Basic Microsoft Excel skills will be required for some of your Glenn College courses and may prove very helpful for others. This four-part tutorial addresses getting started with Excel, performing calculations and functions, conducting data analysis, and basic formatting.

If you are interested in watching a 10-minute video overview of basic Excel data creation and formatting, here is a YouTube video link: [https://www.youtube.com/watch?v=8L1OVkw2ZQ8](https://www.youtube.com/watch?v=8L1OVkw2ZQ8).

### Tutorial 3/4: Statistical Analysis in Excel

This document covers (click each link to skip to that section)

- [Installing Excel Add-Ins](#)
- [Creating a Summary Statistics Table](#)
- [Hypothesis Testing](#)

### Installing Excel Add-Ins

To create tables and conduct hypothesis testing, you will need the Data Analysis Toolpak Add-In. Here are Microsoft’s detailed instructions for installing this Add-In: [https://support.office.com/en-us/article/Load-the-Analysis-ToolPak-6a63e598-cd6d-42e3-9317-6b40ba1a66b4](https://support.office.com/en-us/article/Load-the-Analysis-ToolPak-6a63e598-cd6d-42e3-9317-6b40ba1a66b4)

Once the Add-In has been installed, you should be able to see it on the right hand side of the new Data tab.

In Excel, open the file with your data.

**For this tutorial use Tutorial_Blackbirds.xlsx.** This is synthetic data about the number of Red Winged Blackbirds spotted in different Metro Parks over 24 randomly selected days.
Creating a Summary Statistics Table

Click the **Data tab** on the “ribbon” of options. Click **Data Analysis** in the Analysis group.

Choose **Descriptive Statistics**. Click **OK**.

For the Input Range, use the mouse to select all of the data but not the headers ... A2:C25.

On the same Descriptive Statistics window, under Output options, click **Summary statistics** and **Confidence Level for Mean** (95%). Click **OK**.

A new worksheet, Sheet1, will appear with three tables of raw summary statistics. It will look like this:
Excel does not include the column headings from the data set. It is a good practice to copy/paste these in so that if time passes between creating the table and coming back to it, you will be able to connect the summary statistics with the correct variable.

In addition to the inconsistent decimals, the row labels in this table are in triplicate. This table needs major reformatting before it can be included in a policy brief, paper, or memo.

For information on formatting cells, rows, and columns, refer to Tutorial 4 of 4: Formatting included in this set of four tutorials.

Hypothesis Testing

The means for Inniswood, Three Creeks, and Slate Run are 20.83, 26.17, and 21.63 respectively. We would like to know if the difference in these means is statistically significant.

The null hypothesis and alternative hypotheses:

\[ H_0: \text{The mean number of red winged blackbirds at Inniswood} = \text{the mean number of red winged blackbirds at Three Creeks} \]

\[ H_a: \text{The mean number of red winged blackbirds at Inniswood and Three Creeks are not equal (they are different).} \]

To test this hypothesis in Excel return to the raw data worksheet.

On the Data tab, click Data Analysis in the Analysis Group.

Scroll to the bottom of the list and select the appropriate test. Click OK.
Variable 1 Range should be the data from first column (Inniswood) and Variable 2 Range should be the second column (Three Creeks). Click OK.
Another new worksheet appears with the results of the test. If you were using the Tutorial_Blackbirds.xlsx data, yours should look like this:

![Worksheet Image]

Look at the t Stat... can we reject the null hypothesis, that the mean number of red winged blackbirds at these two parks are the same? In this case, the t-statistic 3.199 > 1.96, so the hypothesis that the two means are equal is rejected.

Repeat this test for Inniswood and Slate Run.

Check that the table in your t-test output worksheet matches the one below:

![Worksheet Image]

In this case, the t Stat is 0.32 < 1.96. The null hypothesis of equal mean blackbirds between these two parks cannot be rejected.
Since the means between Inniswood and Slate Run are equal and the means between Inniswood and Three Creeks are not, it is not necessary to test the difference in means between Slate Run and Three Creeks... we know the means will be different! [Nevertheless, if you would like the practice, go head and conduct the test. The t Stat should be 2.165.]

For more information about conducting statistical analysis in Microsoft Excel, consult Microsoft’s support page, Use the Analysis ToolPak to perform complex data analysis.


Here is a video about calculating descriptive statistics in Excel (including installing the Excel Add-In):
https://www.youtube.com/watch?v=5MFjwM6K55g

   Note, this video specifies the descriptive statistics appear in the same worksheet as the raw data, rather than in a new, separate worksheet.