Final Project

# Members

* Katie Hill
  + kathleen.hill@tamu.edu
* Asmita Nagila
  + anagila@tamu.edu
* Jerrin Wiley
  + [jerrin@tamu.edu](mailto:jerrin@tamu.edu)
  + 832-588-4018

# Dataset

* Need to choose from <https://archive.ics.uci.edu/datasets>
  + Must have ≥ 30 features and be suitable for supervised learning (labeled data)
  + Ideas:
    - Predict students' dropout and academic success
      * https://archive.ics.uci.edu/dataset/697/predict+students+dropout+and+academic+success

# Report

## Data Description

* Background of the data set
* Describe the data set (number of variables, types of variables, etc.)

## Goal

* What is your goal?
  + Classification
* What are you interested in finding out from the data?
  + Classification of student (dropout, enrolled, and graduate)
* Whether it is a classification or regression problem—explain.

## Methods

* Describe the two methods that you have chosen to analyze the data
  + For each method, provide a paragraph of general description/summary of the method in your own words.
* Explain why you choose these techniques.

## Comparisons and findings

* Compare these methods based on the data
  + Describe how you compare the two methods and explain why you choose to compare them in this way
  + Discussion and suggestions
* Report any interesting findings of the data

# Timeline

* Initial meeting: **Done**
* Select dataset and goals: **Done**
* Select 2 supervised learning methodologies: 11/17
* Code Complete: 11/25
* All information gathered and consolidated: 11/29
* Rough Draft: 12/4
* All team members sign off on final draft: 12/6
* Final due date: **12/7**