

Introduction to R

Calibration Curve- Thermometer Example

- 1) Organize your data
 - a. Observations = rows
 - b. Variables = columns
- 2) Create an Excel file with the columns and rows you want. I named my file CalibrateTemp.xlsx. I include units in my variable names.

Point	TempC	mm
Freezing	0	20
Boiling	100	125
Human	37	57

- 3) Open R Studio (not R!)
- 4) Download the ggplot2 package (*This is similar to installing an app on your phone. It doesn't mean it's constantly running.*)
 - a. Go to the bottom right screen
 - b. Click Packages Tab
 - c. Select Install. Type "ggplot2". Then press install.
- 5) Start a new R Script or R markdown file.
 - a. File > New file > R Script
- 6) Activate ggplot2. (*This is similar to clicking on the logo of the app to make it run.*)
 - a. Go to the top left screen (your blank R script)
 - b. Type the code `library(ggplot2)`
 - c. Press the green arrow near the top (Run).
- 7) Import dataset from Excel
 - a. Go to the top right screen.
 - b. Click "Import Dataset"
 - c. Choose "From Excel"
 - d. "Browse" to identify the file.
- 8) Check that the data look correct.
 - a. Go to the top right screen for the file name, # of observations, and # of variables
 - b. Go to the top left screen for the data themselves.
- 9) Assign your file to an easy variable (Why? less typing, more reusable code, don't accidentally overwrite your data)
 - a. `mydata <- CalibrateTemp`
 - b. You should see 'mydata' appear in the top right screen under 'CalibrateTemp'.
- 10) Map the axes of your graph
 - a. `ggplot(data=mydata)`
 - b. You should see an empty graph appear in the bottom right under the "Plot" tab



- 11) Add the points.
 - a. `ggplot(data=mydata) +
 geom_point(mapping=aes(x=mm, y=TempC))`
 - b. You should see the scatter plot of three points at 20,0; 125,100; and 57,37.
- 12) Change the size of the points.
 - a. `ggplot(data=mydata) +
 geom_point(mapping=aes(x=mm, y=TempC), size=5)`
- 13) Add axis labels.
 - a. `ggplot(data=mydata) +
 geom_point(mapping=aes(x=mm, y=TempC), size=5) +
 labs(x= "mm of Ethanol", y = "Temperature (degrees Celsius)")`
- 14) Above the code that generates the graph, create a linear model (lm) and assign it to a variable. (Put the **y** axis first.)
 - a. `model <- lm(mydata$TempC ~ mydata$mm)`
- 15) Look at the summary of the model to identify the slope and y-intercept.
 - a. `summary(model)`
 - b. You should see Coefficients. Find -18.24 and 0.949. These are your y intercept and slope. Use these to add a line to your graph.
- 16) Add a line.
 - a. `ggplot(data=mydata) +
 geom_point(mapping=aes(x=mm, y=TempC), size=5) +
 labs(x= "mm of Ethanol", y = "Temperature (degrees Celsius)") +
 geom_abline(intercept= -18.24, slope=0.949)`
- 17) Use the Help tab to find out how to change the color of the dots.
 - a. Go to the bottom right screen.
 - b. Go to the Help tab.
 - c. In the search window, type "geom_point"
 - d. Use what you know to figure out how to make the points green.
- 18) Export plot to save your graph as an image or a .pdf or copy it to the clipboard to paste into another program.
 - a. In bottom right screen, go to Plots tab, Export pull down menu, and make your choice.
 - b. Paste your graph into your digital lab notebook.
- 19) Save your R script for future use of the code
- 20) Use the ggplot cheat sheet, make different types of graphs and layer one on top of another. (This is one of the big advantages to using R!)