

Major Project PowerBI

Source link:

<https://www.kaggle.com/datasets/brendan45774/test-file>

Data Description:

PassengerId: The unique ID assigned for each passenger aboard the ship

Survived: Whether the passenger survived or not;

0=No, 1=Yes

Pclass: Ticket class;

1=First class, 2=Second class, 3=Third class

Name: Name of the passenger

Sex: Gender of the passenger

Age: Age at the time of boarding the vessel

SibSp: Number of siblings or spouses accompanying the passenger

Parch: Number of parents or children accompanying the passenger

Ticket: Ticket number/ID

Fare: Passenger fare for the journey

Cabin: Cabin number assigned to the passenger

Embarked: The port from where the passer embarked the ship;

C=Cherbourg, Q=Queenstown, S=Southampton

Requirements:

1. Create a custom column named **Survived_category** which contains a value “Yes”/”No” if the value in **Survived** is 1/0 respectfully.
2. Create a Stacked Bar chart to display the Survival Status by Sex. What do you observe?
3. Create Pie Charts to display the Survival Status by Ticket Class for all the classes. What do you observe?
4. Replace all the nulls in Age with a value of -1.
5. Create a custom column Age_category to group the following Age ranges into the following categories:
 - a. $0 \leq \text{Age} < 21$ – “Teen”
 - b. $21 \leq \text{Age} < 41$ – “Young Adult”
 - c. $41 \leq \text{Age} < 61$ – “Adult”
 - d. $\text{Age} \geq 61$ – “Old”
 - e. $\text{Age} = -1$ – “Null”
6. Create a Clustered Column Chart to display the Survival Status by Age using the categories in Age_category. What do you observe?