Part A: For each of the following scenarios list the type of t-test that should be used (2-sample or paired)

1. Is global pathology score different for people who are younger than 85 when they die compared to people who are older than 85?

2. Is the mean years of education among males higher than females?

3. Does the number of major depressive episodes change before and after medical treatment in a group of 25 people with bipolar disorder?

4. Does BMI differ for people who exercise 2 or more times week compared to people who exercise less than 2 times per week?

Part B: Use the FHS dataset (frmgham2.csv) to complete the work below:

1. Read in the frmgham2.csv dataset
    a. Create a new dataset (subset of the larger dataset) including only PERIOD=1

2. Let's compare ages between those with prevalent CHD and those without

    a. Generate descriptive statistics (mean, sd, median, min, max, q1, q3) for age by PREVCHD diagnosis status. There are multiple ways to do this.
        i. Try it by sub-setting the data

        ii. Try it by using a 'by' statement
            - Example: by(FHS1$AGE, FHS1$PREVCHD, summary)
            - Note: by(*continuous*, *categorical*, *function*)

b. Use a hypothesis test to see if there is a difference in average ages between those with prevalent CHD and those without? Be sure to: define your null and alternative hypotheses, state the test statistic, p-value, and state a final conclusion.

c. Interpret your results – based on the spread of the data and the results of the ttest, is this difference in age meaningful?