Your final report should have the following components and follow the APA style as much as possible.

* **Research problem (10pts)**: Describe the task you want to achieve. What is the outcome of interest? What are you trying to predict? Why is it important? What are the potential benefits of having a predictive model for this outcome? Discuss potential applications of such a model.
* **Description of the data (15pts):** Describe core features of the data, any additional features you produced from existing features and how, basic descriptive statistics about these features, and any missing data analysis you conduct. The description should be sufficiently clear that the instructor understands all the variables included in your modeling.

The data was obtained from kaggle: <https://www.kaggle.com/aaronschlegel/austin-animal-center-shelter-intakes-and-outcomes?select=aac_intakes_outcomes.csv>. It was originally provided by the Austin Animal Center in Austin, Texas. The data include information about the intake and outcome of the animal, and details on the type and condition of the animal.

* **Description of the models (15pts):** List at least three different modeling approaches you apply to this dataset. Describe each model, why the given model was selected, which hyperparameters to be optimized and how. Also, discuss how you plan to evaluate model performance.

Three types of modeling approaches were explored: linear regression, bagged trees, and random forests.

* **Model fit (20pts):** Provide the results of your model evaluation. Compare and contrasts results from different fits, including a discussion of model performance. Discuss your final model selection and the evidence that led you to this selection. If it is a classification problem, how did you choose a cut-off point for binary predictions? Did you consider different cut-off points?
* **Data visualization (5pts):** Include at least two plots (or more) to help communicate your findings. The plots may be of initial data explorations, fits of individual models, and plots displaying the performance of competing models.
* **Discussion/Conclusion (25pts):** Discuss and summarize what you learned. Which variables were the most important in predicting your outcome? Was this expected or surprising? Were different models close in performance, or were there significant gaps in performance from different modeling approaches? Are there practical/applied findings that could help the field of your interest based on your work? If yes, what are they?
* **Reproducibility (10pts):** Provide a link to the GitHub repo at the beginning of your report as a note.

The final report should be no longer than 10,000 words (approximately 15 single-spaced pages), excluding the tables, figures, and references.