Abalones - Answer Key for Homework 1

Melinda Higgins - Instructor for N741 Spring 2020

Feb 7, 2020

A Glimpse of the Abalone Dataset

Use the `glimpse()` function from the `tibble` package to take a peak at the **abalone** dataset.

```
## Observations: 4,177
## Variables: 10
## $ id            <dbl> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16...
## $ sex           <chr> "M", "M", "F", "M", "I", "I", "F", "F", "M", "F", "F"...
## $ length        <dbl> 0.455, 0.350, 0.530, 0.440, 0.330, 0.425, 0.530, 0.54...
## $ diameter      <dbl> 0.365, 0.265, 0.420, 0.365, 0.255, 0.300, 0.415, 0.42...
## $ height        <dbl> 0.095, 0.090, 0.135, 0.125, 0.080, 0.095, 0.150, 0.12...
## $ wholeWeight   <dbl> 0.5140, 0.2255, 0.6770, 0.5160, 0.2050, 0.3515, 0.777...
## $ shuckedWeight <dbl> 0.2245, 0.0995, 0.2565, 0.2155, 0.0895, 0.1410, 0.237...
## $ visceraWeight <dbl> 0.1010, 0.0485, 0.1415, 0.1140, 0.0395, 0.0775, 0.141...
## $ shellWeight   <dbl> 0.150, 0.070, 0.210, 0.155, 0.055, 0.120, 0.330, 0.26...
## $ rings         <dbl> 15, 7, 9, 10, 7, 8, 20, 16, 9, 19, 14, 10, 11, 10, 10...
```

Summary of abalone dimensions

1. Select three variables from the **abalone** dataset: *length*, *diameter*, and *height*.
2. Run the `summary()` function to get some basic descriptive statistics.

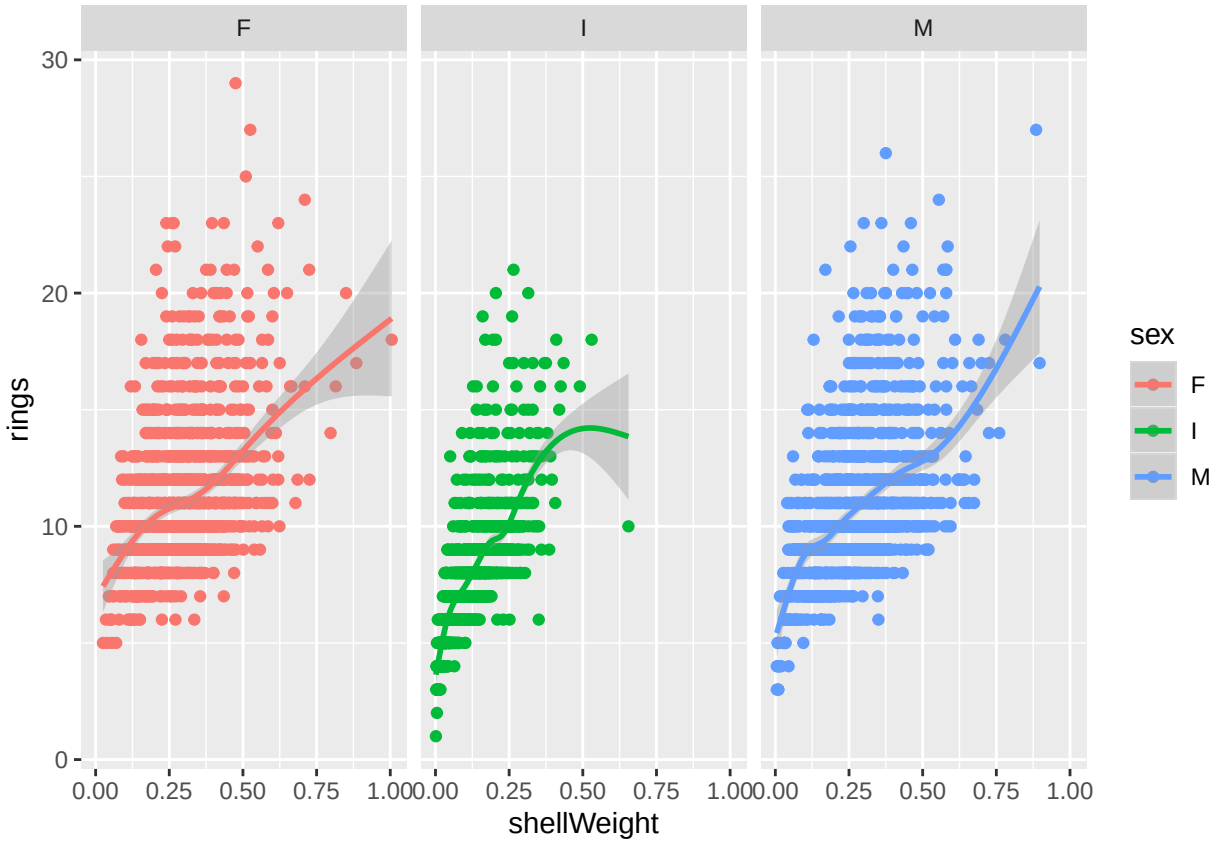
```
##      length         diameter          height      
##  Min.   :0.075   Min.   :0.0550   Min.   :0.0000  
##  1st Qu.:0.450   1st Qu.:0.3500   1st Qu.:0.1150  
##  Median :0.545   Median :0.4250   Median :0.1400  
##  Mean   :0.524   Mean   :0.4079   Mean   :0.1395  
##  3rd Qu.:0.615   3rd Qu.:0.4800   3rd Qu.:0.1650  
##  Max.   :0.815   Max.   :0.6500   Max.   :1.1300
```

Plot of number of rings by shell weight

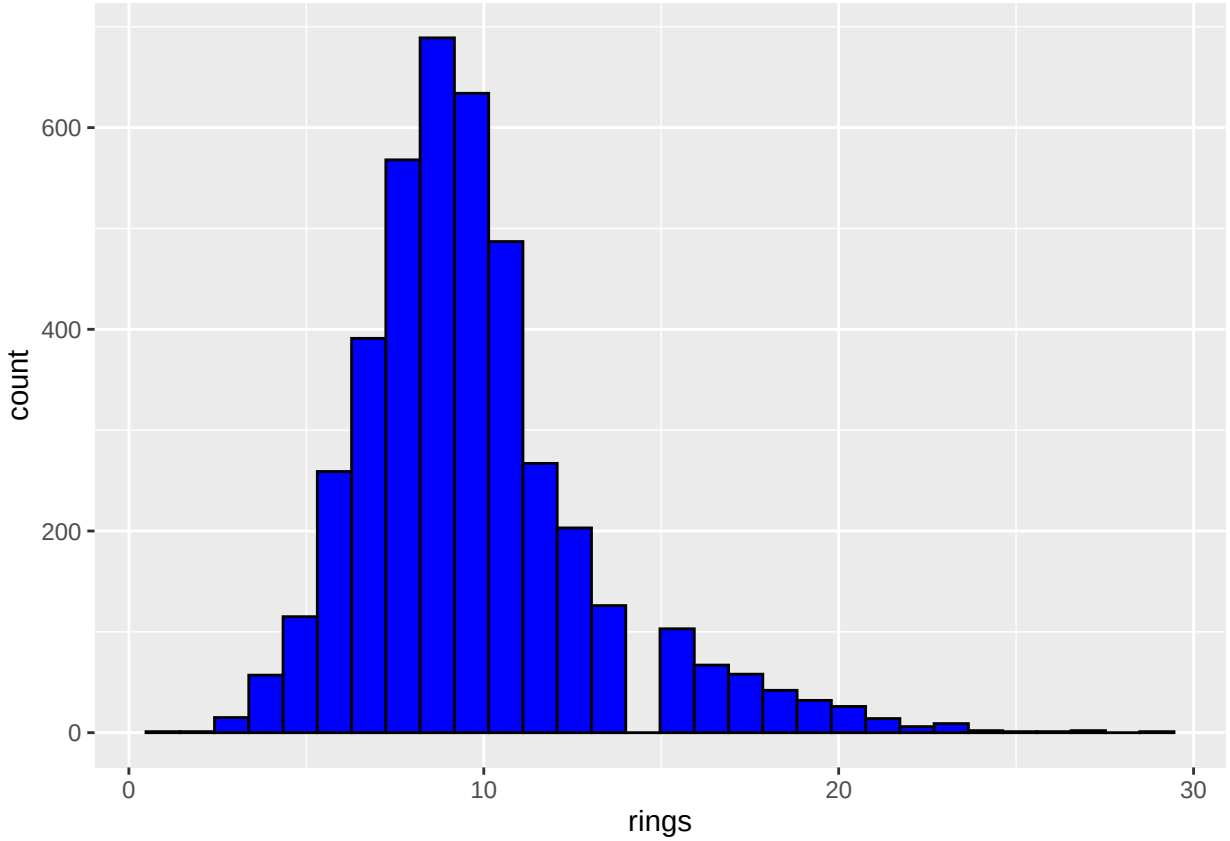
Make a scatterplot of the number of rings (on y-axis) by shell weight (on x-axis) and put plots in 3 separate panels by sex (i.e., male, female and infant) and add a smoothed fit line to each plot.

Note: The number of rings range from 1 to 29.

Also the mean of the abalone heights is `mean(abalone$height, na.rm=TRUE)` and the standard deviaiton of heights is `sd(abalone$height, na.rm=TRUE)`.


F
I
M
30
20
10
0
rings
0.00
0.25
0.50
0.75
1.00
shellWeight
sex
F
I
M

Histogram of abalone rings



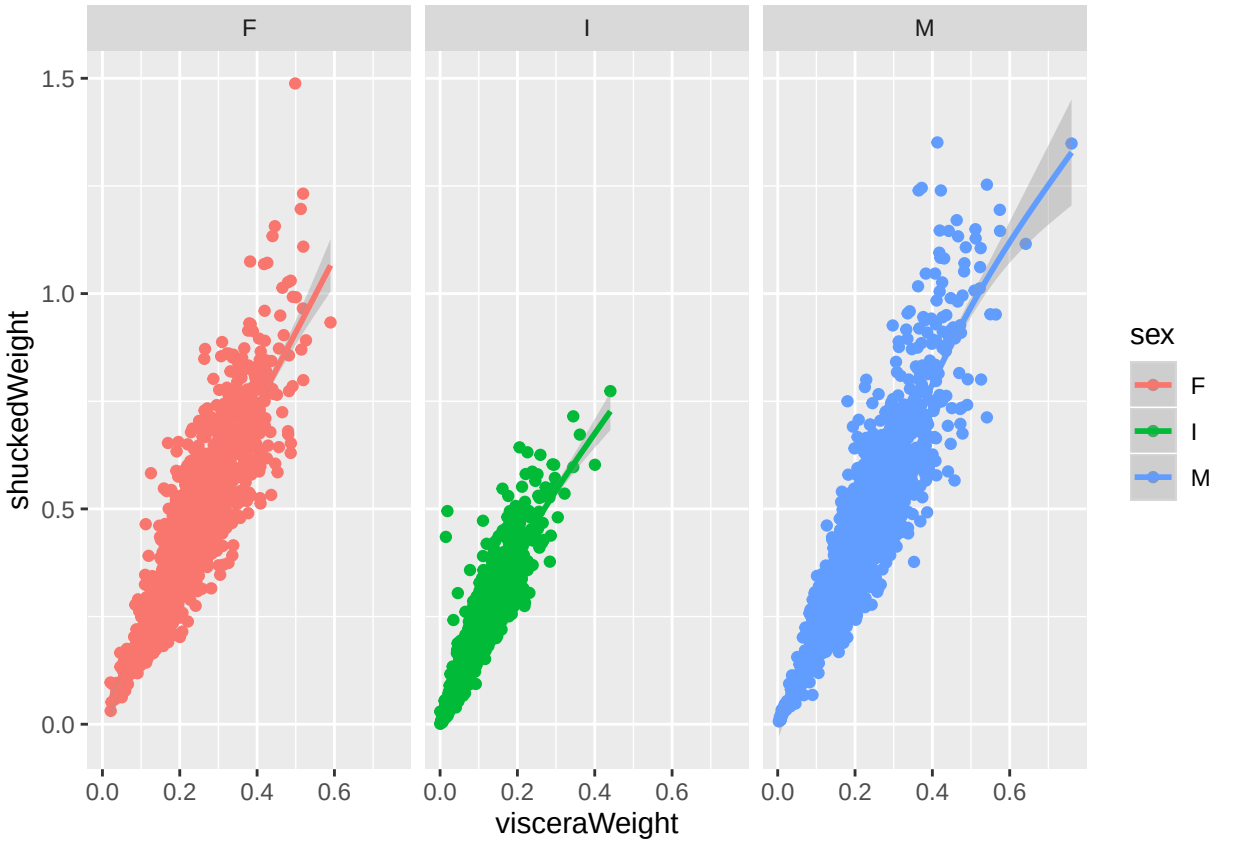
The distribution looks close to normal with slightly more peaked in the middle (positive kurtosis) and with some slight skewness to the right (longer tail to the right)> There is also an odd skipped histogram bar around rings=15, This may be due to the default bin width.

Regression of abalone rings by shell weight

We can save regression model results, then put the model coefficients from the `summary()` function into a nice table using the `kable()` function from the `knitr` package.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	6.462117	0.0771464	83.76431	0
shellWeight	14.535675	0.2790823	52.08382	0

Scatterplot of shucked weight (y-axis) by viscera weight (x-axis)



There appears to be strong correlation between viscera weight and shucked weight. The slopes of the lines appear to be similar between the 3 sexes.