# Project Code Musk Oxen

Jack Nowacek

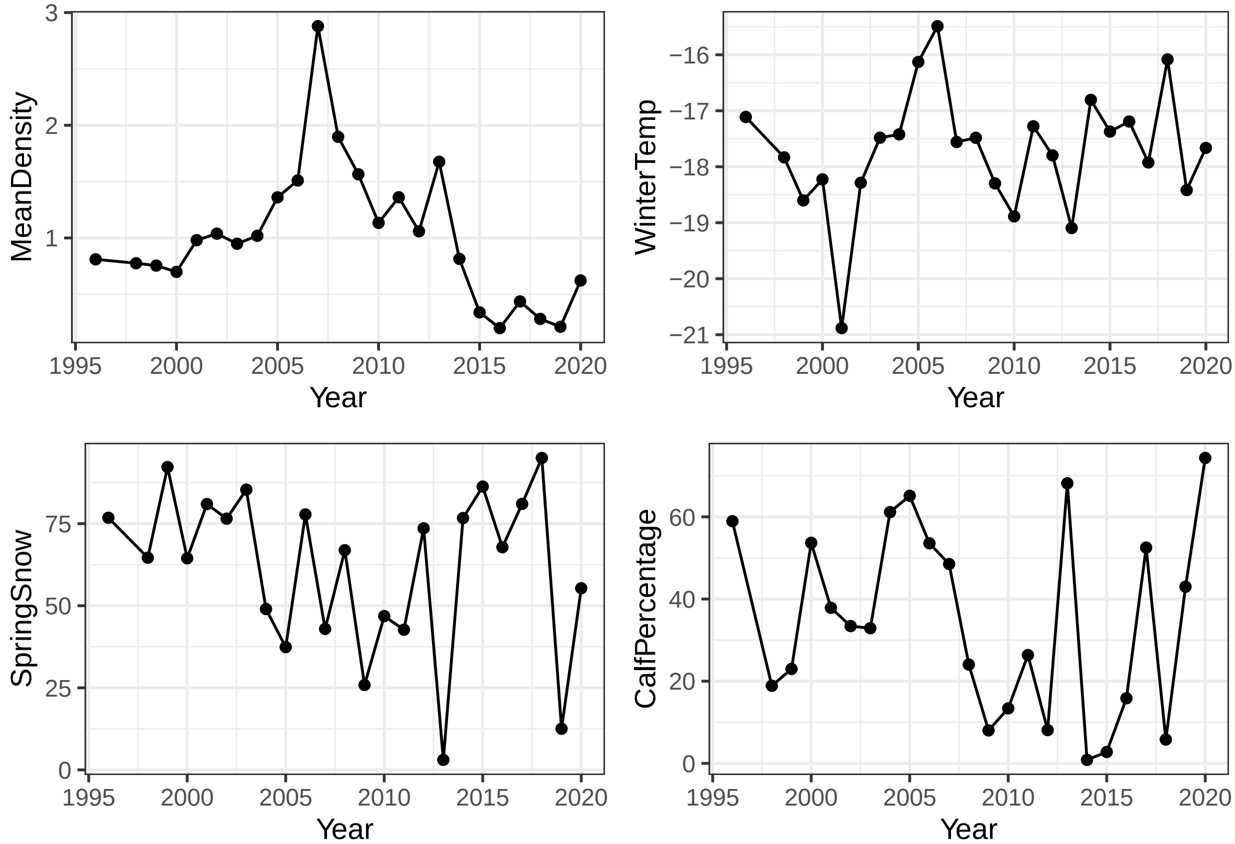
2023-10-19

## Packages

```
data <- read_excel("~/Documents/Copenhagen/Polar_Bio/MuskOxenData.xlsx")
data <- data |>
  select(!Remarks) |>
  na.omit()
```

## EDA

```
p1 <- ggplot(data = data, aes(x = Year, y = MeanDensity)) +
  geom_point() +
  geom_path() +
  geom_abline() +
  theme_bw()

p2 <- ggplot(data = data, aes(x = Year, y = WinterTemp)) +
  geom_point() +
  geom_path() +
  geom_abline() +
  theme_bw()

p3 <- ggplot(data = data, aes(x = Year, y = SpringSnow)) +
  geom_point() +
  geom_path() +
  geom_abline() +
  theme_bw()

p4 <- ggplot(data = data, aes(x = Year, y = CalfPercentage)) +
  geom_point() +
  geom_path() +
  geom_abline() +
  theme_bw()

# p1
# p2
# p3
# p4

grid.arrange(p1, p2, p3, p4, nrow = 2)
```




```
# data |>
#   ggplot(aes(x = Year, group = 1)) +
#   geom_path(aes(y = CalfPercentage)) +
#   geom_point(aes(y = CalfPercentage)) +
#   geom_path(aes(y = MeanDensity)) +
#   geom_point(aes(y = MeanDensity)) +
#   theme_linedraw()
```

Basic trend plots of the four major variables over the range of the data.

## Linear Regressions

```
density_reg <- lm(CalfPercentage ~ MeanDensity, data = data)
density_summ <- summary(density_reg)
density_summ$r.squared
```

```
## [1] 0.0414453
```

```
temp_reg <- lm(CalfPercentage ~ WinterTemp, data = data)
temp_summ <- summary(temp_reg)
temp_summ$r.squared
```

```
## [1] 0.0002001326
```

```r
snow_reg <- lm(CalfPercentage ~ SpringSnow, data = data)
snow_summ <- summary(snow_reg)
snow_summ$r.squared
```

```
## [1] 0.1264935
```

```r
all_reg <- lm(CalfPercentage ~ MeanDensity +
                WinterTemp + SpringSnow, data = data)
all_summ <- summary(all_reg)
all_summ$r.squared
```

```
## [1] 0.1351996
```

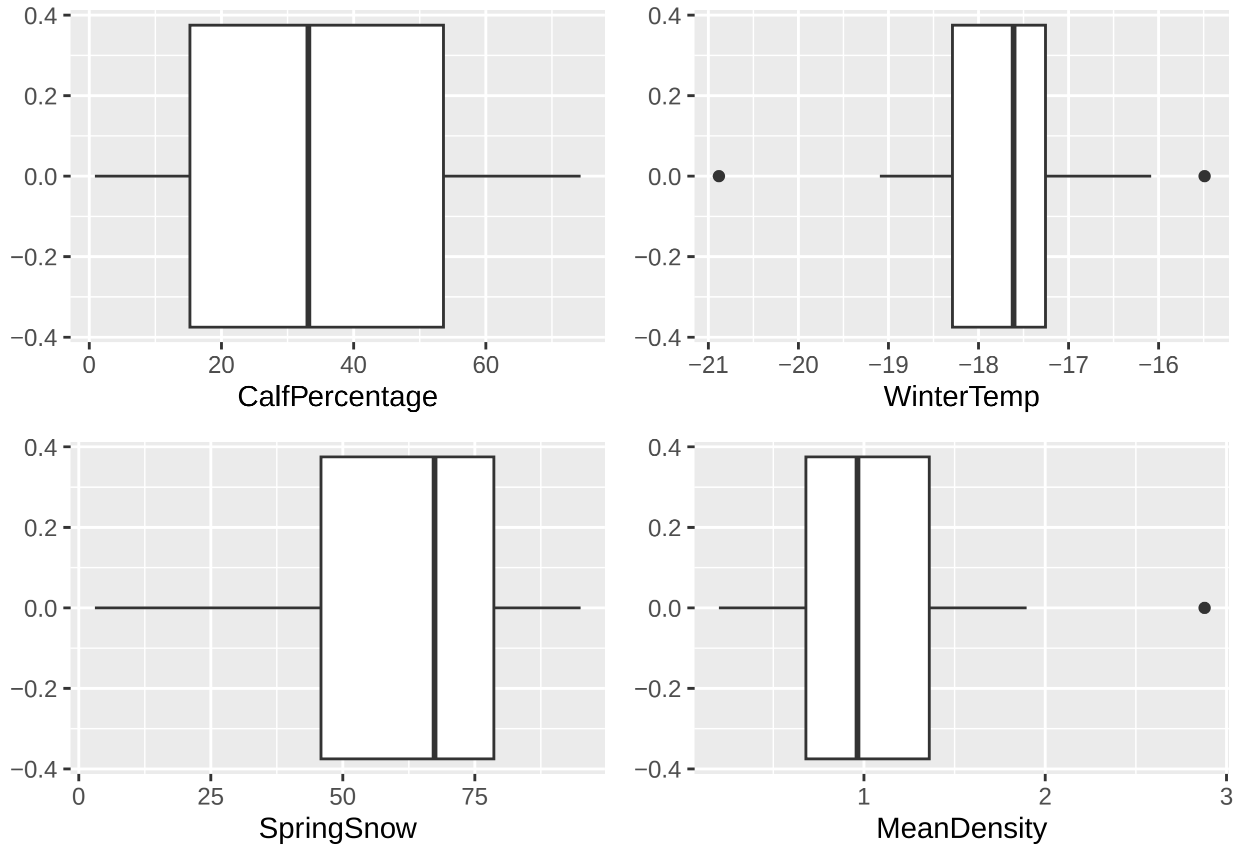
```r
density_to_springsnow <- lm(MeanDensity ~ SpringSnow, data = data)
summary <- summary(density_to_springsnow)
summary$r.squared
```

```
## [1] 0.1404236
```

Standard linear regressions between calf percentage and the three explanatory variables, then a multiple linear regression between calf percentage and all three variables. Finally a regression where spring snow predicts mean density.

## Outlier Investigation

```r
g1 <- data |>
  ggplot(aes(CalfPercentage)) +
  geom_boxplot()

g2 <- data |>
  ggplot(aes(WinterTemp)) +
  geom_boxplot()

g3 <- data |>
  ggplot(aes(SpringSnow)) +
  geom_boxplot()

g4 <- data |>
  ggplot(aes(MeanDensity)) +
  geom_boxplot()

grid.arrange(g1, g2, g3, g4, nrow = 2)
```




### Linear Regressions No Outliers

```
# no_outliers_data <- data |>
#   subset(id = c(5, 10, 11))

data_no_out <- data[-c(5, 10, 11, 13, 14, 15, 16),]

# Removes data from years with outliers in the data and those where a disease
# outbreak affected the population.

density_reg2 <- lm(CalfPercentage ~ MeanDensity, data = data_no_out)
density_summ2 <- summary(density_reg2)
density_summ2$r.squared
```

```
## [1] 0.1034047
```

```
temp_reg2 <- lm(CalfPercentage ~ WinterTemp, data = data_no_out)
temp_summ2 <- summary(temp_reg2)
temp_summ2$r.squared
```

```
## [1] 0.07583693
```

```r
snow_reg2 <- lm(CalfPercentage ~ SpringSnow, data = data_no_out)
snow_summ2 <- summary(snow_reg2)
snow_summ2$r.squared
```

```
## [1] 0.3451069
```

```r
snow_reg2 <- lm(SpringSnow ~ MeanDensity, data = data_no_out)
snow_summ2 <- summary(snow_reg2)
snow_summ2$r.squared
```

```
## [1] 0.1072841
```

```r
all_reg2 <- lm(CalfPercentage ~ MeanDensity +
                 WinterTemp + SpringSnow, data = data_no_out)
all_summ2 <- summary(all_reg2)
all_summ2$r.squared
```

```
## [1] 0.3680844
```

## Winter Temp to Spring Snow Idea

```r
snowy_data <- data |>
  select(!c("Year", "MeanDensity", "CalfPercentage")) |>
  mutate(SpringSnow=lead(SpringSnow)) |>
  slice(1:n()-1)

snowy_reg <- lm(SpringSnow ~ WinterTemp, data = snowy_data)
snowy_summ <- summary(snowy_reg)
snowy_summ$r.squared
```

```
## [1] 0.04887347
```

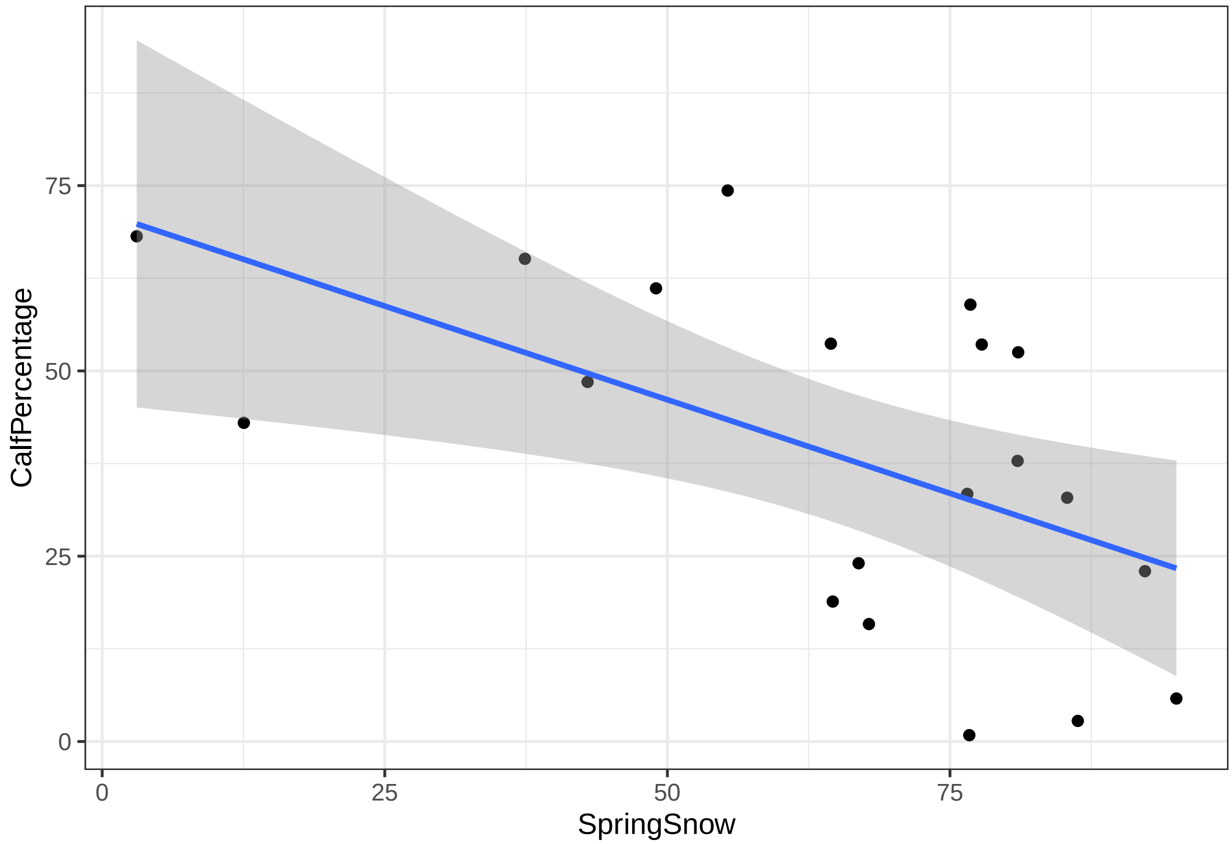
Failed, no correlation between previous winter temperature and the snow cover of the next year.

Calf percentage (y) vs spring snow cover (x) match x with y based on year, fit between the two variables. Use this type of plot to explore relationship between the other variables

## Plots of Major Relationships

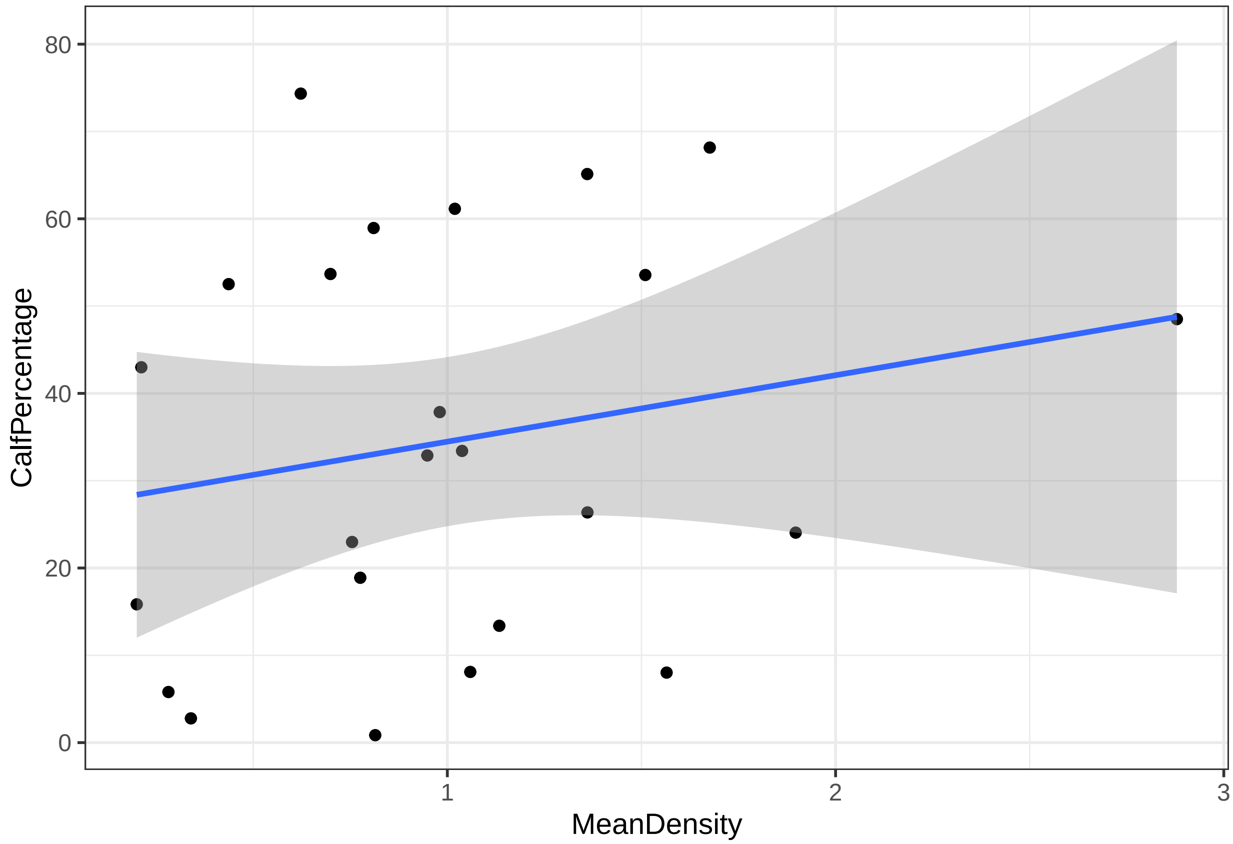
```r
data |>
  slice(1:12, 17:n()) |>
  ggplot(aes(SpringSnow, CalfPercentage)) +
  geom_point() +
  geom_smooth(method = lm) +
  theme_bw()
```

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



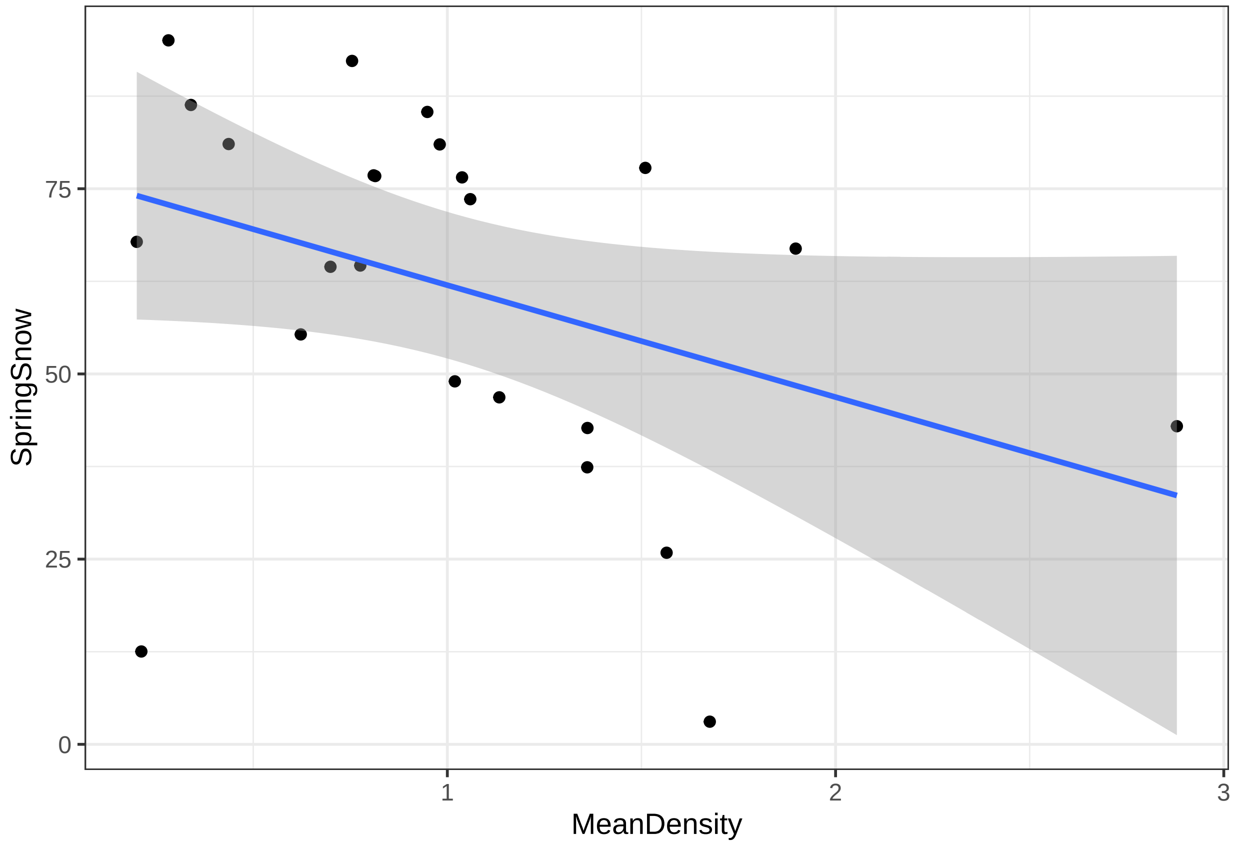

```r
data |>
  ggplot(aes(MeanDensity, CalfPercentage)) +
  geom_point() +
  geom_smooth(method = lm) +
  theme_bw()
```

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




```
data |>
  ggplot(aes(MeanDensity, SpringSnow)) +
  geom_point() +
  geom_smooth(method = lm) +
  theme_bw()
```

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




Is calf production density-dependent?

Contrary to logical consideration, we found that the density of animals and the calf percentage had an extremely weak positive correlation. We expected that, with increased density of adults, would come increased interactions and therefore increased calf percentage. With an $R^2$ value of 0.1034047, there appears to be no significant numerical link between these two variables.

include graph

Are winter conditions (snow and temperature) affecting muskox reproduction or density?

Yes they do affect reproduction, the strongest correlation in the dataset was between snow cover and calf percentage. While the $R^2$ value of 0.3451069 is not very strong by typical standards, it is around three times stronger than the other explanatory variables. As for density, we did not find a strong correlation between the two variables, the $R^2$ value was 0.1072841.

If so, what could be the mechanism?

Theories: Snow cover affects grazing, affects the female's ability to provide nutrients to her fetus, higher chance of being re-absorbed instead of carried to term. Spring snow is predicted by the temp of the previous winter? Is that what is in the paper?

Which other factors could contribute to the observed patterns in density and reproduction?

Anything, need evidence and ideas from the papers.

EOF