

1. Download and install [R](#) and then [RStudio](#). Then install some packages in R, using the following code.

```
install.packages(c("tidyverse", "ape")) # copy and paste this line
```

Here's an example of loading an R package,

```
suppressMessages(library(tidyverse))
```

Ensure these two R packages are properly installed by loading each, very little should happen, but mainly no errors should occur.

2. Replicate our Lil' Analys with your own variable or, at max 2, variables. Hence, you should have at least

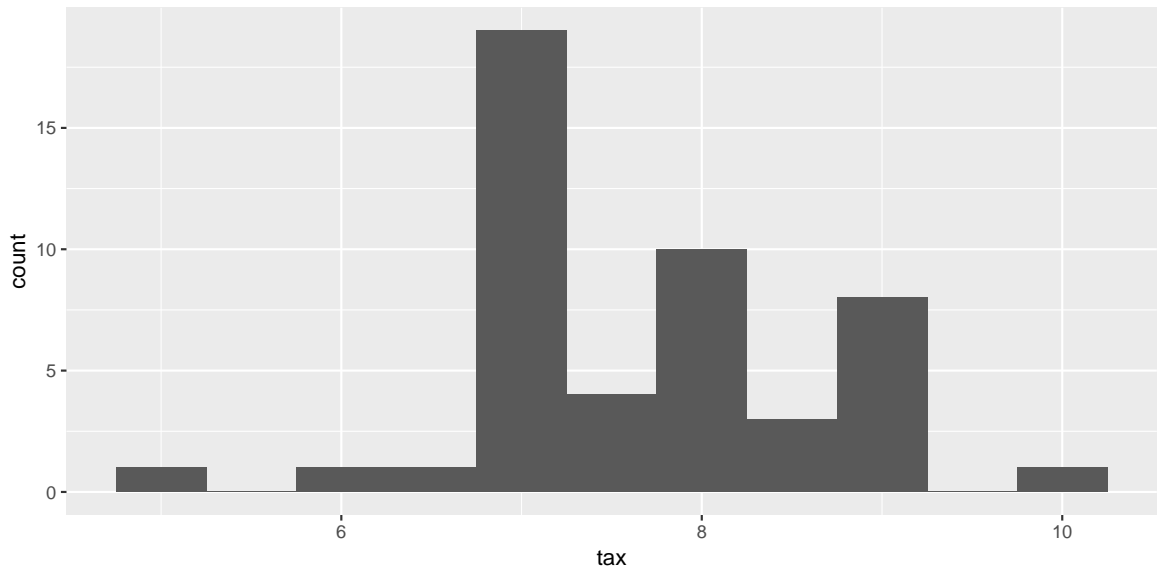
I'll use the dataset named 'petrol' found at [<https://roualdes.us/data/petrol.csv>](petrol). This is a dataset about petrol consumption by state in the year 1980. Petrol consumption is the response variable, while petrol tax, 'tax', is the explanatory variable.

```
df <- read.csv("https://roualdes.us/data/petrol.csv")
df %>% # see http://seananderson.ca/2014/09/13/dplyr-intro.html
  summarise(mean(tax), mean(consumption),
            sd(tax), sd(consumption),
            median(tax), median(consumption),
            IQR(tax), IQR(consumption))

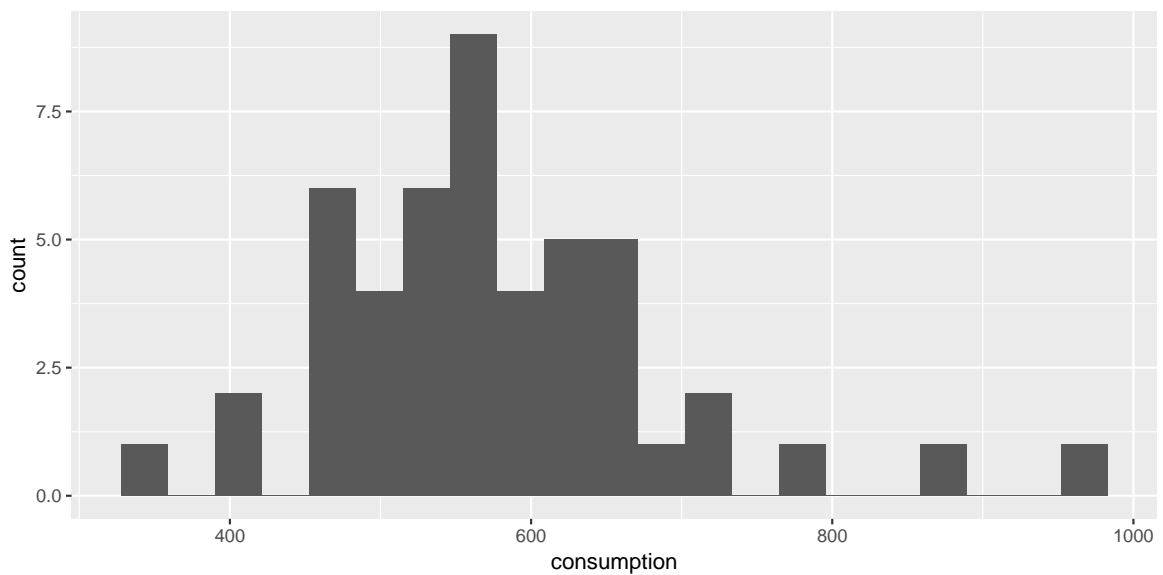
##   mean(tax) mean(consumption)   sd(tax) sd(consumption)
## 1  7.668333          576.7708 0.9507698      111.8858
##   median(tax) median(consumption) IQR(tax) IQR(consumption)
## 1           7.5           568.5      1.125      123.25
```

By the calculations of mean and median above, we see that petrol taxes are applied across the states fairly symmetrically (relatively little or no skew). However, consumption clearly differs by state. This is likely because states vary in population so much. Histograms confirm our analysis of skew.

```
ggplot(df, aes(tax)) + geom_histogram(bins=11)
```



```
ggplot(df, aes(consumption)) + geom_histogram(bins=21)
```



To show how petrol tax and consumption are related, we can use a scatterplot. Because we expect taxes to, at least partially, explain consumption, we put 'tax' on the x-axis. As expected, as the petrol tax goes up, consumption appears to go down.

```
ggplot(df, aes(tax, consumption)) + geom_point()
```

