

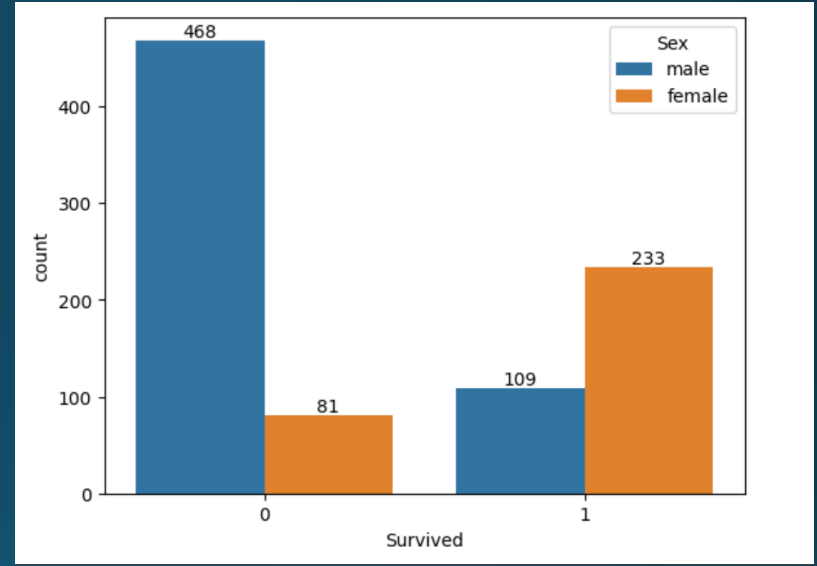
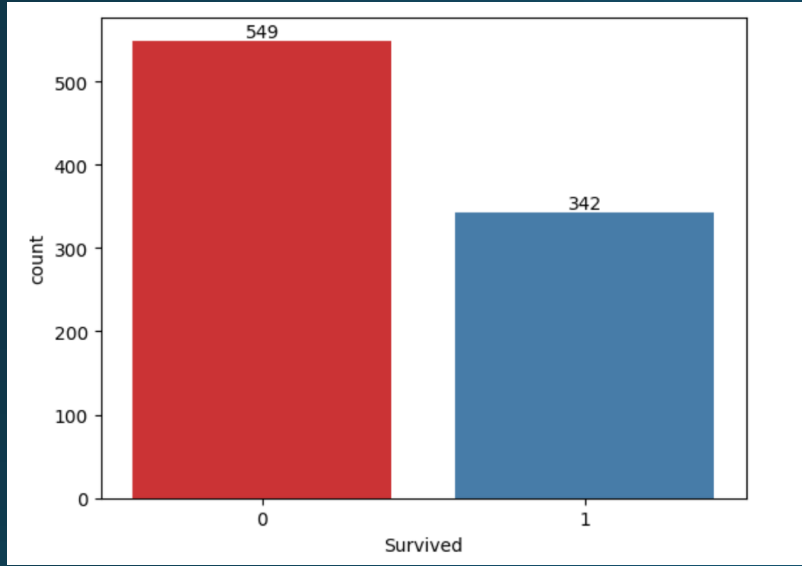
# Report on Titanic Disaster Survival Using Logistic Regression

PREPARED BY  
SHAKIL AHAMMED

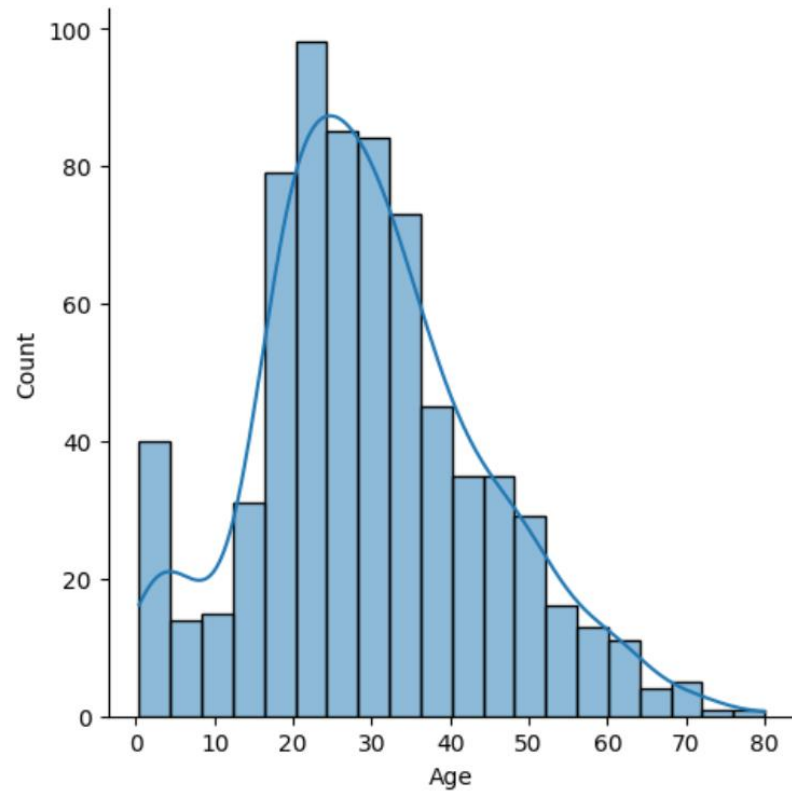
# Introduction

In this project, we will be working with the titanic dataset. Our goal is to train a logistic regression model that will be able to predict the chances of survival in disaster.

# Insights



Total survived 549 person and dead 342. Among them male survived 109 and dead 468, female survived 233 and dead 81. Most number of survival is female.



Age distribution of passenger of titanic.

# Building a Model

```
from sklearn.model_selection import train_test_split

x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.33, random_state=42)

from sklearn.linear_model import LogisticRegression

lr=LogisticRegression()

lr.fit(x_train,y_train)

▼ LogisticRegression
LogisticRegression()

predict=lr.predict(x_test)
predict
```

At first we divided the dataset into test part and train part using train test split method in scikitlearn library, then we train a logistic regression model.

```
lr.score(x_test, y_test)
```

```
0.7932203389830509
```

```
from sklearn.metrics import classification_report
```

```
print(classification_report(y_test, predict))
```

	precision	recall	f1-score	support
0	0.80	0.86	0.83	175
1	0.78	0.69	0.73	120
accuracy			0.79	295
macro avg	0.79	0.78	0.78	295
weighted avg	0.79	0.79	0.79	295

**Precision :** Precision is the ratio of correctly predicted positive observations to the total predicted positive observations. It is also called accuracy score. It's a good score(79%).

**Recall :** Recall is the ratio of correctly predicted positive observations to the all observations in actual class  
**F1 score - F1 Score** is the weighted average of Precision and Recall.