Titanic Passenger Data Analysis

This presentation delves into the analysis of passenger data from the Titanic tragedy. We'll explore survival trends, passenger demographics, and ticket class distribution using Python for data preprocessing and Power BI for insightful visualizations.

By Sonam S

Data Cleaning and Preprocessing

Dataset

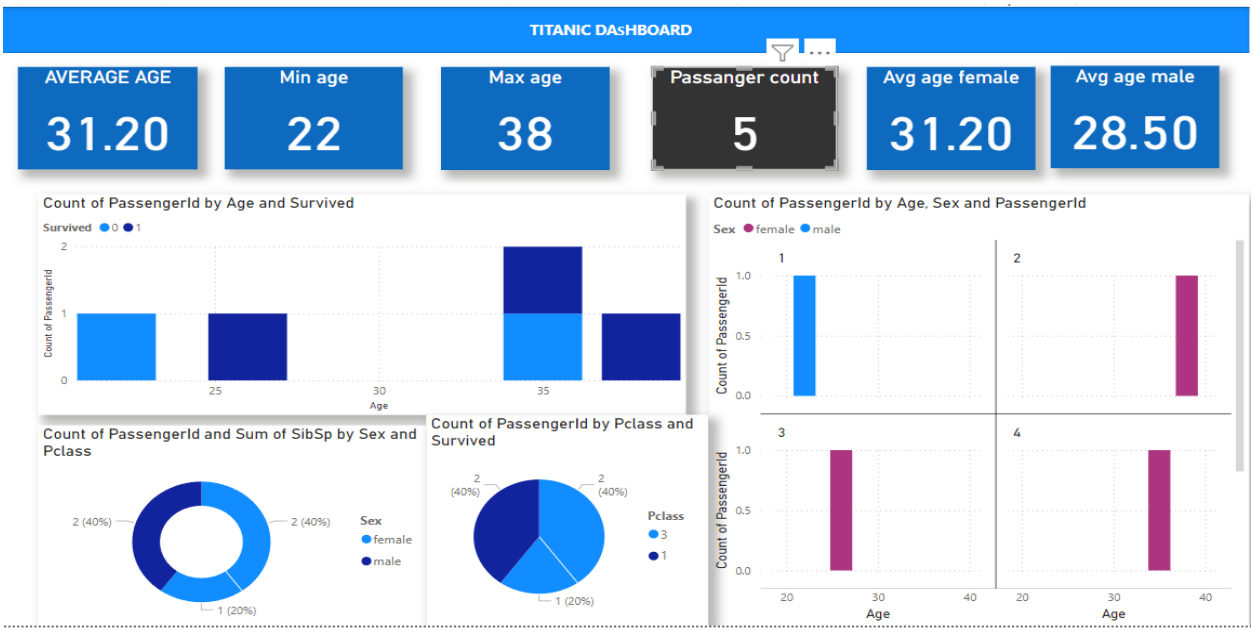
We'll use the Titanic dataset, which contains detailed information about passengers, including their age, gender, class, ticket fare, cabin, embarkation port, and survival status.

Preprocessing Steps

The data will be cleaned and preprocessed using Python. This includes handling missing values (e.g., age and cabin), formatting categorical variables, creating new metrics like family size, and validating data consistency to ensure accuracy in analysis.

Cleaned Dataset

PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
6	0	3	Moran, Mr. James	male		0	0	330877	8.4583		Q
18	1	2	Williams, Mr. Charles Eugene	male		0	0	244373	13		S
20	1	3	Masselmani, Mrs. Fatima	female		0	0	2649	7.225		C
27	0	3	Emir, Mr. Farred Chehab	male		0	0	2631	7.225		C
29	1	3	O'Dwyer, Miss. Ellen "Nellie"	female		0	0	330959	7.8792		Q
30	0	3	Todoroff, Mr. Lalio	male		0	0	349216	7.8958		S
33	1	3	Glynn, Miss. Mary Agatha	female		0	0	335677	7.75		Q
37	1	3	Mamee, Mr. Hanna	male		0	0	2677	7.2292		C
43	0	3	Kraeff, Mr. Theodor	male		0	0	349253	7.8958		C
46	0	3	Rogers, Mr. William John	male		0	0	S.C./A.4. 23567	8.05		S
47	0	3	Lennon, Mr. Denis	male		1	0	370371	15.5		Q
48	1	3	O'Driscoll, Miss. Bridget	female		0	0	14311	7.75		Q
49	0	3	Samaan, Mr. Youssef	male		2	0	2662	21.6792		C
65	0	1	Stewart, Mr. Albert A	male		0	0	PC 17605	27.7208		C
66	1	3	Moubarek, Master. Gerios	male		1	1	2661	15.2458		C
77	0	3	Staneff, Mr. Ivan	male		0	0	349208	7.8958		S
78	0	3	Moutal, Mr. Rahamin Haim	male		0	0	374746	8.05		S
83	1	3	McDermott, Miss. Brigdet Delia	female		0	0	330932	7.7875		Q
88	0	3	Slocovski, Mr. Selman Francis	male		0	0	SOTON/OQ 392086	8.05		S
96	0	3	Shorney, Mr. Charles Joseph	male		0	0	374910	8.05		S
102	0	3	Petroff, Mr. Pastcho ("Pentcho")	male		0	0	349215	7.8958		S
108	1	3	Moss, Mr. Albert Johan	male		0	0	312991	7.775		S
110	1	3	Moran, Miss. Bertha	female		1	0	371110	24.15		Q
122	0	3	Moore, Mr. Leonard Charles	male		0	0	A4. 54510	8.05		S
127	0	3	McMahon, Mr. Martin	male		0	0	370372	7.75		Q
141	0	3	Boulos, Mrs. Joseph (Sultana)	female		0	2	2678	15.2458		C
155	0	3	Olsen, Mr. Ole Martin	male		0	0	Fa 265302	7.3125		S



Visualizations

1. **Survival Trends Over Time**
    - Visualization: Line Chart
    - Purpose: Analyze trends in survival rates over different time periods (if applicable, based on events or data like rescue phases).
2. **Passenger Segmentation**
    - Visualization: Pie Chart
    - Purpose: Show proportions of passenger groups based on survival (e.g., Survivors vs. Non-Survivors) or other segments (e.g., passenger class or age groups).

3.Survival Analysis by Group

- Visualization: Matrix/Table

- Purpose: Highlight survival rates across different groups (e.g., passenger classes, genders, or embarkation points). Include total passengers, survivors, and survival rates.

4.Top Factors Influencing Survival

- Visualization: Bar Chart
- Purpose: Display key factors (e.g., passenger class, age group, gender) with their survival rates, highlighting the top predictors of survival.

3. **Survival Trends Over Time**
    - Visualization: Line Chart
    - Purpose: Analyze trends in survival rates over different time periods (if applicable, based on events or data like rescue phases).
4. **Passenger Segmentation**
    - Visualization: Pie Chart
    - Purpose: Show proportions of passenger groups based on survival (e.g., Survivors vs. Non-Survivors) or other segments (e.g., passenger class or age groups).

3.Survival Analysis by Group

- Visualization: Matrix/Table
- Purpose: Highlight survival rates across different groups (e.g., passenger classes, genders, or embarkation points). Include total passengers, survivors, and survival rates.

4.Top Factors Influencing Survival

- Visualization: Bar Chart
- Purpose: Display key factors (e.g., passenger class, age group, gender) with their survival rates, highlighting the top predictors of survival.