**Principles of Data Science (5530)-Assignment 1**

**Name: Aishwarya Musuku**

**ID: 16344653**

**1.Answer:**

Stage 1- Data collection:

In this step, raw data is collected, fetched, and saved.

**Folder structure:**

**A black screen with white text

Description automatically generated**

◦The gathered information is kept in a.csv file format to facilitate analysis.

◦The dataset's metadata is contained in the README.txt file.

Stage 2-Data processing:

Here, data that has been collected or retrieved will undergo preprocessing to make it suitable for analysis.

A screenshot of a computer

Description automatically generated

◦Various preparation procedures are executed, such as assigning suitable labels to the columns and transforming the categorical values into numerical values. The final, uncontaminated data is then stored in a different file for analysis.

Folder Structure:

A screenshot of a computer

Description automatically generated

Stage 3-Data Analysis:

Data analysis is carried out on the dataset by executing the scripts in order to identify any links or patterns between the fields or observations.

A screenshot of a computer

Description automatically generatedA screenshot of a computer screen

Description automatically generated

◦At this stage, the cleaned data is analyzed.

◦The plots of two distinct analyses done on the frailty dataset are shown below.

◦Data of weight on the Y-axis and frailty on the X-axis are plotted on a box plot graph.

A diagram of a couple of rectangular objects

Description automatically generated with medium confidence

**Analysis Outcome:** The box plot unequivocally demonstrates that those with frailty symptoms have weaker grips than those without.

Age is plotted on the Y-axis and frailty is plotted on the X-axis in a box plot graph.

A diagram of a relationship between age and fracturing

Description automatically generated

**Analysis Outcome:** It is evident from the box plot that individuals exhibiting or experiencing frailty symptoms are limited to the age range of 22 to 35. This suggests that the age range of 22 to 35 is when frailty symptoms are most commonly seen.

**Folder Structure:**

**A screenshot of a computer

Description automatically generated**