**SUMMARY OF TITANIC DATASET**

This assignment utilized the Titanic dataset from Kaggle, which comprises 12 columns and 891 rows.

I tested three hypotheses using chi-square tests based on this dataset.

|  |  |  |  |
| --- | --- | --- | --- |
|  | HYPOTHESES #1 | HYPOTHESES #2 | HYPOTHESES #3 |
| Hypotheses  H0:  H1: | **H0:** The class of passenger does not affect the survival rate.  **H1:** The class of passenger affects the survival rate. | **H0:** The gender of passenger does not affect the survival rate.  **H1:** The gender of passenger affects the survival rate. | If age is younger that 15 “Children”, Otherwise called “Adult”  **H0:** The age class does not affect the survival rate.  **H1:** The age class affects the survival rate. |
| Categorical Tables: |  |  |  |
| Plots: |  |  |  |
| It appears that there is no statistically significant association between survival chance and gender, pcalss, and age. | | |
| Chi Square Test | Hyp1\_Chi2: 102.89  Hyp1\_P-Value: 0.00000  Hyp1\_Def: 2 | Hyp2\_Chi2: 260.72  Hyp2\_P-Value: 0.00000  Hyp2\_Def: 1 | Hyp3\_Chi2: 15.56  Hyp3\_P-Value: 0.00008  Hyp3\_Def: 1 |
| Result | Significant level (5%) > P-Value (0.00000)  Reject the null hypothesis  The passenger class does affect the survival rate | Significant level (5%) > P-Value (0.00000)  Reject the null hypothesis  The passenger gender does affect the survival rate | Significant level (5%) > P-Value (0.00008)  Reject the null hypothesis  The passenger age does affect the survival rate. |