Class 1 Homework

Load the following libraries:

```
library(tidyverse)
library(kableExtra)
library(gcookbook)
```

Get heightweight data *(built in with the gcookbook package)*, and put it in a dataframe *(the gcookbook:: syntax tell r which package the function or data you want is coming from. This normally is not necessary, unless packages have functions or data with the same name)*:

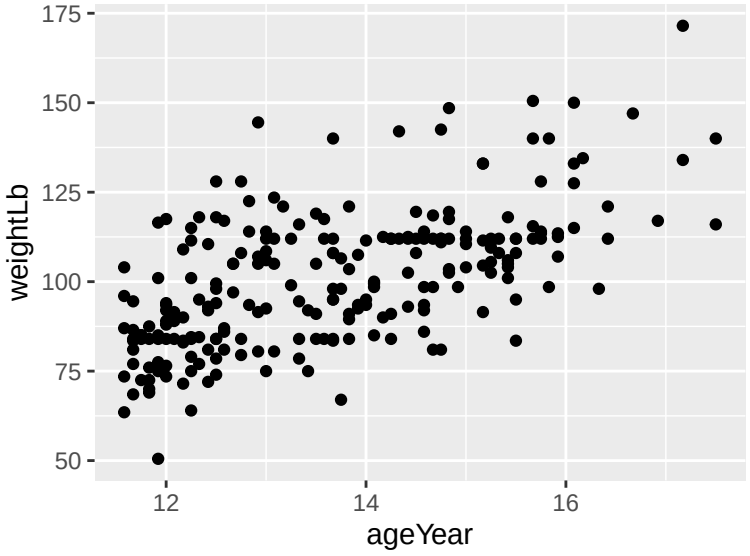
```
dfHW <- heightweight

head(dfHW)
```

```
  sex ageYear ageMonth heightIn weightLb
1   f   11.92      143     56.3     85.0
2   f   12.92      155     62.3    105.0
3   f   12.75      153     63.3    108.0
4   f   13.42      161     59.0     92.0
5   f   15.92      191     62.5    112.5
6   f   14.25      171     62.5    112.0
```

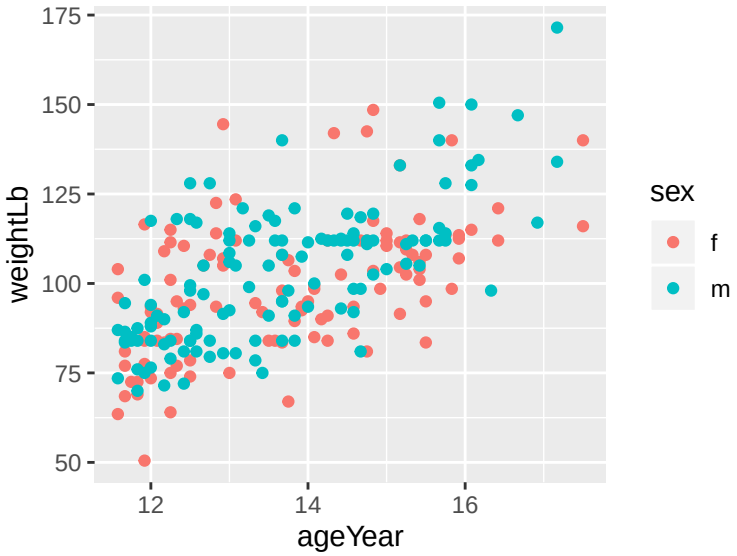
Let's take a quick look at the distribution of female and male observations. Use count to find the totals:

```
# A tibble: 2 x 2
  sex       n
  <fct> <int>
1 f       111
2 m       125
```

Now, Create a scatter plot of Weight vs Age in Years:



Change the scatter plot to show sex in different colors:

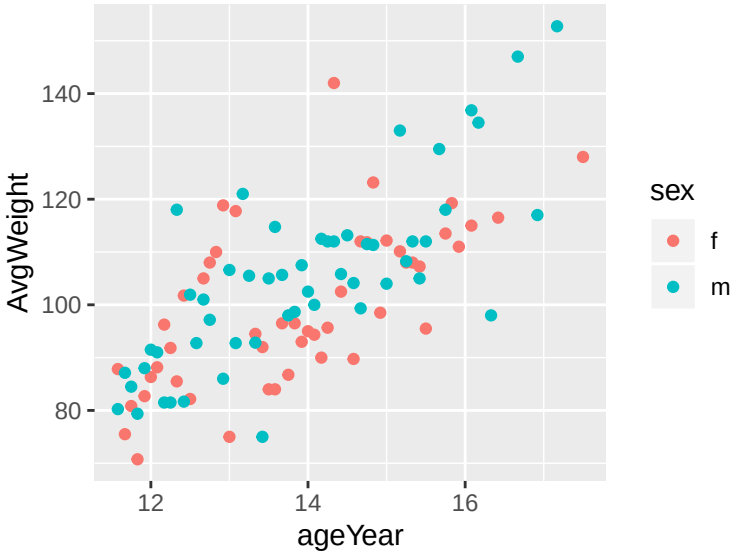


What does that tell us? Is that a little hard to see?

Let's make it easier to see. Create a new dataframe, and group the data by sex and ageYear, Then summarise the means of heightIn. It should look like the following:

```
# A tibble: 102 x 3
# Groups:   sex [2]
   sex   ageYear AvgWeight
   <fct>   <dbl>     <dbl>
 1 f        11.6      87.8
 2 f        11.7      75.5
 3 f        11.8      80.8
 4 f        11.8      70.8
 5 f        11.9      82.7
 6 f        12        86.3
 7 f        12.1      88.2
 8 f        12.2      96.2
 9 f        12.2      91.8
10 f        12.3      85.5
# ... with 92 more rows
```

Now plot that and see what it looks like. And add a geom_smooth regression line to see if we can pick up a trend.



A little easier to see tendency. Do you think there's a tedency? Add a geom_smooth() *(use se = F)*