Applied Data Science Final Assignment
Data Science Project Report

Instructions

For this final assignment, you are required to work on a Data Science problem and submit a detailed report on your approaches, findings and results. You can work on solving the problem posed by this Kaggle competition. You also have the option of taking up your own Data Science problem as the final project. But, you will be required to submit a one page proposal on the problem you are trying to solve, the data used, your approach of the problem and the impact of your solution. Please get your project approved by the Professor or one of the Associate Instructors before you get started with the project. Note that projects that have not been approved will not be graded. If you choose to work on the Airbnb Kaggle problem you do not have to submit a proposal.

Your report should be a short paper, between two (2) and four (4) pages with the following sections: introduction, data, methods, results, discussion, conclusion. Submit your report as a pdf file with your name and email in the header, single-spaced and font no larger than 12pt. Explanations of what is required within each section are below. Place all tables and graphs in an appendix at the end of the document. You should download the required data and may use any algorithm(s) you wish for analysis. Though no consideration will be given to the effectiveness of your model, you should report how well your model predicts the test data. This assignment will be graded on the correctness of your decisions and the overall quality of the report. You do not need to submit your results to Kaggle, though you may do so if you wish.

Outline

The introduction is where you frame your report. Tell the reader why they care about the results you are about to present and why the question you will be answering is important. Specific to this task, why is it meaningful that one can predict the first-booking location of a new Airbnb customer?

The data section is where you describe your data. It is often helpful to have charts and tables in this section that display your data exploration. Explain what the data is and anything about that data that is especially interesting.

In **the methods section**, explain the method(s) you use to analyze the data. Discuss how the method works, why it was well suited for your data, and how you applied it.

The results section is where you describe what the analysis resulted in. Again, it is generally appropriate to use charts or tables here to illustrate important findings.

The discussion section is the place to explain your findings. Why do you think you found the results you did and what do you think they mean? This is generally the only section where it is appropriate to make educated guesses about the phenomena you observed.

Finally, **the conclusion** should provide the analysis' answer to the question posed in the introduction along with a brief description of why the analysis answered the way it did, which should be consistent with your discussion section. Additionally, you may wish to posit questions raised by your analysis or areas for future analysis. In data analysis reports, it is also common to have a recommendation.