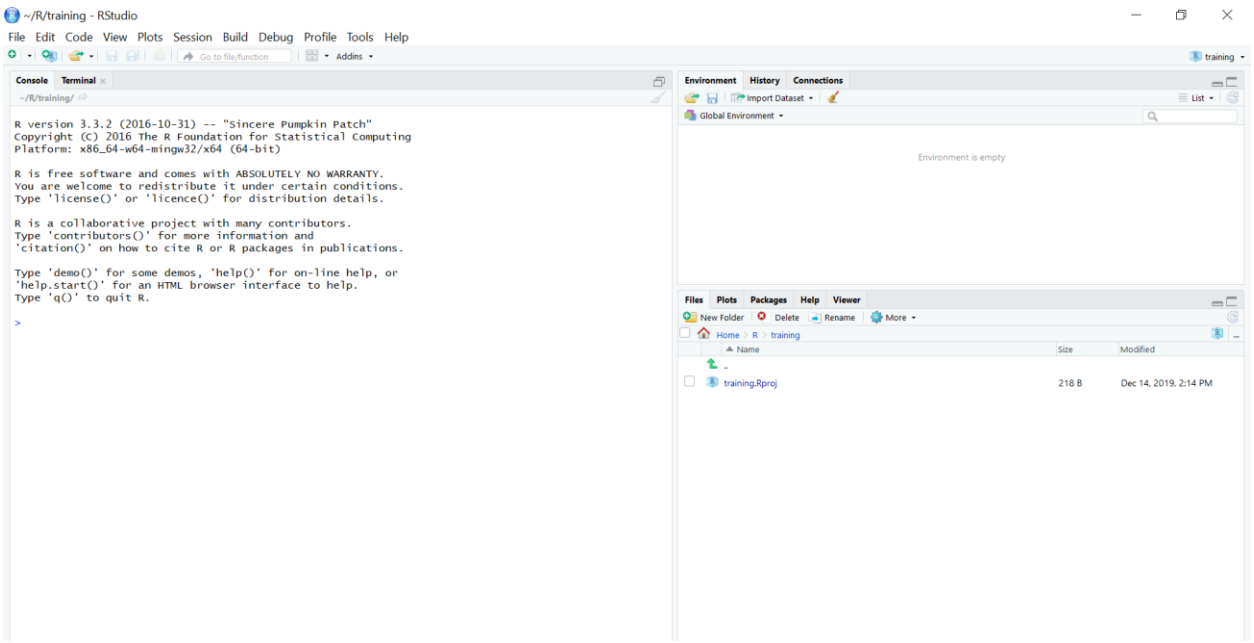
Mini-Tutorials for Creating Graphs in R
Normal Probability Plots

The examples below are intended to instruct you to create statistical graphs in R with minimal initial training in R. You should be able to follow the example codes to obtain graphs by modifying the included code. Some examples (and a key) will be included at the end of the document for practice. The screenshots I will show of the environment use R Studio, which is a free program you can find online. Other R environments will look different, but probably have similar functionality.

When you open up a new project environment in R Studio, it looks like this.



The command line environment is on the left. Images when we construct them will appear on the bottom right. As we add variables, they will appear in the list at the top right (name, dimensions and samples will display, which is useful for checking that you didn't skip entries when entering data by hand).

We are going to start by creating a normal probability plot graph from raw data. Later, we'll do an example with normal probability plots from residuals.

Copy the commands shown into the command line.

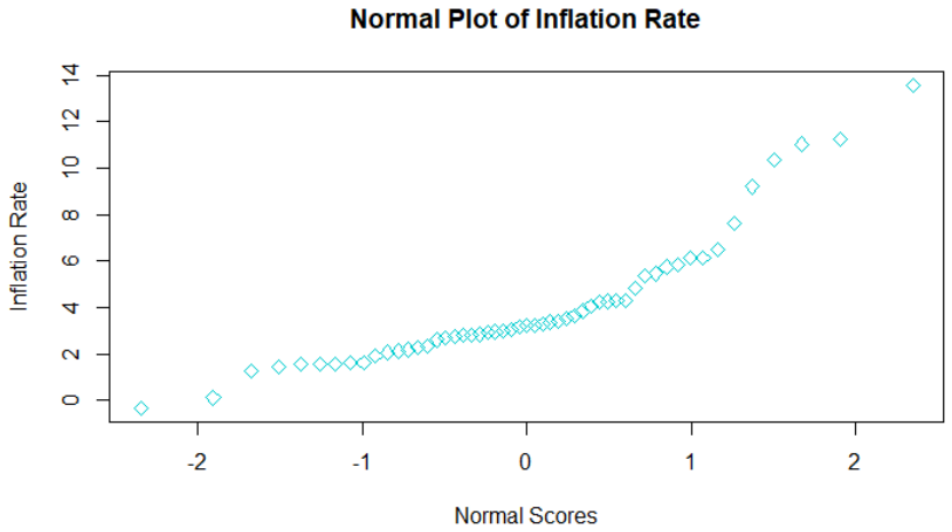
Step 1. Enter the data to be plotted in vector form.

```
InflationRate=c(1.59, 3.01, 2.78, 4.27, 5.46, 5.84, 4.30, 3.27, 6.16, 11.03, 9.20, 5.75, 6.50,
7.62, 11.22, 13.58, 10.35, 6.16, 3.22, 4.30, 3.55, 1.91, 3.66, 4.08, 4.83, 5.39, 4.25, 3.03, 2
```

.96, 2.61, 2.81, 2.93, 2.34, 1.55, 2.19, 3.38, 2.83, 1.59, 2.27, 2.68, 3.39, 3.24, 2.85, 3.85, -0.34, 1.64, 3.16, 2.07, 1.47, 1.62, 0.12, 1.26, 2.13)

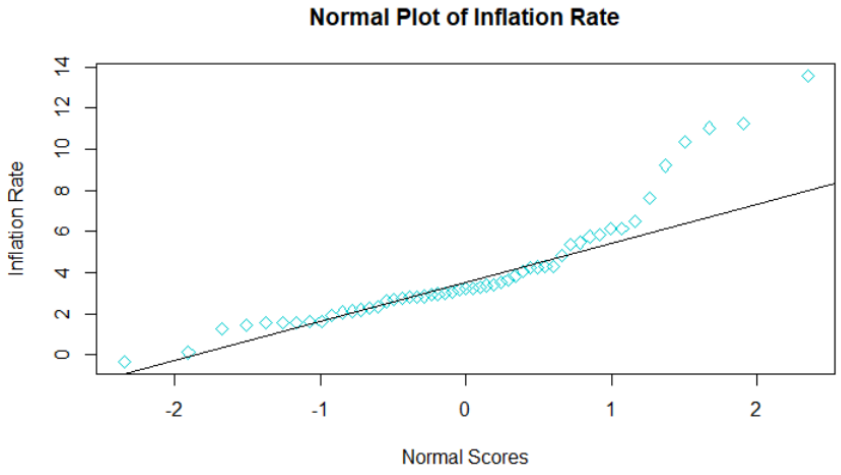
Step 2. Graph the normal plot.

```
qqnorm(InflationRate, ylab="Inflation Rate", xlab="Normal Scores", main="Normal Plot of Inflation Rate", col="darkturquoise", pch=5)
```



Then we need to add the line to assess whether the data is approximately normal or not.

```
qqline(InflationRate)
```



Most of this looks good, but the top end is problematic.

In the practice, we'll look at a normal probability plot of residuals from the wine/heart data (see the scatterplot tutorial for the regression details).

Practice.
This data was analyzed in the practice for the scatterplot data. We'll set up the initial steps quickly and then just look at the normal plot of the residuals.

```
nickel=c(5.9, 2.8, 5.3, 4.6, 4.6, 3.5, 3.8, 5, 5.3, 4.2)
tens_streng=c(948, 859, 921, 909, 915, 876, 828, 964, 964, 900)
```

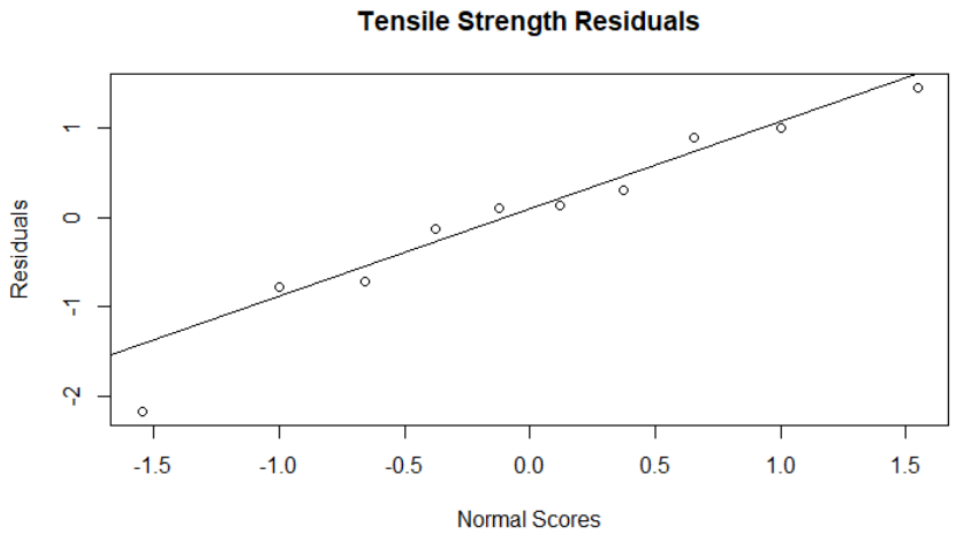
Solutions.
Start with the regression line, and then calculate the residuals.

```
strength.lm = lm(tens_streng ~ nickel)

strength.stdres = rstandard(strength.lm)

qqnorm(strength.stdres, ylab="Residuals", xlab="Normal Scores", main="Tensile Strength Residuals")

qqline(strength.stdres)
```



You can see that these residuals stick much more closely to the line, and so suggest the data is distributed normally.