

# Project Proposal

*Adriana Souza, Roger Filmeyer, Lydia LaSeur*

*March 25, 2018*

## **Who is in your group? What roles do you see each person fulfilling?**

Lydia LaSeur, Roger Filmeyer, and Adriana Souza. Everyone will contribute equally at the beginning stages until we have a more definite plan and can specialize on the parts that each of us are most interested in.

## **What is your data? Write a short (1 paragraph) description of the data. List any variables and their units. Describe the units of analysis (e.g. people, countries, etc.). State the dimensions of the data and the source of the data.**

Our data was kindly provided to us by the Admissions Office of New College of Florida. We have 3 datasets: *Enrolled Student Data*, *Admitted Student Data*, and *Denied Student Data*. Here is a breakdown of the variables:

- **Ref:** Internal reference number for the student.
- **First:** First name.
- **Middle:** Middle name.
- **Last:** Last name.
- **App Student Type:** In these sets it tends to be “Freshman.”
- **Round:** Semester of application
- **Decision:** Accepted or Denied.
- **Course Selection Rating:** Variable that tries to capture the rigor of the students’ high school education curriculum, 1-8.
- **Final NCF GPA:** Internal calculation of GPA based on NCF’s appraisal of the high school’s curriculum rigor; mostly to avoid inflated GPAs.
- **ACT Composite**
- **SAT I Superscore (CR + M) – Verified**
- **SAT R Superscore**
- **Essay Rating:** Internal rating of the application essay for New College.

## **What is your goal? Describe in broad terms (1 paragraph) the story you plan to tell with this data or the problem you intend to solve.**

Our goal is to standardize the way students are being admitted by creating a very simple model that weights each of the application elements according to what New College is looking for in terms of applicants. We would also like to create visualizations that help tell this story, beyond the formula (e.g. a visual profile on the average/accepted/denied applicant at New College).

## **Who is your audience? Who would want to understand this story or solve this problem?**

We have two separate audiences: the people who want to use the formula we come up with for weighting each element and, ideally, prospective students who would like to apply. The students could have a quick, visual representation of what is required to get in, for example. This is open to discussion.

**What types of graphs will you likely use? Describe the specific groups of variables you will include in each graph and suggest one or more options for each grouping.**

Depending on the exact story we decide would be most useful for prospective students, the type of graph will vary. However, the variables we will use are `Final_NCF_GPA`, all of the SAT and ACT variables, and `Decision`. There are many examples online of visualizations for accepted students in colleges.

**What level of interactivity will you incorporate?**

We would like the visualization to be interactive but simple; hovering over for more details, etc.

**What software will you use?**

We will use a combination of RStudio, Python, Tableau, and whatever else may come up that we find useful.