**Midterm Project Report**

**Introduction to the Business Decision Process**

**Name: Abhignu Reddy Alety Student ID: 11443224**

**About Dataset:**

The dataset is regarding 11 clinical features for predicting heart disease events.

<https://www.kaggle.com/fedesoriano/heart-failure-prediction>

**Data set Description:**

The dataset is based on 11 clinical features for predicting heart disease events. It has 918 rows and 12 columns

The following are the attributes of the data set which will be using in the project:

Date

Age: age of the patient

Sex: Sex of the patient

Chest pain type: [Typical angina, ATA, NAP]

Resting BP: resting blood pressure

Cholesterol: Serum cholesterol

Fasting BS: Fasting blood sugar

Resting ECG: resting electrocardiogram

Max HR: maximum heart rate achieved

Exercise Angina: exercise-included angina

Old peak: oldpeak= ST [Numeric value measured in depression]

ST\_Slope: the slope of the peak exercise ST segment

Heart disease: output class

**Research questions:**

Using the above dataset to perform the following analysis:

Q1. Devices manufactured by TSI Electronics to detect heart attacks have life spans that have a normal distribution with a standard deviation of 3000 hours and a mean life span of 15,000 hours. If a monitor is selected at random, find the probability that the life span of the monitor will be more than 13,000 hours. Round your answer to four decimal places.

Q2. The mean of people with heart failures due to Cholesterol is 198.7996 with a standard deviation of 3.610215.

If a sample of 69 people were randomly selected, what is the probability that the sample mean would differ from the true mean by less than 0.5. Round your answer to four decimal places.

Q3. A Doctor examines 42 people for heart failure due to Cholesterol. The mean cholesterol for the sample data is 142.56 with a standard deviation of 1.2102. Determine the 80% confidence. Interval for the population mean cholesterol concentration. Assume the population is approximately normal.

Find the critical value that should be used in constructing the confidence interval. Round your answer to three decimal places.

Q4. The director of research and development is testing a new drug. She wants to know if there is evidence at the 0.02 level that the more concentration of cholesterol in the system results in heart failure. For a sample of 82 patients, the mean of cholesterol concentration in the system was 152.856. Assume the population standard deviation is 2.8025. Find the p-value of the test statistic. Round your answer to four decimal places.