One Sample T-test

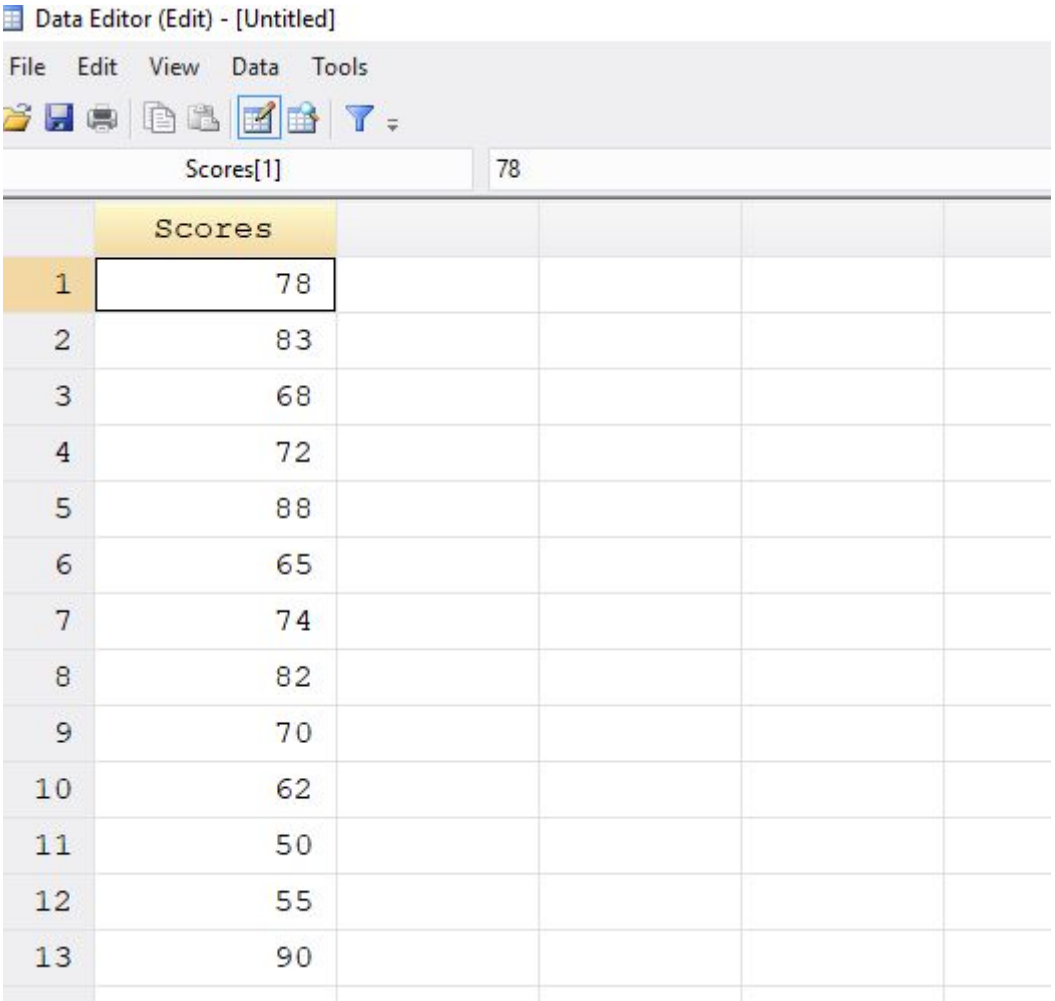
A One sample t-test tests the mean of a single group against a known mean.

Here we will use a sample set of 20 students' scores against a class mean score.

Load data

.import excel "student_t_test\data\t-test2.xls", sheet("One-sample t-test") firstrow

//Browse the imported dataset
.Browse



	Scores
1	78
2	83
3	68
4	72
5	88
6	65
7	74
8	82
9	70
10	62
11	50
12	55
13	90

Let’s summarize data

```
. summarize Scores

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
      Scores |         20       70.15    10.91245        50         90

.
```

As seen above, the mean is 70.15 but our class mean is 79.

Question
Is there a statistically significant difference between the sample mean from 79 ?

Hypothesis
H0: There's no difference between the sample mean from 79
Ha: There's a statistically significant difference between the sample mean from 79

The level of significance
alpha = 0.05

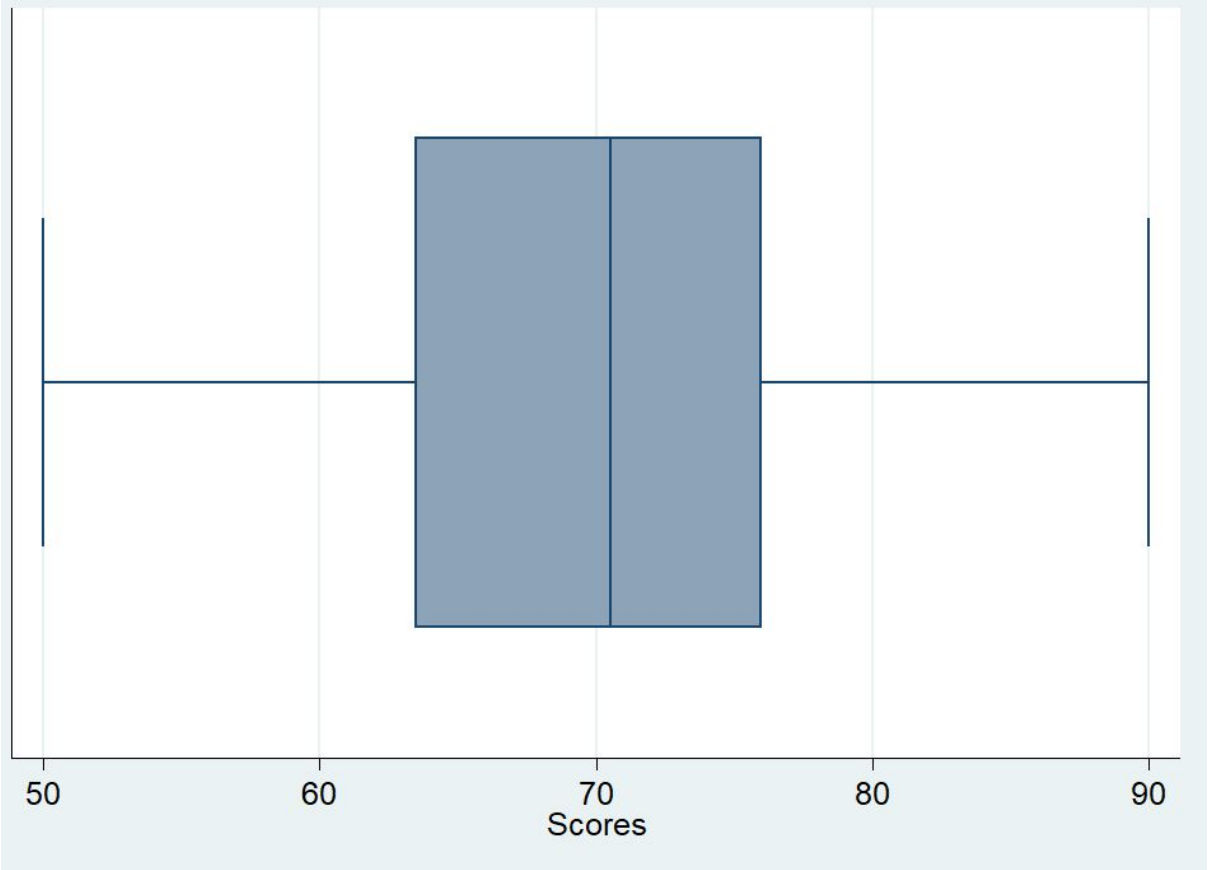
Assumptions
Determine if data meets requirements to perform a one samples t-test.

Assumption #1: Your Test variable should be measured on a continuous scale.
Assumption #2: You should have independence of observations.
Assumption #3: There should be no significant outliers.
Assumption #4: Your dependent variables should be approximately normally distributed.

CHECK FOR OUTLIERS

Check outliers by plotting a boxplot

.graph hbox Scores



As seen from the above boxplot we don't have outliers from our sample.

NORMALITY TEST

Normality Law test using Skewness Kurtosis test for normality

H0: The data follows a normal distribution.

Ha: The data does not follow a normal distribution.

```
. sktest Scores

                    Skewness/Kurtosis tests for Normality
                                                      ------ joint ------
    Variable |        Obs  Pr(Skewness)  Pr(Kurtosis) adj chi2(2)    Prob>chi2
-------------+---------------------------------------------------------------
      Scores |         20      0.9085        0.9546        0.02         0.9918

.
```

Normality Law test using Shapiro-Wilk W test for normal data

H0: The data follows a normal distribution.

Ha: The data does not follow a normal distribution.

```
. swilk Scores

                    Shapiro-Wilk W test for normal data

    Variable |        Obs       W           V          z       Prob>z
-------------+-------------------------------------------------------
      Scores |         20     0.97242       0.653     -0.860    0.80504

.
```

Normality Law test using Shapiro-Francia W' test for normal data

H0: The data follows a normal distribution.

Ha: The data does not follow a normal distribution.

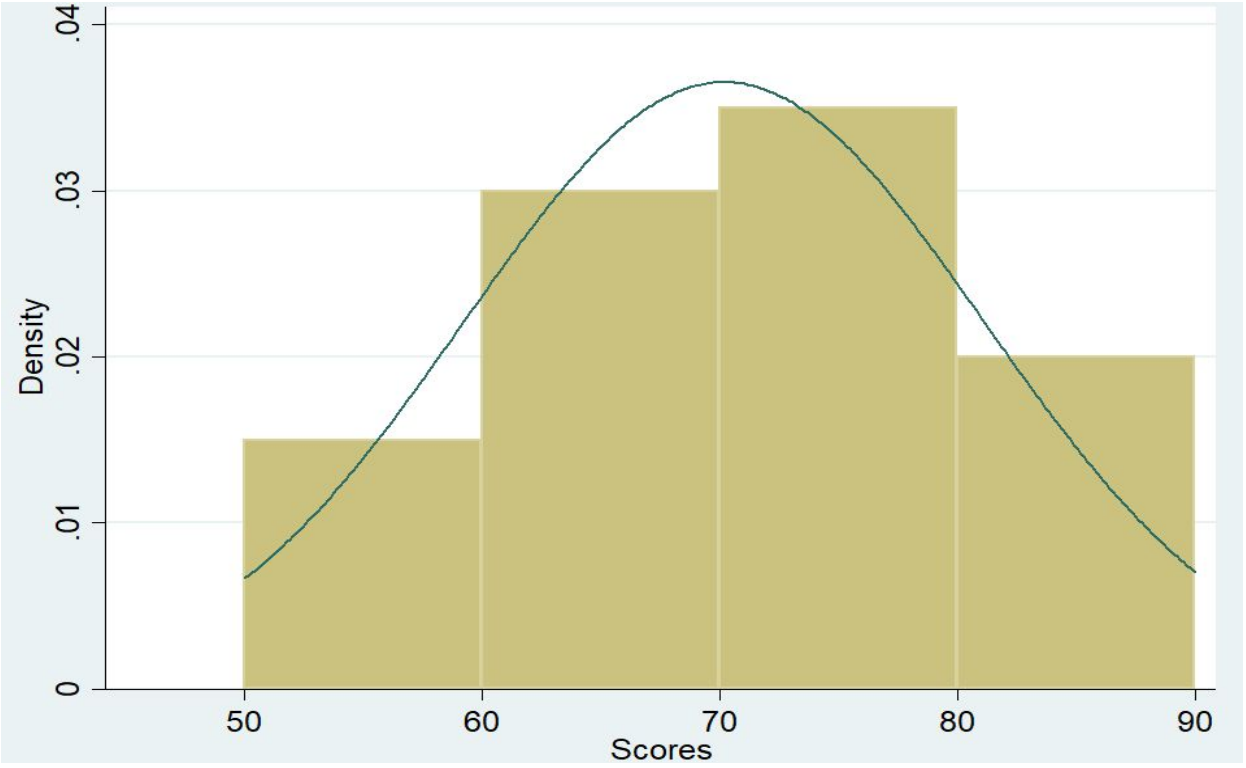
```
. sfrancia Scores

                  Shapiro-Francia W' test for normal data

    Variable |        Obs       W'          V'         z       Prob>z
-------------+-------------------------------------------------------
      Scores |         20     0.97840       0.568     -1.012    0.84420

.
```

As you see in all the above 3 tests (off course you don't need all the 3), they all are NOT significant and so we have no evidence to reject the H0, that states that the data follow a normal distribution. We can plot a histogram to visualize;



One Sample T-test

With all data requirements for One Sample T-test satisfied, let us not run the test.

```
. ttest Scores ==79

One-sample t test

Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
  Scores |      20       70.15    2.440098    10.91245    65.04282    75.25718
---------+--------------------------------------------------------------------
    mean = mean(Scores)                                           t =  -3.6269
Ho: mean = 79                                   degrees of freedom =       19

    Ha: mean < 79                Ha: mean != 79                 Ha: mean > 79
 Pr(T < t) = 0.0009         Pr(|T| > |t|) = 0.0018          Pr(T > t) = 0.9991

.
```

```
. summarize Scores

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
      Scores |         20       70.15    10.91245        50         90

.
```

The test (t(9)=-3.63, p=0.0018) is significant at alpha=0.05. Thus we have enough evidence to reject the null hypothesis in favour of the alternative hypothesis that the sample mean score of 70.15±10.91 is significantly different (lower) compared to the general mean score of 79.00 for this test in the class.

The average score of the sample is about 8.85 units lower than the overall hypothesized average.