

Write a paragraph for each visualization tool you use. What was easy? Difficult? Where could you see the tool being useful in the future? Did you have to use any hacks or data manipulation to get the right chart?

Description of the **Technical achievements** you attempted with this visualization. Some ideas include interaction, such as mousing over to see more detail about the point selected.

Description of the **Design achievements** you attempted with this visualization. Some ideas include consistent color choice, font choice, element size (e.g. the size of the circles).

Python

Pandas + Matplotlib + Seaborn

Python is considered as a popular programming language that has strong visualization capabilities. Matplotlib is the most widely used and the first library for plotting in Python. Seaborn is another Python visualization library based on the Matplotlib library.

Technical achievements

One of the technical achievements was filtering the species we wanted and making a subset. At first, the legend was not shown in the correct part and was on the points of the scatterplot. I tried to solve this by using the 'plt.legend' function and by specifying its location ('loc' command) to be upper right.

Design achievements

Assigning colors to the species used in the graph. Styling with Seaborn to show a grid. Adding axes labels for increasing the plot interpretability.

R

R + ggplot2 + R Markdown + RStudio

R is considered as a programming language for statistical computing and the visualization of data. R Markdown is a format of document that compiles to html or PDF and lets us include the output of R code in the document. ggplot 2 package is an open-source popular visualization package used in R. RStudio is another tool I used for working with the R programming language. I used ggplot2 in R to represent values for two different numeric variables and observe their relationship. 'geom_point()' function was also used for the functions for color and size.

Technical achievements

One of the technical achievements was to filter the specific columns we wanted for the visualization. Doing the color-coding based on the penguin species and encoding the size of points with the 'bill length' variable were other technical achievements.

Design achievements

Providing x and y axis labels for a more informative visualization. Differentiating different colors for the different species we had. Adding the alpha for transparency to the points while maintaining their color.

Matlab

Matlab is a programming language for analyzing and designing systems. Also, Matlab enables us to create custom charts and can be considered as a biovisualization tool for biomedical engineering fields. By the use of the 'gscatter(x,y,g)' command, I could plot a scatter plot of x and y grouped by g.

For visualizing this plot in Matlab, I used the 'figure()' command and specified the plot shape and size. Then, for creating the scatter plot, the gscatter() command was used and I specified the x (flipper length) and y (body mass) axis and differentiate the dots by different species. I tried to make the colors similar to the template we had by the use of RGB values for 'Adelie', 'Gentoo', and 'Chinstrap' species. Then it created the plot for me, but I had to add x and y labels. I also added the grid to be on and added a legend at the end. I found the process of visualizing this plot by Matlab not the easiest way, but a doable one by considering specifications and knowing the required commands.

Technical achievements

Creating a scatter plot using 'gscatter' function for seeing the relationship between flipper length and body mass of the penguins. Enabling a grid for this scatter plot for helping in a more precise analysis. Configuring Matlab plot with different dimensions to find the best choice.

Design achievements

Differentiating the penguins by their species with the specific colors was a little challenging. Adding axis labels and title for a more informative plot for presenting. Adding legends for the interpretability of the plot. Adding a grid for a better layout.

Tableau

Tableau is a data visualization tool that is used for data analysis and business intelligence. Tableau can be used when we are working with larger datasets. It has advanced visualization capabilities and I found it easy to understand, implement, and work with as a powerful visualization tool. It also takes much less time to create graphs with Tableau.

Technical achievements

I load the 'penguin.csv' file to Tableau, select the scatter plot, drag 'Flipper Length' to the columns and 'Body Mass' to the rows section. A challenge that I faced while plotting this graph was that after placing a measure in columns and rows shelf, I saw 3 single points (representing 3 species). Although I wanted all the values to show on the graph. I recognized the points as showing the sum of all values for both measures. So, I used the Analysis part of Tableau at the top of it and clicked on Aggregate measures.

Design achievements

Trying to assign colors similar to the colors used in the template. I also specified the shape to be a circle. I also didn't know how to imply 'bill length' to the graph. I figured out that I can drag 'bill length' to the size shelf in the Marks section. My challenge was how we could color-code the species in the plot. I identified that I have to drag 'Species' columns to the color section. I also did not know how to change the ranges of the x and y axes. So I tried double-clicking an axis to open the edit axis and selected a custom option and added the ranges manually.

D3

JavaScript + D3

Javascript is a widely used programming language that is mostly used in web development and works with html files. D3.js is an open source JavaScript library for visualizing data.

I think plotting this graph with D3 was the most challenging and time-consuming one. I used the template of plotting a scatter plot from the D3 website as my reference.

Technical achievements

The technical achievement I had was to be able to first give the input and load the penguin.csv data in JavaScript. I used several ways and with the use of d3.csv function, I was able to read the dataset. After loading the dataset, I only tried to modify the code I had to make the graph as similar to the reference image I had. Alao, I learned another way to open a new terminal in the visual code studio by writing 'python3 -m http.server'. With this method, I was able to run my local web server and the output of my code.

Design achievements

I tried to modify the x and y axes data ranges to be consistent with the reference image I had. I learned to use the 'ticks()' command to specify the ranges on the x and y axis. I added x and y labels for a better understanding of the plotted graph. Using 'bill length' for circle sizes was another implemented design. Also, plotting different species with a specified color was another design achievement I had regarding this plot.

Flourish

Flourish can be used for creating interactive content and visualizations. Flourish visualizations are useful for web articles and when we want to have a creative data visualization. I found this software very helpful, easy to use, and understand. Flourish also has a high potential to form different creative visualizations in much less time.

Technical achievements

I load the 'penguin.csv' file to Flourish, select the scatter plot, and assign the columns to the x and y axis. I assigned the x and y values by selecting the corresponding columns. My challenge was how to plot the species in the graph. So, I figured out I have to use the part named 'color' and specify the species column. Actually this was my achievement to utilize the 'species' column to color-code the points based on different species in a way that each species will be represented by a distinct color.

Design achievements

Adjusting the dot colors to the colors to the template on the Dot styles section.e we had. Setting the scatter plot to show data points as dots (circles). Setting the opacity of dots to be 0.8 in the Dot styles section. Adding a legend to the graph. Adding the x and y labels in a way that x-label is at the bottom of the graph and y-label be vertical at the left side of the graph.