# Activity 1.3: Using Excel to Visualize & Interpret Earthquake Data

## Data Visualization

Data visualization is a very important component of data analysis. Graphing data helps to reveal patterns and relationships that might not be obvious from simply reading through the raw data itself. At the same time, poorly constructed or intentionally misleading visualizations can mask these relationships or even lead the reader to draw false conclusions from the data. For these reasons it is important to construct visualizations carefully, choosing the r graph type and being mindful not to obscure or distort data.

## Learning Objectives

* Develop proficiency with basic Excel tools
* Learn how plate tectonics and earthquakes relate to each other
* Compare two different tectonic environments using descriptive statistics of earthquake data
* Quantify attributes and visualize earthquake data in order to evaluate whether data supports a proposed evidence-based conclusion

## Deliverables: Upload a .pdf document to the GitHub repository that includes responses to Questions 1-3, two plots for Task 3, and two screenshots from Google Earth.



**Task 1: Download Earthquake Data**

* + 1. Go to the USGS earthquake archive:<http://earthquake.usgs.gov/earthquakes/search/>
    2. Click on the orange box on the right-hand side entitled ‘Draw Rectangle on Map.’
    3. Zoom in to South America and click and drag to create a rectangle around the country of Chile including the Chile trench off the west coast of the continent (see picture). You can use the white circles on the edges of your rectangle to adjust as needed.
    4. Click ‘Use this Region’ to proceed (see arrow)
    5. Under ‘Magnitude’ select 4.5+ button to collect data from only moderate-to-large earthquakes. (*Why? Because small,*

*shallow earthquakes can be generated by a number of processes including minor adjustments in the crust and even human activities. By limiting our data to larger earthquakes, we are filtering out “noise” to better resolve the picture of larger tectonic forces that trigger moderate- to-large earthquakes)*

* + 1. Under ‘Date & Time,’ change the date and time parameters for the search to be between 2016-01-01 0:00 to 2016-12-31 23:59
    2. Click the box with a plus-sign next to ‘Output Options’
    3. Under ‘Format’, select ‘CSV’. Leave all other settings at their defaults.
    4. Click ‘Search’ at the bottom. Save your CSV file with a useful name.
    5. Without changing any of the settings, change output format to ‘KML’
    6. Under ‘KML-Specific Options’ select ‘Color by Depth’ instead of ‘Color by Age.’ Be sure to leave ‘Animated’ unselected. (This might be a feature you want to explore in the future.)
    7. Click ‘Search’ again. Save your KML file with a useful name.
    8. Repeat these steps for a divergent boundary.

**Checkpoint:** At this point, you should have **2 CSV files** and **2 KML files**, one each for a convergent and divergent boundary.

## Task 2: Descriptive Statistics

Now you have downloaded two CSV files of your earthquake data for the regions you selected. Keep in mind that a comma separated values (.csv) file differs from an Excel (.xls) file in the sense that CSV is much simpler format than an Excel file. Most of newer versions of Excel are well suited to reading CSV files and you may not even notice a difference. However, **after opening your CSV file in Excel you should ‘Save As’ and select ‘Excel Workbook’** so that you do not lose any of the functionality required for analyzing and visualizing the data properly.

## Question 1: What are the descriptive statistics for the depth and magnitude of the earthquakes both environments? (Complete the table below with your answers)

*Hint*: Refer to the basic functions you used in **Activity 1.2**. Find empty cells at the bottom or side of your worksheet to make your calculations. Because the number of rows in your spreadsheet is likely large it will be easier to specify the data range manually (e.g. by typing ‘D2:D501’) (if my spreadsheet had 501 rows) rather than trying to click and drag your data selection.

The table below will help you organize your calculations.

## Be sure to include units for your depths!

**Convergent Boundary Divergent Boundary Depth Magnitude Depth Magnitude**

**Maximum**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**Minimum**

**Mean**

**Median**

**Std. Dev.**

**Question 2**: **How do the basic descriptive statistics of the two datasets compare? How would you explain these differences with respect to the tectonic process operating at each location?**

**Checkpoint:** At this point, you should have used your **2 CSV files** to fill in the table and answer Question 2 above.

## Task 3: Visualize the data using Excel

Now it’s time to put your data to work by representing them graphically. Create an X-Y marked scatter plot of the relationship between earthquake longitude and depth.

1.Open your convergent boundary earthquake workbook. 2.Select the depth and longitude columns.

*Hint*: You can select entire columns by clicking on the column index letter. To select adjacent columns just click and drag across the column indices you wish to select. You can also hold ‘shift’ and click adjacent columns to select multiple columns.

**★Protip**: To select multiple *non*-adjacent columns, hold the ‘control’ (PC) or ‘command’ (Mac) keys while you click.

1. Using the ‘Insert > Chart’ dropdown menu, or tab, find the X-Y scatterplot option. A chart should appear on your spreadsheet.
2. Format your chart to allow readers to better understand it:
   1. Change the title to something more informative e.g. (Earthquake Depth vs. Longitude - Convergent Boundary)
   2. Add axis titles (‘Longitude’ on x axis, ‘Depth’ on y axis) and units when relevant. See link for instructions: [Add or Remove Titles in a Chart](https://support.office.com/en-us/article/add-or-remove-titles-in-a-chart-4cf3c009-1482-4908-922a-997c32ea8250?redirectSourcePath=%252fen-us%252farticle%252fAdd-axis-titles-to-a-chart-in-Office-2016-1bc8473d-e7d6-49e6-bb1c-3b10d0e4be3f&ui=en-US&rs=en-US&ad=US)

**★Protip**: A couple of other formatting changes will greatly improve the appearance of this plot:

* In general, numerical axes should start at zero - manipulating the range of an axis is a classic trick to distort the appearance of a graph and change its message. In this case however, because longitude is an arbitrary construct (distance from an arbitrary meridian on the Earth’s surface), we can change the axis range to avoid excess white space**.**
* Reversing the vertical axis would be logical in this case as the axis depicts depth below the Earth’s surface.
* If you would like to experiment with these changes the following link will get you started: [Change the scale of the vertical (value) axis in a chart](https://support.office.com/en-us/article/Change-the-scale-of-the-vertical-value-axis-in-a-chart-05973661-e56a-4486-a9f3-f9ce41df0021)

4.Repeat these steps to make a graph for your divergent boundary earthquake data.

## Save your plots, or copy and paste the plots into your assignment document.

**Checkpoint:** At this point, you should have completed the descriptive statistics table, answered Questions 1 and 2 above, and created and formatted two scatterplots and pasted them into your document.

## Task 4: Visualize the data using Google Earth

1. Open the KML for the convergent boundary you saved earlier (you may have to look in your downloads folder); the file will open in Google Earth. You might need to zoom in a bit before the data appear.
2. Now you can really see where these earthquakes lie on the globe. For some earthquakes you can click on it to glean even more information about the seismic event at the individual scale.
3. Pay attention to the pattern of earthquake depths and their position relative to the trench. Try to integrate this with the pattern in you scatter plot of longitude vs. depth to get a 3D picture of what is going on.
4. Do the same for your divergent boundary KML.

## Save a screenshot from Google Earth for each boundary type (convergent and divergent) and include these two images in your submission.

**Question 3: What conclusions can you draw from your visualizations?**

*Hint*: Think about the difference in the depth of earthquakes between the two environments and about what kind of tectonic activity is happening within and between both regions. You can also refer to your descriptive statistics as well, and as always: Don’t forget units!

**Checkpoint:** At this point, you should have completed the descriptive statistics table, answered Questions 1-3 above, created and formatted two scatterplots and pasted them into your document, and created and pasted two Google Earth Screenshots.