1. In workflow we have three steps i.e., Data Collection, Data Processing, Data Analyzing &Visualization

A screenshot of a computer

Description automatically generated

* Data Collection:

In this step I prepared a CSV file from the given data and stored it under rawdata project structure to perform Analysis

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated

* Data Processing

In this step I did data cleaning by modifying the Frailty feature from categorical to 0 and 1 for better analysis on the data.

Clean data is stored under cleandata folder.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

* Data Analyzing &Visualization

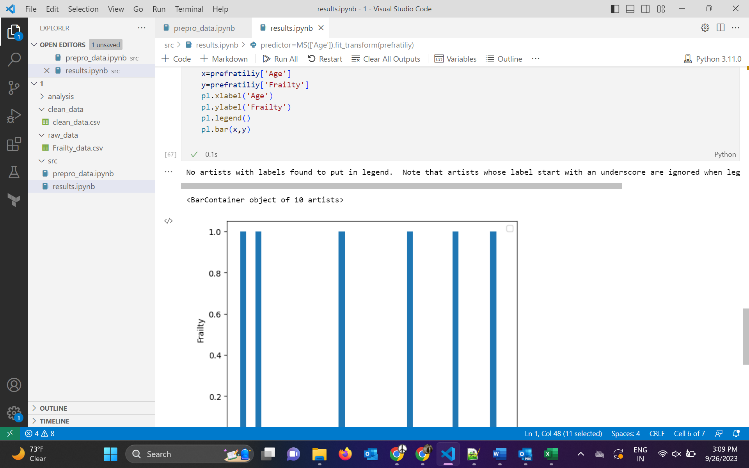
In this step of workflow, after cleaning the data I have again loaded the clean data and performed linear regression on the features frailty, age frailty as predictor and age as response variable. After fitting the data in OLS model, I found that age is having p-value as 0.818 which is greater than 0.5 so age is insignificant to frailty.

I have also visualized the features age and frailty using bar plot.

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated