UDACITY

Investigate a Dataset

A part of the Data Analyst Nanodegree Program

PROJECT REVIEW

NOTES

SHARE YOUR ACCOMPLISHMENT!

Meets Specifications

You put in a lot of effort and it clearly shows in the project report.
Congratulations on a very good project report.
Feel free to share it with your friends or post it as a blog.
I hope you work on more exciting and challenging Data Investigation projects in future. All the best for your future submissions. Good Luck!!

Code Functionality

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

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.

The project correctly uses Pandas series and dataframes throughout the project. Good use of inbuilt functions like groupby.

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

The notebook has well-commented code, the variable names are apt and related to the project. Good use of functions like read_csv, get_average to avoid repetitive code

Quality of Analysis

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

The project has clearly raised more than one good specific questions and addressed them later on in the analysis.

Data Wrangling Phase

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

Good work in implementing the Data Wrangling Phase.

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.

The project correctly investigates the questions mentioned from multiple angles.
These visualizations are relevant to the questions asked in the project.

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.

There is no limit to the visualizations that can be done for a project, but I can see that you have more than two plots as part of EDA. Good job on the visualizations.

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.

Conclusions, as well as limitations of the datasets, is well documented.

Communication

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

Correct reasoning has been given for the visualizations and statistical summary made in the notebook.

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

The plots provide a title, a clear mention of labels on X & Y axis. 👍

⬇ DOWNLOAD PROJECT

RETURN TO PATH

Student FAQ