Module 7 Live Coding Assignment: Data Visualization Strategies

Throughout this assignment, please remember to import all the necessary libraries at the beginning of every exercise.

Question 1

On the same graph, plot the functions $\sin(t)$, $\cos(t)$, and $\tan^{-1}(t)$ where t is a vector from -10π to 10π with spacing 0.02. Label the x-axis and the y-axis "x-coordinate" and "y-coordinate" respectively and set the title of the graph to be "Some trigonometric functions".

Question 2

Set a random seed equal to "38472" and generate two arrays "x" and "y" with 15 entries each. Set the size of the marker "s" equal to 90. Plot two subplots with the defined data next to each other. Set the marker for the first plot to be a filled circle and a "plus sign" for the second one. Repeat the exercise above, but now set the sequence of colours for your scatter plot to be $\sqrt{x^2 + y^2}$.

Question 3

For this question, you will need to use the "seaborn" library. Load the seaborn dataset *tips* and visualize the first 5 rows of this dataset. Set the style of your plot to "ticks" and the colour palette to "pastel". Display a nested boxplot to show the bills and the tips by day and time, and display whether the server was a smoker or not for each case. Repeat the same exercise, but this time, draw a set of horizontal bars.

Question 4

Load the dataset *titanic* from seaborn and visualize the first 5 rows of this dataset. Set the style of your plot to "darkgrid". Plot a scatter plot by passing the column names "fare" and "age" to the "x" and "y" axes, respectively. Make sure your plot shows the regression line.

Question 5

Load the dataset *iris* from seaborn and visualize the first 10 rows of this dataset. Set the style

of your plot to "darkgrid" and your palette to "muted". Use the pandas library to un-pivot the columns "species" in your dataset to "measurements". Produce a swarmplot by passing the column names "measurements" and "value" to the "x" and "y" axes, respectively.

Question 6

Set a random seed equal to 2973 and define a random matrix "M" with dimension 10×15 and print it to screen. Produce two heatmaps of "M" using both the seaborn and the plotly library. Set the title on both plots to be "Matrix M", the x-axis to be labeled "x values" and the y-axis to be labeled "y-values".

Question 7

Import the dataset "tips" from plotly and set it equal to a variable called tipsData. Visualize the first 5 lines of your dataset. Produce a *multiple* violin plot from the tipsData dataset that has the tip amount on the y-axis and on the x-axis smoker and non-smoker workers grouped by sex.