

# Interactive Data Visualizations for Bike Sharing Data

Silvana Belegu

November 2024

## 1 Overview

This project was about implementing 2 visualization methods in react using D3 libraries. The project was realised using the steps below:

1. Set the stage: fork the git repo and start the report on Latex
2. Decide the second visualization which will be implemented.

To make an informed decision on the visualization method that is going to be selected, let's have a closer look at the dataset. The bike sharing dataset contains multivariate data capturing various factors that influence bike rental patterns over time. Each data object represents a specific time interval (such as an hour or a day) and includes the following key variables; datetime, season, holiday (yes/no), working day (yes/no), weather situation, temperature, humidity, wind speed, etc.

The dataset has information on how temporal and weather conditions affect bike rentals, so the visualizations will be to emphasize the correlation of the bike rental amount with temperature (scatterplot) and on the other side there will be a multi-line chart which will show the correlation between the time of the day and the bike rental amount.

3. Implement the scatterplot and make it brushable
4. Implement the heatmap visualization and make it brushable as well
5. Synchronize the interactions with each other
6. Finish the report and push everything in the repo

## Scatterplot and Multi-Line Chart: Visual Design Justification

### 1. Scatterplot

#### Visual Encoding

- **Axes:**
  - **X-axis:** Encodes *Temperature (°C)* as continuous numerical data.
  - **Y-axis:** Encodes *Number of Rented Bikes* as continuous numerical data.
- **Points:**
  - Each point represents an observation in the dataset.
  - **Color:** Encodes *Visibility levels* using a continuous gradient, aiding in visualizing environmental impacts on bike rentals.
  - **Size:** Uniform size minimizes clutter and ensures readability.

### User Tasks

- Explore the relationship between temperature and bike rental counts.
- Identify clusters, outliers, and trends in the data.
- Investigate the influence of visibility on bike rentals using the color encoding.
- Perform dynamic brushing to filter points and synchronize with the multi-line chart.

### Pros and Cons

#### Pros

- Effectively visualizes relationships between two continuous variables with minimal clutter.
- **Color encoding:** Enhances depth by adding visibility as a third variable.
- **Interactivity:** Brushing enables filtering and linking with the multi-line chart.

#### Cons

- Overlap of points in high-density regions may obscure trends.
- The color gradient may be less effective for users with color vision deficiencies.
- Brushing highlights regions but does not inherently summarize broader trends.

## 2. Multi-Line Chart

### Visual Encoding

- **Axes:**
  - **X-axis:** Encodes *Hour of the day (0-23)* as discrete time intervals.
  - **Y-axis:** Encodes *Number of Rented Bikes* as continuous numerical data.
- **Lines:**
  - Each line represents one of the four **seasons** (*Spring, Summer, Fall, Winter*).
  - **Color:** Each season is assigned a distinct color for clear differentiation.

### User Tasks

- Compare rental patterns across seasons and time intervals.
- Identify peak hours for specific seasons.
- Explore brushed data from the scatterplot and examine their seasonal or hourly patterns.

### Pros and Cons

#### Pros

- **Temporal trends:** Captures hourly and seasonal patterns effectively.
- **Color-coding:** Differentiates seasonal trends with accessibility-friendly colors.
- Brushing enables filtering and synchronization with the scatterplot.
- The legend makes it clear which line belongs to each season.

## Cons

- Overlapping lines during peak hours may reduce clarity.
- Scalability is limited; adding more categories (e.g., months) may clutter the chart.
- Small screens or constrained chart width may obscure detailed trends.

## 2 Summary

Feature	Scatterplot	Multi-Line Chart
<b>X-axis</b>	Temperature (continuous)	Hour of the day (discrete)
<b>Y-axis</b>	Number of Rented Bikes (continuous)	Number of Rented Bikes (continuous)
<b>Color Encoding</b>	Visibility (continuous gradient)	Seasons (categorical color scheme)
<b>Interactivity</b>	Brushing, hovering, and filtering	Brushing, hovering, and filtering
<b>Strengths</b>	Shows multi-variable relationships clearly	Highlights temporal trends across seasons
<b>Weaknesses</b>	May suffer from point overlap	Aggregation of information hinders accessing individual points

Table 1: Comparison of visual encodings for scatterplot and multi-line chart.

## 3 Conclusion

The combination of scatterplot and multi-line chart designs provides complementary insights into the Seoul Bike Sharing dataset. The scatterplot excels at exploring multi-variable relationships, while the multi-line chart captures temporal patterns and seasonal differences. By integrating synchronized brushing, both visualizations enable dynamic filtering, ensuring a cohesive and interactive exploration experience. Probably due to the fact that I have aggregated the amount of bikes rented for each hour in a season, I couldn't find a way to synchronize the two brushes with each other. Future enhancements could address overlapping elements and improve color accessibility to cater to a broader range of users.