Homework 5

For this homework you will create an R Markdown file and output (HTML file) and upload both to wolfware. Be sure to include text explaining your thought process/what you are doing with your questions.

The purpose of this homework is to get practice with summarizing data.

Data

We'll use the horseshoe crab data set (available in the homework link).

About the data:

- 173 mating female crabs
- y: whether the female crab has a "satellite" — male crab that group around the female and may fertilize her eggs
- satell: number of satellites
- color: female crab's color (2 = "light", 3 = "medium", 4 = "dark", and 5 = "darker")
- spine: spine condition (1 = "both good", 2 = "one worn or broken", and 3 = "both worn or broken")
- weight: female crab weight (g)
- width: female carapace width (cm)

Notes:

- Multiple delimiters between values are present
- Convert the three variables used in the plots below to factors after reading in the data (this also gives a convenient way to rename their values using `levels`)
- You may get an extra column of `NULL`s, just remove that

Tasks

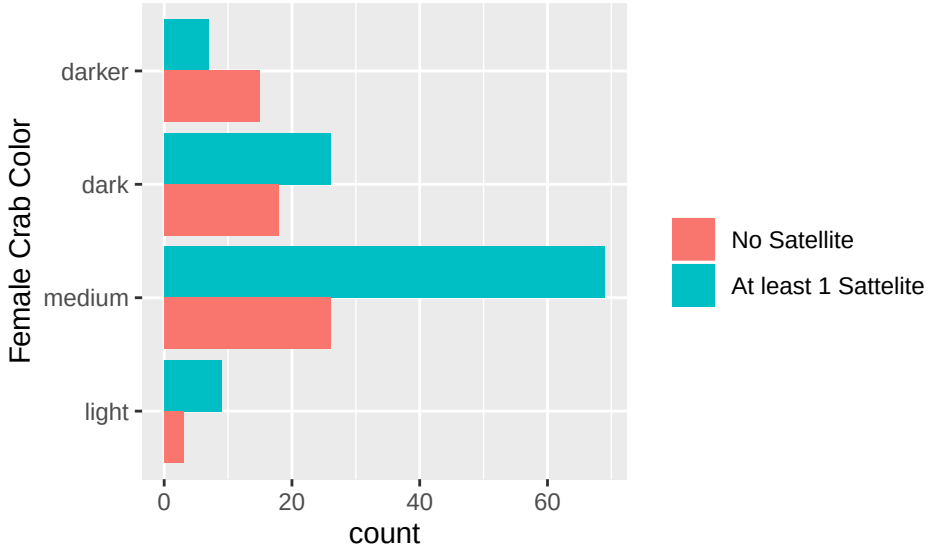
1. Read in the data and modify the variables as mentioned in the notes. Print the updated data object out.

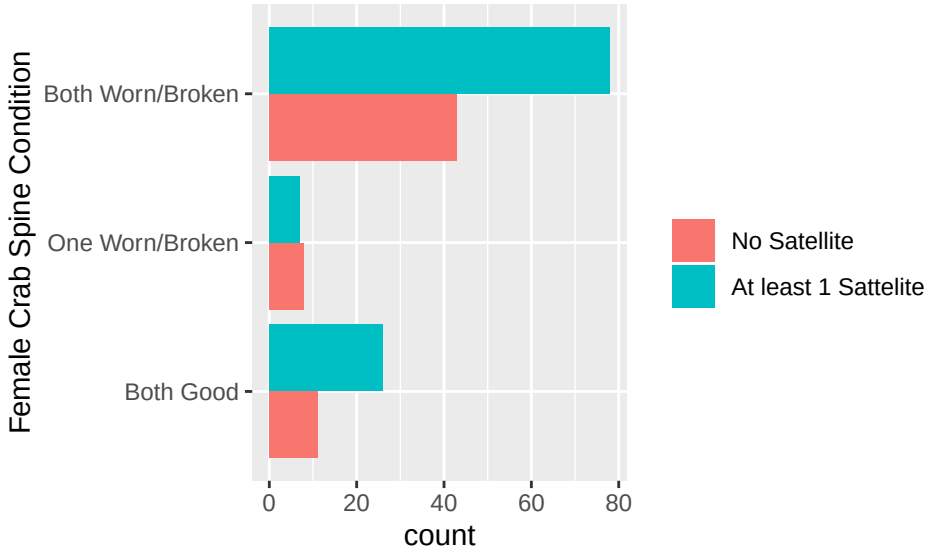
```
## # A tibble: 173 x 6
##    color  spine            width satell weight y
##    <fct>  <fct>            <dbl>  <dbl>  <dbl> <fct>
##  1 medium Both Worn/Broken  28.3      8   3050 At least 1 Sattelite
##  2 dark   Both Worn/Broken  22.5      0   1550 No Satellite
##  3 light  Both Good         26        9   2300 At least 1 Sattelite
##  4 dark   Both Worn/Broken  24.8      0   2100 No Satellite
##  5 dark   Both Worn/Broken  26        4   2600 At least 1 Sattelite
##  6 medium Both Worn/Broken  23.8      0   2100 No Satellite
##  7 light  Both Good         26.5      0   2350 No Satellite
##  8 dark   One Worn/Broken   24.7      0   1900 No Satellite
##  9 medium Both Good         23.7      0   1950 No Satellite
## 10 dark   Both Worn/Broken  25.6      0   2150 No Satellite
## # ... with 163 more rows
```

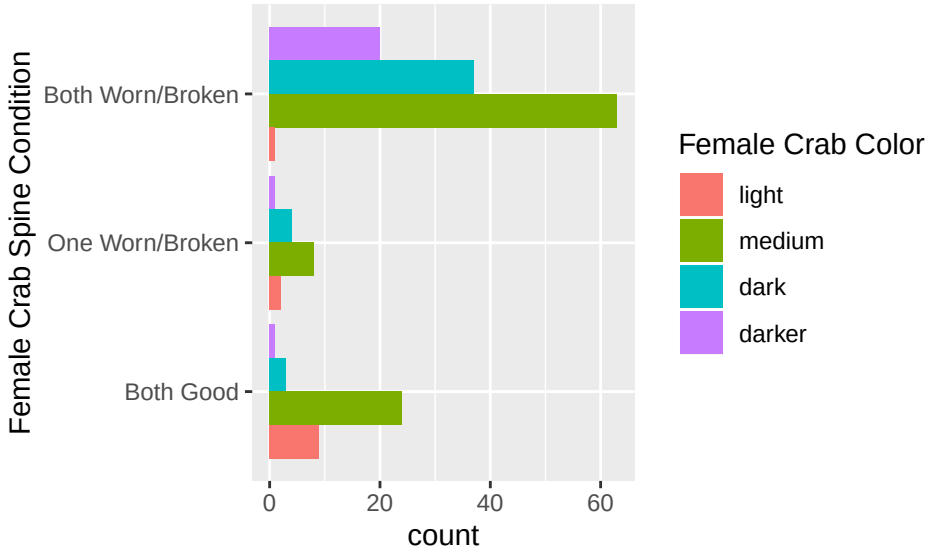
2. Create a two-way contintency table between the satellite and spine variables. Write text describing what two of the numbers in the table represent.

3. Create a three way table between the color, spine, and satellite variables. Output the table. Then, using that table object, print out a two-way table between spine and satellite for crabs with 'darker' color. Write text describing what one of the numbers in the table represents.

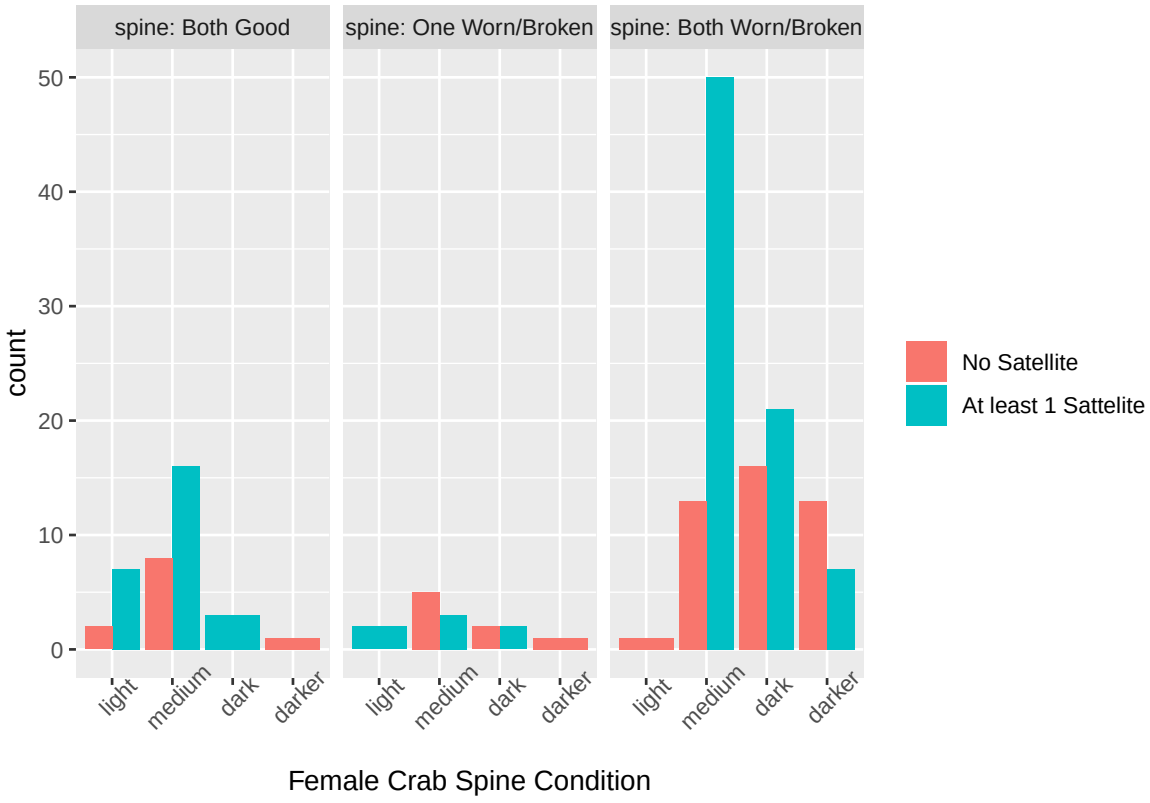
4. Recreate the side-by-side bar plots below (I'm not worried about the sizes of the plots, although they look a little better smaller). Write text that comments on any patterns you see in one of the plots.







5. Recreate the set of side-by-side bar plots below. You'll need to look up how to orient the x-axis lables to 45 degrees. Write text that comments on any patterns you see.

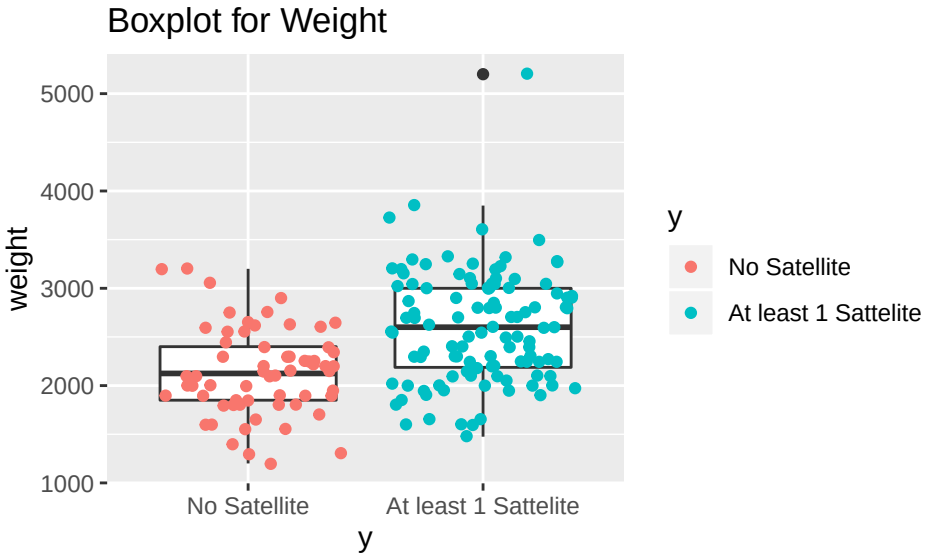


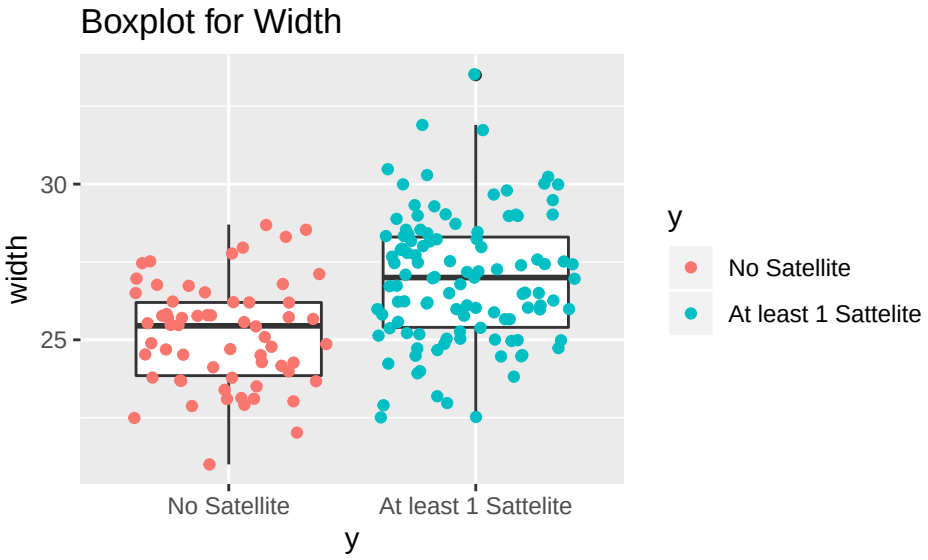
6. Recreate the summary statistics below. Write text that interprets one of the sets of summary stats.

```
## # A tibble: 8 x 6
## # Groups:   color [4]
##   color  y                    Avg    Sd Median   IQR
##   <fct>  <fct>              <dbl> <dbl>  <dbl> <dbl>
## 1 light  No Satellite        2525   152.   2600  138.
```

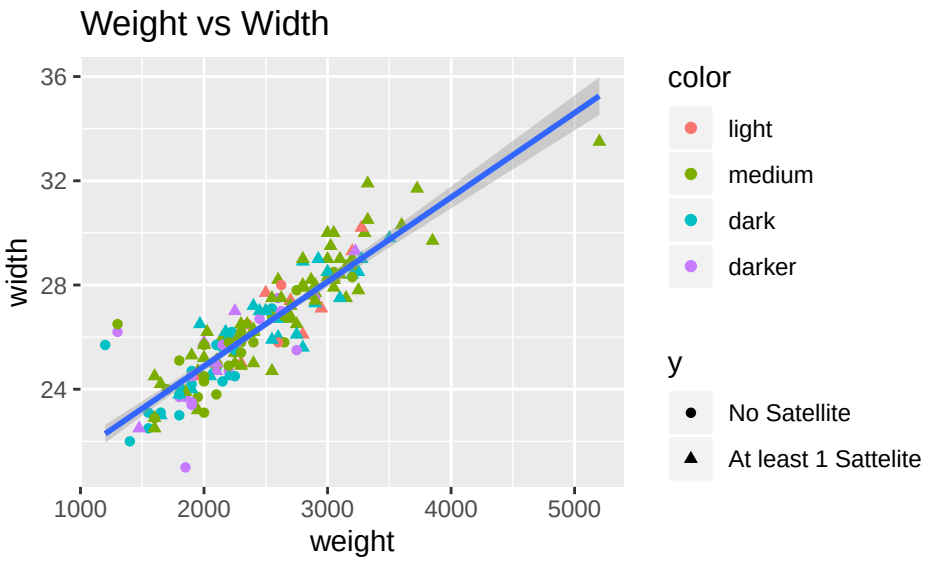
```
## 2 light  At least 1 Sattelite 2664.  442.   2700  650
## 3 medium No Satellite         2242.  482.   2200  512.
## 4 medium At least 1 Sattelite 2649.  615.   2700  800
## 5 dark   No Satellite         1907.  348.   1900  462.
## 6 dark   At least 1 Sattelite 2571.  495.   2575  719.
## 7 darker No Satellite         2162.  436.   2150  650
## 8 darker At least 1 Sattelite 2200   523.   2100  200
```

7. Recreate the boxplots below. Write text that comments on any patterns you see.





8. Report the correlation between the weight and width variables. Then recreate the scatterplot below. Write text that comments on any patterns you see.


Weight vs Width
36
32
28
24
width
1000
2000
3000
4000
5000
weight
color
light
medium
dark
darker
y
No Satellite
At least 1 Sattelite