Assignment 4: Using Data Visualization for Problem Solving

This assignment will require the use of data visualization techniques in order to investigate specific business or social problems. In this specific case, your task is to creatively combine the available data with the available visualization tools to answer a particular question by applying the scientific method. Try to strike a balance between the best charts/graphs that allow answering the specific question and those that have more power to communicate certain ideas.

Case description

You are an analyst at a healthcare consulting company and your new client is a large hospital chain that wants to deploy a business intelligence solution for its research institution. This institution collects, on a regular basis, data about patients that received treatment in one of their hospitals, but due to lack of appropriate analytical solutions and a shortage of qualified medical analysts, the scientists could not make a significant progress in obtaining useful insights. As part of the engagement, you are assigned the role to assist the scientists in analyzing the data about the patients and provide the information to the medial research team. In addition, you would like to showcase the potential of BI tools in advancing medical research and also to give training to the analysts from the research team.

The scientists from the research institution have a list of hypotheses that they would like to test using the existing patient data. These hypotheses are based on the theories and findings from the latest medical literature. The team shared with you the data about patients with heart diseases (HPS.HEART data set in SAS VA application) and you are to test the following hypotheses:

- H1: The weight and cholesterol levels are correlated
- H2: Men are usually more obese than women

- H3: Women usually smoke less than men, but their cholesterol level is higher
- H4: The blood pressure is higher for people with higher cholesterol levels

In addition, you were asked to provide some distinctive characteristics for people who suffered from coronary heart disease and make some assumptions regarding the potential underlying causes of this illness.

Use all the available data to test the hypotheses mentioned above and also provide some key facts regarding the patients who suffered from coronary heart disease. For each hypothesis use at least one chart/graph and for the heart disease research use at least three charts/graphs. Discuss your findings briefly in a write-up.

Note: please use only the data provided in the case and don't consult any external resources regarding the questions mentioned above. Your answers are supposed to be a high-level overview and they should be focused mainly on the possibilities of using BI for problem solving.

Once finished, submit your write-up as a single file to your instructor (via email attachment or CMS/dropbox) or print and submit the document as required.