HEART DISEASES

Your heart is one of your body's most important organs. Essentially a pump, the heart is a muscle made up of four chambers separated by valves and divided into two halves. The general term used to cover malfunctions of the heart is Heart Disease, or sometimes Cardiac Disease ("Cardiac" is a Latin term for the heart). Although heart disease can occur in different forms, there is a common set of core risk factors that influence whether someone will ultimately be at risk for heart disease or not. We start our discussion of heart disease by describing these common risk factors, and then move on to cover specific conditions.

Borrowing words from the UIC Machine Learning Repository description, This Report is a landmark survey of Hungarian Institute of Cardiology. The report has been widely appreciated in Heart institutes with the aim of helping individuals, institutes, insurance companies make better policies aimed at improving the quality of life of its residents & being a pseudo indicator for measuring the better health for individuals.

I am planning to use the Hungarian Institute of Cardiology Reports 2011,2012 for the people residing in one of the four countries Cleveland Ohio, Hungary, Switzerland, and the VA Long Beach (California). There are 14 Variables collected - Age, Sex, Chest pain, Resting Blood Pressure, serum cholesterol, fasting blood sugar, resting electrocardiographic results, maximum heart rate achieved, exercise induced angina, ST depression induced by exercise relative to rest, slope, restckm, number of major vessels, angiographic disease status, thal.

I also believe that this would be an interesting avenue to use techniques involved in longitudinal studies. If it turns out that the study is a longitudinal study, then I would like to explore the possibility of either using thal score as the variable for conducting a univariate analysis on a summary of the multivariate response. It is also likely that there would be high correlation between the variables & I would have to think of conducting a separate univariate analysis on each of the several summaries (if other variables are significant) or conducting a multivariate analysis on several summaries

I am planning to address above problems Using T-tests, Anova, non-parametric t-tests, univariate and multi-variate analysis

1. Heart diseases in different genders
2. Heart diseases in different age groups
3. Heart diseases in healthy people with low blood pressure and low depression
4. Heart diseases wrt to chest pain, serum cholesterol, and maximum heart rate achieved
5. Geographical locations and its heart disease effects
6. Does angiographic disease status help to detect proneness of Heart Disease
7. How good is insurance for a person after age 50?

I do believe that there would be interesting insights which simple exploration plots would provide & while the findings may not be causal, they will surely provide an idea of how heart diseases are seen in different people of different age groups across different cities & also answering, where do the people with most heart diseases belong.