U D A C I T Y

Log

PROJECT

Explore and Summarize Data

A part of the Data Analyst Nanodegree Program

PROJECT REVIEW	CODE REVIEW	NOTES

Meets Specifications

SHARE YOUR ACCOMPLISHMENT

Code Functionality

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

✓	**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.**
	Well Done for demonstrating the use of functions that reduce repetitions and simplify the code.

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.**

✓	**The code uses formatting techniques in a consistent and effective manner to improve code readability. All lines are shorter than 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 analysis makes use of different chart types, including univariate, bivariate and multivariate exploration to explores and investigate many aspects of 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.**
	The discussion between code block includes relevant questions and interesting findings.

✓	**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 analysis follows a logical flow where the results of one analysis lead to another.

It is excellent that you are using the boxplots to depict the outliers for the continues features.

Also great that you use the polynom map to set the borders, I think the information is much clearer.

✓ **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.**
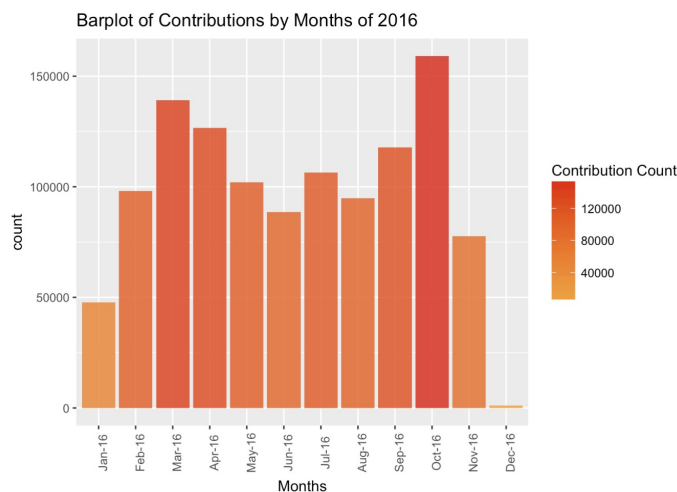
The analysis includes many figures that depict comparison, trends and relations between features.

✓ **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.**
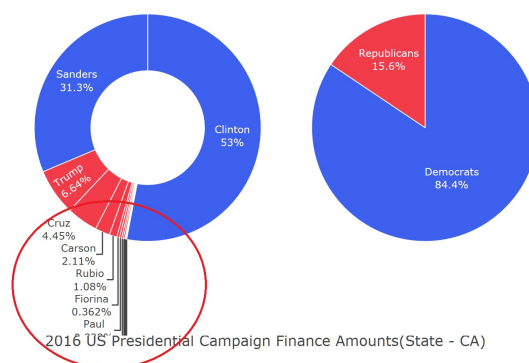
Most of the charts are well done, so I only have few comments here,

For the maps. please consider using log scale color map, that will increase the contrast, especially for the small values.

Please consider avoiding double coding on a single figure. For example, the y-axis and the color in the figure here below represent the same feature,

A bar plot will be more appropriate than a pie chart to depict count distribution,

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 figure in the final plot section represents the analysis done in the exploration 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.

✓   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).

For final plot 1, I strongly encourage you using a log scale so the small values will be visible.

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

✓   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.

Please consider including more relevant statistics to quantify and support the findings and insights of the final plot section.

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

Student FAQ