Chapter 3

# Applying the Concepts: Excel Instructions

## Interpret a Histogram

1. Open the countries.csv dataset in Excel.
2. Select the data.
   1. Find the column header for life\_expectancy in Column F.
   2. Select the data (rows 1 through 177 in Column F only).
3. Insert a histogram.
   1. Navigate to the Excel ribbon at the top, then go to the Insert tab.
   2. Select Insert Statistic Chart→Histogram.
4. Adjust the histogram (optional).
   1. Right click on the horizontal axis for bin size adjustment.
   2. Use Format Axis to modify bin width.
   3. Add axis titles via Chart Design→Add Chart Element→Axis Titles.

## Describe a Density Plot

Currently not available in Excel.

## Work with a Box Plot

1. Open the countries.csv dataset in Excel.
2. Select the data.
   1. Find the column header for mean\_years\_of\_schooling in Column G. Select all the data in that column.
3. Insert a box plot.
   1. Navigate to the Excel ribbon at the top, then go to the Insert tab.
   2. Select Insert Statistic Chart→Box and Whisker.

## Interpret a Violin Plot

Currently not available in Excel.

## Ponder a Word Cloud

Currently not available in Excel.

## Contemplate a Box Plot with a Categorical Variable

1. Open the countries.csv file in Excel.
2. Confirm that column K contains whr\_score and column D contains continent.
3. Filter continents to include only Africa, Asia, Europe and Oceania.
   1. Go to the 'Data' tab on the Excel ribbon at the top of the window.
   2. Select cells A1:P1
   3. Click on the 'Filter' button to enable dropdown arrows on the column header.
   4. Click the dropdown arrow in the continent column header in Column D.
   5. A list of all unique values in the column will appear.
   6. Uncheck the 'Select All' option to clear all selections.
   7. Scroll through the list and find "Africa", “Asia”, “Oceania” and “Europe” and check the boxes next to them.
   8. Click 'OK' or press 'Enter'.
4. Click on the column headers for both the whr\_score (column K) and continent (Column D) to select all the data in those columns.
5. Insert a box plot.
   1. With Columns K and D highlighted, go to the Insert tab on the Excel ribbon.
   2. Select Insert Statistic Chart→Box and Whisker.
6. Set box plot options.
   1. Right click on any box plot.
   2. Select Format Data Series…
   3. In the right-hand panel, be sure that all options are unchecked except “Show outlier points.”
   4. In the right-hand panel, choose “Inclusive median” to include the median in the calculation if the number of values in the data is odd.
7. Excel does not allow us to change the box plots to three different colors.

## Interpret a Violin Plot with a Categorical Variable

Currently not available in Excel.

## Interpret a Scatter Plot

1. Open the countries.csv dataset in Excel.
2. Confirm that column F contains life\_expectancy and column N contains gni\_per\_capita.
3. Insert a scatter plot.
   1. Go to the Insert tab on the Excel ribbon.
   2. Click on the Scatter (X,Y) icon in the Charts group.
   3. Choose the first scatter chart option, a simple scatter chart with no lines.
   4. Right-click the inside of the plotting area.
   5. Click “Select Data”
      1. Under “Legend Entries (Series)”, click “Edit”
      2. In Series name type: “Life Expectancy by GNI per Capita”
      3. In Series X values type:

=countries!$N$2:$N$177

* + 1. In Series Y values type:

=countries!$F$2:$F$177

* + 1. Click Ok
  1. Click Ok

1. Add a trendline to the scatter plot.
   1. To add a trendline, click the chart, select Chart Elements (plus sign button), and select Trendline.
2. Adjust the axes.
   1. Right-click on each axis and select Format Axis to adjust scale options, such as the minimum and maximum values, to better display your data.

## Explore a Line Chart

1. Open the lex\_long.csv dataset in Excel.
2. Filter the data.
   1. Highlight the country column (Column A) and year column (Column B).
   2. In the Excel ribbon, under the Home tab, navigate to Sort & Filter→Filter. A small arrow will appear on the right-hand side of the column header in Columns A and B.
   3. Filter to show only observations where country is USA.
      1. Click the arrow in the column header in Column A.
      2. Deselect the box next to Select All.
      3. Select the box next to USA. Click OK.
   4. Highlight the year column (Column B).
   5. Filter to show only observations where year is between 1880 and 2020.
      1. Click the arrow in the column header in Column B.
      2. Select Number Filters → Between…
      3. In the top box type 1880.
      4. In the bottom box type 2020.
3. Insert the line chart.
   1. Highlight the exact cells for life\_expectancy (Columns C1 to C55606).
   2. In the Excel ribbon, go to the Insert tab.
   3. Click the Line Chart icon under the group of Charts.
   4. Choose line chart.
4. Add year to the line chart.
   1. Right-click inside the line chart.
   2. Choose Select Data…
   3. Click the Add button
   4. Click OK
   5. Click Edit in the far right panel.
   6. Click the up arrow
   7. Select Cells B1 through B55606
   8. Click the down-arrow button to save.
   9. Click OK
   10. Click OK

## Describe a Bubble Chart

1. Open the countries.csv dataset in Excel.
2. Select the data.
   1. Select the mean\_years\_of\_schooling column (column G), the whr\_score column (Column K), and the population column (Column L).
3. Insert a bubble chart.
   1. Go to the Insert tab on the Excel ribbon.
   2. Click on the Insert Scatter (X, Y) or Bubble Chart drop-down menu.
   3. Choose Bubble.
4. Customize the bubble chart.
   1. Once the bubble chart is created, you may need to adjust the data series to ensure that:
      1. mean\_years\_of\_schooling is set as the X-axis (horizontal axis) values,
      2. whr\_score is set as the Y-axis (vertical axis) values,
      3. population is set as the bubble size.
   2. To do this, right click on the chart, choose Select Data, and then Edit the series to match the appropriate columns.
5. Add axis titles and legend.
   1. Click on the chart, then select Chart Elements (plus sign button) on the right side or top of Excel.
   2. Check the boxes for Axis Titles and Data Labels.
   3. Click on the default axis title text and enter your desired titles for both axes.
6. Format bubble sizes.
   1. Right click on one of the bubbles and select Format Data Series.
   2. Adjust the Scale Bubble Size to make the bubbles appropriate for your chart.
7. Create a bubble size legend.
   1. Bubble charts in Excel do not automatically create a legend for bubble size. You can create a manual legend by inserting shapes (circles) of different sizes next to the chart and labeling them with text boxes to represent different population values.

## Interpret a Heatmap

Currently not available in Excel.

## Interpret a Geographic Map

1. Open the countries.csv dataset in Excel.
2. Select the Data:
   1. Select the columns with the country names and the life\_expectancy values. Note that this will produce a map of all countries, not just those in North America.
3. Insert a Filled Map.
   1. Go to the Insert tab on the Excel ribbon.
   2. Click on Maps and choose Filled Map. If "Filled Map" is unavailable, your version of Excel might not support it, or you may need to enable this feature.
4. Adjust the map settings.
   1. Once the map is created, Excel will attempt to automatically recognize and match each country with its geographical location on the map.
   2. If Excel does not recognize some country names, it may prompt you to correct them.