

# Report

## 1. Introduction

In this project are employed two visualizations to represent multivariate data from a bike-sharing dataset: a scatterplot with 2D brush interaction and parallel coordinates.

## 2. Scatterplot

The scatterplot allows us to visualize the relationship between two variables from the dataset. In this visualization:

- **Axes:** The x-axis and y-axis represent two continuous variables of interest, such as the number of bikes rented per hour and the temperature.
- **Points:** Each point in the scatterplot represents a single observation, meaning a specific hour in the dataset.
- **2D Brush interaction:** The interaction allows users to select and highlight regions of the scatterplot. Selected points are highlighted.

## 3. Parallel Coordinates

Parallel coordinates are used to visualize multivariate data by plotting each variable on a separate horizontal axis and connecting observations with lines:

- **Axis:** Each axis represents a different variable from the dataset.
- **Lines:** Each line crossing the parallel coordinates plot represents an observation, showing its value across all the variables.

## 4. Justification of Visual Design

### 4.1. Scatterplot

- **Data properties:** Scatterplots are effective for representing the relationship between two quantitative variables, making it easy to spot trends, clusters, and outliers.
- **User Tasks:** User can explore pairwise relationship and distributions, and the 2D brush interaction enhances this by allowing detailed selection of data points.

### 4.2. Parallel Coordinates

- **Data properties:** This visualization is suitable for high-dimensional data, enabling the observation of relationship across variables simultaneously.

- User Tasks: Users can identify patterns, correlations, and outliers across multiple variables.

## 5. Pros and Cons of design Choices

### 5.1. Scatterplot

Pros:

- Simplicity: Easy to understand and interpret for pairwise relationships.
- Interactivity: Enhances user engagement and allows for precise data selection.

Cons:

- Dimensionality: Limited to two variables at a time, which may not capture the complexity of multivariate relationships.
- Clutter: With high number of points, the scatterplot can become cluttered and hard to interpret.

### 5.2. Parallel Coordinates

Pros:

- Multivariate Analysis: Capable of displaying and comparing multiple variables simultaneously.
- Pattern Identification: Effective for spotting correlations and trends across several dimensions.

Cons:

- Complexity: Can be difficult for users unfamiliar with the technique to understand.
- Overlapping lines: With large datasets, lines can overlap, making it hard to distinguish individual observations.

## 6. Conclusion

The combination of scatterplot with 2D brush interaction and parallel coordinates allow for a comprehensive exploration of the bike-sharing dataset. This visual design balances simplicity and depth.