Sean Toon

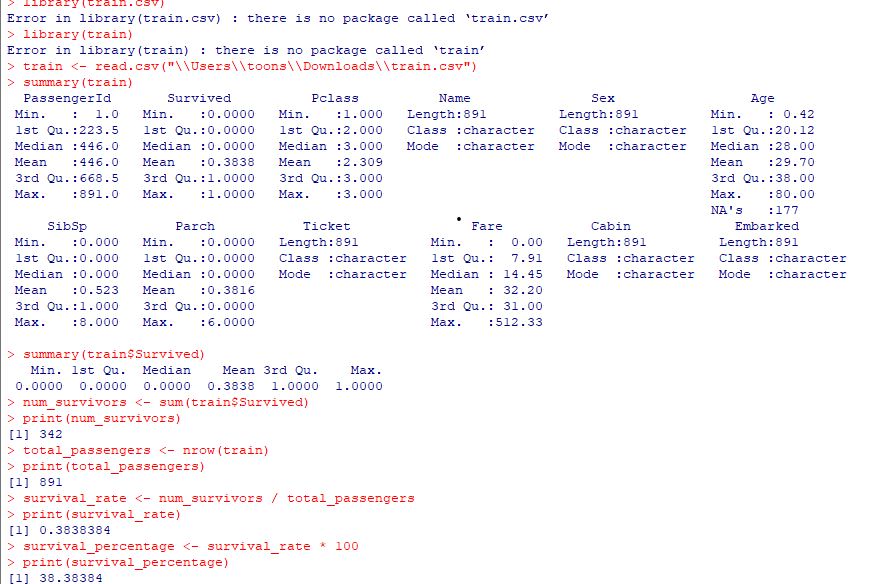
IT-460-18500-M01 Machine Learning 2024 C-5 (Sep - Oct)

9/28/2024

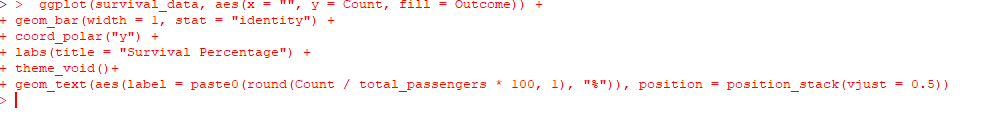
Module Three Lab and Report – R Data Structures and Visualization

The data set used for this lab is a training set that consists of passenger features related to the disaster of the Titanic. Some of the information that was included is name, age, gender, socio-economic class and whether the passenger survived the wreck or not.

Firstly, I loaded the data in R just to get another idea about how the data was stored in the csv files and to make sure to use the correct variables further along the lab. Before creating any visualizations, I made a few manual calculations to get some more practice with R and to confirm the data that was provided.



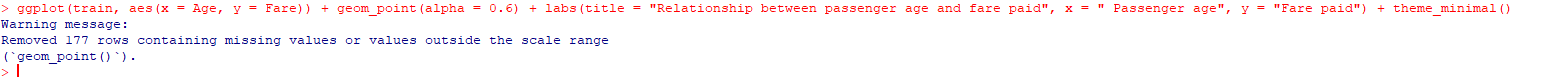
Visual for the above survival percentage commands and the commands used to generate the chart:



A blue and red circle with a red circle

Description automatically generated

Lastly, I showed the relationship between age of passenger and fare paid with the following command in a plot:



A graph with dots and numbers

Description automatically generated