CAAP Statistics Final Project-1
Find a Dataset and Understand the Data

Write your name here

Candidates

- **Here you can list up any datasets you are interested in!**

  - Try to find a dataset you can think of interesting questions

  - In general, numerical variable is easier to deal with than categorical variable

- Potential sources for the datasets are

  - Kaggle datasets

  - Openintro datasets: `data(package="openintro")`

  - UCI Machine Learning Repository

Dataset Name

- **Here you can summarize the reason you choose that particular dataset**

  - Why do you end up with this dataset?

  - Which aspect interest you the most?

  - What might be the potential interesting questions?

  - What might be the potential challenges when working with this dataset?

Understanding the Data

- **Here you try to understand your data.**

  - Write down any useful(might be external information not given by the data source website) information to understand the data

  - Try to understand the meaning of each columns

  - Try to calculate/plot the summary of each variable and see if there is any interesting observation