**Principles of Data Science (5530-0001)-Assignment-1 Name: Mohana Sai Kumar Reddy Bobba**

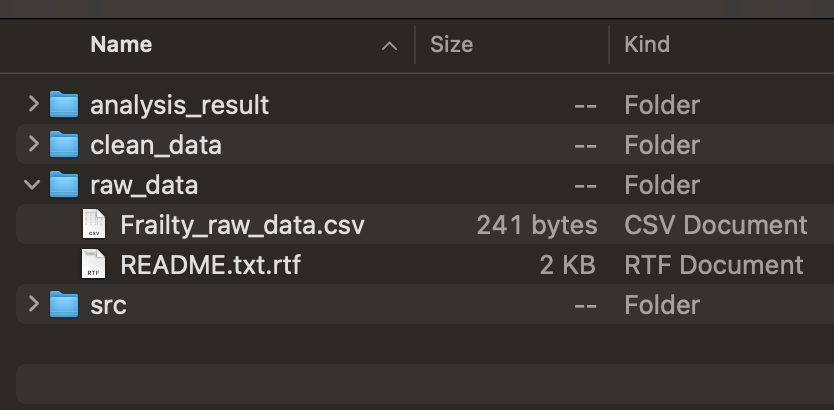
**Id: 16337925**

**1)**

a) Data collection:

The data is collected and stored in a suitable format.

**Folder structure:**



* Gathered the data and stored it in a file with .csv format in order to perform the analysis.
* README.txt file contains the metadata about the dataset.

Text

Description automatically generated

[README.txt file content]

Stage 2 – Data Processing:

Gathered data will be processed here in order to transform the data to be analysis ready.

* Preprocessing of data is performed using the scripts.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

* Initially the data is loaded into the script and then certain preprocessing steps have been performed like renaming with suitable column names and converting the categorical values into numerical values making them more efficient for analysis. Validated if there are any records with NA occurrences. And the final obtained clean data which is ready to perform analysis is stored as a separate file.

**Folder Structure:**

**Background pattern

Description automatically generated**

**Stage 3 – Data Analysis:**

* Data analysis is performed by running the scripts to find the pattern of relation between the fields of the dataset.

**A screenshot of a computer

Description automatically generated**

* **In this stage analysis is performed on the processed clean data.**
* **Two sample t-test analysis is done between weight and frailty and results are plotted.**

**Chart, box and whisker chart

Description automatically generated**

* **box plot graph is generated with the X-axis having Frailty and the Y-axis having Grip strength data.**

**Chart, box and whisker chart

Description automatically generated**

**Analysis Result:** The box plot representation clearly shows that greater the Grip Strength, the stronger will be the individual person.

**Folder Structure:**

**A screenshot of a computer screen

Description automatically generated with medium confidence**