U D A C I T Y

PROJECT

Investigate a Dataset

A part of the Data Analyst Nanodegree Program

PROJECT REVIEW

NOTES

SHARE YOUR ACCOMPLISHMENT! 🐦 f

Meets Specifications

In general, this is a very strong submission! Congrats on passing on your first try. I've stated a list of items that I like regarding your submission.

- You have a good sentence structure and overall, the entire submissions is easy to follow along.
- You've correctly used adjusted revenue for your analysis - this is a surprisingly common mistake amongst students.

Do read up more on other functionalities in Pandas, such as the split-apply-combine recipes (https://pandas.pydata.org/pandas-docs/stable/groupby.html) to better enable you to aggregate data along various columns.

It was a joy reading your submission, and all the best for your next project!

Code Functionality

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

All code appears to be functional and error free.

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.

Good use of vectorized operations in Pandas (checking for missing values, creating new columns based on other columns).

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

Variable names are descriptive, code is overall easy to parse and understand.

Quality of Analysis

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

4 questions were listed, and there was enough depth for each one to be thoroughly addressed.

Data Wrangling Phase

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

Very interesting questions listed!

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.

Great use of the histograms for the 1d analysis, and the line graphs for your time-series analysis.

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.

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.

No causality is implied, and you have been cautious about stating that the analysis is mostly correlations. Good job 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.

⤓ DOWNLOAD PROJECT

RETURN TO PATH

Student FAQ