EDA review

Rienk Heins

9/26/2020

Review on EDA of Linda de Vries

Introduction

Pros and cons of the introduction, containing the introduction, missing data and variation sections.

Pros:
- Clearly stated research question.
- Stating why elimination diet and nutrition data was used.
- Stating there is missing data.
- Stating which variables are looked at as important for the research.
- Dataset is shown.

Cons:
- Grammar, some sentences are not correct.
- No records of how the data was acquired, which study? Who worked on it?
- No real explanation of what Eosinophilic Esophagitis is, it is clear that it has to do with allergic reactions because of the rest of the introduction but it isn't clear what EoE it on itself.
- No real explanation of the variables that will be used in the EDA, why are they important to look at? Is there correlation? Is there other research that shows they may be important?

Plots

Pros and cons of the results section containing the plots.

Pros:
- All plots have names on axis, a title and when needed a legend.
- Brief explanation of plots
- Use of different plots

Cons:
- No explanation why the plot was made, why is the data relevant?
- Only 4 plots while there seem to be a lot of variables in the data.
- Units are missing.
- Explanation of what is seen in the plot is very short, when stated that there is effect no explanation is given why there is effect, or what type of effect.

Final advise

The EDA contains most of the things it needs to contain but there is a lot of documentation missing. Things like explanation of why certain decisions are made, what the data contains and the information that the

plots give. Also since there is data missing it is really important to give information about where the data is acquired and who worked on it as when someone would want to repeat the research they would need the data and might want to contact the makers of the original study so they can acquire the missing data when available. Further there are some mistakes in grammar where the sentences don't really make sense in the way they are constructed. A tip might be to let someone else read it as spelling mistakes are easily made and will be detected faster by another person critically looking at the file.