

# Lab 05 - Data Summary

WILD 502 - Jay Rotella

## Bring in the Data

```
library(tidyverse)
library(GGally)
library(RMark)
sq <- convert.inp("http://www.montana.edu/rotella/documents/502/SwampSquirrels.inp",
                 group.df = data.frame(age = c("juv", "older")),
                 covariates = c("birth", "tail"))
head(sq, 4)
```

```
##      ch freq age birth tail
## 1:3 100000  1 juv  22.5 13.5
## 1:5 100000  1 juv  30.0 24.0
## 1:7 101000  1 juv  20.5 15.0
## 1:9 100000  1 juv  24.5 24.0
```

```
# select just the covariates that will be used in modeling
sq.covs <- sq %>%
  select(birth, tail, age)
head(sq.covs, 4)
```

```
##      birth tail age
## 1:3  22.5 13.5 juv
## 1:5  30.0 24.0 juv
## 1:7  20.5 15.0 juv
## 1:9  24.5 24.0 juv
```

## Some useful Summary Statistics

Here, we have 2 continuous covariates and 1 categorical.

```
summary(sq.covs)
```

```
##      birth      tail      age
## Min.   :15.00  Min.   :10.00  juv  :393
## 1st Qu.:18.62  1st Qu.:13.50  older:357
## Median :22.00  Median :17.50
## Mean   :22.42  Mean   :17.48
## 3rd Qu.:26.00  3rd Qu.:21.50
## Max.   :30.00  Max.   :25.00
```

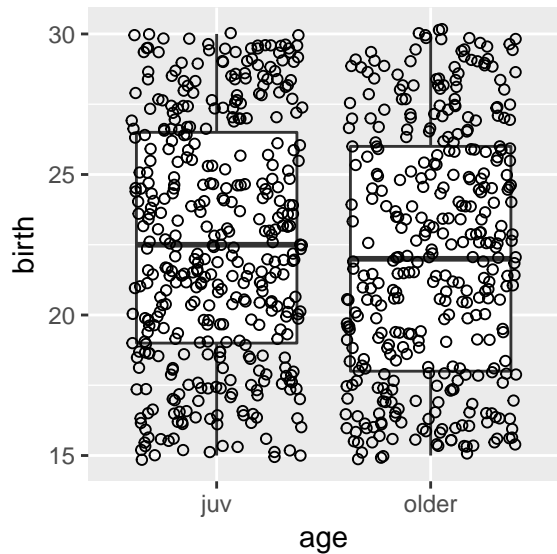
```
# get SDs for continuous covariates
apply(sq.covs[, 1:2], 2, sd)
```

```
##      birth      tail
## 4.333023 4.515280
```

```
ggpairs(sq.covs, aes(color = age, alpha = 0.2))
```



```
ggplot(sq.covs, aes(x = age, y = birth)) +
  geom_boxplot() +
  geom_point(pch = 21, position = position_jitter())
```



```
ggplot(sq.covs, aes(x = age, y = tail)) +
  geom_boxplot() +
  geom_point(pch = 21, position = position_jitter())
```

