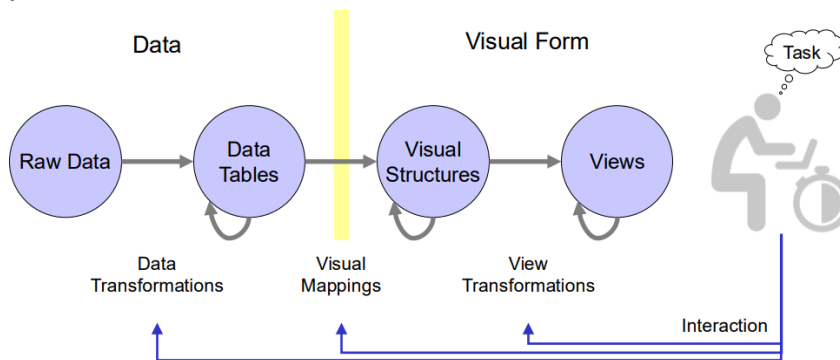


InfoVis A2

Task 1

1)



Raw data: This is the data when it's not formatted or stripped.

Data transformations: The process of cleaning, filtering and structuring the raw data.

Data tables: This is the organized data, with rows often represented as cases/items/objects and columns as variables/attributes. There are also values and meta data.

Visual mapping: Trying to map the data in the table to 2D and sometimes 3D.

Visual structures: Techniques to improve the visual representation for faster and clearer/better understanding, like spatial substrate (axes/coordinates/layout), visual elements (points, lines, regions) and graphical properties (color, size, orientation).

View transformations: Interactions with the visualisation like, zoom, cut, move, pop-up windows.

View: What the user finally sees on the screen.

Interaction: The user interacts with the data. For example the user can adjust the view by zooming or panning, change the mapping from bar chart to scatter plot, filter the data.

The goal is to better understand the data and even see connections. A good visualisation will help the user to faster interpret the data, convey more distinctions, be less distracted and this hopefully leads to less errors.

The model's greatest strength is its flexibility. It doesn't tell you what to build, it provides the blueprint for how to build it systematically. If a visualisation isn't good enough the model can help pinpoint where the error is. If the data is too messy, look at data transformations, is the visual confusing? fix the visual mapping.

2)

Supported interactions are:

Filtering: Define a minimum and maximum threshold to filter out data points.

Dynamic Querying: The visualization updates in real-time.

Brushing: Selecting a range on a slider can "highlight" corresponding data points in another linked view.

Some improvements could be to show the data points within the slider. This would help the user to better understand the data before filtering, like where the data is located within the slider or how many data points are there. This can either be done with a line, small bars or just significant values like smallest/lowest, average, biggest/highest as points/bars/colors.

Pros:

- 1) Users immediately know how it works.
- 2) It prevents users from entering "invalid" data.
- 3) They provide Immediate Feedback.

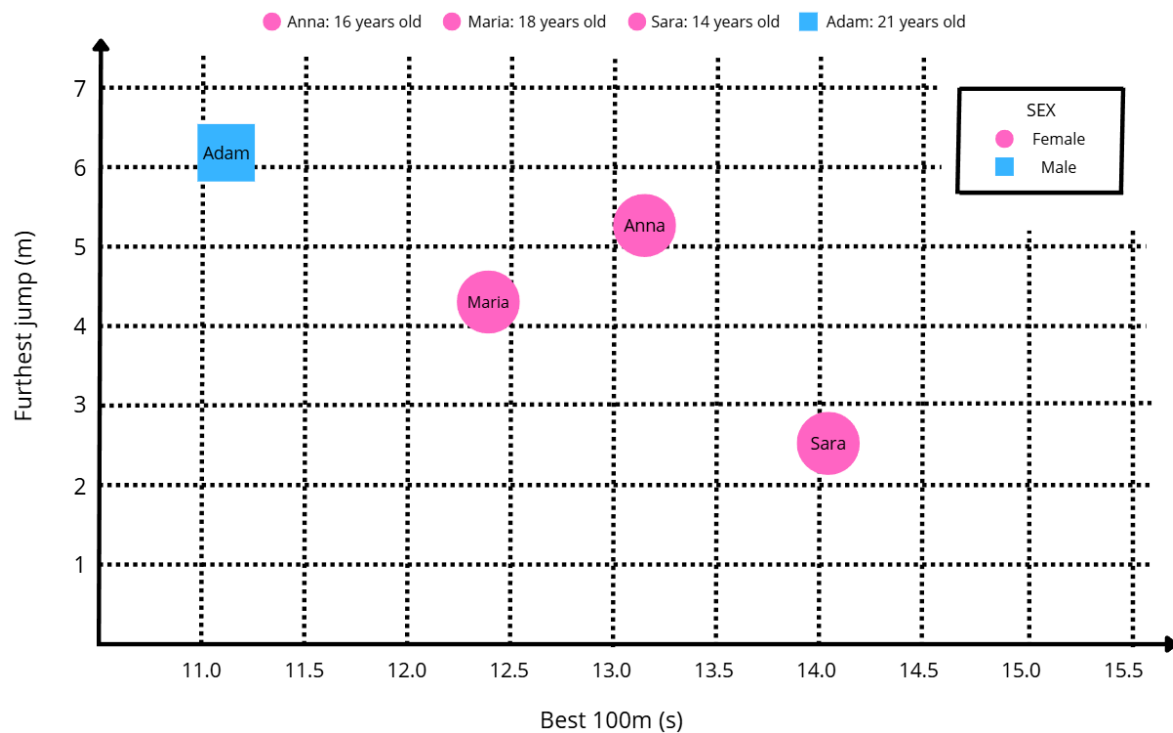
Cons:

- 1) It can be difficult to select an exact value, depending on how the slider is constructed.
- 2) The data is hidden unless some improvements are made to the slider.

3)

It's called brushing and linking. This technique helps highlight the data points which in turn helps the user to more easily see connections between multiple dimensions. By highlighting data points in multiple dimensions it becomes easier to identify their characteristics which can be good if looking at outliers.

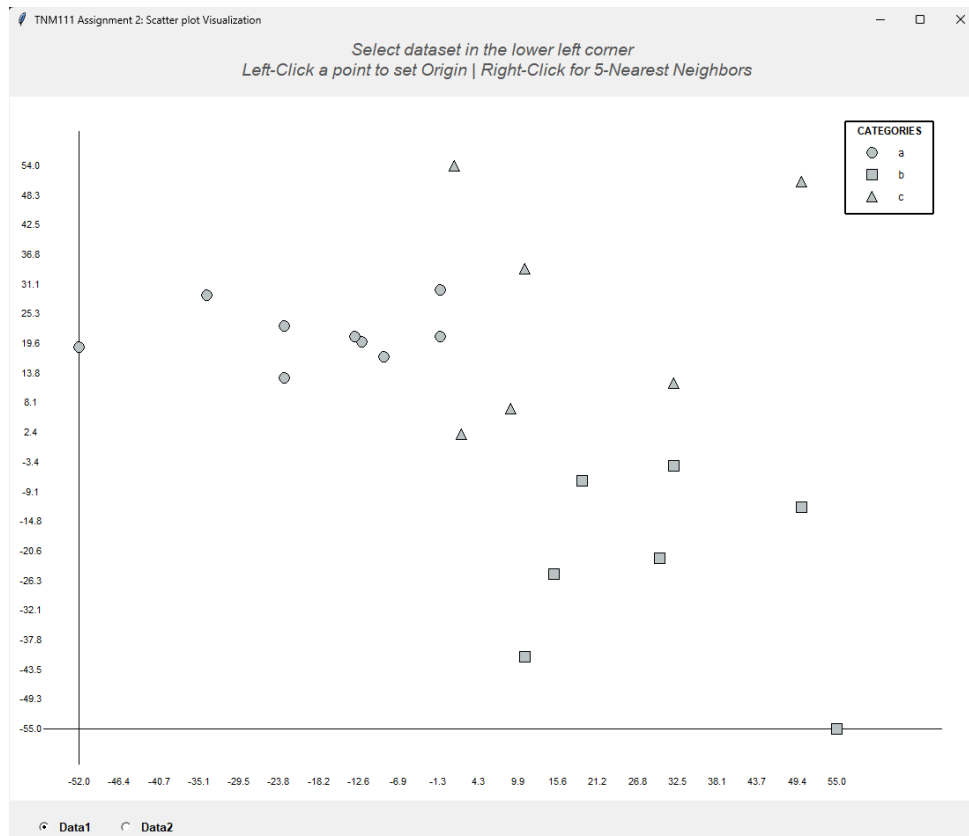
Task 2



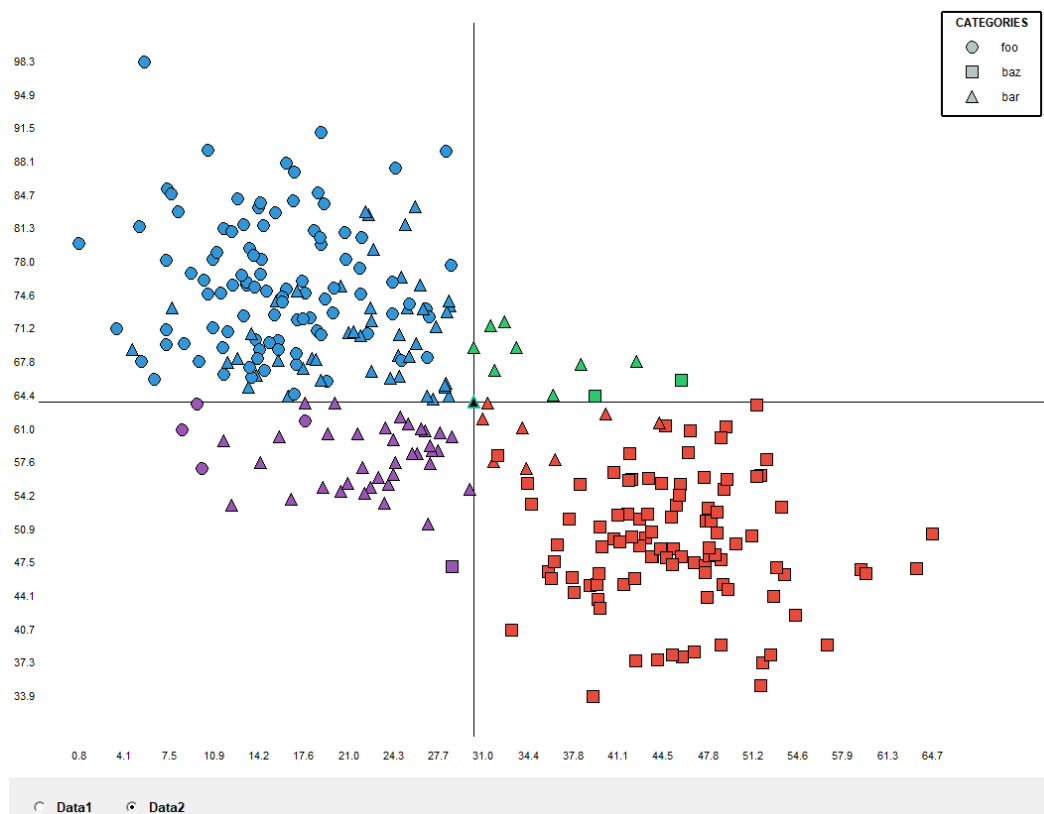
The values aren't shown in a way that you can precisely determine their values but this visualisation makes it not as crowded with lines and you often get within 0.1 from the exact value. The attribute for sex is visualised with both color and shape to make it clear while furthest jump and best 100m are displayed with the help of axis. The age of the contestant was hard to implement in a well visualised way so they are displayed at the top of the plot.

Task 3

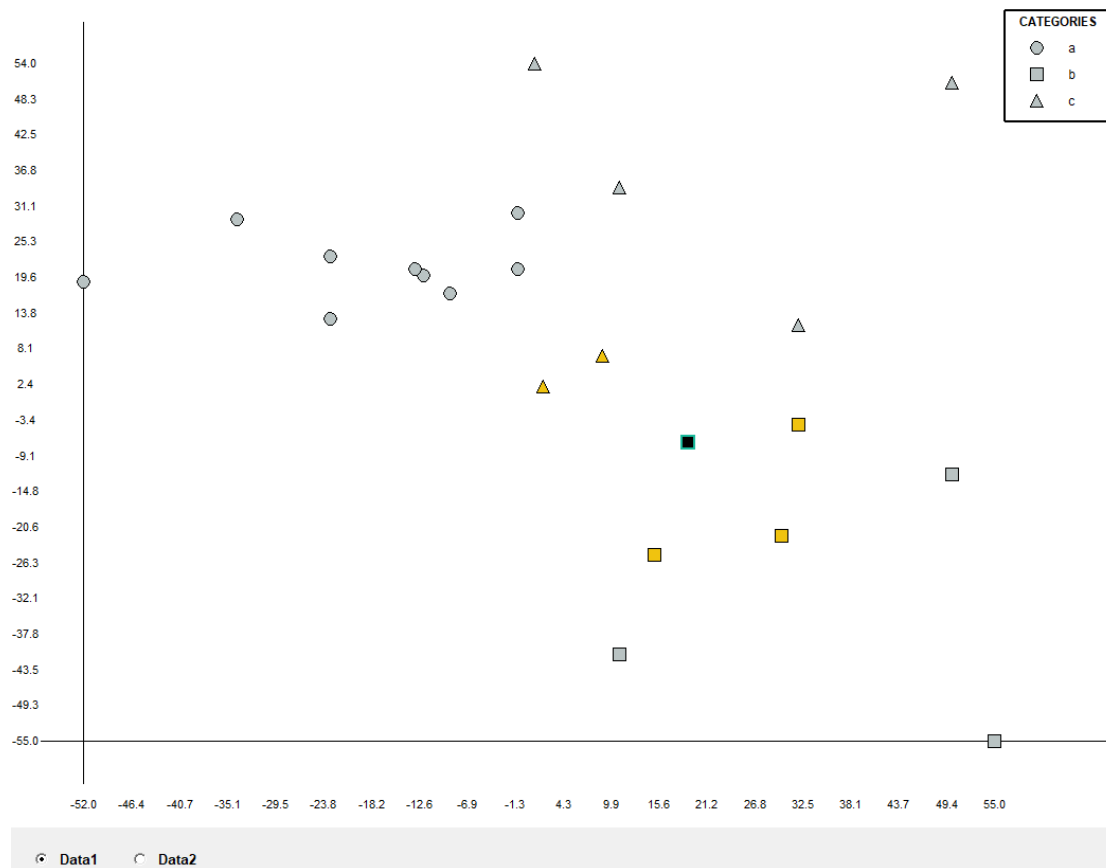
No click:



Left click:



Right click:



In the scatter plot there are different shapes for the different categories which helps the user to quickly distinguish between them (using shapes like this would not work if there are too many categories). The axis displays 20 values for a balance between being able to read exact values and not being too crowded. Being able to interact with the plot helps the user to better understand the data. By left clicking the mouse to set a data point to the new origin the user will be able to better see where the data point lines up on the axis as well as the data will be divided into quadrants with a different color for each, helping the user more easily see if a data point is bigger or smaller than the one selected. By right clicking the mouse the user will see the 5 closest neighbors to the data point to better show similarity.