UDACITY

< Return to "Data Analyst Nanodegree" in the classroom

DISCUSS ON STUDENT HUB

Investigate a Dataset

REVIEW
HISTORY

Meets Specifications

Greetings Student,
This was a good implementation and I congratulate you for passing all rubric items with this submission. It was delightful reviewing your work as it was well-thought-out. I encourage you to keep up the good work as it will make you a great Data Analyst. Way to go!

For any queries, you can ask on Knowledge Portal as well [https://knowledge.udacity.com/]
Stay Safe

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.

Excellent work using NumPy and Pandas library to facilitate the work for this submission.

Some of the few Pandas built-in methods that are very useful for exploring variables in this project:
• Boolean-Indexing: http://pandas.pydata.org/pandas-docs/stable/indexing.html#boolean-indexing
• Group-by: http://pandas.pydata.org/pandas-docs/stable/groupby.html

• Value-Counts: https://chrisalbon.com/python/data_wrangling/pandas_dataframe_count_values/
• Series.map: https://pandas.pydata.org/pandas-docs/stable/generated/pandas.Series.map.html
• Working-with-text-data: https://pandas.pydata.org/pandas-docs/stable/text.html

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

Good Job adding function for sql query.

Quality of Analysis

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

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'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.

The report makes use of different chart type to explore and depict the insights and the results of the analysis

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.

Communication

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

As a future data analyst, It is very important to communicate the results adequately; however, it is also very important to describe each activity, analysis, or graph. This will allow your audience to understand what you are doing and how you are doing it. Moreover, the reasoning makes your work organized, formal, and sophisticated.

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

Awesome! The plots are well labeled and easy to interpret

⤓ DOWNLOAD PROJECT

RETURN TO PATH

Rate this review