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

Sustainable urban traffic relies heavily on the expansion of cycling. The state of Baden-Württemberg recognizes this, explicitly promoting cycling as a key component of its long-term strategy (Ministerium für Verkehr Baden-Württemberg, 2016). To translate these strategies into infrastructure, planners require a good understanding of traffic behavior. Specifically, they need to know not just how many people are cycling, but when and why.

Trip purposes are typically inferred from spatiotemporal traffic patterns. While previous studies have successfully used hourly and daily counts to distinguish utilitarian from recreational traffic, they often rely on rule-based approaches (Miranda-Moreno et al., 2013). These methods, however, struggle to capture hybrid usage patterns common in urban environments.

Furthermore, cycling demand is highly sensitive to external factors. Precipitation, temperature, and public holidays can influence traffic patterns (Liu et al., 2015; Rudloff et al., 2014; Cools et al., 2007). Since commuters are more weather-resilient than leisure riders, these effects vary by station type (Heinen et al., 2011). Therefore, we aim to understand how these factors influence the demand profiles of different station types in the city of Heidelberg, Germany.

To overcome the limitations of rule-based classification, we present a flexible, data-driven approach. Using hourly counting data from 15 stations in Heidelberg , we first derive features to quantify traffic shapes across daily, weekly, and seasonal timescales. We then apply k-means clustering to identify different usage patterns (*utilitarian*, *recreational*, and *mixed*) and analyze their sensitivity to weather and public holidays.

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 assessed by treating days with fewer than 12 hourly observations as missing and by flagging zero-count days at high-traffic stations (median $>$ 500 cyclists/day) as probable sensor failures. Table 1 reports data quality and traffic statistics for all 15 stations. Stations are identified by short IDs (S1–S15) and referenced consistently throughout the paper. The standard deviation (Std) captures day-to-day variability, while the 95% confidence interval margin (CI) quantifies the uncertainty of the estimated daily mean.

Table 1. Data quality and traffic statistics for Heidelberg bicycle counting stations, ranked by data availability.

ID	**Station**	**Mean** [bikes/d]	**Std**	**CI** [±]	**Avail.** [%]
S1	Plöck	4180	1832	62	91.5
S2	Kurfürstenanlage	1257	630	26	91.8
S3	Ernst-Walz-Br.	6179	3344	101	100.0
S4	Gaisbergstraße	4239	1651	57	94.7
S5	Schlierb. Landstr.	786	389	20	60.2
S6	Ziegelh. Landstr.	872	389	20	58.3
S7	Rohrbacher Str.	2426	1251	48	95.7
S8	Theodor-Heuss-Br.	7902	2865	108	100.0
S9	Mannheimer Str.	1964	845	28	83.2
S10	Ernst-Walz-Br. alt	3240	1517	75	99.9
S11	Hardtstraße	1690	640	30	79.4
S12	Berliner Str.	705	245	11	92.3
S13	Bahnstadtprom.	2759	1203	67	63.7
S14	Eppelheimer Str.	909	452	20	95.6
S15	Liebermannstr.	4101	1456	57	71.9

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.):

Figure Patterns, Indices

- *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:

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

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.

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

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

$$q_{90} = \text{quantile}_{0.9}(I_m)$$
$$q_{10} = \text{quantile}_{0.1}(I_m)$$

the SDI is defined as

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

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

Temporal Clustering

To account for temporal variability in bicycle traffic, k-means clustering is performed on sliding two-year windows with monthly shifts. Feature vectors are recomputed separately for each window.

For each station s and usage type u, the empirical cluster membership probability is defined as

$$P(u \mid s) = \frac{1}{N_s} \sum_{i=1}^{N_s} \mathbb{I}\big(c_i(s) = u\big),$$

where $c_i(s)$ denotes the cluster assignment in the i-th run and N_s the number of valid runs.

Table 2. Weather class definitions for the event-based analysis.

Class	$T_{\max}$ [°C]	P_{day} [mm]
L	< 10	$= 0$
M	[10, 20)	(0, 5]
H	≥ 20	> 5

Since k-means produces unlabeled partitions, cluster semantics are assigned post hoc by ordering clusters according to a centroid-level *utilitarian score*

$$\text{DPI} + \text{WSD} - \text{SDI},$$

with lower scores labeled as *recreational*, intermediate as *mixed*, and higher as *utilitarian*.

4. Results

Low-condition days (L) serve as baseline, and relative changes in mean daily bicycle counts are computed as

$$\Delta_X = \frac{\bar{C}_X - \bar{C}_L}{\bar{C}_L} \times 100, \quad X \in \{M, H\}.$$

Table 3. Effect of temperature on station usage across usage types.

	Weekdays			**Weekends**		
Type	L	Δ_M	Δ_H	L	Δ_M	Δ_H
Mix	1447	31%	72%	948	49%	116%
Rec	1034	17%	36%	801	26%	59%
Util	3543	18%	43%	2525	26%	65%

Table 4. Effect of precipitation on station usage across usage types.

	Weekdays			**Weekends**		
Type	L	Δ_M	Δ_H	L	Δ_M	Δ_H
Mix	2100	−17%	−37%	1587	−19%	−46%
Rec	1510	−9%	−22%	1275	−12%	−29%
Util	4810	−8%	−23%	3823	−13%	−29%

Table 5. Effect of public holidays on station usage by usage type.

Type	No holiday	Holiday	Δ
Mix	1648	953	−28.6%
Rec	1260	728	−42.2%
Util	3238	2043	−42.8%

Maybe calc delta of different holidays

- describe temporal shift, describe why this is expected and why, PLOT: city,
- describe holidays impact, describe behaviour (similar as weekend) PLOT: show shift

Figure 1. todo

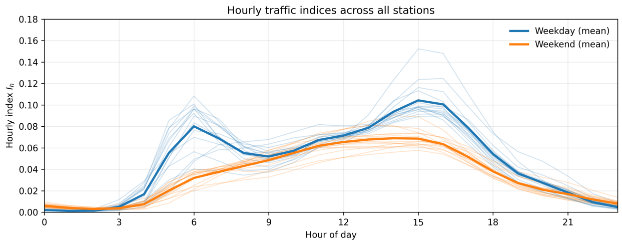
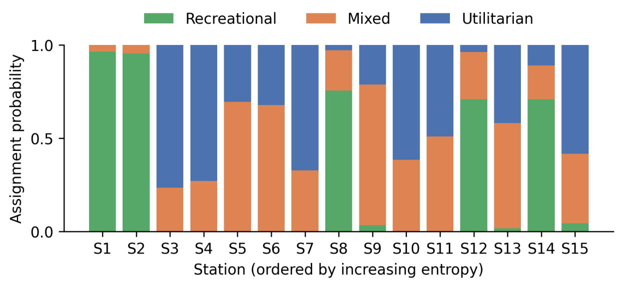


Figure 2. todo



- 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

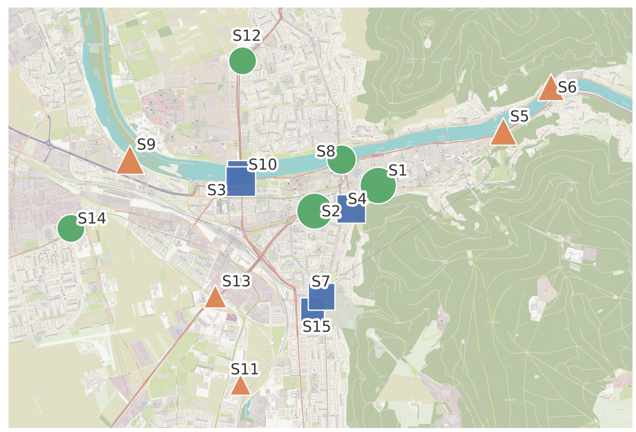


Figure 3. Spatial distribution of bicycle counting stations in Heidelberg. Marker shape indicates the dominant usage type (utilitarian, recreational, mixed), while marker size reflects the dominance probability.

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.

*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.