TITANIC DATABASE PROJECT PLAN AND OUTCOME

1. Data Cleaning and Preparation:

Handled missing values and irrelevant data. Some passengers missed age figures, so I had to exclude them from the summary as they might be both less than 1 year old or any number above.

Performed data type conversions and normalisation for easier analysis.

Created a new table focusing on the most relevant columns for analysis (e.g., survival, age, class, gender) using VLookup function.

2. Data Transformation

Created new features and derived columns to improve analysis (e.g., grouping passengers by age ranges, calculating family size).

Filtered data for specific analysis purposes, such as focusing on survival rates by class and sex.

3. Interactive Dashboard Creation

Leveraged Pivot Tables and Pivot Charts to summarise and visualise key metrics such as survival rates, age distributions, and class demographics.

Integrated slicers for dynamic filtering of data by various categories like gender, number of relatives on board, and age group.

Created a clean and interactive dashboard for easy exploration of insights.

4. Key insights into the data:

As anticipated, the survival rate was significantly higher for first-class passengers compared to those in second and third class (63% versus 47% and 24%, respectively). Additionally, women accounted for 60% to 80% of survivors across all classes.

Survival rates declined as the number of relatives on board increased. Passengers with 1-3 relatives had a survival rate of 50% or higher, while those with 4 or more relatives, as well as those traveling alone, had notably lower survival rates, ranging from 0% to 33%.

The surnames Baclini and Carter were most frequently found among the survivors.