Usage Patterns of Bicycle Counting Stations in Heidelberg and the Influence of External Factors

Julian Jurcevic [*1] **Martin Eichler** [*2] **Simon Rappenecker** [*3] **Tarik Eker** [*4]

Abstract

As cycling becomes key in sustainable urban mobility, understanding spatiotemporal traffic patterns is essential for infrastructure planning. While previous studies have classified bicycle traffic, they often rely on predefined, rule-based approaches that may fail to capture hybrid usage behaviors. Addressing this limitation, we present a data-driven approach to classify urban bicycle traffic using hourly data from counting stations in Heidelberg, Germany. We derive special features to quantify the shape of traffic patterns across various timescales. Subsequent k-means distinguishes distinct usage patterns: *utilitarian*, *recreational*, and *mixed*. Finally, we investigate the influence of external factors, such as weather and public holidays, which varies significantly across different station types.

1. Introduction

Cycling demand in urban areas varies strongly across locations and over time. Previous studies have shown that bicycle traffic exhibits characteristic *daily*, *weekly*, and *seasonal* patterns that reflect underlying trip purposes such as commuting or leisure travel (Miranda-Moreno et al., 2013).

External conditions play a key role in shaping cycling behaviour. Cycling demand is highly sensitive to weather: precipitation and low temperatures reduce cycling activity, while warm conditions increase usage (Liu et al., 2015; Rudloff et al., 2014). These weather effects are heterogeneous, as frequent cyclists tend to be less affected by adverse conditions than occasional users (Heinen et al., 2011). Public holidays are associated with lower overall traffic volumes and changes in daily traffic patterns (Cools et al., 2007).

In parallel, cycling has gained increasing policy relevance. In Baden-Württemberg, long-term strategies explicitly promote cycling as a central component of sustainable urban mobility (Ministerium für Verkehr Baden-Württemberg, 2016), highlighting the need for data-driven analyses of cycling infrastructure usage.

Against this background, we investigate how urban bicycle counting stations can be characterised using daily, weekly, and seasonal traffic patterns, and how weather conditions and public holidays affect cycling demand across usage types. Using multi-year hourly data from permanent counting stations in Heidelberg, we derive interpretable features and apply unsupervised k-means clustering to distinguish *utilitarian*, *recreational*, and *mixed-use* stations and assess their sensitivity to external factors.

2. Data

The analysis is based on bicycle traffic counts complemented by weather and public holiday data. In the following, we describe the data sources, preprocessing steps, and quality checks.

2.1. Bicycle Counting Data

The bicycle counting stations in Heidelberg are densely and centrally distributed across the urban area. For our analysis, hourly bicycle counts from 15 permanent monitoring stations installed between May 2014 and January 2020 were used. The stations record counts at an hourly resolution using Eco-Counter systems equipped with ZELT inductive loops (Eco-Counter, 2025). These inductive loops are embedded in the pavement and detect cyclists based on characteristic electromagnetic signatures. Most stations are capable of recording directional flows (inbound and outbound).

The counting data is publicly available via the MobiData BW open data platform (Nahverkehrsgesellschaft Baden-Württemberg mbH (NVBW), 2025). Data quality was as-

[*]Equal contribution [1]Matrikelnummer 7167089, MSc Computer Science [2]Matrikelnummer 6009076, MSc Computer Science [3]Matrikelnummer 6324777, MSc Machine Learning [4]Matrikelnummer 5668988, MSc Computer Science. Correspondence to: JJ <julian-steffan.jurcevic@student.uni-tuebingen.de>, ME <martin.eichler@student.uni-tuebingen.de>, SR <simon.rappenecker@student.uni-tuebingen.de>, TE <tarik.eker@student.uni-tuebingen.de>.

Project report for the "Data Literacy" course at the University of Tübingen, Winter 2025/26 (Module ML4201). Style template based on the ICML style files 2025.

sessed with respect to missing hourly observations and false-zero days, defined as days with a total recorded count of zero.

uality: Add
ative meassures

2.2. Weather Data

Weather data were used as external factors influencing bicycle counts.

Hourly weather data were obtained from the Open-Meteo API (Open-Meteo, 2025). The data cover the period from 2013 to 2024 and provides hourly weather information based on numerical weather models from national meteorological services.

s not used,

As nearly all bicycle counting stations in Heidelberg lie within the same Open-Meteo grid cell, a single city-wide hourly weather time series was used for all stations.

Locally recorded meteorological data from the counting stations were excluded due to inconsistencies identified in comparison with official weather data.

2.3. Holiday Data

Public holidays and school vacation periods were included to model calendar effects on bicycle traffic. Holiday data for Baden-Württemberg were obtained from the Mehr Schulferien API (meh, 2025), providing start and end dates for public holidays and school vacations.

uality: Add
res that had to
en to clean up

3. Method

Previous works often rely on predefined rules to categorize bicycle traffic (Miranda-Moreno et al., 2013). However, such an approach may fail to capture hybrid usage patterns. We propose a more data-driven, implicit approach. By extracting features that quantify the 'shape' of daily, weekly, and seasonal profiles, we map each station into a multi-dimensional feature space. This allows us to use unsupervised k-means clustering (Lloyd, 1982) to discover different station types.

In a subsequent step in chapter 4, we assign one of the following labels to each cluster (cf. Fig.):

Patterns, In-

- *Utilitarian*: Traffic showing strong commuting patterns, with two distinct peaks in traffic volume during the morning and afternoon hours.
- *Recreational*: Traffic dominated by leisure activities. These locations exhibit a single peak in traffic volume around midday or early afternoon.
- *Mixed*: Locations that serve a dual purpose, showing characteristics of both utilitarian and recreational usage.

To characterise the counting stations, we derive three distinct features that capture traffic patterns across different timescales: daily, weekly and seasonal to exploit the distinct pattern visible in Figure xyz .

Figure pattern

Double Peak Index (DPI)

Weekday hourly profiles of some stations exhibit a double-peak structure, typically associated with morning and evening commuting. The DPI quantifies this behaviour by identifying dominant morning (5–10 h) and evening (14–20 h) peaks and relating their magnitudes to the average midday level (8–14 h). Stations with clear, balanced commuting peaks yield high DPI values, whereas flat or single-peak profiles result in low scores.

Formally, let p_m and p_e denote the magnitudes of the morning and evening peaks at hours h_m and h_e, and let m be the average midday traffic level. The DPI is defined as

$$\mathrm{DPI} = \max(S \cdot Y \cdot D,\, 0)$$

where S, Y and D models strength, symmetry and distance, respectively:

$$\begin{aligned} S &= \frac{(p_m - m) + (p_e - m)}{2} \\ Y &= 1 - \frac{|p_m - p_e|}{\max(p_m, p_e)} \\ D &= \min\left(\frac{|h_e - h_m|}{10},\, 1\right) \end{aligned}$$

Weekend Shape Difference (WSD)

Differences between weekday and weekend hourly traffic patterns provide an additional discriminator between usage types. To capture this effect, we compare the shape of the weekday and weekend hourly profiles. Let $\mathbf{p}^{wd}$ and $\mathbf{p}^{we}$ denote the weekday and weekend hourly profiles, normalised to sum to one.

$$\mathrm{WSD} = \left\| \frac{\mathbf{p}^{wd}}{\sum_h p_h^{wd}} - \frac{\mathbf{p}^{we}}{\sum_h p_h^{we}} \right\|_2$$

Seasonal Drop Index (SDI)

Finally, we consider long-term patterns. Seasonality provides a discriminator between leisure and utilitarian-oriented stations. The SDI quantifies the relative decline between high and low-usage months. Let I_m denote the monthly index values of a station. Using upper and lower quantiles to ensure robustness against outliers

$$\begin{aligned} q_{90} &= \mathrm{quantile}_{0.9}(I_m) \\ q_{10} &= \mathrm{quantile}_{0.1}(I_m) \end{aligned}$$

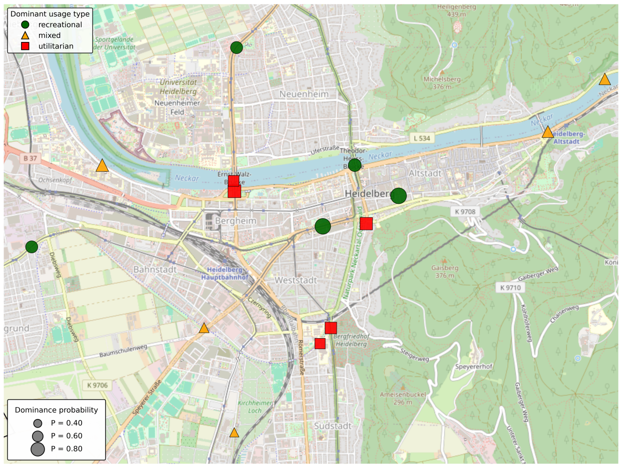

Figure 1. Spatial distribution of bicycle counting stations in Heidelberg. Marker shape indicates the dominant usage type (utilitarian, recreational, mixed), while marker size reflects the dominance probability.

the SDI is defined as

$$\mathrm{SDI} = \frac{q_{90} - q_{10}}{q_{90}}$$

High values indicate strong seasonal variation, whereas low values correspond to relatively stable, year-round usage.

4. Results

- describe temporal shift, describe why this is expected and why, PLOT: city,
- describe holidays impact, describe behaviour (similar as weekend) PLOT: show shift
- include also plot of different public holidays (my own one), this is just to have a funny fact (Vater Tag) PLOT: funny plot, na werden zu viele plots
- describe weather impact, obvious PLOT: weather

5. Conclusion

- summarize
- limitations
- problems
- statements that can be made

Contribution Statement

Add this, see original template

References

Mehr schulferien – ferien- und feiertagsdaten für deutschland. https://www.mehr-schulferien.de/, 2025. Accessed: 2025-12-29.

Cools, M., Moons, E., and Wets, G. Investigating effect of holidays on daily traffic counts: Time series approach. *Transportation Research Record: Journal of the Transportation Research Board*, (2019):22–31, 2007. doi: 10.3141/2019-04.

Eco-Counter. Zelt evo – inductive loop for mixed traffic. https://de.eco-counter.com/losungen/produkte/zelt-evo-selective, 2025. Accessed: 2025-12-27.

Heinen, E., Maat, K., and van Wee, B. Day-to-day choice to commute or not by bicycle. *Transportation Research Record: Journal of the Transportation Research Board*, (2230):9–18, 2011. doi: 10.3141/2230-02.

Liu, C., Susilo, Y. O., and Karlström, A. The influence of weather characteristics variability on individual's travel mode choice in different seasons and regions in sweden. *Transport Policy*, 41:147–158, 2015. doi: 10.1016/j.tranpol.2015.01.001.

Lloyd, S. Least squares quantization in pcm. *IEEE Transactions on Information Theory*, 28(2):129–137, 1982. doi: 10.1109/TIT.1982.1056489.

Ministerium für Verkehr Baden-Württemberg. Radstrategie baden-württemberg. Technical report, Ministerium für Verkehr Baden-Württemberg, Stuttgart, Germany, 2016.

Miranda-Moreno, L. F., Nosal, T., Schneider, R. J., and Proulx, F. Classification of bicycle traffic patterns in five north american cities. *Transportation Research Record*, 2339(1):68–79, 2013. doi: 10.3141/2339-08. URL https://doi.org/10.3141/2339-08.

Nahverkehrsgesellschaft Baden-Württemberg mbH (NVBW). Bicycle counting data repository (api v2). https://mobidata-bw.de/fahrradzaehldaten/v2/, 2025. Accessed: 2025-12-27.

Open-Meteo. Historical Weather API Documentation. https://open-meteo.com/en/docs/historical-weather-api, 2025. Accessed: 2025-12-27.

Rudloff, C., Leodolter, M., Bauer, D., Auer, R., Brög, W., and Kehnscherper, K. Influence of weather on transport demand: A case study from the vienna region. Technical report, Austrian Institute of Technology (AIT), Vienna, Austria, August 2014. Technical report.