COMP 333 — Lab Assignment 4

Motivation

The purpose of this assignment is for you to perform data analytics, develop a story, and create a four-slide presentation of your story.

It is similar to Lab 12 for story telling and visualization, however, it uses the Titanic dataset and considers only the child passengers.

Assignment

Analyse the Titanic dataset to determine which factors affect the survival rate of children. Consider factors such as age, sex, embarkation port, and size of the family that was travelling; if such information is available. Focus on creating a story.

Create a Jupyter notebook using Python and its libraries, to analyse the Titanic dataset using EDA to develop a story about the survival of children aboard the Titanic. Markup the notebook so it contains the content of your four-slide story embedded in the notebook.

Your notebook should clearly show your analysis, the presentation content, and the visualizations.

It is to focus on a simple EDA of the Titanic data, using Jupyter markup for improved commentary in your notebook, and to focus on creating visualizations that are easily readable and good communicators of your story.

Deliverable

Your deliverable is a zipped directory containing the completed ipynb notebook showing all computation, output, and plots; and the pdf of your four-slide presentation.

The four "slides" of your story should stand out as obvious parts of your notebook.

There must be exactly four slides: Situation, Problem, Solution, Next Steps as in the SPSN approach.