Creation of a visualization interface for bicycle rental data

Diogo Almeida and Edgar Antunes

University of Coimbra

Abstract. The objective of this study is to observe the impact of weather conditions on bicycle rental in the city of Washington DC using tools that reinforce how visual analysis can transform raw data into strategic decisions for more sustainable cities.

Keywords: Bike Rental · Data Visualization · GeoPandas · Pandas · Plotly

1 Introdução

A public bicycle system, also called as a bike-sharing system is a service in which bicycles are made available for shared use. It allows people to borrow a bike from one of the stations belonging to the system and return it at another station of the sharing system. The public bicycle system began in Amsterdam in 1965 [6][8], but the expansion of the system to the world was due to the development of IT technology in the 2000s. The development of technology has made it easier and more convenient for people to rent a bike, and mobile IT devices have allowed people to find a nearby station where they can rent a bike. In 2014, public bike-sharing systems were distributed over 50 countries on five continents and included 712 cities [6][9][1].

Bicycle rental is becoming one of the most sustainable and efficient forms of urban mobility, being an excellent alternative as sustainable, affordable and healthy transportation. Bicycles are an increasingly widespread means of transportation in many cities around the world, facilitating sustainable and environmentally friendly urban mobility.

The concept of bicycle rental is simple: citizens use bicycles according to their needs, paying a fee or making a commitment, without the responsibilities of owning them. Its evolution does not occur uniformly over time, and there are many external factors that affect this service, the main one being weather conditions.

This study aims to observe how changes in the weather, and other possible factors such as the time of year being more or less touristy, affect bicycle rental patterns in the urban environment by creating a visual interactive interface that allows the user to explore different tools. These types of studies are important for better planning and improving the efficiency of these types of services.

2 Related Work

While developing this project, we consulted innumerable sources of inspiration in the field of data visualization applied to urban mobility. One of the main references was the CityVis platform, which aggregates a curated collection of interactive visualizations related to urban data.

This interface allowed us to observe concrete examples of good practice in the visual representation of spatial data, particularly with regard to the use of interactive maps, the temporal representation of mobility patterns and the integration of contextual data, such as points of interest (POIs) while exploring the effects of weather in bike rental, for scientific purposes. Our course lecturer was also a fundamental key, guiding the focus of the work, namely the idea of POIs into the analysis, enriching the interpretation of bike rental patterns. This suggestion proved to be particularly relevant, as it allows demand to be contextualized according to the characteristics of the surrounding territory.

Design decisions were thus informed by these visual and conceptual references, ensuring an approach based on both practical examples and academic guidance. [3][7]

3 Dataset

We used mainly 2 datasets: one for the bike rental information ("capitalbikeshare"), containing information about the duration of the trip, the starting and ending date, the starting and ending station, bike number and member type [2] (https://capitalbikeshare.com/system-data) and one containing the geodata necessary for the usage of for POIs, making use of the Geofabrik project[4] (https://download.geofabrik.de/north-america/us.html).

The "capitalbikeshare" dataset, cited above, aims to answer the following questions: Where do Capital Bikeshare riders go? When do they ride? How far do they go? Which stations are most popular? What days of the week are most rides taken on? (...) They invite developers, engineers, statisticians, artists, academics and other interested members of the public to use the data we provide for analysis, development, visualization or whatever else moves you[2].

Initially, we wanted to present data for all the months of 2025, but due to its weight, our demonstration only showed for the month of January.

The Geofabrik project was created out of the conviction that free geodata created by projects like OpenStreetMap will become increasingly attractive for commercial uses and is a free, community-maintained data like that produced by the OpenStreetMap project which is a real alternative to the offerings of the trade's top players. Such data is not only free of charge, but also comes with fewer license restrictions than other offers[4].

We used this dataset, merged with the "capitalbikeshare", to show the POIs in the city of Washington DC.

4 Exploratory Data Analysis (EDA)

Pandas, GeoPandas and Plotly libraries were used, in particular the plotly.graph module, as well as graphical representations such as scatter plots, among other visual approaches.

The Pandas library was firstly used to load, clean and structure the data from the bike-sharing system in Washington, D.C. It made it possible to perform essential operations such as monthly aggregations, handling missing values, filtering relevant columns and transforming data types. With GeoPandas, we incorporated geospatial data into the analysis process, namely the geographical location of bicycle stations, and cross-reference it with shapefiles of the city to spatially contextualize the observed patterns.

For the graphical representation, we used scatter plots with plotly.graph_objects, which made it possible to create high-quality interactive visualizations. These were crucial for identifying correlations and temporal patterns. For example, when visualizing the number of trips per month, a reduction in use during the colder months became apparent, raising the hypothesis that weather conditions have a significant impact on demand for bicycles.

5 Design

The design was developed based on principles of usability and visual clarity prioritizing an intuitive experience. The choices reflect a balance between functionality and aesthetics, with iterations based on empirical feedback. The inclusion of customizable themes and a modular structure allows adaptation to various usage scenarios.

We used the Global Contentious Politics public dashboard as a model to create our interface. `https://dashboard-glocon.ku.edu.tr/turkey-protest-data`

As *"Less is More"*, our interface shows a clean and efficient look that avoids superfluous elements to reduce the cognitive load, such as the single sidebar containing all controls (filters, themes, layers) and limited color palette, with predefined gradients (blue → red) for representing quantitative data, suggested by the teacher.

We implemented the option to switch between a dark theme, figure 2, and a light theme, figure 1, (for accessibility to reduce eyestrain) and user preference (dark mode for low-light environments, for example). We also choosed the gradient blue - red based on color coding for ordinal data (e.g. temperature or intensity of use), where blue (cold/low) and red (hot/high) are intuitive (semantic color theory), see figure 3.

We also worked on a visual hierarchy highlighting primary data, using bicycle markers in red (high salience) vs. POIs in soft colors, such as museums in dark blue, and making use of *size* and *opacity* for, figure 4, data density creating larger markers in areas of high concentration.

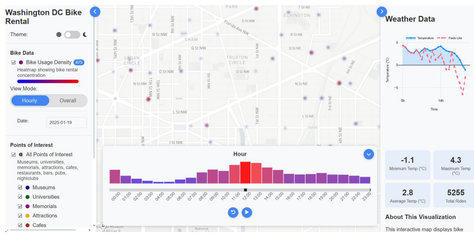

Fig. 1. Overall look to our Interface in Light mode

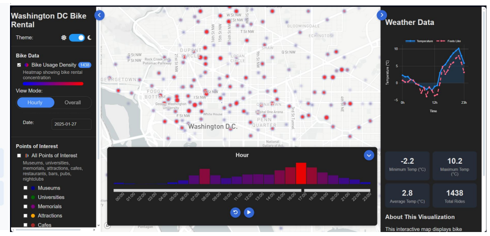

Fig. 2. Overall look to our Interface in Dark mode

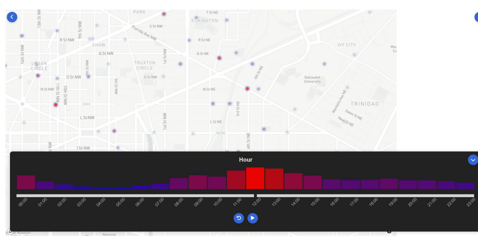

Fig. 3. Data density markers

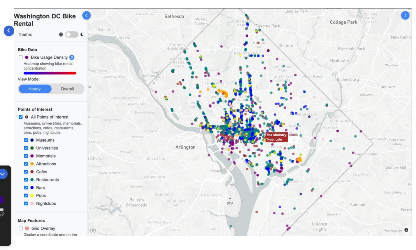

Fig. 4. Poins of Interest (POIs) represented on the map's dashboard

If we take a look to the image below, figure 5, we can see that the sidebar was modeled on analyzed projects (e.g. Plotly dashboards and GIS interfaces such as QGIS), ensuring a recognizable pattern in which reusable components such as sliders and dropdowns that follow Plotly's design system for familiarity.

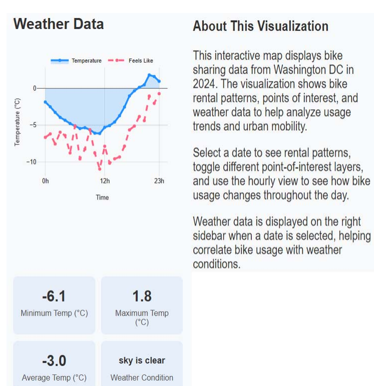

Fig. 5. Side Bar's dashboard in Light mode

6 Implementation

To implement the tools, we created 2 main files: main.py (Python file) and main.html (CSS file).

The implementation first establishes Washington DC's geographic reference point by loading the weather dataset and extracting the latitude (dc_lat) and longitude (dc_lon) from its first entry. These values serve as the spatial anchor for all subsequent map visualizations and calculations:

```
    weather_df = pd.read_csv('datasets/weather.csv')
    dc_lat = weather_df['lat'].iloc[0]
    dc_lon = weather_df['lon'].iloc[0]
```

Listing 1.1. Loading weather data

Interface's code makes use of plotly.graph_objects to create an interactive map of Washington, DC with layers for bicycle data, points of interest (POIs) and a reference grid. For example, check the following code listing that shows spatial density patterns of bike rentals when combined with user interactions, using a layer of red markers for bike data:

```
        bike_trace_index = len(fig.data)
        fig.add_trace(go.Scattermapbox(
            lat=[],
            lon=[],
            mode='markers',
            marker=dict(size=10, color='red'),
            name='Bike Data',
            visible=False
        ))
```

Listing 1.2. example code for Geographic Visualization

We used Plotly's Scattermapbox to overlay multiple data layers on a Mapbox base map, including Bike rental data, shown in listing 1.2 (with red markers showing start locations of bike trips, dinamically filtered by date/hour) Points of Interest (POIs), shown in listing 1.3, that are categorized markers such as museums, universities, etc with custom colors and tooltips and a Reference Grid that consists of Latitude and Longitude lines for spatial orientation (hidden by default).

```

# Layer 1: Bike Data (hidden until user interaction)
fig.add_trace(go.Scattermapbox(
    lat=[], lon=[],
    mode='markers',
    marker=dict(color='red', size=10),
    visible=False  # Initially hidden
))
```

Listing 1.3. example code to create a figure showing Bike Data

```
# Layer 2: POIs (e.g., museums)
fig.add_trace(go.Scattermapbox(
    lat=museum_df['lat'], lon=museum_df['lon'],
    marker=dict(color='darkblue', opacity=0.7),
    text=museum_df['name'],  # Tooltip content
    visible=False
))
```

Listing 1.4. example code to create a figure showing Points of Interest (POIs)

Listing number 6 shows that the bicycle rental data is organized into a hierarchical temporal-spatial structure based on a daily level containing total counts and all geographic coordinates and on a hourly level, providing time-binned distributions of rental locations.

```
bike_data_by_date[date] = {
                'count': len(date_df),
                'lat': date_df['start_lat'].tolist(),
                'lng': date_df['start_lng'].tolist(),
                'hourly': hourly_data  # Add hourly data
            }
```

Listing 1.5. Creating data by hour

This dual-resolution storage enables both macro-level date comparisons and micro-level hourly pattern analysis while maintaining efficient memory usage through list-based coordinate storage.

The next layer loads and processes POIs from a shapefile. It first filters for relevant categories (museums, cafes, etc.) and fills missing names with "Unknown". The script extracts latitude/longitude coordinates from the geometric data, then creates map layers for each category with: category-specific colors (blue for museums, green for universities), fixed marker size (10), and interactive tooltips showing names. All layers start hidden (‘visible=False‘) for user-controlled display, with consistent visual styling (70% opacity) and logical grouping in the legend. The color-coding follows intuitive conventions for easy recognition.

Finally, our code converts all visualization data to JSON format for web deployment.

7 Evaluation

To evaluate the performance of our interface, we inquired a group of people that said to be at least "**a bit familiar**" with bike rental services or "**not at all familiar**", using a Google Forms sheet created by us and the professor.

On figure 6, we can check the answers to the question "On a scale from 1 to 5, how intuitive at first sight was the interface?". Answers spanned from **3** to **5**.

At first glance, these numbers show that our interface is quite intuitive, possibly with a few details to improve. With 47.9% of respondents answering **3** and 57.1% answering **4**, making an average of **3.57** in the classification table of how intuitive it is, we can fairly classify our interface as **moderately intuitive**.

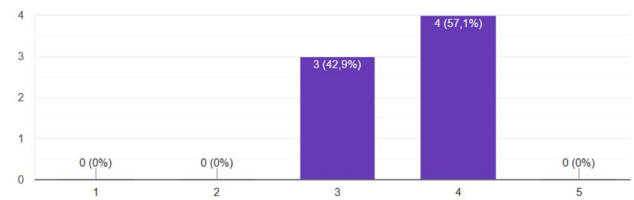

Fig. 6. "On a scale from 1 to 5, how intuitive at first sight was the interface?"

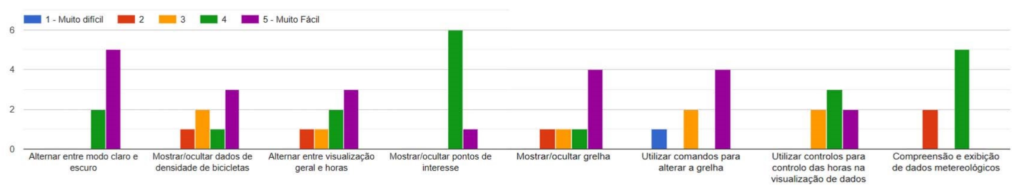

Fig. 7. "Evaluate the completion of these tasks"

Next, we asked them to evaluate the performance of tasks, seen on figure 7:

Let's start by paying attention to the "Switch between light and dark mode" and "Show/hide points of interest" questions. All the surveyed answered as "**easy**" or "**very easy**".

The rest of the answers, there is now some dispersion, specially if we look to the "Show/hide data from bike density", "Switch between overview and hours" and "Show/hide grid" answers.

The last tasks did not performed so well because answers show a not so homogeneous dispersion. This is probably explained by the surveyed group, that is divided by "**a bit familiar**" and "**not at all familiar**" with bike rental services and we could improve the tools to be more intuitive.

Despite all of that, results end up showing a majority of "**easy**" or "**very easy**" answers. When asked a general first opinion about the interface, the group responded in general as being compact and straight to the point or simple and intuitive yet effective. One surveyed stated that design could be improved. No further clarification.

When asked about the clarity of bike density showed, 28.6% of the group rated it with a **5**, and 71.4% with a **4**, averaging a **4.29** (on a scale from 1 to 5). We can conclude that we did well by representing bike rental density. Results showed in figure 8:

Regarding the POIs, 71.4% of the group answered that they were easy to distinguish and the remaining 28.6% said they were not easy to distinguish.

71.4% also rated with **4** in terms of how useful is the timeline for visualizing changes throughout the chosen day. The remaining rated with a **5** (14.3%) and a **3** (14.3%).

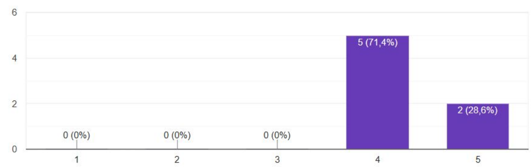

Fig. 8. "How clear is the visualization of bike rental density to show popular areas?"

Finally, we asked users to rate on a scale of 1 to 5, how would they rate the overall usability of the view: 1 user rated it with a **3** and the remaining (85.7%) with a **4**.

Overall we rate our platform with a light **4**

Additionally, the surveyed left some comments that could help us improve our interface. With suggestions such as "dark mode should also change the color of the map, since it takes up most of the screen", "it should be possible to read information by hovering the cursor over density data", "allow combining filters to show only journeys in terms of time and at stations close to POIs or tool to compare two time periods simultaneously" or even grid modification commands. One user also stated that "The usefulness of this interface is for a very specific audience who only wants to see data on where bikes are rented. It's not very easy to associate with the weather".

8 Conclusion

The interface, rated as **intuitive** (4/5 by users), allowed these insights to be explored through temporal filters and layers of POIs. However, it could be enriched with direct overlay of weather data, heat maps for density and time comparison tools (e.g. July vs. January), or even some small details, such showing information on the spot or allow a the map to change colors, based on user preferences.

This interface reinforces how visual analysis can transform raw data into strategic decisions for more sustainable cities, highlighting the importance of user-centered design (simplicity + personalization) and the integration of multiple contexts, such as climate, infrastructure, habits, etc.

This adaptable framework enables analysis of urban mobility systems globally through simple dataset replacement while retaining its core functionality. The solution offers standardized spatiotemporal visualization and contextual data integration (e.g., transit hubs, events), allowing non-technical users to generate localized insights. This creates a versatile toolkit for data-driven transport planning across diverse cities.

This work was greatly assisted by the translation platform Deepl Translator [5].

References

1. Ahillen, M., Mateo-Babiano, D., Corcoran, J.: Dynamics of bike sharing in washington, dc and brisbane, australia: implications for policy and planning. International Journal of Sustainable Transportation **10**(5), 441–454 (2016)
2. Capital Bikeshare: System data - capital bikeshare. `https://capitalbikeshare.com/system-data` (2025), accessed: 2025-05-14
3. CityVis: Cityvis: A curated collection of city data visualizations. `https://www.cityvis.io/collection/` (2024), accessed: 2025-05-14
4. Geofabrik: North america / united states data - geofabrik. `https://download.geofabrik.de/north-america/us.html` (2025), accessed: 2025-05-14
5. GmbH, D.: Deepl translator. Disponível em: `https://www.deepl.com/translator` (2023), `https://www.deepl.com/translator`, acesso em: 18 mai. 2023
6. Kim, K.: Investigation on the effects of weather and calendar events on bike-sharing according to the trip patterns of bike rentals of stations. Journal of Transport Geography **66**(December 2017), 309–320 (2018). `https://doi.org/10.1016/j.jtrangeo.2018.01.001`, `https://doi.org/10.1016/j.jtrangeo.2018.01.001`
7. OpenAI: Chatgpt: Gpt-4 language model. `https://chat.openai.com` (2025), accessed: 2025-05-14
8. Shaheen, S., Guzman, S., Zhang, H.: Bikesharing in europe, the americas, and asia. Transportation Research Record (2143), 159–167 (2010). `https://doi.org/10.3141/2143-20`
9. Shaheen, S.A., Martin, E.W., Chan, N.D., Cohen, A.P., Pogodzinski, M.: Public bikesharing in north america during a period of rapid expansion: Understanding business models, industry trends, and user impacts. Technical report, Mineta Transportation Institute (2014)