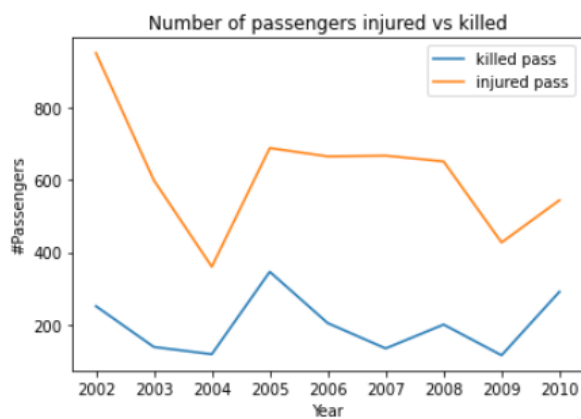


Assignment - 1

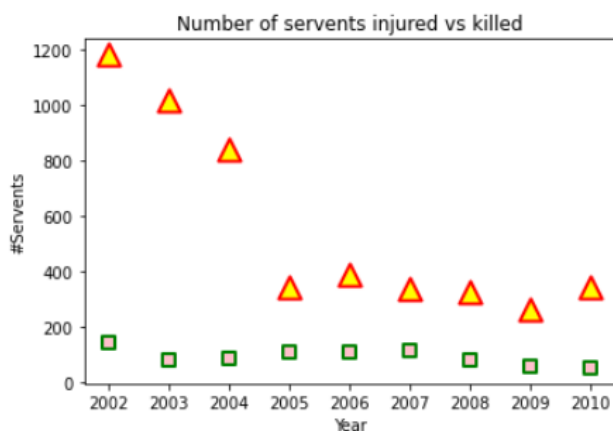
This is the assignment page for DS200. I have taken a public dataset from data.gov.in for the analysis done in this assignment.

The data is related to the number of railway employers and passengers who were killed due to various railway accidents such as unprecedented movements of railway vehicles, failures of railway equipments, train accidents and other miscellaneous accidents.

The bar plot shown below indicates the number of passenger who were killed vs who were injured. The trends of the two plots keep varying similarly showing that there is an equal chance of either getting injured by a railway accident or getting killed.

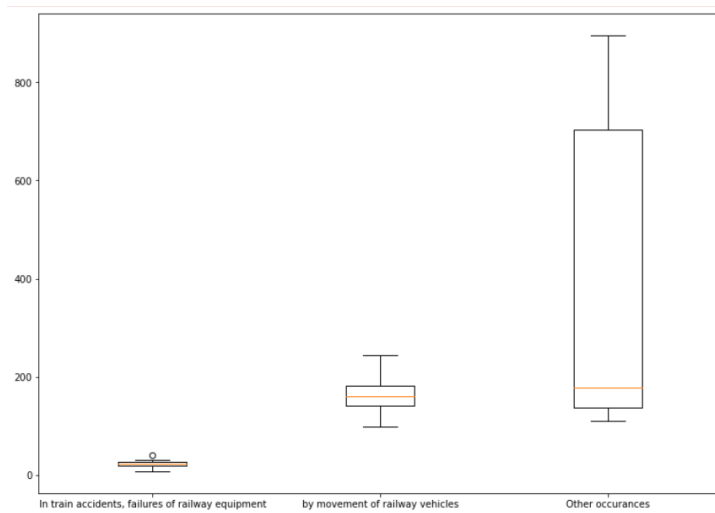


The following plot shows a scatter plot indicating the injury and deaths of railway servants. There is a decreasing trend over the injuries in each year, indicating that measures are being taken to prevent accidents and there is increasing awareness in the community regarding such accidents.



We further analyse the data for the injuries to predict the particular cause of it. From the box plot below it is clear that injuries are caused by unusual occurrences on railway premises not connected

with the movement of railway vehicles, which is indicated by the third bar showing the maximum range.



Thus, we analysed the risks involved in railway accidents and we were able to conclude the reason that causes the railway servants to get injured or killed.