PROJECT

Investigate a Dataset

A part of the Data Analyst Nanodegree Program

PROJECT REVIEW

CODE REVIEW

NOTES

SHARE YOUR ACCOMPLISHMENT!

Meets Specifications

Congratulations for meeting all specifications for this project! I added some comments and suggestions below, I hope they'll be of use. Keep up the good work!

Code Functionality

✓

All code is functional and produces no errors when run. The code given is sufficient to reproduce the results described.

Awesome

It's great to see your code cells were run in order from start to finish, as this ensures the code will run correctly without the use of "future" variables or functions.

✓

The project uses NumPy arrays and Pandas Series and DataFrames where appropriate rather than Python lists and dictionaries. Where possible, vectorized operations and built-in functions are used instead of loops.

Awesome

I don't think there was anything against the rubrics in the calculations flagged by the previous reviewer (in my understanding, only loops, lists and dictionaries are *verboten*), but it's nice to see you used more idiomatic Pandas expressions in this version of your report. :)

✓

The code makes use of functions to avoid repetitive code. The code contains good comments and variable names, making it easy to read.

Suggestion

When it comes to functions, it's better (and more pythonic) to use docstrings instead of comments to describe them and explain what they do.

Quality of Analysis

✓

The project clearly states one or more questions, then addresses those questions in the rest of the analysis.

Suggestion

When it comes to functions, it's better (and more pythonic) to use docstrings instead of comments to describe them and explain what they do.

Data Wrangling Phase

✓

The project documents any changes that were made to clean the data, such as merging multiple files, handling missing values, etc.

Suggestion

If you are interested, this website presents some alternatives on dealing with missing data that can be useful.

Exploration Phase

✓

The project investigates the stated question(s) from multiple angles. At least three variables are investigated using both single-variable (1d) and multiple-variable (2d) explorations.

Suggestion

Another interesting variable combination to analyze when studying survival rates is *gender* and *class*. For instance, would a male passenger in the 1st class have a higher or lower survival chance then a female passenger in the 3rd class?

Comment

`Fare` is a particularly tough variable to deal with in this data set. Does it correspond to the individual value paid by each passenger, or to the value paid for the ticket as a whole - which might include people traveling together (such as family members)? Can you think of a way to figure this out?

✓

The project's visualizations are varied and show multiple comparisons and trends. Relevant statistics are computed throughout the analysis when an inference is made about the data.

At least two kinds of plots should be created as part of the explorations.

Awesome

Great job dealing with the issue pointed out by the previous reviewer and properly interpreting the plots based on whether they present *counts* or *rates* of passengers.

Conclusions Phase

✓

The results of the analysis are presented such that any limitations are clear. The analysis does not state or imply that one change causes another based solely on a correlation.

Awesome

Good job discussing some of the limitations of your analysis in the concluding section of your report, and also suggesting some future work that can be done to expand your analysis here.

Communication

✓

Reasoning is provided for each analysis decision, plot, and statistical summary.

✓

Visualizations made in the project depict the data in an appropriate manner that allows plots to be readily interpreted.

Comment

> We can see that we have more male passengers on board than females (almost double).

But the plot from cell 18 seems to show the opposite - more female than male passengers! I don't think this merits another resubmission, but are you sure you got the legend right?

DOWNLOAD PROJECT