UDACITY

PROJECT

Explore and Summarize Data

A part of the Data Analyst Nanodegree Program

PROJECT REVIEW

CODE REVIEW

NOTES

SHARE YOUR ACCOMPLISHMENT!

Meets Specifications

Congratulations! You've done an excellent work. I find your project is well organised and easy to follow. It depicts many interesting findings as well. I also leave a few suggestions which I think is helpful for your future project, which you can refer below. Good luck on your next project.

Code Functionality

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

All code is 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.

Functions are used to avoid repetitive code. Excellent!

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.

Complex code is explained with comments.

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

The code is formatted well. Most lines are shorter than 80 characters.

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

Markdown syntax is correctly used in the RMD.

Quality of Analysis

The project appropriately uses univariate, bivariate, and multivariate plots to explore most of the expected relationships in the data set.

Different types of plots are included in the project, which depict many aspects of the dataset.

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

Discussions are placed between different plots in your project, which makes it easier to follow your explanation.

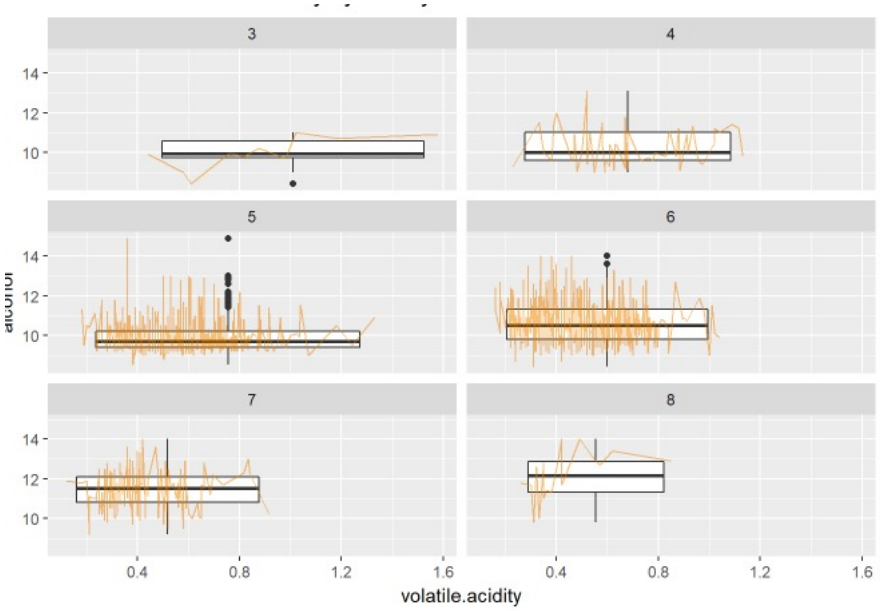
It's great to see you arrange box plots and histograms together for easily identifying potential outliers.

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

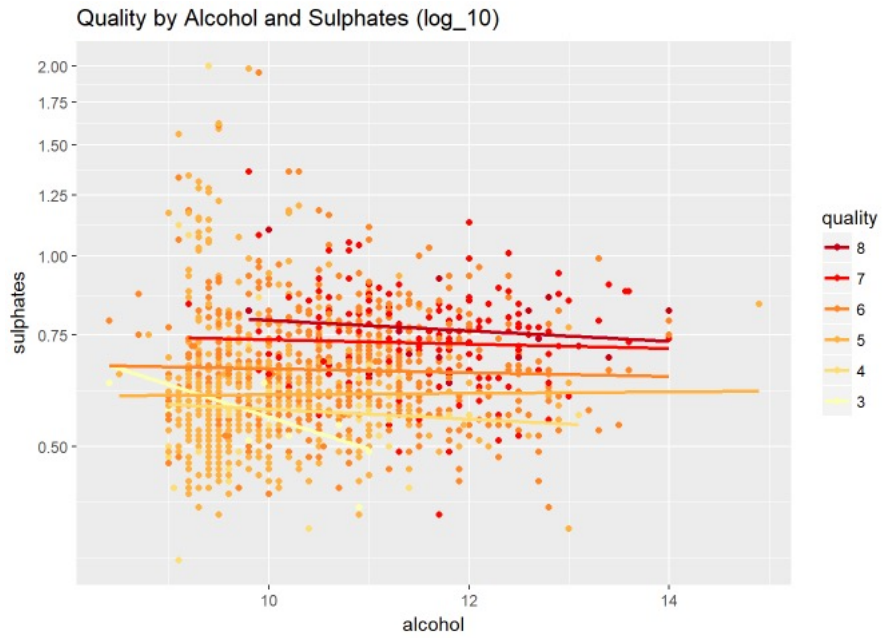
The project contains at least 20 visualizations. The visualizations are varied and show multiple comparisons and trends. Relevant statistics (e.g. mean, median, confidence intervals, correlations) are computed throughout the analysis when an inference is made about the data.

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.

Visualisations are well done in this submission, so I only have few comments.



Box plots are more effective when exploring relationship between categorical (e.g.: quality) and numerical variables. If we take a look at a subplot, you will find both x and y variables are numerical, which brings difficulty to interpret the distribution horizontally. What does that mean of width of the box? In this case, scatterplots are more appropriate.



Good job adding regression lines and choosing a sequential colour palette.

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.

Three plots are included in Final Plots and Summary section that depict interesting finding from previous analysis.

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.

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).

All final plots are plotted in an appropriate way.

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

Final plots are labeled well.

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.

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.

DOWNLOAD PROJECT

RETURN TO PATH

Rate this review