

Infinity

Description:

You will be given the opportunity to resubmit your assignment for full points if you address my comments. **Please email add to this document with your responses colored in red** as to how you addressed each of my comments. Here is a generic example of how this might look (this comment is not specific to any submission):

The interpretation of your p-value in the “xx” section is not fully correct. The p-value is NOT the probability that the null hypothesis is true.

The p-value interpretation has been revised to indicate that it represents the probability of observing something as far (or farther) away from the assumed average.

Be sure to include this response to review inside of the R project that you submit.

Comments:

(Do not reply to this one comment:) Excellent work on this project. All the key elements are there, unlike many initial submissions. Please see the comments below that should be easy for you to fix.

From the assignment description: **“You will submit a pdf copy of your package vignette THAT INCLUDES the URL to your GitHub repository.”** While you did provide an assignment comment that let me grade this, please also add this URL to your package vignette.

We agree. Now we have added the GitHub repository link in the description.

You have some pieces of this vignette that takes forever to run. This makes it difficult to check your package or re-compile the vignette. My guess it is some of your ML models. The vignette should run quickly so I suggest that you adjust things in one of the following two ways:

1. Make the sample problem smaller, using only a subset of the total values.
2. Pre-compute some of the results and save them as an R data object that you can then call to show the results. Then, take the computationally heavy portions of your vignette and make them “eval=FALSE” so that the code is still shown (and explained) but not computed each and every time.

We agree. Now we have pre-run all the models and data and saved inside vignettes as “.rda” files. Now we use “eval = FALSE” inside these specific r-chunks and later we recall the saved outputs using separate R-chunks (These codes are not showing when the document is generated).

Please eliminate the pie chart from your explore_visuals function. There is no need to replace it with anything else. Pie charts are a poor representation of proportions and I don’t think a pie chart is necessary in this case since you only have two categories.

We agree. Now we have removed the pie-chart.

Please adjust your spam/ham bar chart visuals to show all the words, you may need to plot less words to make this work. I also recommend switching the x and y axis so that the words appear on the y-axis instead of the x-axis. Lastly, I recommend using a dot chart instead of a bar chart as it represents the same thing with less “ink”.

We agree. Now we use dot charts instead of bar charts.