- What is the problem you are attempting to solve?

I am trying to solve the problem of how to predict heart disease. I have family members who have been diagnosed with heart disease and I wanted to help find the predictors to see if they may get it. About **610,000 people** die of heart disease in the United States every year–that's **1 in every 4 deaths**. Also, heart disease is the leading cause of death for both men and women. **More than half** of the deaths due to heart disease in 2009 were in men.This isn't a fool-proof way to determine if someone will get heart disease but checking via data can certainly give us a better idea of what to look for to help prevent heart disease.

- How is your solution valuable?

If we can determine the factors that contribute to heart disease, we can find ways to prevent heart disease. Currently, heart disease is the number 1 killer of adults.If we can reduce that in any way, it's a win for humanity.

- What is your data source and how will you access it?

I am retrieving my data from Kaggle who got it from the University of California, Irvine.

- What techniques from the course do you anticipate using?

Supervised and unsupervised learning.

- What do you anticipate to be the biggest challenge you'll face?

A potential hurdle is that my data may not accurately represent the population seeing as my sample size is rather small (1000 rows).