

Ozan Tastekin

Analyse von Bahndaten aus der Schweiz:
Pünktlichkeit des öffentlichen Verkehrs

Öffentlicher Verkehr ist für viele Menschen ein Teil ihres Alltags. Dabei ist es ihnen wichtig, dass die Busse und Bahnen rechtzeitig ankommen [1,2]. Das Verkehrsmittel kann sich aus vielen Gründen verspäten, zum Beispiel wegen einer Oberleitungsstörung, wegen Stau auf den Straßen oder wegen Personalausfällen [3].

In diesem Projekt wird auf weitere Gründe für Verspätungen im öffentlichen Verkehr eingegangen und Erkenntnisse aus Daten des Verkehrsbundes Zürich (VBZ) ins Licht gestellt.

Daten:

- Fahrzeiten Soll und Ist-Vergleich aus Zürich:**
Vergleich zwischen tatsächlichen und geplanten Ankunfts- und Abfahrtszeiten
Dazu gibt es Verbindungstabellen: **Haltestellen & Haltepunkte**
- Passagierfrequenzen der Hardbrücke-Haltestelle:**
Zählung an zwei Gleisen durch Sensor von ein- und aussteigenden Passagieren
- Stündliche Wetterdaten aus Zürich:**
Messungen von **3 Wetterstationen** in Zürich über Temperatur, Luftfeuchtigkeit, Regendauer, etc.

Datenbereinigung:

- Spaltenbereinigung**
Doppelt gespeicherte Spalten & solche, die keine zusätzliche Information geben entfernt
- Duplikate entfernen**
Doppelt gespeicherte Zeilen entfernt
- Leere Werte auffüllen**
Einzelne fehlende Werte aufgefüllt
- Fehlende Datenpunkte auffüllen**
Ganze fehlende Datenpunkte aufgefüllt

H1: Im Jahr 2022 gibt es einen positiven linearen Zusammenhang zwischen ein- und aussteigenden Fahrgästen und aufgebauten Verspätungen an der Hardbrücke Haltestelle bei den Linien 33, 72, 83 und 8.

Durchführung:

- Identifizierung relevanter Haltepunkte und zugehöriger Passagierzahlen
- Fahrzeiten aggregieren auf 5 Minuten
- Verspätungen ausrechnen
- Spearman’s Rangkorrelation anwenden

Ergebnisse:

- $r = 0,1529$: Sehr leichte positive Korrelation, nicht stark genug für Annahmekriterium => Hypothese wird abgelehnt
- Siehe Abbildung 1

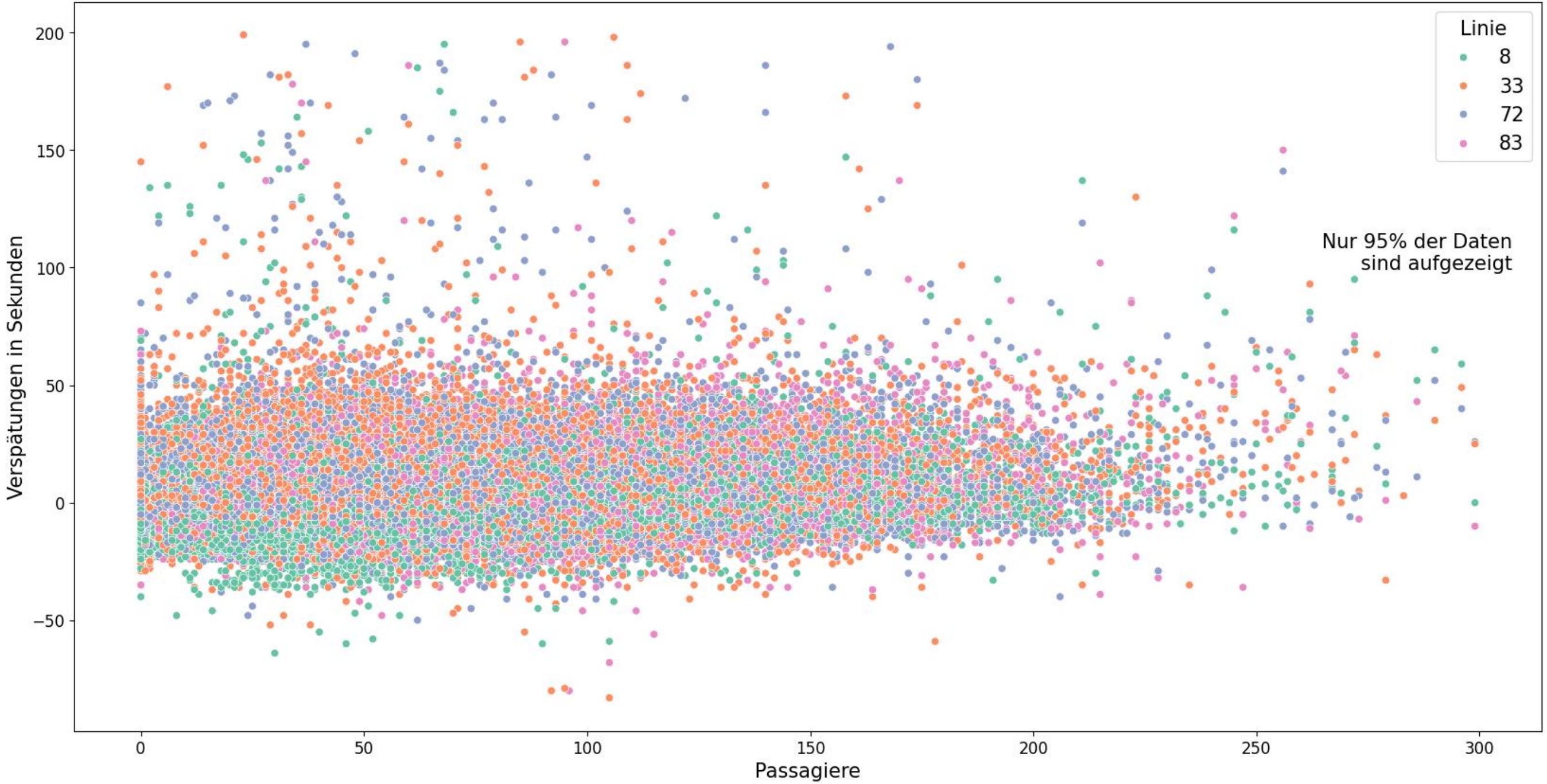


Abbildung 1: Korrelation zwischen Passagierfrequenzen und Verspätungen an der Hardbrücke Haltestelle

H2: In der Mehrheit der Jahre von 2016 bis 2022 verzeichneten Busse in Stunden mit mehr als 30 Minuten Regen überdurchschnittlich hohe aufgebaute Verspätungen.

Durchführung:

- Filtern nach Fahrzeiten Soll und Ist Vergleiche von Bussen
- Regendauer pro Stunde den Haltepunkten zuordnen:

H2a) Maximale Regendauer	H2b) Nächste Wetterstation
Maximale Regendauer der 3 Messungen der einzelnen Wetterstationen bestimmen und jedem Haltepunkt zuordnen	Nächste Wetterstation zu je einem Haltepunkt bestimmen und dessen Messung dem Haltepunkt zuordnen

- Fahrzeiten aggregieren auf 1 Stunde
- Verspätungen ausrechnen
- Mann-Whitney U Test anwenden

Ergebnisse:

- In beiden Versionen ist das Annahmekriterium nicht bei der Mehrheit der Jahre zutreffend => Hypothese wird abgelehnt
- Siehe Tabelle 1

	H2a) Maximale Regendauer		H2b) Nächste Wetterstation	
Jahr	p-Wert	< 0,05?	p-Wert	< 0,05?
2016	0,3679	✗	$0,8 \cdot 10^{-4}$	✓
2017	0,9999	✗	1,0	✗
2018	0,9933	✗	1,0	✗
2019	1,0	✗	1,0	✗
2020	$0,2 \cdot 10^{-59}$	✓	0,9601	✗
2021	1,0	✗	1,0	✗
2022	$0,25 \cdot 10^{-4}$	✓	0,0942	✗

Tabelle 1: Ergebnisse des Mann-Whitney U Tests von H2a) und H2b)

[1] Arnoud Mouwen. 2015. Drivers of customer satisfaction with public transport services. Transportation Research Part A: Policy and Practice 78 (2015), 1–20.
[2] Dea Van Lierop, Madhav G Badami, and Ahmed M El-Geneidy. 2018. What influences satisfaction and loyalty in public transport? A review of the literature. Transport Reviews 38, 1 (2018), 52–72.
[3] Enikő Nagy and Csaba Csiszár. 2015. Analysis of delay causes in railway passenger transportation. Periodica Polytechnica: Transportation Engineering 43, 2 (2015), 73–80.