

# BIO395W Lab Notebook on Phenotypic Selection

*Ars Ghani*

*9/11/2019*

## **BIO395 Lab Notebook**

*Arranged chronologically from 9/6/19 to the end of semester.*

---

### **Friday, 6 September 2019**

*One-on-one meeting with Nancy and Michaela*

Discussed expectations for credit-hour commitments, final paper, and the project. The paper will be peer-reviewed and about 6 papers and will have incremental writing steps that will be punctuated by reviews from Michaela and Nancy. The final paper will be due (by e-mail) to Dr. Larracuente when the semester ends to ensure BIO395 has been completed properly.

*Outline & Annotated Bibliography: Friday, November 1*

First Draft: Monday, November 25

*Peer Review: Wednesday, December 2*

Final draft: Friday, December 13

Rough sketch of this project: 1) Clean the phenotype data and filter it,  
2) This phenotype data will be analyzed by the Mann-Kendall test,  
3) A fitness response (such as clutch size or life span) will be regressed with the statistically significant phenotypes,  
4) Direction and strength of the selection will be inferred,  
5) Grant write-up will be based on this data.

### **Saturday, 7 September 2019**

Read phenotypic selection papers including the famous Lande & Arnold (1983) and Kingsolver & Pfennig (2007). These papers will guide this research project as they focus on the inheritance and patterns of phenotypic selection. There were certain parts of these papers that piqued my interest:

“Phenotypic selection in nature is common and can be measured in the field in real time. In particular, directional selection is often sufficiently strong to cause substantial evolutionary change in a relatively short period” (Kingsolver and Pfennig, 2007).

This forms the basis of our project: measuring the phenotypic changes in populations of FLSJ: because hopefully any or all phenotypic selection in the Jays will be evident through the data collected on that certain phenotype. We can then regress fitness as a function of an individual phenotype and that will give us valuable insight on this phenomenon.

“Directional selection favors larger body size. This pattern contrasts with the pattern for other morphological traits, which tend to experience positive and negative directional selection with equal frequency. Moreover, bigger organisms are generally fitter, regardless of whether larger body size enhances survival, fecundity, or mating success. In fact, directional selection favoring

larger body size is sufficiently strong to explain Cope's rule, the widespread tendency for lineages to evolve toward larger body size" (Kingsolver and Pfennig, 2007).

This actually is very exciting and fascinating to me- I would never think that a larger size would translate to being fitter, even though it seems very plausible. This is one of the reason that a phenotype that I would like to focus on would be *body weight* and some other numerical trait which can correlate with a larger size such as *head length*.

## Thursday, 12 September 2019

Started and finished making a spreadsheet which would be the stepping-stone to digitizing the metadata. It contains a separate sheet for each of the list in the FSLJ demography data with its corresponding data frame and columns. The best part about this whole endeavor is that this will serve as a dictionary and reference point for any future analyses. This is because it breaks down each and every data marker of the demography data and also gives a small description of what it is and why it is included. The spreadsheet can be found here: [Link](#)

## Friday, 13 September 2019

*One-on-one meeting with Nancy and Michaela*

Got more clarity on the project and approval for the spreadsheet. The spreadsheet is now more detailed and will be filled in within a week. It would be pertinent to choose the phenotypes for analysis during this time since it is the first time we are really diving into the demography data.

```
#Using this for figuring out the type of each data point:  
#str()  
#typeof()  
#This usually gives a type such as integer, date, or character.
```

## Monday, 16 to Thursday, 19 September 2019

Around 10 hours were spent accessing the myriad of data contained in the FLSJ demography data and digitizing it. It personally gave me a lot of time to look at each data entry personally, explore the serious nooks of different tables, and decide upon phenotypes for analysis by looking at data available for them. Since this is a collaborative effort, certain parts of our spreadsheet remain unfilled but for the most part it seems to be done.

Next steps: *Getting the list of phenotypes confirmed by Michaela and Nancy,*

Data filtering step for each phenotype,

\*Knowing how PLE digitization will tie in with this project.

## Thursday, 19 to Friday, 27 September 2019

This was a very productive week especially due to the clear expectations set in the last week's meeting. Right after last week's meeting, the following goals for this week were made:

- Complete the metadata sheet and highlight description for everything we are unsure about,
- Complete lit search: 20 papers and 4 books reference minimum on phenotypes, fitness, and other things our project is focusing on,

- Make a list of phenotypes, split it between the group members, and also put in where it could be found with code,
- Put in why we think this phenotype is important and how it will act in the population.

The metadata sheet was completed. Next group meeting, we can go over description of stuff we were confused about. The metadata sheet can be found here: [Link](#)

A lit search was also completed, and in doing so I think I realized what I really want to look into specifically: Cope's Rule in the FLSJ. I have thoroughly looked in Cope's Law and I am sure it will make for a great and interesting dive into the FLSJ phenotypes. The lit search can be found here: [link](#)

The phenotypes were also divided:

#### **Phenotypes that I am doing:**

- Weight (access\$metrics)
- HeadLength (access\$metrics)
- BillDepth (access\$metrics)
- BillWidth (access\$metrics)
- TailLength (access\$metrics)
- HeadBreadth (access\$metrics)
- Culmen (access\$metrics)

#### **Phenotypes that Michaela is doing:**

- Parasite (access\$metrics)
- Manus (access\$metrics)
- Tarsus (access\$metrics)
- Primary7 (access\$metrics)
- WingCord (access\$metrics)
- Nares (access\$metrics)

The week ended with me choosing the discussion paper for the lab meeting, and Michaela and I pondering upon how each and every chosen phenotype will act in population to form sort of mental models and early hypotheses.

## **October and Fall Break:**

Filled with more lit searches and more importantly finalization of phenotypes.

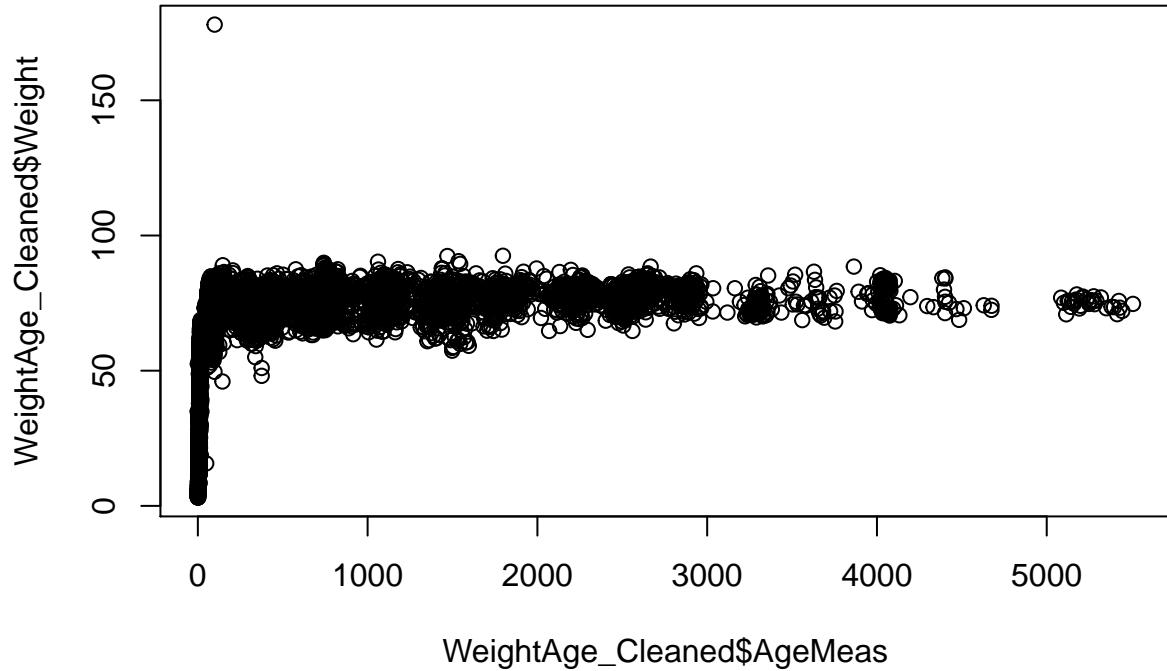
```
metadata <- load("/cloud/project/all_tables_v2.rdata")
library(ggplot2)
```

## **Filtering Each Phenotype Alone.**

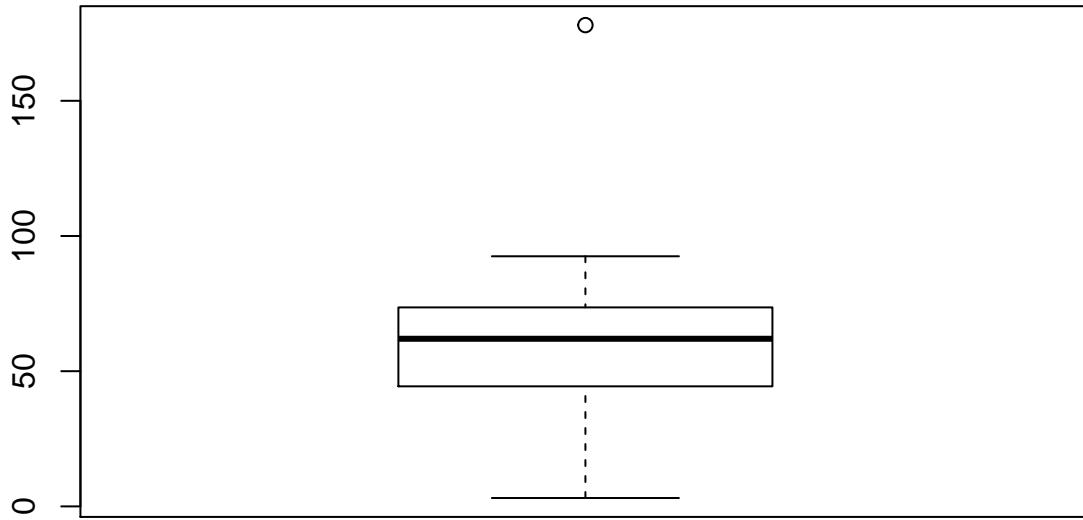
The phenotypes I chose were Weight, HeadLength, BillDepth, BillWidth, TailLength, HeadBreadth, Culmen, Nares. My general protocol for filtering these phenotypes is: 1) Remove all data entries that do not have a date of birth and/or a date of measurement because then it is impossible to determine the age of the bird, 2) Only Demo tract data was used, 3) Data only after 1990 was used, 4) Keep the ID of the bird. This will allow us to merge them together later when calculating correlations between phenotypes. 5) Remove very obviously erroneous data points. 6) Also the ages were categorised in brackets: • day 0 to 11: nestling, • >day 11 post hatch: fledgelings, • >day 70 post hatch: juveniles, • >365 post hatch: yearlings (with multiples)

Filtering Weight Alone:

```
WeightAge <- (access$metrics[, c('Weight', 'AgeMeas', 'MeasDate', "ID", 'est.hatch', 'USFWS')]) # Store  
##View(WeightAge) # ##Viewing the variable  
WeightAge_Cleaned <- WeightAge[which(!is.na(WeightAge$Weight) & WeightAge$AgeMeas != ""), ] # Removing  
##View(WeightAge_Cleaned) # ##Viewing cleaned data  
plot(WeightAge_Cleaned$Weight~WeightAge_Cleaned$AgeMeas) # Just a quick scatterplot to see where we start
```



```
boxplot(WeightAge_Cleaned$Weight)
```



```
outliersweight<-boxplot(WeightAge_Cleaned$Weight, plot=FALSE)$out # This gives me the exact value of the outliers  
print(outliersweight)
```

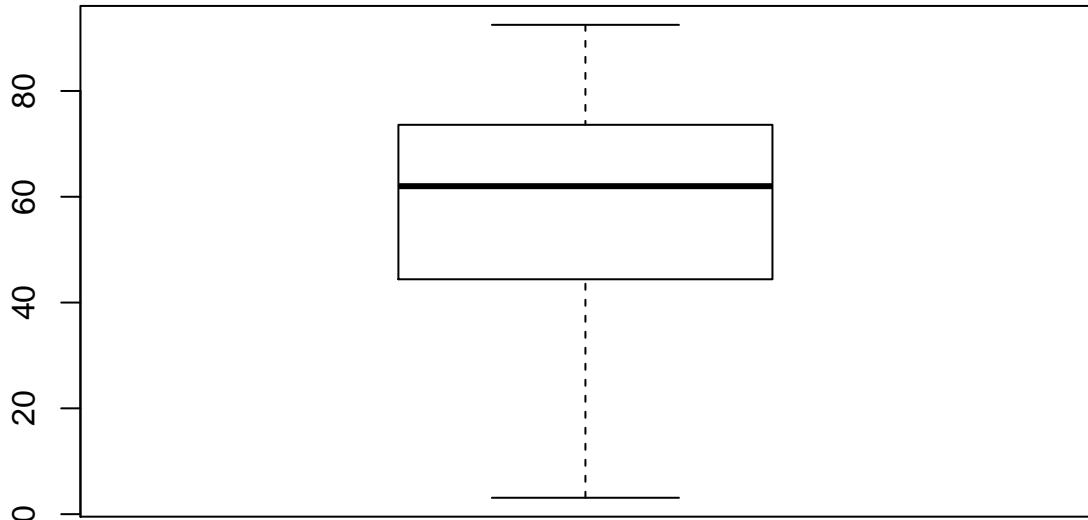
```
## [1] 178
```

```
WeightAge_Cleaned[which(WeightAge_Cleaned$Weight %in% outliersweight),] # First you need find in which
```

```
##      Weight AgeMeas     MeasDate      ID est.hatch       USFWS
```

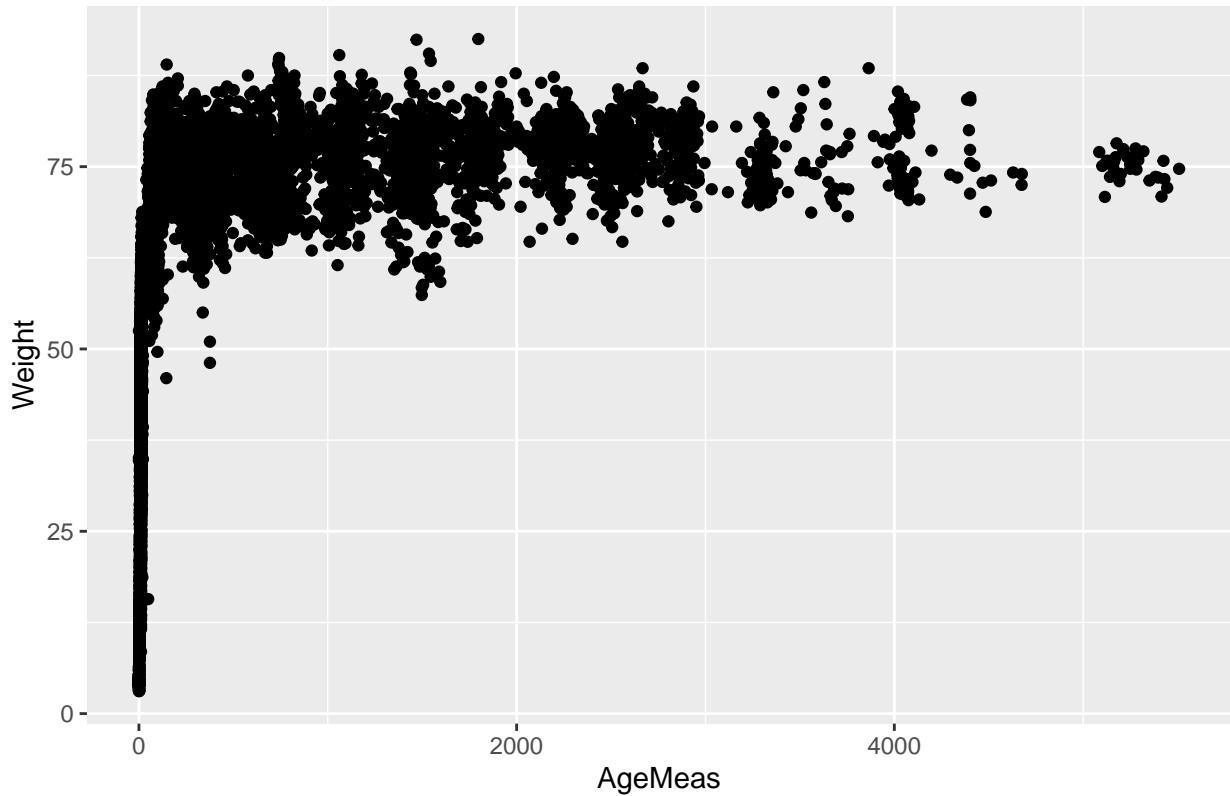
```
## 1643     178      99 2001-07-06 1644 2001-03-29 1513-50904
```

```
WeightAge.Cleaned<-WeightAge_Cleaned[!(WeightAge_Cleaned$Weight %in% c(178)), ] # The outlier is in the  
##View(WeightAge.Cleaned)  
boxplot(WeightAge.Cleaned$Weight)
```



```
ggplot(data = WeightAge.Cleaned, aes(x = AgeMeas , y = Weight)) + geom_point() + ggttitle("Weight vs Age")
```

Weight vs Age



```
# There are 11,750 entries.
```

Filtering Headlength Alone:

```

HeadlengthAge <- (access$metrics[, c('HeadLength', 'AgeMeas', 'MeasDate', "ID", 'est.hatch', 'USFWS')])  

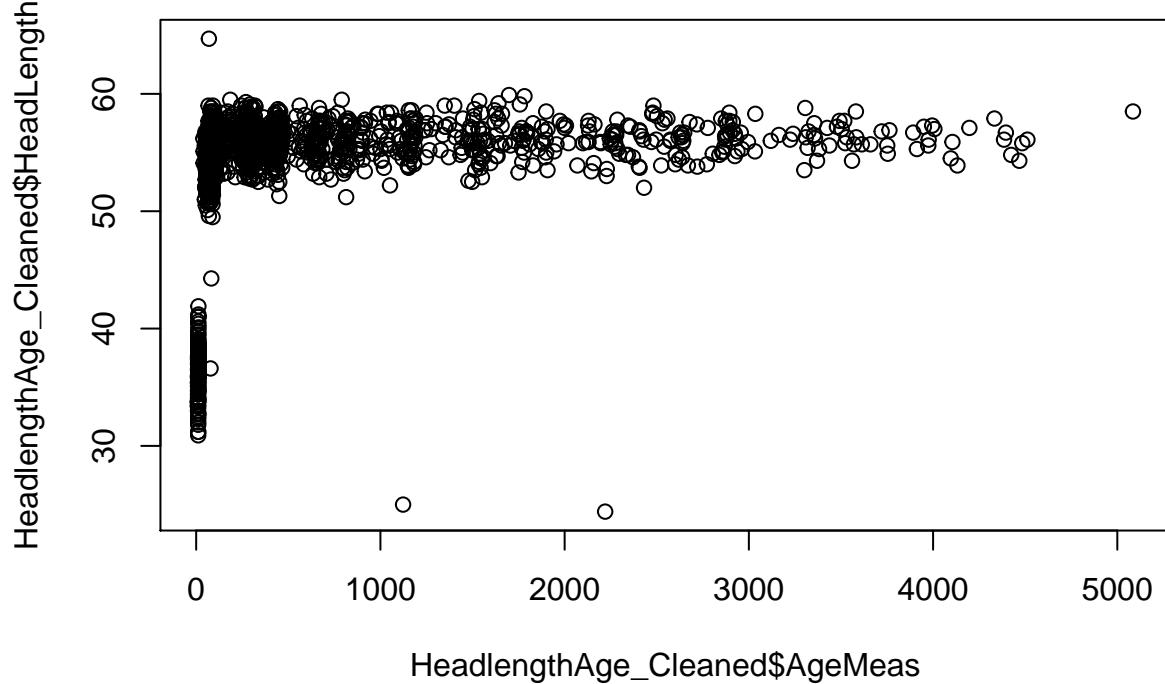
##View(HeadlengthAge) # ##Viewing the variable  

HeadlengthAge_Cleaned <- HeadlengthAge[which(!is.na(HeadlengthAge$HeadLength) & HeadlengthAge$AgeMeas != 0),]  

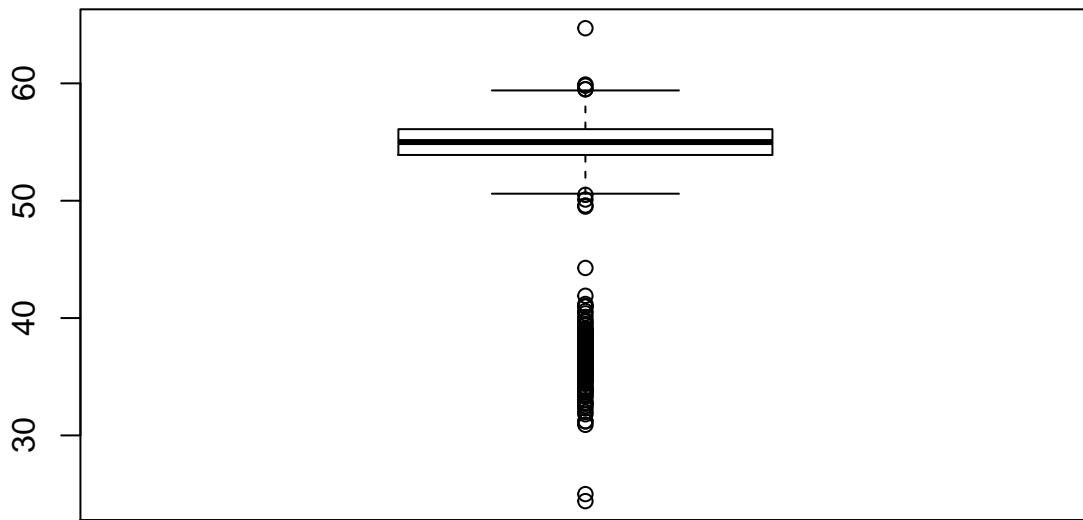
##View(HeadlengthAge_Cleaned) # ##Viewing cleaned data  

plot(HeadlengthAge_Cleaned$HeadLength~HeadlengthAge_Cleaned$AgeMeas) # Just a quick scatterplot to see if there is a relationship

```

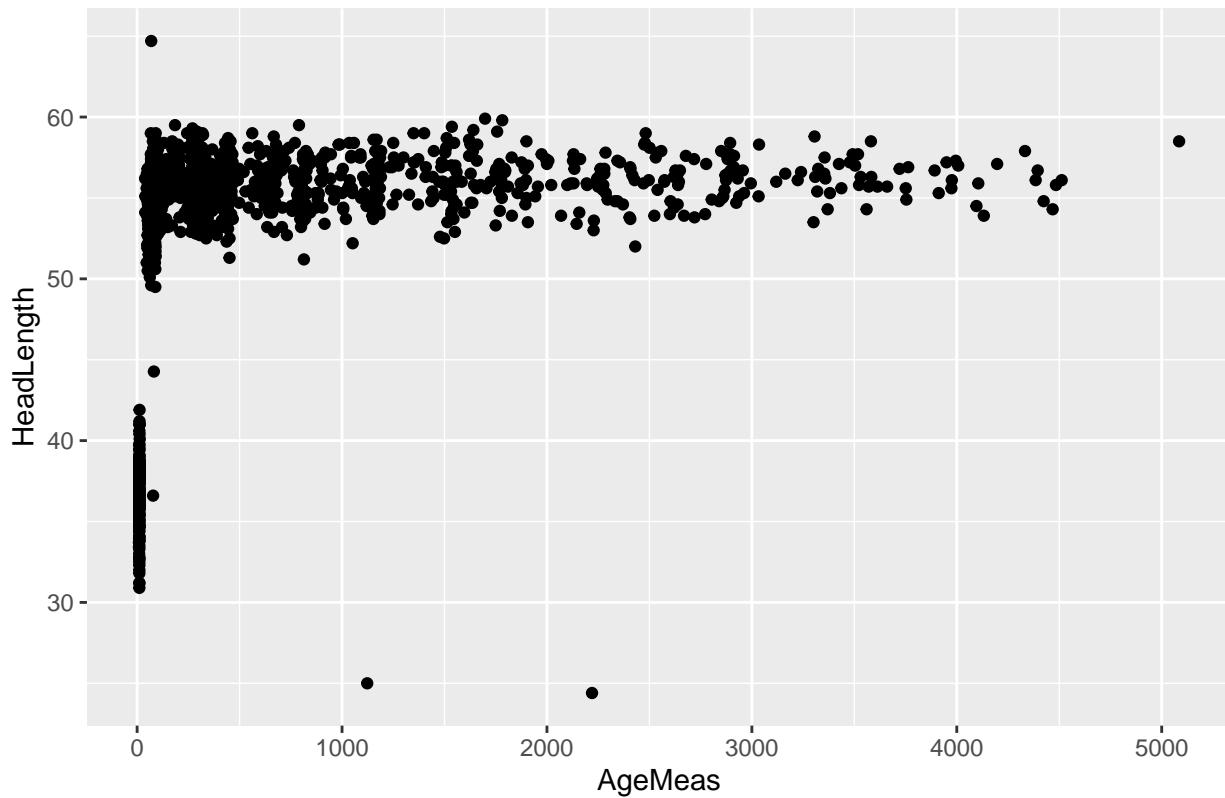


```
boxplot(HeadlengthAge_Cleaned$HeadLength) # No erroneous data points are obvious.
```



```
ggplot(data = HeadlengthAge_Cleaned, aes(x = AgeMeas , y = HeadLength)) + geom_point() + ggtitle("HeadLength vs AgeMeas")
```

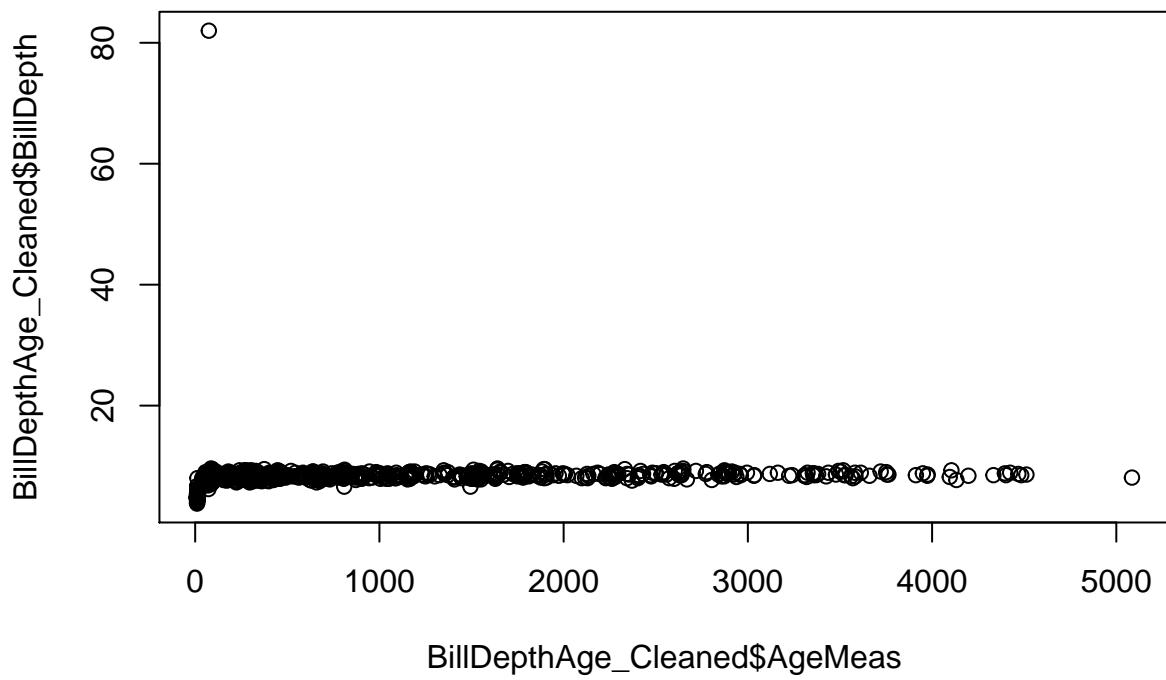
## Head length vs Age



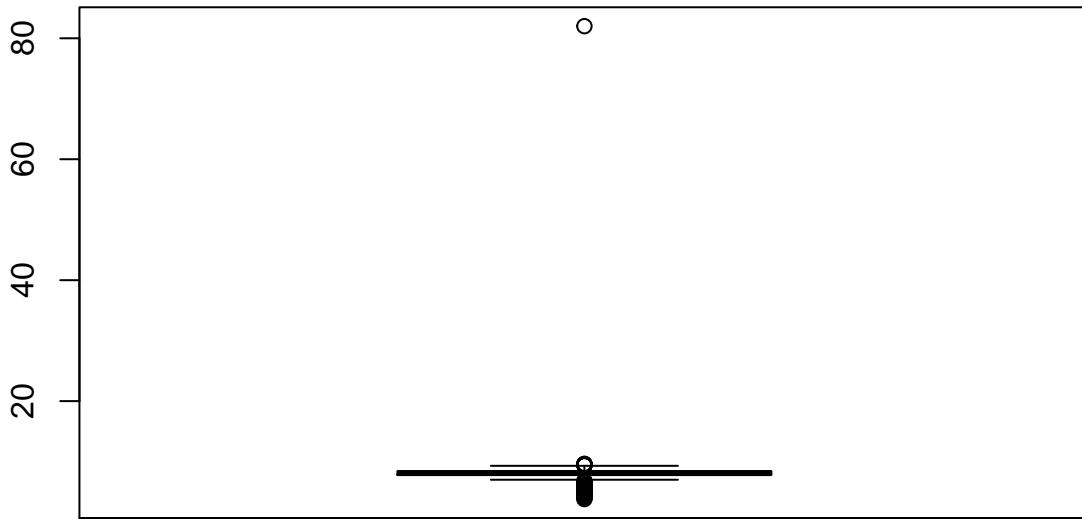
```
# There are 2,711 entries.
```

Filtering BillDepth Alone:

```
BillDepthAge <- (access$metrics[, c('BillDepth', 'AgeMeas', 'MeasDate', "ID", 'est.hatch', 'USFWS')])  
##View(BillDepthAge) # ##Viewing the variable  
BillDepthAge_Cleaned <- BillDepthAge[which(!is.na(BillDepthAge$BillDepth) & BillDepthAge$AgeMeas != "")]  
##View(BillDepthAge_Cleaned) # ##Viewing cleaned data  
plot(BillDepthAge_Cleaned$BillDepth~BillDepthAge_Cleaned$AgeMeas) # Just a quick scatterplot to see whe
```



```
boxplot(BillDepthAge_Cleaned$BillDepth) # A huge outlier noticed at ~80. Let's remove it.
```



```
outliersbilldepth<-boxplot(BillDepthAge_Cleaned$BillDepth, plot=FALSE)$out # This gives me the exact values
print(outliersbilldepth) # We can see it as 82.
```

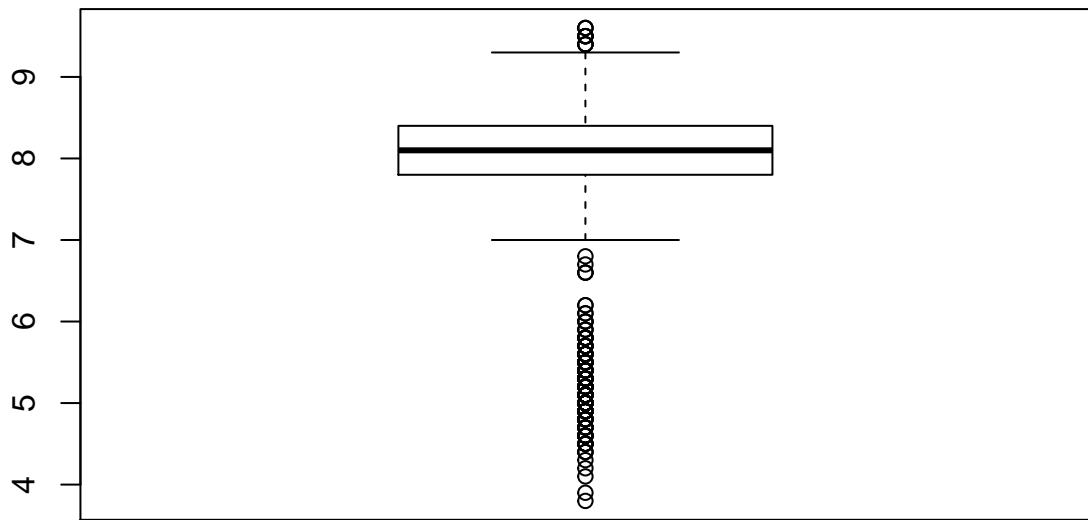
```
## [1] 6.2 6.6 6.8 9.5 5.3 6.0 5.3 4.9 5.0 5.8 5.0 6.0 6.1 5.0
## [15] 4.9 4.8 4.9 4.9 5.8 5.0 5.1 4.8 4.7 4.1 5.5 5.0 5.2 5.7
## [29] 5.4 5.7 4.8 5.4 5.6 4.8 4.9 5.0 5.2 5.3 5.2 5.2 5.2 5.5
## [43] 5.5 5.1 5.9 5.2 4.6 4.7 3.8 4.8 4.6 4.6 4.6 4.6 4.8 5.4
## [57] 4.9 4.4 5.1 4.8 4.8 4.5 4.9 4.6 4.9 4.5 4.9 4.5 5.4 5.3
## [71] 5.2 3.9 4.6 5.1 5.0 5.0 5.4 4.6 5.0 5.5 5.2 4.9 5.3 5.2
## [85] 5.6 4.7 4.9 5.0 5.0 5.0 5.3 5.4 5.8 6.1 5.6 5.0 5.0 5.4
## [99] 5.4 4.4 4.9 5.2 4.9 5.2 5.3 5.4 5.5 5.5 5.3 5.0 5.3 5.7
## [113] 5.0 5.1 5.3 5.5 5.6 5.9 5.5 5.0 5.1 5.1 5.1 5.3 5.8 5.4
## [127] 9.6 5.8 5.0 5.7 5.3 4.7 4.8 5.4 5.4 5.7 5.7 5.4 5.3 4.4
## [141] 4.5 5.0 4.6 4.8 6.0 6.0 6.2 4.9 9.5 5.6 5.2 5.3 4.9 5.3
```

```

## [155] 5.4 5.2 5.3 4.3 4.5 5.1 4.9 5.3 5.8 5.5 5.6 5.1 5.4 5.4
## [169] 5.5 5.8 5.2 5.4 5.5 5.3 4.8 5.2 5.0 5.3 5.6 5.7 4.8 5.3
## [183] 5.4 5.6 5.1 5.3 5.3 5.5 5.5 4.8 5.5 4.8 4.8 4.9 5.4 5.2
## [197] 5.4 5.3 5.3 5.1 5.0 5.3 5.1 5.1 5.2 5.3 5.1 4.6 4.7 5.5
## [211] 5.4 5.2 4.6 5.3 4.7 4.9 5.0 5.3 4.9 4.8 4.6 4.6 4.9 5.2
## [225] 4.8 5.0 4.9 5.3 4.5 4.7 4.9 4.9 5.0 4.8 5.5 5.0 4.7 5.0
## [239] 4.8 5.0 4.8 5.3 5.1 5.2 5.4 5.8 5.6 5.5 4.2 5.0 5.2 4.6
## [253] 5.2 5.0 5.3 5.1 5.3 6.6 5.9 5.4 6.7 5.0 4.9 5.3 9.5 9.4
## [267] 6.6 9.4 4.9 5.2 5.3 9.6 9.5 9.4 9.5 9.4 9.6 82.0

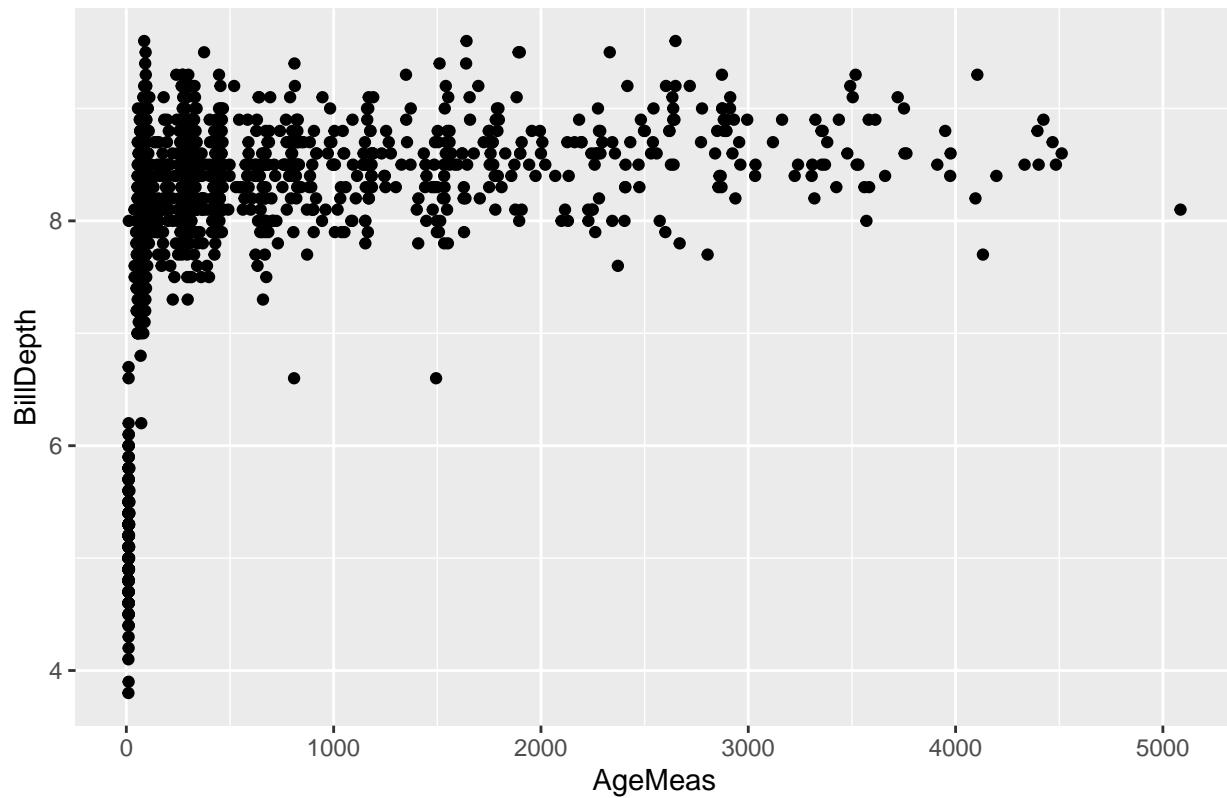
BillDepthAge.Cleaned<-BillDepthAge_Cleaned[!(BillDepthAge_Cleaned$BillDepth %in% c(82)), ]
boxplot(BillDepthAge.Cleaned$BillDepth)

```



```
ggplot(data = BillDepthAge.Cleaned, aes(x = AgeMeas , y = BillDepth)) + geom_point() + ggtitle("BillDep
```

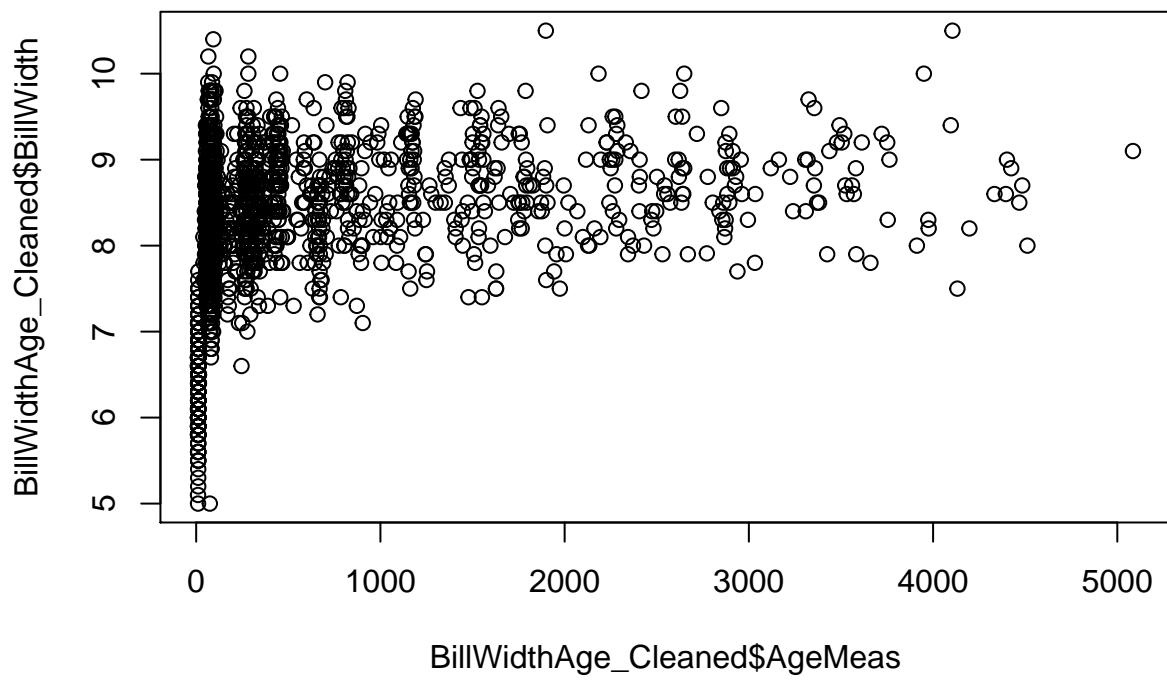
## BillDepth vs Age



```
# There are 2,760 entries.
```

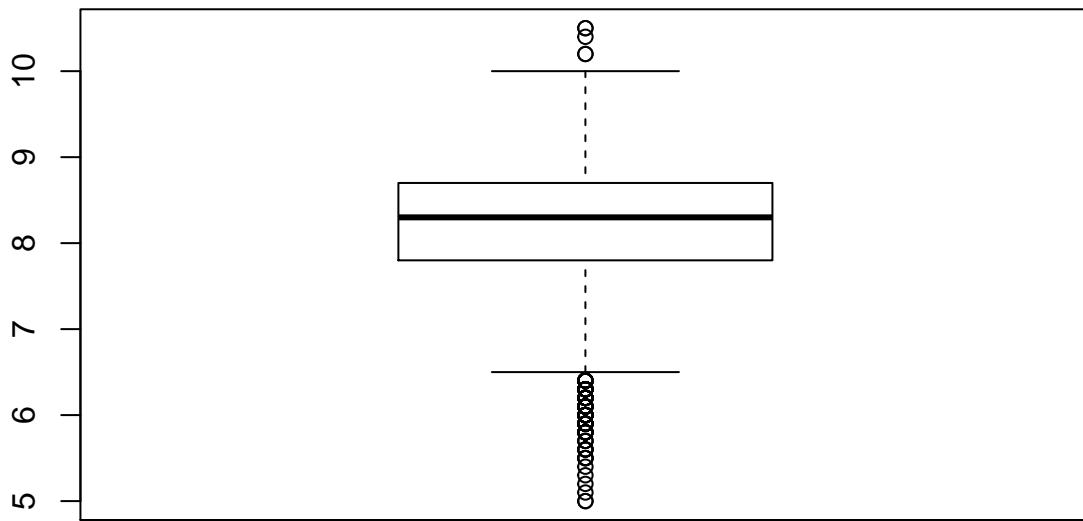
Filtering BillWidth Alone:

```
BillWidthAge <- (access$metrics[, c('BillWidth', 'AgeMeas', 'MeasDate', "ID", 'est.hatch', 'USFWS')])  
##View(BillWidthAge) # ##Viewing the variable  
BillWidthAge_Cleaned <- BillWidthAge[which(!is.na(BillWidthAge$BillWidth) & BillWidthAge$AgeMeas != "")]  
##View(BillWidthAge_Cleaned) # ##Viewing cleaned data  
plot(BillWidthAge_Cleaned$BillWidth~BillWidthAge_Cleaned$AgeMeas) # Just a quick scatterplot to see whe
```



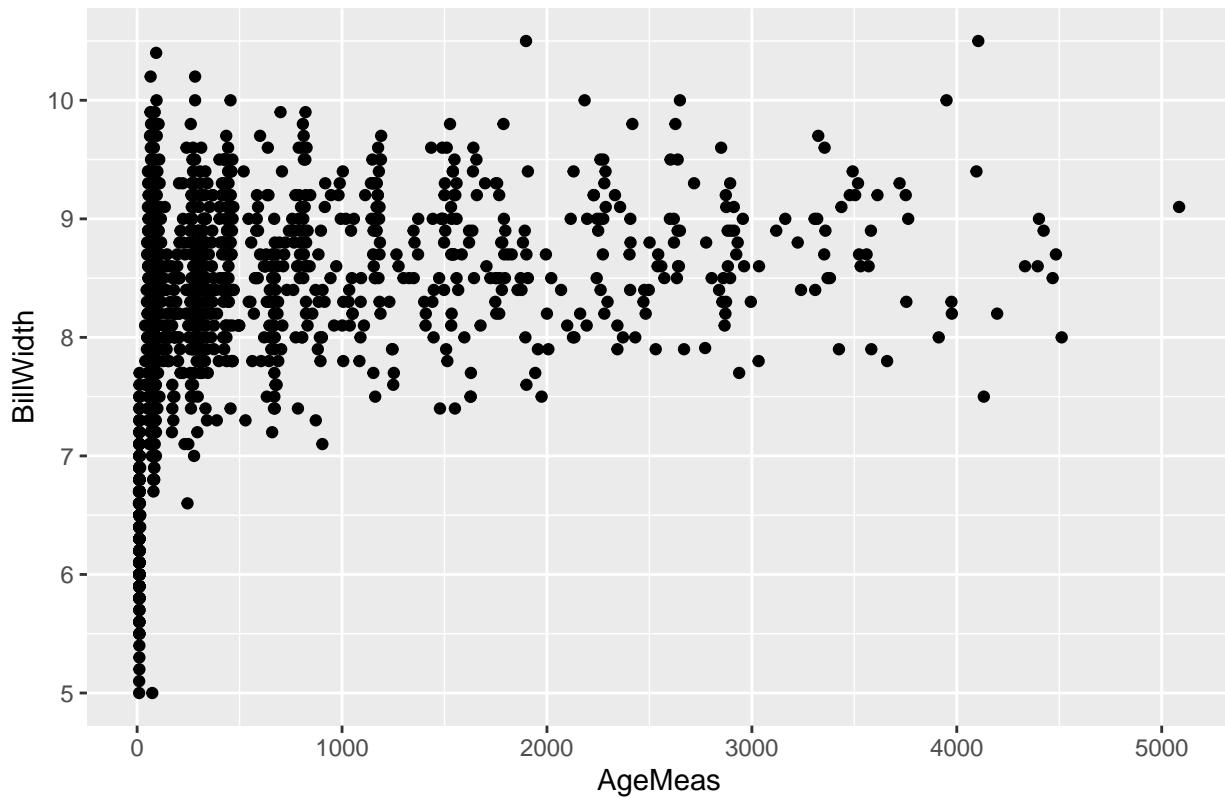
BillWidthAge\_Cleaned\$AgeMeas

```
boxplot(BillWidthAge_Cleaned$BillWidth) # No erroneous data detected.
```



```
ggplot(data = BillWidthAge_Cleaned, aes(x = AgeMeas , y = BillWidth)) + geom_point() + ggtitle("Bill Wi")
```

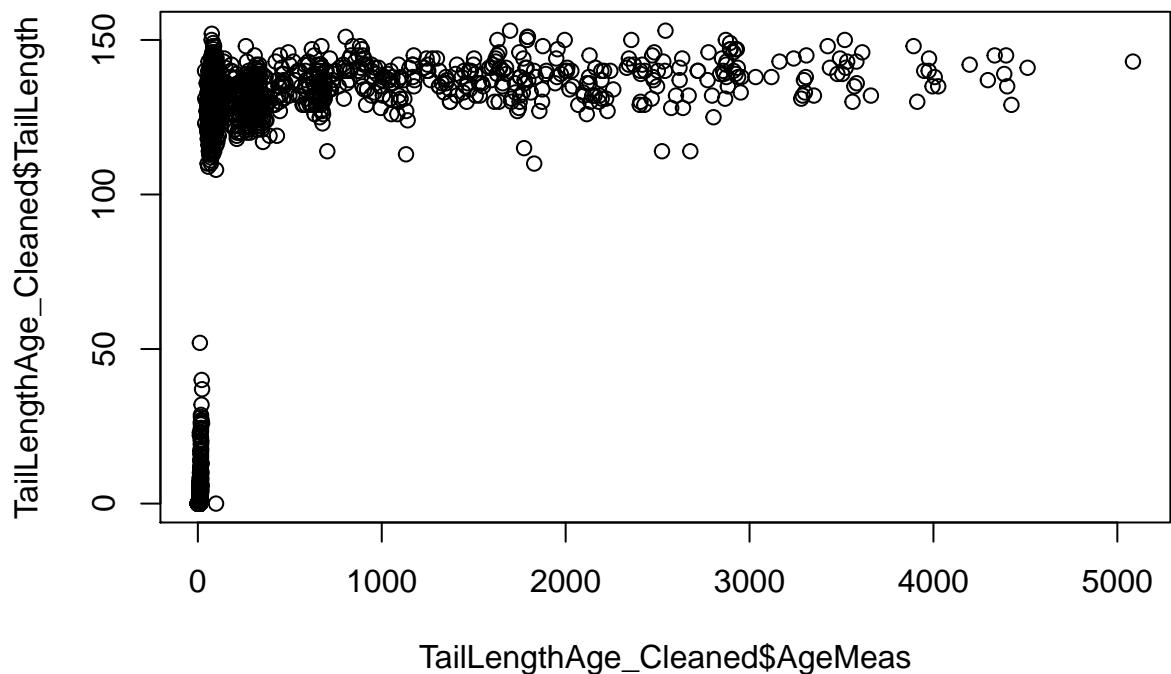
## Bill Width vs Age



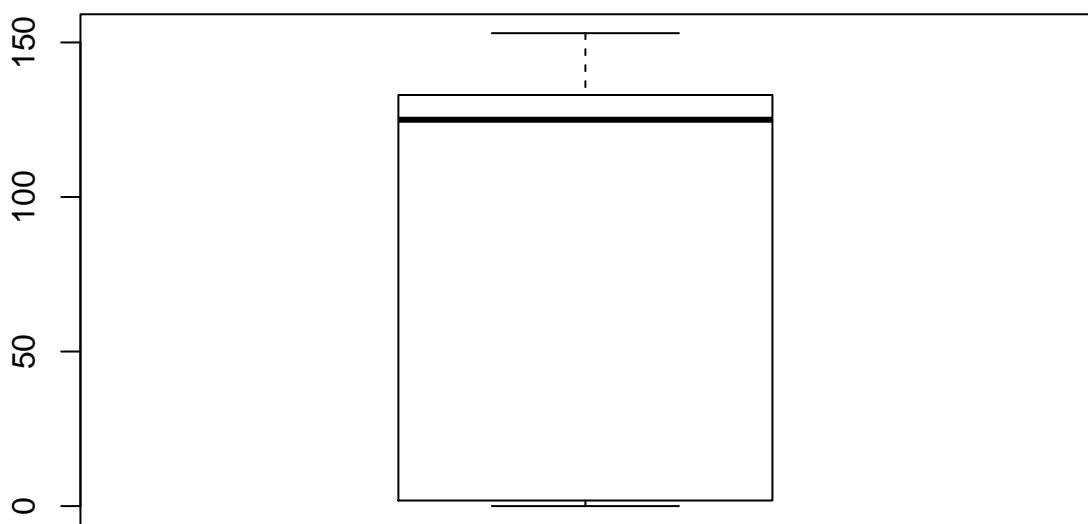
```
# There are 2,758 entries.
```

Filtering TailLength Alone:

```
TailLengthAge <- (access$metrics[, c('TailLength', 'AgeMeas', 'MeasDate', "ID", 'est.hatch', 'USFWS')])  
##View(TailLengthAge) # ##Viewing the variable  
TailLengthAge_Cleaned <- TailLengthAge[which(!is.na(TailLengthAge$TailLength) & TailLengthAge$AgeMeas != 0),]  
##View(TailLengthAge_Cleaned) # ##Viewing cleaned data  
plot(TailLengthAge_Cleaned$TailLength~TailLengthAge_Cleaned$AgeMeas) # Just a quick scatterplot to see if there is a relationship
```

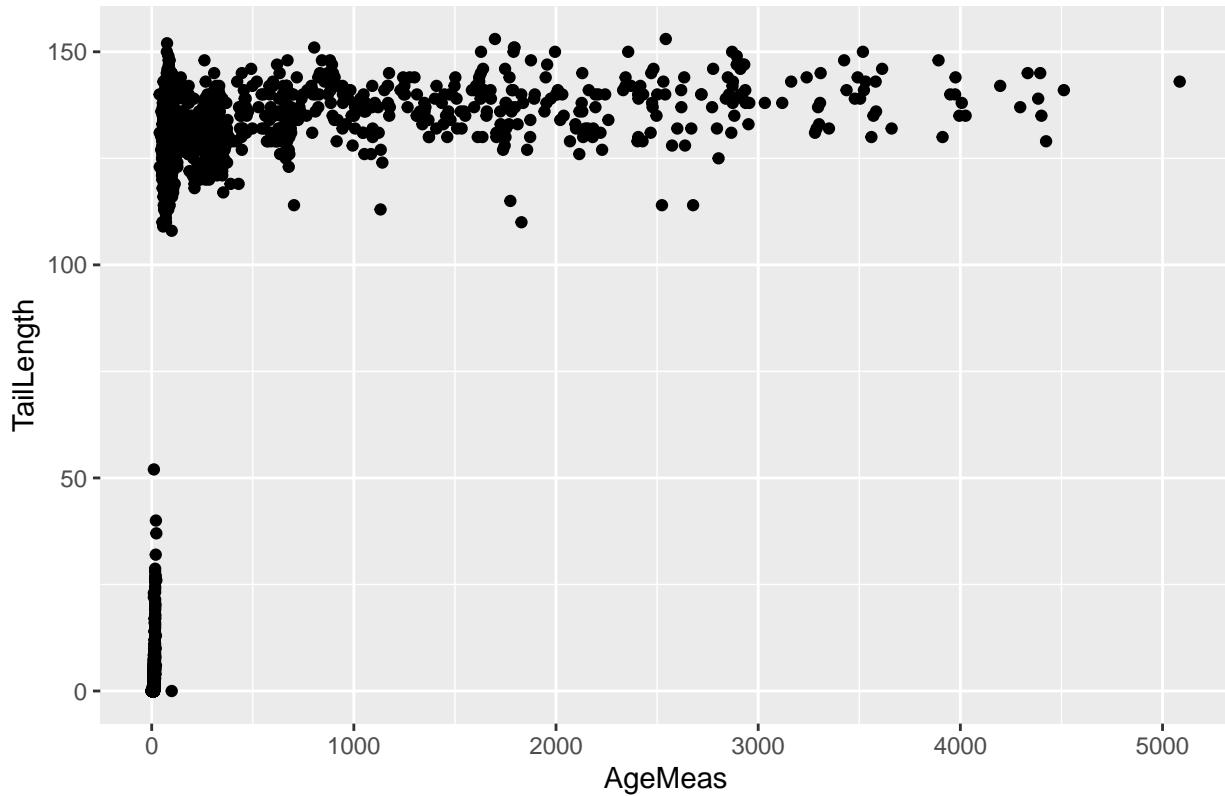


```
boxplot(TailLengthAge_Cleaned$TailLength)
```



```
ggplot(data = TailLengthAge_Cleaned, aes(x = AgeMeas , y = TailLength)) + geom_point() + ggttitle("Tail Length vs Age Measured")
```

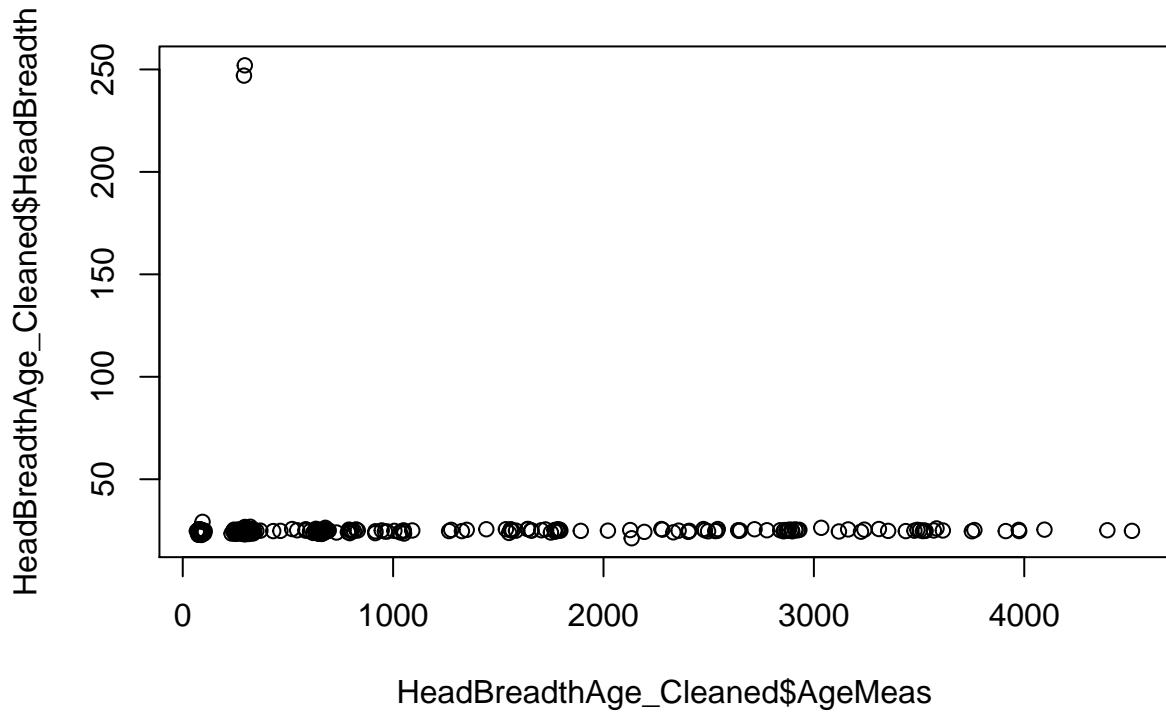
## Tail Length vs Age



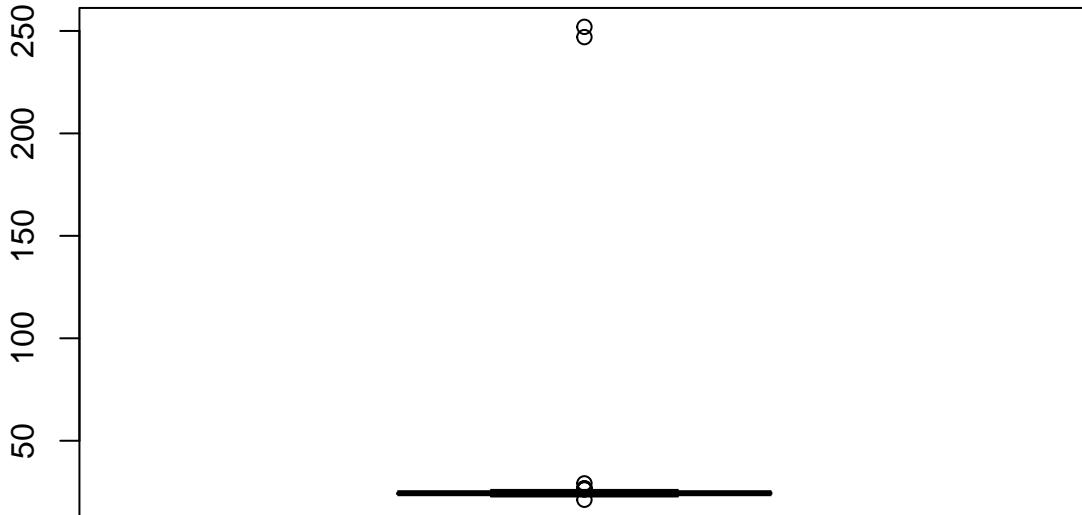
```
# There are 4,429 entries.
```

Filtering HeadBreadth Alone:

```
HeadBreadthAge <- (access$metrics[, c('HeadBreadth', 'AgeMeas', 'MeasDate', "ID", 'est.hatch', 'USFWS')])  
##View(HeadBreadthAge) # ##Viewing the variable  
HeadBreadthAge_Cleaned <- HeadBreadthAge[which(!is.na(HeadBreadthAge$HeadBreadth) & HeadBreadthAge$AgeMeas != 0),]  
##View(HeadBreadthAge_Cleaned) # ##Viewing cleaned data  
plot(HeadBreadthAge_Cleaned$HeadBreadth~HeadBreadthAge_Cleaned$AgeMeas) # Just a quick scatterplot to see if it worked
```



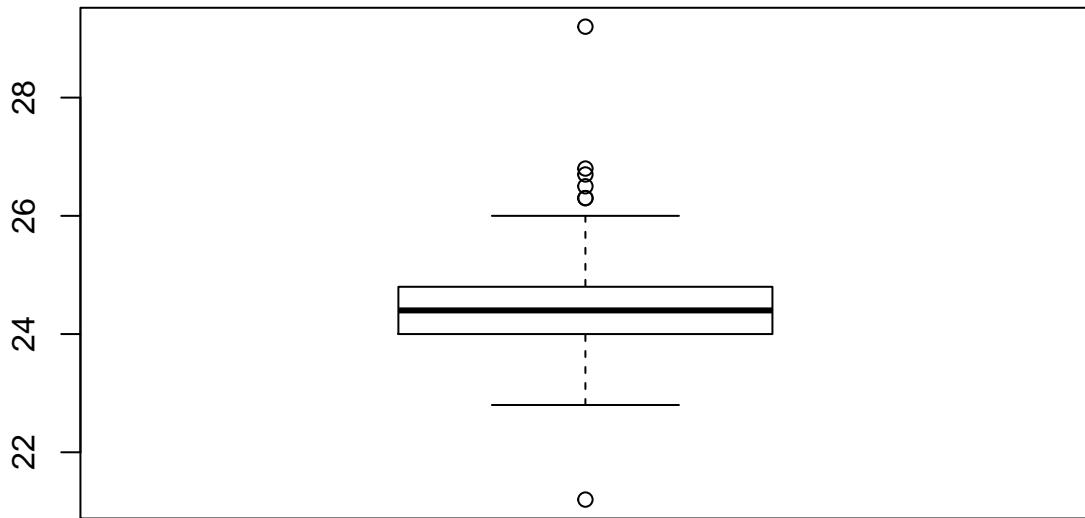
```
boxplot(HeadBreadthAge_Cleaned$HeadBreadth) #Two outliers/errors can be seen.
```



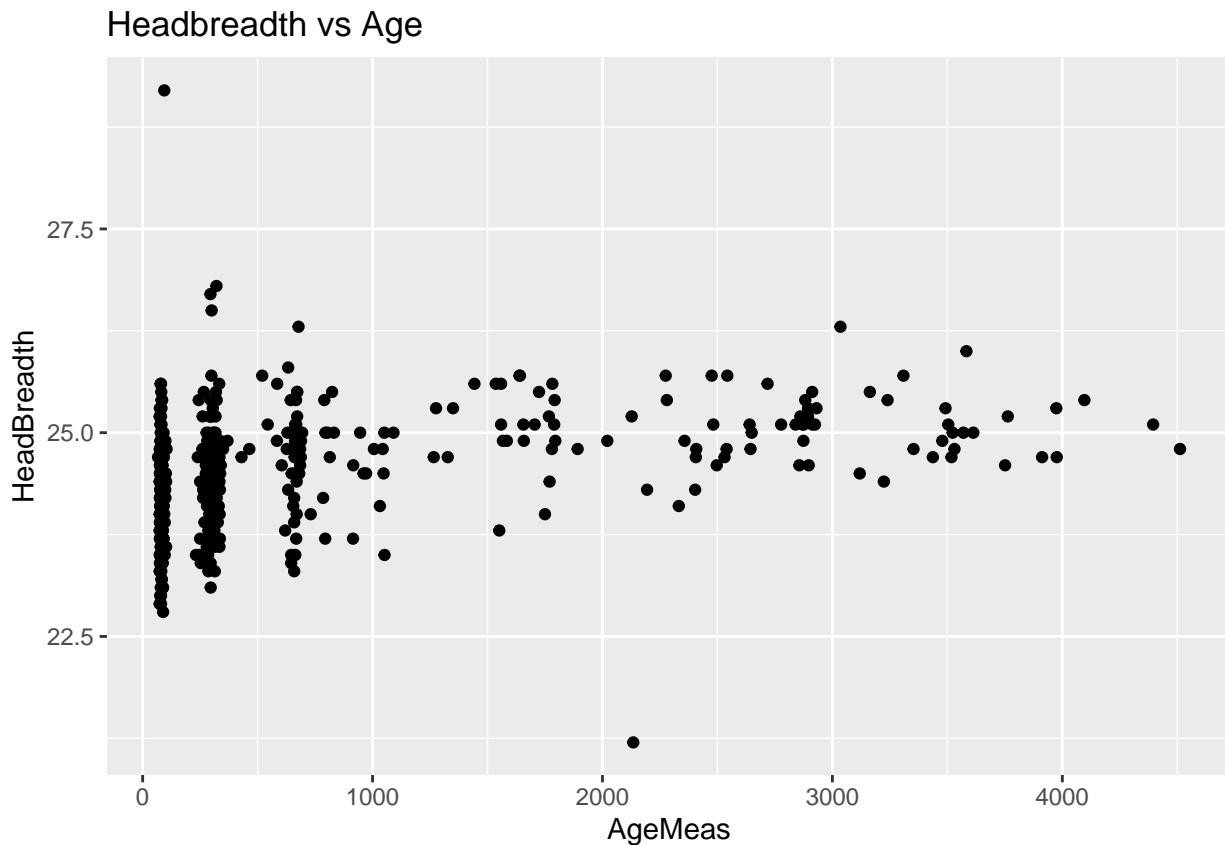
```
outliersheadbreadth<-boxplot(HeadBreadthAge_Cleaned$HeadBreadth, plot=FALSE)$out # This gives me the exact values
print(outliersheadbreadth) # read them off as 252 and 247.
```

```
## [1] 26.8 26.5 252.0 26.7 247.0 29.2 26.3 21.2 26.3
```

```
HeadBreadthAge.Cleaned<-HeadBreadthAge_Cleaned[!(HeadBreadthAge_Cleaned$HeadBreadth %in% c(247,252))]
```



```
##View(HeadBreadthAge.Cleaned)
ggplot(data = HeadBreadthAge.Cleaned, aes(x = AgeMeas , y = HeadBreadth)) + geom_point() + ggtitle("Headbreadth vs Age")
```

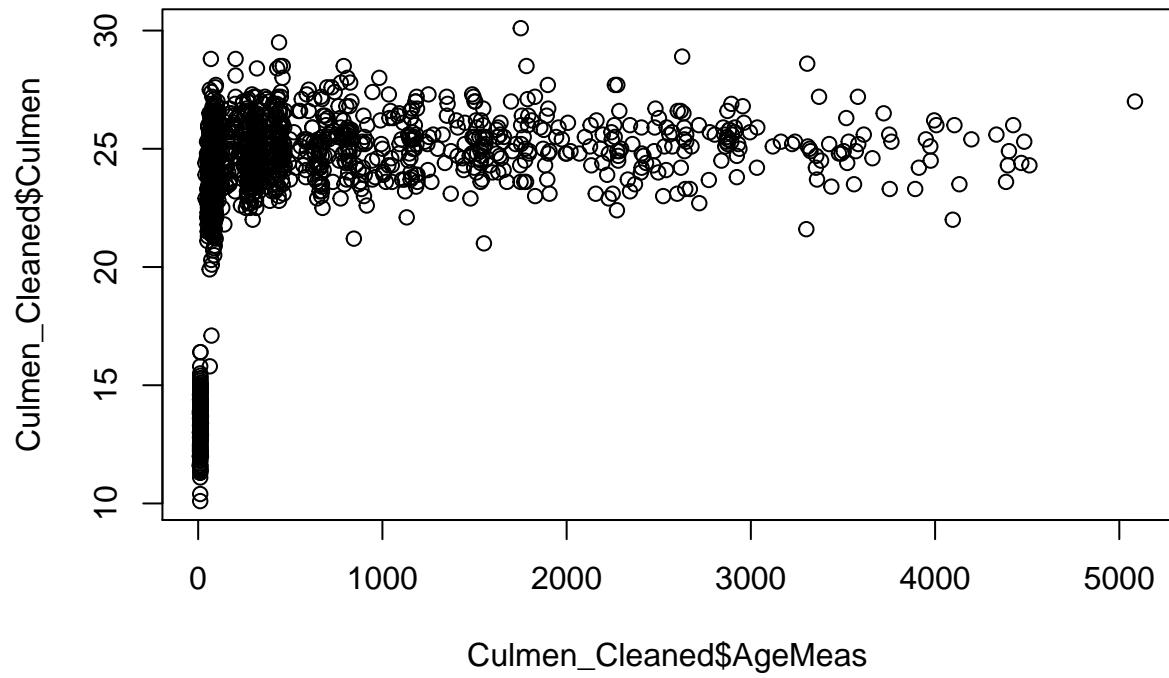


```
# There are 694 entries.
```

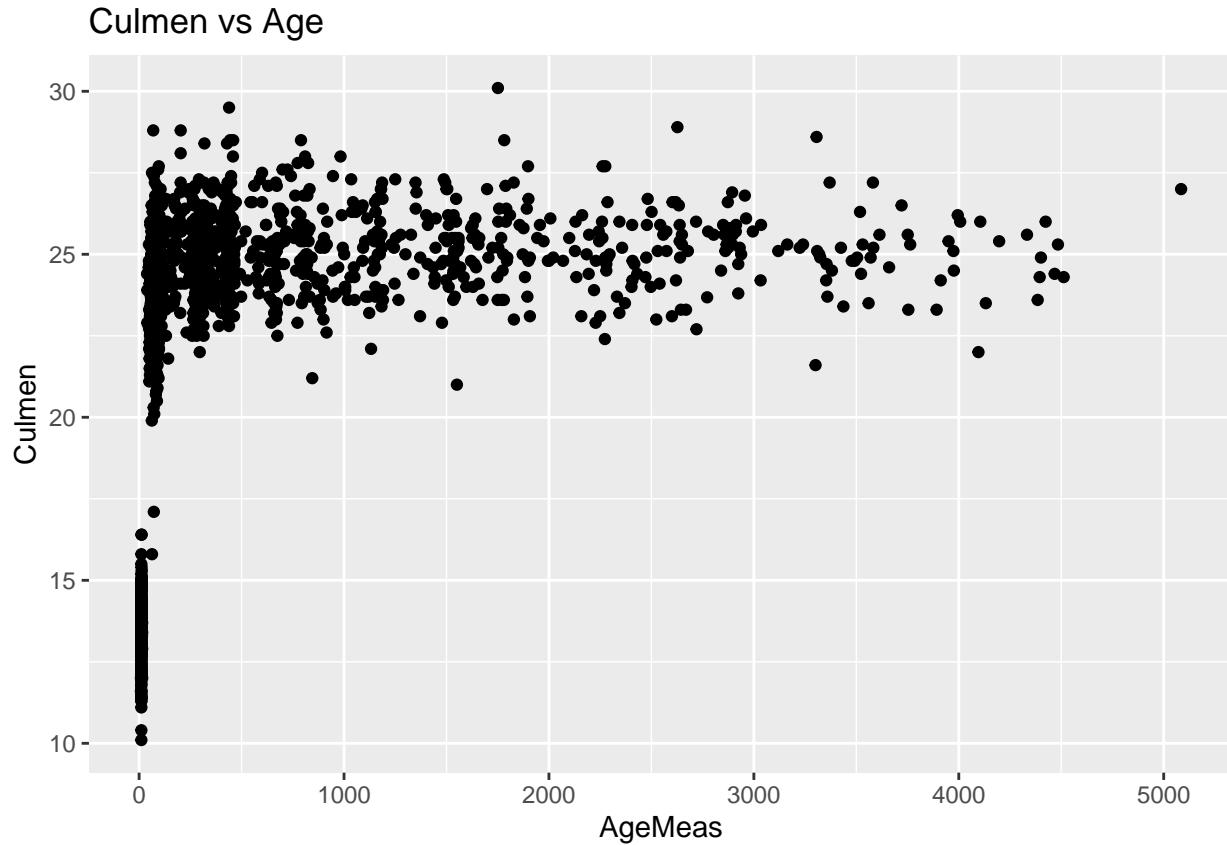
Filtering Culmen Alone:

```
CulmenAge <- (access$metrics[, c('Culmen', 'AgeMeas', 'MeasDate', "ID", 'est.hatch', 'USFWS')]) # Storing the metrics
##View(CulmenAge) # ##Viewing the variable
Culmen_Cleaned <- CulmenAge[which(!is.na(CulmenAge$Culmen) & CulmenAge$AgeMeas != ""), ] # Removing any rows with missing values
##View(Culmen_Cleaned) # ##Viewing cleaned data
```

```
plot(Culmen_Cleaned$Culmen~Culmen_Cleaned$AgeMeas) # Just a quick scatterplot to see where we stand at
```



```
ggplot(data = Culmen_Cleaned, aes(x = AgeMeas , y = Culmen)) + geom_point() + ggtitle("Culmen vs Age")
```

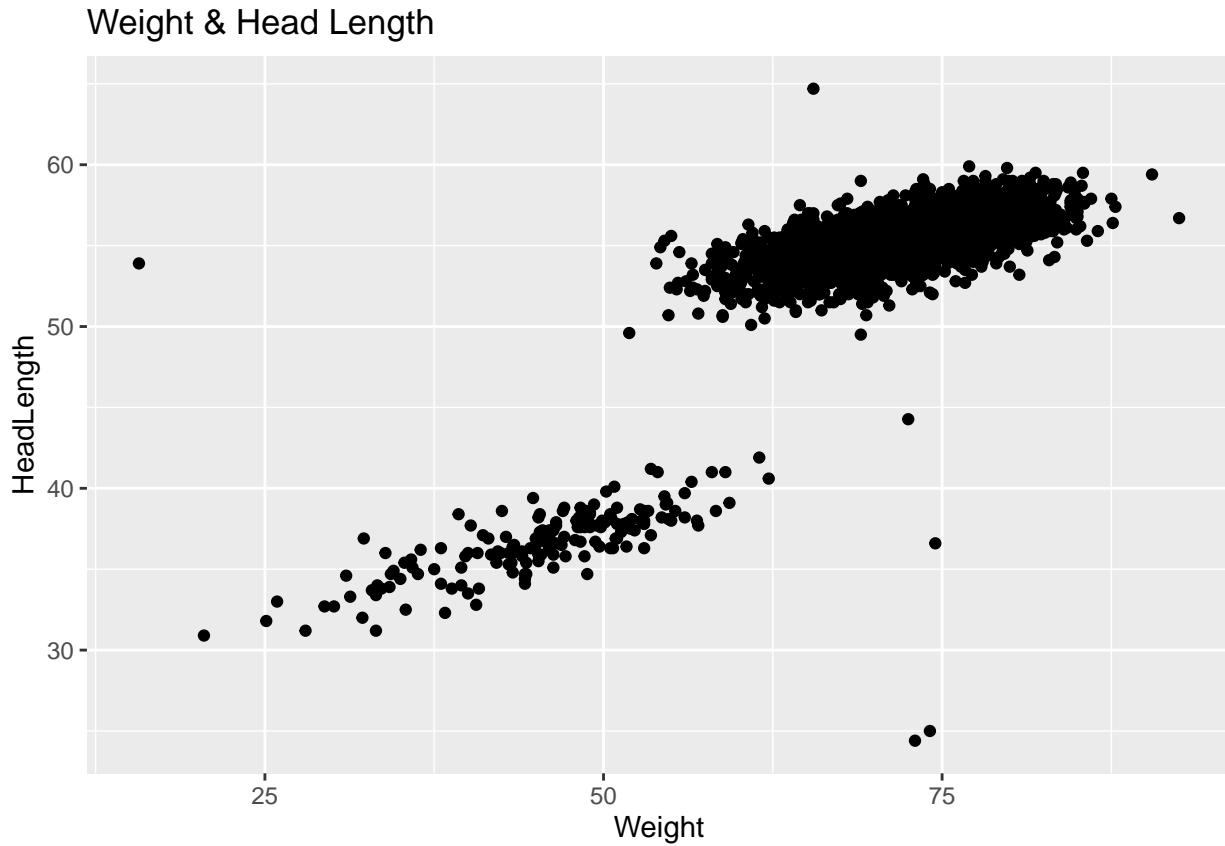


```
# There are 2,731 entries.
```

## Filtering Each Phenotype in Pairs

Filtering Weight and HeadLength:

```
WeightHL <- merge(WeightAge.Cleaned, HeadlengthAge_Cleaned, by = "ID") # Since every Jay has a unique ID  
##View(WeightHL)  
ggplot(data = WeightHL, aes(x = Weight, y = HeadLength)) + geom_point() + ggtitle("Weight & Head Length")
```



```
cor(WeightHL$Weight,WeightHL$HeadLength)
```

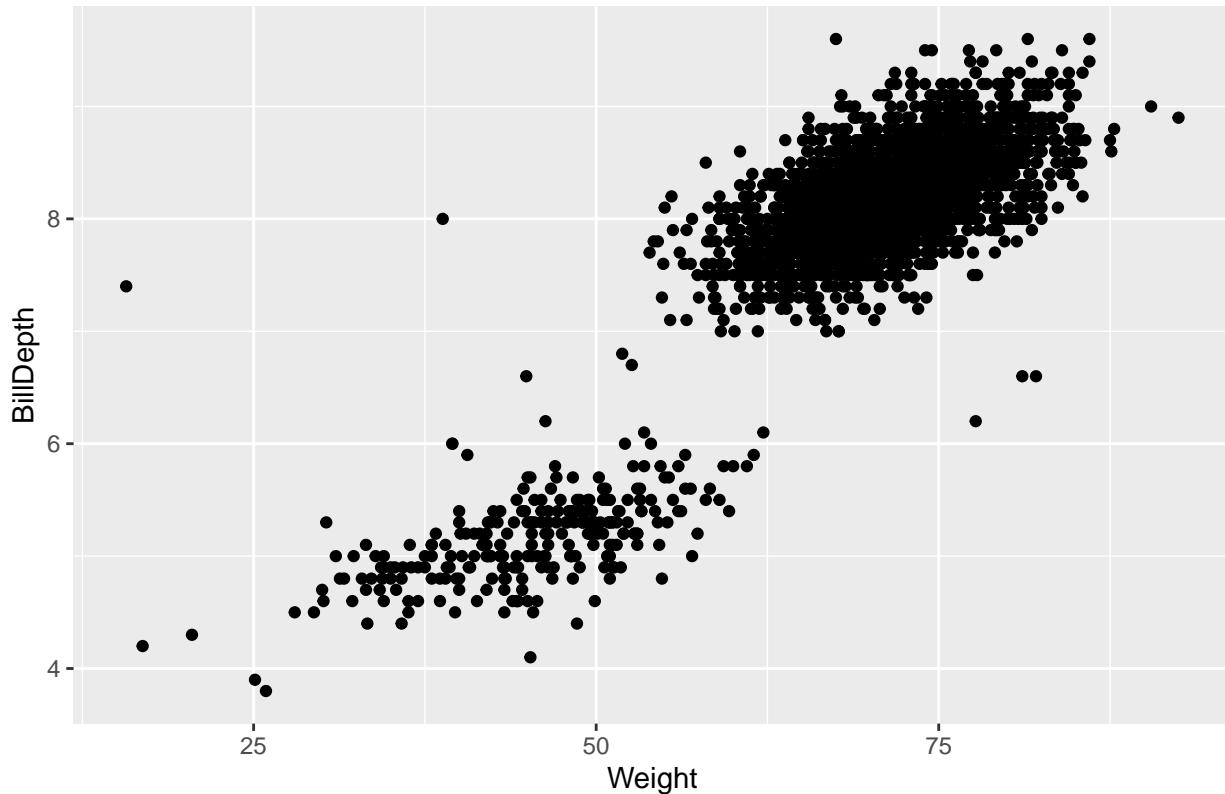
```
## [1] 0.8094054
```

```
# There are 2,693 entries.
```

Filtering Weight and BillDepth:

```
WeightBD <- merge(WeightAge.Cleaned, BillDepthAge.Cleaned, by = "ID") # Since every Jay has a unique ID  
##View(WeightBD)  
ggplot(data = WeightBD, aes(x = Weight, y = BillDepth)) + geom_point() + ggtitle("Weight & Bill Depth")
```

## Weight & Bill Depth



```
cor(WeightBD$Weight,WeightBD$BillDepth)
```

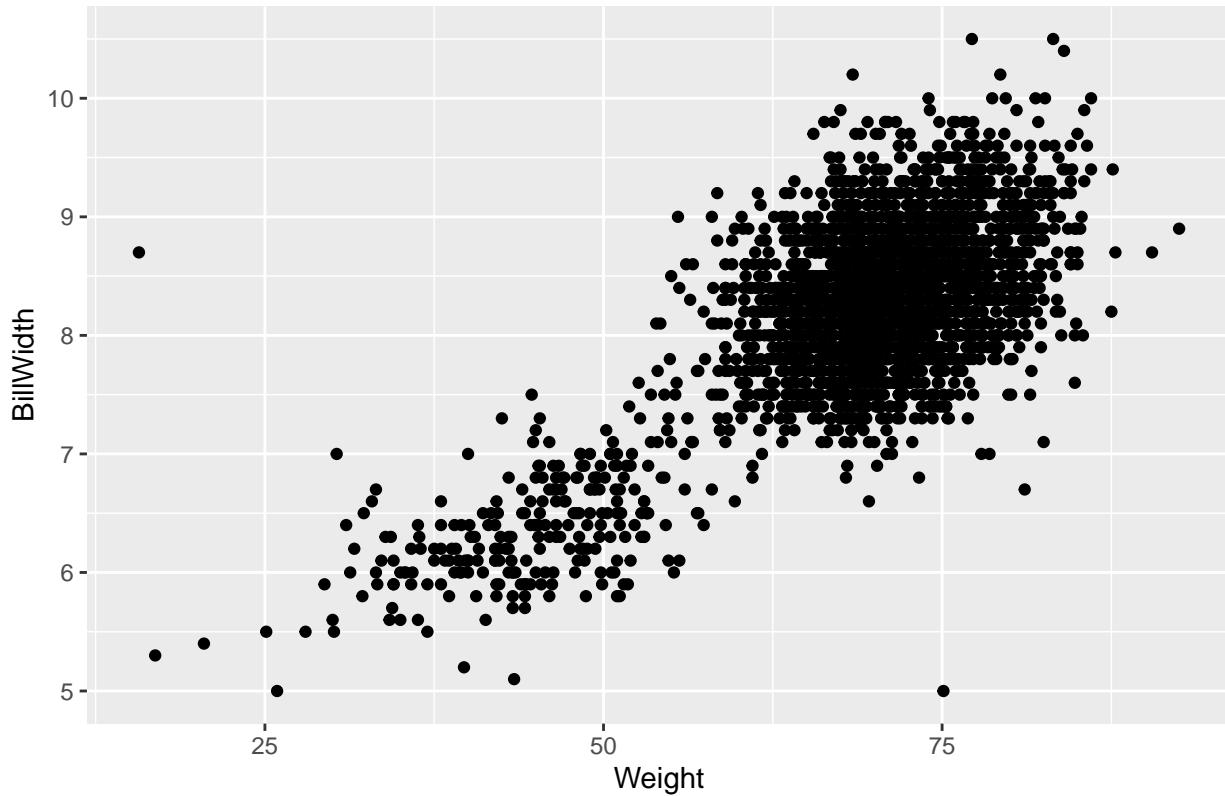
```
## [1] 0.8576739
```

```
# There are 2,743 entries.
```

Filtering Weight and Bill Width:

```
WeightBW <- merge(WeightAge.Cleaned, BillWidthAge_Cleaned, by = "ID") # Since every Jay has a unique ID
##View(WeightBW)
ggplot(data = WeightBW, aes(x = Weight, y = BillWidth)) + geom_point() + ggtitle("Weight & Bill Width")
```

## Weight & Bill Width



```
cor(WeightBW$Weight,WeightBW$BillWidth)
```

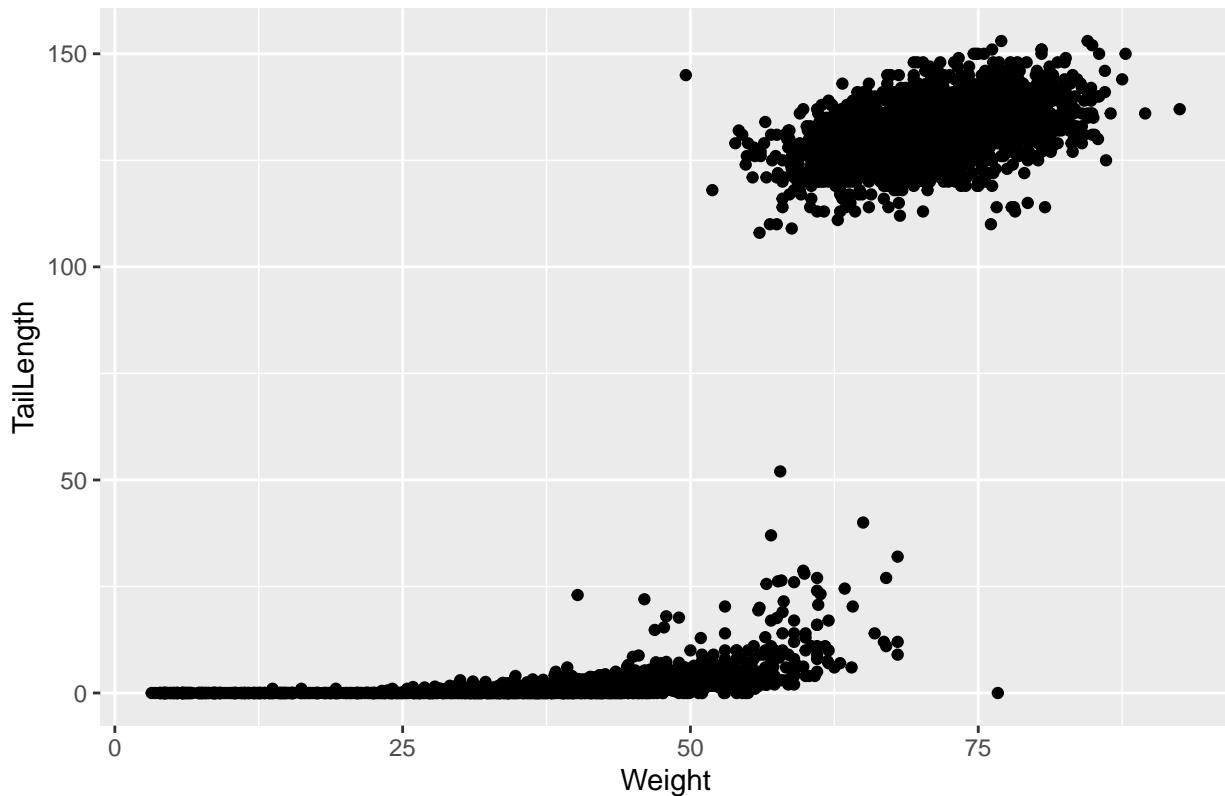
```
## [1] 0.7143414
```

```
# There are 2,742 entries.
```

Filtering Weight and Tail Length:

```
WeightTL <- merge(WeightAge.Cleaned, TailLengthAge_Cleaned, by = "ID") # Since every Jay has a unique ID
##View(WeightTL)
ggplot(data = WeightTL, aes(x = Weight, y = TailLength)) + geom_point() + ggttitle("Weight & Tail Length")
```

## Weight & Tail Length



```
cor(WeightTL$Weight,WeightTL$TailLength)
```

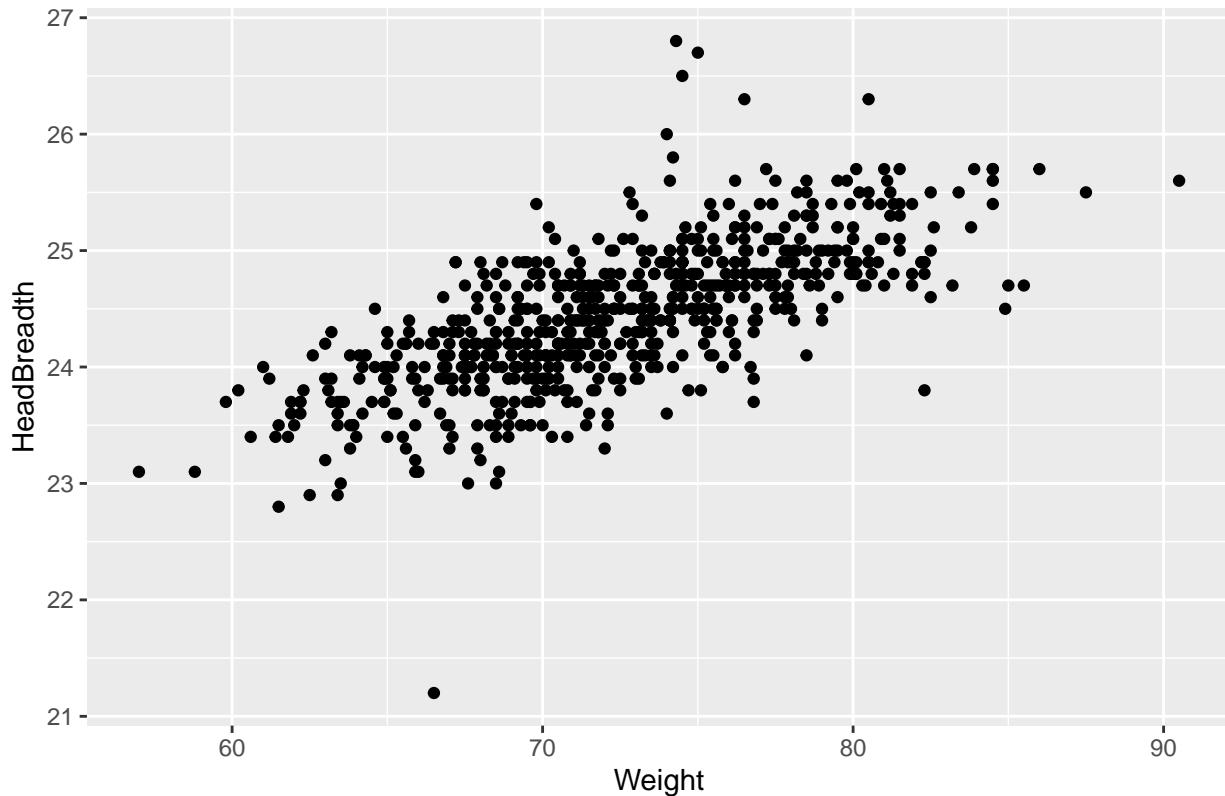
```
## [1] 0.8570235
```

```
# There are 4,415 entries.
```

Filtering Weight and Head Breadth:

```
WeightHB <- merge(WeightAge.Cleaned, HeadBreadthAge.Cleaned, by = "ID") # Since every Jay has a unique ID
##View(WeightHB)
ggplot(data = WeightHB, aes(x = Weight, y = HeadBreadth)) + geom_point() + ggtitle("Weight & Head Breadth")
```

## Weight & Head Breadth



```
cor(WeightHB$Weight,WeightHB$HeadBreadth)
```

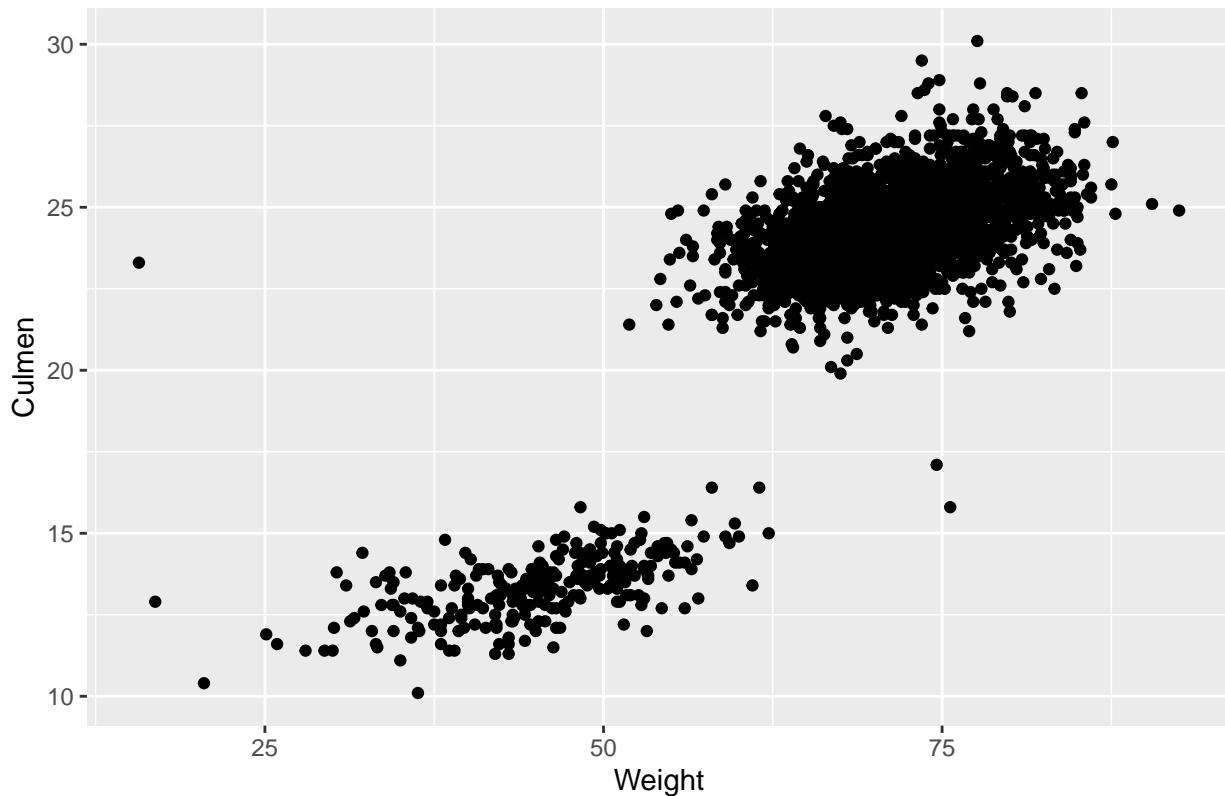
```
## [1] 0.7047766
```

```
# There are 691 entries.
```

Filtering Weight and Culmen:

```
WeightCulmen <- merge(WeightAge.Cleaned, Culmen_Cleaned, by = "ID") # Since every Jay has a unique ID i  
##View(WeightCulmen)  
ggplot(data = WeightCulmen, aes(x = Weight, y = Culmen)) + geom_point() + ggtitle("Weight & Culmen")
```

## Weight & Culmen



```
cor(WeightCulmen$Weight,WeightCulmen$Culmen)
```

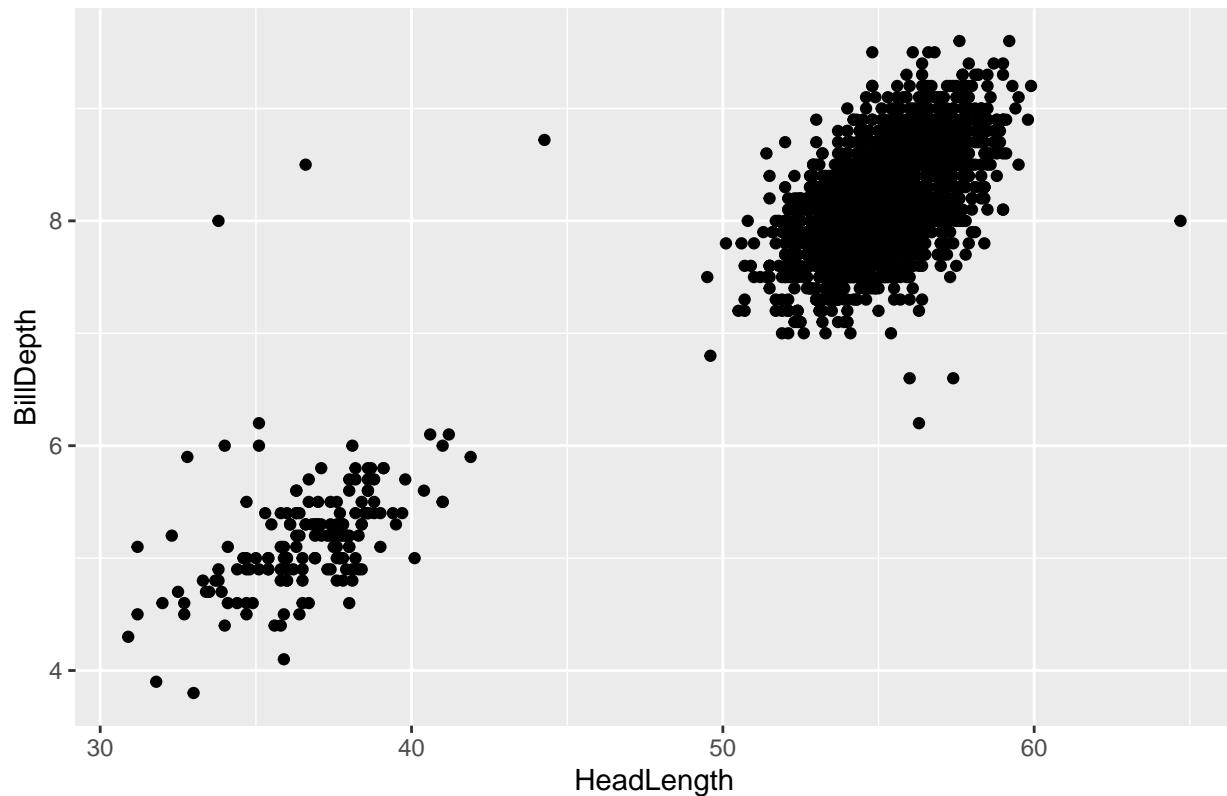
```
## [1] 0.836248
```

```
# There are 2,714 entries.
```

Filtering HeadLength and BillDepth:

```
HLBD <- merge(HeadlengthAge_Cleaned, BillDepthAge.Cleaned, by = "ID") # Since every Jay has a unique ID
##View(HLBD)
ggplot(data = HLBD, aes(x = HeadLength, y = BillDepth)) + geom_point() + ggtitle("Head Length & Bill Depth")
```

## Head Length & Bill Depth



```
cor(HLBD$HeadLength, HLBD$BillDepth)
```

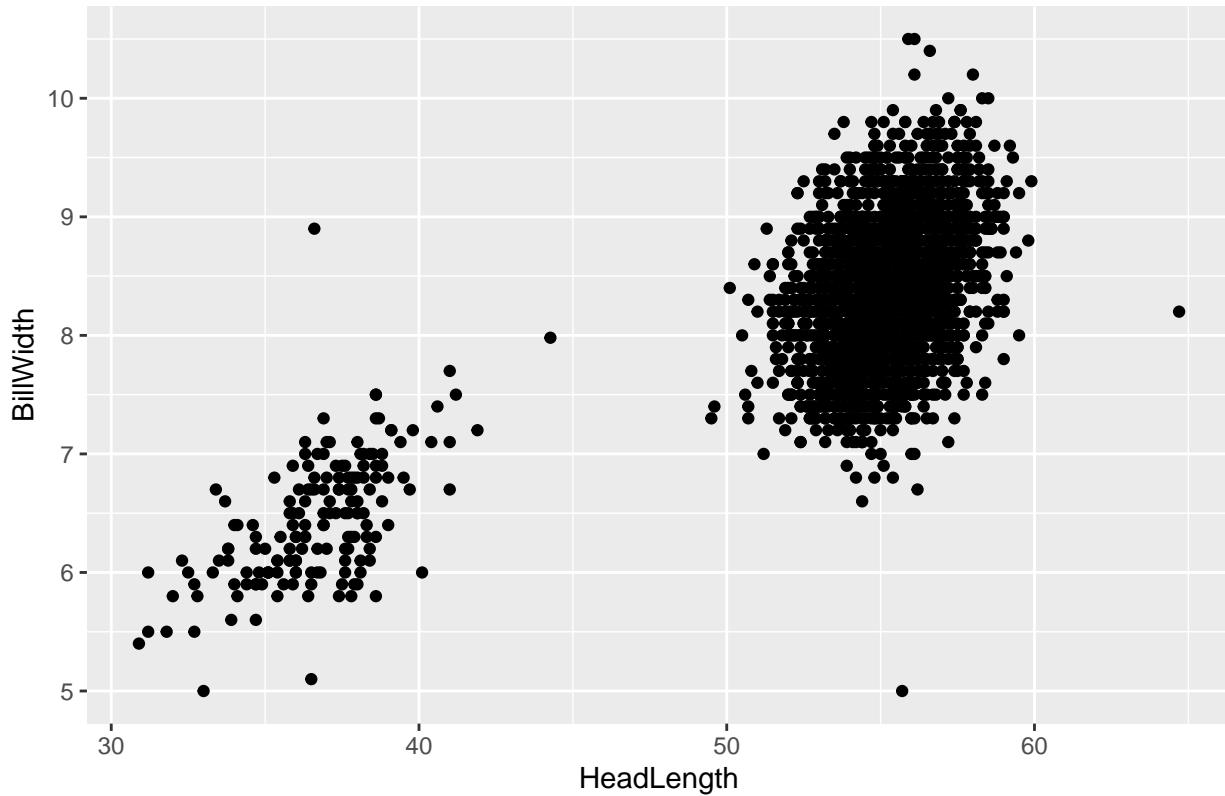
```
## [1] 0.8962257
```

```
# There are 2,624 entries.
```

Filtering HeadLength and BillWidth:

```
HLBW <- merge(HeadlengthAge_Cleaned, BillWidthAge_Cleaned, by = "ID") # Since every Jay has a unique ID
##View(HLBW)
ggplot(data = HLBW, aes(x = HeadLength, y = BillWidth)) + geom_point() + ggtitle("Head Length & Bill Width")
```

## Head Length & Bill Width



```
cor(HLBW$HeadLength, HLBW$BillWidth)
```

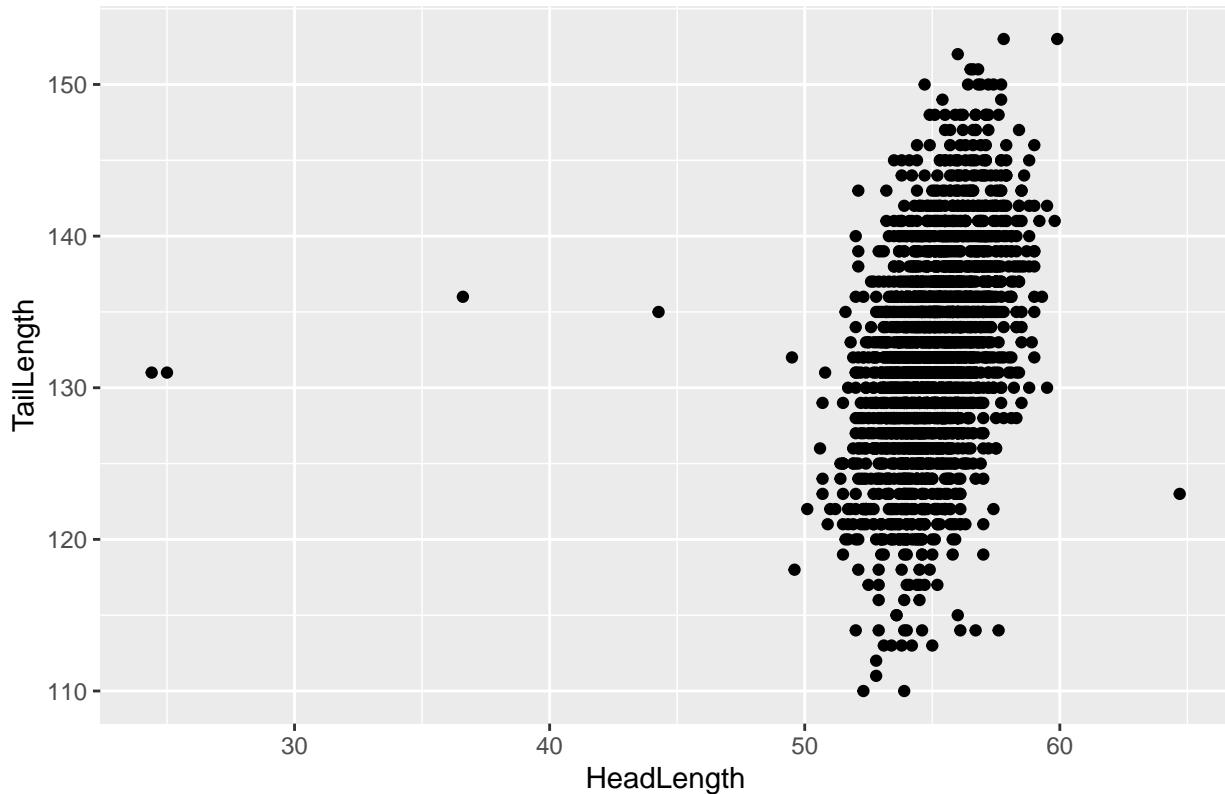
```
## [1] 0.6836355
```

```
# There are 2,623 entries.
```

Filtering HeadLength and TailLength:

```
HLTL <- merge(HeadlengthAge_Cleaned, TailLengthAge_Cleaned, by = "ID") # Since every Jay has a unique ID
##View(HLTL)
ggplot(data = HLTL, aes(x = HeadLength, y = TailLength)) + geom_point() + ggttitle("Head Length & TailLength")
```

## Head Length & TailLength



```
cor(HLTL$HeadLength, HLTL$TailLength)
```

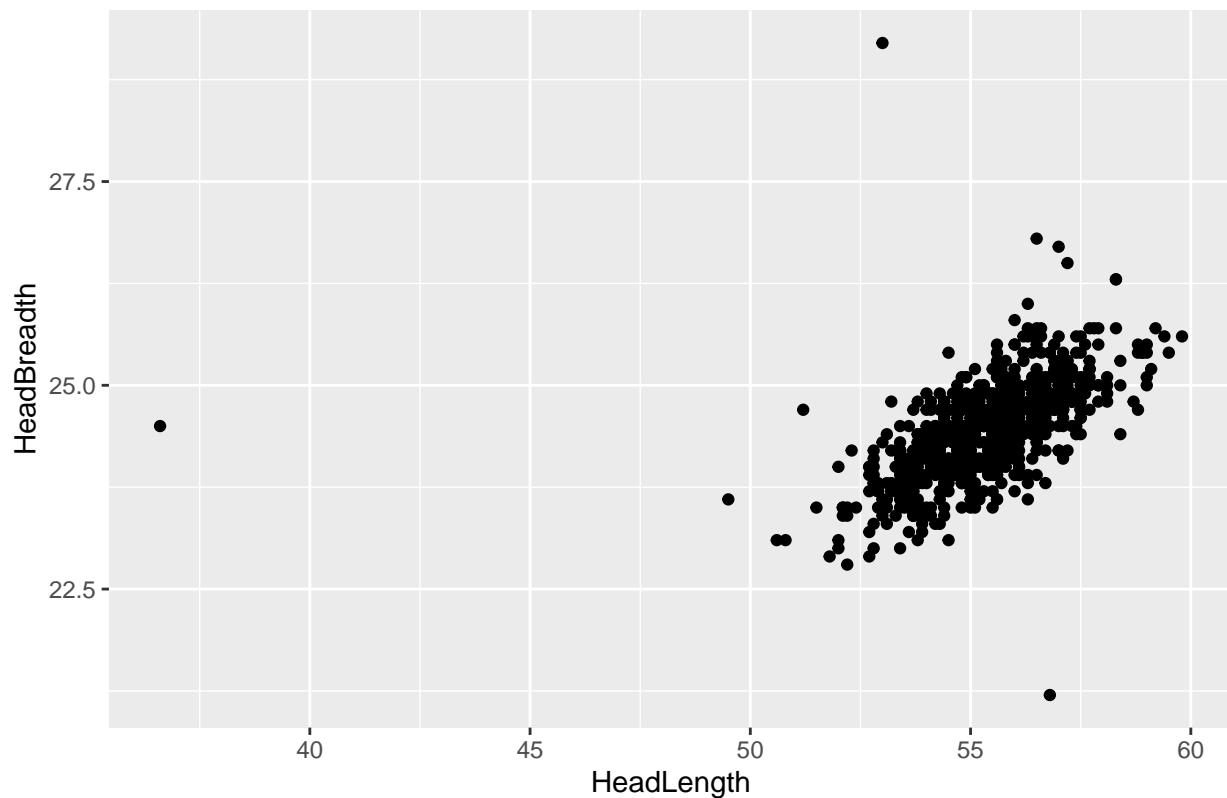
```
## [1] 0.3701413
```

```
# There are 2,218 entries.
```

Filtering HeadLength and Head Breadth:

```
HLHB <- merge(HeadlengthAge_Cleaned, HeadBreadthAge.Cleaned, by = "ID") # Since every Jay has a unique ID
##View(HLHB)
ggplot(data = HLHB, aes(x = HeadLength, y = HeadBreadth)) + geom_point() + ggttitle("Head Length & Head Breadth")
```

## Head Length & Head Breadth



```
cor(HLHB$HeadLength, HLHB$HeadBreadth)
```

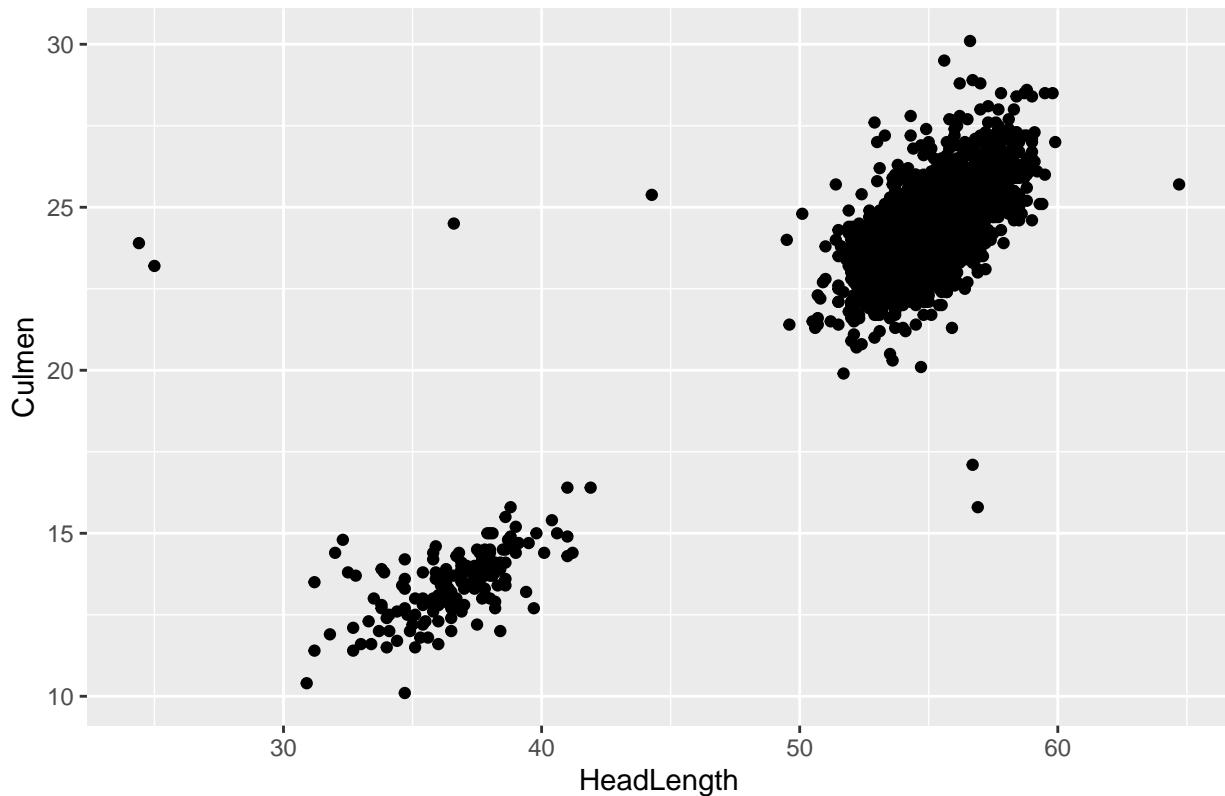
```
## [1] 0.5664638
```

```
# There are 694 entries.
```

Filtering Head Length and Culmen:

```
HLCulmen <- merge(HeadlengthAge_Cleaned, Culmen_Cleaned, by = "ID") # Since every Jay has a unique ID i  
##View(HLCulmen)  
ggplot(data = HLCulmen, aes(x = HeadLength, y = Culmen)) + geom_point() + ggtitle("Head Length & Culmen")
```

## Head Length & Culmen



```
cor(HLCulmen$HeadLength, HLCulmen$Culmen)
```

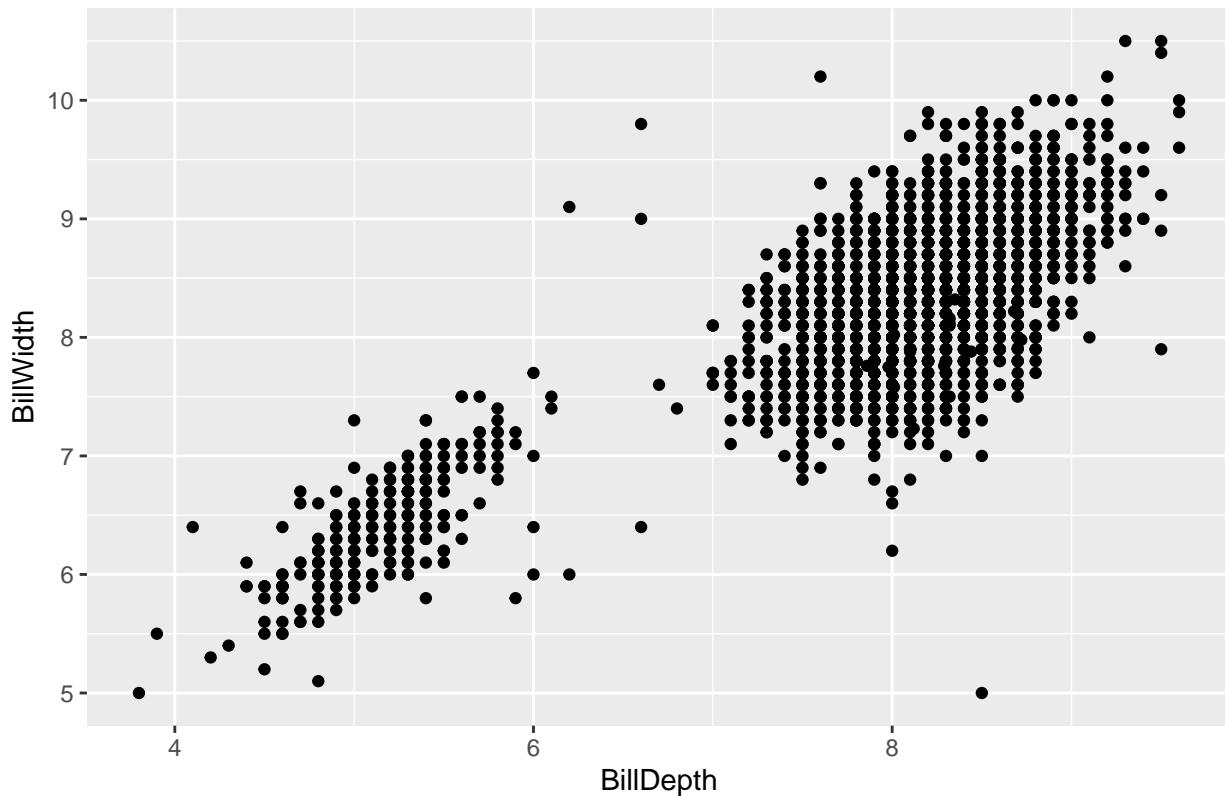
```
## [1] 0.9230694
```

```
# There are 2,597 entries.
```

Filtering Bill Depth and Bill Width:

```
BDBW <- merge(BillDepthAge.Cleaned, BillWidthAge_Cleaned, by = "ID") # Since every Jay has a unique ID
##View(BDBW)
ggplot(data = BDBW, aes(x = BillDepth, y = BillWidth)) + geom_point() + ggttitle("Bill Depth & Bill Width")
```

## Bill Depth & Bill Width



```
cor(BDBW$BillDepth,BDBW$BillWidth)
```

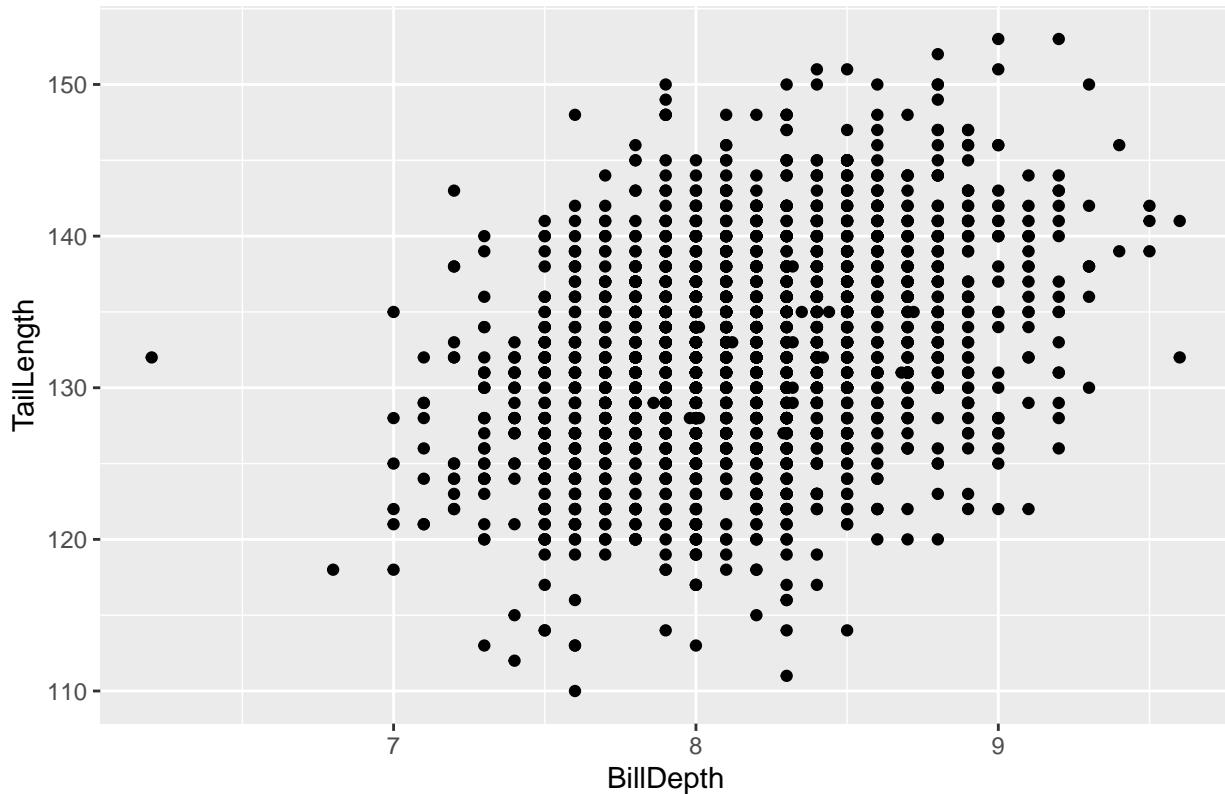
```
## [1] 0.8002042
```

```
# There are 2,757 entries.
```

Filtering Bill Depth and Tail Length:

```
BDTL <- merge(BillDepthAge.Cleaned, TailLengthAge_Cleaned, by = "ID") # Since every Jay has a unique ID
##View(BDTL)
ggplot(data = BDTL, aes(x = BillDepth, y = TailLength)) + geom_point() + ggtitle("Bill Depth & Tail Length")
```

## Bill Depth & Tail Length



```
cor(BDTL$BillDepth,BDTL$TailLength)
```

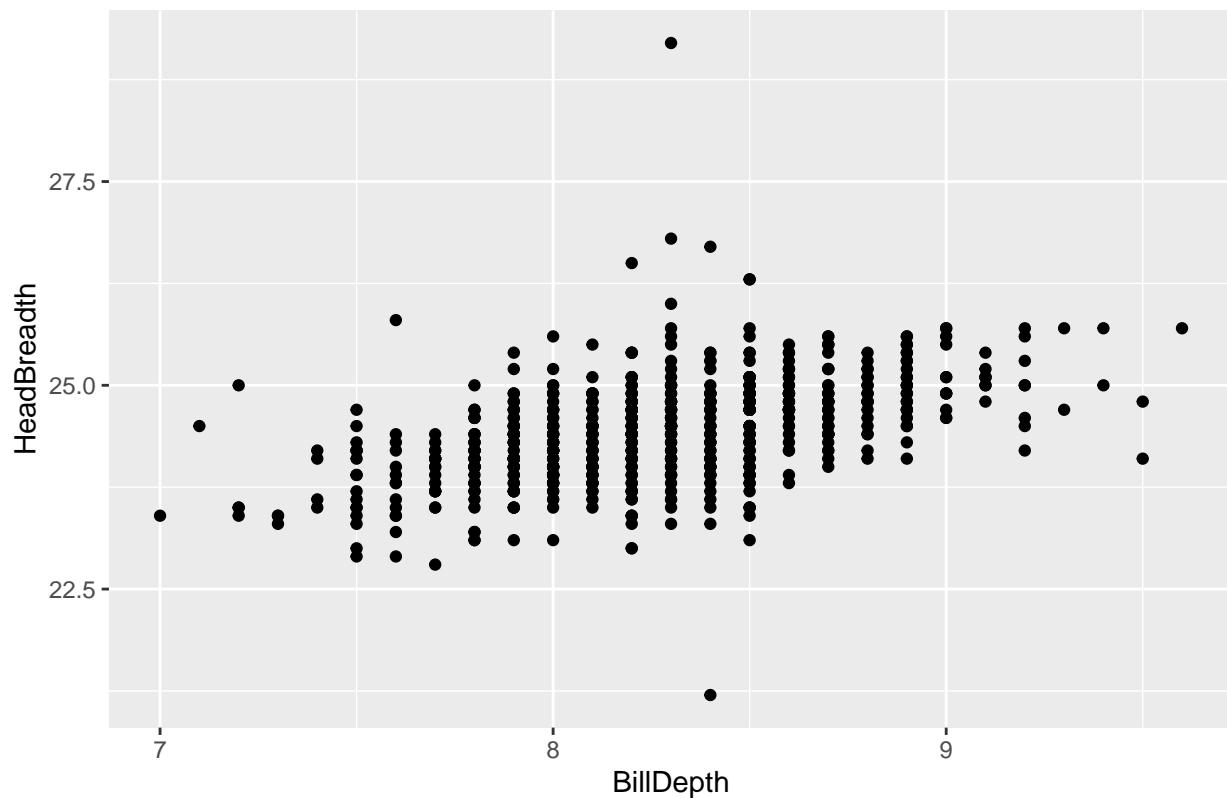
```
## [1] 0.340756
```

```
# There are 2,171 entries.
```

Filtering Bill Depth and Head Breadth:

```
BDHB <- merge(BillDepthAge.Cleaned, HeadBreadthAge.Cleaned, by = "ID") # Since every Jay has a unique ID
##View(BDHB)
ggplot(data = BDHB, aes(x = BillDepth, y = HeadBreadth)) + geom_point() + ggttitle("Bill Depth & Head Breadth")
```

## Bill Depth & Head Breadth



```
cor(BDHB$BillDepth,BDHB$HeadBreadth)
```

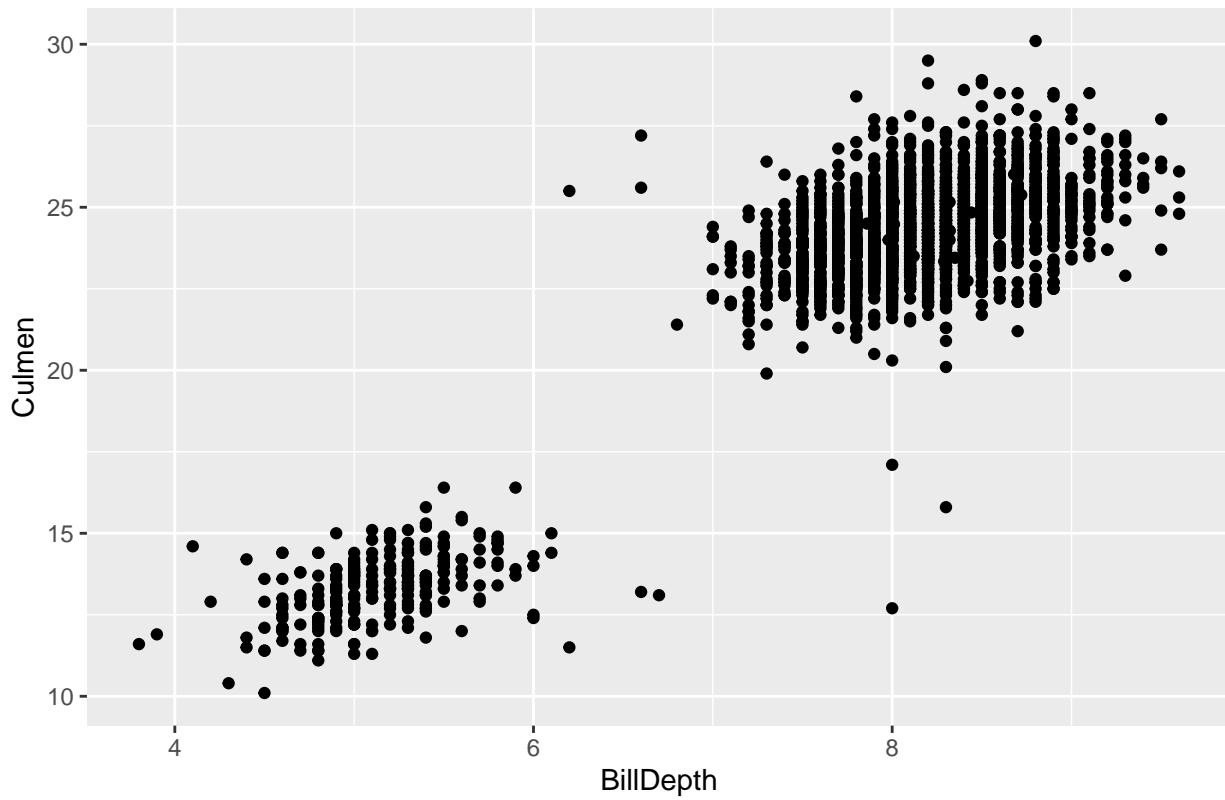
```
## [1] 0.4971986
```

```
# There are 690 entries.
```

Filtering Bill Depth and Culmen:

```
BDCulmen <- merge(BillDepthAge.Cleaned, Culmen_Cleaned, by = "ID") # Since every Jay has a unique ID in  
##View(BDCulmen)  
ggplot(data = BDCulmen, aes(x = BillDepth, y = Culmen)) + geom_point() + ggtitle("Bill Depth & Culmen")
```

## Bill Depth & Culmen



```
cor(BDCulmen$BillDepth,BDCulmen$Culmen)
```

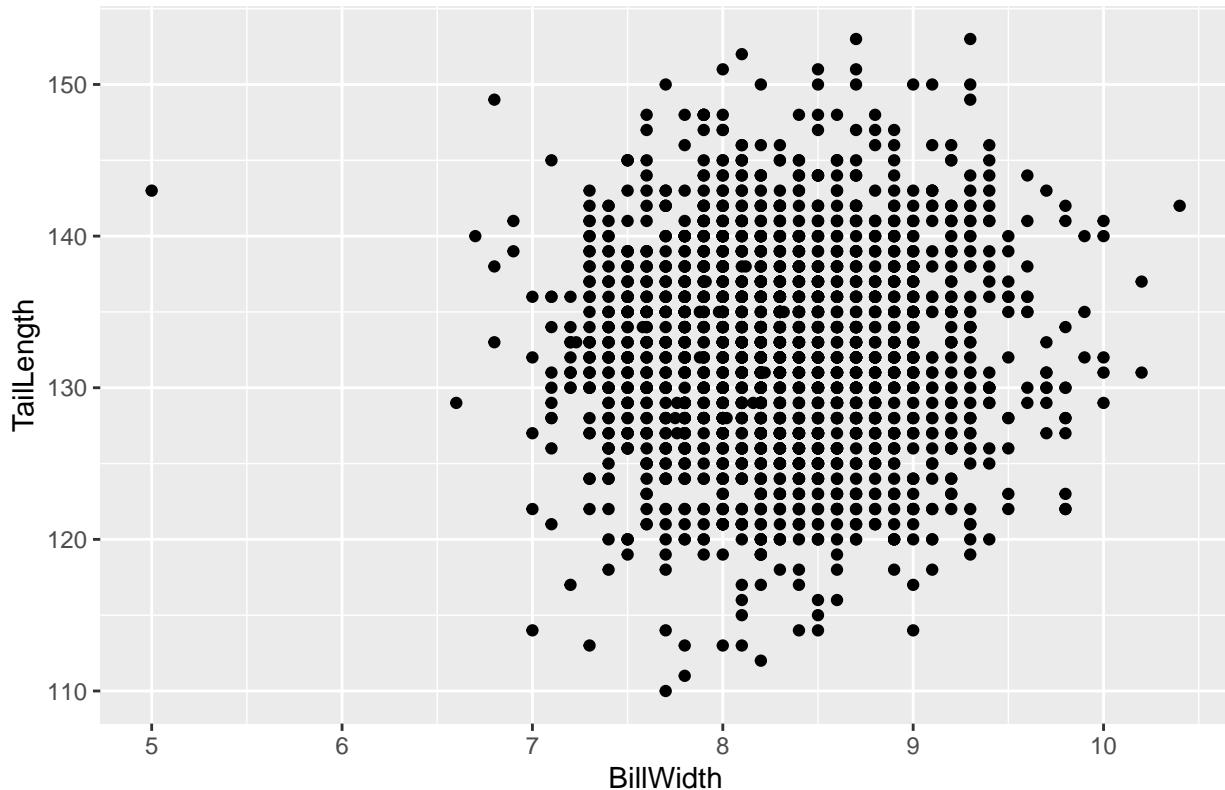
```
## [1] 0.9052125
```

```
# There are 2,698 entries.
```

Filtering Bill Width and Tail Length:

```
BWTL <- merge(BillWidthAge_Cleaned, TailLengthAge_Cleaned, by = "ID") # Since every Jay has a unique ID
##View(BWTL)
ggplot(data = BWTL, aes(x = BillWidth, y = TailLength)) + geom_point() + ggtitle("Bill Width & Tail Length")
```

## Bill Width & Tail Length



```
cor(BWTL$BillWidth, BWTL$TailLength)
```

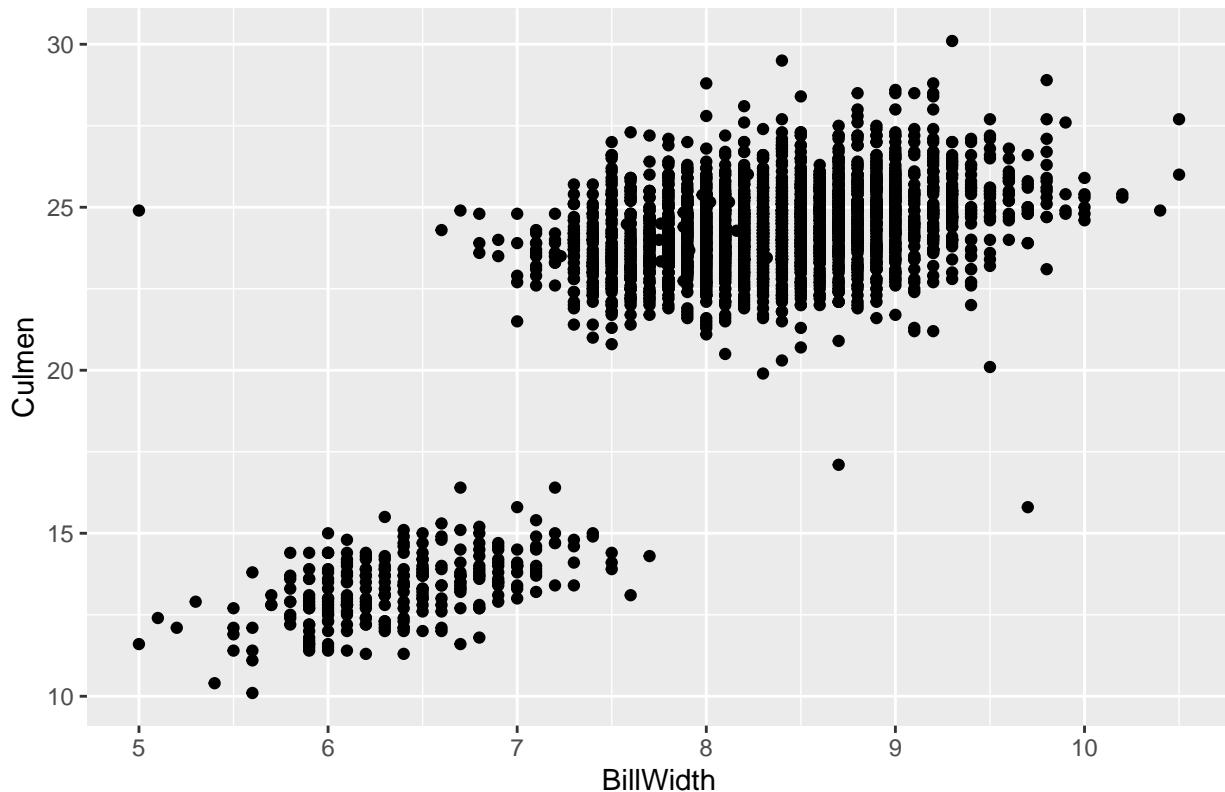
```
## [1] 0.01333621
```

```
# There are 2,170 entries.
```

Filtering Bill Width and Culmen:

```
BWCulmen <- merge(BillWidthAge_Cleaned, Culmen_Cleaned, by = "ID") # Since every Jay has a unique ID in  
##View(BWCulmen)  
ggplot(data = BWCulmen, aes(x = BillWidth, y = Culmen)) + geom_point() + ggtitle("Bill Width & Culmen")
```

## Bill Width & Culmen



```
cor(BWCulmen$BillWidth,BWCulmen$Culmen)
```

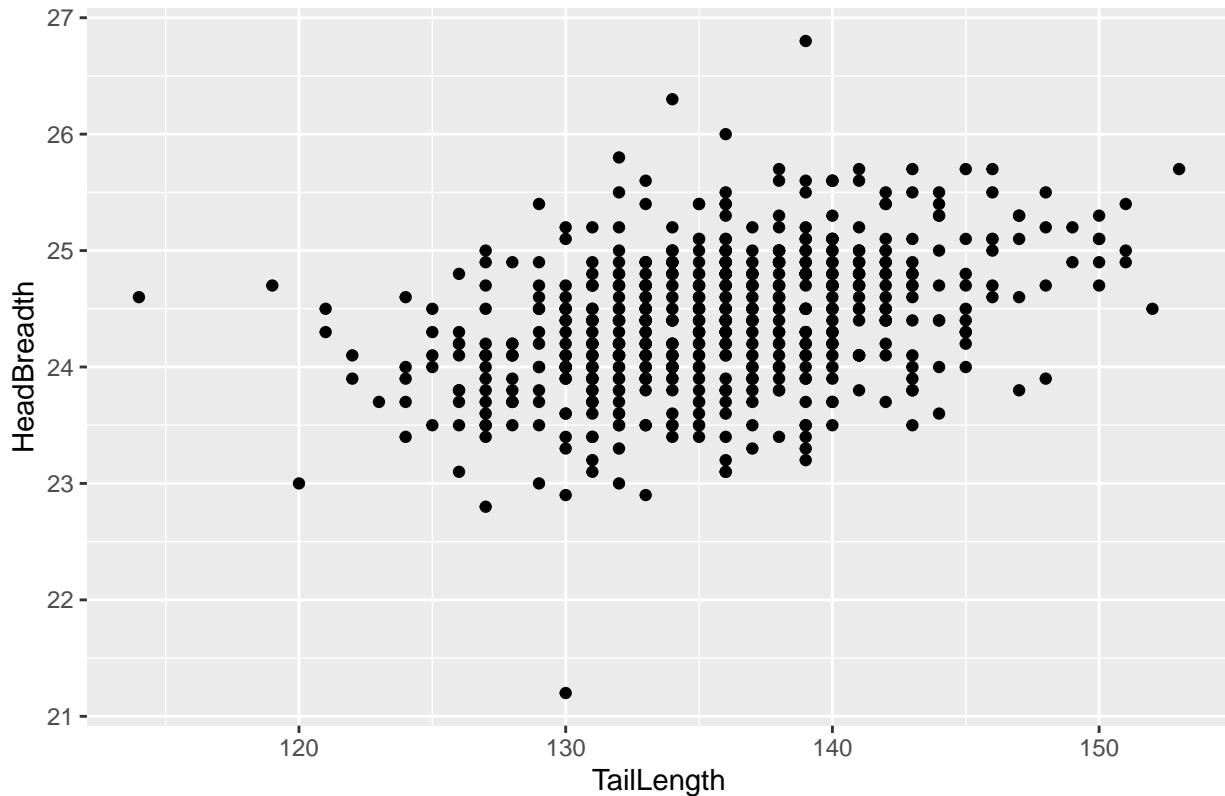
```
## [1] 0.7464234
```

```
# There are 2,697 entries.
```

Filtering Tail Length and Head Breadth:

```
TLHB <- merge(TailLengthAge_Cleaned, HeadBreadthAge.Cleaned, by = "ID") # Since every Jay has a unique ID
##View(TLHB)
ggplot(data = TLHB, aes(x = TailLength, y = HeadBreadth)) + geom_point() + ggtitle("Tail Length & Head Breadth")
```

## Tail Length & Head Breadth



```
cor(TLHB$TailLength, TLHB$HeadBreadth)
```

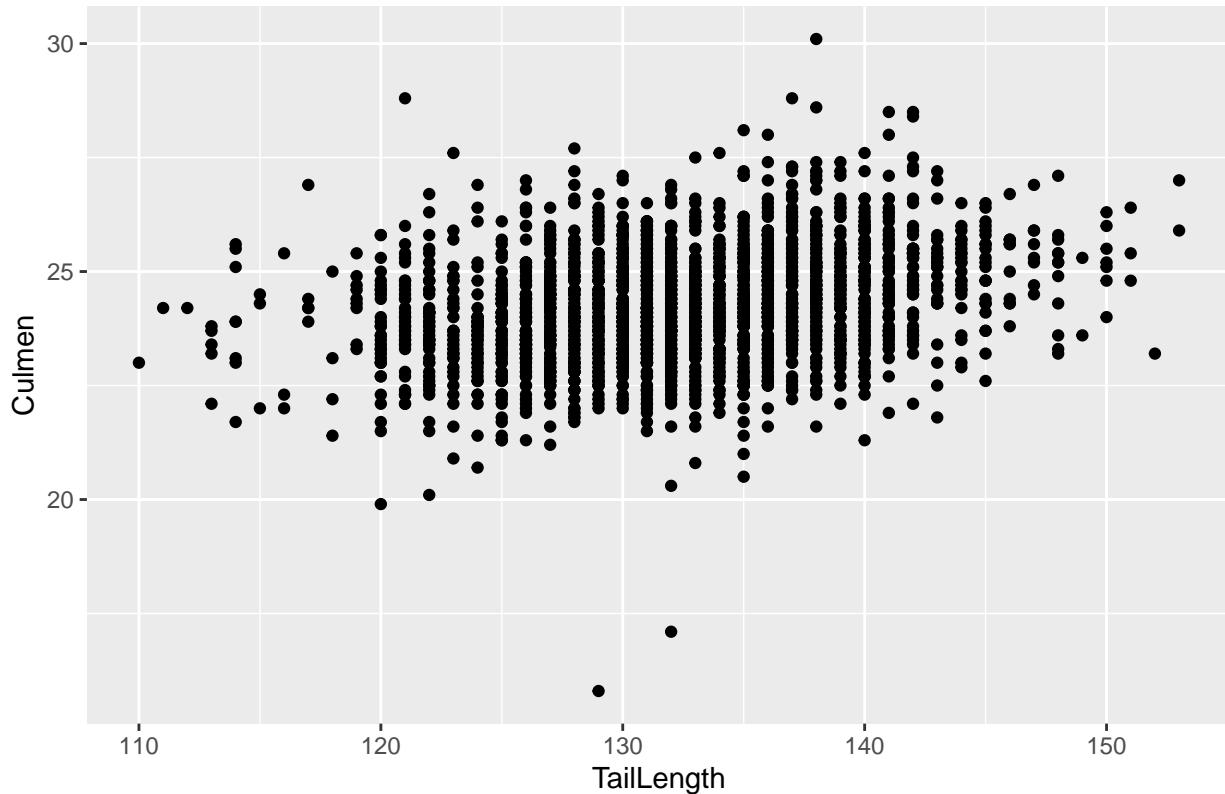
```
## [1] 0.3945056
```

```
# There are 624 entries.
```

Filtering Tail Length and Culmen:

```
TCulmen <- merge(TailLengthAge_Cleaned, Culmen_Cleaned, by = "ID") # Since every Jay has a unique ID i  
##View(TCulmen)  
ggplot(data = TCulmen, aes(x = TailLength, y = Culmen)) + geom_point() + gtitle("Tail Length & Culmen")
```

## Tail Length & Culmen



```
cor(TLCulmen$TailLength, TLCulmen$Culmen)
```

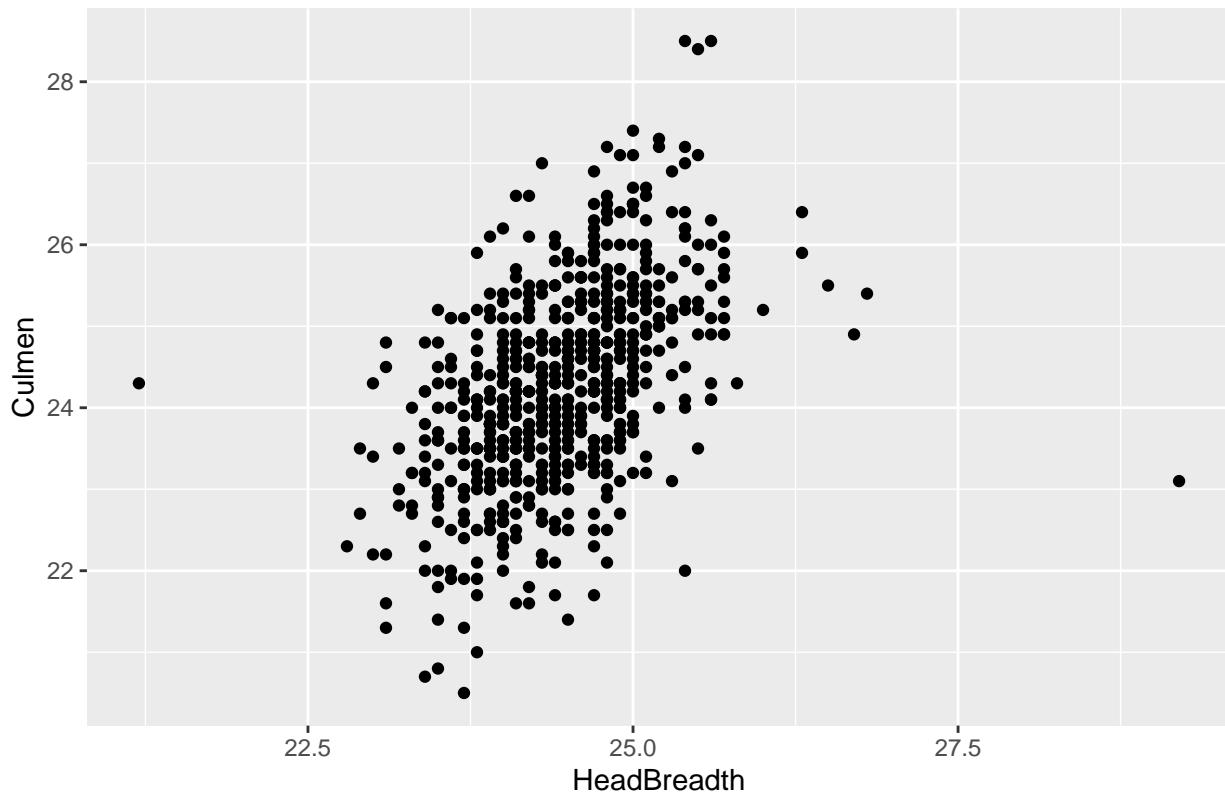
```
## [1] 0.2861985
```

```
# There are 2,149 entries.
```

Filtering Head Breadth and Culmen:

```
HBCulmen <- merge(HeadBreadthAge.Cleaned, Culmen_Cleaned, by = "ID") # Since every Jay has a unique ID
##View(HBCulmen)
ggplot(data = HBCulmen, aes(x = HeadBreadth, y = Culmen)) + geom_point() + ggtitle("Tail Length & Culmen")
```

## Tail Length & Culmen



```
cor(HBCulmen$HeadBreadth,HBCulmen$Culmen)
```

```
## [1] 0.4819437
```

```
# There are 693 entries.
```

The procedure to ensure that we are controlling for birds only in demo and not in the peripheral territories is as follows (in psuedocode): Go to the list all.breeders, Ensure that the following columns have loaded properly: ID, TerrYr. This is to make sure that the ID of the Jay can be matched with the territory and Year and thus the data can be properly filtered. The last 2 digits of TerrYr are the years the data was collected in.

TLDR: all.breeders->filter out ID, and TerrYr column. Use merge function to merge the this new list with each phenotype using ID as an argument.

```
PerTer<-all.breeders[, c('TerrYr', 'ID', 'Tract', 'Year', 'Terr')] #Make a new list of just these columns
##View(PerTer) #This has some "South" in it. Lets remove that. Also there are no PLEs in this.
PerTer_Cleaned<-na.omit(PerTer) # Removes all NA
##View(PerTer_Cleaned)
PerTer_Cleaned1<-PerTer[PerTer$Tract != "South",] # Removes all South tracts.
##View(PerTer_Cleaned1)
```

**Now to ensure that we are controlling for the correct year because Demo study tract stabilised around 1990.**

```
PerTer_Year<-PerTer_Cleaned1[PerTer_Cleaned1$Year >= "1990",]
###View(PerTer_Year)
```

Importing a list of peripheral territories to be removed. This list will be done against our PerTer\_Year and will remove any matches from the Terr column.

```
#toremovePT<-read.csv("/cloud/project/PTNancy1.csv")
##View(toremovePT)
remove <- c("ATNT", "BEAR", "CHGR", "CHIK", "CIRB", "CIRC", "COTG", "EAGB", "EAGR", "FTZP", "GRVR", "GR
finalphenotype<-PerTer_Year[!PerTer_Year$Terr %in% remove,] #Removal sucessful!
##View(finalphenotype)
```

Now hopefully, this can be merged with our phenotypes, and although our number will decrease, we will get controlled data.

Trying to merge weight with the new controlled `for` periphery data along with marking the age of the Jays

```
#File name: WeightAge.Cleaned
#11,750 entries

WeightMerge<-merge(WeightAge.Cleaned, finalphenotype, by = "ID")
##View(WeightMerge)
WeightMerge$AgeCategory<-99 #adding a new column to prepare for categorical ages such as juveniles, yearlings
WeightMerge$AgeCategory[WeightMerge$AgeMeas<=11]<- "Nestling"
WeightMerge$AgeCategory[11<WeightMerge$AgeMeas & WeightMerge$AgeMeas<=70]<- "Fledgling"
WeightMerge$AgeCategory[70<WeightMerge$AgeMeas & WeightMerge$AgeMeas<=365]<- "Juveniles"
WeightMerge$AgeCategory[365<WeightMerge$AgeMeas & WeightMerge$AgeMeas<=3650]<- "Yearling"
```

*#The number of entries dropped to 1,524.*

```
# Covariates for weight:
# 1) All other morphometrical dimensions especially lengths.
```

```
#File name: Headlength_Cleaned
#2711 entries

HLMerge<-merge(HeadlengthAge_Cleaned, finalphenotype, by = "ID")
##View(HLMerge)
HLMerge$AgeCategory<-99 #adding a new column to prepare for categorical ages such as juveniles, yearlings
HLMerge$AgeCategory[HLMerge$AgeMeas<=11]<- "Nestling"
HLMerge$AgeCategory[11<HLMerge$AgeMeas & HLMerge$AgeMeas<=70]<- "Fledgling"
HLMerge$AgeCategory[70<HLMerge$AgeMeas & HLMerge$AgeMeas<=365]<- "Juveniles"
HLMerge$AgeCategory[365<HLMerge$AgeMeas & HLMerge$AgeMeas<=3650]<- "Yearling"
```

*#The number of entries dropped to 473.*

```
# Covariates for Head Length:
# 1) All other morphometrical dimensions especially all the other lengths.
```

```
#File name: BillDepthAge.Cleaned
#2760 entries

BDMerge<-merge(BillDepthAge.Cleaned, finalphenotype, by = "ID")
##View(BDMerge)
HLMerge$AgeCategory<-99 #adding a new column to prepare for categorical ages such as juveniles, yearlings
```

```

BDMerge$AgeCategory[BDMerge$AgeMeas<=11] <- "Nestling"
BDMerge$AgeCategory[11<BDMerge$AgeMeas & BDMerge$AgeMeas<=70] <- "Fledgling"
BDMerge$AgeCategory[70<BDMerge$AgeMeas & BDMerge$AgeMeas<=365] <- "Juveniles"
BDMerge$AgeCategory[365<BDMerge$AgeMeas & BDMerge$AgeMeas<=3650] <- "Yearling"

```

*#The number of entries dropped to 471.*

*# Covariates for Bill Depth:*

*# 1) All other morphometrical dimensions especially all the other lengths.*

```

#File name: BillWidthAge_Cleaned
#2758 entries

```

```

BWMerge<-merge(BillWidthAge_Cleaned, finalphenotype, by = "ID")
##View(BWMerge)

```

*HLMerge\$AgeCategory<-99 #adding a new column to prepare for categorical ages such as juveniles, yearlings, nestlings*

```

BWMerge$AgeCategory[BWMerge$AgeMeas<=11] <- "Nestling"

```

```

BWMerge$AgeCategory[11<BWMerge$AgeMeas & BWMerge$AgeMeas<=70] <- "Fledgling"

```

```

BWMerge$AgeCategory[70<BWMerge$AgeMeas & BWMerge$AgeMeas<=365] <- "Juveniles"

```

```

BWMerge$AgeCategory[365<BWMerge$AgeMeas & BWMerge$AgeMeas<=3650] <- "Yearling"

```

*#The number of entries dropped to 471.*

*# Covariates for Bill Width:*

*# 1) All other morphometrical dimensions especially all the other lengths.*

```

#File name: TailLengthAge_Cleaned
#4429 entries

```

```

TLMerge<-merge(TailLengthAge_Cleaned, finalphenotype, by = "ID")
##View(TLMerge)

```

*TLMerge\$AgeCategory<-99 #adding a new column to prepare for categorical ages such as juveniles, yearlings, nestlings*

```

TLMerge$AgeCategory[TLMerge$AgeMeas<=11] <- "Nestling"

```

```

TLMerge$AgeCategory[11<TLMerge$AgeMeas & TLMerge$AgeMeas<=70] <- "Fledgling"

```

```

TLMerge$AgeCategory[70<TLMerge$AgeMeas & TLMerge$AgeMeas<=365] <- "Juveniles"

```

```

TLMerge$AgeCategory[365<TLMerge$AgeMeas & TLMerge$AgeMeas<=3650] <- "Yearling"

```

*#The number of entries dropped to 569*

*# Covariates for Tail length:*

*# 1) All other morphometrical dimensions especially all the other lengths.*

```

#File name: HeadBreadthAge.Cleaned
#694 entries

```

```

# HBMerge<-merge(HeadBreadthAge.Cleaned, finalphenotype, by = "ID")
# ##View(HBMerge)

```

*# HBMerge\$AgeCategory<-99 #adding a new column to prepare for categorical ages such as juveniles, yearlings, nestlings*

```

# HBMerge$AgeCategory[HBMerge$AgeMeas<=11] <- "Nestling"

```

```

# HBMerge$AgeCategory[11<HBMerge$AgeMeas & HBMerge$AgeMeas<=70] <- "Fledgling"

```

```

# HBMerge$AgeCategory[70<HBMerge$AgeMeas & HBMerge$AgeMeas<=365] <- "Juveniles"

```

```

# HBMerge$AgeCategory[365<HBMerge$AgeMeas & HBMerge$AgeMeas<=730] <- "Yearling"

```

```

# HBMerge$AgeCategory[730<HBMerge$AgeMeas & HBMerge$AgeMeas<=1095] <- "Second Yearling"

```

```

# HBMerge$AgeCategory[1095<HBMerge$AgeMeas & HBMerge$AgeMeas<=1460] <- "Third Yearling"

```

```

# HBMerge$AgeCategory[1460<HBMerge$AgeMeas & HBMerge$AgeMeas<=1825] <- "Fourth Yearling"
# HBMerge$AgeCategory[1825<HBMerge$AgeMeas & HBMerge$AgeMeas<=2190] <- "Fifth Yearling"
# HBMerge$AgeCategory[1825<HBMerge$AgeMeas & HBMerge$AgeMeas<=2190] <- "Fifth Yearling"
# HBMerge$AgeCategory[2190<HBMerge$AgeMeas & HBMerge$AgeMeas<=2555] <- "Sixth Yearling"
# HBMerge$AgeCategory[2555<HBMerge$AgeMeas & HBMerge$AgeMeas<=2920] <- "Seventh Yearling"
# HBMerge$AgeCategory[2920<HBMerge$AgeMeas & HBMerge$AgeMeas<=3285] <- "Eighth Yearling"
# HBMerge$AgeCategory[3285<HBMerge$AgeMeas & HBMerge$AgeMeas<=3650] <- "Ninth Yearling"

#No entries because IDs don't intersect.

# Covariates for Head Breadth:
# 1) All other morphometrical dimensions especially all the other lengths.

#File name: Culmen_Cleaned
#694 entries

CMerge<-merge(Culmen_Cleaned, finalphenotype, by = "ID")
##View(CMerge)
CMerge$AgeCategory<-99 #adding a new column to prepare for categorical ages such as juveniles, yearling
CMerge$AgeCategory[CMerge$AgeMeas<=11] <- "Nestling"
CMerge$AgeCategory[11<CMerge$AgeMeas & CMerge$AgeMeas<=70] <- "Fledgling"
CMerge$AgeCategory[70<CMerge$AgeMeas & CMerge$AgeMeas<=365] <- "Juveniles"
CMerge$AgeCategory[365<CMerge$AgeMeas & CMerge$AgeMeas<=3650] <- "Yearling"

#473 enteries

# Covariates for Head Breadth:
# 1) All other morphometrical dimensions especially all the other lengths.

```

---

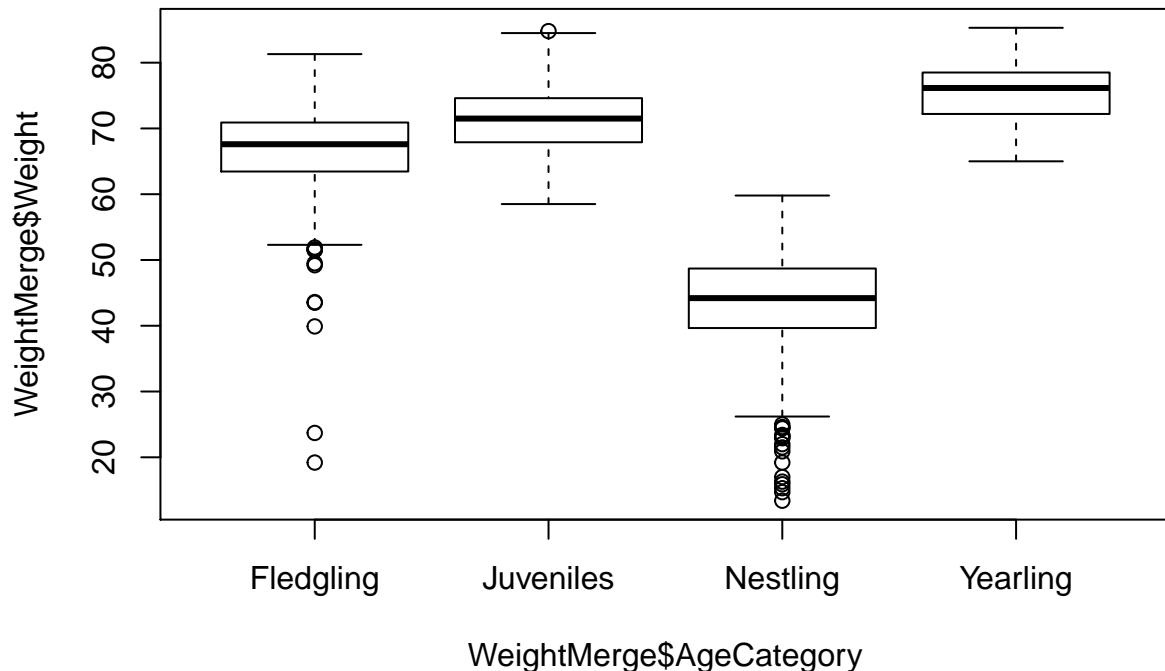
## Mann-Kendall Tests On The Phenotypes

I am confused here: after my summer project, one of the feedbacks I got was that I should control for age when determining the trends in phenotypes. This makes sense because we don't have repeated readings of the same birds and some birds would be bigger and some would be smaller, making for an automatic trend but with age as a confounding factor. Thereby, I have decided to control for each subset of age aka Nestling, Juveniles, Yearlings, and Fledglings. But first, I have run the Mann-Kendall test for the entire population.

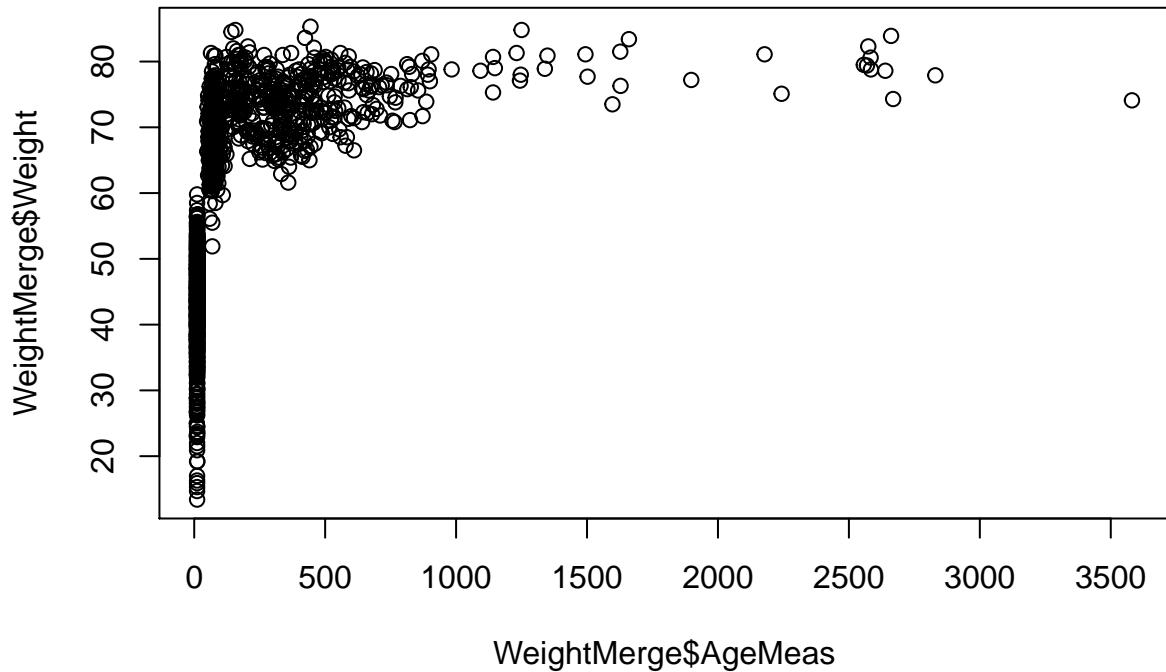
```
#install.packages("Kendall")
library(Kendall)
#require(Kendall)
```

On the entire data:

```
##View(WeightMerge)
boxplot(WeightMerge$Weight~WeightMerge$AgeCategory) # It is so clear how there is a trend in the
```



```
#whole population as nestlings have the lowest medians and the highest medians are of Yearlings.
plot(WeightMerge$AgeMeas, WeightMerge$Weight)
```



```

resultaggregate<-MannKendall(WeightMerge$Weight) # Running the Mann-Kendall on the WHOLE
#population.
print(resultaggregate)

## tau = 0.217, 2-sided pvalue =< 2.22e-16
summary(resultaggregate)

## Score = 251367 , Var(Score) = 393670432
## denominator = 1159084
## tau = 0.217, 2-sided pvalue =< 2.22e-16

```

We get a p-value of less than 0.05 shows a presence of monotropic trend. The Tau statistic is not that high which implies that the trend, although present, is not that high.

### On specific age groups:

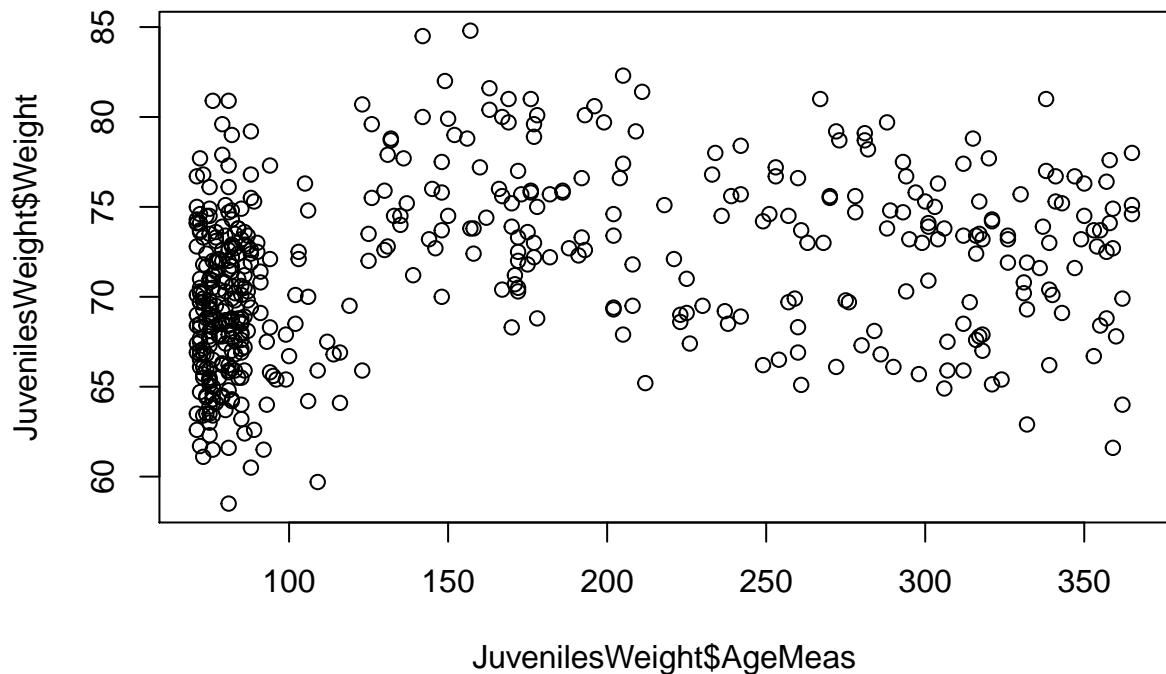
```

library(plyr)
count(WeightMerge$AgeCategory) #Gives a count of all the age categories we have.

##          x freq
## 1 Fledgling 172
## 2 Juveniles 483
## 3 Nestling  663
## 4 Yearling  206

# We can see that we have the most amount of Nestlings, followed by Juveniles, Yearlings,
# and Fledglings (in that order).
# We can check for a trend phenotype by breaking it down in these constituent values.
JuvenilesWeight<-subset(WeightMerge, AgeCategory=="Juveniles") #Subsetting just Juveniles
plot(JuvenilesWeight$Weight~JuvenilesWeight$AgeMeas)

```



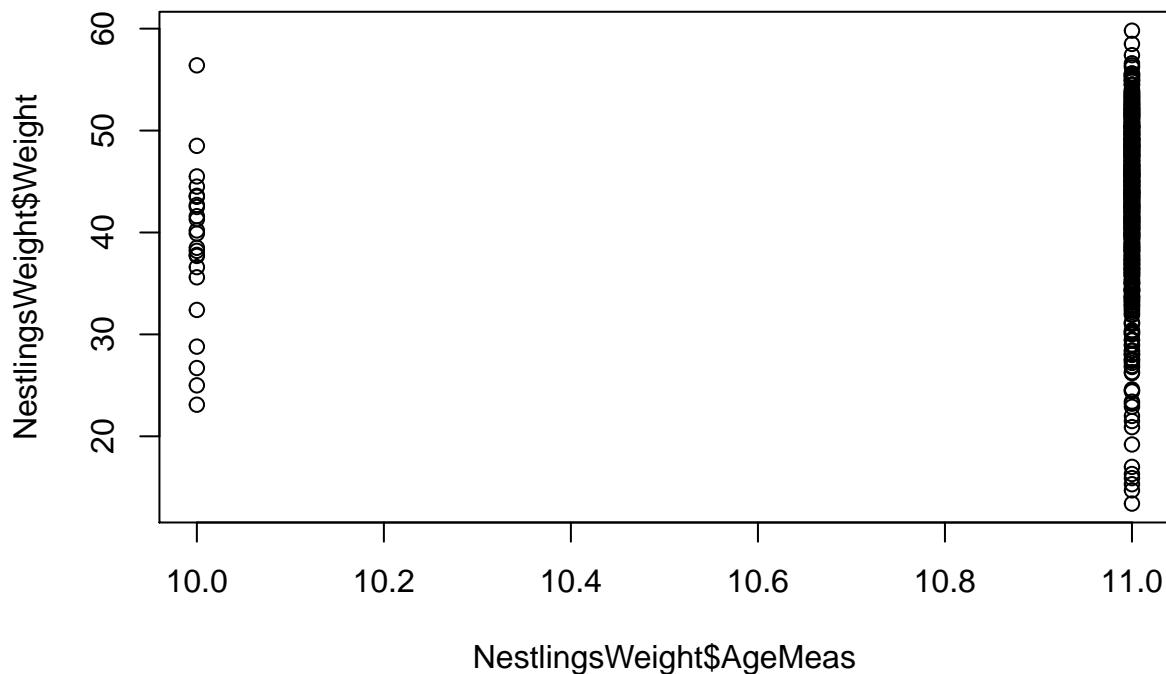
```
##View(JuvenilesWeight)
resultsJuveniles<-MannKendall(JuvenilesWeight$Weight)
print(resultsJuveniles)
```

```
## tau = 0.117, 2-sided pvalue =0.0001334
summary(resultsJuveniles)
```

```
## Score = 13538 , Var(Score) = 12557541
## denominator = 116076
## tau = 0.117, 2-sided pvalue =0.0001334
```

Also significant.

```
# We can see that we have the most amount of nestlings, followed by Juveniles, Yearlings, and Fledgling
# We can check for a trend phenotype by breaking it down in these constituent values.
NestlingsWeight<-subset(WeightMerge, AgeCategory=="Nestling")
plot(NestlingsWeight$Weight~NestlingsWeight$AgeMeas)
```



```
##View(NestlingsWeight)
resultsNestlings<-MannKendall(NestlingsWeight$Weight)
print(resultsJuveniles)
```

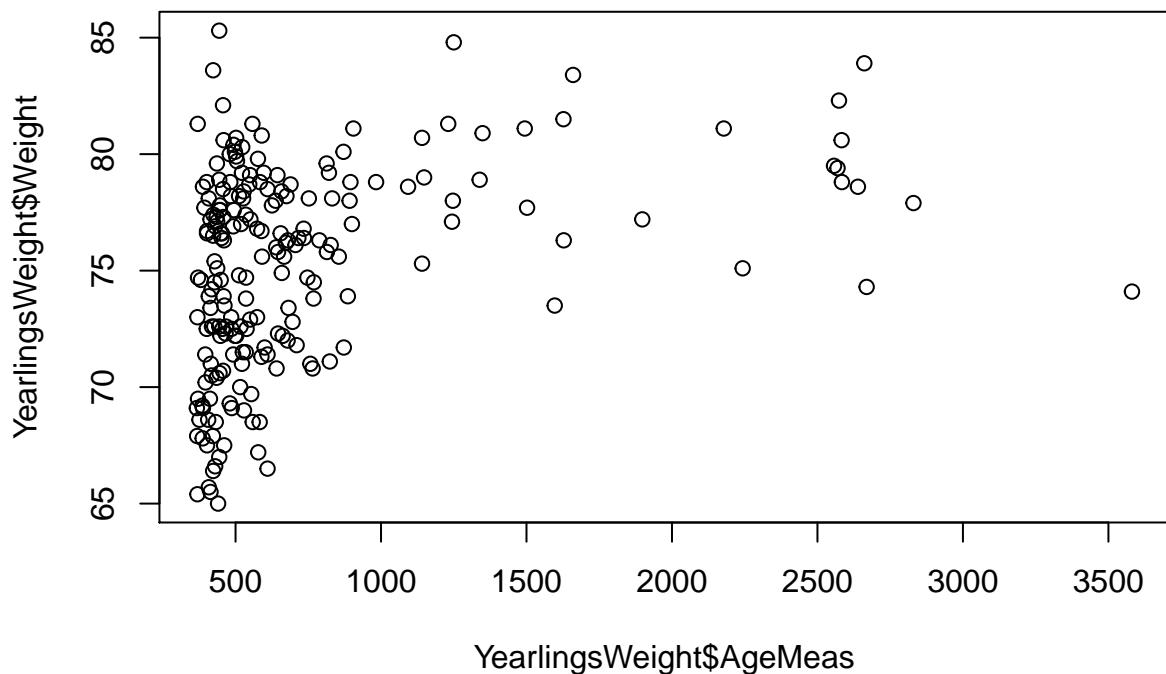
```
## tau = 0.117, 2-sided pvalue =0.0001334
summary(resultsJuveniles)
```

```
## Score = 13538 , Var(Score) = 12557541
## denominator = 116076
## tau = 0.117, 2-sided pvalue =0.0001334
```

Also significant.

*# We can see that we have the most amount of nestlings, followed by Juveniles, Yearlings, and Fledgling*  
*# We can check for a trend phenotype by breaking it down in these constituent values.*

```
YearlingsWeight<-subset(WeightMerge, AgeCategory=="Yearling")
plot(YearlingsWeight$Weight~YearlingsWeight$AgeMeas)
```



```
##View(YearlingsWeight)
resultsYearlings<-MannKendall(YearlingsWeight$Weight)
print(resultsYearlings)
```

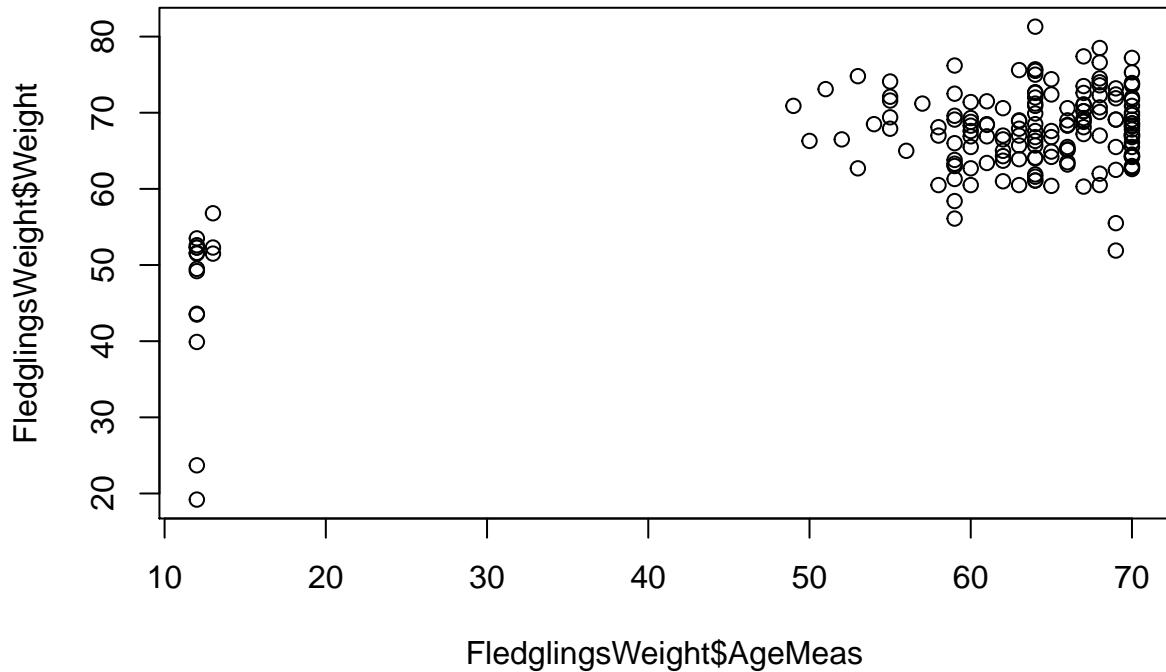
```
## tau = -0.154, 2-sided pvalue =0.0010231
summary(resultsYearlings)
```

```
## Score = -3249 , Var(Score) = 978143.7
## denominator = 21044.88
## tau = -0.154, 2-sided pvalue =0.0010231
```

Also significant.

*# We can see that we have the most amount of nestlings, followed by Juveniles, Yearlings, and Fledgling*  
*# We can check for a trend phenotype by breaking it down in these constituent values.*

```
FledglingsWeight<-subset(WeightMerge, AgeCategory=="Fledgling")
plot(FledglingsWeight$Weight~FledglingsWeight$AgeMeas)
```



```
##View(FledglingsWeight)
resultsFledgling<-MannKendall(FledglingsWeight$Weight)
print(resultsFledgling)
```

```
## tau = -0.109, 2-sided pvalue =0.033984
summary(resultsFledgling)

## Score = -1602 , Var(Score) = 570170
## denominator = 14667.95
## tau = -0.109, 2-sided pvalue =0.033984
```

Also significant.

So my questions are: 1) Is dividing the dataset in their respective categories the correct way to control for age? 2) Does this look ok so far- so that it may be replicated for other phenotypes?

And on 7 November 2019, Dr. Chen indicated that breaking down phenotypes by age classes and then running Mann-Kendall on those is indeed the best way forward to find monotonic trends. That being said, the following is Mann-Kendall on all phenotypes, followed by a table summarising that information.

```

{r} # library(plyr) # boxplot(HLMerge$HeadLength~HLMerge$AgeCategory)
# count(HLMerge$AgeCategory) # It is interesting to notice
that there are no nestlings that had their head lengths measured.
# JuvenilesHL<-subset(HLMerge, AgeCategory=="Juveniles") #
Subsetting as age categories # FledglingsHL<-subset(HLMerge,
AgeCategory=="Fledgling") # YearlingsHL<-subset(HLMerge, AgeCategory=="Y
# #NestlingsHL<-subset(HLMerge, AgeCategory=="Nestling") #
resJuvHL<-MannKendall(JuvenilesHL$HeadLength) # resFlHL<-MannKendall(Fle
# resYearHL<-MannKendall(YearlingsHL$HeadLength) # #resNestHL<-MannKenda
# summary(resJuvHL) # summary(resFlHL) # summary(resYearHL) #
#print(resNestHL) #

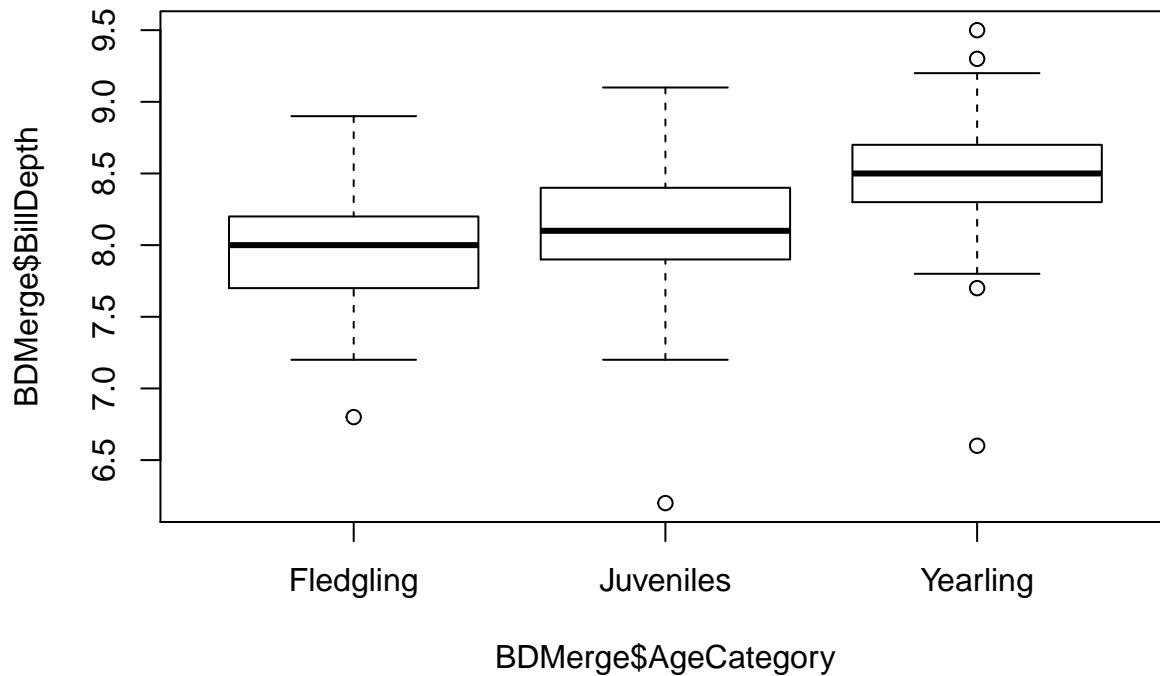
```

None of the age categories showed significant evidence of having a monotonic trend associated with head length.

```

library(plyr)
boxplot(BDMerge$BillDepth~BDMerge$AgeCategory)

```



```

#count(BDMerge$AgeCategory) # It is interesting to notice that there are no nestlings that had their bi
JuvenilesBD<-subset(BDMerge, AgeCategory=="Juveniles") # Subsetting as age categories
FledglingsBD<-subset(BDMerge, AgeCategory=="Fledgling")
YearlingsBD<-subset(BDMerge, AgeCategory=="Yearling")
NestlingsBD<-subset(BDMerge, AgeCategory=="Nestling")
resJuvBD<-MannKendall(JuvenilesBD$BillDepth)
resFlBD<-MannKendall(FledglingsBD$BillDepth)
resYearBD<-MannKendall(YearlingsBD$BillDepth)
#resNestBD<-MannKendall(NestlingsBD$BillDepth)
summary(resJuvBD)

```

```

## Score = 1712 , Var(Score) = 1861141
## denominator = 31446.17
## tau = 0.0544, 2-sided pvalue =0.20978
summary(resFlBD)

## Score = -171 , Var(Score) = 398445.7
## denominator = 11179.86
## tau = -0.0153, 2-sided pvalue =0.78769
summary(resYearBD)

## Score = 392 , Var(Score) = 26763.33
## denominator = 1812.887
## tau = 0.216, 2-sided pvalue =0.016846
#summary(resNestBD)

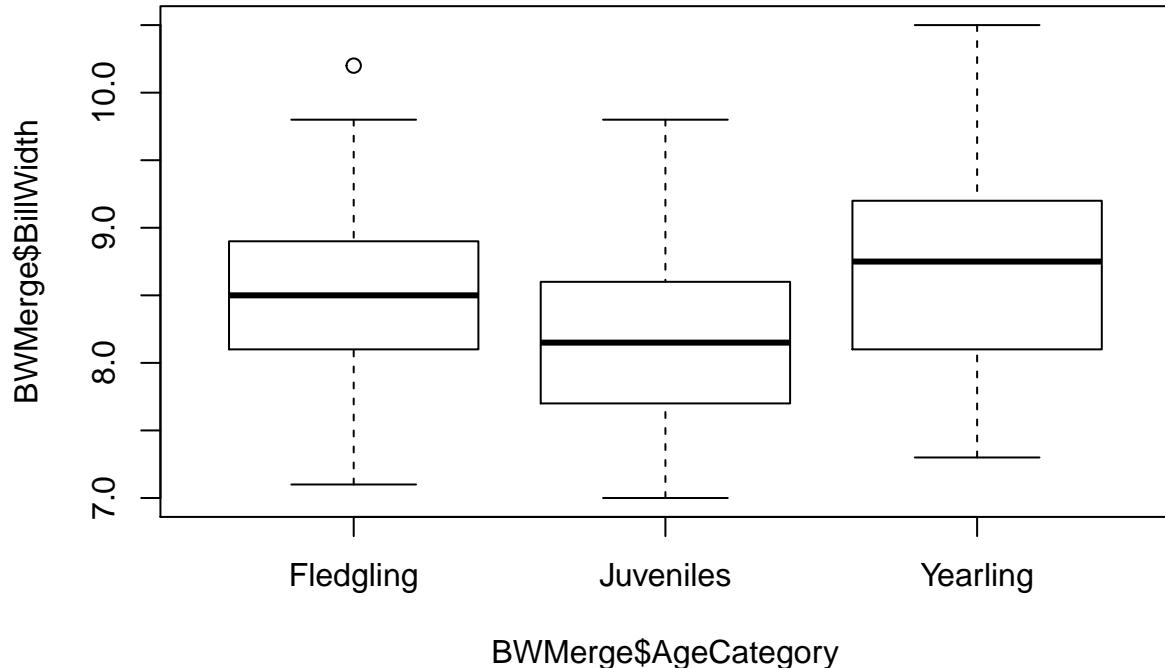
```

Only Yearlings had a p-value which showed significance when compared with Bill Depth.

```

library(plyr)
boxplot(BWMerge$BillWidth~BWMerge$AgeCategory)

```



```

count(BWMerge$AgeCategory) # It is interesting to notice that there are no nestlings that had their bil

##          x freq
## 1 Fledgling 153
## 2 Juveniles 256
## 3 Yearling   62

JuvenilesBW<-subset(BWMerge, AgeCategory=="Juveniles") # Subsetting as age categories
FledglingsBW<-subset(BWMerge, AgeCategory=="Fledgling")
YearlingsBW<-subset(BWMerge, AgeCategory=="Yearling")
NestlingsBW<-subset(BWMerge, AgeCategory=="Nestling")
resJuvBW<-MannKendall(JuvenilesBW$BillWidth)
resFlBW<-MannKendall(FledglingsBW$BillWidth)

```

```

resYearBW<-MannKendall(YearlingsBW$BillWidth)
#resNestBW<-MannKendall(NestlingsBW$BillWidth)
summary(resJuvBW)

## Score = 10666 , Var(Score) = 1869219
## denominator = 31885.27
## tau = 0.335, 2-sided pvalue =< 2.22e-16
summary(resFlBW)

## Score = -170 , Var(Score) = 399782
## denominator = 11305.53
## tau = -0.015, 2-sided pvalue =0.78925
summary(resYearBW)

## Score = 694 , Var(Score) = 26955.33
## denominator = 1846.988
## tau = 0.376, 2-sided pvalue =2.4319e-05
#summary(resNestBW)

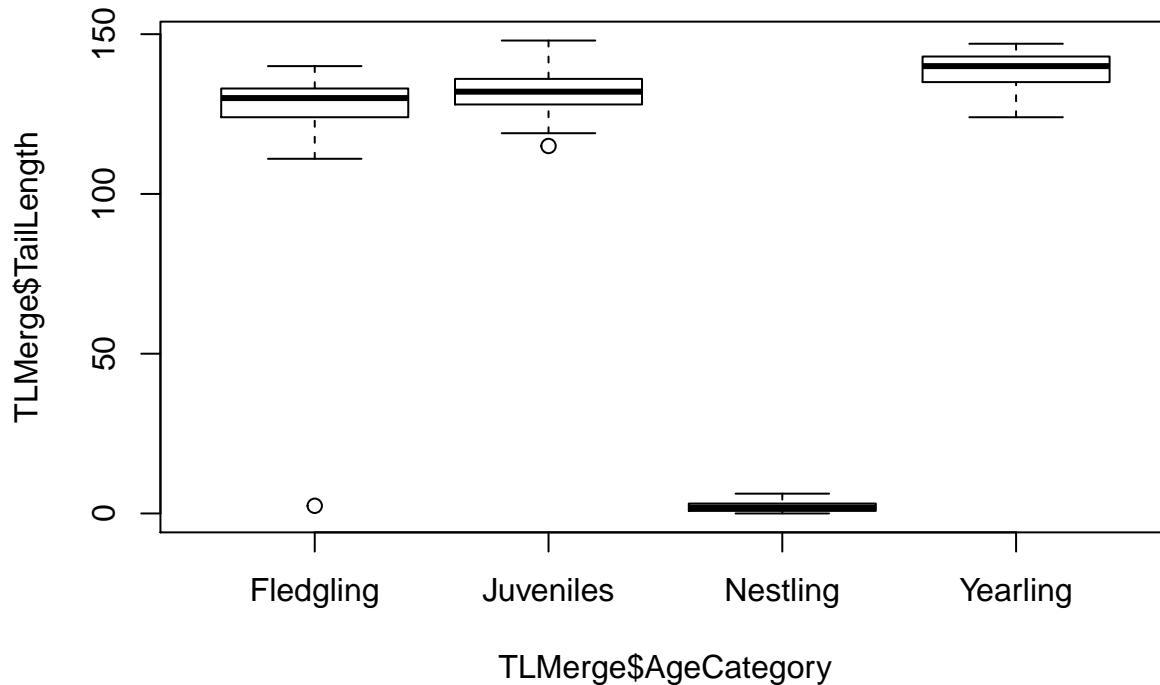
```

Juveniles and Yearlings had a p-value which showed significance when compared with Bill Width.

```

library(plyr)
boxplot(TLMerge$TailLength~TLMerge$AgeCategory)

```



```

count(TLMerge$AgeCategory)

##          x freq
## 1 Fledgling 149
## 2 Juveniles 253
## 3 Nestling 130
## 4 Yearling  37

```

```

JuvenilesTL<-subset(TLMerge, AgeCategory=="Juveniles") # Subsetting as age categories
FledglingsTL<-subset(TLMerge, AgeCategory=="Fledgling")
YearlingsTL<-subset(TLMerge, AgeCategory=="Yearling")
NestlingsTL<-subset(TLMerge, AgeCategory=="Nestling")
resJuvTL<-MannKendall(JuvenilesTL$TailLength)
resFlTL<-MannKendall(FledglingsTL$TailLength)
resYearTL<-MannKendall(YearlingsTL$TailLength)
resNestTL<-MannKendall(NestlingsTL$TailLength)
summary(resJuvTL)

## Score = -8518 , Var(Score) = 1803861
## denominator = 31127.16
## tau = -0.274, 2-sided pvalue =2.2769e-10
summary(resFlTL)

## Score = -454 , Var(Score) = 369789.3
## denominator = 10766.96
## tau = -0.0422, 2-sided pvalue =0.45631
summary(resYearTL)

## Score = -184 , Var(Score) = 5810
## denominator = 651.8497
## tau = -0.282, 2-sided pvalue =0.016358
summary(resNestTL)

## Score = -759 , Var(Score) = 245473
## denominator = 8211.199
## tau = -0.0924, 2-sided pvalue =0.12604

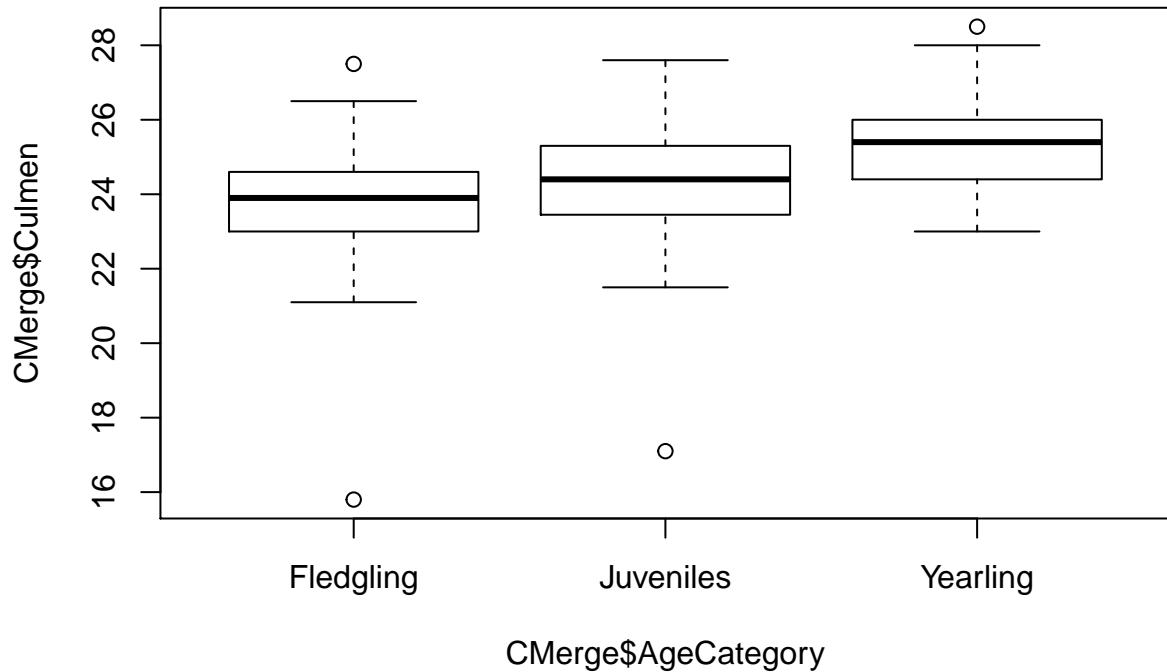
```

Juveniles and Yearlings had a p-value which showed significance when compared with Tail Length.

```

library(plyr)
boxplot(CMerge$Culmen~CMerge$AgeCategory)

```



```

count(CMerge$AgeCategory)

##          x freq
## 1 Fledgling 153
## 2 Juveniles 255
## 3 Yearling  65

JuvenilesCM<-subset(CMerge, AgeCategory=="Juveniles") # Subsetting as age categories
FledglingsCM<-subset(CMerge, AgeCategory=="Fledgling")
YearlingsCM<-subset(CMerge, AgeCategory=="Yearling")
#NestlingsCM<-subset(CMerge, AgeCategory=="Nestling")
resJuvCM<-MannKendall(JuvenilesCM$Culmen)
resFlCM<-MannKendall(FledglingsCM$Culmen)
resYearCM<-MannKendall(YearlingsCM$Culmen)
#resNestCM<-MannKendall(NestlingsCM$Culmen)
summary(resJuvCM)

## Score =  7138 , Var(Score) = 1851363
## denominator =  32016.4
## tau = 0.223, 2-sided pvalue =1.1921e-07
summary(resFlCM)

## Score = -555 , Var(Score) = 401337
## denominator =  11494.74
## tau = -0.0483, 2-sided pvalue =0.38185
summary(resYearCM)

## Score =  301 , Var(Score) = 31113.67
## denominator =  2050.288
## tau = 0.147, 2-sided pvalue =0.088986
#summary(resNestTL)

```

Juveniles and Yearlings had a p-value which showed significance when compared with Culmen.

*So far we have all 4 age categories from weight as significant; Juveniles and Yearlings only for tail length, culmen, bill width; and only Yearlings for Bill Depth.*

Get JAYID and NatalYear from Access\$Demofldg, and then merge with the main data set. Then take the mean of weights from 1990, 1991.. and then run Mann-Kendall on that.

```
NY<-(access$Demofldg[, c('ID', 'NatalYear')])  
WeightNY<-merge(WeightMerge, NY, by="ID")  
WeightNY<-na.omit(WeightNY)  
#count(WeightNY, NatalYear)  
#mean(WeightNY$Weight [WeightNY$NatalYear==1970])  
#mean(WeightNY$Weight [WeightNY$NatalYear==1969])  
MkWeight<-aggregate(Weight~NatalYear, data=WeightNY, FUN=function(WeightNY) c(mean=mean(WeightNY), count=length(WeightNY)))  
#View(MkWeight) #Now we have mean measurements by natal Year and we can run Mann-Kendall on this.  
resWeight<-(MannKendall(MkWeight$Weight))  
summary(resWeight)
```

```
## Score = -603 , Var(Score) = 45909  
## denominator = 2696.997  
## tau = -0.224, 2-sided pvalue =0.00496
```

It is clear with a p-value of 0.00496, and a tau value of -0.224, there is a monotonic downward trend in the weight of these birds.

```
NY<-(access$Demofldg[, c('ID', 'NatalYear')])  
BDNY<-merge(BDMerge, NY, by="ID")  
BDNY<-na.omit(BDNY)  
#count(WeightNY, NatalYear)  
#mean(WeightNY$Weight [WeightNY$NatalYear==1970])  
#mean(WeightNY$Weight [WeightNY$NatalYear==1969])  
MkBD<-aggregate(BillDepth~NatalYear, data=BDNY, FUN=function(BDNY) c(mean=mean(BDNY), count=length(BDNY)))  
#View(MkBD) #Now we have mean measurements by natal Year and we can run Mann-Kendall on this.  
resBD<-(MannKendall(MkBD$BillDepth))  
summary(resBD)
```

```
## Score = 516 , Var(Score) = 42278.67  
## denominator = 2541.961  
## tau = 0.203, 2-sided pvalue =0.012257
```

Also significant with a p-value of 0.012 and a positive tau.

```
NY<-(access$Demofldg[, c('ID', 'NatalYear')])  
BWNY<-merge(BWMerge, NY, by="ID")  
BWNY<-na.omit(BWNY)  
#count(WeightNY, NatalYear)  
#mean(WeightNY$Weight [WeightNY$NatalYear==1970])  
#mean(WeightNY$Weight [WeightNY$NatalYear==1969])  
MkBW<-aggregate(BillWidth~NatalYear, data=BWNY, FUN=function(BWNY) c(mean=mean(BWNY), count=length(BWNY)))  
#View(MkBW) #Now we have mean measurements by natal Year and we can run Mann-Kendall on this.  
resBW<-(MannKendall(MkBW$BillWidth))  
summary(resBW)
```

```
## Score = 267 , Var(Score) = 42277.67  
## denominator = 2541.459  
## tau = 0.105, 2-sided pvalue =0.19578
```

Not significant.

```

NY<-(access$Demofldg[, c('ID', 'NatalYear')])
TLNY<-merge(TLMerge, NY, by="ID")
TLNY<-na.omit(TLNY)
#count(WeightNY,NatalYear)
#mean(WeightNY$Weight [WeightNY$NatalYear==1970])
#mean(WeightNY$Weight [WeightNY$NatalYear==1969])
MkTL<-aggregate(TailLength~NatalYear, data=TLNY, FUN=function(TLNY) c(mean=mean(TLNY), count=length(TLNY)))
#View(MkTL) #Now we have mean measurements by natal Year and we can run Mann-Kendall on this.
resTL<-(MannKendall(MkTL$TailLength))
summary(resTL)

```

```

## Score = -1259 , Var(Score) = 42294.33
## denominator = 2546.482
## tau = -0.494, 2-sided pvalue =9.5344e-10

```

Tail length is significant with a very low p-value and negative tau.

```

NY<-(access$Demofldg[, c('ID', 'NatalYear')])
CulmenNY<-merge(CMerge, NY, by="ID")
CulmenNY<-na.omit(CulmenNY)
#count(WeightNY,NatalYear)
#mean(WeightNY$Weight [WeightNY$NatalYear==1970])
#mean(WeightNY$Weight [WeightNY$NatalYear==1969])
MkC<-aggregate(Culmen~NatalYear, data=CulmenNY, FUN=function(CulmenNY) c(mean=mean(CulmenNY), count=length(CulmenNY)))
#View(MkC) #Now we have mean measurements by natal Year and we can run Mann-Kendall on this.
resC<-(MannKendall(MkC$Culmen))
summary(resC)

```

```

## Score = -894 , Var(Score) = 42275.33
## denominator = 2540.956
## tau = -0.352, 2-sided pvalue =1.4043e-05

```

Culmen has a significant p-value along with the tau being negative.

**Looking at means born on specific years along with age categories to control for age:**

```

# #View(YearlingsWeight)
# #View(WeightNY) # Has Natal Year
# WNYYearling<-merge(WeightNY, YearlingsWeight) # This works. It has no NAs.
# AgWNYYearling<-aggregate(Weight~NatalYear, data=WNYYearling, FUN=function(MkWNYYearling) c(mean=mean(Weight), n=n))
# #View(AgWNYYearling)
# MkWNYYearling<-(MannKendall(AgWNYYearling$Weight))
# print(MkWNYYearling)
# ts(test1)
# test1<-subset(AgWNYYearling, AgWNYYearling$NatalYear>=1990)
# plot(test1)
# testmk<-MannKendall(test1$Weight)
# print(testmk)
# #plot(Weight ~ NatalYear, data=test1)+abline()
# abline(lm(Weight ~ NatalYear, data=test1))
#
# #lm<-(AgWNYYearling$Weight~AgWNYYearling$NatalYear)
# #lm

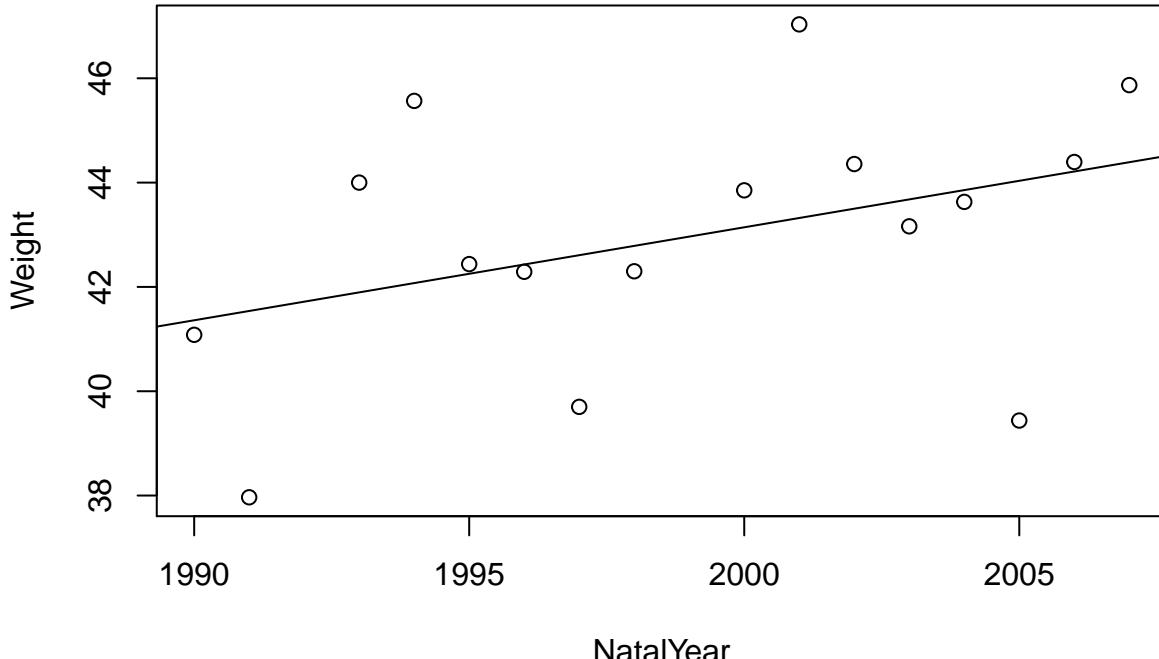
```

```
# #abline(lm(AgWNYYearling$Weight~AgWNYYearling$NatalYear))
# #print(AgWNYYearling$Weight)
# # print(AgWNYYearling$NatalYear)
```

Weight: Nestlings

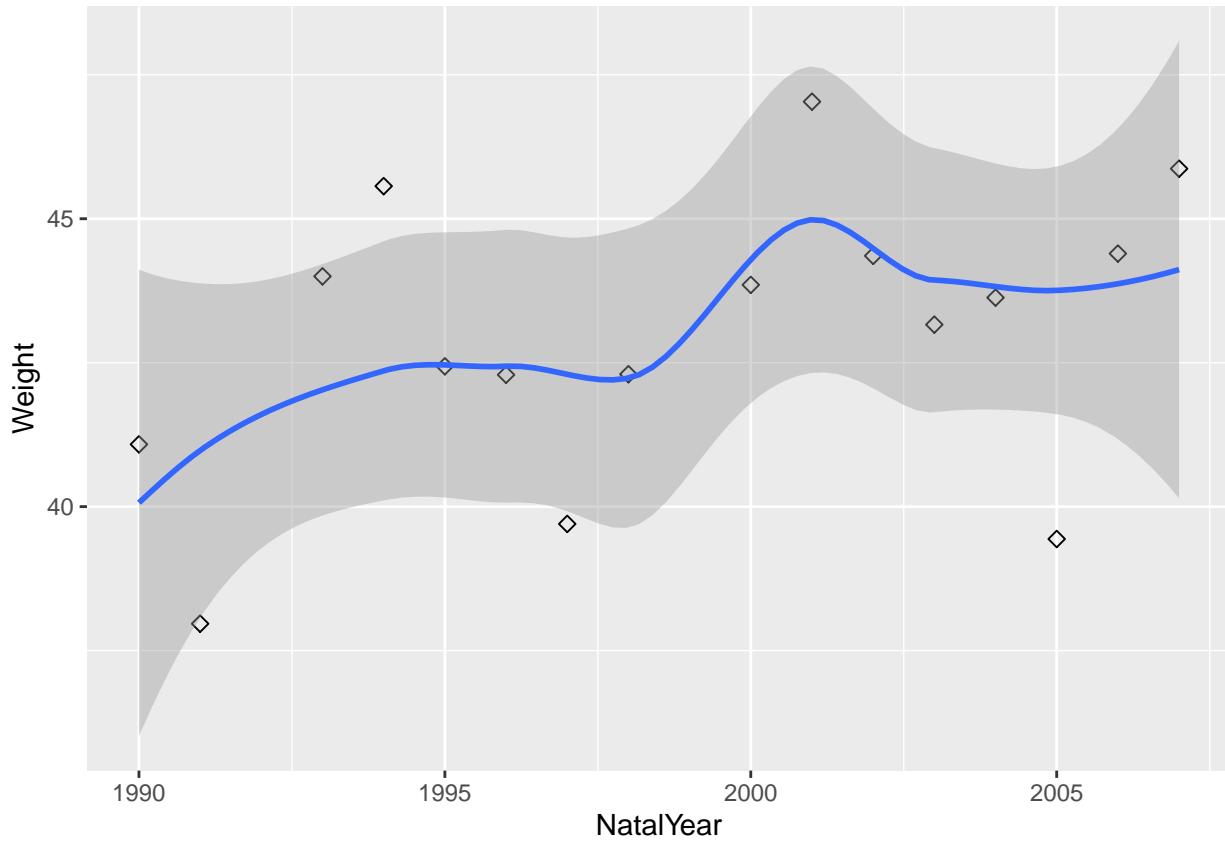
```
##View(NestlingsWeight) # Has all our pertinent data so far
##View(WeightNY) # Has Natal Year
WNYNestlings<-merge(WeightNY, NestlingsWeight) # This works. It has no NAs.
WeightAgN<-aggregate(Weight~NatalYear, data=WNYNestlings, FUN=function(WNYNestlings) c(mean=mean(WNYNestlings), n=nrow(WNYNestlings)))
##View(WeightAgN)
WeightAgN2<-subset(WeightAgN, WeightAgN$NatalYear>=1990) # Only take greater than 1990's
##View(WeightAgF1)
plot(Weight ~ NatalYear, data=WeightAgN2)
lm(Weight ~ NatalYear, data=WeightAgN2)
```

```
##
## Call:
## lm(formula = Weight ~ NatalYear, data = WeightAgN2)
##
## Coefficients:
## (Intercept)    NatalYear
## -313.4183      0.1783
abline(lm(Weight ~ NatalYear, data=WeightAgN2))
```



```
MkWeightAgN1<-MannKendall(WeightAgN2$Weight)
print(MkWeightAgN1)
```

```
## tau = 0.283, 2-sided pvalue = 0.13735
ggplot(WeightAgN2, aes(x=NatalYear, y=Weight)) + geom_point(size=2, shape=23) +geom_smooth()
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



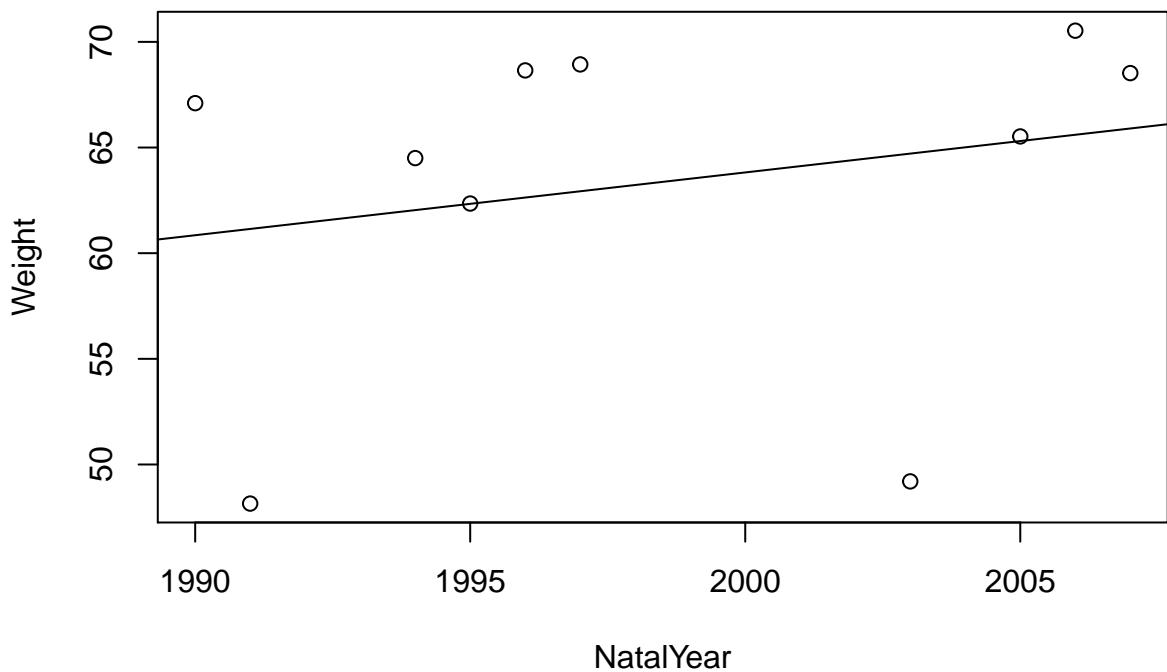
Weight: Fledgling

```

##View(FledglingsWeight) # Has all our pertinent data so far
##View(WeightNY) # Has Natal Year
WNYFledglings<-merge(WeightNY, FledglingsWeight) # This works. It has no NAs.
WeightAgF<-aggregate(Weight~NatalYear, data=WNYFledglings, FUN=function(WNYFledglings) c(mean=mean(WNYFledglings), n=nrow(WNYFledglings)))
##View(WeightAgF)
WeightAgF1<-subset(WeightAgF, WeightAgF$NatalYear>=1990) # Only take greater than 1990's
##View(WeightAgF1)
plot(Weight ~ NatalYear, data=WeightAgF1)
lm(Weight ~ NatalYear, data=WeightAgF1)

##
## Call:
## lm(formula = Weight ~ NatalYear, data = WeightAgF1)
##
## Coefficients:
## (Intercept)    NatalYear
##      -530.5039       0.2972
abline(lm(Weight ~ NatalYear, data=WeightAgF1))

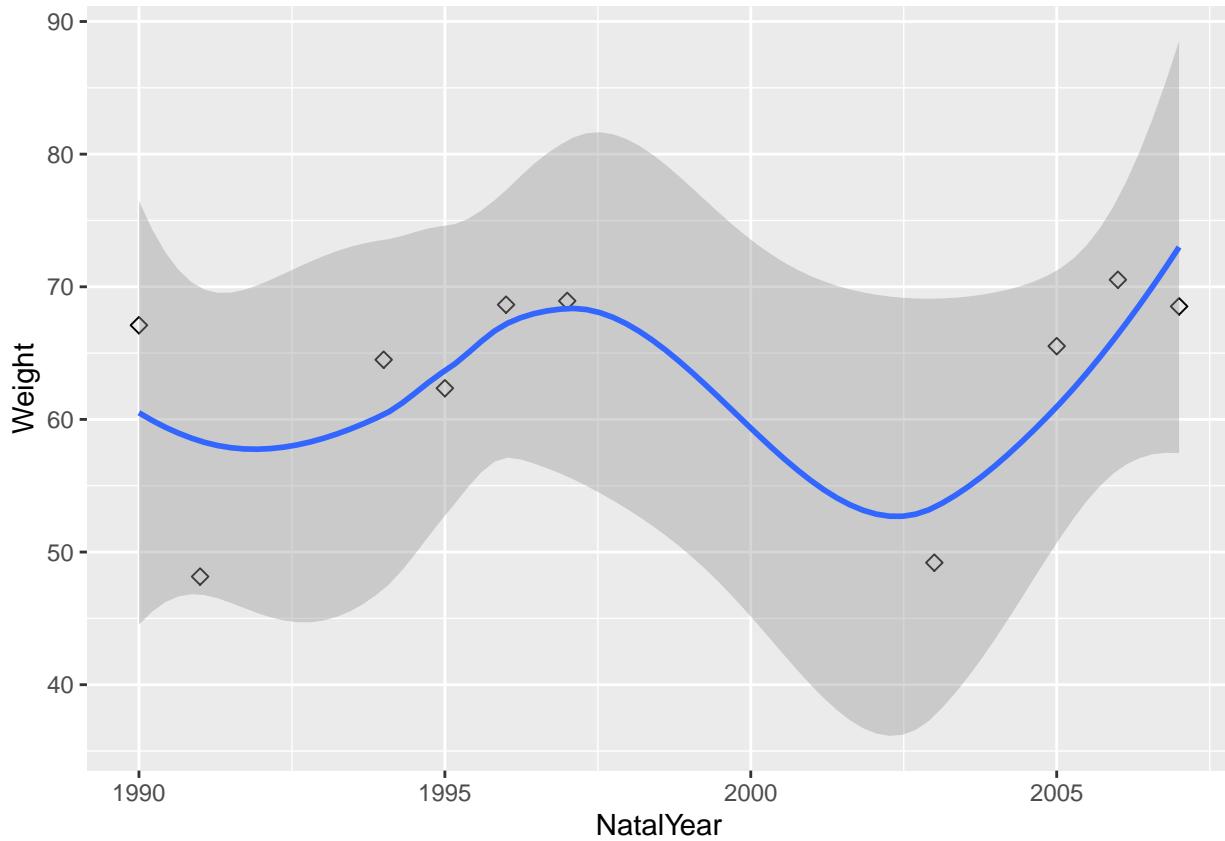
```



```
MkWeightAgF1<-MannKendall(WeightAgF1$Weight)
print(MkWeightAgF1)

## tau = 0.333, 2-sided pvalue =0.2105
ggplot(WeightAgF1, aes(x=NatalYear, y=Weight)) + geom_point(size=2, shape=23) +geom_smooth()

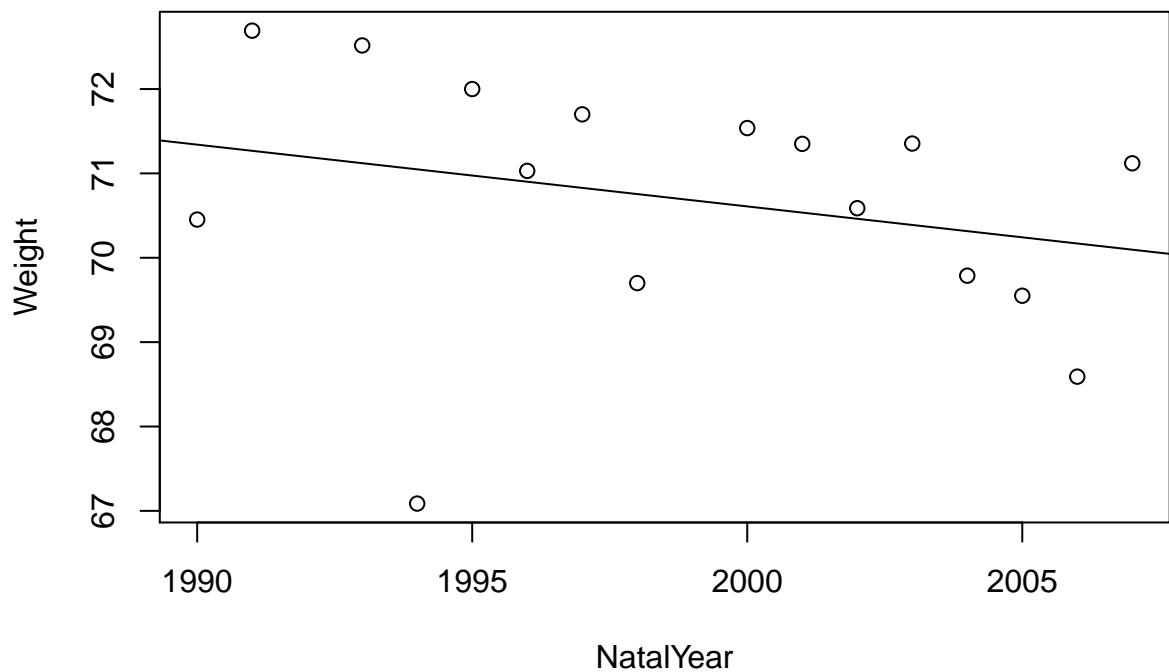
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



Weight: Juvenile

```
#View(JuvenilesWeight) # Has all our pertinent data so far
#View(WeightNY) # Has Natal Year
WNYJuvenile<-merge(WeightNY, JuvenilesWeight) # This works. It has no NAs.
WeightAgJ<-aggregate(Weight~NatalYear, data=WNYJuvenile, FUN=function(WNYJuvenile) c(mean=mean(WNYJuvenile), n=n))
##View(WeightAgJ)
WeightAgJ1<-subset(WeightAgJ, WeightAgJ$NatalYear>=1990) # Only take greater than 1990's
##View(WeightAgJ1)
plot(Weight ~ NatalYear, data=WeightAgJ1)
lm(Weight ~ NatalYear, data=WeightAgJ1)

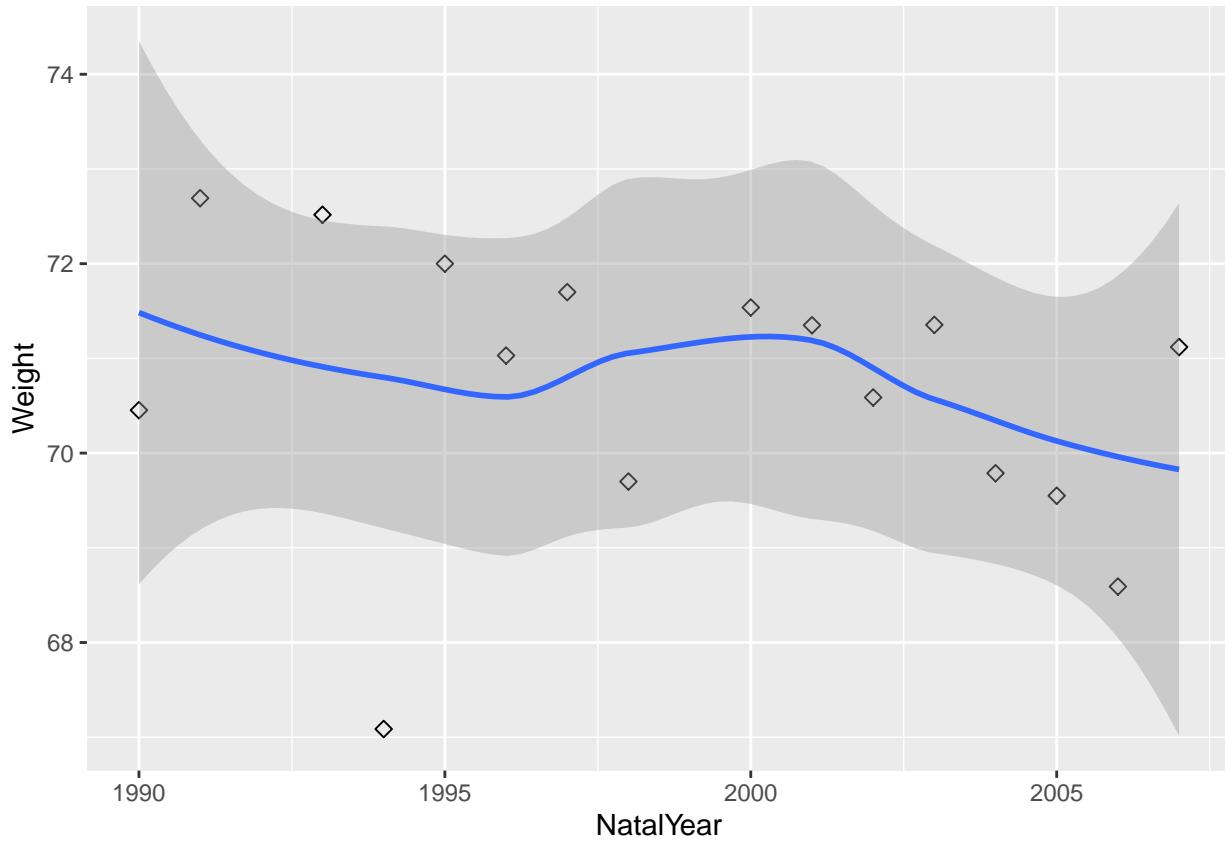
##
## Call:
## lm(formula = Weight ~ NatalYear, data = WeightAgJ1)
##
## Coefficients:
## (Intercept)    NatalYear
##      217.06818   -0.07323
abline(lm(Weight ~ NatalYear, data=WeightAgJ1))
```



```
MkWeightAgJ1<-MannKendall(WeightAgJ1$Weight)
print(MkWeightAgJ1)

## tau = -0.35, 2-sided pvalue = 0.064903
ggplot(WeightAgJ1, aes(x=NatalYear, y=Weight)) + geom_point(size=2, shape=23) +geom_smooth()

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



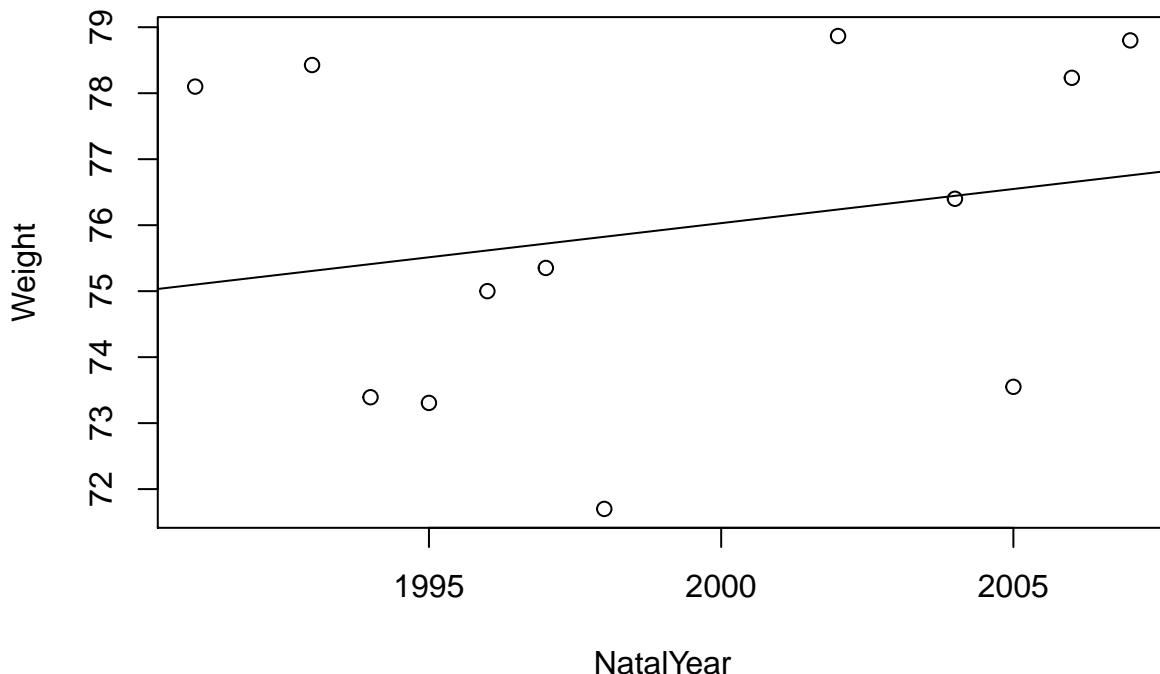
Weight: Yearling

```

##View(YearlingsWeight) # Has all our pertinent data so far
##View(WeightNY) # Has Natal Year
WNYYearling<-merge(WeightNY, YearlingsWeight) # This works. It has no NAs.
WeightAg<-aggregate(Weight~NatalYear, data=WNYYearling, FUN=function(MkWNYYearling) c(mean=mean(MkWNYYearling), lower=quantile(MkWNYYearling, 0.025), upper=quantile(MkWNYYearling, 0.975)))
##View(WeightAg)
WeightAg1<-subset(WeightAg, WeightAg$NatalYear>=1990) # Only take greater than 1990's
##View(WeightAg1)
plot(Weight ~ NatalYear, data=WeightAg1)
lm(Weight ~ NatalYear, data=WeightAg1)

##
## Call:
## lm(formula = Weight ~ NatalYear, data = WeightAg1)
##
## Coefficients:
## (Intercept)    NatalYear
##      -130.9450        0.1035
abline(lm(Weight ~ NatalYear, data=WeightAg1))

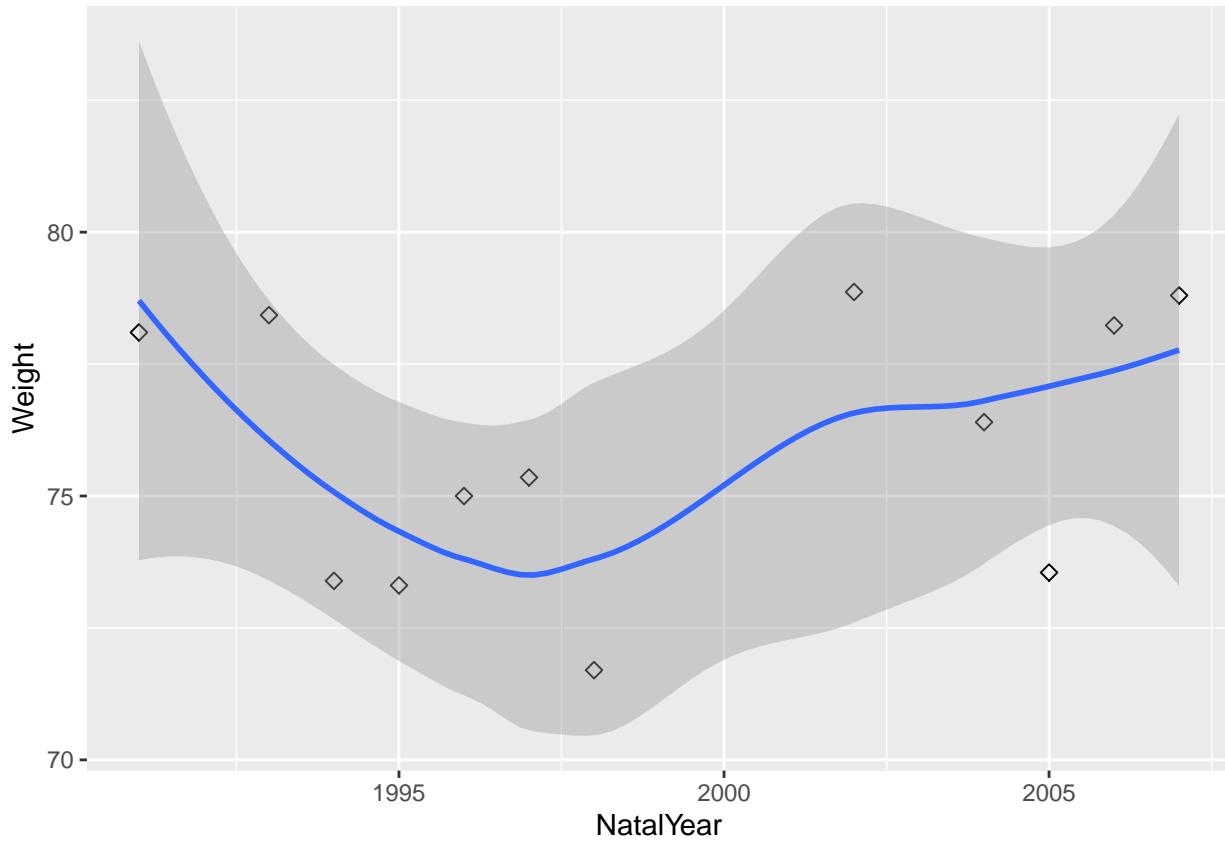
```



```
MkWeightAg1<-MannKendall(WeightAg1$Weight)
print(MkWeightAg1)

## tau = 0.182, 2-sided pvalue =0.45067
ggplot(WeightAg1, aes(x=NatalYear, y=Weight)) + geom_point(size=2, shape=23) +geom_smooth()

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



Making a 2x2 display of weight change for research presentation.

```
par(mfrow=c(2,2))
plot(Weight ~ NatalYear, data=WeightAgN2, xlab = "Year Born (Nestling)", ylab="Mean Weight")
lm(Weight ~ NatalYear, data=WeightAgN2)

##
## Call:
## lm(formula = Weight ~ NatalYear, data = WeightAgN2)
##
## Coefficients:
## (Intercept)    NatalYear
## -313.4183      0.1783
abline(lm(Weight ~ NatalYear, data=WeightAgN2))

plot(Weight ~ NatalYear, data=WeightAgF1, xlab = "Year Born (Fledgling)",ylab="Mean Weight")
lm(Weight ~ NatalYear, data=WeightAgF1)

##
## Call:
## lm(formula = Weight ~ NatalYear, data = WeightAgF1)
##
## Coefficients:
## (Intercept)    NatalYear
## -530.5039      0.2972
abline(lm(Weight ~ NatalYear, data=WeightAgF1))

plot(Weight ~ NatalYear, data=WeightAgJ1, xlab = "Year Born (Juvenile)",ylab="Mean Weight")
lm(Weight ~ NatalYear, data=WeightAgJ1)
```

```

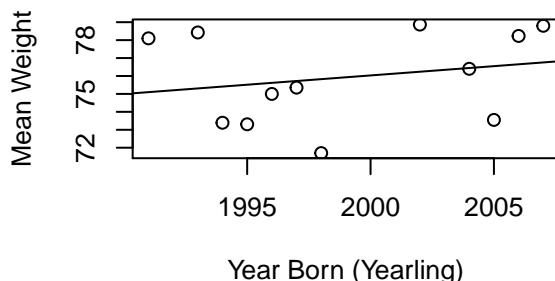
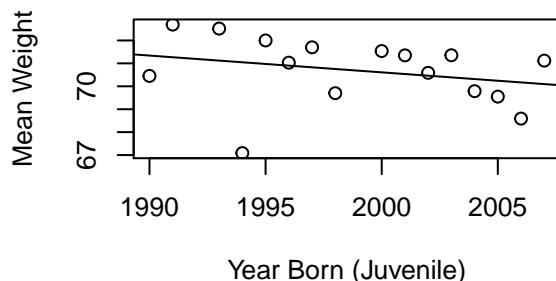
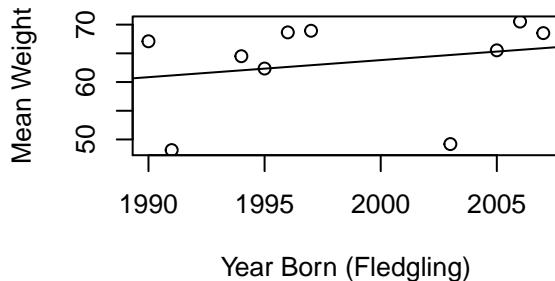
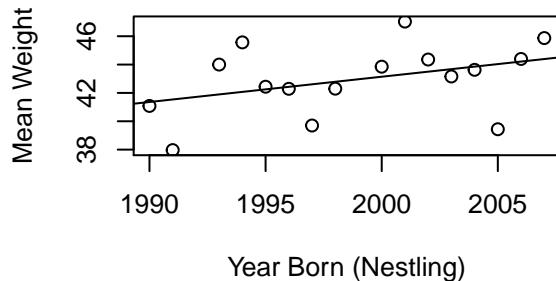
## 
## Call:
## lm(formula = Weight ~ NatalYear, data = WeightAgJ1)
## 
## Coefficients:
## (Intercept)    NatalYear
## 217.06818     -0.07323
abline(lm(Weight ~ NatalYear, data=WeightAgJ1))
plot(Weight ~ NatalYear, data=WeightAg1, xlab = "Year Born (Yearling)", ylab="Mean Weight")
lm(Weight ~ NatalYear, data=WeightAg1)

```

```

## 
## Call:
## lm(formula = Weight ~ NatalYear, data = WeightAg1)
## 
## Coefficients:
## (Intercept)    NatalYear
## -130.9450      0.1035
abline(lm(Weight ~ NatalYear, data=WeightAg1))

```

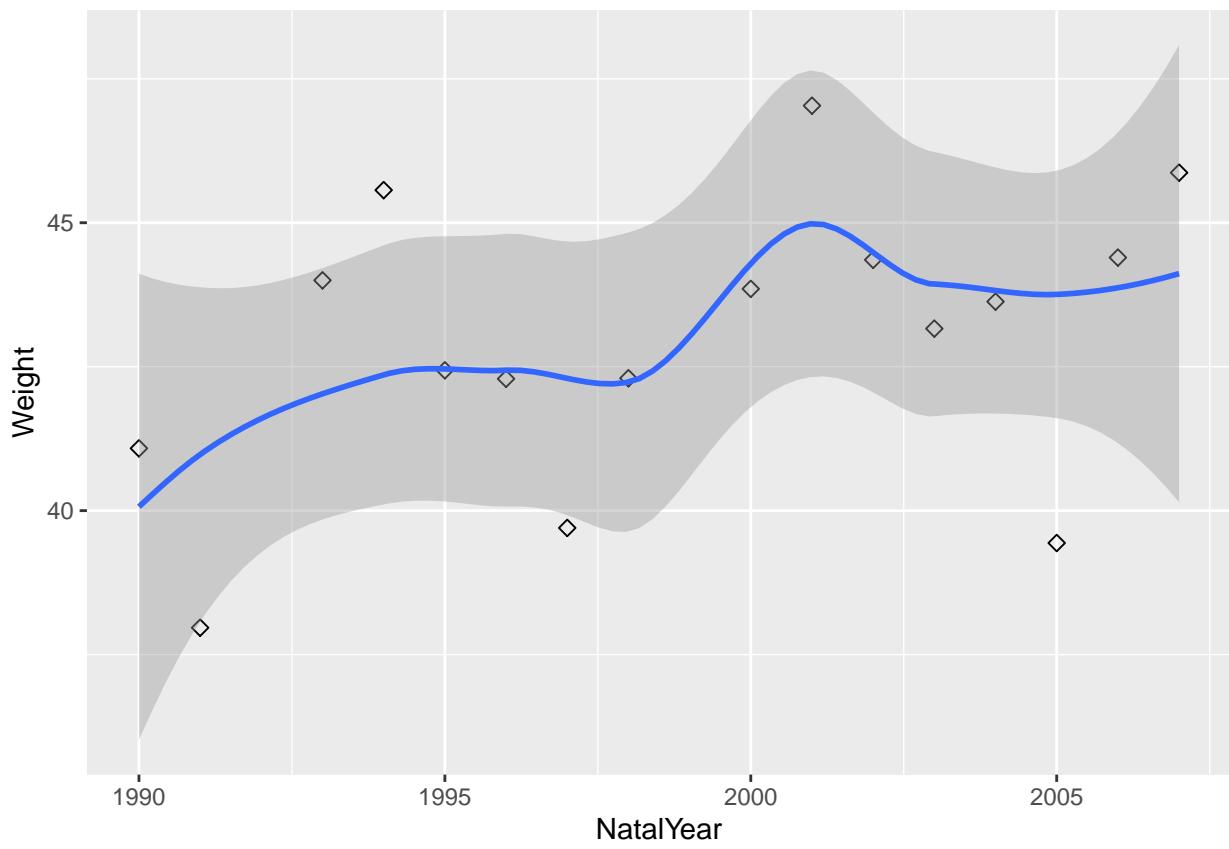


```

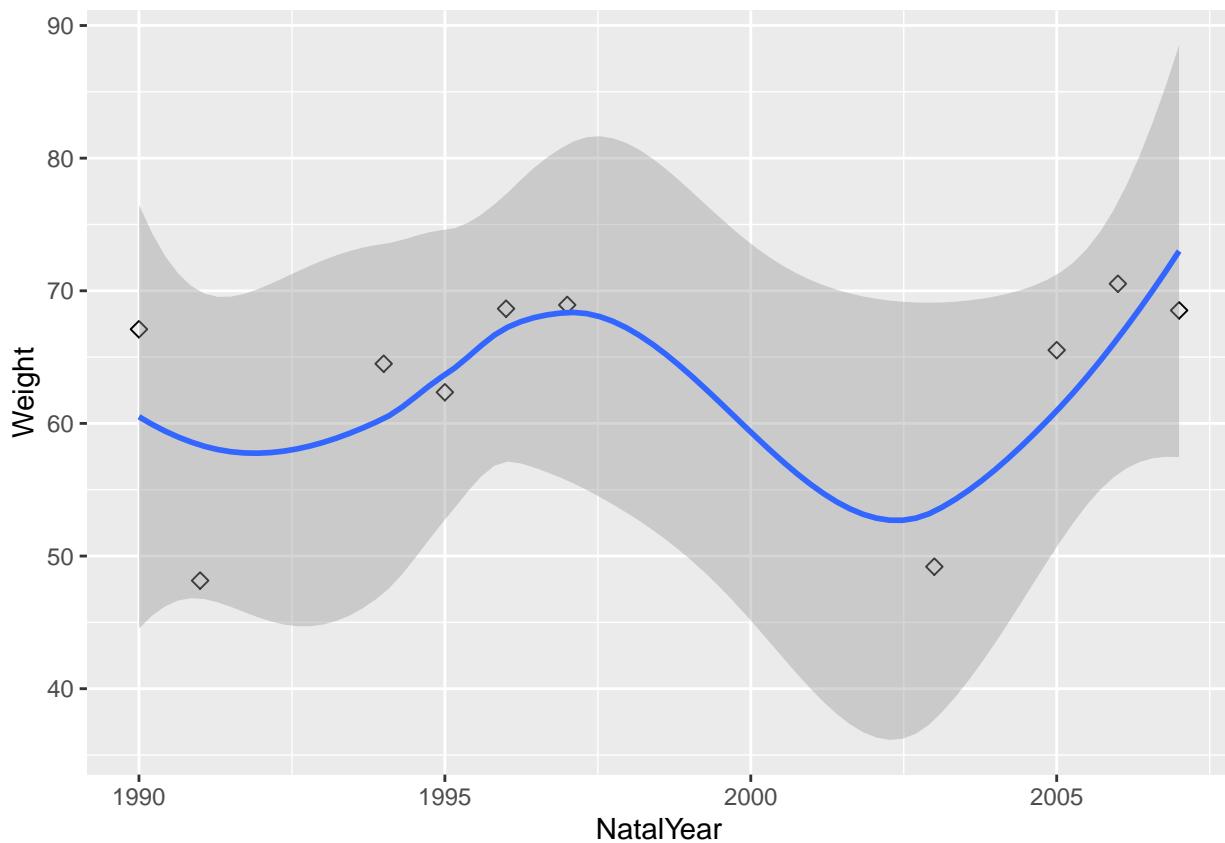
par(mfrow=c(2,2))
ggplot(WeightAgN2, aes(x=NatalYear, y=Weight)) + geom_point(size=2, shape=23) +geom_smooth()

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

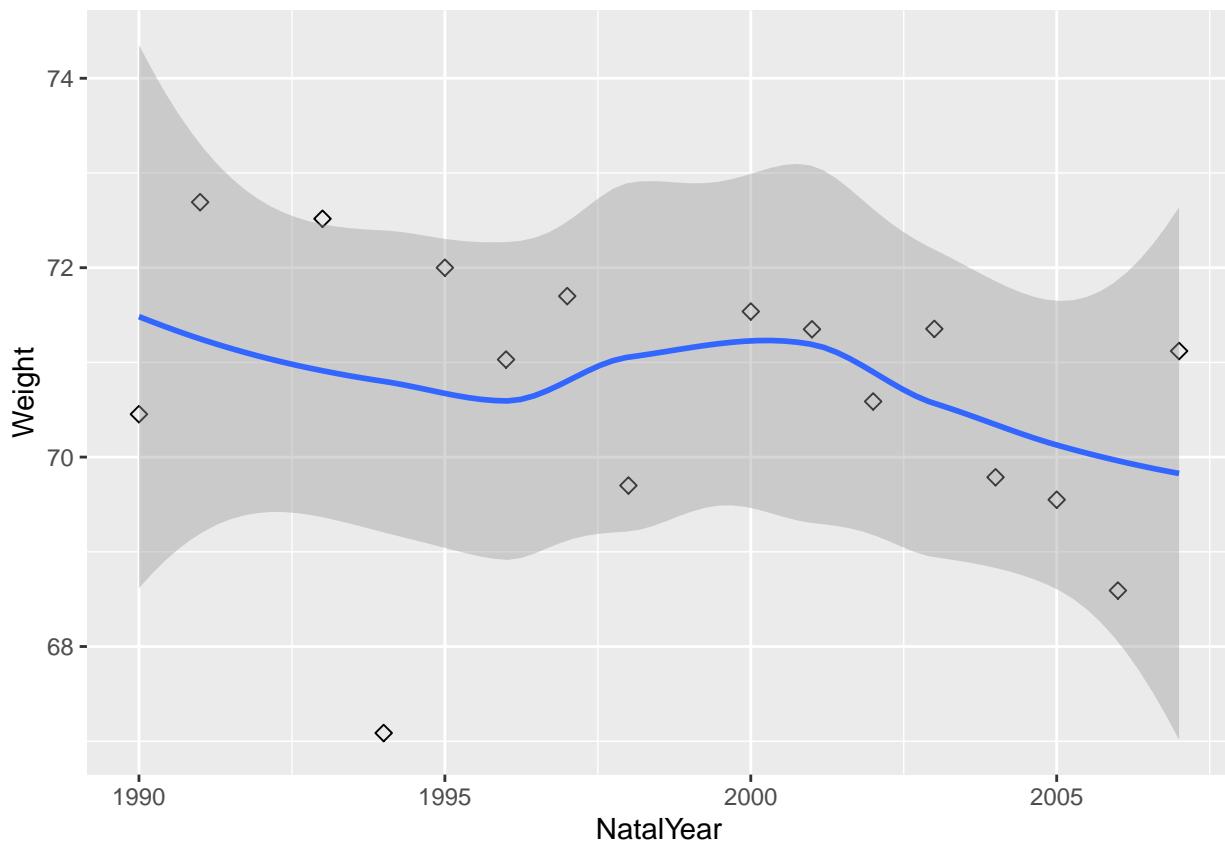
```



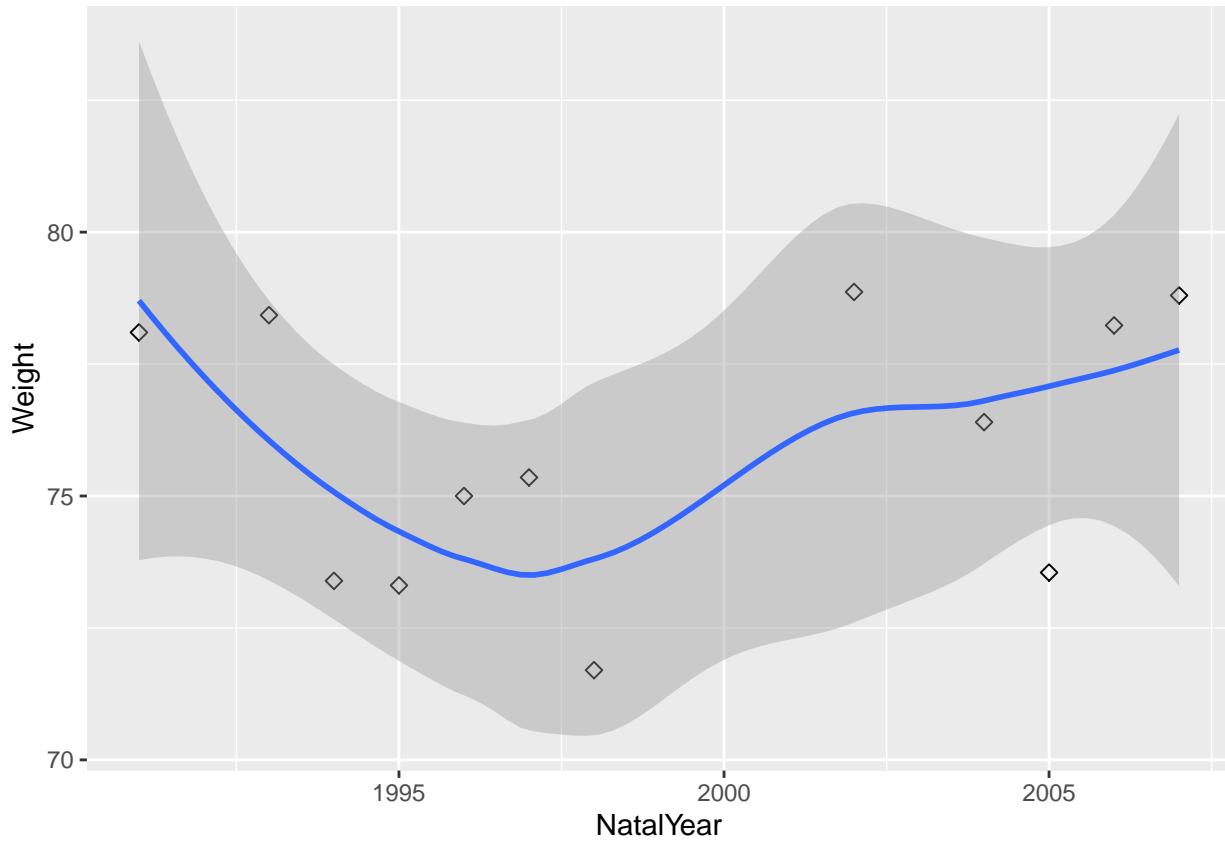
```
ggplot(WeightAgF1, aes(x=NatalYear, y=Weight)) + geom_point(size=2, shape=23) +geom_smooth()  
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



```
ggplot(WeightAgJ1, aes(x=NatalYear, y=Weight)) + geom_point(size=2, shape=23) +geom_smooth()  
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



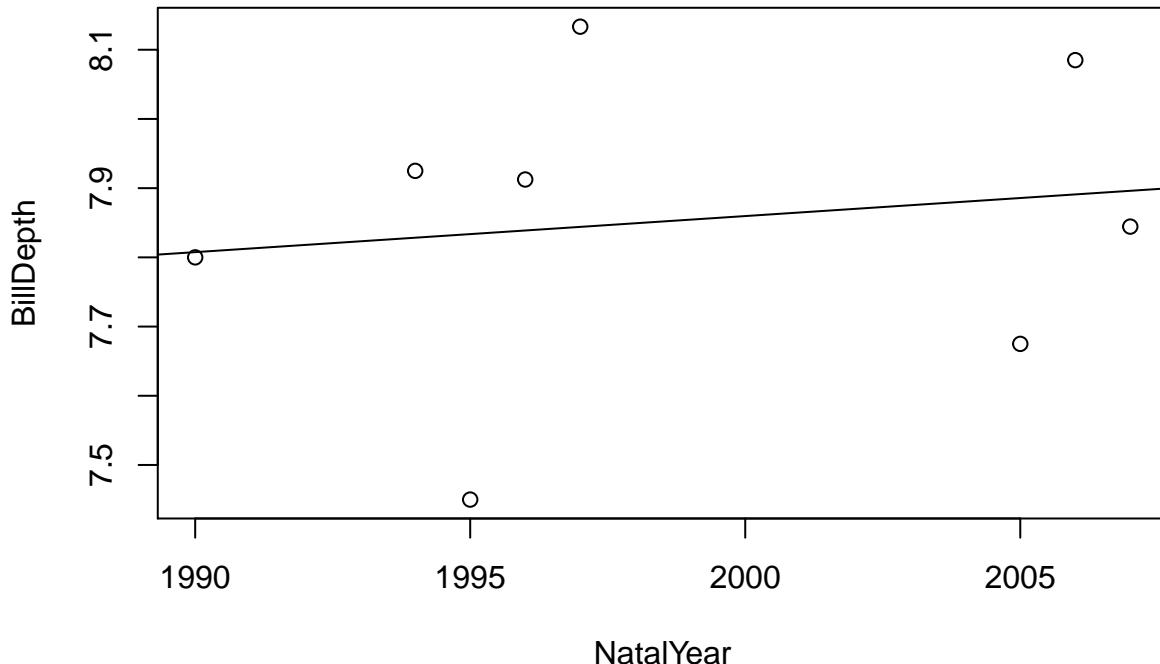
```
ggplot(WeightAg1, aes(x=NatalYear, y=Weight)) + geom_point(size=2, shape=23) +geom_smooth()  
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



#Important to remember that there are no Nestlings for Bill Depth. Bill Depth: Fledglings

```
#View(FledglingsBD) # Has all our pertinent data so far
#View(BDNY) # Has Natal Year
BDF<-merge(FledglingsBD, BDNY) # This works. It has no NAs.
BDAgF<-aggregate(BillDepth~NatalYear, data=BDF, FUN=function(BDF) c(mean=mean(BDF))) # Aggregates the m
##View(BDAgF)
BDAgF1<-subset(BDAgF, BDAgF$NatalYear>=1990) # Only take greater than 1990's
##View(BDAgF1)
plot(BillDepth ~ NatalYear, data=BDAgF1)
lm(BillDepth ~ NatalYear, data=BDAgF1)
```

```
##
## Call:
## lm(formula = BillDepth ~ NatalYear, data = BDAgF1)
##
## Coefficients:
## (Intercept)    NatalYear
## -2.565075     0.005212
abline(lm(BillDepth ~ NatalYear, data=BDAgF1))
```



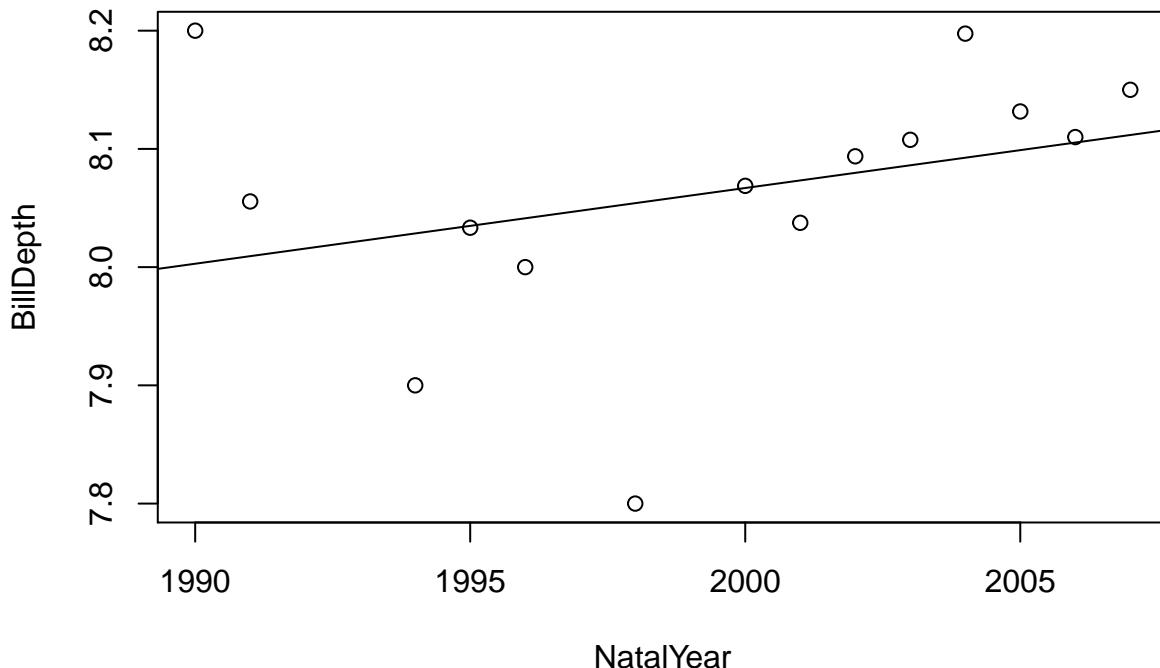
```
BDAgF2<-MannKendall(BDAgF1$BillDepth)
print(BDAgF2)
```

```
## tau = 0.143, 2-sided pvalue =0.71052
```

Bill Depth: Juveniles

```
#View(JuvenilesBD) # Has all our pertinent data so far
#View(BDNY) # Has Natal Year
BDJ<-merge(JuvenilesBD, BDNY) # This works. It has no NAs.
BDAgJ<-aggregate(BillDepth~NatalYear, data=BDJ, FUN=function(BDJ) c(mean=mean(BDJ))) # Aggregates the mean
##View(BDAgF)
BDAgJ1<-subset(BDAgJ, BDAgJ$NatalYear>=1990) # Only take greater than 1990's
##View(BDAgJ1)
plot(BillDepth ~ NatalYear, data=BDAgJ1)
lm(BillDepth ~ NatalYear, data=BDAgJ1)
```

```
##
## Call:
## lm(formula = BillDepth ~ NatalYear, data = BDAgJ1)
##
## Coefficients:
## (Intercept)    NatalYear
## -4.744859      0.006406
abline(lm(BillDepth ~ NatalYear, data=BDAgJ1))
```



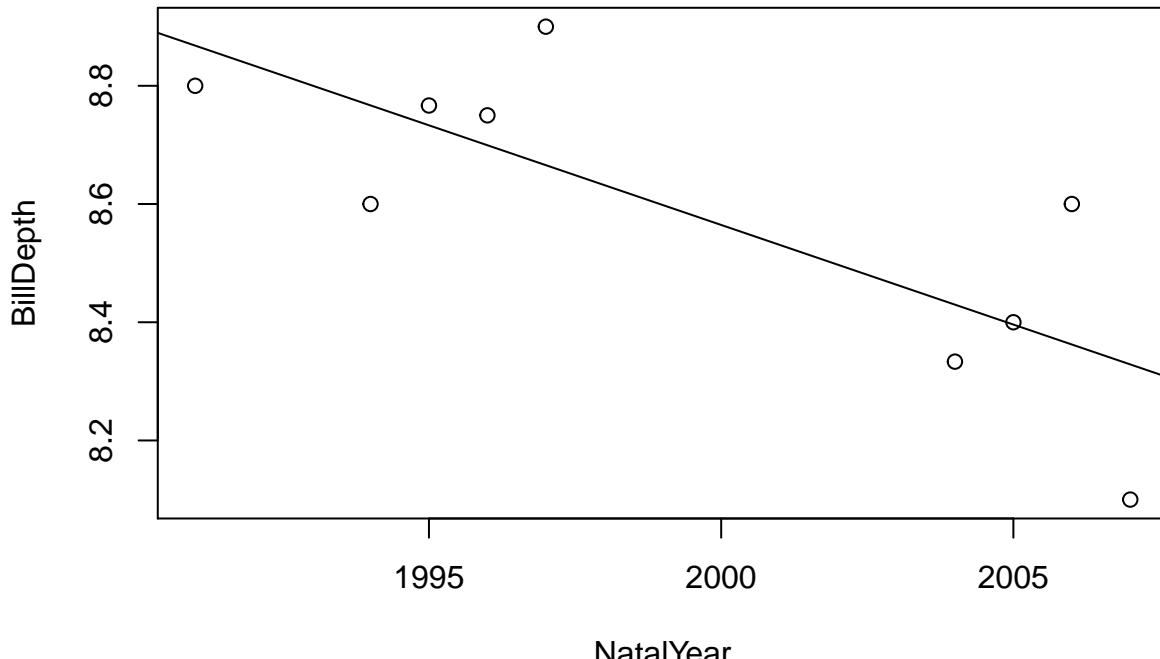
```
BDAgF2<-MannKendall(BDAgJ1$BillDepth)
print(BDAgF2)
```

```
## tau = 0.407, 2-sided pvalue =0.048745
```

Bill Depth: Yearling

```
#View(YearlingsBD) # Has all our pertinent data so far
#View(BDNY) # Has Natal Year
BDY<-merge(YearlingsBD, BDNY) # This works. It has no NAs.
BDAgY<-aggregate(BillDepth~NatalYear, data=BDY, FUN=function(BDY) c(mean=mean(BDY))) # Aggregates the mean
##View(BDAgY)
BDAgY1<-subset(BDAgY, BDAgY$NatalYear>=1990) # Only take greater than 1990's
##View(BDAgY1)
plot(BillDepth ~ NatalYear, data=BDAgY1)
lm(BillDepth ~ NatalYear, data=BDAgY1)
```

```
##
## Call:
## lm(formula = BillDepth ~ NatalYear, data = BDAgY1)
##
## Coefficients:
## (Intercept)    NatalYear
##      75.98421     -0.03371
abline(lm(BillDepth ~ NatalYear, data=BDAgY1))
```



```
BDAgY2<-MannKendall(BDAgY1$BillDepth)
print(BDAgY2)
```

```
## tau = -0.479, 2-sided pvalue =0.093493
```

```
#Important to remember that there are no Nestlings for Bill Width Bill Width: Fledglings
```

```
#View(FledglingsBW) # Has all our pertinent data so far
```

```
#View(BWNY) # Has Natal Year
```

```
BWF<-merge(FledglingsBW, BWNY) # This works. It has no NAs.
```

```
BWAgF<-aggregate(BillWidth~NatalYear, data=BWF, FUN=function(BWF) c(mean=mean(BWF))) # Aggregates the means
```

```
##View(BWAgF)
```

```
BWAgF1<-subset(BWAgF, BWAgF$NatalYear>=1990) # Only take greater than 1990's
```

```
##View(BWAgF1)
```

```
plot(BillWidth ~ NatalYear, data=BWAgF1)
```

```
lm(BillWidth ~ NatalYear, data=BWAgF1)
```

```
##
```

```
## Call:
```

```
## lm(formula = BillWidth ~ NatalYear, data = BWAgF1)
```

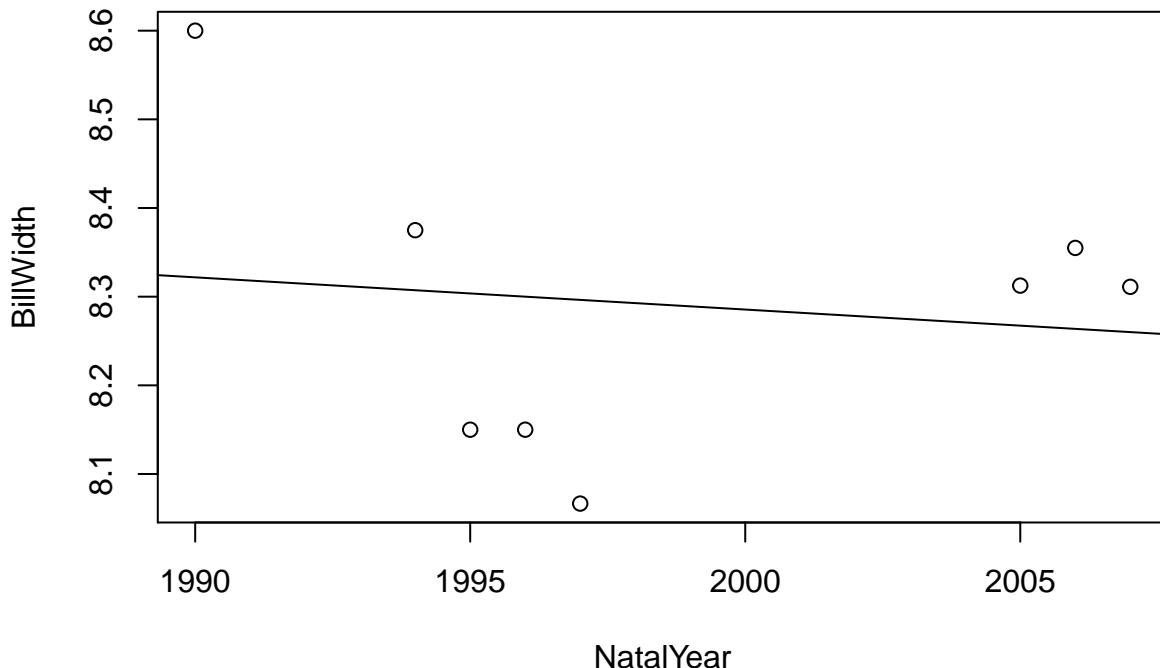
```
##
```

```
## Coefficients:
```

```
## (Intercept) NatalYear
```

```
## 15.547401 -0.003631
```

```
abline(lm(BillWidth ~ NatalYear, data=BWAgF1))
```



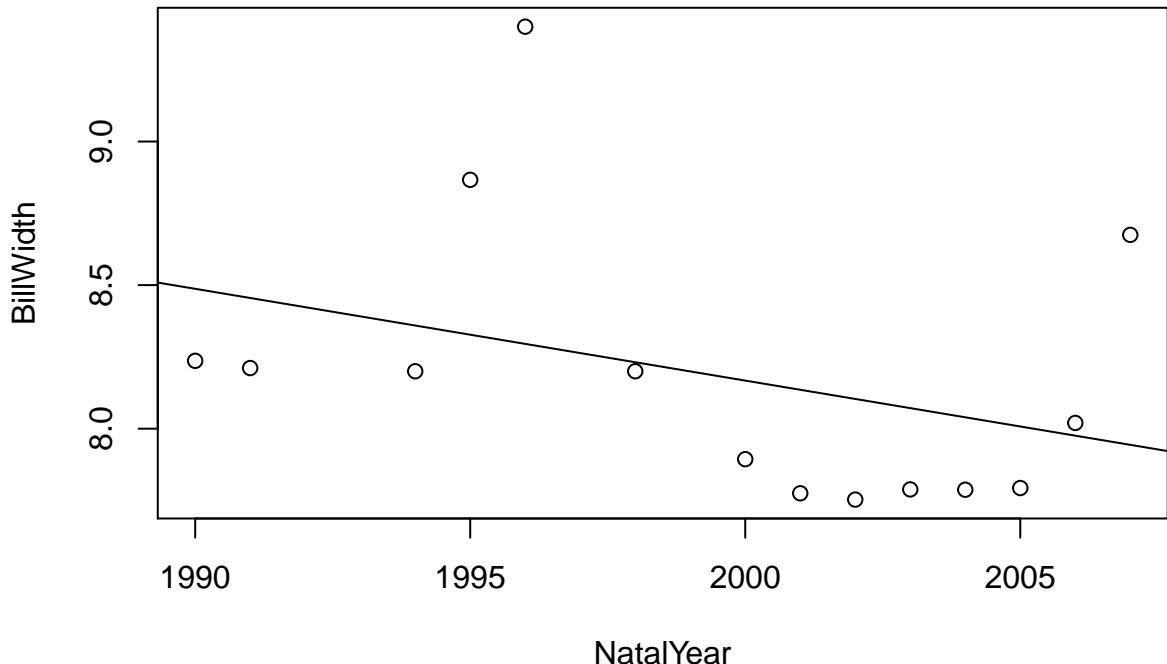
```
BWAgF2<-MannKendall(BWAgF1$BillWidth)
print(BWAgF2)
```

```
## tau = -0.255, 2-sided pvalue =0.45443
```

Bill Width: Juvenile

```
#View(JuvenilesBW) # Has all our pertinent data so far
#View(BWNY) # Has Natal Year
BWJ<-merge(JuvenilesBW, BWNY) # This works. It has no NAs.
BWAj<-aggregate(BillWidth~NatalYear, data=BWJ, FUN=function(BWJ) c(mean=mean(BWJ))) # Aggregates the mean
##View(BWAj)
BWAj1<-subset(BWAj, BWAj$NatalYear>=1990) # Only take greater than 1990's
##View(BWAj1)
plot(BillWidth ~ NatalYear, data=BWAj1)
lm(BillWidth ~ NatalYear, data=BWAj1)

##
## Call:
## lm(formula = BillWidth ~ NatalYear, data = BWAj1)
##
## Coefficients:
## (Intercept)    NatalYear
##      72.09230     -0.03196
abline(lm(BillWidth ~ NatalYear, data=BWAj1))
```



```
BWAgJ2<-MannKendall(BWAgJ1$BillWidth)
print(BWAgJ2)
```

```
## tau = -0.287, 2-sided pvalue =0.17048
```

Bill Width: Yearling

```
#View(YearlingsBW) # Has all our pertinent data so far
```

```
#View(BWNY) # Has Natal Year
```

```
BWY<-merge(YearlingsBW, BWNY) # This works. It has no NAs.
```

```
BWAgY<-aggregate(BillWidth~NatalYear, data=BWY, FUN=function(BWY) c(mean=mean(BWY))) # Aggregates the mean
```

```
##View(BWAgY)
```

```
BWAgY1<-subset(BWAgY, BWAgY$NatalYear>=1990) # Only take greater than 1990's
```

```
##View(BWAgY1)
```

```
plot(BillWidth ~ NatalYear, data=BWAgY1)
```

```
lm(BillWidth ~ NatalYear, data=BWAgY1)
```

```
##
```

## Call:

```
## lm(formula = BillWidth ~ NatalYear, data = BWAgY1)
```

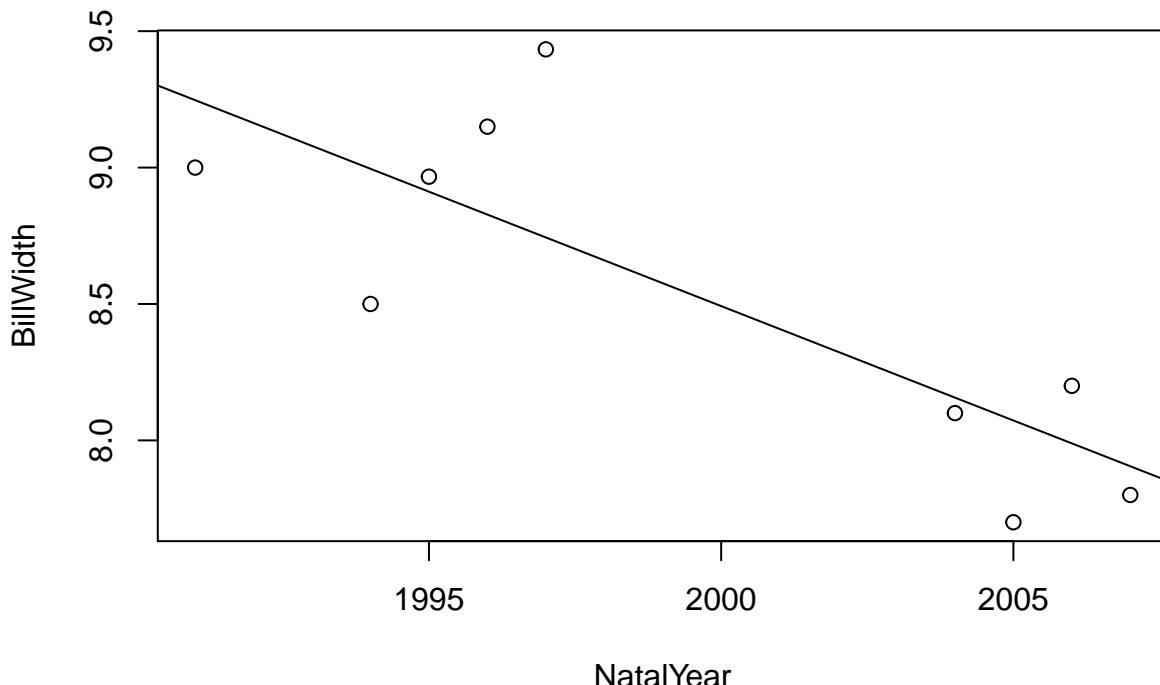
```
##
```

## Coefficients:

```
## (Intercept) NatalYear
```

```
## 176.21818 -0.08386
```

```
abline(lm(BillWidth ~ NatalYear, data=BWAgY1))
```



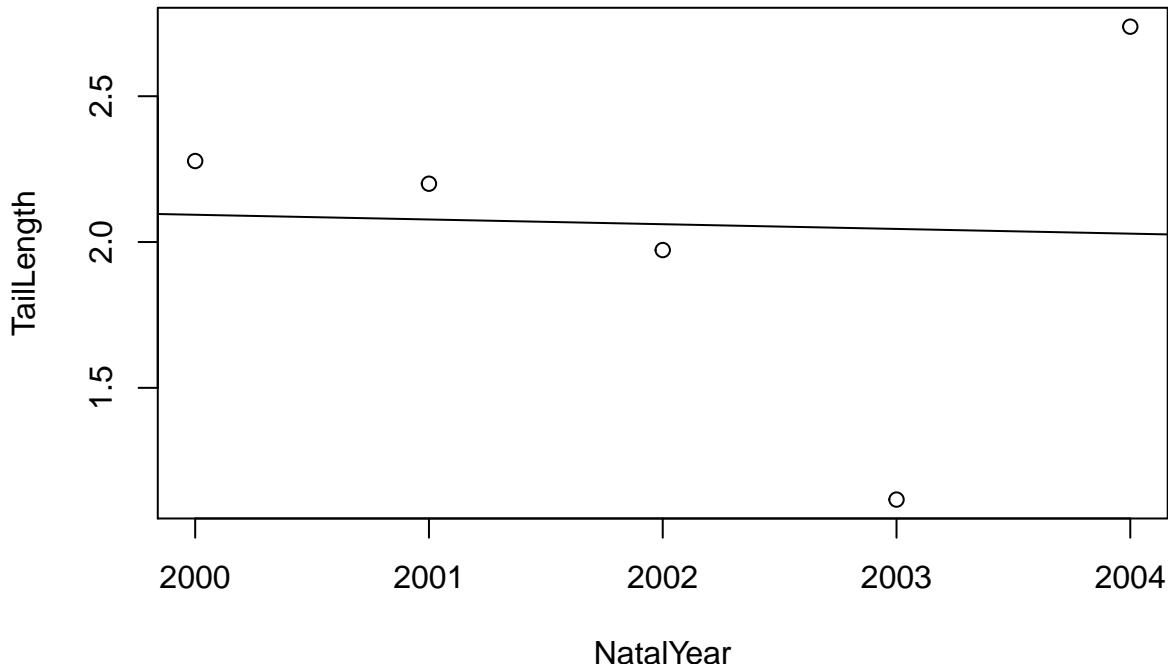
```
BWAgY2<-MannKendall(BWAgY1$BillWidth)
print(BWAgY2)
```

```
## tau = -0.389, 2-sided pvalue =0.17531
```

Tail Length: Nestling

```
#View(NestlingsTL) # Has all our pertinent data so far
#View(TLN) # Has Natal Year
TLN<-merge(NestlingsTL, TLNY) # This works. It has no NAs.
TLAGN<-aggregate(TailLength~NatalYear, data=TLN, FUN=function(TLN) c(mean=mean(TLN))) # Aggregates the
##View(TLAGN)
TLAGN1<-subset(TLAGN, TLAGN$NatalYear>=1990) # Only take greater than 1990's
##View(TLAGN1)
plot(TailLength ~ NatalYear, data=TLAGN1)
lm(TailLength ~ NatalYear, data=TLAGN1)
```

```
##
## Call:
## lm(formula = TailLength ~ NatalYear, data = TLAGN1)
##
## Coefficients:
## (Intercept)    NatalYear
##      34.45156     -0.01618
abline(lm(TailLength ~ NatalYear, data=TLAGN1))
```



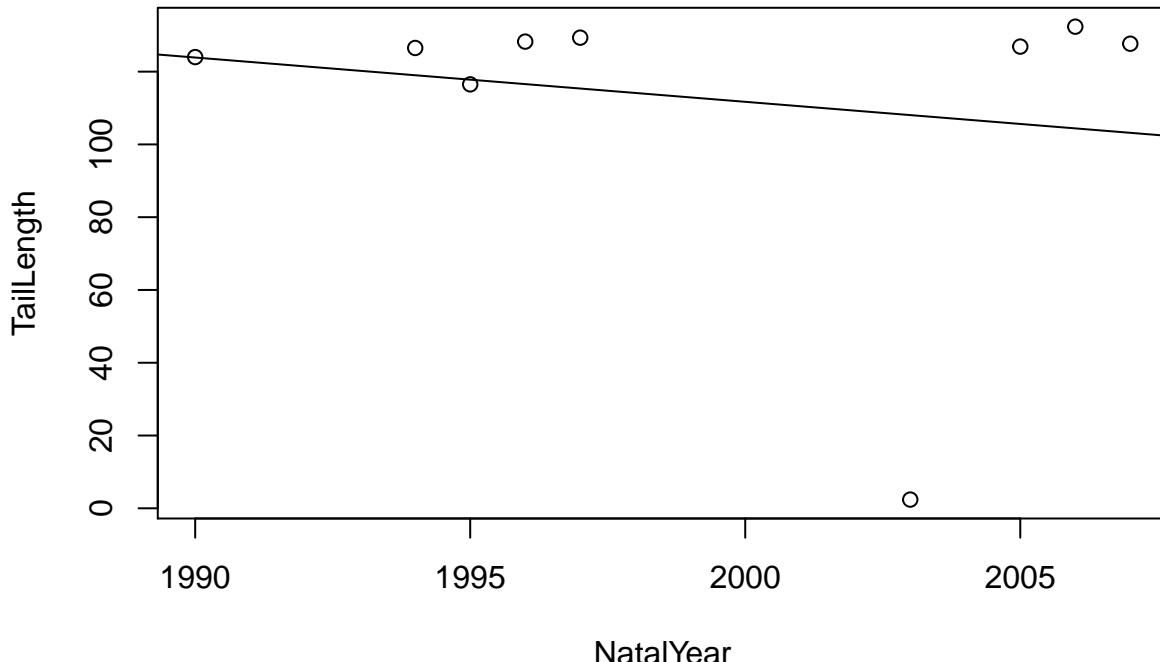
```
TLAgN2<-MannKendall(TLAgN1$TailLength)
print(TLAgN2)
```

```
## tau = -0.2, 2-sided pvalue =0.8065
```

Tail Length: Fledglings

```
#View(FledglingsTL) # Has all our pertinent data so far
#View(TLNY) # Has Natal Year
TLF<-merge(FledglingsTL, TLNY) # This works. It has no NAs.
TLAgF<-aggregate(TailLength~NatalYear, data=TLF, FUN=function(TLF) c(mean=mean(TLF))) # Aggregates the data
##View(TLAgF)
TLAgF1<-subset(TLAgF, TLAgF$NatalYear>=1990) # Only take greater than 1990's
##View(TLAgF1)
plot(TailLength ~ NatalYear, data=TLAgF1)
lm(TailLength ~ NatalYear, data=TLAgF1)

##
## Call:
## lm(formula = TailLength ~ NatalYear, data = TLAgF1)
##
## Coefficients:
## (Intercept)    NatalYear
##      2545.905     -1.217
abline(lm(TailLength ~ NatalYear, data=TLAgF1))
```



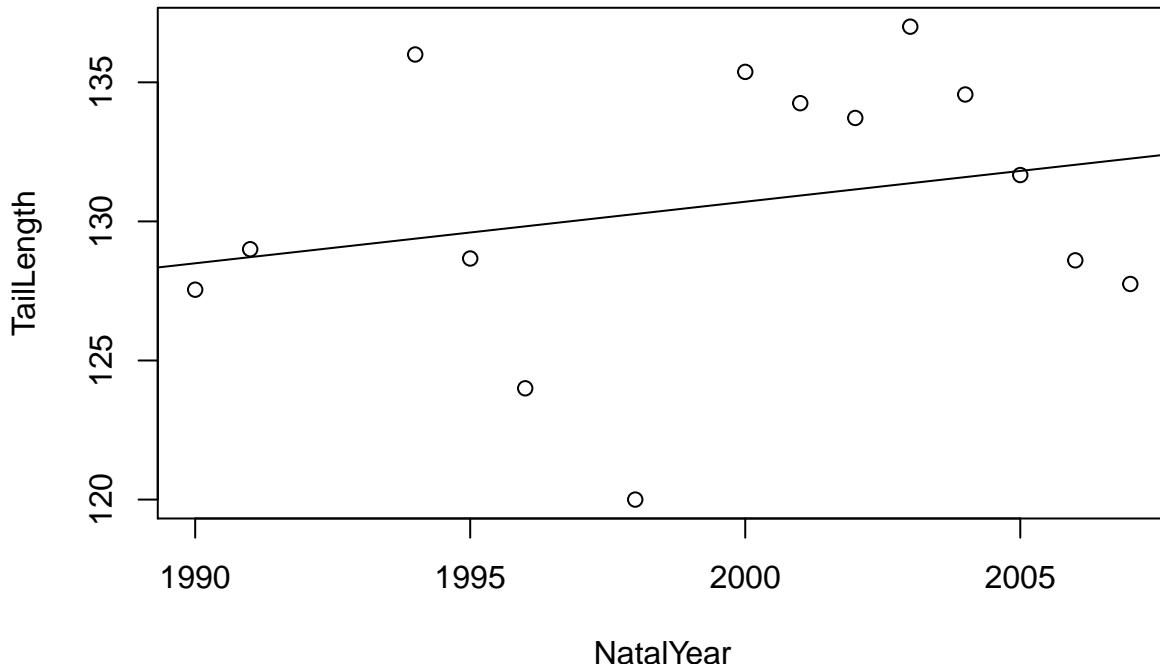
```
TLAgF2<-MannKendall(TLAgF1$TailLength)
print(TLAgF2)
```

```
## tau = 0.333, 2-sided pvalue =0.25145
```

Tail Length: Juvenile

```
#View(JuvenilesTL) # Has all our pertinent data so far
#View(TLNY) # Has Natal Year
TLJ<-merge(JuvenilesTL, TLNY) # This works. It has no NAs.
TLAgJ<-aggregate(TailLength~NatalYear, data=TLJ, FUN=function(TLJ) c(mean=mean(TLJ))) # Aggregates the data
##View(TLAgJ)
TLAgJ1<-subset(TLAgJ, TLAgJ$NatalYear>=1990) # Only take greater than 1990's
##View(TLAgJ1)
plot(TailLength ~ NatalYear, data=TLAgJ1)
lm(TailLength ~ NatalYear, data=TLAgJ1)
```

```
##
## Call:
## lm(formula = TailLength ~ NatalYear, data = TLAgJ1)
##
## Coefficients:
## (Intercept)    NatalYear
## -312.0464      0.2214
abline(lm(TailLength ~ NatalYear, data=TLAgJ1))
```



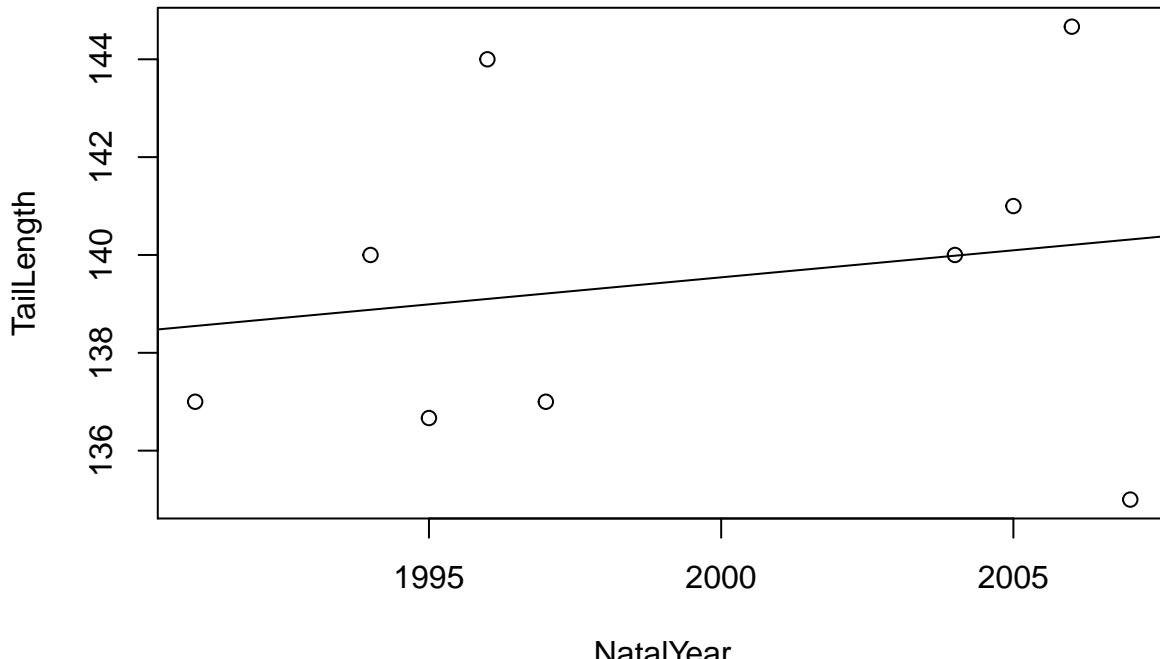
```
TLAgJ2<-MannKendall(TLAgJ1$TailLength)
print(TLAgJ2)
```

```
## tau = 0.011, 2-sided pvalue =1
```

Tail Length: Yearling

```
#View(YearlingsTL) # Has all our pertinent data so far
#View(TLNY) # Has Natal Year
TLY<-merge(YearlingsTL, TLNY) # This works. It has no NAs.
TLAgY<-aggregate(TailLength~NatalYear, data=TLY, FUN=function(TLY) c(mean=mean(TLY))) # Aggregates the data
##View(TLAgY)
TLAgY1<-subset(TLAgY, TLAgY$NatalYear>=1990) # Only take greater than 1990's
#View(TLAgY1)
plot(TailLength ~ NatalYear, data=TLAgY1)
lm(TailLength ~ NatalYear, data=TLAgY1)
```

```
##
## Call:
## lm(formula = TailLength ~ NatalYear, data = TLAgY1)
##
## Coefficients:
## (Intercept)    NatalYear
##      -81.4883       0.1105
abline(lm(TailLength ~ NatalYear, data=TLAgY1))
```



```
TLAgY2<-MannKendall(TLAgY1$TailLength)
print(TLAgY2)
```

```
## tau = 0.171, 2-sided pvalue =0.59816
```

Making a 2x2 display for research presentation.

```
par(mfrow=c(2,2))
plot(TailLength ~ NatalYear, data=TLAgN1, xlab = "Year Born (Nestling)", ylab="Mean Tail Length")
lm(TailLength ~ NatalYear, data=TLAgN1)
```

```
##
## Call:
## lm(formula = TailLength ~ NatalYear, data = TLAgN1)
##
## Coefficients:
## (Intercept)    NatalYear
##      34.45156     -0.01618
```

```
abline(lm(TailLength ~ NatalYear, data=TLAgN1, xlab = "Year Born (Nestling)", ylab="Mean Tail Length"))
```

```
## Warning: In lm.fit(x, y, offset = offset, singular.ok = singular.ok, ...):
##   extra arguments 'xlab', 'ylab' will be disregarded
```

```
plot(TailLength ~ NatalYear, data=TLAgF1)
lm(TailLength ~ NatalYear, data=TLAgF1)
```

```
##
## Call:
## lm(formula = TailLength ~ NatalYear, data = TLAgF1)
##
## Coefficients:
## (Intercept)    NatalYear
##      2545.905     -1.217
```

```

abline(lm(TailLength ~ NatalYear, data=TLAgF1, xlab = "Year Born (Nestling)", ylab="Mean Tail Length"))

## Warning: In lm.fit(x, y, offset = offset, singular.ok = singular.ok, ...) :
##   extra arguments 'xlab', 'ylab' will be disregarded
plot(TailLength ~ NatalYear, data=TLAgJ1)
lm(TailLength ~ NatalYear, data=TLAgJ1)

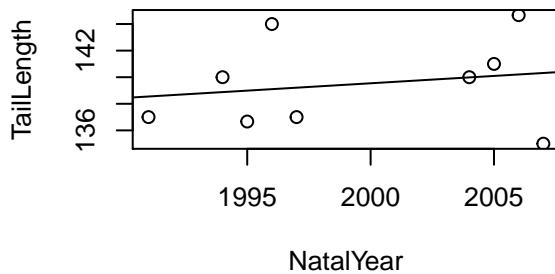
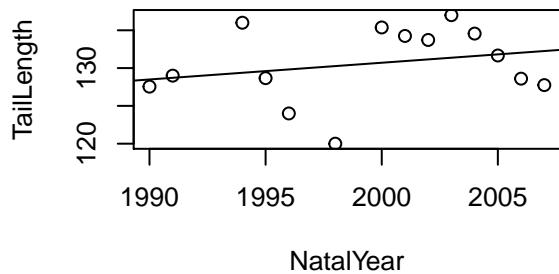
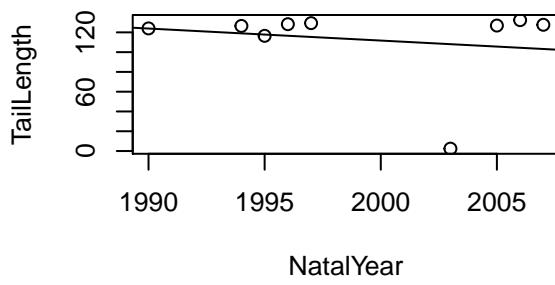
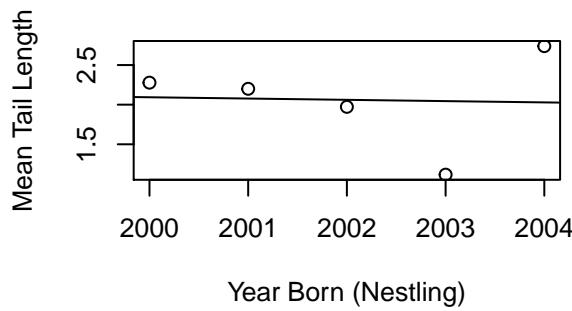
##
## Call:
## lm(formula = TailLength ~ NatalYear, data = TLAGJ1)
##
## Coefficients:
## (Intercept)    NatalYear
## -312.0464      0.2214
abline(lm(TailLength ~ NatalYear, data=TLAgJ1, xlab = "Year Born (Nestling)", ylab="Mean Tail Length"))

## Warning: In lm.fit(x, y, offset = offset, singular.ok = singular.ok, ...) :
##   extra arguments 'xlab', 'ylab' will be disregarded
plot(TailLength ~ NatalYear, data=TLAgY1)
lm(TailLength ~ NatalYear, data=TLAgY1)

##
## Call:
## lm(formula = TailLength ~ NatalYear, data = TLAGY1)
##
## Coefficients:
## (Intercept)    NatalYear
## -81.4883      0.1105
abline(lm(TailLength ~ NatalYear, data=TLAgY1, xlab = "Year Born (Nestling)", ylab="Mean Tail Length"))

## Warning: In lm.fit(x, y, offset = offset, singular.ok = singular.ok, ...) :
##   extra arguments 'xlab', 'ylab' will be disregarded

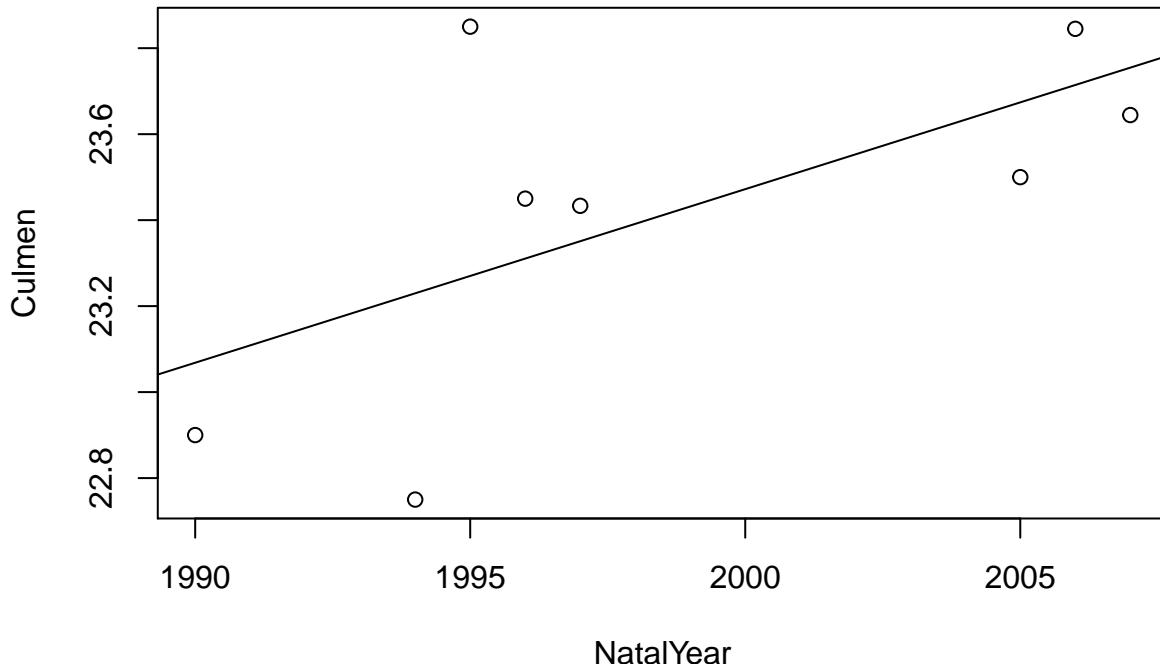
```



#Important to remember that there are no Nestlings for Bill Width Culmen: Fledglings

```
#View(FledglingsCM) # Has all our pertinent data so far
#View(CulmenNY) # Has Natal Year
CF<-merge(CulmenNY, FledglingsCM) # This works. It has no NAs.
CAgF<-aggregate(Culmen~NatalYear, data=CF, FUN=function(CF) c(mean=mean(CF))) # Aggregates the mean for
##View(CAgF)
CAgF1<-subset(CAgF, CAgF$NatalYear>=1990) # Only take greater than 1990's
##View(CAgF1)
plot(Culmen ~ NatalYear, data=CAgF1)
lm(Culmen ~ NatalYear, data=CAgF1)
```

```
##
## Call:
## lm(formula = Culmen ~ NatalYear, data = CAgF1)
##
## Coefficients:
## (Intercept)    NatalYear
##      -57.28337     0.04038
abline(lm(Culmen ~ NatalYear, data=CAgF1))
```



```
CAgF2<-MannKendall(CAgF1$Culmen)
print(CAgF2)
```

```
## tau = 0.429, 2-sided pvalue =0.17355
```

Culmen: Juvenile

```
#View(JuvenilesCM) # Has all our pertinent data so far
#View(CulmenNY) # Has Natal Year
```

```
CJ<-merge(CulmenNY, JuvenilesCM) # This works. It has no NAs.
```

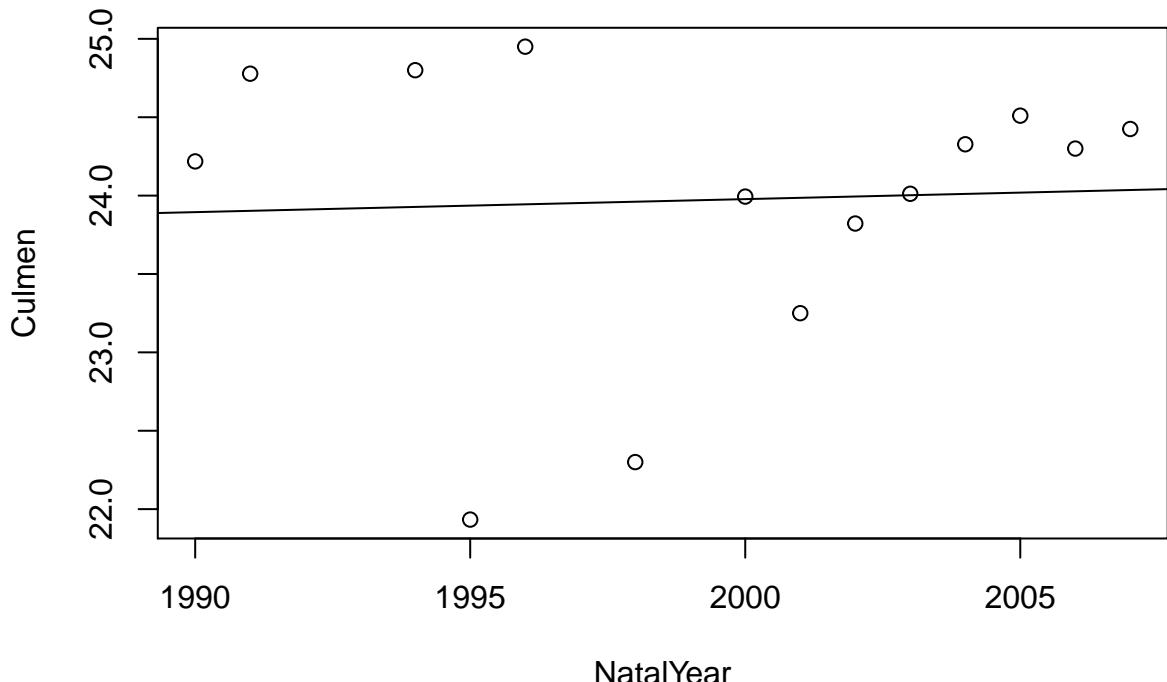
```
CAgJ<-aggregate(Culmen~NatalYear, data=CJ, FUN=function(CJ) c(mean=mean(CJ))) # Aggregates the mean for
##View(CAgJ)
```

```
CAgJ1<-subset(CAgJ, CAgJ$NatalYear>=1990) # Only take greater than 1990's
```

```
#View(CAgJ1)
```

```
plot(Culmen ~ NatalYear, data=CAgJ1)
lm(Culmen ~ NatalYear, data=CAgJ1)
```

```
##
## Call:
## lm(formula = Culmen ~ NatalYear, data = CAgJ1)
##
## Coefficients:
## (Intercept)    NatalYear
##      7.385151     0.008296
abline(lm(Culmen ~ NatalYear, data=CAgJ1))
```



```
CAGJ2<-MannKendall(CAGJ1$Culmen)
print(CAGJ2)
```

```
## tau = 0.121, 2-sided pvalue = 0.58407
```

Culmen: Yearling

```
#View(YearlingsCM) # Has all our pertinent data so far
#View(CulmenNY) # Has Natal Year
```

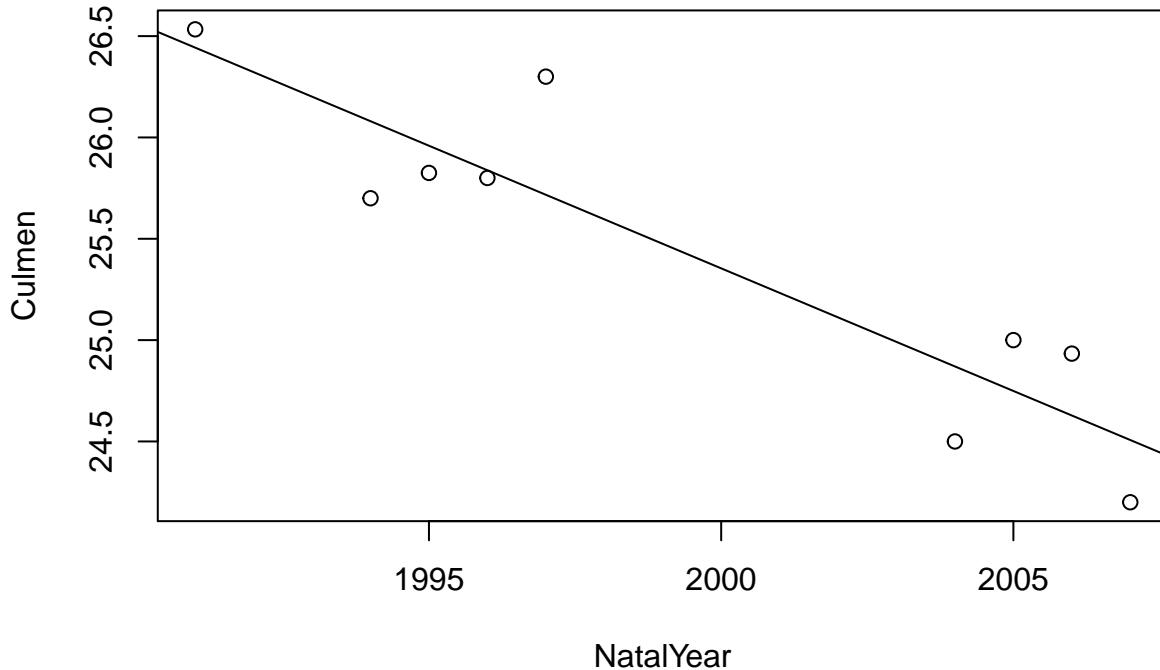
```
CY<-merge(YearlingsCM, CulmenNY) # This works. It has no NAs.
```

```
CAGY<-aggregate(Culmen~NatalYear, data=CY, FUN=function(CY) c(mean=mean(CY))) # Aggregates the mean for
#View(CAGY)
```

```
CAGY1<-subset(CAGY, CAGY$NatalYear>=1990) # Only take greater than 1990's
#View(CAGY1)
```

```
plot(Culmen ~ NatalYear, data=CAGY1)
lm(Culmen ~ NatalYear, data=CAGY1)
```

```
##
## Call:
## lm(formula = Culmen ~ NatalYear, data = CAGY1)
##
## Coefficients:
## (Intercept)    NatalYear
##      267.365     -0.121
abline(lm(Culmen ~ NatalYear, data=CAGY1))
```



```
CAGY2<-MannKendall(CAGY1$Culmen)
print(CAGY2)

## tau = -0.611, 2-sided pvalue =0.028568
```

Preparing for Regression Analysis:

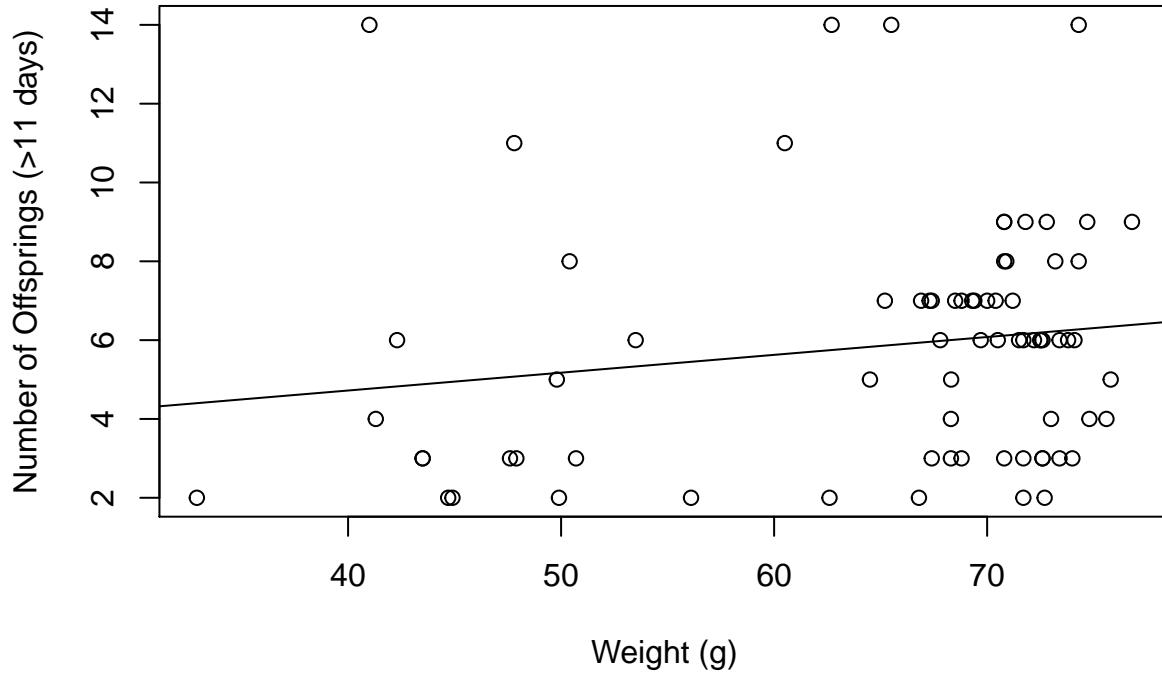
Weight with Overall Reproductive Success Looking at dead Jays and the amount of day 11's they managed to produce.

```
#Using hy.stats to see parents
#View(hy.stats)
R1<-merge(hy.stats, WeightNY, by="USFWS") #This has a lot of duplicates for some reason. Removing those
R2<-R1[!duplicated(R1$USFWS), ]
R3<-aggregate(USFWS~mom, data=R2, FUN=function(R2) c(length=length((R2)))) # Counts the number of fledglings
#We may require to change the names of the columns as one of them is now called mom and should be called
colnames(R3)<-c("USFWS", "Offsprings")
R4<-merge(R3, WeightMerge, by="USFWS")
##View(WeightMerge)
#View(R4)
plot(R4$Offsprings~R4$Weight, xlab="Weight (g)", ylab="Number of Offsprings (>11 days)", main="Female Jays Reproductive Success")
WeightModel<-lm(R4$Offsprings~R4$Weight)
print(WeightModel)

##
## Call:
## lm(formula = R4$Offsprings ~ R4$Weight)
##
## Coefficients:
## (Intercept)    R4$Weight
##           2.91215     0.04524
```

```
abline(lm(R4$Offsprings~R4$Weight))
```

## Female Jays' Lifetime Reproductive Success



```
summary(WeightModel)
```

```
##  
## Call:  
## lm(formula = R4$Offsprings ~ R4$Weight)  
##  
## Residuals:  
##      Min       1Q   Median       3Q      Max  
## -4.201 -2.323 -0.185  1.119  9.233  
##  
## Coefficients:  
##             Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 2.91215  2.13276  1.365   0.176  
## R4$Weight    0.04524  0.03237  1.397   0.167  
##  
## Residual standard error: 3.04 on 72 degrees of freedom  
## Multiple R-squared:  0.02641,   Adjusted R-squared:  0.01288  
## F-statistic: 1.953 on 1 and 72 DF,  p-value: 0.1666
```

Piecing it all together: 1) Replicate the access\$metrics dataset and work with that from now. Immediately, merge it with the Jay database to get other pertinent information by using USFWS. 2) Filter out the non-Demo categories and NatalYear<1990. The run the loop to categorise them by age. (Ensure that Nestling==11) *MAKE A TABLE WITH ALL THE PHENOTYPES AND THEIR SAMPLE SIZES* 3) Aggregate the means by natal year and run Mann-Kendall. Use the Loess lines and ggplot2() to illustrate it. 4) Merge hy.stats to run a logistic regression on whether or not the Jays survived as a function of the phenotype of interest. (Use this: <https://towardsdatascience.com/implementing-binary-logistic-regression-in-r-7d802a9d98fe>) 5) Write

the paper! #GrandTable: All filtering steps except for final removal of Jays whose age is unknown (has 10442 entries). This one can be used for Regression. #GrandTable1: All filtering steps including the removal of Jays whose age is unknown (has 7571 entries). This one can be used for MannKendall.

```
GrandTable<-access$metrics # Replicating access$metrics
GrandTable<-merge(GrandTable, JAYS, by="USFWS") # Merging with JAYS to get a increased breadth and dept.
remove <- c("ATNT", "BEAR", "CHGR", "CHIK", "CIRB", "CIRC", "COTG", "EAGB", "EAGR", "FTZP", "GRVR", "GR")
GrandTable<-GrandTable[!GrandTable$Terr %in% remove,] # Removal sucessful!
# Now to ensure that we are controlling for the correct year because Demo study tract stabilised around
GrandTable<-GrandTable[GrandTable$NatalYear >= "1990",]
# Now we age them. All day 11s get the Nestling label. Everything below that is removed. Everything above
GrandTable$AgeCategory<-NA #adding a new column to prepare for categorical ages such as juveniles, year
GrandTable$AgeCategory[GrandTable$AgeMeas==11]<- "Nestling"
GrandTable$AgeCategory[11<GrandTable$AgeMeas & GrandTable$AgeMeas<=70]<- "Fledgling"
GrandTable$AgeCategory[70<GrandTable$AgeMeas & GrandTable$AgeMeas<=365]<- "Juveniles"
GrandTable$AgeCategory[365<GrandTable$AgeMeas & GrandTable$AgeMeas<=3650]<- "Yearling"
#This gives an age category to every Jay in dataframe. Whichever Jay does not have a Age Category is NA
#Writing a function to remove all NAs from age and naming it GrandTable1. GrandTable remains the original
completeFun <- function(data, desiredCols) {
  completeVec <- complete.cases(data[, desiredCols])
  return(data[completeVec, ])
}
GrandTable1<-completeFun(GrandTable, "AgeCategory")
```

Counting and making tables of phenotypes:

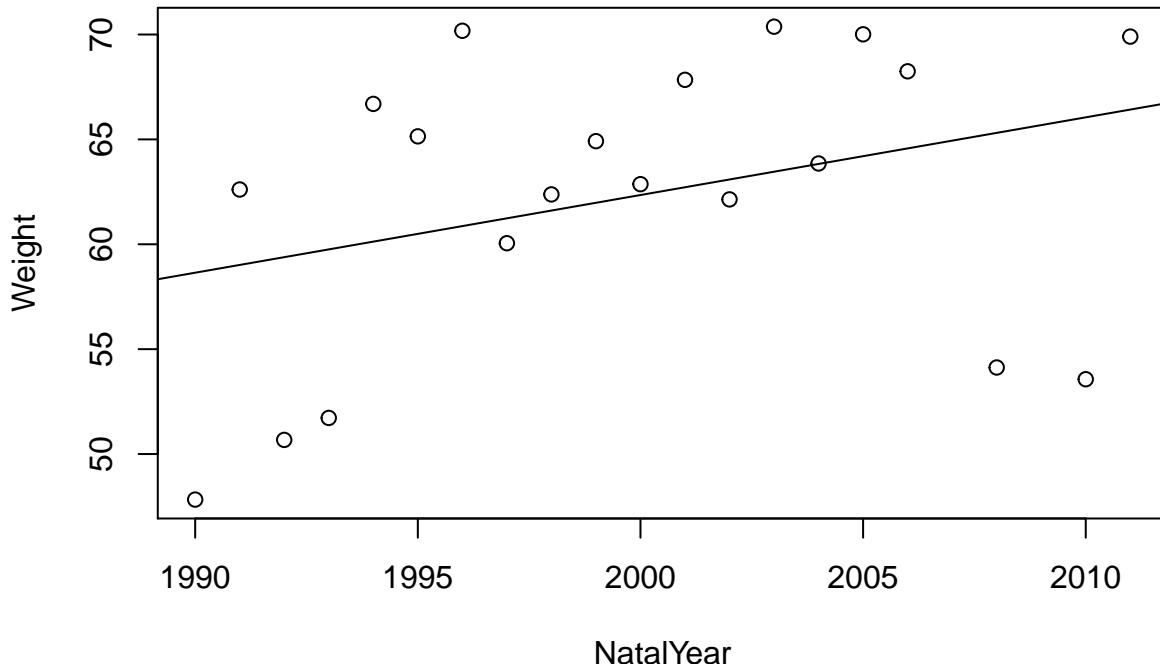
```
#Starting with weight
library(Kendall)
GT1WF<-subset(GrandTable1, AgeCategory=="Fledgling")
GT1WF<-GT1WF[,c("Weight", "NatalYear")]
GT1WFm<-aggregate(Weight~NatalYear, data=GT1WF, FUN=function(GT1WF) c(mean=mean(GT1WF)))
x<-MannKendall(GT1WFm$Weight)
summary(x)

## Score = 50 , Var(Score) = 950
## denominator = 190
## tau = 0.263, 2-sided pvalue = 0.11189

plot(Weight ~ NatalYear, data=GT1WFm)
lm(Weight ~ NatalYear, data=GT1WFm)

##
## Call:
## lm(formula = Weight ~ NatalYear, data = GT1WFm)
##
## Coefficients:
## (Intercept)    NatalYear
## -678.3600      0.3704

abline(lm(Weight ~ NatalYear, data=GT1WFm))
```



```
#GT1WF<-na.omit(GT1WF$Weight)
GT1WN<-subset(GrandTable1, AgeCategory=="Nestling")
GT1WN<-na.omit(GT1WN$Weight)
GT1WJ<-subset(GrandTable1, AgeCategory=="Juveniles")
GT1WJ<-na.omit(GT1WJ$Weight)
GT1WY<-subset(GrandTable1, AgeCategory=="Yearling")
GT1WY<-na.omit(GT1WY$Weight)
```

```
#Starting with BillDepth
GT1BDF<-subset(GrandTable1, AgeCategory=="Fledgling")
GT1BDF<-na.omit(GT1BDF$BillDepth)
GT1BDN<-subset(GrandTable1, AgeCategory=="Nestling")
GT1BDN<-na.omit(GT1BDN$BillDepth)
GT1BDJ<-subset(GrandTable1, AgeCategory=="Juveniles")
GT1BDJ<-na.omit(GT1BDJ$BillDepth)
GT1BDY<-subset(GrandTable1, AgeCategory=="Yearling")
GT1BDY<-na.omit(GT1BDY$BillDepth)
```

```
#Starting with BillWidth
GT1BWF<-subset(GrandTable1, AgeCategory=="Fledgling")
GT1BWF<-na.omit(GT1BWF$BillWidth)
GT1BWN<-subset(GrandTable1, AgeCategory=="Nestling")
GT1BWN<-na.omit(GT1BWN$BillWidth)
GT1BWJ<-subset(GrandTable1, AgeCategory=="Juveniles")
GT1BWJ<-na.omit(GT1BWJ$BillWidth)
GT1BWY<-subset(GrandTable1, AgeCategory=="Yearling")
GT1BWY<-na.omit(GT1BWY$BillWidth)
```

```
#Starting with TailLength
GT1TLF<-subset(GrandTable1, AgeCategory=="Fledgling")
GT1TLF<-na.omit(GT1TLF$TailLength)
GT1TLN<-subset(GrandTable1, AgeCategory=="Nestling")
GT1TLN<-na.omit(GT1TLN$TailLength)
```

```

GT1TLJ<-subset(GrandTable1, AgeCategory=="Juveniles")
GT1TLJ<-na.omit(GT1TLJ$TailLength)
GT1TLY<-subset(GrandTable1, AgeCategory=="Yearling")
GT1TLY<-na.omit(GT1TLY$TailLength)

#Starting with Culmen
GT1CF<-subset(GrandTable1, AgeCategory=="Fledgling")
GT1CF<-na.omit(GT1CF$Culmen)
GT1CN<-subset(GrandTable1, AgeCategory=="Nestling")
GT1CN<-na.omit(GT1CN$Culmen)
GT1CJ<-subset(GrandTable1, AgeCategory=="Juveniles")
GT1CJ<-na.omit(GT1CJ$Culmen)
GT1CY<-subset(GrandTable1, AgeCategory=="Yearling")
GT1CY<-na.omit(GT1CY$Culmen)

#Starting with HeadLength
GT1CF<-subset(GrandTable1, AgeCategory=="Fledgling")
GT1CF<-na.omit(GT1CF$HeadBreadth)
GT1CN<-subset(GrandTable1, AgeCategory=="Nestling")
GT1CN<-na.omit(GT1CN$HeadBreadth)
GT1CJ<-subset(GrandTable1, AgeCategory=="Juveniles")
GT1CJ<-na.omit(GT1CJ$HeadBreadth)
GT1CY<-subset(GrandTable1, AgeCategory=="Yearling")
GT1CY<-na.omit(GT1CY$HeadBreadth)

PhenotypeTable <- matrix(c(2825,0,0,946,0,0,0,3771,685,419,417,471,409,418,1,2820,1766,1217,1216,1181,1
colnames(PhenotypeTable) <- c("Nestling","Fledgling","Juvenile","Yearling", "Row Total")
rownames(PhenotypeTable) <- c("Weight","Bill Depth","Bill Width","Tail Length","Culmen","Head Length",
PhenotypeTable <- as.table(PhenotypeTable)
PhenotypeTable

##          Nestling Fledgling Juvenile Yearling Row Total
## Weight      2825       685     1766    2270    7546
## Bill Depth     0        419     1217     419    2055
## Bill Width     0        417     1216     419    2052
## Tail Length   946       471     1181     282    2880
## Culmen        0        409     1180     437    2026
## Head Length    0        418     1219     442    2079
## HeadBreadth     0         1      456     142     599
## Column Total  3771      2820     8235    4411   19237

#Lets make a prettier table
#Install the relevant libraries - do this one time
install.packages("data.table")
install.packages("dplyr")

## Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/3.6'
## (as 'lib' is unspecified)
install.packages("formattable")

## Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/3.6'
## (as 'lib' is unspecified)
install.packages("tidyverse")

## Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/3.6'

```

```

## (as 'lib' is unspecified)
#Load the libraries
library(data.table)
library(dplyr)

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:data.table':
## 
##     between, first, last

## The following objects are masked from 'package:plyr':
## 
##     arrange, count, desc, failwith, id, mutate, rename, summarise,
##     summarise

## The following objects are masked from 'package:stats':
## 
##     filter, lag

## The following objects are masked from 'package:base':
## 
##     intersect, setdiff, setequal, union

library(formattable)
library(tidyr)
#Set a few color variables to make our table more visually appealing
customGreen0 = "#DeF7E9"
customGreen = "#71CA97"
customRed = "#ff7f7f"

#FPhenotypeTable<-formattable(PhenotypeTable)
#FPhenotypeTable
library(knitr)
kable(PhenotypeTable)

```

	Nestling	Fledgling	Juvenile	Yearling	Row Total
Weight	2825	685	1766	2270	7546
Bill Depth	0	419	1217	419	2055
Bill Width	0	417	1216	419	2052
Tail Length	946	471	1181	282	2880
Culmen	0	409	1180	437	2026
Head Length	0	418	1219	442	2079
HeadBreadth	0	1	456	142	599
Column Total	3771	2820	8235	4411	19237