VeloViz: Representing the Traffic Load Between Vélo’v Stations

Amanda Belande[1]
Master Architecture de l’information
École normale supérieure de Lyon

Ka Hei Cheung[2]
Master Architecture de l’information
École normale supérieure de Lyon

Henri-Louis Sergent[3]
Master Architecture de l’information
École normale supérieure de Lyon

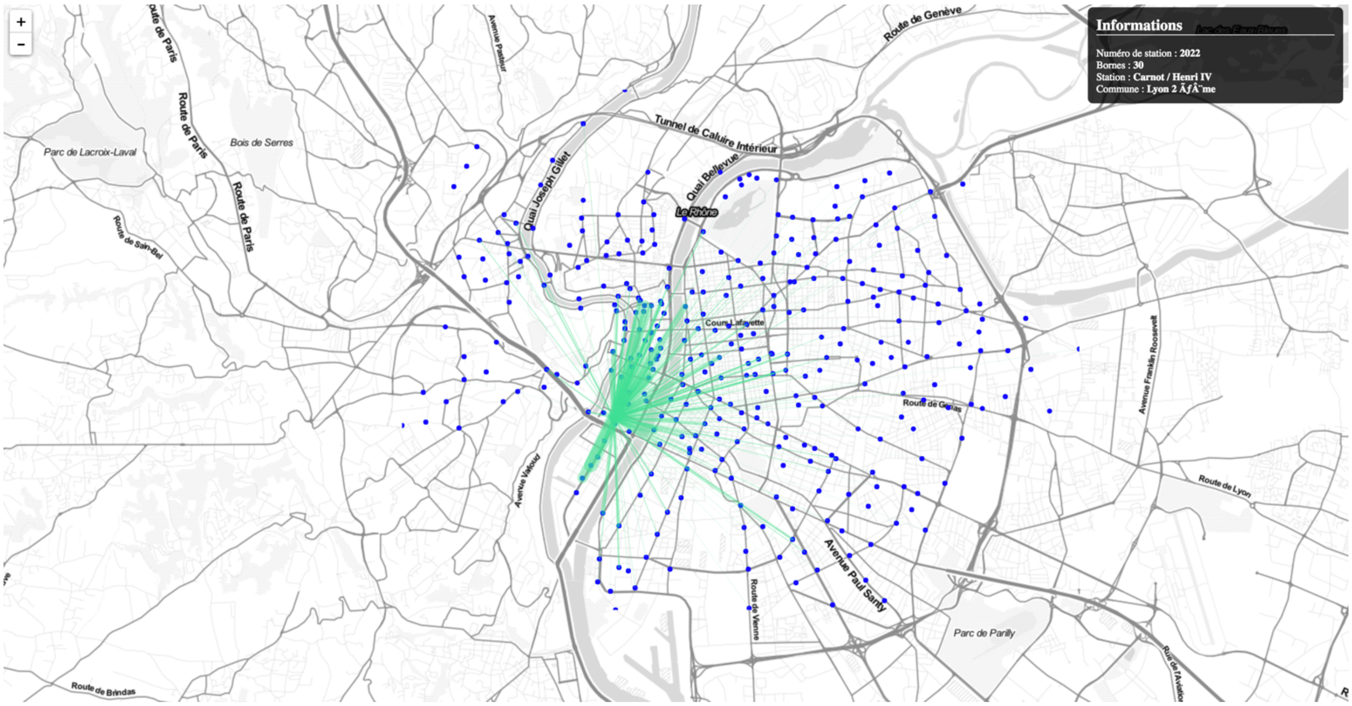

Figure 1: Screenshot of VeloViz

ABSTRACT

In this project, we aim to study bike hire traffic patterns of Lyon’s bike hire service—Vélo’v—through its subscribers’ socio-demographic data and the recorded bike trip information. The visualisation is based on the data set from 2011 and shows bike station locations and trips on a geographical map of Lyon. This visualisation uses D3.js, Leaflet map library and the data provided by JCDecaux, the company operating the service. Whilst the majority of other visualisations mainly rely on open data (stations and bike availability) provided by the service operator, this visualisation offers new insights on the study of bike hire traffic patterns with an exclusive set of anonymised user data.

1. INTRODUCTION

Vélo’v is a bike-sharing system in the city of Lyon. Launched in 2005, it was one of the first bike sharing systems implemented in France [1]. It started with 2,000 bikes and 200 stations and has increased to 4,000 bikes spread over 348 stations located in Lyon and Villeurbanne [2]. Since 2013, JCDecaux, which operates Vélo’v in partnership with the Greater Lyon administration, has released static data on the locations of Vélo’v stations, the number of bike stands at each station and information on the usage of the service from the seven previous days. It also provides access to real-time data on bike and bike stand availability through an API [3].

For this project, we will use a data set on cycling trips in 2011 made available by JCDecaux and Lyon authorities as part of the research project ANR ‘VEL’INNOV’ [4]. The project was carried out by an interdisciplinary research team in Lyon studying the use of bike-sharing services. Following discussions with a professor who participated in the project, we were able to gain access to the data set for learning purposes on the condition that the visualisation is not publicly accessible. The data set includes service subscribers’ information, such as gender, age, place of residence and type of membership (yearly subscription), and cycling trip information between Vélo’v stations. Crossing information of demographics with bike hire service use, our visualisation will specifically focus on trips by yearly subscribers, illustrating over 4 out of 7 million recorded trips in 2011.

The main advantage of this project is the ability to show bike use in relation to a socio-demographic perspective, which would not have been possible with the general open data set. A second advantage comes from the availability of a ready-to-use data set which will allow us to focus on the analysis and the visualisation. Nevertheless, drawbacks still exist in regard to the use of this data

[1] amanda.belande@ens-lyon.fr [2] ka-hei.cheung@ens-lyon.fr [3] henri-louis.sergent@ens-lyon.fr

set. The first one concerns the freshness of the data. Our data set is over five years old which may limit the insights we could contribute to discussions on smart cities and bike-sharing systems. Secondly, the data set only covers the year of 2011. Comparison with another period of time for studying any pattern evolution is hence not possible. Yet, there is still sufficient data to create an original visualisation of users' habits over the year, on hourly, daily, weekly, monthly, and seasonal basis. In addition, previous visualisations were static—We hope that, through an interactive visualisation, new insight could be explored from the same data set.

We expect our visualisation could interest various public groups. It could help urban planning services (city council and Grand Lyon) better grasp the movement of Lyon inhabitants and the effect of bike hire service on their commutes and way of living. It could also improve city development and inhabitants' quality of life by supporting the expansion of eco-friendly modes of transportation. Second, it can provide a different external perspective to transport service providers (JCDecaux and/or TCL) even if they have probably studied thoroughly their own service data. For example, they could find out whether adding a new station to the network around an existing one is a good idea, or whether one station is worth maintaining with regard to its traffic or location. Our visualisation could also show some outliers (unusually short or long trips between stations). Finally, it can raise awareness amongst Vélo'v users and the general public on some interesting facts about the way this service is used across the city. When using currently available applications, which mainly only show bike station locations and availability of hireable bikes and docks, their knowledge of Vélo'v traffic may be restricted to personal use. Through our visualisation, they could perceive how the traffic is interconnected with stations and obtain a general overview.

In order to visualise traffic patterns, we first explored different design hypotheses using the Five Design Sketches method [4]. Then, we developed the chosen design using the aforementioned data set, containing station information and the Vélo'v subscribers' movement using D3.js. Finally, we explore some bike hire traffic patterns in the resulting visualisation which are discussed later in the article.

2. Related Work

While searching for data sets on bike sharing, we came across various studies and articles that inspired us. Team datahub's GitHub website recommended by the professors, demonstrated that an in-depth visualisation including socio-demographic data could be an interesting idea if we could find a detailed data set with data on users [5]. Our search for this information led us to an online article by Professor Luc Merchez on a static visualisation of usage statistics with whom we met [6]. He introduced us to the work carried out by former intern Dominique Pitt who had used the same data set to develop a geographical information system to analyse the original data set through queries [7].

In addition, we looked at visualisations of bike-sharing systems in other cities such as the following ones to find out what visualisations would be the most relevant and retained the ones from of Boston [8] and Moscow [9] bike-sharing services.

Andy Woodruff's interactive map [10] of trips in Boston was particularly interesting. It plays with the width of lines to represent trip volumes between two stations. He also uses two colours to differentiate trips to and from a selected station. What was interesting in Urbica's visualisation of traffic in Moscow was the use of a dark background map on which they superimposed colours to represent stations and traffic load. Design being an important part of visualisation, we needed to start thinking of how best to represent data about bike trips and these examples allowed us to test different design options and improve the final visualisation.

3. Project Description

3.1. Design

At the beginning of the project we explored some early leads for our interface. We made some sketches (hand-drawn & Balsamiq Mockups™) to help us shape our work.

We initially imagined two interfaces. The first interface illustrated the relation between Vélo'v traffic and time (Figure 2 below). We would adopt a geographical map as the base layer, on which we mark the location of Vélo'v stations. Lines would be added over the map and the station locations to show the displacements effectuated by travellers. Hence, the denser the lines between two stations, the more frequent the traffic between them. A tab bar could be added to offer the possibility to switch between different time periods, namely, day, week, month, and season. A timeline slider could also be available for viewing the variation between corresponding time intervals. On the right-hand side, filters could enable users to further explore traffic patterns by varying available data on Vélo'v subscribers' profiles, for instance, gender and age, and journey duration.

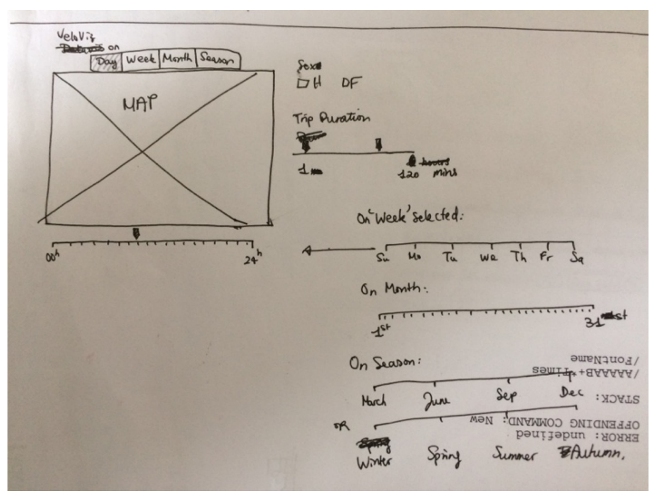

Figure 2: Sketch of the map interface

We had imagined a second interface using a chord diagram showing Vélo'v traffic between areas in Lyon and Villeurbanne. The stations would be grouped into zones (Lyon arrondissements and Villeurbanne) and laid along the perimeter to form a circular space, in which lines would show the traffic flux between stations across two zones or two stations.

We finally chose the map design because it seemed more relevant to use spatial coordinates as they were present in the station data set.

3.2. Realisation

Whilst we try to realise our objective, we encountered, mostly, technical difficulty in regard to manipulating the data set and visualising its contents. With the available resources in the given time frame, we gradually readjusted our project scope whilst maintaining our principal intention—visualising bike hire service traffic of Lyon.

We familiarised ourselves with the original data set and did some preliminary work in order to prepare it for the visualisation process. This process included dropping columns (verification keys, duplicates, non-relevant data, bike ID numbers) to make it more easily readable. We kept the data on points of origin and destination (stations out/in), the time values and the user trips we needed. The realisation process then followed following steps:

3.2.1. Geographical Map Implementation

We used a Leaflet and OpenStreetMap tile (zoomable) base as the first layout of information. We explored amongst different themes, from multicolour to black and white, before coming to an informatively clear yet unobtrusive map, so as to underline the Vélo'v station and traffic information which we wanted to provide.

3.2.2. Bike Station Location Visualisation

The second step consists of extracting geographical coordinates of Vélo'v bike stations and adding them on the map. To achieve this, we created a layout on top of the map and used a loop to run through the station data file to extract all the station coordinates. We then plotted these coordinates with the use of D3.js (Figure 3 below). Each tiny blue dot represents a bike station in service in 2011.

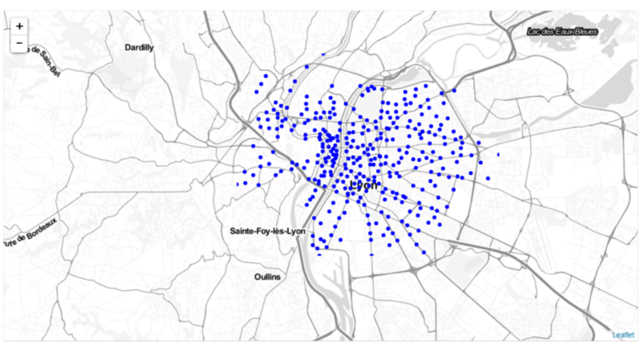

Figure 3: Locations of Vélo'v stations (blue dots) superimposed on the background map

3.2.3. Bike Station Information Visualisation

Whilst the blue dots precisely indicate bike stations' location, the display of further information about each station could be useful for locating a desired station and understanding. This drove us to the third step of adding an information box providing details of each bike station, namely station ID, number of docks, station name and the zone to which it belongs.

We adopted the mechanism of mouse hover. When the pointer moves above a blue dot, corresponding items of information mentioned above would be sought in the bike station data file and be displayed in a pop-up box next to the blue dot. We later adjusted the position of this box and put it statically in the top corner for making this extra information subtler (Figure 4 opposite).

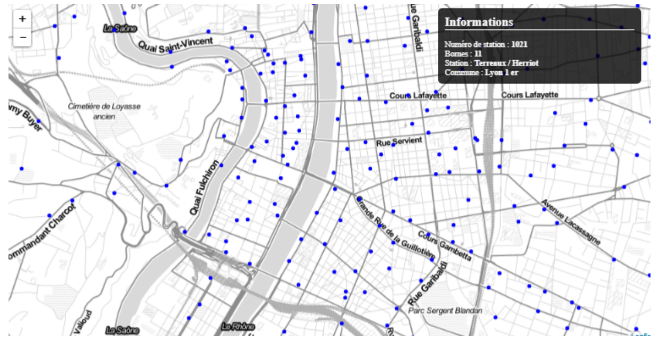

Figure 4: Map with the information box in the upper-right corner

3.2.4. Bike Traffic Visualisation

The last step involves plotting recorded cycling journeys in 2011. Given that the volume of the concerned data file is massive (around 4 million journeys), several actions were taken to lighten the calculation burden when loading the corresponding data file, manipulating the data, and visualising the information.

Although we had already removed unnecessary columns in the given data file prior to our realisation process, the file was still too heavy to load, both locally and on the web. We, unfortunately, conceded to remove the socio-demographic aspect of our data file—and later temporal aspect, too—so as to create an aggregated overview of all cycling journeys in 2011. We first created a co-occurrence matrix to record the sum of cycling journeys between two stations.

For the sake of easier technical manipulations with JavaScript, the co-occurrence matrix table was replaced with a linear table which stated, in each entry row, the departure bike station ID, the arrival bike station ID, and the sum of recorded journeys. This linear table was then converted in a JSON file in which arrival station IDs and corresponding sums of journeys were nested under departure station IDs (Figure 5 on the following page). With this additional treatment, we were able to construct our visualisation on a data file which is drastically reduced from around 1 GB to 5 MB, facilitating both our realisation and prospective users' browsing.

We added another layer on top of our geographical map and station locations. On this layer, lines are drawn from a selected (by hovering) departure bike station to all bike stations to which cycling journeys are recorded in 2011. The thickness of lines correspond directly and proportionally to the traffic—the thicker the lines are, the higher number of journeys existed from the selected station to the relative destination.

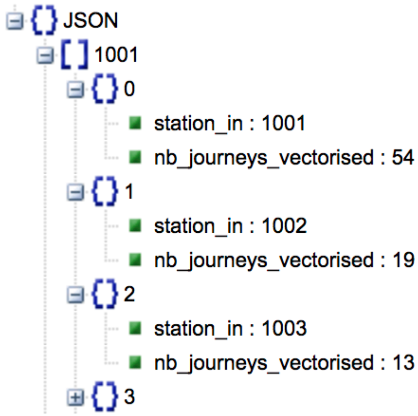

Figure 5: Aggregated, vectorised journey counts in JSON

In conclusion, the final result of our visualisation is a zoomable OpenStreetMap of Lyon with dots representing Vélo'v stations in 2011. This map offers bike station details in an information box located in the upper right corner of the interface when users hover over a selected station. Simultaneously, lines from this station are drawn towards all stations to which cycling journeys existed in 2011. These lines visually represent the variance of aggregated bike traffic in this year between stations.

4. Discussion

Compared to the original expectations, the resulting visualisation is not using temporal nor socio-demographic data. Instead, because of technical and time constraints, we focused our study on the yearly importance of the number of trips from one station to others.

This visualisation allows users to hover over one station and see the aggregated traffic to the other as it rises with the thickness of turquoise lines. Despite its lack of relative accuracy, it allows users to swiftly and visually associate the main related traffic patterns for each station (Figure 6 opposite).

It is also possible to discover some outliers, for example between *Debourg* and *Place de l'École* stations. Despite the fact that there is only a distance of 600 m (7 minutes of walk) between the two stations, there is a great amount of Vélo'v traffic (see Figure 7 opposite). It could be interesting to study this specificity (maybe with hypotheses related to the proximity of the *École Normale Supérieure* campus).

5. Conclusion

This visualisation using dots and lines with variable width is efficient to reveal key patterns on a large data set (4 million trips). These aggregated trips allow an instantaneous view of how the stations are related in terms of traffic during one year and are meaningful even without socio-demographic data. It allows an exploration of the flow for each Vélo'v station and can help identify main hubs, some local patterns and uncanny traffic on some areas.

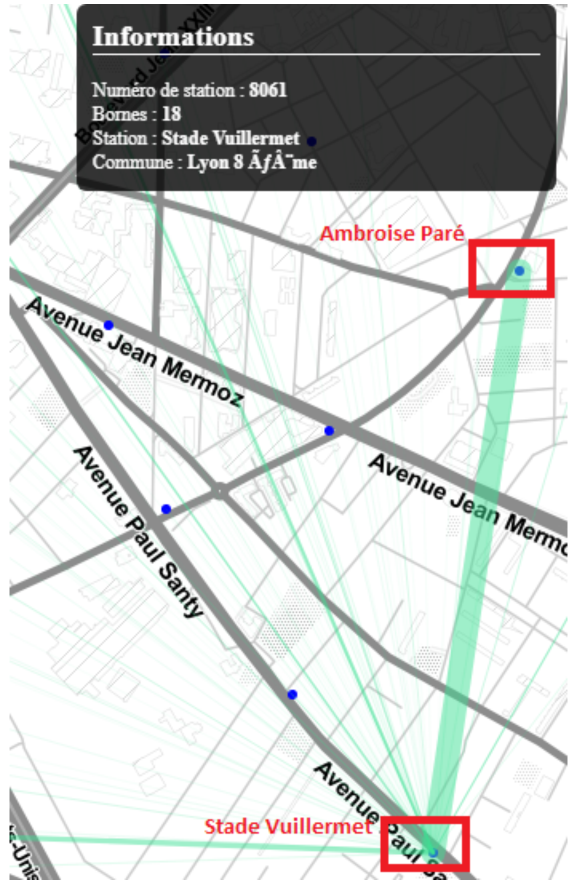

Figure 6: Significant traffic between stations *Stade Vuillermet* and *Ambroise Paré* (marked in red) shown by a thick turquoise line connecting the two stations

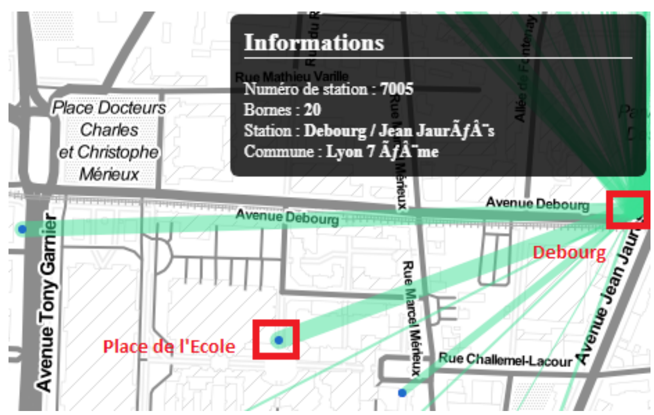

Figure 7: Traffic between Debourg and Place de l'École stations (marked in red)

References

[1] Wikipédia. Liste des systèmes de vélos en libre-service en France. Available online at https://fr.wikipedia.org/wiki/Liste_des_syst%C3%A8mes_de_v%C3%A9los_en_libre-service_en_France, November 2017.

[2] Métropole de Lyon, Vélo'v : Mode d'emploi. Available online at https://velov.grandlyon.com/fr/mode-demploi.html, 2017.

[3] JCDecaux, Open data: the latest on the public distribution of bike-share data, available online at http://www.jcdecaux.com/blog/open-data-latest-public-distribution-bike-share-data, October 2016.

[4] Jonathan C. Roberts, Chris Headleand, and Panagiotis D. Ritsos, "Sketching Designs Using the Five Design-Sheet Methodology", Visualization and Computer Graphics, IEEE Transactions on Visualization and Computer Graphics, vol. 22, 1, IEEE, pp. 419–428, 31 Janurary 2016.

[5] Agence nationale de la recherche, Projet VEL'INNOV, available online at http://www.agence-nationale-recherche.fr/Projet-ANR-12-SOIN-0001, 2012.

[6] Merchez L. Carte à la une : la fréquentation des stations vélo'v à Lyon, partout et à tous âges ?. Available online at http://geoconfluences.ens-lyon.fr/informations-scientifiques/a-la-une/carte-a-la-une/carte-a-la-une-les-stations-velo2019v-a-lyon-partout-a-tous-ages, December 2013.

[7] Pitt P., Merchez L. and Mathian H., "VISUVELOV : visualisation des flux des vélos en libre-service à Lyon", M@ppemonde, N° 115(3-2014) available online at https://mappemonde-archive.mgm.fr/num43/fig14/fig14301.html, 2014.

[8] Metropolitan Area Planning Council, Hubway Data Visualisation Challenge. Available online at http://hubwaydatachallenge.org/, 2012.

[9] Urbica. Bicycles in the city. Available online at https://medium.com/the-data-experience/bicycles-in-the-city-f9529d918388, 2015.

[10] Andy Woodruff, "Hubway Trip Explorer", in Hubway Datavisualisation Challenge 2012, available online at http://bostonography.com/hubway-map/, 2012.