


[Back to Data Analyst Nanodegree](#)

Investigate a Dataset

REVIEW

HISTORY

Meets Specifications

Excellent work overall! To further practice data analysis, I would encourage you to try Kaggle competitions, Kaggle is a good source to start from. I have provided some suggestions which are provided as tips for future analysis. Stay Udacious 

Code Functionality



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

The code is running perfectly well and there are no errors in the notebook.



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 has been made.

Suggestion – Please refer this [link](#) and this [link](#) for more on vectorized operations if want to learn more on this topic

Awesome review!



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.
Suggestion- Please refer this [link](#) to know more about how to optimize your 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.
Suggestion - A nice [article](#) to for additional reading - **The hardest thing about Data Science is asking the right question**

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 - I would encourage you to read this [article](#) which suggests some ways to handle the null values in the data for future projects. Since handling null values is a very important part of data analysis

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.
Suggestion – Please refer this [link](#) to know more about how to explore data using visualizations

- ✓ 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.
Suggestion – Please refer this [link](#) to know more about varied visualizations and how they can be used

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.

Good job on discussing the limitation of current analysis and indicating that exploring the positive correlation might give more information for the analysis.
Suggestion – Please refer this [link](#) to learn more about how to draw conclusions from your analysis

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.
Suggestion – Please refer this [article](#) to know how to effectively present and communicate your work in data science

- ✓ 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. 👍
Suggestion – Please refer this [link](#) to know more about how to create effective visualizations

 [DOWNLOAD PROJECT](#)

[RETURN TO PATH](#)