Assignment 0: IDS latex template

Name1 (s1234567)
Name2 (s7654321)

Group 1234

December 28, 2021

Abstract

This is the template for writing reports for the weekly assignments. The page 5 pages. You are expected not to change the font size or style or margin width. As in certain conference submissions, if the template of your submission deviates significantly from this, your submission might incur negative points. So please follow this template and read the instructions.

Introduction

This template was made for the course introduction to data science [2]. In this section we explain some of the conventions which should be followed for answering theory questions, and how you should organize the report when explaining the experimental part of the assignment. Here are some general rules of the report who have to hand in:

- You are expected to **read and follow** instructions via Nestor and Github. Adhere to the rules and conventions of this report template.

- The answers to the assignment questions should have number codes as the questions being answered. Students are requested not to change the number codes. For example, if one is writing the answer of question 2.1 sub part (a), the answer should have the tag 2.1(a). If the answer to the question 2.1(a) is tagged as II.(i)(a) or 2.1.A or (ii).I.(a) or anything which is different from the convention followed in numbering the assignment questions, the answers might be missed and not graded.

- Do not exceed the page number limit of 5pages , single column.

- Do not change the font size and page font.

- Do not change the page margins.

- Provide snippets of codes in the main report only when **explicitly** asked for it. Else provide code in Appendices only if you really want to share.

- We ask for the codes to be submitted separately anyway. We will run these codes and only if they produce the same figures as those in your report, will your report be validated.

The following part guides you on how you should organize the report when answering the practical parts of the assignment, where you have to perform experiments.

1. Start with the correct number and letter tag of the question of the assignment.

2. Provide a **Motivation** for the experiments [*Had to do it for IDS assignment* or *because the question asked me/us to do* will not be accepted as motivation].

3. Mention the problem statement, where you point out exactly which sub-sets of the issues associated with the *Motivation* you are trying to investigate, and why.

4. Give a brief overview of what are the available resources and tools which have have in the past and recent past been applied to find solutions to the problems you are trying to handle.

5. Explain how your experiment report is organized, i.e., you mention that:
   - the first part is the **Introduction** to the topic, which contains the motivation, problem statement and literature review briefly.
   - the second part is the **Methodology** where you have explained briefly the n techniques you have used in the concerned experiment.
   - the third part is the **Results** where you have put your figures, tables, graphs, etc.
   - the fourth part is **Discussion** where you have discussed the results.
   - the report ends with a **Conlcusion** where you have concluded about your findings

1   Methodology

1. Give a brief overview (in 1 or 2 sentences) of how you arranged this section,- *name* the techniques that the reader will read, and what kind of a dataset you got and from where.

2. in <u>separate sub-sections</u> now *describe* :
   - the different techniques and their pros and cons wrt the problem you are trying to address using them (**remember to cite your references**);
   - the dataset (no. of dimensions, types of variables in contains, whether labelled or not, and if yes how many classes are there).
   - how you cleaned the dataset and the preprocessing technique you used.
   - your experimental setup (you may also show it as a flowchart or block dia
   - the evaluation measures you used (accuracy/ precision/ AUC/ etc)

2   Results

- In this section you should have all the figures and tables.
- Please ensure that the report shows your figures and tables exactly where you originally intended it to be. The package called *Float* is quite useful for this purpose.
- Please check that your figures and tables contain **title, and caption**.
- Graphs of any kind should have their axes explicitly defined by you. Also please mention the unit of measurement and keep the tick values of graphs.
- If 2 or more figures are to be compared then please ensure that they are in the <u>same scale</u>, for not just the benefit of your TAs and instructors, but also your own benefit to visualize and interpret your results better.
- If your figure contains colour codes, please use <u>legends</u> and <u>colorbars</u> to help the reader.
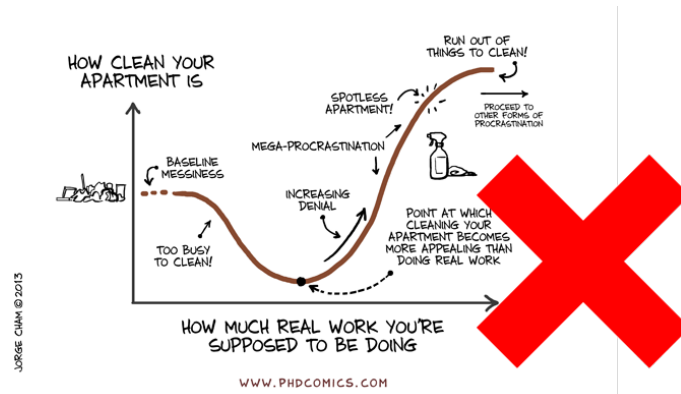

Figure 1: This figure lacks the axes tick labels and units of measurement. We kindly request you to provide not just axes labels but also axes measurements and units of measurements.

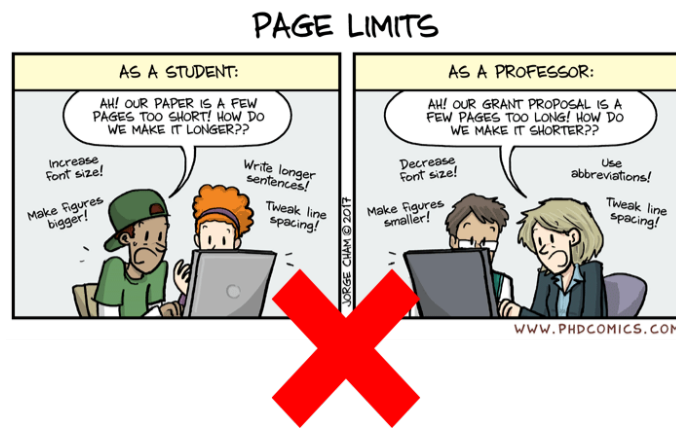Humour aside, please also avoid the following.

Figure 2: Please try neither. If your figure size is so small that the readability and resolution are compromised, then so shall your grades be. If you have been able to keep your report concise and yet been able to describe what you wanted to tell then keep it like that.

If the titles and axes labels are not readable, they will be considered as absent.
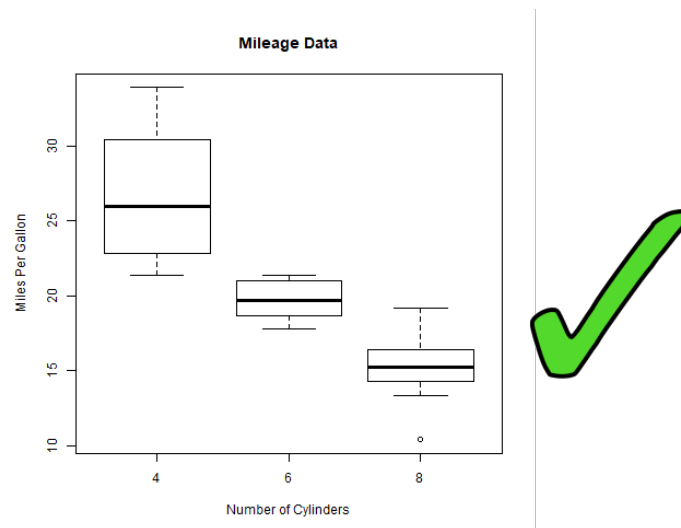

Figure 3: Boxplot example from mtcars data.

Figure 3 was created by the script from tutorialspoint [1]. The above is given as an example of an ideal figure. Notice how it contains proper and visible axes, axes labels, captions, and title. The figure has good resolution. Produce figures of such quality (with respect to resolution, axes information, label information, units, etc).

3 Discussion

Unfortunately due to lack of clairvoyance and omniscience the TAs will not be able to decipher from just your tables and figures what you are trying to suggest, why you did a particular experiment, whether your findings are good or bad and if so why. Therefore you are supposed to:

1. provide explanation of the results
2. guide the TAs and instructors by explaining your results here
3. explain why a plot or figure supports your hypothesis or not
4. which technique is better based on which figures and tables

Please refer to your figures and tables when discussing them.

4 Conclusion

What is the take home message from this experiment you performed? That is, which techniques do you think work better and why. If you had more time which are some things you would have liked to investigate? Did the findings

from the experiments you performed support your initial hypothesis? If not, what did you learn from this analysis about the data and about the question whose answer you tried to find with this experiment.

5 Conclusion

What is the take home message from this experiment you performed? That is, which techniques do you think work better and why. If you had more time which are some things you would have liked to investigate? Did the findings from the experiments you performed support your initial hypothesis? If not, what did you learn from this analysis about the data and about the question whose answer you tried to find with this experiment.

References

[1] tutorialspoint boxplots. https://www.tutorialspoint.com/r/r_boxplots.htm. Accessed: 2017-08-31.

[2] M.F. Lungu K. Bunte. Introduction to data science, 2017.