IBM SPSS

# # 1 (Data Analytics) Q1:

A football coach is frustrated with his team’s lack of speed. He measures each player’s 30-yard dash speed and then sends all of them to a speed and agility camp. He then measures their times again after. The data is below. Is there sufficient evidence to say that the camp helped the players speed? Run a test.

|  |  |
| --- | --- |
| Before | After |
| 5.88 | 5.76 |
| 5.89 | 5.21 |
| 3.41 | 3.35 |
| 5.79 | 5.89 |
| 4.92 | 4.78 |
| 4.96 | 4.54 |
| 4.95 | 4.79 |
| 4.66 | 4.88 |
| 5.56 | 5.08 |
| 5.73 | 5.65 |
| 5.05 | 5.1 |
| 4.44 | 4.79 |
| 4.27 | 4.78 |
| 4.87 | 4.95 |
| 5.55 | 4.98 |

Is there evidence that the team gets significantly better performance after camp? Use a 0.05 level of significance and a test.

1. Write an appropriate hypothesis test for this situation and state the appropriate testing procedure.
2. Compute the necessary summary statistics for the test in part (a).
3. Perform the t-test and report the p-value.
4. Interpret your results in the conclusion

## 1). Manual

#### Write an appropriate hypothesis test for this situation and state the appropriate testing procedure.

Since we are comparing the results of same players (before and after performance of a camp in terms of speed and agility), we will use a paired sample t-test.

In this test we consider two different hypothesis.

1. Null Hypothesis (H0): The mean of players time before and after the camp are equal

H0: μ\_before = μ\_after

1. Alternative Hypothesis (H1): The mean of players time before and after the camp are different

H1: μ\_before ≠ μ\_after

#### Compute the necessary summary statistics for the test in part (a).

|  |  |  |
| --- | --- | --- |
| **metrics** | **Formulas** | **Calculations** |
| Number of Samples (n) | count(sample) | 15 |
| Mean Before camp (Xb) | sum of players before time/n | 5.062 |
| Mean After camp (Xa) | sum of players after time/n | 4.969 |
| Mean of differences (Xbar) | Xb-Xa | 0.093 |
| Standard deviation of differences (Sbar) | Excel formula =STDEV(range of difference in time) | 0.336 |
| Degree of freedom (df) | n-1 | 14 |

1. ***Perform the t-test and report the p-value.***

**Formulas:**

t-Statistic = Xbar – 0/Sbar/square root of n

T Critical value: t statistic with degree of freedom (n-1)

P value: p value for the t-statistic

1. To calculate T Critical value, I used the below website: https://goodcalculators.com/student-t-value-calculator/

Where we need to enter Degree of freedom and Significance level value.

1. To calculate P value, I used the below website: https://[www.socscistatistics.com/pvalues/tdistribution.aspx](http://www.socscistatistics.com/pvalues/tdistribution.aspx)

Here we need to enter the t-statistic value, Degree of freedom and Significance level value.

In this website, we can compute p-value of one and two-tailed hypothesis.

### Calculations:

Significance value: 0.05

t-Statistic = 1.075

T Critical value: t statistic with degree of freedom (n-1) : 1.761 P value: p value for the t-statistic:

a). P-value one-tile: 0.1502 b). P-value two-tile: 0.3005

#### Interpret your results in the conclusion

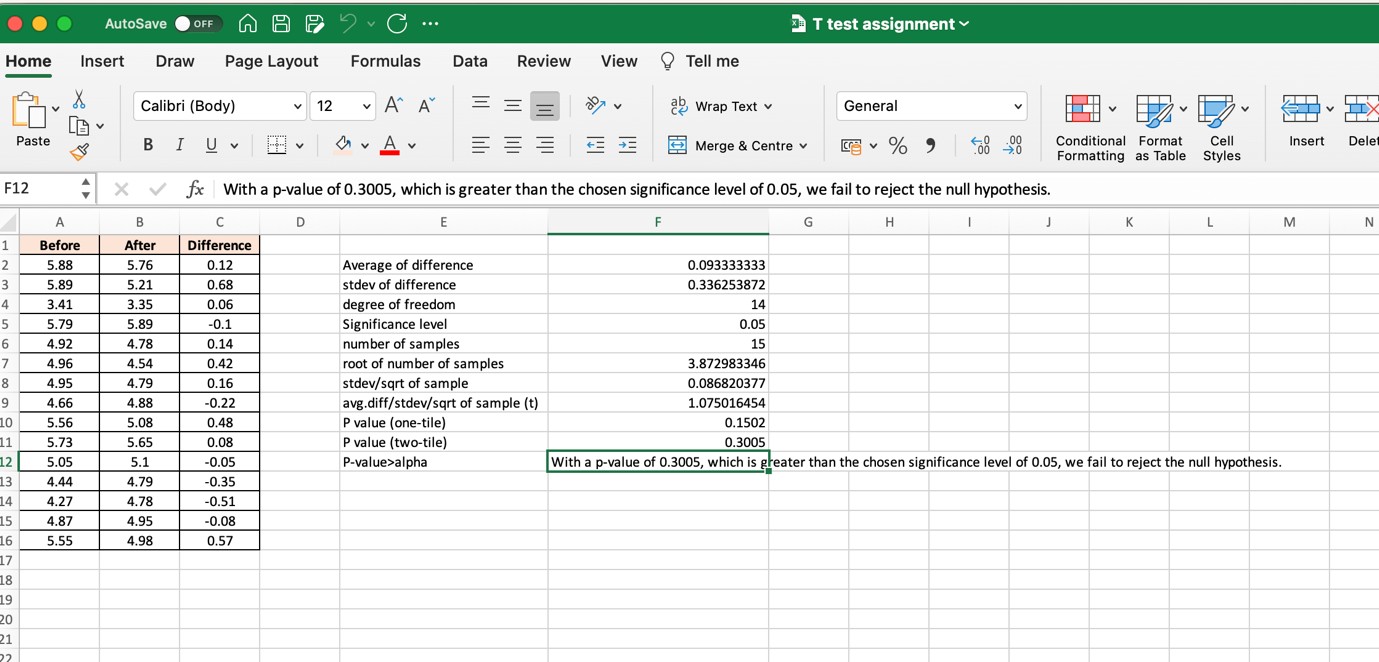
With a p-value of 0.3005, which is greater than the chosen significance level of 0.05, we fail to reject the null hypothesis.

Therefore, based on this analysis there isn't sufficient evidence to conclude that the speed and agility camp had a statistically significant effect on improving the players speed in the 30 yard dash.

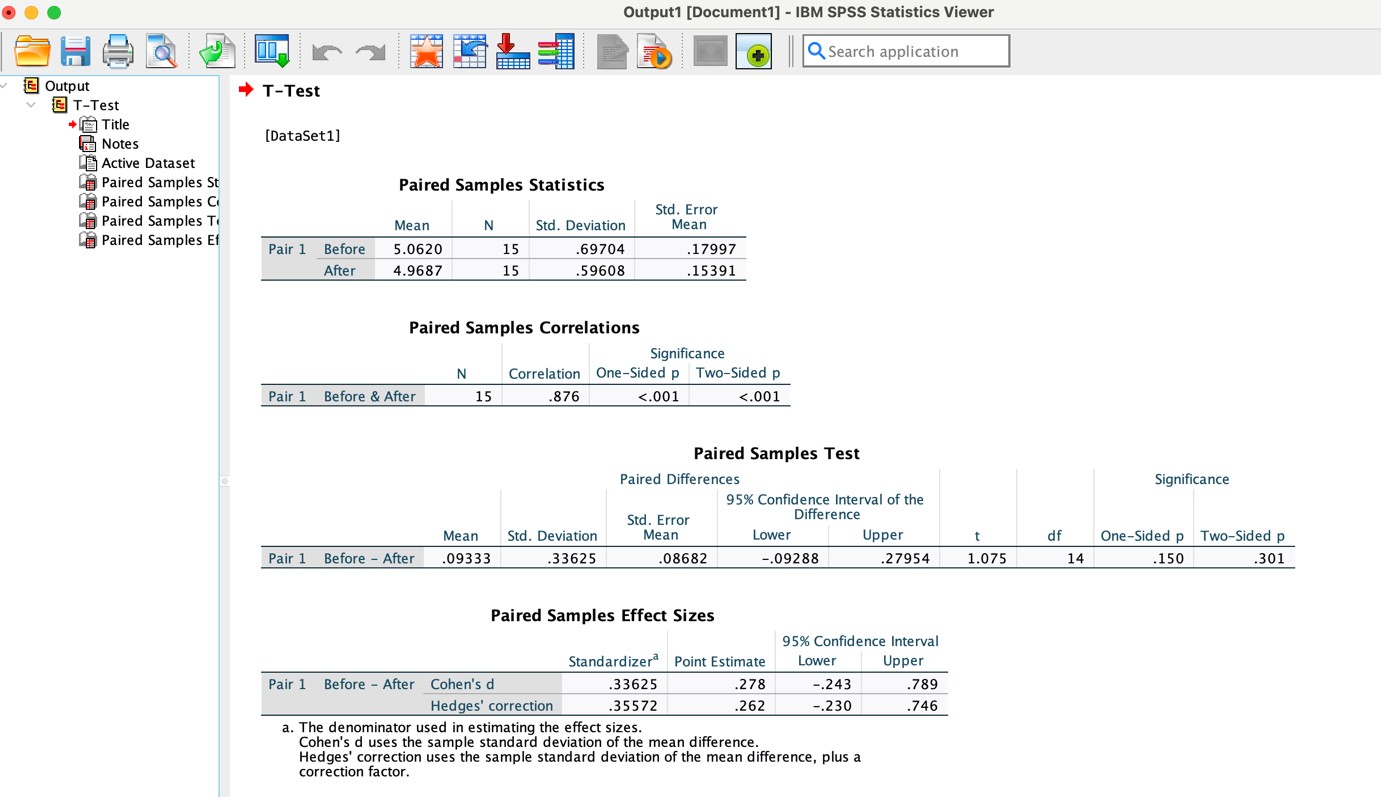
The data does not support the claim that the camp led to a significant change in the average speed of the players.

## Excel

Here I calculated the P value using excel and attached in below.



## SPSS



1. ***Python***

i). import scipy.stats as stats

ii). before\_camp = [5.88, 5.89, 3.41, 5.79, 4.92, 4.96, 4.95, 4.66, 5.56, 5.73, 5.05,

4.44, 4.27, 4.87, 5.55]

after\_camp = [5.76, 5.21, 3.35, 5.89, 4.78, 4.54, 4.79, 4.88, 5.08, 5.65, 5.1, 4.79,

4.78, 4.95, 4.98]

iii). t\_stat, p\_value = stats.ttest\_rel(before\_camp, after\_camp) iv). alpha = 0.05

1. t\_stat

output: 1.0750164544594143

1. p\_value

output: 0.30054603822799947

1. if p\_value < alpha:

print("There is significant evidence to say that the camp helped improve speed.") else:

print("There is not enough evidence to say that the camp helped improve speed.")

output: There is not enough evidence to say that the camp helped improve speed.

# Q2:

A teacher wants to test the effectiveness of a new textbook. She believes that this new textbook is easier to read, and that her students should have better grades on their tests this year than they have in the past. She took a random sample of test scores from last year’s classes, and then a random sample of test scores from this year’s classes. Assume normal populations for both years. Test her theory at α= 0.05.

Note: You can solve this question by any (SPSS, EXCEL or Manually) (3 Marks)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Old book | | | | | New book | | | | |
| 75 | 80 | 94 | 70 | 60 | 94 | 62 | 86 | 89 | 80 |
| 72 | 85 | 83 | 88 | 59 | 96 | 88 | 88 | 79 | 75 |
| 70 | 68 | 55 | 92 | 52 | 94 | 84 | 86 | 78 | 64 |

## 1). Manual

#### Write an appropriate hypothesis test for this situation and state the appropriate testing procedure.

Here the two samples are statistically independent and randomly selected from independent groups, we will use two sample t-test.

Xo= Mean test score from class using old text books *Xn*= Mean test score from class using old new books Hypothesis:

Null Hypothesis (H0): X0-Xn = 0, there is no difference between mean test scores from the last year using old and this year using new text books

Alternative Hypothesis (Ha): X0-Xn<0, the mean test scores of last year are lower than the mean scores from this year.

#### Compute the necessary summary statistics for the test in part (a).

|  |  |  |
| --- | --- | --- |
| **metrics** | **Formulas** | **Calculation s** |
| Number of samples with old books (No) | count(old book sample) | 15 |
| Number of samples with new books (Nn) | count(new book sample) | 15 |
| Mean with old books (Xo) | sum of players before time/n | 5.062 |
| Mean with new books (Xn) | sum of players after time/n | 4.969 |
| Standard deviation old books (So) | .=stdev(marks range with old books) | 13.346 |
| Standard deviation new books (Sn) | .=stdev(marks range with new books) | 10.112 |
| degree of freedom | No+Nn-2 | 28 |

1. ***Perform the t-test and report the p-value.***

**Formulas:**

t-Statistic = Xo-Xn/Sqrt of So^2/N0 + Sn^2/Nn

T Critical value: t statistic with degree of freedom (No+Nn-2)

P value: p value for the t-statistic

1. To calculate T Critical value, I used the below website: https://goodcalculators.com/student-t-value-calculator/

Where we need to enter Degree of freedom and Significance level value.

1. To calculate P value, I used the below website: https://[www.socscistatistics.com/pvalues/tdistribution.aspx](http://www.socscistatistics.com/pvalues/tdistribution.aspx)

Here we need to enter the t-statistic value, Degree of freedom and Significance level value.

In this website, we can compute p-value of one and two-tailed hypothesis.

### Calculations:

Significance value: 0.05

t-Statistic = 2.159

T Critical value: t statistic with degree of freedom (n-1) : 1.7011 P value: p value for the t-statistic:

a). P-value one-tile: 0.0198 b). P-value two-tile: 0.0396

#### Interpret your results in the conclusion

With a p-value of 0.0198, which is less than the chosen significance level of 0.05, we have enough evidence to reject the null hypothesis.

Therefore, based on this analysis there is sufficient evidence to conclude that the new text books had a statistically significant effect on improving the grades of the students.