Team Members:

- Justin Hsie, jih2119
- Dayoung Yu, dry2115
- Christian Pascual, cbp2128

Tentative Project Title:

- Predicting chance of admission to Master's programs using application characteristics

Anticipated Data Sources:

- Graduate Admissions Fromfrom Kaggle
    - https://www.kaggle.com/mohansacharya/graduate-admissions

Description of Data

The dataset consists of our response variable **Chance of Admit** and 7 other predictor variables. **Chance of Admit** is a continuous number, taking possible values from 0 to 1, 1 representing a 100% chance of admission. The predictors represent various important factors in the admission process, including: GRE score, TOEFL score, the rating of the university being applied to, statement of purpose strength (SOP), letter of recommendation strength (LOR), and research quality. GRE, TOEFL and university rating are continuous, while SOP, LOR and research quality are categorical.

Planned Analyses:

For our analyses, we plan to create two separate regression models to predict the chance of admission: a normal model and a LASSO model. Before fitting these models, we plan to see how each of the predictors is correlated with the response and to possibly control for any multicollinearity. Since we are performing a LASSO regression, we'll also perform cross-validation to find an optimal tuning parameter.

After creating the models, we'll assess which covariates are significant in predicting the chance of admission. If any are possibly insignificant, we will perform an ANOVA to test if we can remove them. To assess which model has better predictive ability, we'll perform a 10-fold cross-validation on both. At the end, we'll be able to say which factors affect chances of admission and the magnitudes of these effects.