Capstone Two: Project Proposal
Predict whether diabetic patients are readmitted into the hospital

130 hospitals and integrated delivery networks have amassed 10 years of clinical care information for diabetic patients. These patients have spent at least 1-14 days admitted into one of these hospitals. After undergoing the appropriate lab tests and receiving treatment, they are released from the hospital. It is important to know if a patient will be readmitted to the hospital or not and why. Readmission can be attributed to several factors such as, problems with the treatment or the state of the patient. Treatment can be changed in order to prevent readmission.

We are interested in having the ability to predict whether diabetic patients are readmitted into the hospital within 30 days of release, more than 30 days of release, or not at all.

Several factors will determine a successful outcome. Since predictions can't be absolute, **false positive and false negative** are two problems we have to deal with. In our case, a **false positive** is a prediction indicating that a patient will be readmitted into the hospital when they should not, which can lead to unnecessary hospitalization and financial costs. A **false negative** is a prediction indicating that a patient will not be readmitted into the hospital when they should and consequently does not receive the necessary care. I've determined that a false negative is more costly than a false positive so the goal is to keep the false negatives to a minimum and to successfully predict whether patients are readmitted into a hospital with a false negative rate of 5% and below.

The project scope will focus on the data detailed in the csv file. We will examine and compare the laboratory test results, treatment, outcome and the state of patients that were not readmitted to a hospital with those patients that were readmitted.

In order to minimize constraints that may prevent this project from succeeding, we will determine if we have all the relevant data for analysis and whether the data is accurate. We'll also perform some preprocessing, such as dealing with missing values to ensure the data can be properly analyzed, as well as exploratory data analysis.

The relevant Stakeholders that can provide key insight are the Hospital Board or Directors, the President and Administration and the Patient's doctors and caregivers.

Capstone Two: Project Proposal
Predict whether diabetic patients are readmitted into the hospital

The key data source is a csv file which represents 10 years (1999-2008) of clinical care at 130 hospitals and integrated delivery networks. This data has been prepared to analyze factors related to readmission as well as other outcomes pertaining to patients with diabetes.
https://archive.ics.uci.edu/ml/datasets/Diabetes+130-US+hospitals+for+years+1999-2008