

Programmieren
Projekt “Unfallatlas-Analyse” – Bericht

29.06.2025

Ali Ufuk Bozkurt & Janusz Guzman

Der Code behandelt die öffentlich verfügbaren Daten des Unfallatlas für Deutschland (<https://unfallatlas.statistikportal.de/>) in den Jahren 2019–2023, beschränkt auf die Stadt Karlsruhe. Es werden dabei 3248 Unfälle betrachtet. Mithilfe des k-means-Algorithmus, einem Clustering-Verfahren, sollen möglichst alle separaten Orte mit Unfällen ermittelt werden, sodass gefährliche Kreuzungen optisch identifiziert werden können. Diese sollen auf einer interaktiven (mit Folium erzeugten) Karte angezeigt werden. Zusätzlich soll ermittelt werden, ob ein Zusammenhang zwischen dem an Unfallorten geltenden Tempolimit und der dortigen Unfallschwere besteht.

Um beim Clustering-Verfahren optimale Ergebnisse zu ermitteln, gibt es genau eine wichtige Stellschraube: die Anzahl der Cluster. Wie viele sind zu wenig und wie viele zu viel? Es gibt Werte, die die optimale Clusterzahl ermitteln sollen, wie etwa den Silhouette-Score, den wir ebenfalls verwendet haben. Wir haben uns aber schlussendlich für das optisch Richtige entschieden, da eine Karte exportiert werden soll, also die Darstellung zählt – ob eine Karte übersichtlich ist, lässt sich nur optisch beurteilen. Auf der Karte lässt sich schnell verifizieren, ob Cluster-Mittelpunkte (Zeichen für eine gefährliche Kreuzung) an Unfallorten fälschlicherweise fehlen oder sogar mehrmals auftauchen.

Optisch haben wir uns für 308 Cluster – also im Durchschnitt ca. 10,5 Unfälle pro Cluster – entschieden. Der Test mit dem Silhouette-Score ergab eine optimale Clusteranzahl von 919 (Silhouette-Score hierfür: 0,572). Diese halten wir hier für unangemessen hoch, wie man an folgenden Abbildungen sieht:



Abb. 1 & 2: Kreuzung Oststadtkreisel bei “optimaler” Clusteranzahl von 919 (links – fälschlicherweise wie zwei Kreuzungen behandelt, mit zwei Clusterzentren) und Clusteranzahl von 308 (rechts – korrekt und ungefähr in der Mitte der Kreuzung)

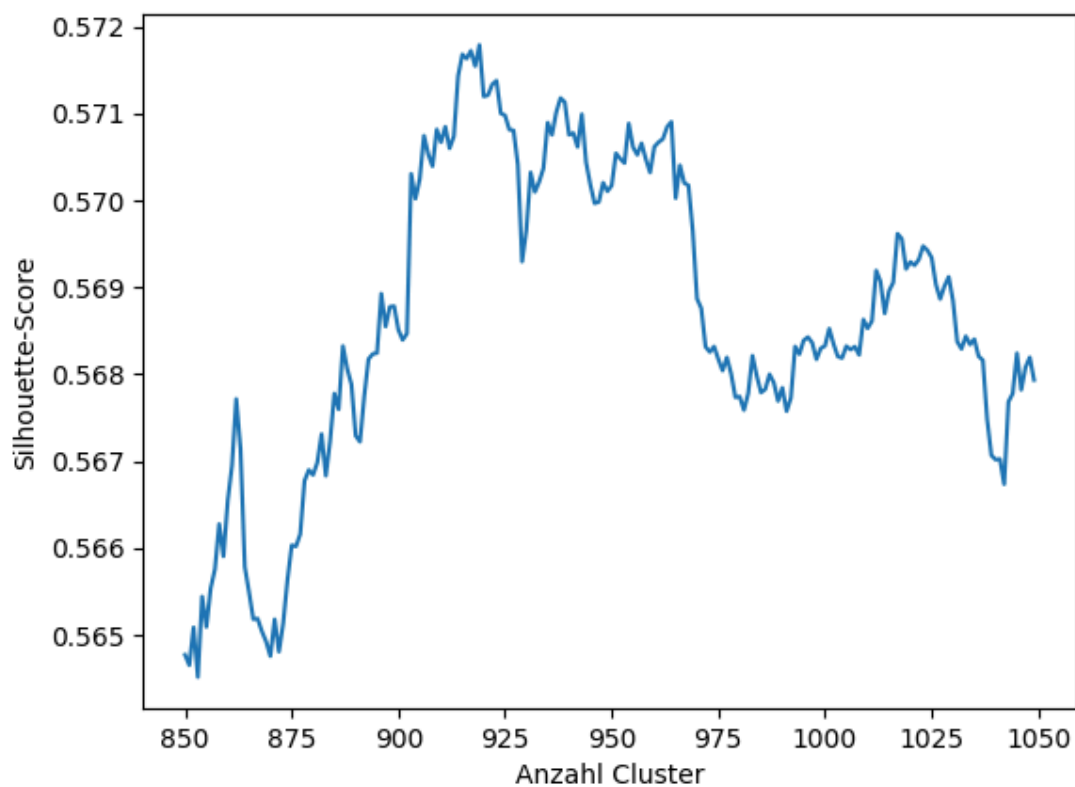


Abb. 3: Visualisierte Ergebnisse des Silhouette-Score-Tests im Bereich der höchsten Scores

Für die Fragestellung nach der Korrelation zwischen Tempolimit und Unfallschwere haben wir zur Ermittlung des an den jeweiligen Unfallorten geltenden Tempolimits OpenStreetMap-Daten (wo vorhanden) aus einer Overpass-API-Abfrage angewendet. Um Probleme (insbesondere mit der API) zu vermeiden, haben wir weniger Unfalldaten (nur die neuesten vom Jahr 2023) genommen, da diese nach wie vor alle Jahreszeiten und Situationen abdecken. Es wurden also 1150 Unfälle betrachtet.

Die Unfälle mit gleichen Tempolimits (bspw. 30, 50, 80 km/h) wurden nach ihrer Unfallschwere (UKATEGORIE) gezählt. Damit die einzelnen Geschwindigkeiten miteinander vergleichbar sind, haben wir uns nur auf die Verhältnisse (Unfälle mit Toten [1] / Schwerverletzten [2] / Leichtverletzten [3]) und nicht auf absolute Zahlen (welches Tempolimit am häufigsten bei einem schweren Unfall herrschen würde) konzentriert. Es ergibt sich folgendes Bild:

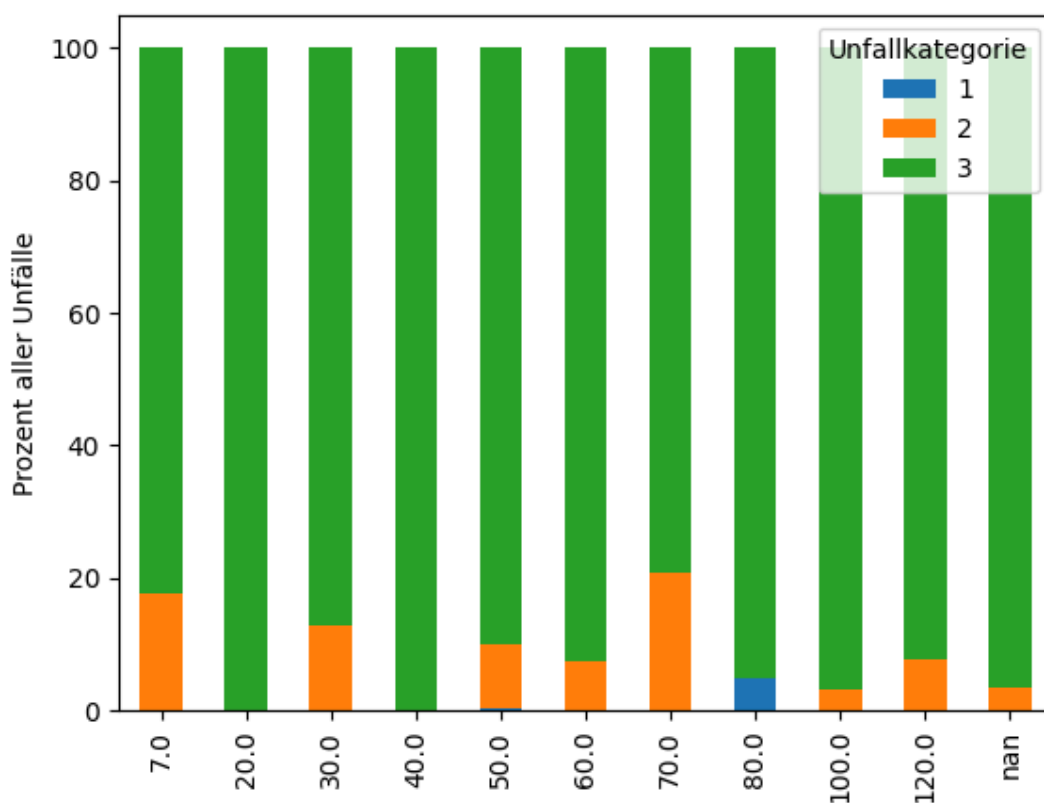


Abb. 4: Unfallschwere (prozentuell) nach Tempolimit. nan = keines

Daraus lässt sich keine eindeutige Tendenz erkennen, da die meisten Unfälle mit Schwerverletzten und Toten in Karlsruhe bei 70 km/h, die zweitmeisten solcher Unfälle allerdings bei Schrittgeschwindigkeit entstehen. Wir führen dies darauf zurück, dass hohe Tempolimits weitestgehend dort herrschen, wo Radfahrer und Fußgänger nicht fahren dürfen, und bei diesen Geschwindigkeiten im Durchschnitt ein bis (beim Zusammenstoß mit einem anderen Fahrzeug) zwei Autofahrer (schwer) verletzt werden. Bei geringen Tempolimits hingegen sind Autofahrer sicher. Diese herrschen aber auch selten dort, wo Autofahrer alleine sind. Dort wird im Durchschnitt stattdessen ein Fußgänger oder Radfahrer (schwer) verletzt.

Es lässt sich somit kein optimales Tempolimit empfehlen. Dies wäre aber ohnehin unmöglich, da es von vielen weiteren und wichtigen Bedingungen abhängt, sodass es kein optimales Tempolimit für alle Situationen geben kann.