U U D A C I T Y

Explore and Summarize Data

A part of the Data Analyst Nanodegree Program

PROJECT REVIEW
CODE REVIEW
NOTES

SHARE YOUR ACCOMPLISHMENT! 🐦 f

Meets Specifications

Code Functionality

All code is functional (e.g. No Error is produced and RMD document is not prevented from being knit.)
Your code is fully functional.

The project almost never uses repetitive code where a function would be more appropriate. The code references variables by name instead of using constants or column numbers.
Customized functions are used to avoid repetitive codes. Good job!

Project Readability

All complex code is adequately explained with comments. It is always clear what the code is doing and how and why any unusual coding decisions were made.
I strongly suggest you to add some comments for some changes subject to the dataset, which will improve the readability of you code.
As you have passed this rubric from previous reviewer, I didn't fail you here.

The code uses formatting techniques in a consistent and effective manner to improve code readability. All lines are shorter than 80 characters.
All lines are under 80 characters.

Markdown syntax is used in the RMD file to improve readability of the knitted file.

Quality of Analysis

The project appropriately uses univariate, bivariate, and multivariate plots to explore most of the expected relationships in the data set.
The report include different chart type that explore many aspects about the data set.

Questions and findings are placed between blocks of R code regularly so it is clear what the student was thinking throughout the analysis.

In the univariate section, you may consider to state the statistics, and expand the discussion about the outliers (if there is any) for each feature in the data set. At this stage you can also remove some of the outliers if you find it justified, that will make the following analysis more robust.

Reasoning is provided for the plots made throughout the analysis. Plots made follow a logical flow. Comments following plots accurately reflect the plots' contents.

The project contains at least 20 visualizations. The visualizations are varied and show multiple comparisons and trends. Relevant statistics (e.g. mean, median, confidence intervals, correlations) are computed throughout the analysis when an inference is made about the data.

The analysis include many figures that depict comparison, trends and relation between features.

Suggestion

I want to encourage you to include the relevant statistics in the discussion under each chart. For example, you can quantify the distribution of each feature with the mean median and quartiles. For the bivariate section , the correlation values, again, please consider to include the relevant correlation values in the discussion under each chart (you did so to many figures).

Visualizations made in the project depict the data in an appropriate manner that allows plots to be readily interpreted. Choice of plot type, variables, and aesthetic parameters (e.g. bin width, color, axis breaks) is appropriate.

Final Plots and Summary

The project includes a Final Plots and Summary section containing three plots and commentary. All plots in this section reflect what has been explored in the main body of the analysis.

The final plot section include 3 figures that represents the analysis done in the exploratory section.

The plots are well chosen and the plots fulfill at least 2 of the criteria. The plots are varied and reveal interesting trends and relationships.

The three plots are well chosen, which are varied and can reveal interesting trends and relationships.

All plots have appropriately selected variables and are plotted in a way that accurately conveys the data/information (i.e findings in Final Plot 1 do not depend on the findings of Final Plot 2).

Three plots are properly selected. The findings of each plot don't depend on those of the rest plots.

All plots are labeled appropriately (axis labels, plot titles, axis units) and can be read and interpreted easily. Plots are scaled appropriately.

All plots are labeled appropriately with axis labels, plot titles, axis units.

The reasoning and findings from each plot are explained and the text about each plot is descriptive enough to stand alone. Comments reflect the contents of the plots that they are associated with.

Reflection

The project includes a Reflection section discussing the analysis performed.

A reflection section is included. Great job!

The section reflects on how the analysis was conducted and reports on the struggles and successes throughout the analysis. The section provides at least one idea or question for future work. The section explains any important decisions in the analysis and how those decisions affected the analysis.

Great job!
The reflection section shows your deep understanding of the dataset and the careful analysis performed.

Suggestion

For future work, it will definitely help if we collect more more variables. What feature would you like to add in your dataset? Will adding more samples helps to get a better conclusion as well?

⬇ DOWNLOAD PROJECT

RETURN TO PATH

Student FAQ

For future work, it will definitely help if we collect more more variables. What feature would you like to add in your dataset? Will adding more samples helps to get a better conclusion as well?

⬇ DOWNLOAD PROJECT