PROJECT

Explore and Summarize Data

A part of the Data Analyst Nanodegree Program

PROJECT REVIEW | CODE REVIEW | NOTES

Meets Specifications

Congratulations on passing the project on the first attempt! Great work , The report is excellent and looks perfect. Best of luck for the rest of the nanodegree.

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.

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.

There are handful of lines which are above 80 character's.per line. Use the R Studio warning line to limit it , This would enable better code readability.

https://support.rstudio.com/hc/en-us/community/posts/207625357-Toggle-80-character-warning-line

Here's the link which can help you.

✓ 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.

Superb work with the analysis. Well done!

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

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

Superb work. Report is well done!

✓ 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.

Visualization's are perfect. Well done on this part. You can even plot boxplot's for specific prosper rating in comparison to loan amount on the basis of Delinquent list.

But i have to be honest , Visualization's are polished and are done well.

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

When plotting for Loan Amount , use axis label as `Loan Amount ($)` it is important to specify the units but since it's the only part and the entire report looks good and polished I am going to leave this as a suggestion.

✓ 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.