

Data Visualization of Titanic

Name:

Student ID:

Github:

<https://github.com/adeelsardar/datavisualization>

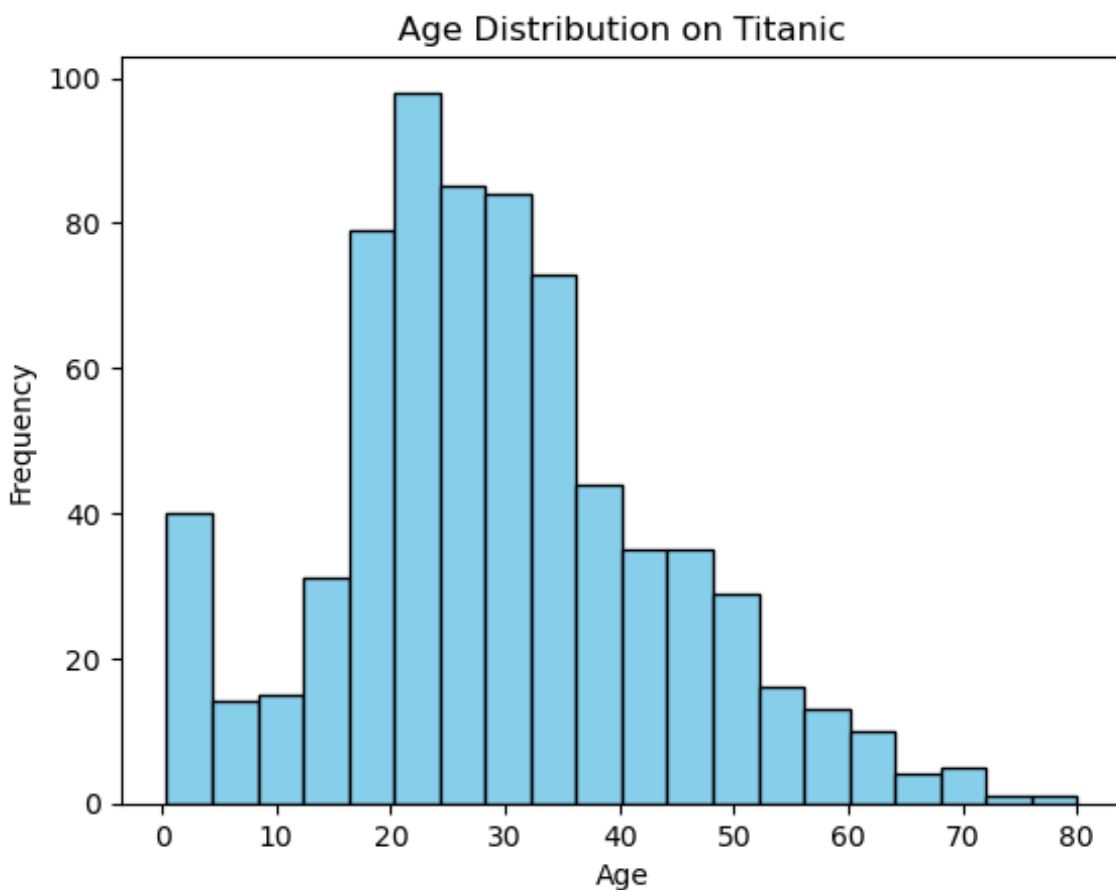
Dataset:

<https://www.kaggle.com/competitions/titanic/data>

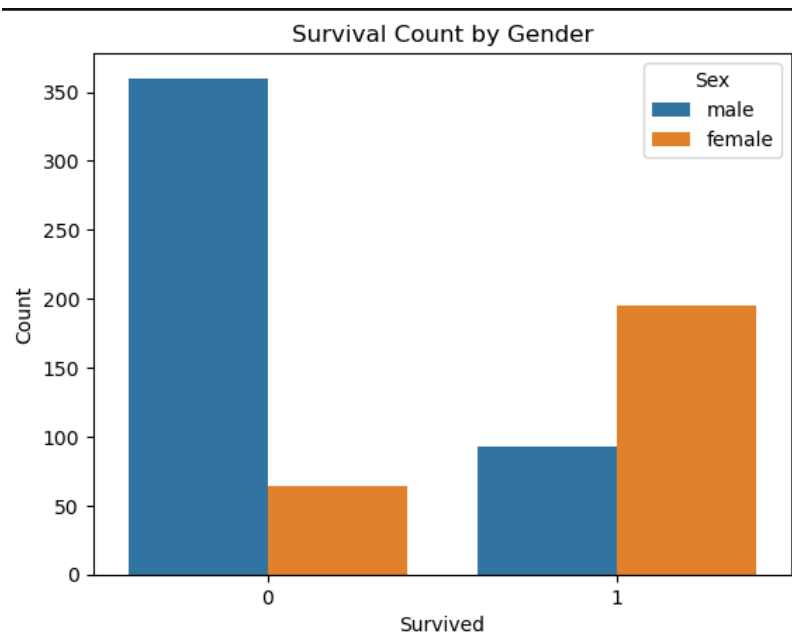
Report of Titanic Dataset:

1. In this Task we downloaded the dataset from of titanic from Kaggle
2. First of all we loaded the dataset into data frame using pandas.
3. Then we made the function for getting the descriptive summary of dataset.
4. In the next step we cleaned the dataset by removing non-numeric columns and dropping rows with 0 or Nan values.
5. Then we plotted different graphs for visualization.

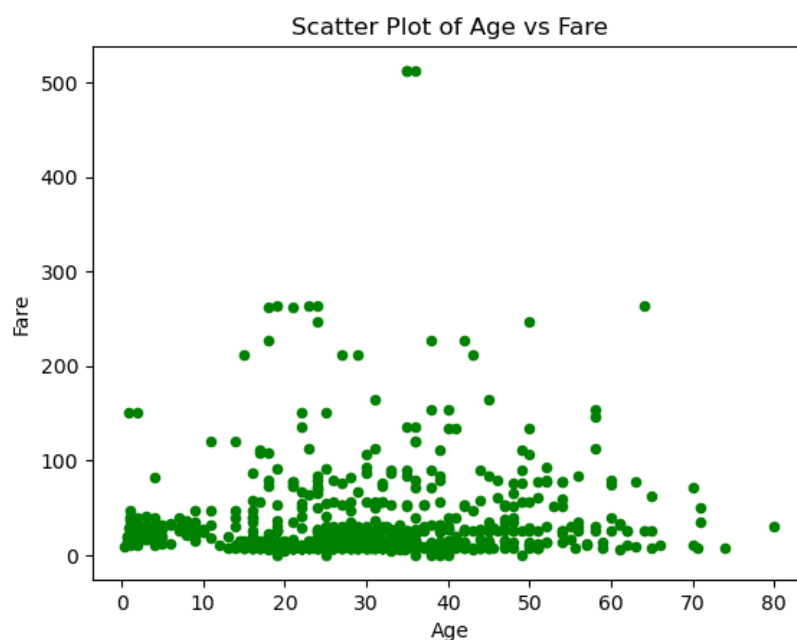
Graphs:



This is the first graph we plotted of age distribution which showing the age distribution of people present in titanic and their frequency.(y axis frequency showing how many in quantity)
e.g older people with age greater than 65 are less in quantity which show less people of age greater than travelling on titanic.



This graph showing the survivors 0 and 1 showing survived or not survived this graph showing less quantity of male survived and more female survived on titanic after the incident. Count on y-axis showing how many survived or died.



The graph is showing a scatter plot that represents the relationship between two variables. Each point on the plot represents an individual data point that shows the fare paid by a person of a certain age. From the graph, it looks like there is no clear, strong correlation between age and fare paid, since the points are widely dispersed and do not follow a specific pattern or trend.