Visually Analysing Urban Mobility: Results and Insights from Three Student Research Projects

Till Nagel[1]

Abstract

Since the digitalization of urban mobility, such systems generate large and heterogenous data sets. Handling these, however, is still a major challenge. To help people making sense of urban data, skills in data literacy are needed. In this paper, we present results and insights from a set of mobility visualization projects, students from computer science and design developed over one semester. The resulting prototypes visualize (a) user-generated bike trajectories, (b) population movement, and (c) public transit passengers. After a brief justification on why we find urban mobility a practical domain for teaching data literacy, we introduce the course and describe its set and setting. For each project, we report on the challenge, the data, the visualization prototype, and the impact these projects had. Our contribution is twofold: in the description of the problem-oriented course and its results, and in the reflection of our approach and our lessons learnt.

Keywords Urban data · Education · Research-based learning · Functional prototypes · Visual analysis

Visuelles Analysieren urbaner Mobilität – Ergebnisse und Erkenntnisse aus drei studentischen Forschungsprojekten

Zusammenfassung

Seit der Digitalisierung urbaner Mobilität erzeugen solche Systeme große und heterogene Datensätze. Deren Handhabung stellt jedoch noch immer eine große Herausforderung dar. Um Menschen dabei zu helfen, mit urbanen Daten sinnvoll umzugehen, sind Fähigkeiten im Bereich der Data Literacy erforderlich. In diesem Artikel stellen wir Ergebnisse und Erkenntnisse aus einer Reihe von Mobilitätsvisualisierungsprojekten vor, die von Informatik- und Designstudierenden in einem Semester entwickelt wurden. Die daraus resultierenden Prototypen visualisieren (a) nutzergenerierte Fahrradtrajektorien, (b) Bevölkerungsbewegungen und (c) Fahrgäste des öffentlichen Nahverkehrs. Nach einer kurzen Erläuterung, warum wir urbane Mobilität als sinnvolles Feld für die Vermittlung von Datenkompetenzen betrachten, stellen wir den Kurs vor und beschreiben dessen Aufbau und Umgebung. Für jedes Projekt berichten wir über die Aufgabenstellung, die Daten, den Visualisierungsprototypen und die Auswirkung dieser Projekte. Mit diesem Artikel leisten wir zwei Beiträge: zum einen die Beschreibung des problemorientierten Kurses und seiner Ergebnisse, zum anderen die Reflexion unseres Ansatzes und der daraus gezogenen Lehren.

1 Introduction

Data literacy been called the next frontier in the open data movement (Veeckmann et al. 2017). For the "benefit of students, employers, and society, data literacy must be recognized as a necessary civic skill" (Risdale et al. 2015). Learning data visualization can help to increase such skills. One of the main goals is supporting learners of different backgrounds to acquire knowledge, gain skills, and practice

✉ Till Nagel
t.nagel@hs-mannheim.de

1 Mannheim University of Applied Sciences, Mannheim, Germany

competencies. As active learning supports achieving deeper understanding (Kerren et al. 2008), problem-based data-driven courses are useful.

Good data sets for teaching should have appropriate complexity and size, be from the real world, and be relevant to the student's interest (Lo et al. 2019). A fitting data set stimulates exploring a wider range of visualization techniques, without hindering rapid progress. Using real-world data enables working on actual problems and heightens the meaningfulness of the results. Furthermore, it elicits working on typical challenges such as missing metadata, or merging different sources. Lastly, students can better connect to the problem if the data is from their environment. Data related to the student's living situation stimulate curiosity and interest.

Many urban mobility data sets fit these requirements. In recent years, large amounts of movement data are digitally collected from various sources (Zhao et al. 2016). In the projects we are presenting in this paper, these range from volunteered geographic information such as GPS-tracked cycling trips, to governmental databases such as population register, to sensor data such as automatic passenger counts. With their combination of spatial and temporal attributes, they are a specifically good choice. The familiarity of these space and time dimensions facilitates engagement of non-experts (Fechner and Kray 2014). Geospatial information is often presented in map-based visualizations which help to contextualize the data in an understandable manner (Degbelo et al. 2016).

Thus, the goal of using real-world urban mobility data in a project-based visualization course is twofold: firstly, it is a fitting topic to teach visualization concepts and second, it allows students to work on actual problems and communicate their findings to different stakeholders.

2 Visualization in a Research-Based Course

In a course at the Mannheim University of Applied Sciences, students investigate how interactive geovisualizations can support users to explore and understand urban mobility patterns. The course is part of the Mannheim Model of Data Literacy Education (MODAL). In this programme, students of all fields develop and deepen their knowledge and skills in collecting, managing, analysing, visualising, interpreting, evaluating, and applying data (Balz et al. 2018). MODAL provides data literacy skills on three levels. The course presented in this paper is at the third level, where students carry out data-oriented research projects to deepen computational thinking competencies. Learning objectives are gaining foundational knowledge of data analysis methods, extended understanding of visualization techniques, and applied skills in user-centred design.

Each semester, the course runs in collaboration with external partners. These do not act as clients, but as research partners providing questions, domain expertise, and most importantly a data set. Partners can come from industry, research, administration, or civic society.[1] Outcome for the partners lie in the ideas and the prototype created by the students.

One of the core aspects of teaching data visualization is the interdisciplinary aspect (Dykes et al. 2010). The course was developed as part of the computer science elective module system, but has been open to students from the design department from the very beginning. Two semesters after inauguration, the course became part of the design department's module catalogue. The trans-faculty course is designed for small groups with shifting compositions of the students' backgrounds. Starting with six students from the computer science department and two students from the design department in summer 2017, the group consisted of varying fractions, mostly with a majority in computer science.

This resulted in groups capable of handling the various tasks in creating visualization systems, from exploring related work, to participating in co-creation workshops, to collecting requirements, to creating paper prototypes, to preparing and analysing data, to designing mock-ups, to implementing the functional prototype.

The course starts with lectures and guest talks by partners to lay the foundations in both data visualization and the specific domain. Over the semester, students employ an iterative approach ranging from exploratory data analysis (EDA), to collaboratively collecting visualization ideas. Using design thinking in data visualization, education has been shown as beneficial (Roberts et al. 2016). Students design and develop a functional visualization prototype addressing a real-world issue.

The course aims to stimulate enquiry-based learning and encourages students to explore the data in an open-ended manner. Besides the functional prototype, course outcomes are the documentation of the design process, descriptions of the system's features, and of the project-specific insights gained through the visualization.

3 Three Cases of Student Works in the Field of Urban Mobility Visualization

In the following, we present three selected projects from the years 2016 to 2018. All focus on data-driven visual analysis of mobility and movement behavior in an urban setting.

---

[1] See https://infovis-mannheim.de for partners and projects.

Fig. 1 Major outcome of each student research project was a functional prototype. These systems visualize **a** user-generated bike trajectories, **b** population movement, and **c** public transit passengers on a large interactive surface

3.1 Via Velox

Via Velox is an interactive system visualizing rides tracked and shared by cyclists through a mobile app (Kaya et al. 2018). It consists of a map and a calendar view side by side, and enables to interactively filter spatial and temporal properties of rides to allow analysing bicycle traffic.

3.1.1 Data

The volunteered geographic information come from cyclists using a custom mobile app, developed for iOS by students of a partner university. The GPS tracks consist of a series of geo-locations and corresponding timestamps representing ride trajectories. For privacy reasons, the first and last meters of each track are being truncated. We simplified the trajectories in such a fashion to reduce the size of the data without
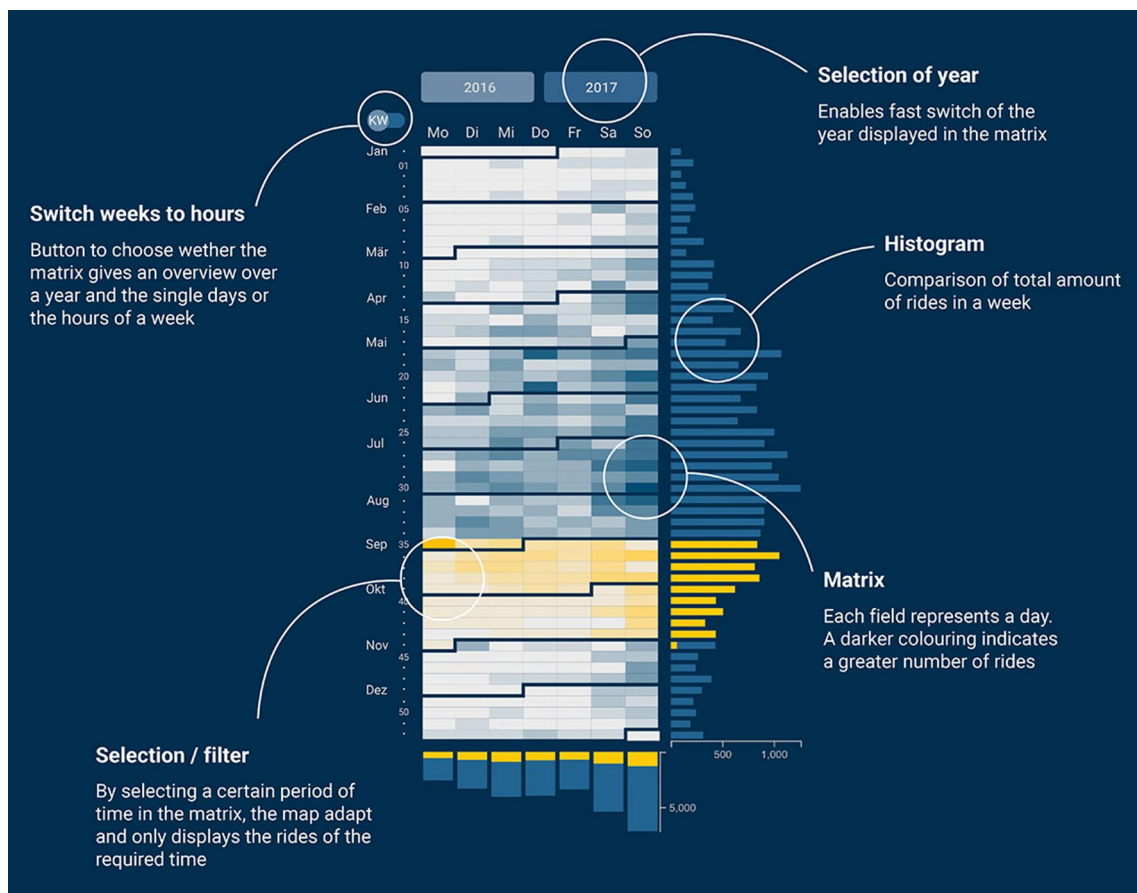

Fig. 2 Interactive calendar heatmap with two histograms on each axis visualizing the density of bike trips over different temporal granularities
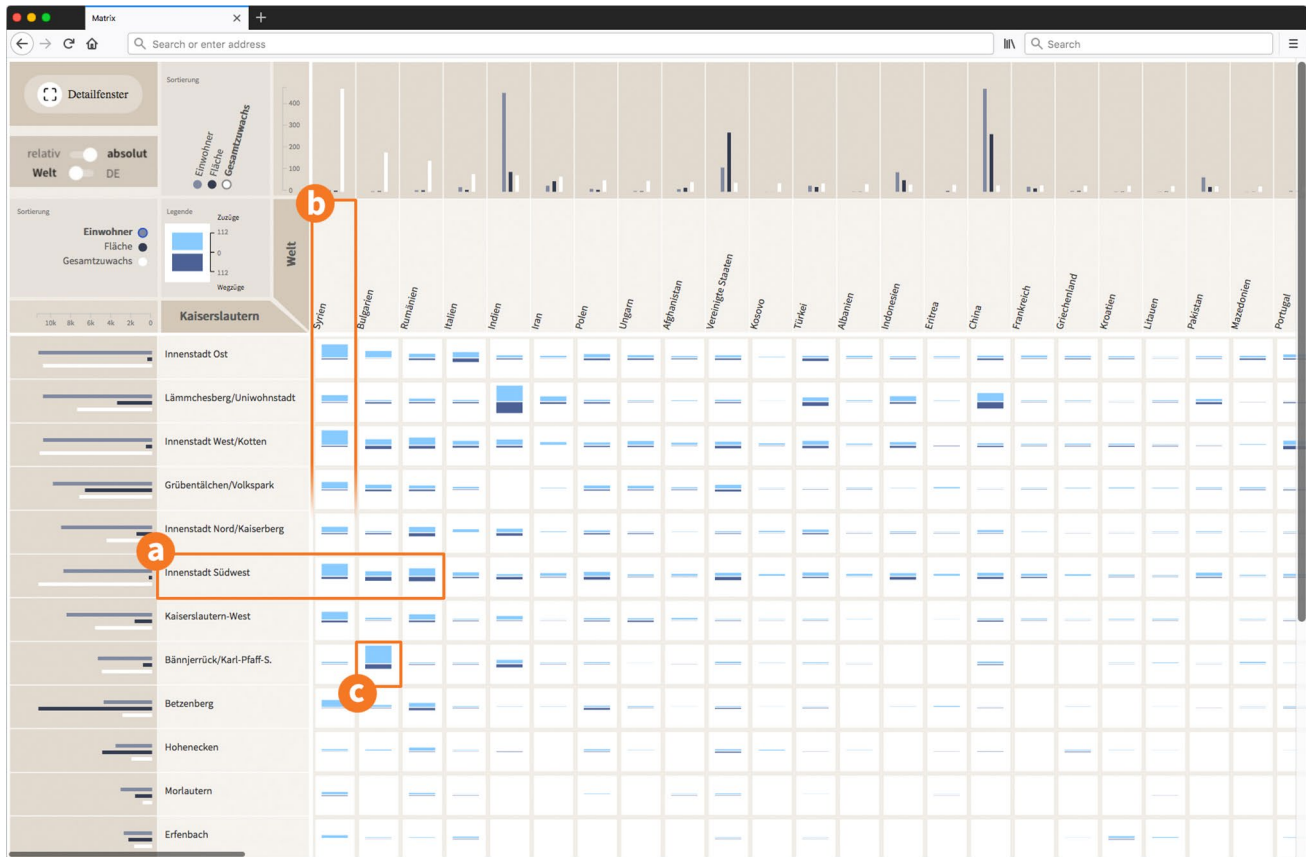
Fig. 3 Interactive matrix showing population movement

limiting its expressiveness (Douglas and Pecker 1973). We did not road-match the tracks to incorporate desire paths such as cycling through parks, crossing squares, or other offbeat routes.

3.1.2 Prototype

The map view shows the paths of all bike trips (in blue), or all currently selected ones in yellow (Fig. 1a, left). Users can zoom and pan the map to view areas of interest, or enter place names for the map to dynamically show the matching place, with a zoom level fitting the place type (e.g. POI or neighbourhood). Users can select locations to filter bike trips which started or ended in their vicinity by simply clicking on the map. A translucent yellow circle is displayed whose size encompasses the geographic area to filter. Through direct manipulation, users can adapt these filters. By dragging the border of the circle, people can change the filter radius to highlight more or fewer cycling paths.

On the right side, the temporal distribution of bike ride frequency is visualized as a calendar heatmap (Fig. 2), with darker cells indicating higher amounts of rides. Two histograms on both axes show aggregated frequency of trips for each weekday, as well as for each week, respectively, each hour. Through summarization, these histograms allow a more direct comparison of trends over different temporal granularities.

Users can select time spans via the calendar by clicking on a single cell, or by dragging the mouse over the matrix. After selecting specific time ranges, the histograms in the margins get updated to highlight the sub-selection, as well. In this way, the highlighted parts of the histogram bars enable comparing the frequency of bike trips in a selected time span with the overall temporal distribution.

The two views are coordinated, i.e. every spatial filtering in the map is mirrored in the calendar view, with all matching trips being displayed over time. The range of the two axes as well as the legend update dynamically.

3.1.3 Impact

One of the goals of the larger project this visualization was part of, was to understand the rides cyclists take in the Mannheim region to plan cycling highways based on

Fig. 4 Visualizes passenger flows as well as time plan and delay data in an extended Marey chart (left). On the lower right, boarding and deboarding passenger counts are displayed as two line charts for both directions

actual demands. To get more people to track and share their rides, we utilized the visualization also as a communication tool to invite citizens to participate. The interactive system accompanied by explanatory posters and videos was on display in our booth at a large cycling festival in Mannheim. Student researchers gave talks at conferences and meetups, and published a poster abstract at the IEEE VIS conference. Based on this, others from the visualization community have re-implemented the calendar heatmap with marginal histograms (e.g. Schwabish 2017; Kruse 2017). Furthermore, it was also one of the exemplar demonstrators for a research proposal.

3.2 WanderVis

WanderVis is a visualization tool for urban administration and city planners to explore and better understand population movement to, from, and within a city (Andres et al. 2018). The main view contains a sortable matrix showing population change in a given time range. Each cell includes a diverging bar chart visualizing people moving to and from urban districts, separated by their citizenship.

3.2.1 Data

The project uses an anonymized extract from the register of residents, spatially aggregated to district level. The original data are in the KOSIS data format.[2] Selected additional data for each district and country such as population and area size come from further official sources.

3.2.2 Prototype

The prototype consists of an interactive matrix with bar glyphs in each cell showing movement to and from districts (rows), subdivided by citizenship (columns). Marginal bar graphs show population, area, and migration delta for each district and country.

The visualization technique for the prototype was chosen in such a way that patterns and outliers stand out visually and are quickly recognizable by the analyst. The interactions make it possible to look at such finds in greater depth. Thus, the prototype supports on different levels: from basal to more complex insights.

---

2 Statistikdatensatz Bevölkerungsbewegungen, 2007.

3.2.3 Insights

In the matrix, three basic shapes can emerge visually: horizontal patterns, vertical patterns, and individual cells. Horizontal patterns highlight districts with special movement characteristics (such as *Innenstadt Südwest*, where people from most of the top three nationalities move in and out, see Fig. 3a). Vertical patterns refer to nationalities with special movement characteristics (e.g. many Syrians move to Kaiserslautern, see Fig. 3b). The glyphs of individual cells are particularly noticeable when the movements in the immediate vicinity differ from those of the cell. Thus, it becomes clear that in contrast to persons of other nationalities, many Bulgarians move to *Bännjerrück* (Fig. 3c). In addition to these basic visual analyses, users can also compare distant glyphs. In Fig. 3, for example, two glyphs are particularly noticeable, indicating that people from India and China move to and from *Lämmchesberg* in higher amounts. Since this is the district where the Technical University is located, it can be assumed that these are students.

3.2.4 Impact

We presented our results at the Lord Mayor's office of Kaiserslautern. The prototype was well received by the experts from different departments, who found the tool insightful and expressed a clear interest in further development. The Lord Mayor praised the prototype as an "easy-to-use and visually pleasing tool, which can bring an enormous relief to city council and administration" (Stadt Kaiserslautern 2018). Local newspapers have reported on the project. It has been referenced as a pilot project in a grant proposal by the partners.

3.3 PaxMotion

This visualization system is designed to support urban transit planners and policy makers to better understand and explore passenger demands. It visualizes boarding and deboarding passengers as well as deviations from the time plan to allow analyzing a possible interplay between these. Thus, temporal characteristics of single stations as well as of clusters of neighbouring stations can be easily seen such as emerging commuting patterns.

3.3.1 Data

The timetable data provided in General Transit Feed Specification (GTFS) format consists of schedules, lines, and geographic information of the tracks. The schedule data have been blended with passenger flow data. The passenger flows are recorded by door sensors with precise entry and exit counts. Our partner provided us with boarding and deboarding numbers for their demonstrator tram line 1 (VRN 2018). Finally, we also incorporated delay data which our partner offers in real time on their website.

3.3.2 Prototype

The interactive dashboard consists of multiple coordinated views. On the left, an extended Marey chart (also known as Ibry chart) displays the stations along the *x*-axis and the time on the *y*-axis, thus showing velocity for each train. The stations ticks are placed according to the geospatial distances between them. In this way, areas with different urban characteristics become visible (e.g. denser distribution of stops in the city center).

We extended the traditional Marey chart in two ways: with station glyphs encoding passenger flow, and optional secondary graphical lines showing delays. Marey charts have been extended before, e.g. to show passenger load along the lines. As our aim was to allow investigating passenger exchange, the glyph encodes passengers entering (turquoise) and exiting (orange) the vehicle. The color coding is consistent in all views. Besides the scheduled time, this view also shows the actual times. In the case of delays, two lines are being shown. The distance between these lines encode the delay in the same scale of the overall time axis. To enable analysts to investigate both global times such as changing headways (visible at a day scale), as well as local delay variations (visible at an hour scale), the whole chart is vertically zoomable through pinching to view different temporal scales.

On the right side, a set of additional views are displayed. A calendar heatmap (Fig. 4, upper right) shows total passenger distribution for a month. Users can browse through different months by swiping, and can select a specific date to investigate. After tapping on a day, the data in all other views are being updated.

At an area shows some summary statistics. It shows mean boarding and deboarding passenger numbers, as well as average delays. If no station is selected in the Marey chart, the average is shown for all stations over the current day. After selecting a station, the mean over all day for this station is shown. When the user taps a single glyph (i.e. a specific vehicle at a specific station), the precise passenger numbers as well as the actual delay are shown.

3.3.3 Impact

The project results have been presented to the CEO and traffic planners of our partner public transit company. Besides commending the visualization system as a helpful and interesting prototype (and publishing multiple

press releases), our partner especially lauded the design process with its co-creation workshops, and the richness of the ideation results. Furthermore, this project won the international CityVis competition 2018 in the "research" category.

4 Discussion

Running this course over the years with different students and different partners helped us refine the curriculum and project setup. In the following, we reflect on some of our lessons learnt.

Problem-oriented, yet open-ended In our case studies, students identified, prioritized, and selected problems together with the experts. The partners were involved by participating in design thinking workshops, and giving feedback throughout. However, they did not predefine possible solutions. These were driven by the data, and by the aim to design visualizations, new and interesting, to the students. While this might result in prototypes not directly practical for day-to-day problems, all partners were highly satisfied and found outcomes from the ideation phase to be valuable for future endeavours.

Data are not perfect In all projects, the data given to us by our partners needed further preparation. While the provided data were mainly according to standards (GPX in Via Velox, KOSIS in WanderVis, GTFS in PaxMotion), issues emerged. The data had not been fully to the specification, parts of the data had been missing, etc. In ViaVelox, the tracking app was developed in parallel and could be deployed only to few volunteers in a beta. We had to get further cycling data from third parties, to test the scalability of our system and visualizations. In PaxMotion, our partner the public transport company did not record the real-time data, themselves. Thus, we had to scrape and store them within the runtime of the project to be able to show delays within our extended Marey chart. We see this as useful practice in finding solutions to typical real-world data preparation challenges.

Trust the students' capabilities Over the semesters, we learned it to be beneficial to trust the student's eagerness to learn, their passion to explore, and their capabilities to create. In an open-ended process with plenty of uncertainties and challenges, students might feel lost from time to time. At these stages, we found it important to reiterate that learning is the major outcome. Interestingly, we experienced this to result in better prototypes. Reinforcing the faith you have in the students fosters ownership and promotes a self-driven approach. You need an environment

which encourages them: a space to work independently, interdisciplinary, and in shifting compositions; a setting to experiment, where open discussions and trial and error are encouraged, and to not fear failure but see it as an opportunity to improve. Lastly, this needs teachers who act not as gatekeepers of knowledge, but as facilitators of learning. Combined with the need to present their final prototypes to the partners, this resulted in mature, well-designed data visualizations.

5 Conclusion

We have demonstrated with three student research projects that undergraduates are capable of designing and developing fully functional visual analytics systems to explore urban mobility. Through brief descriptions of selected outcomes, we have shown that these projects were exceptionally successful. While none of the prototypes are in production, all were well received by the partners as well as by the design and research communities.

We briefly discussed some of our lessons learnt, including that urban mobility is a specifically fitting domain for learners to work on data-driven problems from the real world. Through formal teaching evaluation, we understand that we have met the targeted learning goals. Through informal feedback, we learned that students found this course to be especially fun, exciting, and beneficial to deepening their skills and competencies.

We documented the course setup and the project results to act as basis for further discussion. As a next step, we hope to connect to other educators to collect and exchange best practices, and to create a set of recommendations on how to create a learning environment for successful inquiry-based student projects in the urban data domain.

References

Andres D, Memmel M, Nagel T (2018) WanderVis - Interaktive Visualisierung von Bevölkerungs-bewegungen. Planerin - Fachzeitschrift für Stadt-, Regional- und Landesplanung 4(2018):27–30

Balz D, Bandtel M, Kreim S, Nagel T, Noyon A, Trägner U, Trefs L, Werft W (2018) Mannheimer modell data literacy education (modal), future-skills-tag. Hochschulforum Digitalisierung, Berlin

Degbelo A, Granell C, Trilles S, Bhattacharya D, Casteleyn S, Kray C (2016) Opening up smart cities: citizen-centric challenges and opportunities from GIScience. ISPRS Int J Geo-Inf 5:16

Douglas D, Peucker T (1973) Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. Can Cartogr 10(2):112–122

Dykes J, Keefe D, Kindlmann G, Munzner T, Joshi A (2010) Perspectives on teaching data visualization, IEEE VisWeek 2010 Panel

Fechner T, Kray C (2014) Geo-referenced open data and augmented interactive geo-visualizations as catalysts for citizen engagement. JeDEM 6(1):14–35

Kaya D, Keles B, Nagorny D, Perle P, Pregler P, Rudolf L, Schröder M, Tunali U, Nagel T (2018) ViaVelox—a system to visually analyze GPS-tracked bike rides. In: Poster Abstracts of IEEE VIS 2018

Kerren A, Stasko J, Dykes J (2008) Teaching information visualization. Information visualization: human-centered issues and perspectives. Springer, Berlin, pp 65–91

Kruse A (2017) Bike Sharing in Hamburg, https://www.alexknowsdata.com/post/stadtrad. Accessed 21 Jan 2020

Lo L, Ming Y, Qu H (2019) Learning vis tools: teaching data visualization tutorials. IEEE VIS 2019

Ridsdale C et al (2015) Strategies and best practices for data literacy education. Dalhousie University, Halifax

Roberts JC, Headleand C, Ritsos PD (2016) Sketching designs using the five design-sheet methodology. IEEE Trans Visual Comput Graphics 22(1):419–428

Stadt Kaiserslautern (2018) Wer wohnt wo? Studierende stellen Prototyp vor. Press release, 16 Jan 2018

Schwabish J (2017) Calendar Heatmap in Excel, https://policyviz.com/2017/08/29/calendar-heatmap-excel/. Accessed 21 Jan 2020

Veeckman C, McCrory G, Walravens N (2017) Data literacy for greater civic participation in smart cities—visualising open mobility data. Proceedings of the 21th EMAN Conference, Liège, 2017

VRN Verkehrsverbund Rhein-Neckar (2018) Press release, PR/20–18, 19 June 2018

Zhao K, Tarkoma S, Liu S, Vo H (2016) Urban human mobility data mining: an overview, 2016 IEEE International Conference on Big Data (Big Data), pp 1911–1920