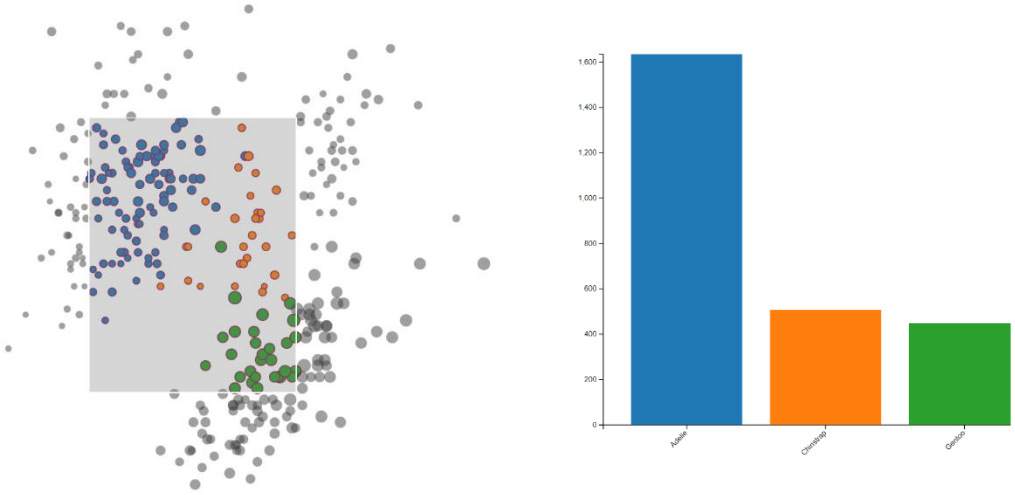
Exercise 8 **(20 points)**

Due: 03.07.2022 8AM

Contributor 1: Mohammad Nomaan Husain
Contributor 2: Anurag Pacholi



Task 1: Linking and Brushing

For this exercise, your task is to extend the scatterplot from exercise 4 with a brushing technique and dynamically create a Barchart based on the given selection.
The bars should display the culmen depth summed up by species based on the selected penguins.
A new selection should trigger:

- coloring the selected and not-selected points as displayed (see the index.css)
- the creation of new bars and scales

Your task is to finish the implementation such that opening the *index.html* shows the scatterplot as depicted in figure above. To finish the implementation, follow the steps described as comments within the dedicated file.

Tips:

- Use the d3-brush package to add a selection rectangle
- You can add event handlers to the brush (similar to previous events), see the possible events in the d3-brush documentation
- Based on the selection, modify the attributes of the points in the scatter plot
- Display a bar cart with the filtered data (you may either create a new one, or display/hide it)

Task 2: Multivariate Data **(6 points)**

Task 2a) **(1 points)**

Using your own words, describe what **multivariate** data is.

Answer: Multivariate data refers to a type of data that contains multiple variables or attributes simultaneously. In other words, it is a collection of observations or measurements that are associated with more than one characteristic.

Task 2b) **(2 points)**

Give one example of a multivariate visualizations techniques for each of the three marks introduced by *Bertin*: **Point**, **Line** and **Area**. Which additional visualization techniques exist?

Answer:
Point: Scatter Plot Matrix
Line: Parallel Coordinate Plot
Area: Mosaic Plot
Other: Stacked Area Chart, Glyph Plots, Tree map chart, Heat Map, etc.

Task 2b) **(3 points)**

1. For a parallel coordinate plot**,** why do visible patterns **depend on the order of dimensions**?
2. Why is this **not** the case for a **scatterplot matrix**?
3. Can you come up with one **possible criterion** to choose the **best ordering** for parallel coordinates (I.e. if multiple orderings are available, which one should I choose)?

Answer:

1. The visibility of patterns in a parallel coordinate plot depends on the order of dimensions because it determines the arrangement and intersection of lines representing data points. When the dimensions are ordered in a particular way, it affects the way the lines cross and align with each other, making certain patterns more visible or pronounced.
2. This is not the case for a scatterplot matrix because each scatter plot within the matrix represents a distinct pair of variables.
3. One possible criterion to choose the best ordering for parallel coordinates is to prioritize the arrangement that maximizes the visibility of meaningful patterns or relationships among the variables. Minimizing crossings, overlapping, and seeking visual consistency would help us to choose the best ordering from available orders.

After completing your answers, export the docx-File to PDF and upload it alongside the source code files.

Submission: Zipped folder including all necessary files and a PDF of the completed written exercise.

Please form a group of **2 Students**. Only 1 member of the group must submit the exercise in ILIAS. Please state the collaborators in the beginning of the document.