Project 2 Info

Project Overview

For project 2, choose a dataset to implement predictive modeling approaches. You should consider at least two different methods (such as decision trees, random forests, linear models, etc..) for your comparison.

The report should contain sections such as introduction, data overview + data visualization, statistical learning procedures, and results and discussion. The following is a rubric that will be used to assess the report.

All items are graded with the following scheme, where some have multipliers:

1. **No Credit**: Criterion was not addressed or was written in a way that was not understandable.
2. **Beginning**: Ideas are not clear and supporting ideas are not presented.
3. **Developing**: Ideas are identified but not well supported and developed or are minimally supported and developed.
4. **Advanced**: Ideas are clearly identified and are adequately supported and developed.

Report generalities	**Points**
Spelling, grammar, writing clarity, paragraphs & section labels	/12
Appendix with complete code	/4
Citations and Acknowledgments	/4
Introduction	**Points**
Project motivation	/4
Research question (Predictive Goal)	/4
Data Overview + Data Visualization	**Points**
Variables with units and descriptive statistics	/4
Titles, Labels, and Captions for Figures	/4
Figure Clarity	/4
Figure Quality	/4
Statistical Learning Procedures	**Points**
Statement and Defense of Loss Function	/8
Define models to fit (with complete notation)	/8
Results and Discussion	**Points**
Comparison of models and defense of model choice	/8
Discuss Predictive Results in the context of the research question	/4
Scope of Inference: how can the results be generalized?	/4
Statement about Uncertainty, how is uncertainty included in your predictions	/4