

# High Performance Data Analytics and Visualisation project

Giorgio Bettonte

November 2024

## 1 Introduction

This work deals with the visualization of the information contained in the dataset "SeoulBikeData". My proposed visualizations relies on three different scatterplots and on a controlbar. In the following sections I will explain the motivation of my choices and how this project can be improved

## 2 Welcome message

The web page contains a very simple welcome message and some information about the usage of the application.

There is the source reference for the code; the explanation of the three different scatterplots and of the used colors.

### Seul Bike Dataset

Code: [Giorgio Bettonte Github](#).

It is possible to select variables with the controlbar.

The third scatterplot allows to brush; the most left-one highlight the selected items, while the central scatterplot displays a zoom on that items  
Colors mapping: Yellow = Summer; Green = Autumn; Red = winter; Blue = Spring

## 3 Components

In this section there is the description of each component in my application.

### 3.1 ControlBar component

This controlbar allows the user to select the numerical variables of the dataset for axis; then the scatterplots will update according to the selection.

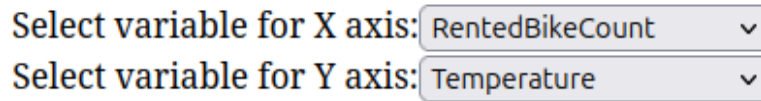


Figure 1: Controlbar for axis

## 3.2 Scatterplots

There are three scatterplots in line; at the begin of the application, they all display the same cloud of points, with different colors. I chose to use only scatterplots since I think that they are as simple as actually effective to transmit information.

As I said above, with the controlbar is possible to select only the numerical variables of the dataset; anyway, each point is characterized also by some categorical variables; to give information about them I decided to assign colors according to the season and to implement the possibility to visualize all the information of a point in the scatterplots simply clicking on it, as it is shown in the figure.

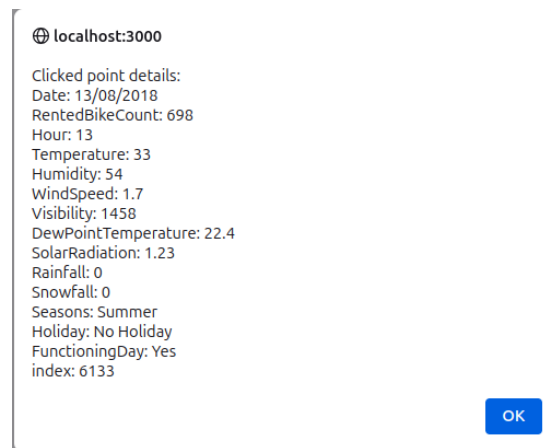


Figure 2: Point details

### 3.2.1 First scatterplot

This scatterplot highlights the items selected on the third scatterplot, while keeping the others in black. In this case, there isn't any zoom.

### 3.2.2 Second scatterplot

This scatterplot is dedicated to the zoom of the items selected on the third scatterplot. I decided to implement this "zoom" because only highlight the

items in the scatterplot is not sufficient to have a clear idea of their reciprocal disposition.

### 3.2.3 Third scatterplot

This scatterplot is positioned on the right extreme of the screen and allow to brush.

At the beginning, this is the only chart engaged with colors.

## 4 Example of flow

Here some images illustrating a possible execution flow:

### 4.1 Selection

In the image [3], some elements are selected using a window.

### 4.2 Highlight

In the image[4],selected items are highlighted in the first plot and zoomed in the second

### 4.3 Change variables

Now, suppose that the user would like to visualize the selected items in a different configuration (different variables for axis). As it is shown in the image [5], we still see the items selected in the previous selection, and how they are dispose with respect to different variables

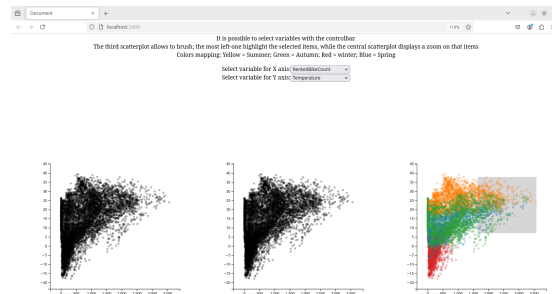


Figure 3: Select elements

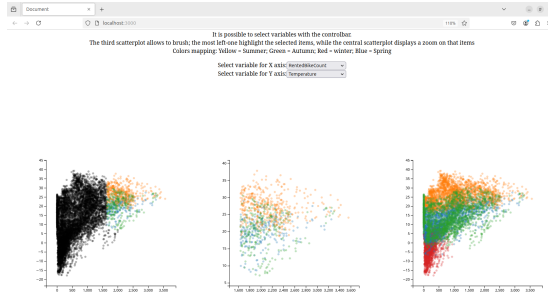


Figure 4: Highlight selected elements

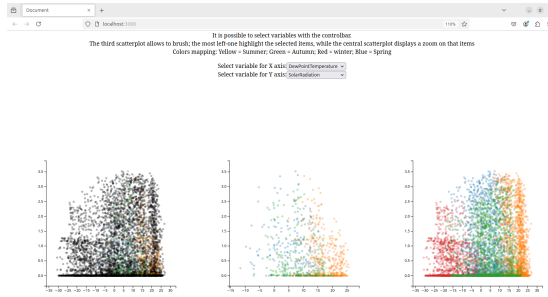


Figure 5: Highlight selected elements in different configuration

## 5 What can be improved?

Currently, the transition induced by the brush selection is a bit slow, but still acceptable. For a larger amount of data, it can be too slow and consequently annoying for an user. Surely, this aspect would be the first one to improve in my opinion.

Secondly, it would be interesting to implement other types of chart and not only scatterplots. For this work, I decide to limit myself only on scatterplots, but it could be interesting to have other representations of data.