

# Activity 3.7 Statistical Analysis with Excel

## Introduction

Engineers use various tools to make their jobs easier. Spreadsheets can greatly improve the accuracy and efficiency of repetitive and common calculations; therefore, engineers often employ spreadsheet applications in their work.

In this activity you will collect data and use Microsoft Excel to perform statistical analyses and create a statistical chart to display your data.

## Equipment

- Engineering notebook
- Completed activity 3.5 Applied Statistics

## Procedure

**IMPORTANT:** This activity requires the use of the Analysis ToolPak, which is an Add-in to Microsoft Excel that is not available by default in the program. To load the Analysis ToolPak follow the instructions of your instructor or read **Analysis ToolPak Loading Instructions**.

**Part 1.** Perform a statistical analysis in Excel of height measurements of students in your class.

1. Open an Excel workbook. On worksheet 1 type "Activity 3.5 Statistical Analysis with Excel" in cell A1. Type your name in cell A2.
2. In your notebook record the height of each student in your class in feet and inches to the nearest *quarter of an inch*.
3. Input the raw data into an Excel worksheet using a separate column (A) for feet and a separate column (B) for inches. Include appropriate data (column) headers.
4. Use a formula to convert each height to decimal feet and place the results in column C and include an appropriate column heading.
5. Format the height measurements in decimal feet to show two decimal places.

|    | File  | Home   | Insert       | Page Layout | Formulas |
|----|---|--------|--------------|-------------|----------|
|    | G27   |        |              |             |          |
|    | A   | B      | C            | D           |          |
| 1  | Activity 1.3.6a Statistical Analysis with Excel |        |              |             |          |
| 2  | Student Name                                    |        |              |             |          |
| 3  | Height of Students                              |        |              |             |          |
| 4  | feet  | inches | decimal feet |             |          |
| 5  | 5   | 6      | 5.50         |             |          |
| 6  | 5   | 4      | 5.33         |             |          |
| 7  | 5   | 9      | 5.75         |             |          |
| 8  | 5   | 10     | 5.83         |             |          |
| 9  | 6   | 2      | 6.17         |             |          |
| 10 | 6   | 4      | 6.33         |             |          |
| 11 | 4   | 9      | 4.75         |             |          |
| 12 | 5   | 9      | 5.75         |             |          |
| 13 | 5   | 7      | 5.58         |             |          |
| 14 | 5   | 6      | 5.50         |             |          |
| 15 | 5   | 11     | 5.92         |             |          |
| 16 | 6   | 0      | 6.00         |             |          |

6. In the cell just below the column of heights in decimal feet, calculate the sum of the height measurements using the SUM function. Format the cell containing the sum to display a box around the number and add the text "Sum =" in the cell to the left of the sum cell. Note that the sum should display two decimal places.

7. Calculate the statistics indicated in the image to the right. Create the text labels in the appropriate cells. Be sure to calculate the **population** standard deviation (STDEV.P) as well as the appropriate Mode function (single or multimodal).

**OPTIONAL:** Use formulas to calculate the standard deviation of your height data.

If your data has more than one mode, use the MODE.MULT function. This function will create an answer in the form of an array. Before typing the function into a cell, highlight multiple cells (vertically), type in the function text, and select the range of values. Then depress **Ctrl/Shift/Enter** keys simultaneously to indicate that an array will be returned.

You can create a simple formula to calculate the range.

8. Create class intervals (value ranges) for a histogram at 0.250 feet intervals that will include **your** minimum and maximum recorded height.

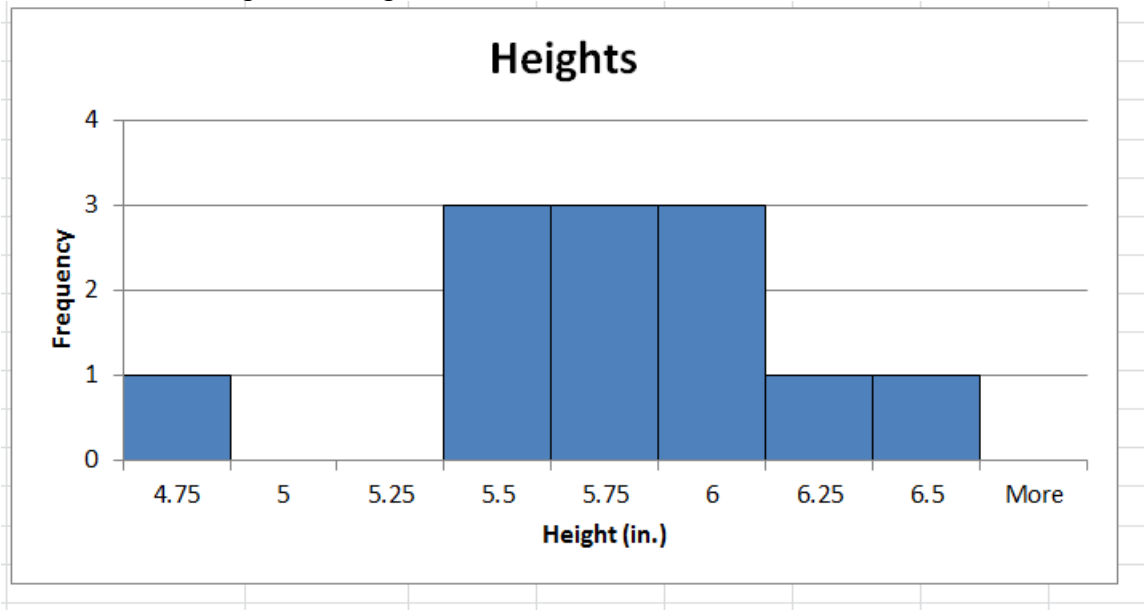
|                        |       |
|------------------------|-------|
|                        |       |
|                        |       |
| Mean                   | 5.701 |
| Mode                   | 5.500 |
|                        | 5.750 |
|                        |       |
| Standard Deviation (P) | 0.399 |
|                        |       |
| Minimum                | 4.750 |
|                        |       |
| Median                 | 5.750 |
|                        |       |
| Maximum                | 6.333 |
|                        |       |
| Range                  | 1.583 |

|       |  |
|-------|--|
| Bins  |  |
| 4.750 |  |
| 5.000 |  |
| 5.250 |  |
| 5.500 |  |
| 5.750 |  |
| 6.000 |  |
| 6.250 |  |
| 6.500 |  |

9. Create a frequency table using the Histogram tool choice in the Data Analysis Tool (Data tab, Analysis panel).

| <i>Bin</i> | <i>Frequency</i> |
|------------|------------------|
| 4.750      | 1                |
| 5.000      | 0                |
| 5.250      | 0                |
| 5.500      | 3                |
| 5.750      | 3                |
| 6.000      | 3                |
| 6.250      | 1                |
| 6.500      | 1                |
| More       | 0                |

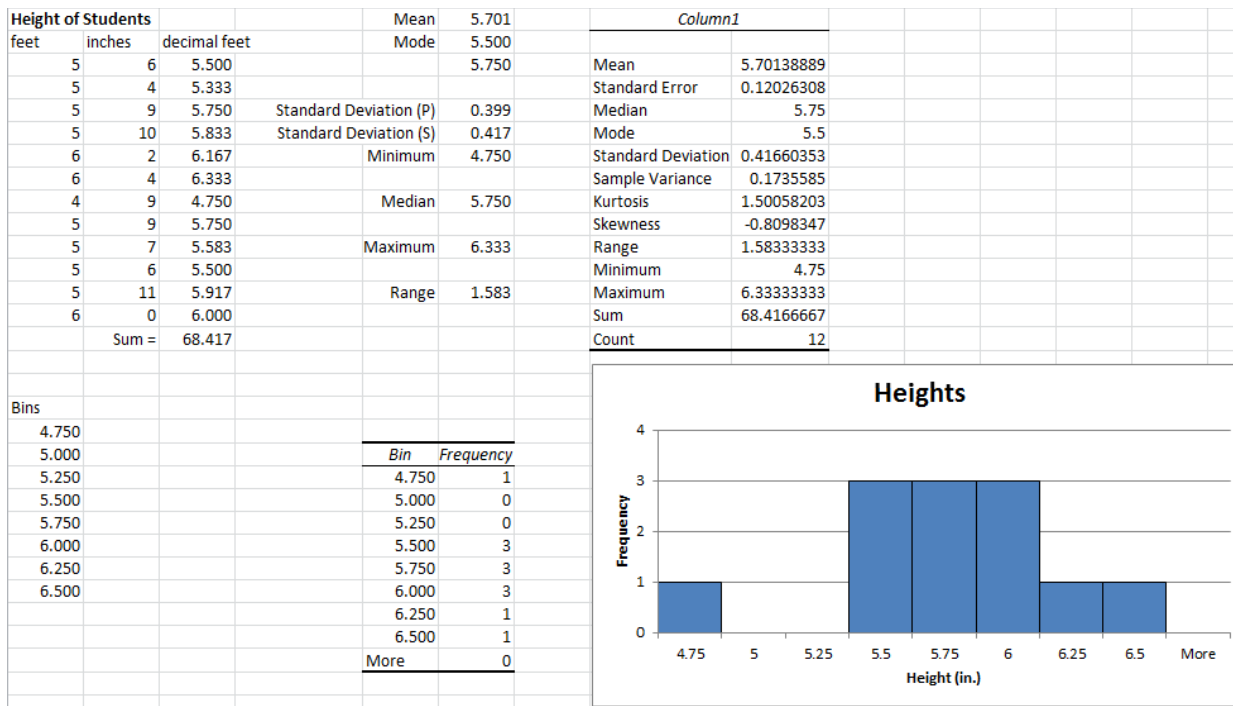
10. Create a histogram using the 2D chart tool. Format the chart as shown below.



11. Calculate the *sample* standard deviation using a function (STDEV.S) and display the output near the population standard deviation calculation.

12. Use the Data Analysis tool to calculate the Descriptive Statistics (Summary Statistics) and place the output data next to your calculated statistics.

13. Print your worksheet. Include all of the elements shown on the worksheet below.



14. Based on your histogram, does the data appear to be normally distributed? Explain your answer.

**Part 2.** Revisit Activity 3.5 Applied Statistics. Review the connector depth data (number 1) and the statistical analysis that you performed on the data. Then address each of the following.

1. Which value of mean did you use to indicate the central tendency of the data for the connector depth (population mean or sample mean) in Activity 3.4? Is there a difference in the numerical value between the population mean and the sample mean?
2. You used the population standard deviation to indicate the variation in the connector depth data. Which value of standard deviation (population or sample) should be used to estimate the variation of the connector depth among all connectors produced? Explain.
3. Calculate, by hand, the sample standard deviation of the connector depth data. You may use the data table that you create in Activity 3.4; however, show the sample standard deviation formula, your substitutions, and the result of your calculation to the appropriate number of decimal places.

- Use formulas or functions in Excel to find the mean, median, mode, population standard deviation, and the sample standard deviation of the connector depth data. Your worksheet should contain entries similar to those shown below.

Note that the data is multimodal, therefore you must use the MODE.MULT function in Excel. The MODE.MULT function will create an answer in the form of a vertical array (list of numbers). Before typing the function (MODE.MULT) into a cell, highlight multiple cells (vertically), type in the function, and select the range of data values. Then depress **Ctrl/Shift/Enter** keys simultaneously to enter the formula and to indicate that an array will be returned.

| Activity 3.6 Statistical Analysis with Excel |      |        |            |                          |
|--|------|--------|------------|--------------------------|
|  | X    | x-mean | (x-mean)^2 |                          |
|  | 3.8  |        |            | Mean =                   |
|  | 3.9  |        |            | Standard Deviation (P) = |
|  | 3.9  |        |            | Standard Deviation (S) = |
|  | 4.0  |        |            |                          |
|  | 4.0  |        |            |                          |
|  | 4.0  |        |            |                          |
|  | 4.1  |        |            |                          |
|  | 4.1  |        |            | Median =                 |
|  | 4.1  |        |            | Mode =                   |
|  | 4.2  |        |            |                          |
|  | 4.3  |        |            |                          |
|  | 4.4  |        |            |                          |
| Sum  | 48.8 |        | 0.0000     |                          |

- How do the statistics presented in Excel compare to the toy connector depth statistics that you calculated by hand?

- Create a histogram using the 2D chart tool. Use the following bins.

| Bins |
|------|
| 3.8  |
| 3.9  |
| 4    |
| 4.1  |
| 4.2  |
| 4.3  |
| 4.4  |

- Does the data appear to be normally distributed? Explain your answer.
- Use the Data Analysis tool to calculate the Descriptive Statistics (Summary Statistics).

9. Print the results of your statistical analysis. Include the Descriptive Statistics and histogram in a format similar to that shown in Part 1 number 13 above.

**Part 3 (Optional).** Revisit Activity 3.6 Instant Challenge: Fling Machine. Use Excel to verify your answers to Conclusion question number 1.

## Conclusion

1. Describe the difference between population standard deviation and sample standard deviation. Describe a scenario in which you would use each.
2. Which type of standard deviation (population or sample) is displayed when the Data Analysis Tool is used within Excel?
3. How do the statistics that you calculated using Excel compare to the statistics that you calculated by hand for the connector depth data?
4. Other than the calculation of statistics, how could an engineer use a spreadsheet application to increase efficiency?