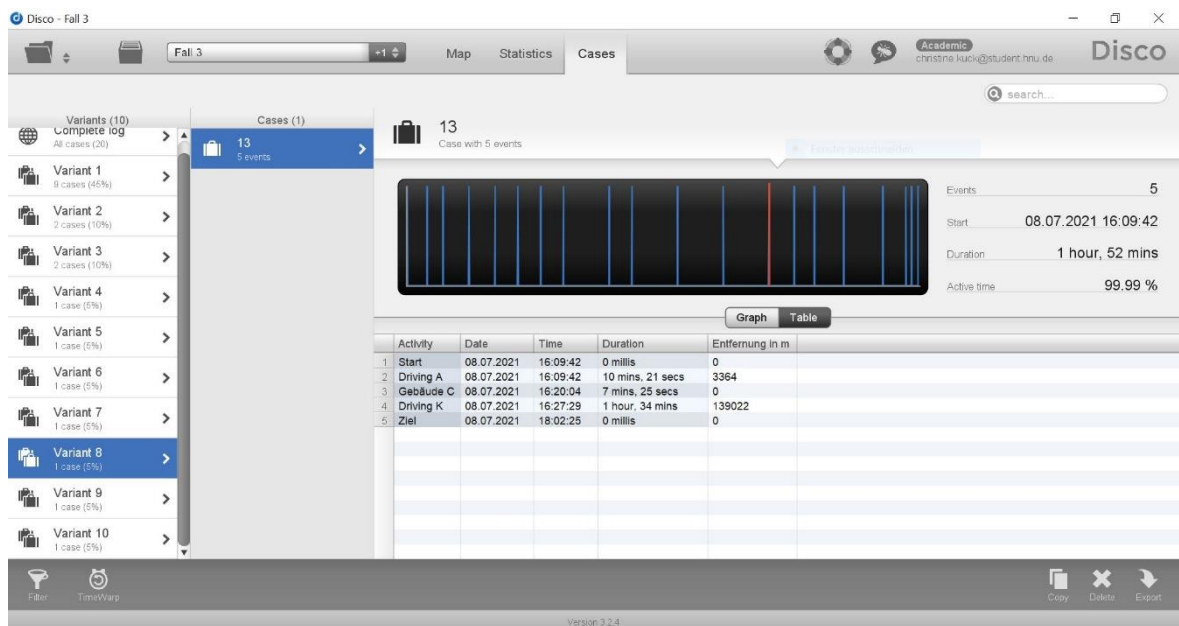


Abbildung 99: Auswertung Fall 3 - Prozessablauf Variante 8. Eigene Darstellung (2022)

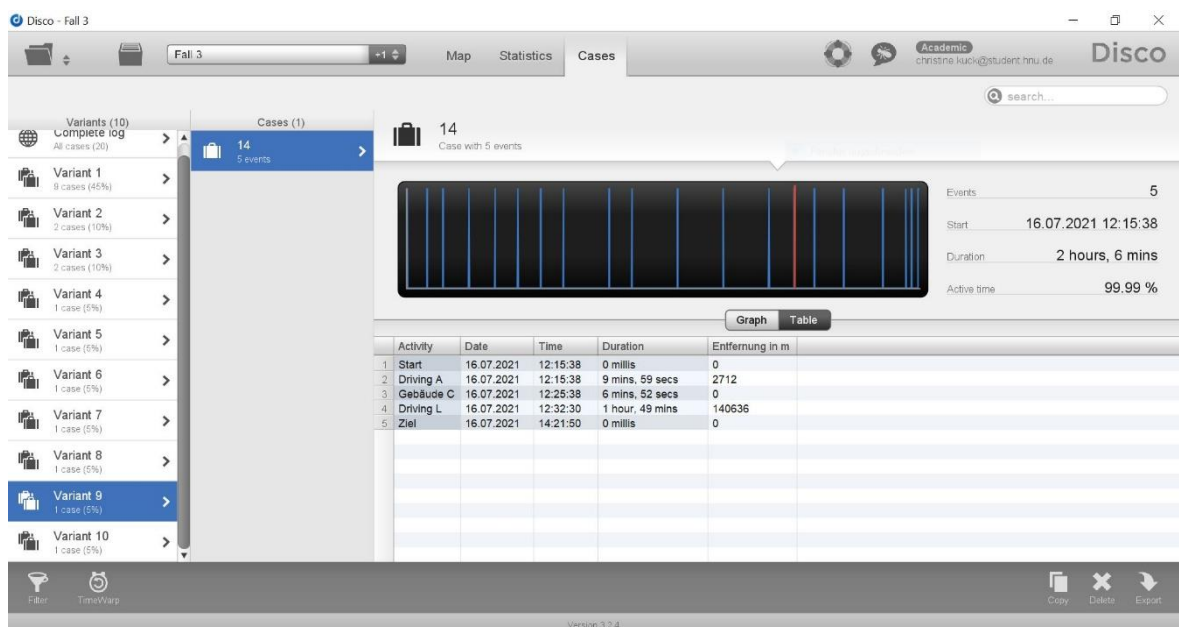


Abbildung 100: Auswertung Fall 3 - Prozessablauf Variante 9. Eigene Darstellung (2022)

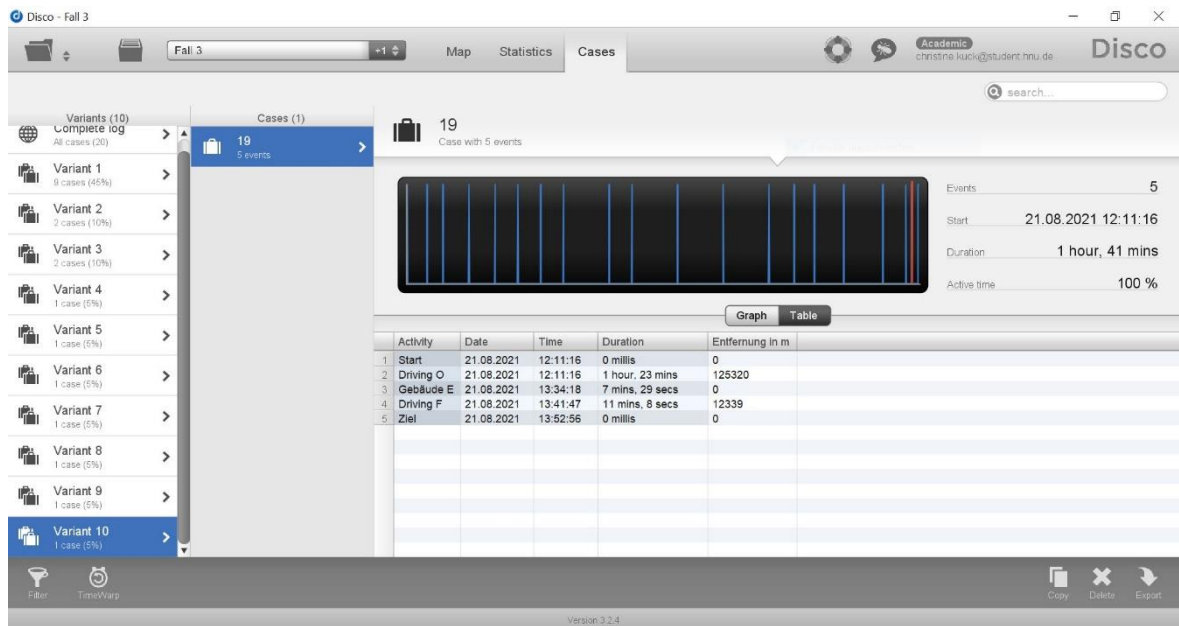


Abbildung 101: Auswertung Fall 3 - Prozessablauf Variante 10. Eigene Darstellung (2022)

Tabelle 25: Auswertung Fall 3 - Ergebnisse KPI V. Eigene Darstellung (2022)

Variante	V
Variante 1	1,090909091
Variante 2	1,235849057
Variante 3	1,033333333
Variante 4	1,426229508
Variante 5	1,330769231
Variante 6	1,141414141
Variante 7	1,111111111
Variante 8	1,142857143
Variante 9	1,235294118
Variante 10	0,990196078

Tabelle 26: Auswertung Fall 3 - Ergebnisse KPIs G & D. Eigene Darstellung (2022)

Case	G (in km/h)	D
1	50,73103448	1,559139785
2	57,39433526	1,550179211
3	74,37716814	1,012544803
4	89,77630435	0,82437276
5	84,1855102	0,878136201
6	67,97605839	1,227598566

7	90,68835165	0,815412186
8	74,80741935	1,111111111
9	89,60413043	0,82437276
10	88,72451613	0,833333333
11	77,11963636	0,985663082
12	88,7083871	0,833333333
13	76,27821429	1,003584229
14	68,26095238	1,129032258
15	86,87431579	0,85125448
16	91,3167033	0,815412186
17	87,61515789	0,85125448
18	86,98357895	0,85125448
19	81,77762376	0,905017921
20	71,13051724	1,039426523

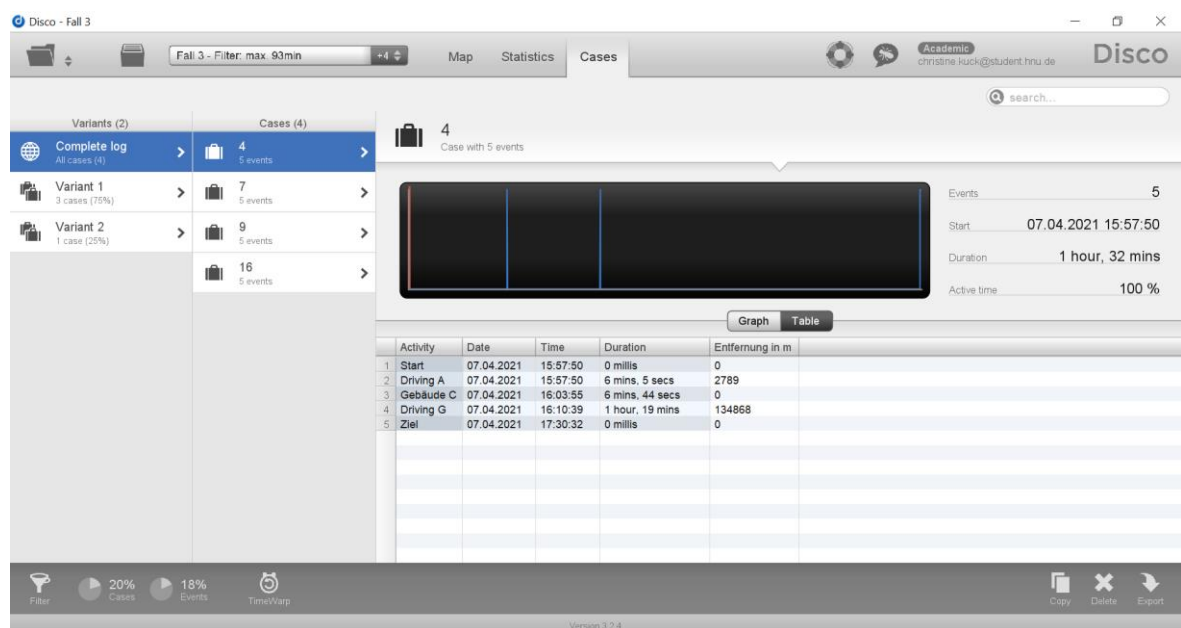


Abbildung 102: Auswertung Fall 3 - Filter: Kürzeste Cases. Eigene Darstellung (2022)

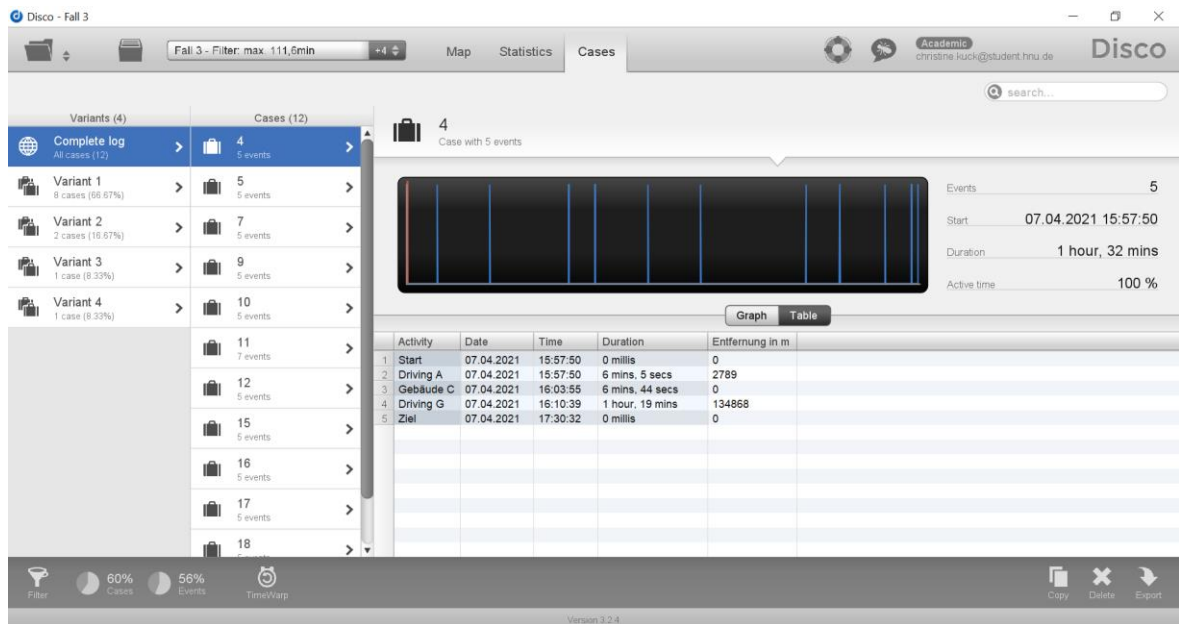


Abbildung 103: Auswertung Fall 3 - Filter: Kürzer als Durchschnitt. Eigene Darstellung (2022)

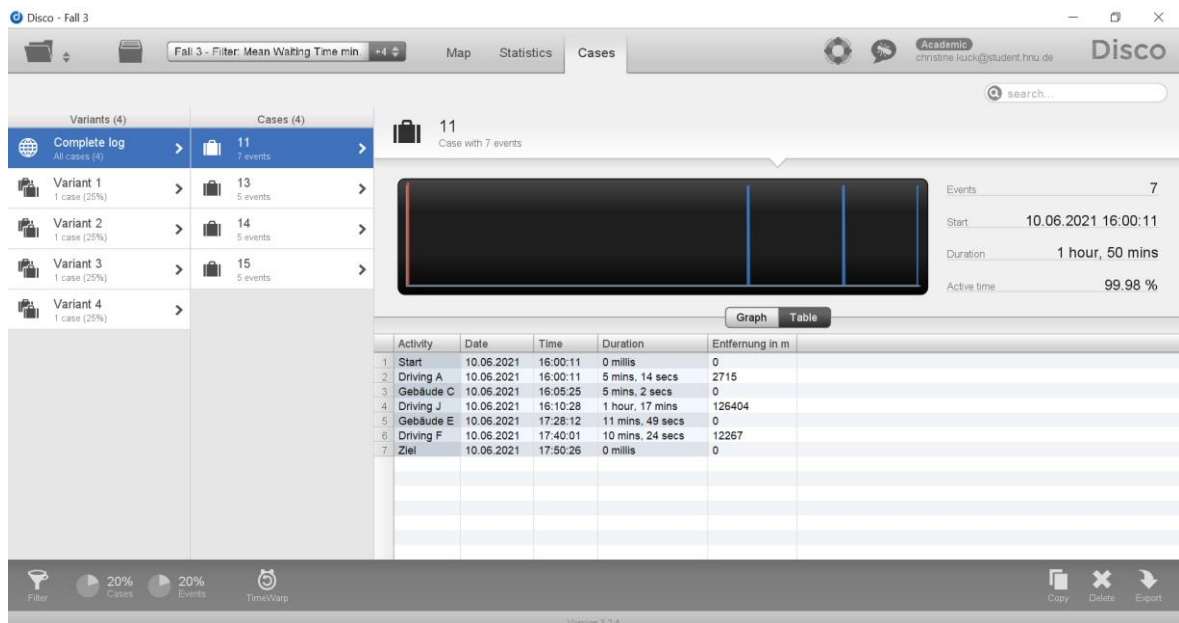


Abbildung 104: Auswertung Fall 3 - Filter: Cases mit Wartezeiten. Eigene Darstellung (2022)

Events	195
Cases	32
Activities	17
Median case duration	51.9 mins
Mean case duration	61.2 mins
Start	08.02.2021 05:31:25
End	16.04.2021 04:59:23

Abbildung 105: Auswertung Fall 4 - allgemeine Informationen. Eigene Darstellung (2022)

Variants (8)	
 Complete log	>
All cases (32)	
 Variant 1	>
22 cases (68.75%)	
 Variant 2	>
3 cases (9.38%)	
 Variant 3	>
2 cases (6.25%)	
 Variant 4	>
1 case (3.12%)	
 Variant 5	>
1 case (3.12%)	
 Variant 6	>
1 case (3.12%)	
 Variant 7	>
1 case (3.12%)	
 Variant 8	>
1 case (3.12%)	

Abbildung 106: Auswertung Fall 4 - Varianten in Prozent. Eigene Darstellung (2022)

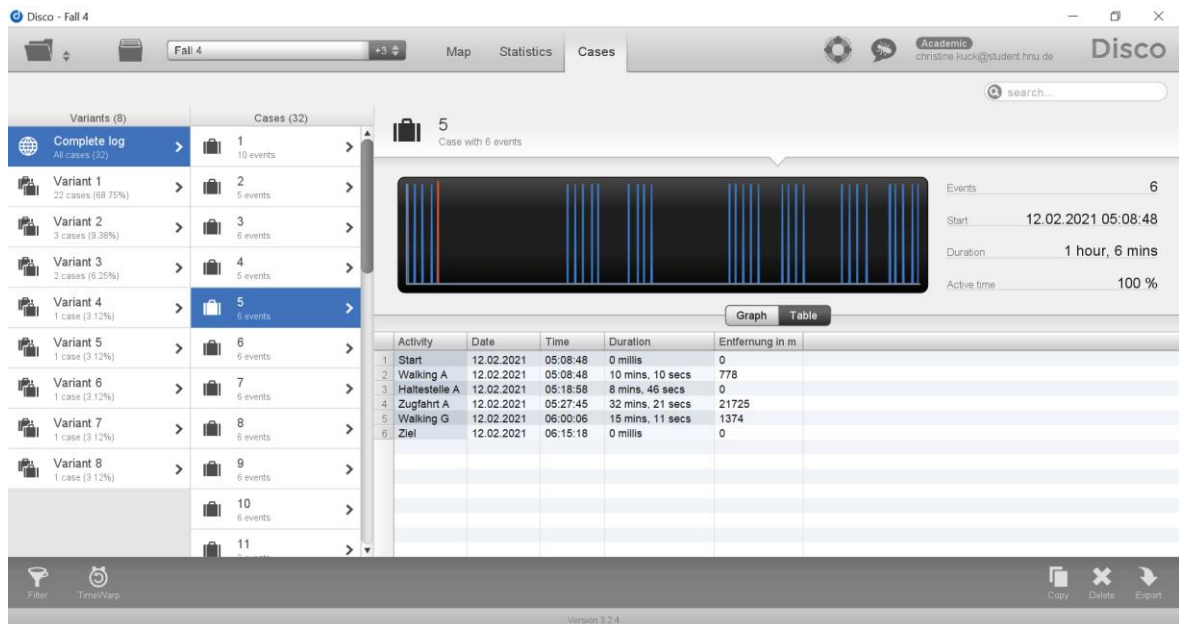


Abbildung 107: Auswertung Fall 4 - Dauer und Entfernung der Cases. Eigene Darstellung (2022)

Cases (32)		Variants (8)	
Variant	Cases	Events	Mean duration
Variant 1	22	6	51 mins, 48 secs
Variant 2	3	6	59 mins
Variant 3	2	7	49 mins, 46 secs
Variant 4	1	10	2 hours, 57 mins
Variant 5	1	5	43 mins, 37 secs
Variant 6	1	6	47 mins, 6 secs
Variant 7	1	5	54 mins
Variant 8	1	5	53 mins, 41 secs

Abbildung 108: Auswertung Fall 4 - Tabelle: „Variants“. Eigene Darstellung (2022)

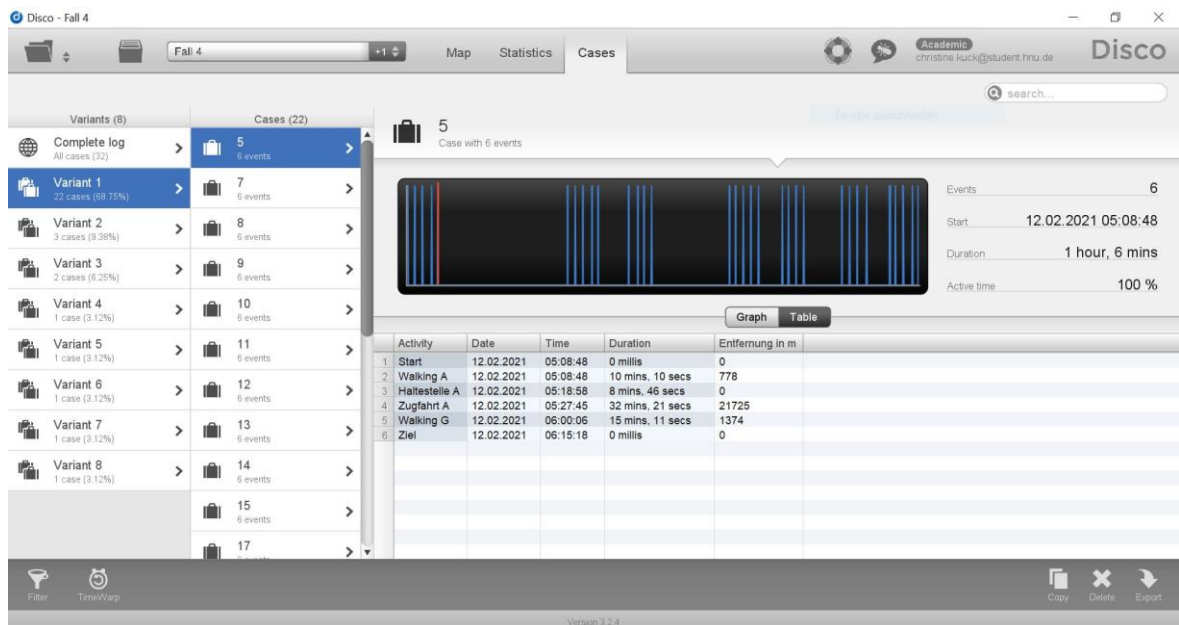


Abbildung 109: Auswertung Fall 4 - Prozessablauf Variante 1. Eigene Darstellung (2022)

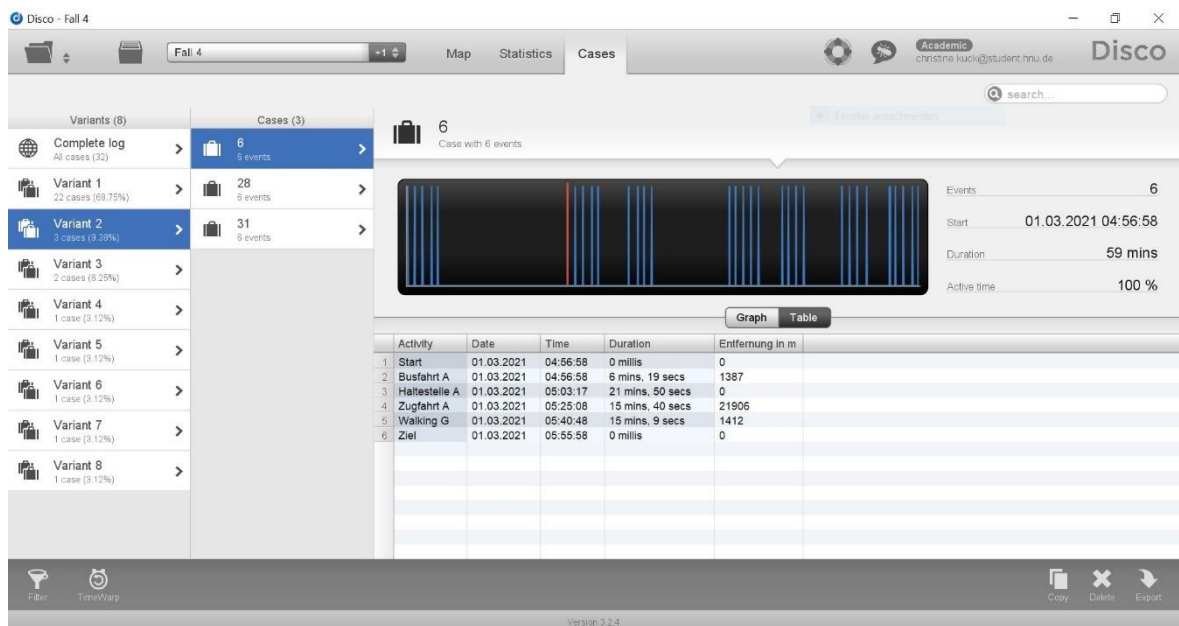


Abbildung 110: Auswertung Fall 4 - Prozessablauf Variante 2. Eigene Darstellung (2022)

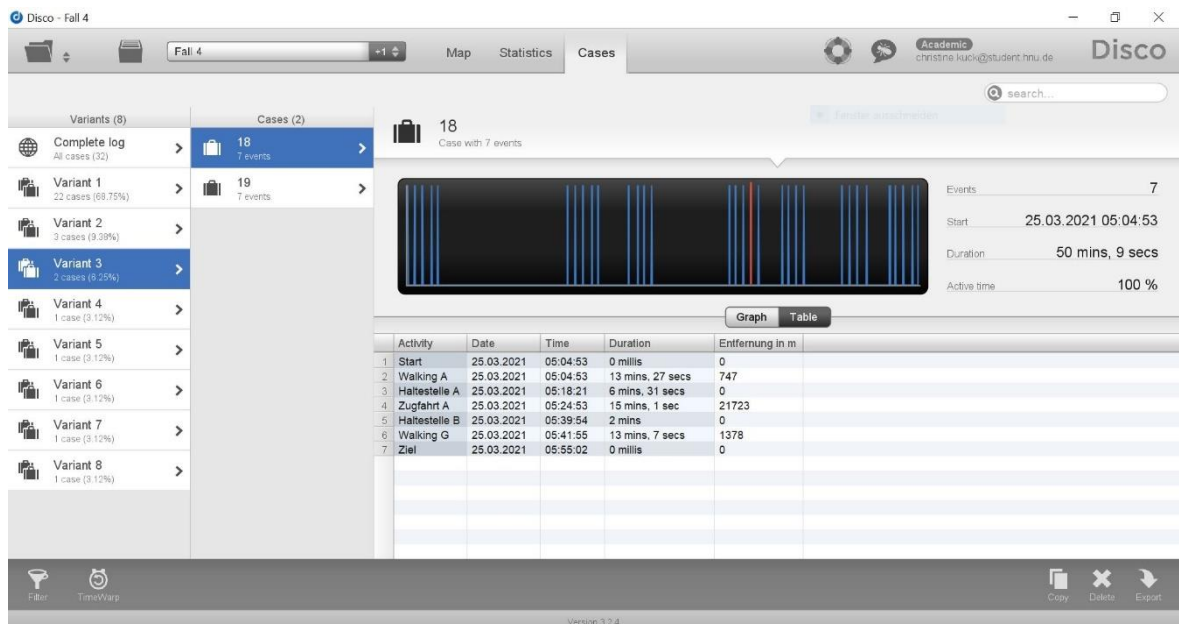


Abbildung 111: Auswertung Fall 4 - Prozessablauf Variante 3. Eigene Darstellung (2022)

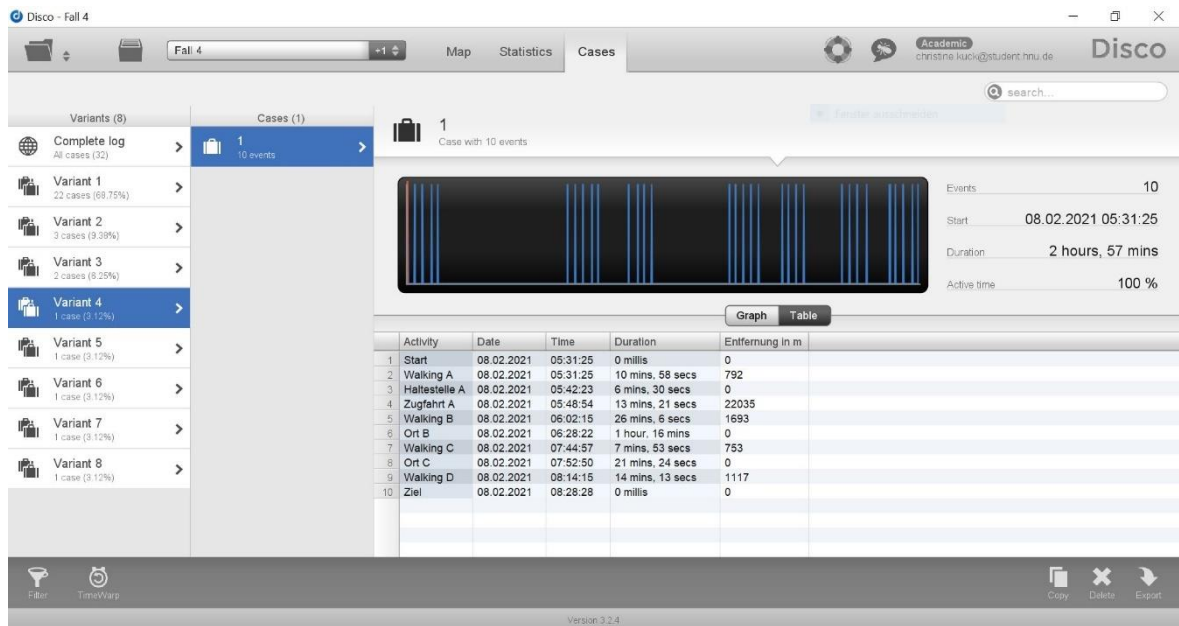


Abbildung 112: Auswertung Fall 4 - Prozessablauf Variante 4. Eigene Darstellung (2022)

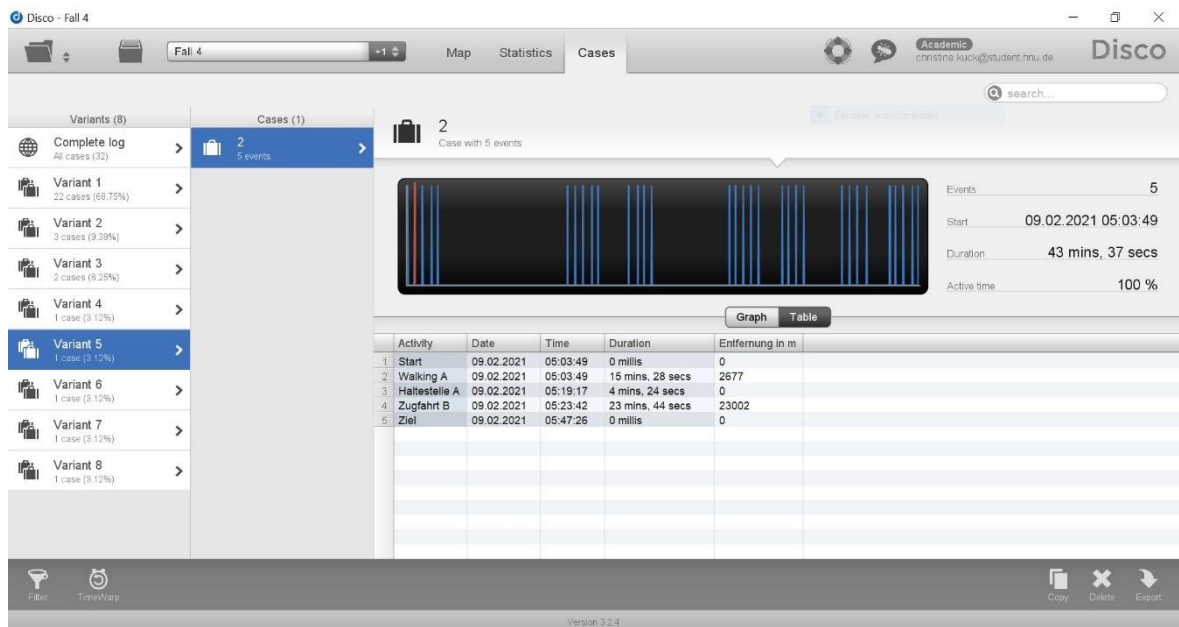


Abbildung 113: Auswertung Fall 4 - Prozessablauf Variante 5. Eigene Darstellung (2022)

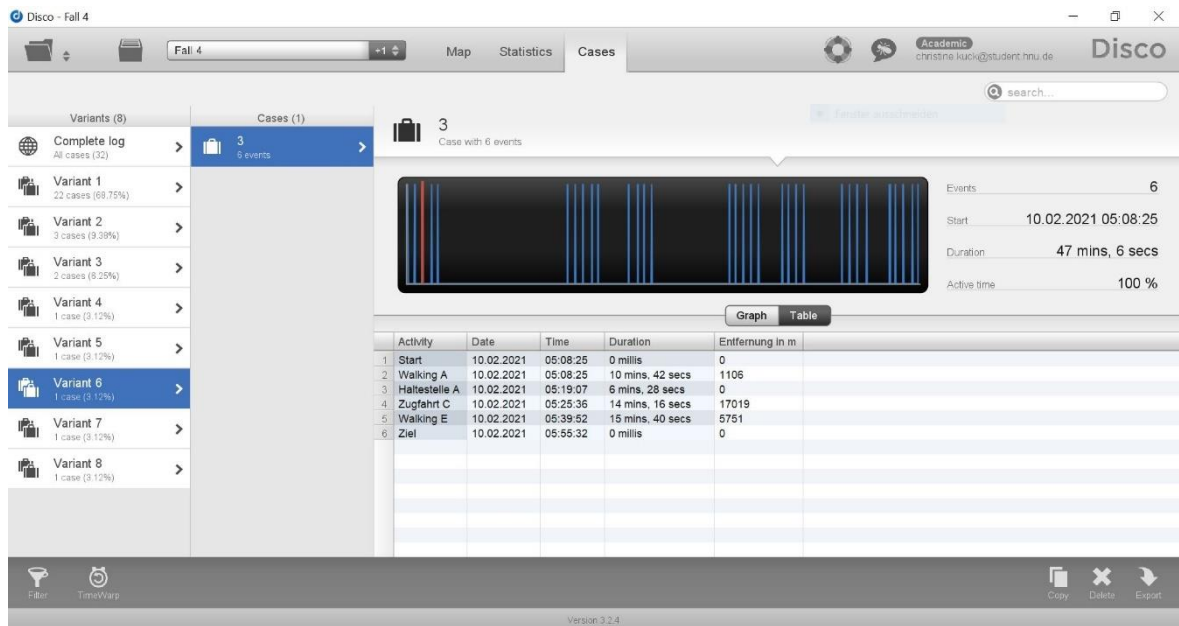


Abbildung 114: Auswertung Fall 4 - Prozessablauf Variante 6. Eigene Darstellung (2022)

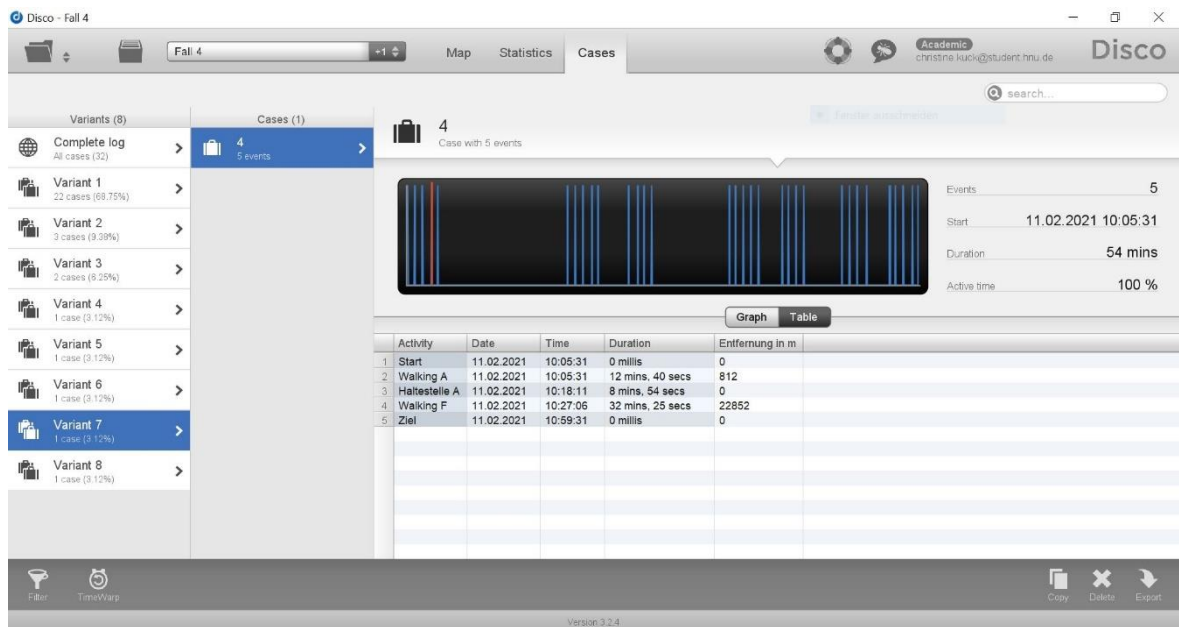


Abbildung 115: Auswertung Fall 4 - Prozessablauf Variante 7. Eigene Darstellung (2022)

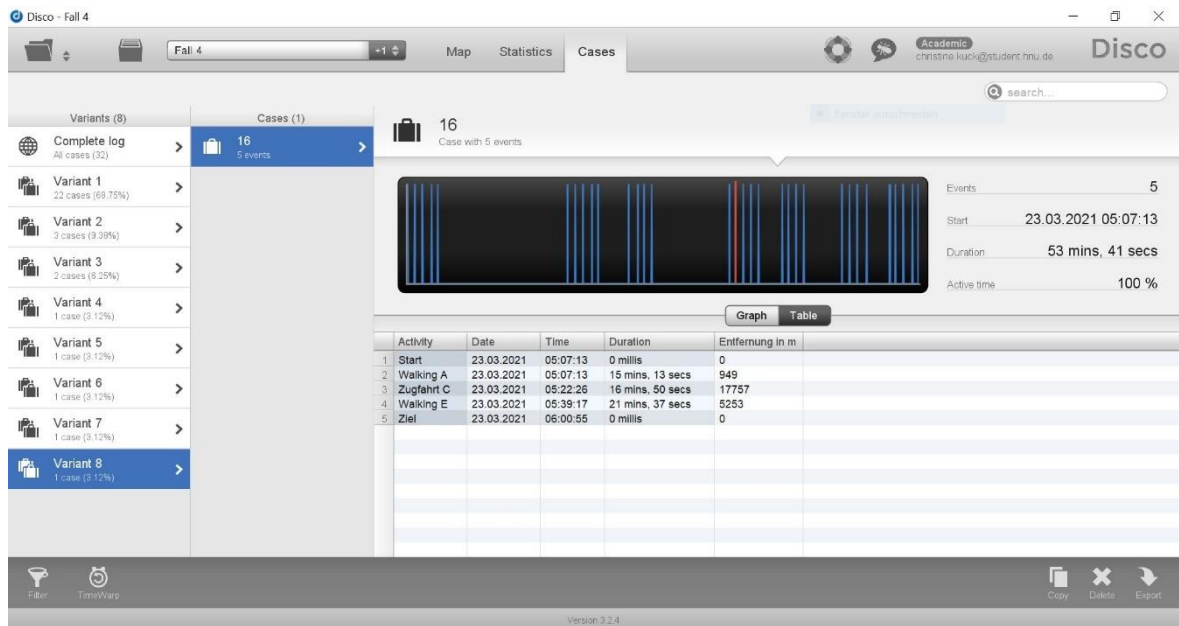


Abbildung 116: Auswertung Fall 4 - Prozessablauf Variante 8. Eigene Darstellung (2022)

Tabelle 27: Auswertung Fall 4 - Ergebnisse KPI V. Eigene Darstellung (2022)

Variante	V
Variante 1	1,32
Variante 2	1,255319149
Variante 3	1,003
Variante 4	2,458333333
Variante 5	0,908680556
Variante 6	0,428181818
Variante 7	0,165137615
Variante 8	0,488030303

Tabelle 28: Auswertung Fall 4 - Ergebnisse KPIs G & D. Eigene Darstellung (2022)

Case	G (in km/h)	D
1	8,945762712	2,892156863
2	35,32457012	0,712690632
3	30,41528662	0,769607843
4	26,29333333	0,882352941
5	21,70636364	1,078431373
6	25,12372881	0,964052288
7	30,38893281	0,826797386

8	29,35261538	0,796568627
9	25,44198993	0,919662309
10	25,02592275	0,951797386
11	25,70056937	0,908769063
12	26,70083256	0,883169935
13	29,42353723	0,819172113
14	30,97459283	0,75245098
15	25,92109422	0,895969499
16	26,77814343	0,877178649
17	29,63600137	0,792755991
18	28,53200399	0,819444444
19	28,85953426	0,806917211
20	27,78227848	0,839052288
21	29,02968435	0,811002179
22	29,11686747	0,81372549
23	28,06124156	0,84667756
24	28,15842733	0,845043573
25	30,43924051	0,774509804
26	26,03061728	0,992647059
27	26,21903314	0,895697168
28	7,15862069	3,316993464
29	28,20926641	0,846405229
30	25,67390791	0,922657952
31	28,04344473	0,847494553
32	29,61748166	0,891067538

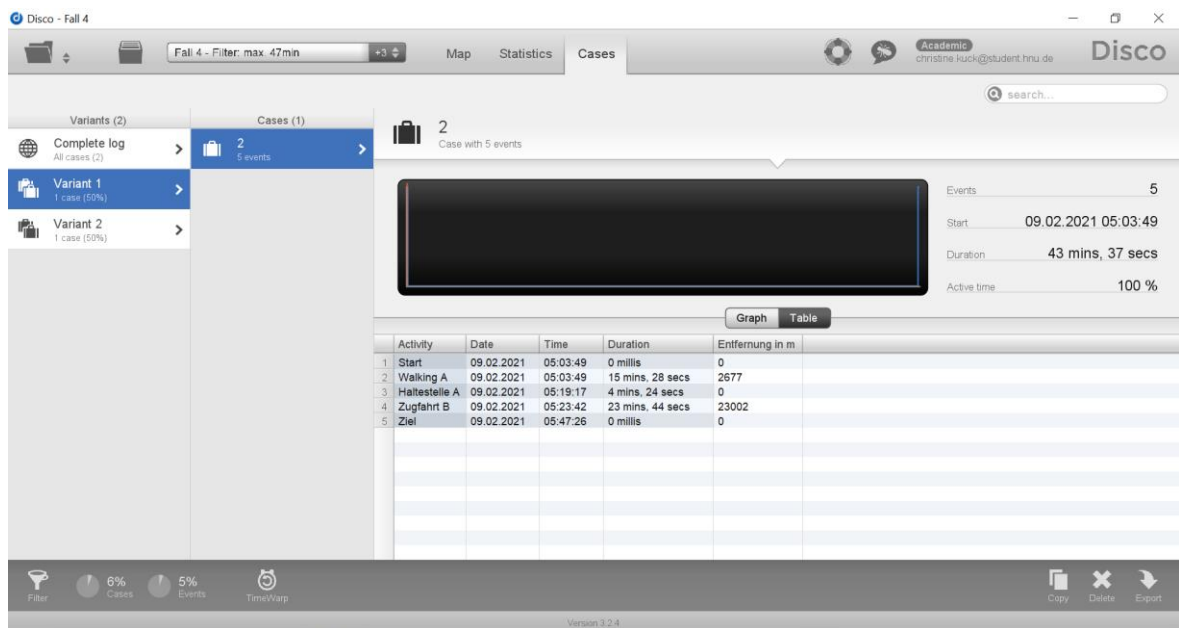


Abbildung 117: Auswertung Fall 4 - Filter: Kürzeste Cases. Eigene Darstellung (2022)

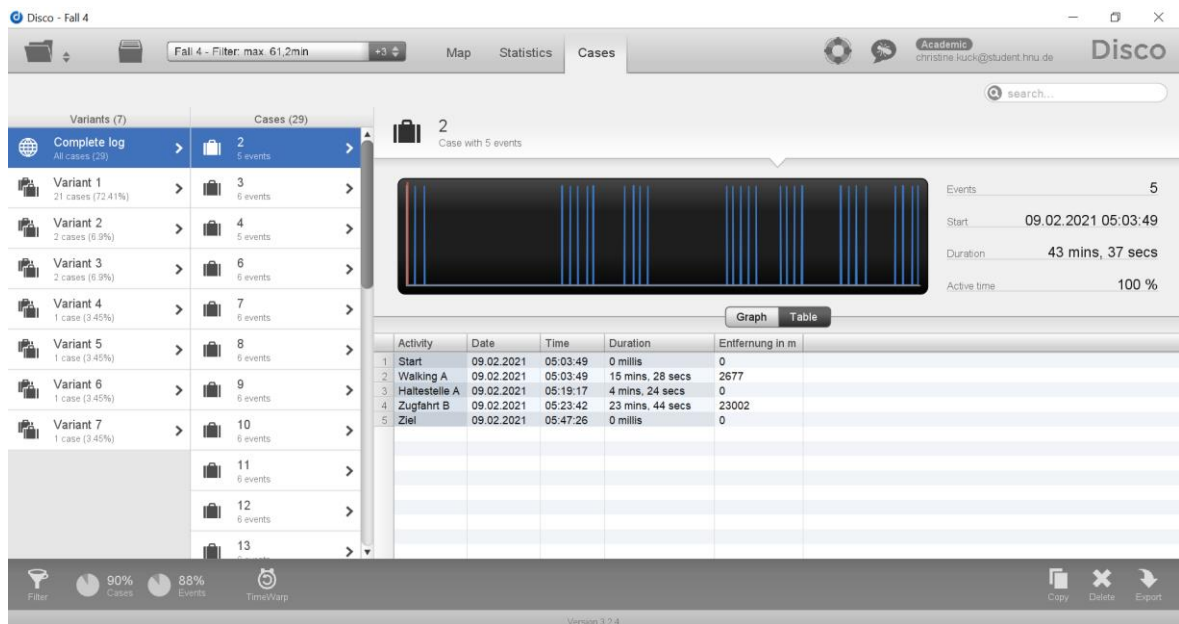


Abbildung 118: Auswertung Fall 4 - Filter: Kürzer als Durchschnitt. Eigene Darstellung (2022)

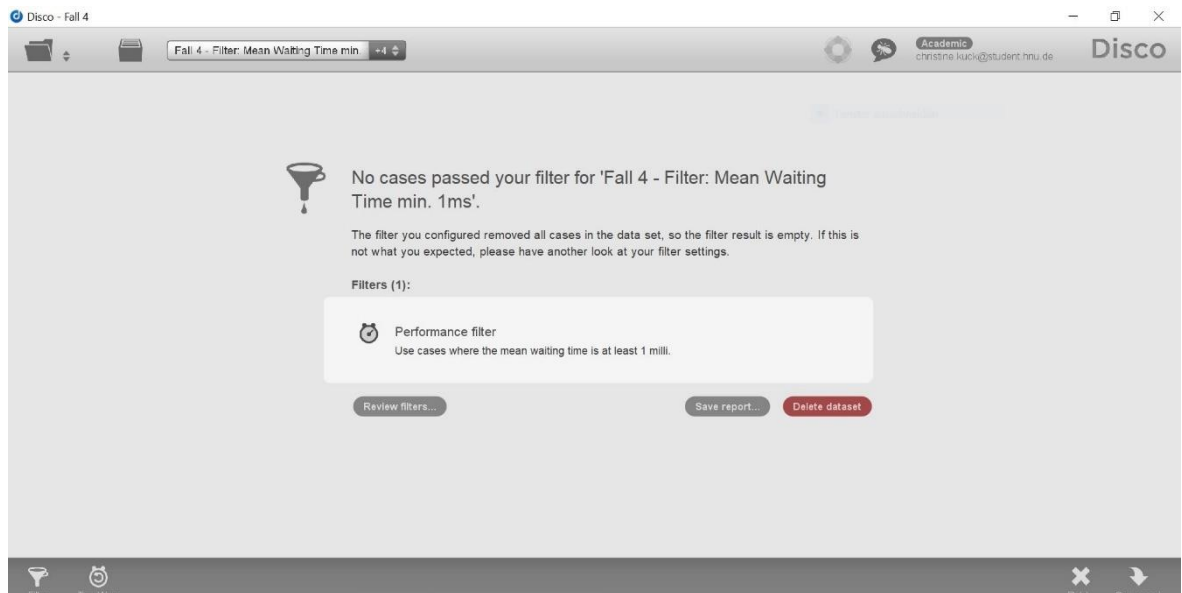


Abbildung 119: Auswertung Fall 4 - Filter: Cases mit Wartezeiten. Eigene Darstellung (2022)

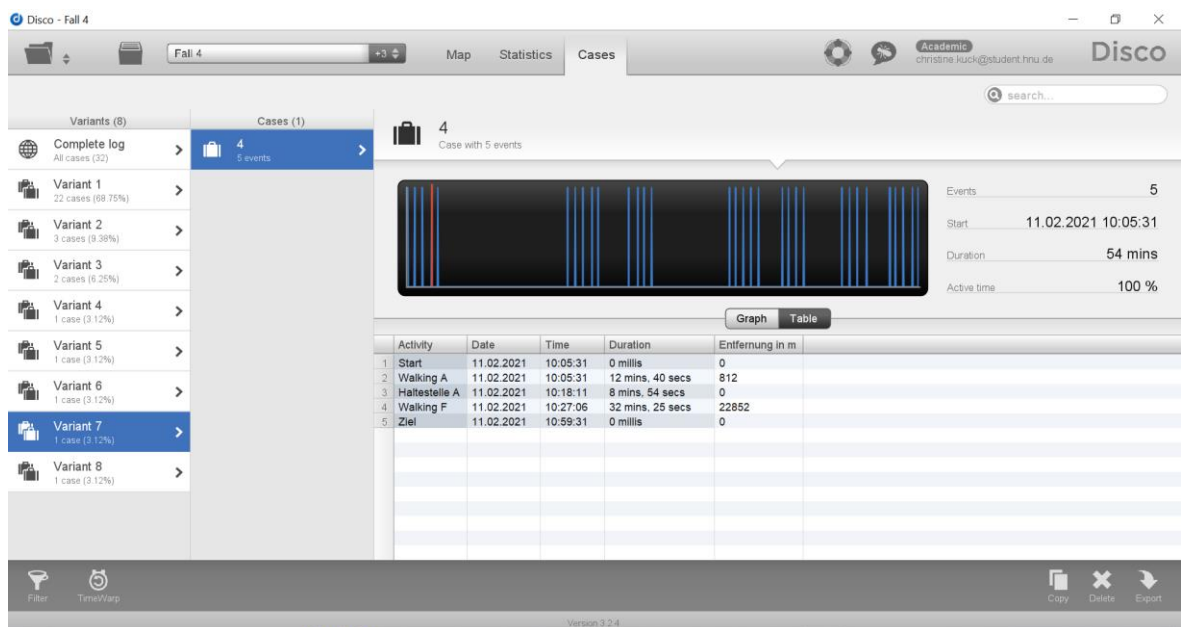


Abbildung 120: Fall 4 - Variante 7: fehlerhafte Aufzeichnung. Eigene Darstellung (2022)

55	Entfernung:	5962	+	484	=	6446	Meter	Zeit:	24 min	19 sek	=	1459	Sekunden	G =	15,90514051	km/h
56	Entfernung:	6328	+	541	=	6869	Meter	Zeit:	25 min	26 sek	=	1526	Sekunden	G =	16,20471822	km/h
57	Entfernung:	6088	+	485	=	6573	Meter	Zeit:	23 min	37 sek	=	1417	Sekunden	G =	16,69922371	km/h
58	Entfernung:	5993	+	690	=	6683	Meter	Zeit:	24 min	10 sek	=	1450	Sekunden	G =	16,59227586	km/h
59	Entfernung:	6237	+	555	=	6792	Meter	Zeit:	23 min	26 sek	=	1406	Sekunden	G =	17,39061166	km/h

Abbildung 121: Auswertung Fall 1 - Berechnung des Mittelwertes der KPI G in Excel. Eigene Darstellung (2022)

33	Case 28:	Entfernung:	27117	+		=	27117	Meter	Zeit:	25 min	25 sek	=	1525	Sekunden	G =	64,01390164	km/h
34	Case 29:	Entfernung:	25993	+		=	25993	Meter	Zeit:	23 min	20 sek	=	1400	Sekunden	G =	66,83914286	km/h
35	Case 30:	Entfernung:	28735	+	197	=	28932	Meter	Zeit:	29 min	14 sek	=	1754	Sekunden	G =	59,38152794	km/h
36	Case 31:	Entfernung:	28863	+	235	=	29098	Meter	Zeit:	33 min	53 sek	=	2033	Sekunden	G =	51,52621741	km/h

Abbildung 122: Auswertung Fall 2 - Berechnung des Mittelwertes der KPI G in Excel. Eigene Darstellung (2022)

Abbildung 123: Auswertung Fall 3 - Berechnung des Mittelwertes der KPI G in Excel. Eigene Darstellung (2022)

Abbildung 124: Auswertung Fall 4 - Berechnung des Mittelwertes der KPI G in Excel. Eigene Darstellung (2022)

Abbildung 125: Auswertung Fall 4 - Ermittlung der Durchschnittsdauer ohne Cases 1 und 28. Eigene Darstellung (2022)

Eidesstattliche Erklärung

Ich versichere, dass ich die vorliegende Abschlussarbeit selbständig angefertigt, nicht anderweitig für Prüfungszwecke vorgelegt, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt, sowie wörtliche und sinngemäße Zitate als solche gekennzeichnet habe und die Überprüfung mittels Anti-Plagiatssoftware dulde.

Nersingen, 25.09.2022

Ort, Datum

Kuck

Unterschrift