# Project 3 Proposal
# Predicting Heart Disease

Bianca Orozco

## Domain

UCI Machine Learning Repository provides databases, domain theories, and data generators used by the machine learning community. I will be using their Heart Disease Dataset, specifically the dataset created by the Cleveland Clinic Foundation. It provides 13 features and a target variable referring to the presence of heart disease in the patient. Target is an integer value from 0 (no heart disease presence) to 4.

## MVP

The analysis will predict the presence of heart disease presence in a patient.

## Heart Disease Data

|     | Variable | Type | Description | Used for Model |
| --- | --- | --- | --- | --- |
| 1. | age | int | age in years | Y |
| 2. | gender | int | 1=male, 0=female | Y |
| 3. | cpain | int | chest pain: 1=typical angina, 2=atypical angina, 3=non-anginal pain, 4=asymptomatic | Y |
| 4. | restbp | int | resting blood pressure | Y |
| 5. | chol | int | serum cholesteral in mg/dl | Y |
| 6. | bsugar | int | fasting blood sugar>120 mg/dl 1=true, 0=false | Y |
| 7. | restecg | int | resting electrocardiographic results | Y |
| 8. | maxhr | int | maximum heart rate achieved | Y |
| 9. | exercisecp | int | exercise induced chest pain | Y |
| 10. | oldpeak | float | ST depression induced by exercise | Y |
| 11. | slope | int | the slope of the peak exercise ST segment | Y |
| 12. | ca | int | number of major vessels colored by flourosopy | Y |
| 13. | thal | int | 3=normal, 6=fixed defect, 7=reversable defect | Y |
| 14. | target | int | diagnosis of heart disease: 0=no presence to 4 | Y |