



Open data and quantitative techniques for anthropology of road traffic

Ajda Pretnar Žagar¹ , Tomaž Hočvar¹  and Tomaž Curk^{1*} 

*Correspondence:

tomaz.curk@fri.uni-lj.si

¹Faculty of Computer and Information Science, University of Ljubljana, Večna pot 113, SI 1000 Ljubljana, Slovenia

Abstract

What kind of questions about human mobility can computational analysis help answer? How to translate the findings into anthropology? We analyzed a publicly available data set of road traffic counters in Slovenia to answer these questions. The data revealed information on how a population drives, how it travels for tourism, which locations it prefers, what it does during the week and the weekend, and how its habits change during the year. We conducted the empirical analysis in two parts. First, we defined traffic profile deviations and designed computational methods to find them in a large data set. As shown in the paper, traffic counters hint at potential causes and effects in driving practices that we interpreted anthropologically. Second, we used hierarchical clustering to find groups of similar traffic counters as described by their daily profiles. Clustering revealed the main features of road traffic in Slovenia. Using the two quantitative approaches, we outlined the general properties of road traffic in the country and identified and explained the outliers. We show that quantitative data analysis only partially answers anthropological questions, but it can be a valuable tool for preliminary research. We conclude that open data are a useful component in an anthropological analysis and that quantitative discovery of small local events can help us pinpoint future fieldwork sites.

Keywords: Quantitative analysis; Open data; Travel habits; Road traffic counters; Road traffic flows; Computational anthropology

1 Introduction

With computational techniques gaining popularity in anthropology, there is no lack of examples successfully mixing quantitative and qualitative approaches. Ethno-mining [1], stitching [2], blending [3], circular mixed methods [4], or hybrid methodologies [5] join the best of both (methodological) worlds into a single research framework to explore human habits and practices. However, could one do predominantly quantitative anthropological research?

Anthropology has been, at least since Malinowski [6], deeply intertwined with long periods spent with the people under research, *i.e.*, fieldwork, and with a particular research technique called participant observation [7]. The researcher participates in the daily life of a community with a heightened awareness of social, cultural, economic, religious, and physical processes. Such participation results not only in detail-rich narratives

© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

called ethnographies, but also in a lived, embodied experience [8]. The researcher in anthropology is often seen as a “research tool” – a medium that translates observations into a coherent and structured form.

Therefore, can one do anthropological research without immediate human contact but solely using semi-big data and computational techniques? To demonstrate how such research can be done, we analyzed a publicly available data set on traffic frequencies on roads in Slovenia. Our aim was to determine the driving patterns and outline Slovenian road traffic. We argue that there is valuable information in traffic data showing how people in a country drive, travel for tourism, which locations do they prefer, what do they do during the week and the weekend, and how the drivers’ habits change during the year. We corroborated these findings with newspaper archive data on local festivities, road closures, and infrastructure changes. Our contribution is an anthropological interpretation of outlier detection, where we establish what is a “typical behavior” by determining the baseline and by looking for various deviations from the baseline as indicators of unusual human habits and practices. We address two research questions, namely how can open data aid anthropological research, and which traffic counters show distinctive patterns? Distinctive patterns were defined as deviant car traffic profiles, by which we aimed to detect high-frequency counters, seasonal increases and decreases, weekly patterns, changes in traffic infrastructure, and detect local festivities, thus creating an image of trends and particularities of the Slovenian car traffic system.

First, we established the typical car traffic behavior, *i.e.*, we looked at how a population uses cars to move around the country. Exploratory analysis showed when traffic peaks occur and where, what are the most frequented locations, and how the traffic changes during the weekend and between different seasons. Second, we used statistical methods to find interesting counters, namely those locations that showed a deviation from typical behavior (*e.g.*, road accidents, traffic congestion, changes in road infrastructure). Unusual traffic patterns showed how traffic changes and hinted at potential causes and effects that we interpreted anthropologically. Third, we used hierarchical clustering to structure the traffic patterns into groups and again tried to explain those groups from the perspective of mobility and traffic as a reflection of social structures.

The analysis in this study is an example of a first phase in computational anthropology applied to publicly available data. Combined with computational techniques, public data can be used by anthropologists to gauge human practices. We show that quantitative data science approaches such as data mining and machine learning are appropriate for the analysis of large data sets (longitudinal and simultaneous phenomena) and for preliminary field research (generating research questions). Afterward, anthropological methods explain the discovered patterns, place the results in social and cultural contexts, and supplement the findings with detailed descriptions (see also [4]). Therefore, data science can be used in anthropology as the first analytical phase, even before fieldwork begins. To encourage the reproducibility and reuse of the data, we provide the data, Jupyter notebooks, and workflows in Orange,¹ an open source machine learning and interactive data visualization tool (see section “Availability of data and materials”).

¹Orange Data Mining, <https://orange.biolab.si>.

2 Related work

Traffic reflects human habits and practices [9, 10]. It is a form of communication where the locations (cities, villages, points of interest) are the emitters and receivers, and the drivers are the signal. As Horta [11] succinctly argues, roads embody a social process in which the collective life of a society emerges to the surface through movement. In our case, car traffic movement is a social process with specific characteristics, intent, and societal implications.

Estevan [12] argues that “mobility” is a quantitative side of the movements performed by people and commodities, while “accessibility” is the qualitative counterpart. It is worth noting that in contemporary data science-based mobility analysis, accessibility is also quantified. Traffic data describe “mobility,” the basic patterns of social movement, quantity, frequency, and (i)regularity. However, glimpses of “accessibility” are also hidden in quantitative data. The activity of individuals is a reflection of social structures – collective preferences, negotiations, restrictions, and affordances. Traffic data is an aggregated individual action, displaying social structures through preferred or shunned locations, handling of (traffic) contingencies and corresponding strategies, and broad habitual patterns of a country.

Cars are an extension of human beings, affording them specific mobility and motility [13]. Assemblages of drivers and cars do not embody the technology, but the relationship between society and technology and how technology is used and appropriated to achieve specific means. The intentions behind driving are demonstrated in the choice of destination, route, and driving time, which are also contained in a numerical form of car traffic data.

Several studies demonstrate the usefulness of quantitative analysis of human mobility data. Palmer *et al.* [14] relate GPS positioning and demographic data to establish intimate patterns of human mobility, showing the activity, segregation, and well-being of research participants through spatial data. Focusing on urban areas, Gallotti *et al.* [15] analyze mobility data from 10 large cities to determine the structure of internal flows. Their analysis uncovers the heterogeneity of flows and internal segregation of cities, which aids in determining future policy responses. With London as their use case, Aslam and Cheng [16] model individual points of interest using the data from the Oyster Smart Card (payment system for the London Underground). Taking it a step further, Zhao *et al.* [17] use high-resolution human mobility data from cell phones. They show that high-resolution data differ significantly from more high-level aggregated mobility data.

Nevertheless, when high-resolution data is not available, even low-resolution data can reveal interesting patterns. Huang and Wong [18] extracted regular individual activity from geotagged Twitter data, showing that low-resolution data can provide valuable insights. Car traffic data from the present study fall into the medium resolution category and, as we show, offer comparable insights into human mobility, for example, the difference in mobility between weekdays and weekends or the dependence of the time of the day on location [14].

In anthropology, quantitative data for analyses of human mobility have been used sparingly. Podjed [19] shows how to combine ethnography with quantitative data analysis to determine the differences in driving styles and the effect of mobile apps on driving behavior. The author used quantitative data to map frequent behaviors, while qualitative data from interviews and the so-called “participant driving” were used to explain them. In the

same volume, Babič [20] presents an ethnographic analysis of traffic as reflected in the media discourse. While Babič's linguistic analysis was qualitative and manual, it could also be quantitative and computational by applying text mining and natural language processing. A combination of the two approaches would likely generate additional insight into the discourse on traffic. Mixed methods work well for various research problems [2, 4], but no research has yet attempted a predominantly quantitative anthropological study. In the paper, we present an experimental quantitative anthropological analysis and explore how to look at car traffic data as a cultural and social phenomenon. Our focus is not on devising new analytical methods but on discovering novel insights in existing data sets. We argue that even a medium-resolution open data set can serve as a helpful starting point for anthropological research.

3 Open data in anthropology

In anthropology, most of the data consist of interviews, field notes, and lived experiences, some of which cannot be translated into tabular data. Some contain personal and sensitive information. Thus, it is no wonder why anthropological open data repositories are rare, but they are not nonexistent.

In 1949, five American universities established the Human Relations Area Files (HRAF), a non-profit organization providing access to major databases on cultural diversity. eHRAF, its main database, is an index of world cultures and human societies based on the identifier codes from the Outline of Cultural Materials (OCM). While eHRAF, the organization's main product with online access to the index, is accessible only with membership privileges, the site also hosts an open database of cross-cultural studies, Explaining Human Culture.² Similarly, AnthroBase³ is a searchable database of anthropological texts, but it is a result of community effort, not institutional engagement.

Both sites offer open access only to texts from previous studies, which is valuable for a meta-review and comparative analyses, but they do not provide data for academic reuse. Archaeology is at the forefront of open data efforts, with several open databases covering biology, geology, archival sources, images, *etc.* Similarly, a metadata search engine Open Language Archives Community (OLAC, [21]) and Digital Research Infrastructure for the Arts and Humanities⁴ provide resources with a specific emphasis on languages and philology. Europeana⁵ is at the forefront of European digitization and open data efforts for cultural heritage. This EU initiative brings together museums, archives, and libraries in a single platform that provides open access to digitized collections. Europeana also provides an API that enables automated browsing and retrieval of data, making it particularly suitable for computational analysis.

The main problems in anthropology concerning open data are privacy issues and sensitive topics [22, 23]. Ethnographic materials often contain personal narratives, with easily identifiable subjects and locations. While this makes data extremely rich in context and thus valuable for qualitative analyses, it also makes it difficult to share. Such data is also hard to anonymize since obfuscation would likely remove too much context and make the

²Explaining Human Culture, <https://hraf.yale.edu/ehc>.

³AnthroBase, <http://www.anthrobase.com/>.

⁴DARIAH, <https://www.dariah.eu/>.

⁵Europeana, <https://www.europeana.eu/en>.

data useless or, conversely, not hide enough information and in this way expose research participants.

Outside of anthropology, there are many services that offer open access data,⁶ for example, OpenDOAR⁷ and Registry of Open Access Repositories (ROAR⁸). Both provide a list of public databases with statistics on the content and upload activity, making it easy to shortlist relevant services. CESSDA⁹ is a database for social sciences, with open access to survey and questionnaire results, demographic data, experiment results, and more. Results of qualitative and multi-method studies are available at the Qualitative Data Repository (QDR¹⁰), whose aim is also to prevent data in social sciences from going to waste.

Most modern governments nowadays offer at least some type of access to public information. Using open data gathered by government entities could be beneficial for a variety of research inquiries, including those from anthropology. In continuation, we present an experiment in which we took publicly available data, quantitatively analyzed them, and tried to interpret the findings from an anthropological point of view. We present some main results and expose the strengths and shortcomings of quantitative analysis in anthropology.

4 Data

The analysis of car traffic flows used the traffic data from the Slovenian Infrastructure Agency [24]. The agency coordinates an extensive network of road traffic counters to measure traffic [25], which are installed on motorways and regional roads. The data is publicly available through an application programming interface (API) providing real-time data and directly from the agency, which stores historical data. We analyzed the historical data for 2015, 2016, and 2017, for which traffic counts were available at hourly intervals. The traffic is reported per type of vehicle. We decided to look at a segment of human mobility, namely cars and motorbikes, excluding buses, lorries, and trucks. Buses were excluded because they follow pre-defined schedules, while freight does not reflect human mobility. Henceforth, when we use the term traffic, we refer to car traffic. Where we refer to motorbike traffic, we state so explicitly. Additionally, traffic counters report traffic for each direction in the case of regional roads and for each lane in the case of motorways, making it necessary to consider the direction of each measurement.

There were 903 traffic counters, each reporting for two directions or lanes. We removed the counters with a significant proportion of missing data, *i.e.*, the counters that did not report data for all twelve months or had a fall-out (all values for a particular month were zero). We ended up with 654 counters with sufficient longitudinal data. For the vast majority of the analysis, we aggregated the data so that each counter reported average traffic frequencies (per month, season, or day of the week). We removed the data in the 10% and 90% quantile when calculating the average, thus excluding extraneous factors that could affect the results. Where we distinguished between weekdays and weekends, we explicitly state it. The weekend days were Saturdays, Sundays, and public holidays. We made this

⁶An overview of OA repositories by the University of Minnesota, <https://libguides.umn.edu/c.php?g=858128>.

⁷OpenDOAR, <http://v2.sherpa.ac.uk/opensoar/>.

⁸Registry of Open Access Repositories, <http://roar.eprints.org/>.

⁹CESSDA, <https://datacatalogue.cessda.eu/>.

¹⁰Qualitative Data Repository, <https://qdr.syr.edu/>.

separation because the traffic patterns were significantly different in these two periods. With all aggregations done, this resulted in several hundred thousand rows of monthly data. To analyze the data by hand, we would have to inspect numerous combinations of counter IDs, days of the week, and directions to find counters with deviant patterns. To avoid manually browsing the counters and arbitrarily deciding what is interesting, we used statistical techniques to measure the deviation of the traffic profile from the baseline.

Sample scripts and workflows are available on GitHub and Figshare (see section “Availability of data and materials”). To replicate the visualizations presented in this study, the workflows can be loaded in Orange [26].

5 Methods

The traffic profile is the distribution of traffic counts over a given period (see Fig. 1a for hourly profiles, and Fig. 2a for monthly profiles). For the needs of this research, we defined the traffic profile deviation of the counter as an unusual monthly, daily (by day of the week), or hourly deviation from the general baseline car traffic trend. Based on initial observations, we defined the baseline car traffic as the mean of the lowest four values that represent the regular daily traffic. Our primary assumption was that car traffic is a reasonably predictable phenomenon that shows distinctive patterns throughout the year. Traffic counters generally showed consistent trends, mainly corresponding to commuter traffic. However, some of them had hourly spikes, seasonal increases, or monthly deviations that

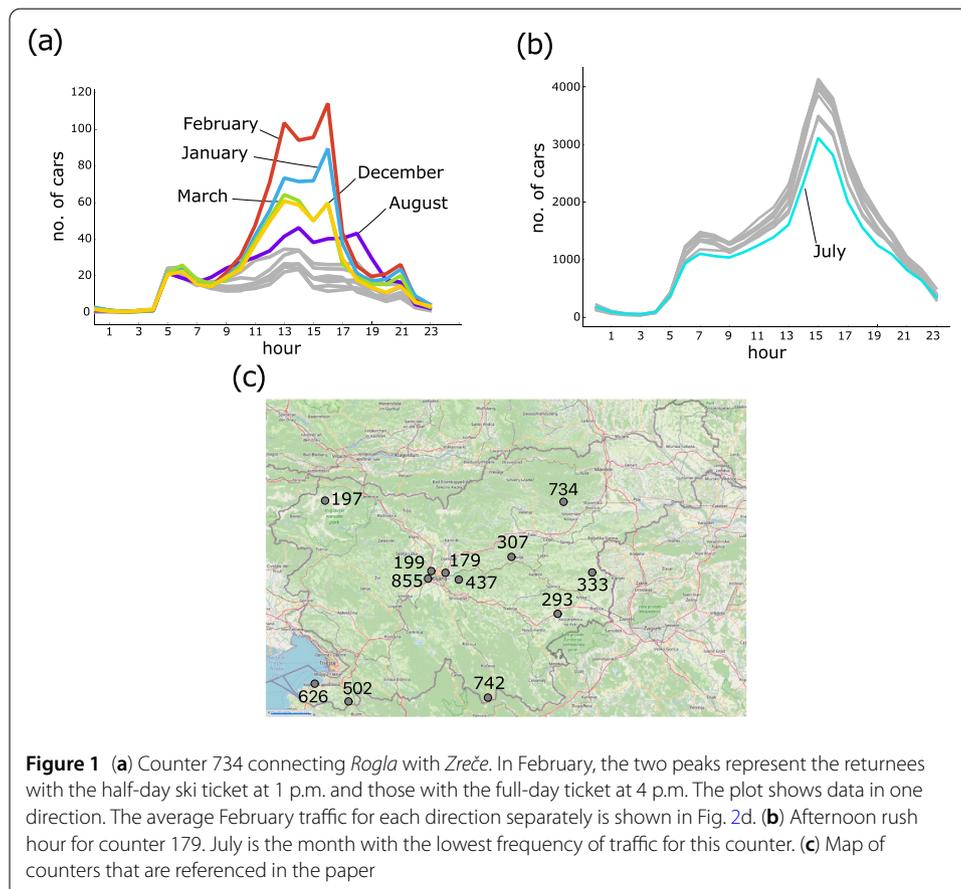
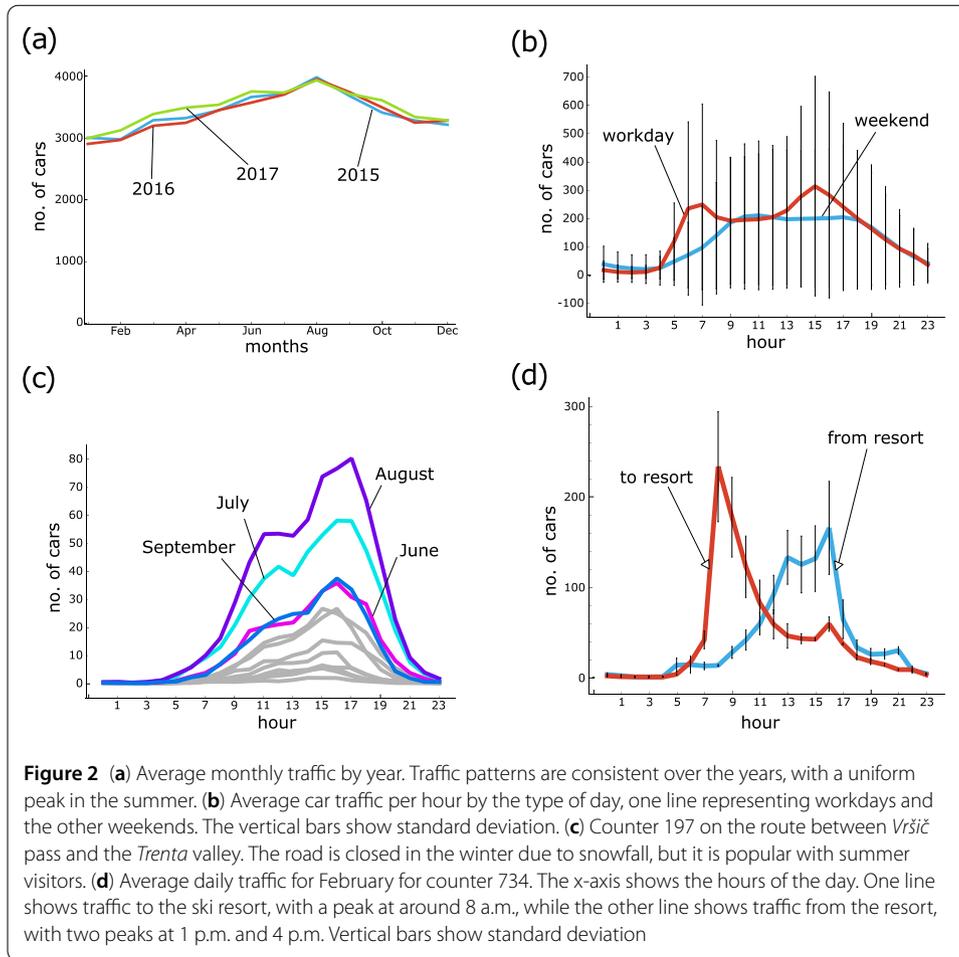


Figure 1 (a) Counter 734 connecting Rogla with Zreče. In February, the two peaks represent the returnees with the half-day ski ticket at 1 p.m. and those with the full-day ticket at 4 p.m. The plot shows data in one direction. The average February traffic for each direction separately is shown in Fig. 2d. (b) Afternoon rush hour for counter 179. July is the month with the lowest frequency of traffic for this counter. (c) Map of counters that are referenced in the paper



need to be explored further. They identified shifts in expected traffic with implications for the immediate locality and sometimes even the entire country.

The deviation of the traffic profile is best seen from the line plot, where each line represents the average car traffic for a given period, and the x-axis indicates the hour of the day (Fig. 1a). The deviation scores correspond to the visual information for each image. The plot is considered deviant: if there is a considerable dispersion of lines, which is typical for seasonal trends; if there is a single outlying line, which is typical for road work; or if there is a peak at a particular hour, which indicated individual events, such as festivals, road works, etc. We defined and applied five statistical measures to score the counters, each corresponding to a particular type of deviation:

- *A: absolute deviation from the baseline.* The absolute deviation from the baseline takes the lowest four values as the baseline commuter car traffic and then computes the difference between the individual car traffic profile and the baseline. The score finds popular destinations and transit areas.
- *B: relative deviation from the baseline.* The relative deviation is similar, but it looks at the ratio of difference. The score finds locally popular destinations.
- *C: coefficient of variation.* The coefficient of variation [27] is a standardized measure of the dispersion of the frequency distribution. In our case, it measures how far apart from each other the car traffic profiles are. Such a measure would capture counters

with high variability across the data. The score finds counters with significantly different profiles, for example, weekly patterns.

- *D: total difference from the baseline.* The difference from the baseline looks at each hour, computes the average car traffic, and then sums the differences of the profiles to the baseline. The final score is the sum of these differences, so such a score would find counters with several high spikes. The score finds both positive and negative deviations, finding, for example, seasonal changes and changes in traffic infrastructure.
- *E: adjusted z-score.* Finally, the adjusted z-score normalizes the data to have zero mean and a variance of one. However, we replaced the mean with the baseline in our case, thus standardizing the data to the assumed everyday car traffic. Z-score finds counters with singular deviations. The score finds smaller peaks in the profiles, for example, deviations related to local festivities.

Finding deviant car traffic profiles is not just an exercise in data handling. Deviant profiles show relations between the drivers and their destinations. In his account of the social life of a road between Albania and Greece, Dalakoglou [28] shows how roads embody both the material aspect and the socio-cultural transformation these technological objects bring to the people. However, the author's choice of the ethnographic site required previous knowledge of the importance of this particular section of the road. Measures of deviation mitigate the need for insider knowledge by providing a way to detect potentially relevant sites objectively and exhaustively.

6 Results

We used the above scores to find patterns in car traffic data. For each score, we proposed a research question that could be explored based on the information obtained. For each method, we sorted the counters by the scores and analyzed the ten unique counters with the highest ranking. We used Jupyter notebooks and Orange [26] to generate visualizations. After finding deviant counters, we performed an extensive online archive search to explain why such patterns occur. The patterns began to fall into several categories, creating an image of trends and particularities of the Slovenian road traffic system.

6.1 High-frequency counters

Our main goal was to find deviant traffic counters, where deviant means a high traffic volume that can be seen as a total increase per month or a spike at a particular hour of the day. The first task was to find high-volume traffic counters by sorting the traffic counts for each direction and observing the highest counters. Unsurprisingly, these were the counters on the ring road around the capital that detect commuters (counters number 179, 199, 855, see Fig. 1c). Their monthly averages also displayed a very predictable and strict weekday schedule, with incoming spikes around 7 a.m. and outgoing spikes around 3 p.m., which designate rush hours. In contrast, weekends showed a greater dispersion of the daily spike and greater variability at different hours of the day. The contrast between the weekend values of counters 179 and 855 is quite stark, with July being the lowest month for 179 and one of the highest for 855 (Fig. 1b). That is because counter 179 is on a road leading outside of the capital and the city center has a low traffic volume in summer. Conversely, counter 855 is located on a major motorway connecting Austria with Croatia and Italy, hence the increase in July and August.

High-frequency counters show popular transit areas or final destinations, reflecting general trends in the country's traffic flow. Highways are transport lifelines, as most commuter traffic passes at least a section of them. The capital, for example, records around 120,000 commuters daily [29], which amounts to 6% of the total population of the country. The initial infrastructure built to facilitate connections between regional capitals, such as Zagreb-Graz and Trieste-Budapest, serves many commuters who come to work in the capital. Due to the dependence of the traffic on highways and the country's position at the top of the Adriatic, the strain on the traffic is particularly dire in the summer. Local traffic is then forced to share the roads with neighboring tourists. Based on this information, the researcher could conduct fieldwork with the commuters to observe in detail the decision-making regarding the mode of transport, the route, and the time of departure.

6.2 Seasonal increase and decrease

Two main motorways criss-cross Slovenia in an X pattern, connecting Hungary with Italy and Austria with Croatia. The intersection is in Ljubljana, with its ring road also serving daily migrants coming to work in the capital. As a transit country, Slovenia experiences high traffic in the summer, mainly in July and August (Fig. 2a), as its northern neighbors take the journey south to Croatia.

The morning rush hour lasts from 6 a.m. to 8:30 a.m., while the afternoon one starts at 2 p.m. and ends around 4 p.m. (Fig. 2b). Rush hour is particularly evident on Ljubljana's ring road, with high incoming traffic in the morning and a high outgoing one in the afternoon. Analyzing rush hours uncovers typical commuting patterns. The workday patterns show the country's economic hubs and when most of the workforce goes to work. Slovenia is still fairly centralized, with Ljubljana registering the highest proportion of workday traffic, with small increases in *Kranj*, *Koper*, and *Novo mesto*. *Maribor* is different, as many inhabitants work in neighboring Austria. The rush hours are conservative, starting and finishing early in the day [30].

However, we aimed to go beyond typical car traffic flows and discover regions and locations where traffic is different from usual. We were interested in seasonal increases, short spikes, and outlying profiles. We wanted to map the landscape of car traffic flows quantitatively and qualitatively, and elicit patterns of behavior that exhibit collective preferences and habits.

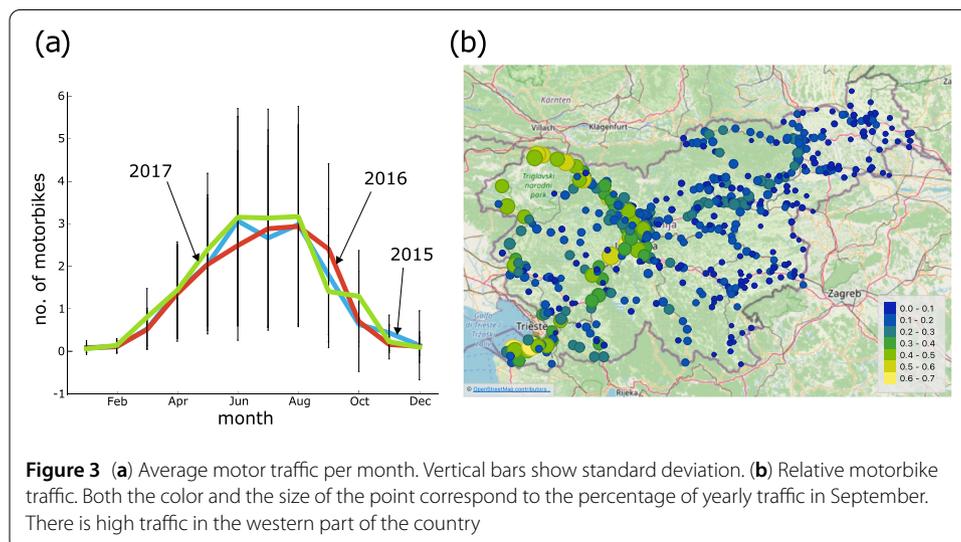
For seasonal increases, we first grouped the data into seasons. March, April, May represented spring, June, July, August represented summer, September, October, November represented autumn, and December, January, February represented winter. The final score was the difference to the mean hourly traffic. In other words, if a counter reported 30% of annual traffic in the winter, but the overall winter traffic was 26%, the score would be 4 percentage points (0.04). Finally, we ranked the counters according to their deviation magnitude as estimated by the difference to the mean and took the top 10 results.

In north-west was counter 197, counting traffic on the road from *Vršič* pass to the *Trenta* valley. This counter had a particularly low baseline, with very little traffic recorded over the winter months (Fig. 2c). The drop was expected, as *Vršič* is Slovenia's highest mountain pass, which is closed in the winter due to heavy snowfall. Considering it is not a vital route connecting villages or cities for the locals, it perfectly reflects the tourists' behavior. The main peaks are around 12 a.m. and 5 p.m., showing that tourists are not early birds and prefer to finish their trip before nightfall. The traffic profile shows incoming traffic from

Vršič to Trenta in the morning and returning traffic in the afternoon. While we cannot claim with certainty that these are the same cars, most people likely decide not to stay in the beautiful but quite remote *Trenta* with fewer capacities. They instead return to *Kranjska Gora*, the local tourist hub.

Counter 197 was consistently ranked at the top for the most deviation measures, namely seasonal deviations and C and D scores. Different scores revealed different information. Score D, for example, found counters with high relative deviation from the mean, where the score was the sum of absolute differences for each hour. This measure considered both positive and negative deviations by summing the absolute values. Moreover, each deviation from the mean added to the score, ranking highly the counters with several deviant months. One such counter is 734, which connects the ski resort *Rogla* with the spa town *Zreče* (Fig. 2d). The increase in traffic is highest in February and January, with significant deviations in March and December as well. Both the deviations and the direction of the traffic show that *Rogla* is predominantly a winter destination. Traffic peaks at 8 a.m. going to the ski resort, and it peaks at 1 p.m. and 4 p.m. going away from the resort. The two peaks in the afternoon correspond to two types of ski passes (half-day and full-day), each expiring exactly at the time of the traffic peak. The reason why February is the most popular month is not only due to the amount of snow but also due to the winter school holidays, which occur at that time. While this destination is a typical winter destination for domestic tourists, it has recently expanded the summer offer, evident from the increase in traffic flow in August.

Seasonal patterns are distinctive for two other entities, motorcycle riders as a type of traffic and border crossings as a type of counter. Motorcycle riders begin their season in March when there is a first significant increase in motorcycle traffic. The trend increases until June and falls slightly in July, probably because few people travel to the seaside by motorbike. The season reaches its peak in August (Fig. 3a and 3b), then falls towards the end of the year. We assume motorcycle riding is a seasonal leisure activity, with a slight hiatus during the summer holidays. As for the border crossings, they are, compared to other counters, at their highest in July and August. The increase starts in June at Slove-

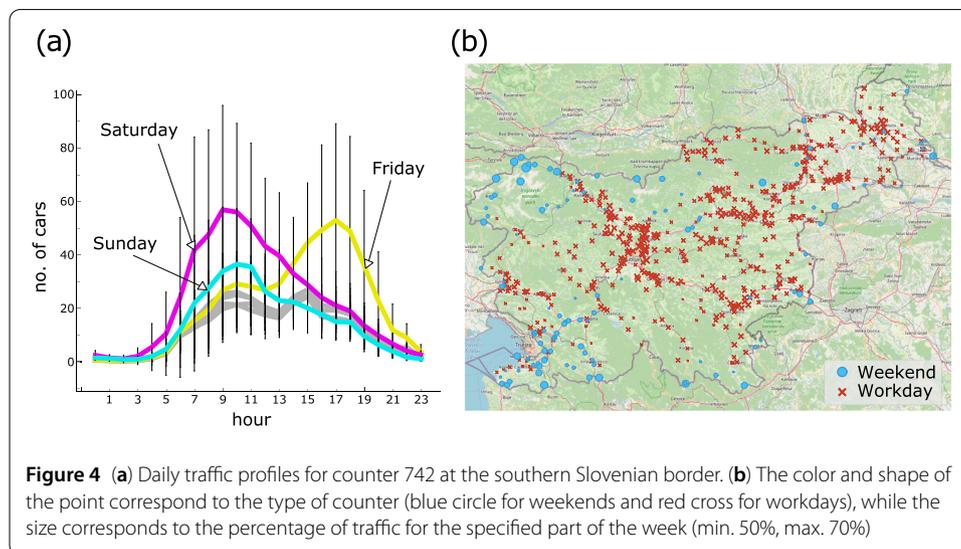


nia’s north-eastern (Hungary) and south-western (Italy) borders, with a similar pattern occurring in September.

Looking at seasonal peaks, we establish that tourists generally stick to major traffic routes, travel outside the main rush hour and frequent different destinations. The western part of the country is popular for tourists, especially motorcyclists, with the eastern part slowly gaining ground (particularly with tailored seasonal offers and investments in infrastructure). Local tourism is still prominent on holidays but tied to traditional destinations (*i.e.*, ski resorts in winter and Croatia in summer). The researcher could use this information to target particular groups, such as motorcyclists directly in the localities they frequent. The participants can be asked about their choice of destination, where they heard about it, whether they considered visiting the eastern part of the country, and so on.

6.3 Weekly patterns

After considering seasonal fluctuations, we also had a look at weekly patterns. Instead of grouping counters by month or season, we grouped them by the day of the week. We were interested in how counters differ by daily averages, so we computed C (coefficient of variation) and D (difference from the baseline) scores for daily deviations. The top-ranked counters were two border crossings, one in *Sočerga* (ID 502) in the southwest and the other connecting *Fara* with *Petrina* (ID 742) in the south (Fig. 4a). These two border crossings had a strikingly similar pattern – people leaving the country on a Friday evening or Saturday morning and returning on a Sunday evening. According to unofficial data, Slovenians had around 110,000 properties in Croatia in 2014,¹¹ which explained the weekend trips across the border. When observing the raw time series for the mentioned counters, we noticed that the traffic peak slowly shifts from Friday to Saturday towards the height of the summer and then back to Friday in September. The peak is likely foreign traffic since tourists from the north start their journey on a Friday and arrive at the Slovenian-Croatian border on Saturday morning. Friday traffic is predominantly due to Slovenian tourists. The researcher could base herself at the very border, interviewing Friday travelers about their



¹¹Slovenians’ Real Estate in Croatia, <https://www.dnevnik.si/1042657187>.

destinations and choice of holiday time frame, and can observe the composition of the passengers.

Since traffic patterns seemed to be highly related to the day of the week, we plotted a map (Fig. 4b) of counters, where we coloured each counter with the part of the week, when the counter recorded the highest proportion of traffic. In other words, if the counter registered more traffic during the weekend than during the week, say in a 60:40 ratio, the counter was tagged as a “weekend counter,” and its size was 0.6. The map shows typical weekend destinations with higher traffic during the weekend than during the week. The western part of the country seems popular for weekend trips, while the capital gets a lot of commuter traffic.

6.4 Changes in traffic infrastructure

Most of the highest-ranked counters have increased traffic in the summer. These counters are popular tourist spots in the mountains and near the sea. Score C highly ranks a counter, which has a spike in the spring. Counter 626 is located at the Slovenian coast, specifically on the scenic route between *Koper* and *Izola*. The route was built in 1837 and is locally known as *Riva lunga* (Long coast). In March 2017, it was permanently closed to car, bus and freight traffic, but it was among the busiest ones before that. Considering the route’s popular location, it was unclear why there was so little traffic in the summer months years before its closure. However, on 5 June 2015, the tunnel *Markovec* in the hinterland was finally opened,¹² enabling the traffic to bypass the coastal road. Longitudinally, this shows significant infrastructural changes relevant both for locals and tourists.

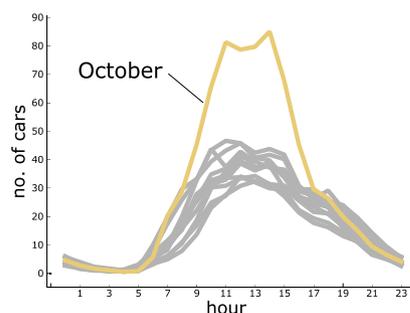
Changes in traffic infrastructure are indicators of economic change, while commuter flows mark the level of metropolization of the city. As new roads are built, new economic centers emerge, and old ones are abandoned, contributing to the region’s decentralization [31]. It marks the shift at an infrastructural and social level, with changing communication and power relations. The example of *Markovec* tunnel showcases the magnitude of change when new roads take over the old ones. It also shows how such a transition can be beneficial for regional development. The new road now carries the burden of traffic (commuters, freight, tourists). In contrast, the local community re-appropriated the old road to become a popular strolling path and a tourist attraction in its own right. Finding this shift in road traffic pinpoints a starting point for future anthropological research – how were the locals socially and economically affected by the construction of the *Markovec* tunnel?

6.5 Local festivities

Most of our exploration focused on observing deviations from the mean, either relative or absolute. Deviations are partially reflected in the coefficient of variation, which measures how dispersed the profiles are relative to the mean. We also wanted to observe smaller peaks in the profiles, so we decided to compute an adjusted z-score, where instead of the deviation from the mean we observed the deviation from the baseline. Z-score [27] is a type of normalization, which puts the data on the same scale. In this way, we bypass the most frequented counters and observe local particularities. We decided to compute the z-score for each hour of each month as the deviation from a baseline. We considered only

¹²Building the *Markovec* tunnel, <https://www.rtvsllo.si/slovenija/foto-gradnja-predora-markovec-je-trajala-predolgo/366795>.

Figure 5 A significant increase in traffic is evident in October when *Šmartno pri Litiji* with counter ID 437 holds the annual Chestnut festival



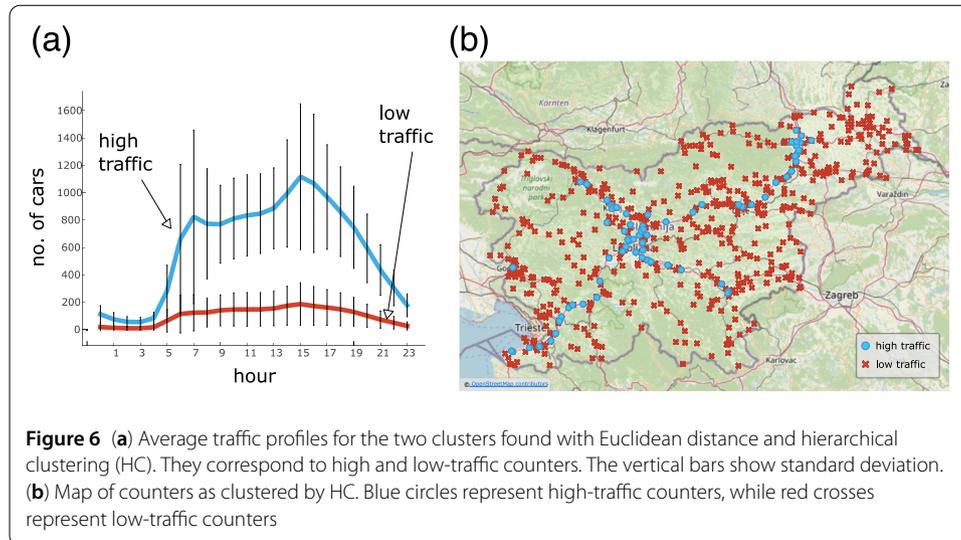
the baseline to observe the deviations, not the actual mean. Finally, we ranked the counters by the highest z-score per month.

Adjusted z-score successfully identified counters with daily peaks that were significantly different from the general traffic distribution. Thus we pinpointed the chestnut festival near *Šmartno pri Litiji* (counter 437), the festival of *Kozjansko* apple (counter 333), and the celebration of Saint Stephen (counter 293). Identifying and measuring the popularity of these local festivities enables the researcher to understand the attractiveness of local events and the importance of regional tourism. In continuation, we present one such event, namely the Chestnut festival.

The highest-ranked counter was 437 between *Zadvor* and *Šmartno pri Litiji*. The latter is a small town, while the former is a suburban section of Ljubljana. The road goes through a quaint little valley full of farms and houses. It is far from a major traffic route, so why the observed increase in traffic? Furthermore, is there a specific time in which it occurred? Looking at the graph (Fig. 5a), a high z-score occurs in October. The increase happens on the weekend and follows a typical incoming-outgoing pattern. People seem to drive from the direction of Ljubljana towards *Šmartno pri Litiji* between 11 a.m. and 2 p.m., while they all seem to be returning around 5 p.m. The counter does not display such behavior for other months or even for working days in October. The reason for this increase is the annual festival of chestnut,¹³ held in the hill-top village of *Janče*. Inhabitants of the capital seem drawn to the nearby celebration of this autumn delicacy, showing the importance of regional festivities for domestic tourism, particularly outside of the summer season.

The findings show the importance of domestic flows for smaller, lesser-known points of interest, which cannot compete with traditional tourist hot spots in the main season. In their modern interpretation, local festivities are a part of the cultural heritage, which benefits the locals socially and economically. As Poljak Istenič [32] argues, the revitalization of these festivals is a good practice for the sustainable development of the peripheral and rural communities. The researcher could use this information to plan the fieldwork, specifically targeting the proposed festivals for observation. Future research could focus on the role of off-season festivities for local communities, with a case study of one of the proposed counters. The computational analysis thus pinpoints relevant research questions and indicates potential field sites. However, the quantitative analysis does not answer why tourists visit these particular festivals, who visit them, and their social and cultural role in local life.

¹³Chestnut Festival, <https://www.dnevnik.si/1042787995>.



6.6 Discovering traffic profiles

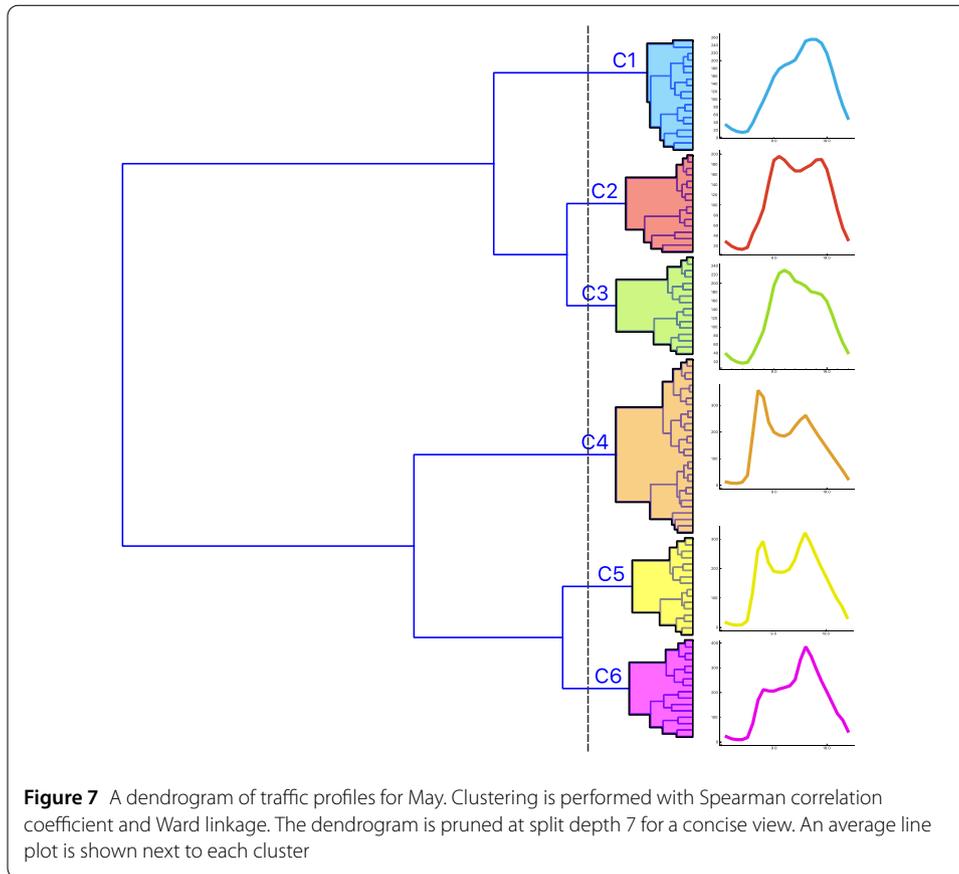
In the second part of the analysis, we used data clustering to discover typical car traffic patterns in the country. We took the data for May, as this is the month before the main tourist season, but still popular with the local tourists. The data reported the percentage of traffic for each counter per hour, with a distinction between weekends and weekdays and the direction of traffic.

We used hierarchical clustering (HC) with Euclidean distance and the Spearman correlation coefficient. We used the Ward linkage and estimated the number of clusters from the dendrogram. Finally, we grouped the counters by cluster labels to compute the average cluster profile and plotted the results in a line plot.

HC with Euclidean distance found two distinct clusters, which corresponded to the high-traffic and low-traffic counters (Fig. 6a). The result was unsurprising since the magnitude of the measured phenomena heavily influenced Euclidean distance calculation. The high-traffic counters were located on the so-called highway cross of the country, the four main motorways joining in the capital (Fig. 6b).

HC with the Spearman correlation coefficient gave us a different image. We used the Spearman correlation coefficient, which measured how similar the shapes of the profiles of traffic data were [33]. The sensible cutoff was between two and six clusters. The six clusters corresponded to the workday and weekend profiles and the morning and afternoon rush hour (Fig. 7). Setting the cutoff at two clusters corresponded to the workday and weekend clusters, meaning there was a significant difference between the two periods. With three clusters, the workday cluster is further split into two groups: the morning and the afternoon rush hour. With four clusters, the weekend profiles were further split into morning and evening traffic.

The first cluster (C1, Fig. 7) contained predominantly weekend profiles (76.16% weekend, 23.84% workday), with a slow increase in traffic between 4 p.m. and 5 p.m. The second cluster (C2) was almost exclusively a weekend cluster (98.92%) and had higher traffic in the early morning. The third cluster (C3) was also a weekend cluster (95.79%), but in this case, the traffic was higher in the wee hours, namely, at midnight and one and two in the morning. The fourth cluster (C4) contained almost exclusively workday profiles (99.78%),



including the morning rush hour, with higher values for 5 a.m., 6 a.m., and 7 a.m. The fifth cluster (C5) was exclusively a workday cluster (100%), with a morning and afternoon spike. The sixth cluster (C6) was also a workday cluster (93.76%), with higher traffic between 2 p.m. and 4 p.m. – the afternoon rush hour.

The results were very similar for all other months. Winter months had less distinct weekend traffic patterns and more distinct patterns during the week. Interestingly, December had a high afternoon peak, which contained about half of the workday and half of the weekend counters, reflecting the afternoon errands and festivities typical for this month. The weekend afternoon spike shifted to later hours in the summer, while the morning spike shifted to earlier hours and became a separate cluster. The morning cluster predominantly consisted of the tourist traffic that transits Slovenia. Finally, the average weekend traffic in July and August overtook the average workday traffic.

When selecting three clusters, regardless of the month chosen, the most typical traffic patterns were the workday *vs.* weekend profiles and the afternoon rush hours. The workday patterns were relatively typical, with spikes at 6 a.m. and 3 p.m. The weekend profile climbed later, at around 10 a.m., continued throughout the day, and dropped at around 4 p.m. The distinctive weekend counters are at the border and more frequently in the western part of the country. Slovenians seem to travel across the border for the weekend frequently. If they stay in the country, the *Primorska* region appears to be most popular (same is true for tourists coming into Slovenia).

7 Discussion

In the paper, we show how to use open public data for anthropological research. Specifically, we show how data science and computational approaches such as data mining and machine learning can be used for finding interesting patterns of human behavior, which are then qualitatively interpreted and explained. The focus is not on data mining and machine learning, which have already proven valuable in anthropological research [2, 4, 34, 35], but on the potential of quantitative analysis in anthropology and the rich information hidden in public data sources.

Simple statistical methods are great for initial data exploration and uncovering interesting patterns. Certain traffic profile deviation scores reveal general traffic trends (scores A and B). Others detect local deviations (C, D, and E scores), such as road work, seasonal festivities, and popular tourist spots. The data is recorded at hourly intervals, so it is easy to observe trends for a single traffic counter or group in different periods. The ability to traverse between different levels of granularity is one of the key benefits of using sensor data for anthropological research. Identifying outlying patterns enables preliminary analysis of the phenomenon (in this case, traffic) and pinpointing potentially relevant field sites.

For a comparative analysis of car traffic patterns, we used clustering. Cluster analysis shows that traffic patterns differ primarily by day of the week and time of day. Patterns are not location-specific, *i.e.*, there is no region where a specific clustering pattern would be predominant. The only exceptions are the western part of the country for motor traffic and the border counters with increases at the end of the week. This approach showed the general structure of the road network, which is the basis for understanding the flow of traffic in a country, not only materially and logistically, but also socially. Clustering reveals patterns of frequent behavior, categorizes human practices, and enables subsequent comparative analysis.

Analysis of road car traffic counters revealed a significant distinction between work-day and weekend traffic patterns. Most people seem to work predominantly from about 7 a.m. to 3 p.m. For the weekend, they go for a trip at around 10 a.m., then stop for lunch at around 1 p.m., and return home by 5 p.m. Many people travel to Croatia for the weekend, leaving on Friday evening or Saturday morning and returning on Sunday evening. Slovenia is a transit country with an increase in car traffic on the highways in summer. That said, summer tourism is well-established in the western part of the country. Some winter destinations are being transformed into year-round destinations, as is the case with the *Rogla* ski resort. Off-season activities are also popular with the locals. In October, people attend different autumnal events that celebrate local produce and traditions, while in December, there is an emphasis on religious festivities and pre-New Year's Eve celebrations.

The present study has its limitations. The data is not related to individual activity but to the aggregation of local mobility patterns, making it impossible to observe individual behavior patterns. It is also impossible to completely discount extraneous factors affecting counter data (such as road closures or local points of interest). These factors were already partially handled by aggregating over a longer time (3 years), but the data could be expanded to an even longer time frame. In some specific cases, however, over-zealous aggregations discount small local events such as the celebration of St. Stephen. Aggregations thus have to be handled case-by-case.

With the identification of general, specific, and local car traffic properties, we have shown a glimpse into public life as seen from the road. The anthropological study of traffic

should highlight universals in behavior as well as cross-cultural differences in intention, interpretation, interaction, and management of risk [36]. This study was able to pinpoint certain universals if we consider universals as general car traffic patterns found throughout the world [37, 38]. The study also revealed intentions and interactions by studying deviations from the general patterns.

However, to understand human mobility, we must also look at cyclists, public transport, pedestrians, and other modes of transportation, including water and air. The road traffic data at our disposal did not include these modes of transport. We would have to use additional tracking systems to acquire such data and not rely solely on publicly available data.

8 Quantitative anthropology as the future?

With the broad availability of open data, it is easier than ever to analyze large amounts of data on human practices computationally. Quantitative data reveals the structure of the phenomena, temporal changes, and interesting outliers, all highly relevant for describing a phenomenon or a community in detail. It shows where, when, how often, and by whom a given phenomenon occurs. Nevertheless, computational analysis insufficiently addresses some of the key questions of anthropology – why and how? It does not offer detail-rich descriptions and context, and it does not and cannot relate correlations to the cause. That said, computational anthropological research is a great starting point and can provide a relevant glimpse into the life of a community.

Quantitative analysis of mobility data provides information on general mobility patterns and seasonal trends. It also reveals outliers, which can serve as a starting point for forming research questions. Detecting interesting locations can also help the researcher narrow down potential research sites. Finally, ethnographic observations of mobility can be supplemented with findings from quantitative studies. Interdisciplinary approaches often provide a richer picture of a phenomenon than a single method [39]. In summary, quantitative analyses for anthropology can: (1) enable traversing between different levels of granularity, (2) pinpoint relevant field sites in the preliminary analysis, (3) show a general structure of the phenomenon, (4) aid in categorizing human practices, (5) enable comparative analysis.

Moreover, open data are a great starting point for quantitative research. Such data can hold information on people's behaviors, practices, and habits. One of the most attractive opportunities in the age of digital data is to find quality in quantity, or in other words, reflections of social structures, cultural patterns, norms, and values in numerical data. Open data are essentially archival. We argue that it benefits research by studying the community in a broader context, detecting rare events, observing temporal cycles and shifts, and mitigating biases.

Analysis of traffic data revealed several strengths and weaknesses of quantitative approaches for anthropological analysis. Computational techniques help to find general patterns in large data sets and interesting deviations [4, 34]. Large data sets can be analyzed on several levels. One level is the overview, taking the entire data set and observing its properties. Such analysis reveals typical behaviors, relations, and hierarchies of the population. In the case of traffic data, the overview would entail nationwide traffic patterns, seasonal fluctuations, and clustering of traffic profiles. Another level is the midway analysis, with the observation of data segments, sub-populations, regions, *etc.* An example from the traffic data analysis would be comparing different parts of the country, looking at the

inbound-outbound traffic of a city, or observing specific periods. The most detailed level is the granular analysis of individual data points or categories. These different levels of analysis are not specific to quantitative data. However, they are much more pronounced in this case since traversing between levels must be explicitly encoded in the data analysis procedure. Finally, quantitative analyses are great at answering “where,” “what,” “when,” and “who.” They provide procedural information by utilizing large data sets of geo-spatial, temporal or personal data.

However, determining “how” and, most importantly, “why” with quantitative data is more challenging. Traffic data, for example, is not very rich in detail. It is aggregated at a 15-minute interval, bound to a specific counter, and contains only indirect data on human habits. To answer “why” a particular behavior occurred, we had to resort to archive data and newspaper clips. The “how” was partially reflected in the routes taken and the time driving occurs, but again, the aggregated data could not accurately reflect individual variability. Generally, qualitative analyses provide explanatory information, typically from smaller samples.

Consequently, we should substantiate open data with qualitative data to achieve a valid ethnographic account and provide a detail-rich explanation of traffic practices. Quantitative analysis does pinpoint broad behaviors that serve as a starting point for further research. However, the process should be explained ethnographically by connecting the materiality of the infrastructure with social actors, that is, the people behind the process.

Anthropological research using open public data sets is not meant to replace ethnographic fieldwork but should be considered a complementary venue with numerous possibilities. It is mainly appropriate for preliminary analysis, as shown by identifying deviant traffic counters. Moreover, it provides insight into the structure of the phenomenon or a system. It shows when an event occurs, where, and how. In other words, it describes the phenomenon in terms of its properties, but, as explained above, it cannot provide the context and the reason why something happens. Nevertheless, as computational analyses are time and resource-efficient and reveal relevant social and cultural aspects, they are a valuable research tool.

Acknowledgements

The authors thank Urška Starc Peceny and Samo Eržen for feedback on the research questions and design addressed in this paper.

Funding

This work was supported by the Slovenian Research Agency [P2-0209] and co-funded by the Slovenian Ministry of Education, Science and Sport, and the European Regional Development Fund [Tourism 4.0 – enriched tourist experience (OP20.03536)]. The funding bodies did not influence the design of the study, the collection, analysis, and interpretation of data, or the writing of the manuscript.

Abbreviations

API, Application Programming Interface; CESSDA, Consortium of European Social Science Data Archives; eHRAF, Human Relations Area Files; EU, European Union; GPS, Global Positioning System; HC, hierarchical clustering; OCM, Outline of Cultural Materials; OLAC, Open Language Archives Community; OpenDOAR, Open Directory of Open Access Repositories; QDR, Qualitative Data Repository; ROAR, Registry of Open Access Repositories.

Availability of data and materials

Jupyter notebook scripts, Orange workflows, and data that support the findings of this study are available on GitHub (<https://github.com/ajdapretnar/traffic-flows>) and in Figshare (<https://doi.org/10.6084/m9.figshare.13259093>).

Declarations

Competing interests

The authors declare that they have no competing interests.

Author contributions

APŽ and TC designed the study. APŽ and TC acquired the data. APŽ wrote most of the data analysis scripts. APŽ, TH and TC analyzed the data. APŽ wrote the first draft. APŽ, TH and TC wrote and reviewed the manuscript. All authors read and approved the final manuscript.

Received: 17 December 2021 Accepted: 31 May 2023 Published online: 11 August 2023

References

1. Anderson K, Nafus D, Rattenbury T, Aipperspach R (2009) Numbers have qualities too: experiences with ethno-mining. In: EPIC 2009 proceedings 2009(1), pp 123–140. <https://doi.org/10.1111/j.1559-8918.2009.tb00133.x>
2. Blok A, Pedersen MA (2014) Complementary social science? Quali-quantitative experiments in a big data world. *Big Data Soc* 1(2):1–6. <https://doi.org/10.1177/2053951714543908>
3. Bornakke T, Due BL (2018) Big – thick blending: a method for mixing analytical insights from big and thick data sources. *Big Data Soc* 5(1):1–16. <https://doi.org/10.1177/2053951718765026>
4. Pretnar A, Podjed D (2019) Data mining workspace sensors. *Contrib Contemp Hist* 59(1):179–197
5. Cury M, Whitworth E, Barfort S, Bochereau S, Browder J, Jonker TR, Kim KS, Krenchel M, Ramsey-Elliott M, Schüür F et al (2019) Hybrid methodology: combining ethnography, cognitive science, and machine learning to inform the development of context-aware personal computing and assistive technology. In: *Ethnographic praxis in industry conference proceedings*, vol 1. Wiley, New York, pp 254–281. <https://doi.org/10.1111/1559-8918.2019.01284>
6. Średniawa B (1981) The anthropologist as a young physicist: Bronisław Malinowski's apprenticeship. *Isis* 72(4):613–620
7. Spradley JP (1980) *Participant observation*. Cengage Learning, Wadsworth
8. Csordas TJ (1990) Embodiment as a paradigm for anthropology. *Ethos* 18(1):5–47
9. Kuipers G (2013) The rise and decline of national habitus: Dutch cycling culture and the shaping of national similarity. *Eur J Soc Theory* 16(1):17–35
10. Podjed D, Bezjak S (eds) (2017) *Research on the road: methodology and practice of studying traffic, driving habits, and sustainable mobility*. ZRC SAZU, Institut za slovensko narodopisje
11. Horta G (2019) From peripatetic anthropology to the ethnography of roads and motorisation in Africa. *Cad Estud Afr* 37:213–233
12. Estevan AE (1994) Contra transporte, cercanía. *Archipiél Cuad Crit Cult* 18:33–42
13. Dant T (2004) The driver-car. *Theory Cult Soc* 21(4–5):61–79
14. Palmer JR, Espenshade TJ, Bartumeus F, Chung CY, Ozgencil NE, Li K (2013) New approaches to human mobility: using mobile phones for demographic research. *Demography* 50(3):1105–1128
15. Gallotti R, Bertagnolli G, De Domenico M (2021) Unraveling the hidden organisation of urban systems and their mobility flows. *EPJ Data Sci* 10(1):1
16. Aslam NS, Cheng T (2018) Smart card data and human mobility. In: *Consumer data research*, p 111
17. Zhao C, Zeng A, Yeung CH (2021) Characteristics of human mobility patterns revealed by high-frequency cell-phone position data. *EPJ Data Sci* 10:5. <https://doi.org/10.1140/epjds/s13688-021-00261-2>
18. Huang Q, Wong DWS (2015) Modeling and visualizing regular human mobility patterns with uncertainty: an example using Twitter data. *Ann Assoc Am Geogr* 105(6):1179–1197
19. Podjed D (2017) Augmented ethnography on the road. In: Podjed D, Bezjak S (eds) *Research on the road: methodology and practice of studying traffic, driving habits, and sustainable mobility*. ZRC SAZU, institut za slovensko narodopisje, pp 15–41. Ljubljana
20. Babič S (2017) Discovering traffic phenomena through language analysis. In: Podjed D, Bezjak S (eds) *Research on the road: methodology and practice of studying traffic, driving habits, and sustainable mobility*. ZRC SAZU, institut za slovensko narodopisje, pp 129–150. Ljubljana
21. Hughes B, Kamat A (2005) A metadata search engine for digital language archives. *D-Lib Mag* 11(2):6
22. Boyd D, Crawford K (2012) Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon. *Inf Commun Soc* 15(5):662–679
23. Pink S, Lanzani D (2018) Future anthropology ethics and datafication: temporality and responsibility in research. *Soc Media Soc* 4(2). <https://doi.org/10.1177/2056305118768298>
24. MZI DRSI (2021) *Promet 2015-2017, CD-ROM*. Ministrstvo za infrastrukturo, Direkcija Republike Slovenije za infrastrukturo, Ljubljana. <https://www.gov.si/zbirke/storitve/narocilo-publikacije-ozroma-zgoscenke-promet>
25. Bickel PJ, Chen C, Kwon J, Rice J, Van Zwet E, Varaiya P (2007) Measuring traffic. *Stat Sci* 22(4):581–597
26. Demšar J, Curk T, Erjavec A, Črt Gorup, Hočevar T, Milutinovič M, Možina M, Polajnar M, Toplak M, Starič A, Štajdohar M, Umek L, Žagar L, Žbontar J, Žitnik M, Zupan B (2013) Orange: data mining toolbox in python. *J Mach Learn Res* 14:2349–2353
27. Everitt B, Skrondal A (2010) *The Cambridge dictionary of statistics*. Cambridge University Press, Cambridge
28. Dalakoglou D (2010) The road: an ethnography of the Albanian–Greek cross-border motorway. *Am Ethnol* 37(1):132–149. <https://doi.org/10.1111/j.1548-1425.2010.01246.x>
29. Bajuk Senčar T (2017) Kultura udobja in razvoj trajnostne prometne infrastruktore. *Glas SED* 57:66–74
30. Anttila T, Oinas T, Tammelin M, Nätti J (2015) Working-time regimes and work-life balance in Europe. *Eur Sociol Rev* 31(6):713–724. <https://doi.org/10.1093/esr/jcv070>
31. Zdanowska N (2015) Metropolisation and the evolution of systems of cities in the Czech Republic, Hungary and Poland since 1950. *Detoupe – Cent Eur J Reg Dev Tour* 7(2):45–64
32. Poljak Istenič S (2013) Družbeno ustvarjanje praznikov: med dediščino, identiteto in trajnostnim razvojem. *Etnol* 23:27–48
33. Ye J, Xiao C, Esteves RM, Rong C (2015) Time series similarity evaluation based on Spearman's correlation coefficients and distance measures. In: Qiang W, Zheng X, Hsu C-H (eds) *Cloud computing and big data*. Springer, Cham, pp 319–331
34. Krieg LJ, Berning M, Hardon A et al (2017) Anthropology with algorithms? An exploration of online drug knowledge using digital methods. *Med Anthropol Theory* 3:21–52. <https://doi.org/10.17157/mat.4.3.458>
35. Pretnar A, Žagar L, Zupan B, Podjed D (2018) Power of algorithms for cultural heritage classification: the case of Slovenian hayracks p 218. *AIUCD*, 2018

36. Rosin TR (2003) Vishnu and the art of motorcycle driving: toward an anthropology of traffic. *Kroeber Anthropol Soc Pap* 89/90:55–74
37. Pasaoglu Kilanc G, Fiorello D, Martino A, Scarcella G, Alemanno A, Zubaryeva A, Thiel C (2012) Driving and parking patterns of European car drivers – a mobility survey. Technical report, European Commission, Publications Office of the European Union, Ljubljana. <https://doi.org/10.2790/7028>
38. Wolday F, Næss P, Tønnesen A (2019) Workplace location, polycentricism, and car commuting. *J Transp Land Use* 12(1):785–810
39. Greene JC, Hall JN (2010) Dialectics and pragmatism: being of consequence. In: *Handbook of mixed methods in social and behavioral research*, pp 119–144

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
