Intro to R

```
a_thing <- 4
another_thing <- 1
another_Thing <- 7

both_things <- a_thing + another_thing
```

Then we created a tiny data set.

```
a_data_thing <- data.frame(x = 2, y = 8)

a_data_thing$x
```

```
## [1] 2
```

How would we print the variable y? Create a new chunk below and type your answer in it.

Write notes for yourself in the white space. Maybe explain to your future self what dollar signs do.

ERASE THIS AND TYPE SOME NOTES HERE

Enough playing around, let's load some data!

```
acitelli <- read.csv("acitelli.csv")
```

Notice that there is now another data object in the top right "environment" pane. If you click on the name of the dataset you can actually look at it. Importantly, you **cannot** change any data, this is by design. We want this behavior but it's hard to get used to!

Next, you want to look at your data. Hint: you can run a single line of code within a chunk with the keyboard shortcut: `ctrl + enter`.

```
head(acitelli)
```

```
##   cuplid  Yearsmar gender self_pos other_pos satisfaction tension simhob
## 1      3  8.202667     -1      4.8       4.6     4.000000     1.5      0
## 2      3  8.202667      1      3.8       4.0     3.666667     2.5      1
## 3     10 10.452667     -1      4.6       3.8     3.166667     4.0      0
## 4     10 10.452667      1      4.2       4.0     3.666667     2.0      0
## 5     11 -8.297333     -1      5.0       4.4     3.833333     2.5      0
## 6     11 -8.297333      1      4.2       4.8     3.833333     2.5      0
```

```
str(acitelli)
```

```
## 'data.frame':    296 obs. of  8 variables:
##  $ cuplid     : int  3 3 10 10 11 11 17 17 21 21 ...
##  $ Yearsmar   : num  8.2 8.2 10.5 10.5 -8.3 ...
```

```
##  $ gender      : int  -1 1 -1 1 -1 1 -1 1 -1 1 ...
##  $ self_pos    : num  4.8 3.8 4.6 4.2 5 4.2 4 4 4.2 4.4 ...
##  $ other_pos   : num  4.6 4 3.8 4 4.4 4.8 3.6 4.4 3.8 4.8 ...
##  $ satisfaction: num  4 3.67 3.17 3.67 3.83 ...
##  $ tension     : num  1.5 2.5 4 2 2.5 2.5 3 2 3.5 2.5 ...
##  $ simhob      : int  0 1 0 0 0 0 -1 0 0 0 ...
```

```
names(acitelli)
```

```
## [1] "cuplid"       "Yearsmar"     "gender"       "self_pos"
## [5] "other_pos"    "satisfaction" "tension"      "simhob"
```

You probably also want descriptive statistics.

```
summary(acitelli)
```

```
##      cuplid         Yearsmar              gender      self_pos
##  Min.   :  3.0   Min.   :-11.214000   Min.   :-1   Min.   :2.600
##  1st Qu.:165.2   1st Qu.: -7.089000   1st Qu.:-1   1st Qu.:4.000
##  Median :313.5   Median : -1.089000   Median : 0   Median :4.200
##  Mean   :282.6   Mean   : -0.000036   Mean   : 0   Mean   :4.186
##  3rd Qu.:401.2   3rd Qu.:  6.077667   3rd Qu.: 1   3rd Qu.:4.400
##  Max.   :485.0   Max.   : 15.036000   Max.   : 1   Max.   :5.000
##    other_pos      satisfaction      tension          simhob
##  Min.   :2.600   Min.   :1.167   Min.   :1.000   Min.   :-1.0000
##  1st Qu.:4.000   1st Qu.:3.333   1st Qu.:2.000   1st Qu.: 0.0000
##  Median :4.200   Median :3.833   Median :2.500   Median : 0.0000
##  Mean   :4.264   Mean   :3.605   Mean   :2.431   Mean   : 0.0777
##  3rd Qu.:4.600   3rd Qu.:4.000   3rd Qu.:3.000   3rd Qu.: 0.2500
##  Max.   :5.000   Max.   :4.000   Max.   :4.000   Max.   : 1.0000
```

The `summary()` function is smart, it will give five-number summaries for numerical variables and counts for categorical vairables, called factors in R. We can use the `as.factor()` function to temporarily change gender from an integer to a factor. This will come in handy later

```
summary(acitelli$gender)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      -1      -1       0       0       1       1
```

```
#the summary function gives counts for factor type variables
summary(as.factor(acitelli$gender))
```

```
##  -1   1
## 148 148
```

We can also select pieces of a data frame. That first number is the row, the second is the column.

```
acitelli[2, 6]
```

```
## [1] 3.666667
```

```
#You try it! Find a numder you want to pull from the dataset.
#acitelli[ ?, ?]
```

If it is instead a single variable, you can also select a piece.

```
acitelli$satisfaction[2]
```

```
## [1] 3.666667
```

In the chunk below, pick out the gender of the person in the 50th case.

```
#try it by referring to the row and column of the data frame.

#try it by referring to the variable, using the dollar sign notation.
```