# **DATA SCIENCE(UCS538)**

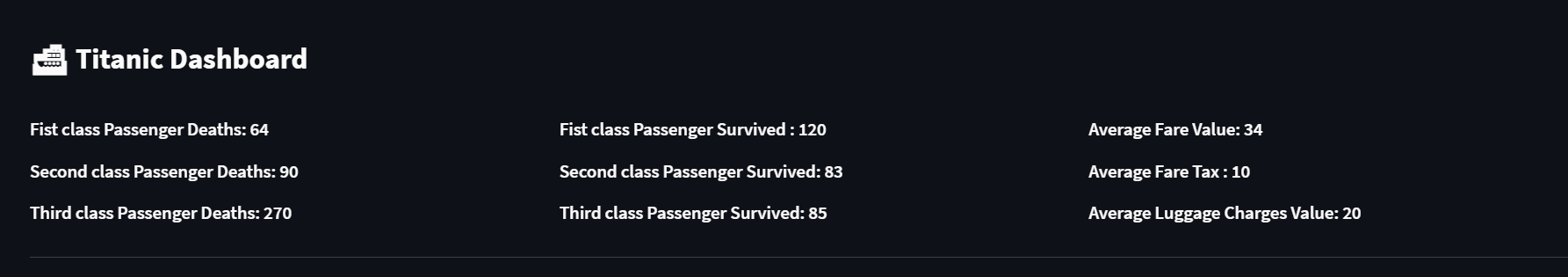
# **PROJECT REPORT**

# ****Submitted by- Paras Bakshi****

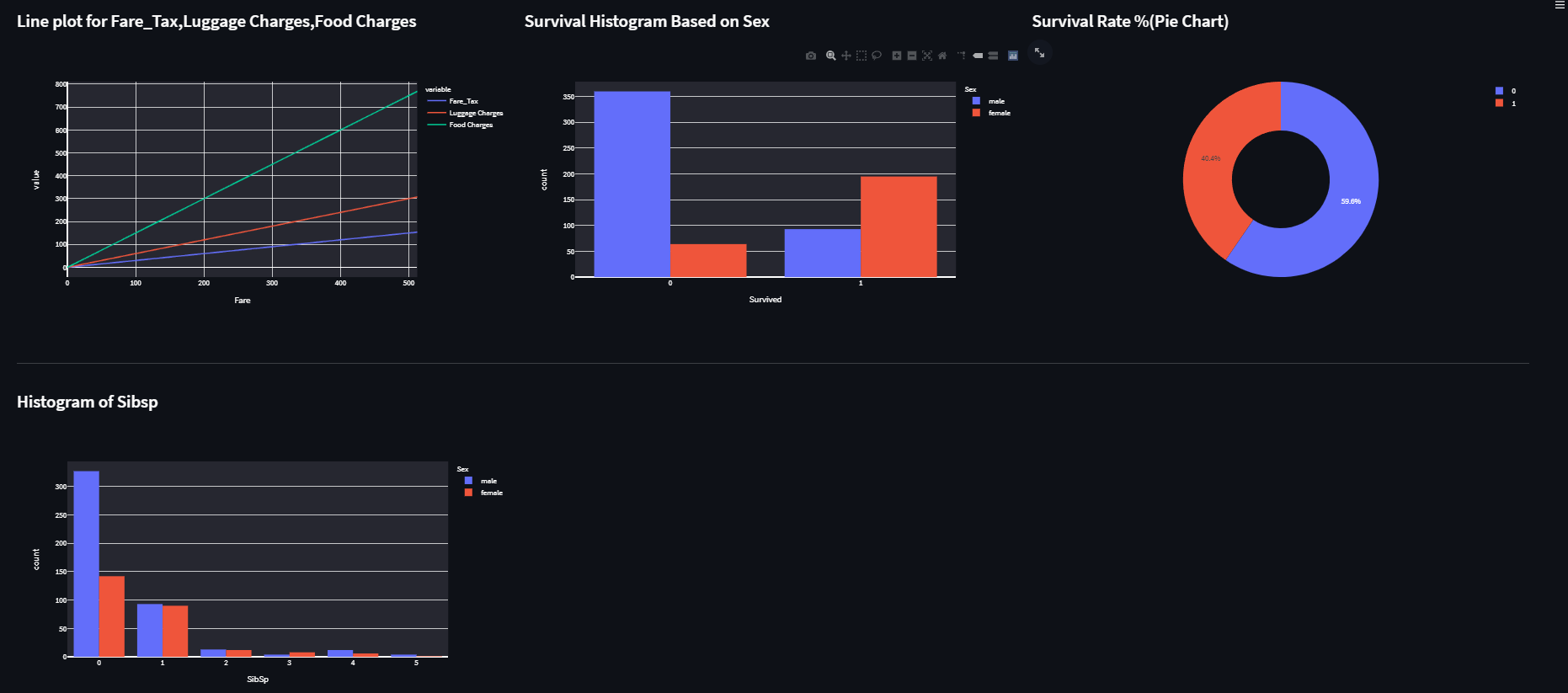
# ****Batch- 3CS5****

# ****Roll No- 101917118****

# 🚢 Titanic Dataset Dashboard Interface







# Dataset

Used titanic dataset from Kaggle

The csv file contains data for 887 of the real Titanic passengers. Each row represents one person. The columns describe different attributes about the person including whether they survived (SS), their age (AA), their passenger-class (CC), their sex (GG) and the fare they paid (XX)

Libraries used

1. Pandas

2. Numpy

3. Plotly(plotly.figure\_factory, plotly.offline, plotly.graph\_objs, plotly.express)

4.Streamlit

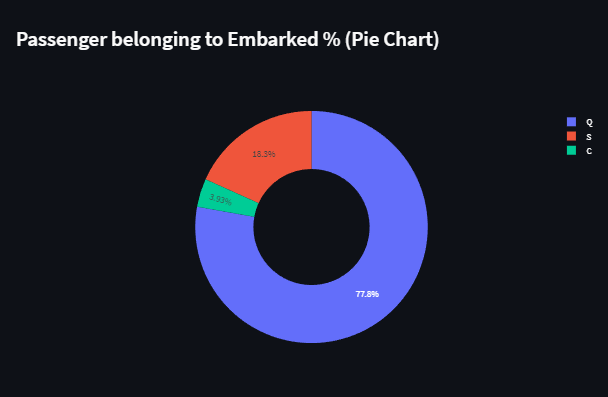
I have explained all the Charts and representation that I used in the dashboard in this report. The Screenshots of the report are as follows:

Calculated observations observations:-

Number of Fist class Passenger Survived : 120  
 Number of Second class Passenger Survived : 83  
 Number of class Passenger Survived : 85

Number of Fist class Passenger Deaths: 64  
 Number of Second class Passenger Deaths: 90  
 Number of Third class Passenger Deaths: 270

Number of Fist class Passenger Deaths: 34  
 Number of Second class Passenger Deaths: 10  
 Number of Third class Passenger Deaths:20

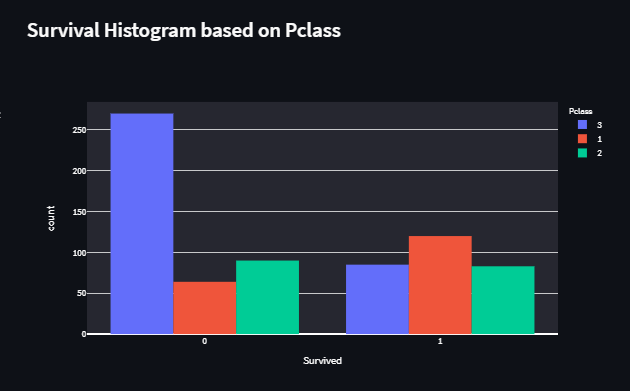


**Passenger belonging to Embarked % (Pie Chart)**

Shows that the highest % of the passengers belong to embarked class Q

That is 77.8 ,next 18.39 from the embarked class S and leaving 3.93 from class C.

### Passenger belonging to Embarked % (Pie Chart)



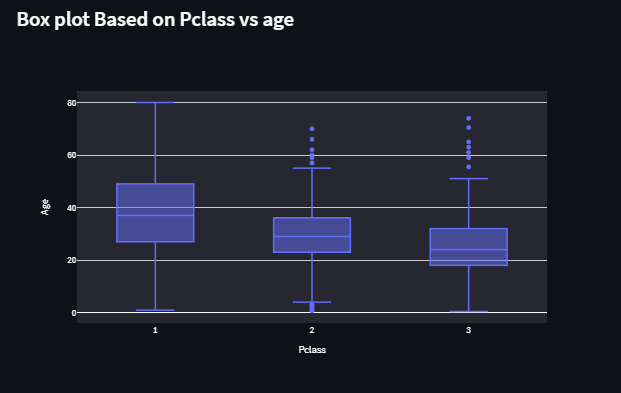
**Survival Histogram based on Pclass**

Number of Fist class Passenger Survived : 120  
 Number of Second class Passenger Survived : 83  
 Number of class Passenger Survived : 85

Number of Fist class Passenger Deaths: 64  
 Number of Second class Passenger Deaths: 90  
 Number of Third class Passenger Deaths: 270

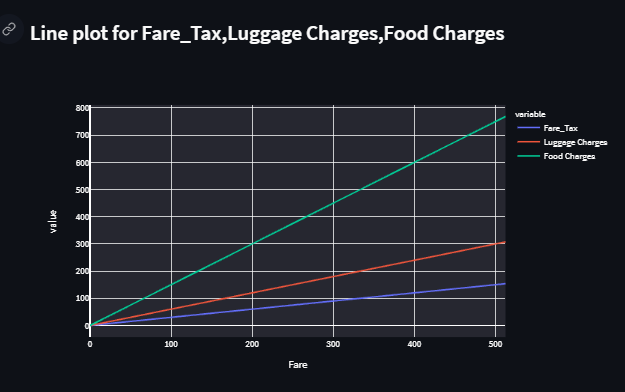
From this we can conclude that the people who travel in the first class

Have the maximum survival rate than class second and third , where as people who travel in the third class have the maximum death rate.



**Box plot Based on Pclass vs age**

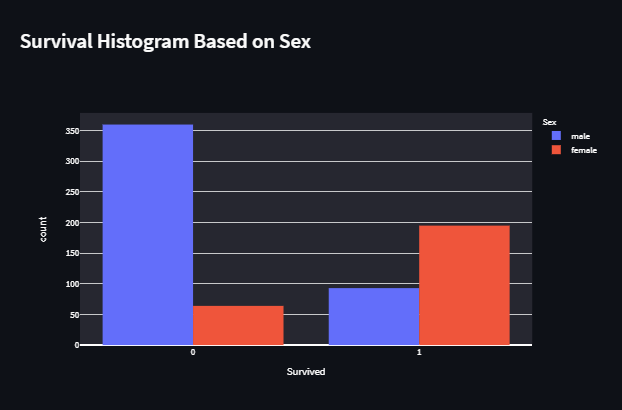
This box plot shows the presence of outliers in case of class 2 and 3 in term of ageWhere as there no presence of outlier in case of class 1.



**Line plot for Fare\_Tax,Luggage Charges,Food Charges**

Shows linear relationship between the Fare\_Tax,Luggage

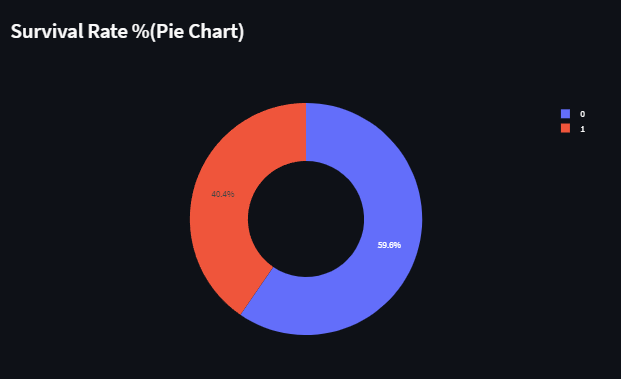
Charges,Food Charges.



**Survival Histogram Based on Sex**

This clearly shows that death rate of male is more than female.

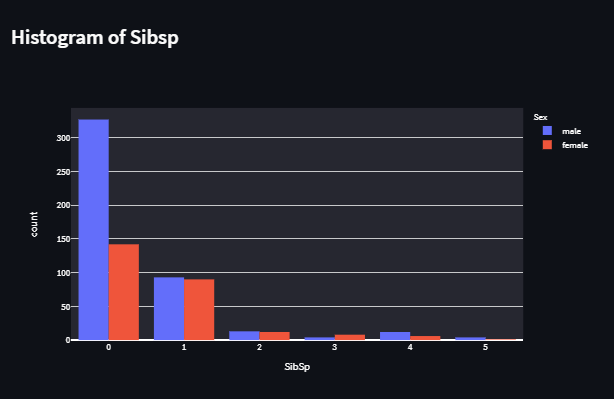
Survival rate of female is more than male.



**Survival Rate %(Pie Chart)**

survival rate=59.6

Death rate=40.49



**Histogram of Sibsp**

**Code**

*# importing libraries***from** numpy **import** double  
**import** pandas **as** pd  
**import** plotly.express **as** px  
**import** streamlit **as** st  
  
*# Assigning page title as Titanic Dashboard and favicon*st.set\_page\_config(page\_title=**"Titanic Dashboard"**, page\_icon=**"🚢"**,  
 layout=**"wide"**)  
  
  
*# function to read dataset from csv file*@st.cache  
**def** read\_data():  
 df = pd.read\_csv(**"train.csv"**)  
 **return** df  
  
  
*# taking input from function*df = read\_data()  
  
*# Assigning page title as Titanic Dashboard*st.title(**"🚢 Titanic Dashboard"**)  
st.markdown(**"##"**)  
  
*# making   
# Passenger belonging to Embarked % (Pie Chart*pie\_chart = px.pie(names=df[**"Embarked"**].unique(), values=df.groupby(**"Embarked"**)[**"Embarked"**].count(), hole=0.5)  
*# Survival Histogram based on Pclass*box\_plot = px.box(df, y=**"Age"**, x=**"Pclass"**)  
*# Box plot Based on Pclass vs age*group\_plot = px.histogram(df, x=**"Survived"**, color=**"Pclass"**, barmode=**'group'**)  
  
*# making 3 cols left\_column, middle\_column, right\_column*left\_column, middle\_column, right\_column = st.columns(3)  
  
*# showing data in left col  
# Number of Fist class Passenger Survived  
# Number of Second class Passenger Survived  
# Number of Second class Passenger Survived***with** left\_column:  
 st.subheader(**f" Fist class Passenger Deaths: {**len(df.loc[(df[**'Pclass'**] == 1) & (df[**'Survived'**] == 0)])**}"**)  
 st.subheader(**f"Second class Passenger Deaths: {**len(df.loc[(df[**'Pclass'**] == 2) & (df[**'Survived'**] == 0)])**}"**)  
 st.subheader(**f"Third class Passenger Deaths: {**len(df.loc[(df[**'Pclass'**] == 3) & (df[**'Survived'**] == 0)])**}"**)  
  
*# showing data in middle col  
# Number of Fist class Passenger Deaths  
# Number of Second class Passenger Deaths  
# Number of Third class Passenger Deaths***with** middle\_column:  
 st.subheader(**f"Fist class Passenger Survived : {**len(df.loc[(df[**'Pclass'**] == 1) & (df[**'Survived'**] == 1)])**}"**)  
 st.subheader(**f"Second class Passenger Survived: {**len(df.loc[(df[**'Pclass'**] == 2) & (df[**'Survived'**] == 1)])**}"**)  
 st.subheader(**f"Third class Passenger Survived: {**len(df.loc[(df[**'Pclass'**] == 3) & (df[**'Survived'**] == 1)])**}"**)  
  
*# showing data in right col  
# Average Fare Value  
# Average Fare Tax   
# Average Luggage Charges Value***with** right\_column:  
 st.subheader(**f"Average Fare Value: {**int(df[**'Fare'**].mean())**}"**)  
 st.subheader(**f"Average Fare Tax : {**int(df[**'Fare\_Tax'**].mean())**}"**)  
 st.subheader(**f"Average Luggage Charges Value: {**int(df[**'Luggage Charges'**].mean())**}"**)  
  
st.markdown(**'---'**)  
  
*# making 3 cols left\_column, middle\_column, right\_column*left\_column, middle\_column, right\_column = st.columns(3)  
  
*# assigning heading to each columns***with** left\_column:  
 st.subheader(**"Passenger belonging to Embarked % (Pie Chart)"**)  
**with** middle\_column:  
 st.subheader(**"Survival Histogram based on Pclass"**)  
**with** right\_column:  
 st.subheader(**"Box plot Based on Pclass vs age"**)  
  
st.markdown(**"---"**)  
*# ploting Passenger belonging to Embarked % (Pie Chart)*left\_column.plotly\_chart(pie\_chart, use\_container\_width=**True**)  
  
*# ploting Survival Histogram based on Pclass*right\_column.plotly\_chart(box\_plot, use\_container\_width=**True**)  
  
*# Box plot Based on Pclass vs age*middle\_column.plotly\_chart(group\_plot, use\_container\_width=**True**)  
  
*# MakingLine plot for Fare\_Tax,Luggage Charges,Food Charges*line\_plot = px.line(df, x=**"Fare"**, y=[**"Fare\_Tax"**, **"Luggage Charges"**, **"Food Charges"**])  
*# Making Survival Histogram Based on Sex*group\_plot1 = px.histogram(df, x=**"Survived"**, color=**"Sex"**, barmode=**'group'**)  
*# Making Survival Rate %(Pie Chart)*pie\_chart1 = px.pie(names=df[**"Survived"**].unique(), values=df.groupby(**"Survived"**)[**"Survived"**].count(), hole=0.5)  
  
*# making 3 cols left\_column, middle\_column, right\_column*left\_column, middle\_column, right\_column = st.columns(3)  
  
*# assigning heading to each columns***with** left\_column:  
 st.subheader(**"Line plot for Fare\_Tax,Luggage Charges,Food Charges"**)  
**with** middle\_column:  
 st.subheader(**"Survival Histogram Based on Sex"**)  
**with** right\_column:  
 st.subheader(**" Survival Rate %(Pie Chart)"**)  
  
st.markdown(**"---"**)  
*# Ploting Line plot for Fare\_Tax,Luggage Charges,Food Charges*left\_column.plotly\_chart(line\_plot, use\_container\_width=**True**)  
  
*# Making Survival Rate %(Pie Chart)*middle\_column.plotly\_chart(group\_plot1, use\_container\_width=**True**)  
  
*# Making Survival Rate %(Pie Chart)*right\_column.plotly\_chart(pie\_chart1, use\_container\_width=**True**)  
  
plot = px.histogram(df, x=**"SibSp"**, color=**"Sex"**, barmode=**'group'**)  
left\_column, middle\_column, right\_column = st.columns(3)  
*# ploting Histogram of Sibsp***with** left\_column:  
 st.subheader(**"Histogram of Sibsp"**)  
  
left\_column.plotly\_chart(plot, use\_container\_width=**True**)