Q1 SLR: beaver dams

2024-12-06

Question 1

In recent decades, beavers have expanded beyond their typical habitats and begun colonizing the Arctic tundra. Jones et al. collected data on surface water area and surveyed beaver dams to investigate how beaver activity influences hydrology in this region. Beaver dams create localized flooding around water bodies, potentially impacting the Arctic's hydrology and carbon cycling.

Using data from the Baldwin Peninsula in Alaska, Jones and colleagues analyzed surface water area and the number of beaver dams over nearly two decades. In this question, evaluate whether the presence of beaver dams is associated with an increase in surface water in the tundra.

Perform a linear regression to assess whether there is a significant relationship between the number of beaver dams and surface water area. Write a concise paragraph summarizing your results and conclusions, ensuring you clearly interpret the regression coefficients and significance.

```
data <- read.csv('Q1-SLR-beaver-dams.csv')
head(data)
```

```
##   year dams.n area.ha
## 1 2002      2     594
## 2 2007      6     610
## 3 2008      7     623
## 4 2009     10     600
## 5 2010     15     618
## 6 2011     31     625
```

Set up regression

```
mod <- lm(area.ha ~ dams.n, data)
summary(mod)
```

```
##
## Call:
## lm(formula = area.ha ~ dams.n, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -12.639  -7.191  -1.006   6.954  14.772
##
## Coefficients:
##             Estimate Std. Error t value             Pr(>|t|)
## (Intercept) 606.00410    4.13835 146.436 < 0.0000000000000002 ***
## dams.n        0.31769    0.08037   3.953              0.00272 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 9.251 on 10 degrees of freedom
## Multiple R-squared:  0.6097, Adjusted R-squared:  0.5707
## F-statistic: 15.62 on 1 and 10 DF,  p-value: 0.002718
```

Get standardized residuals

```
dam.res <- rstandard(mod)
shapiro.test(dam.res)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  dam.res
## W = 0.92993, p-value = 0.3793
```

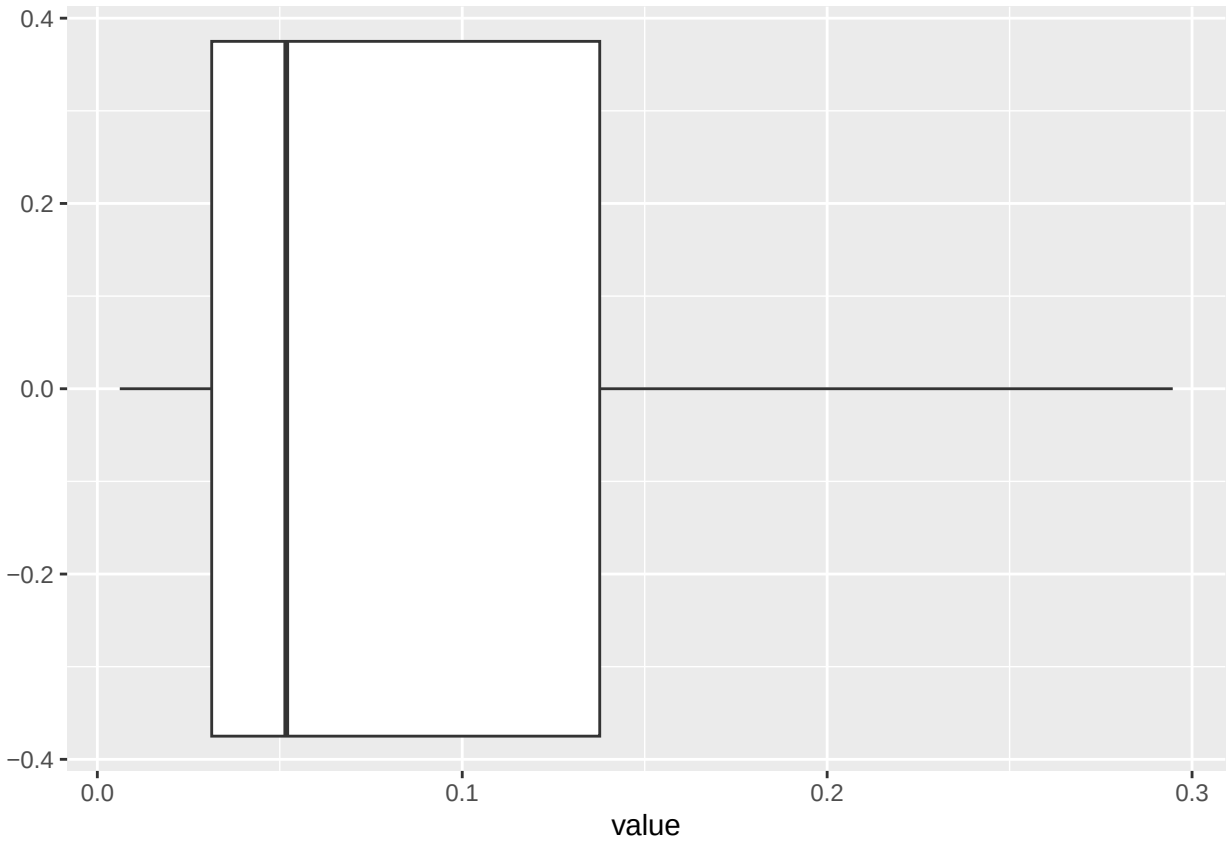
The Breusch-Pagan test to assess homoscedasticity

```
ncvTest(mod)
```

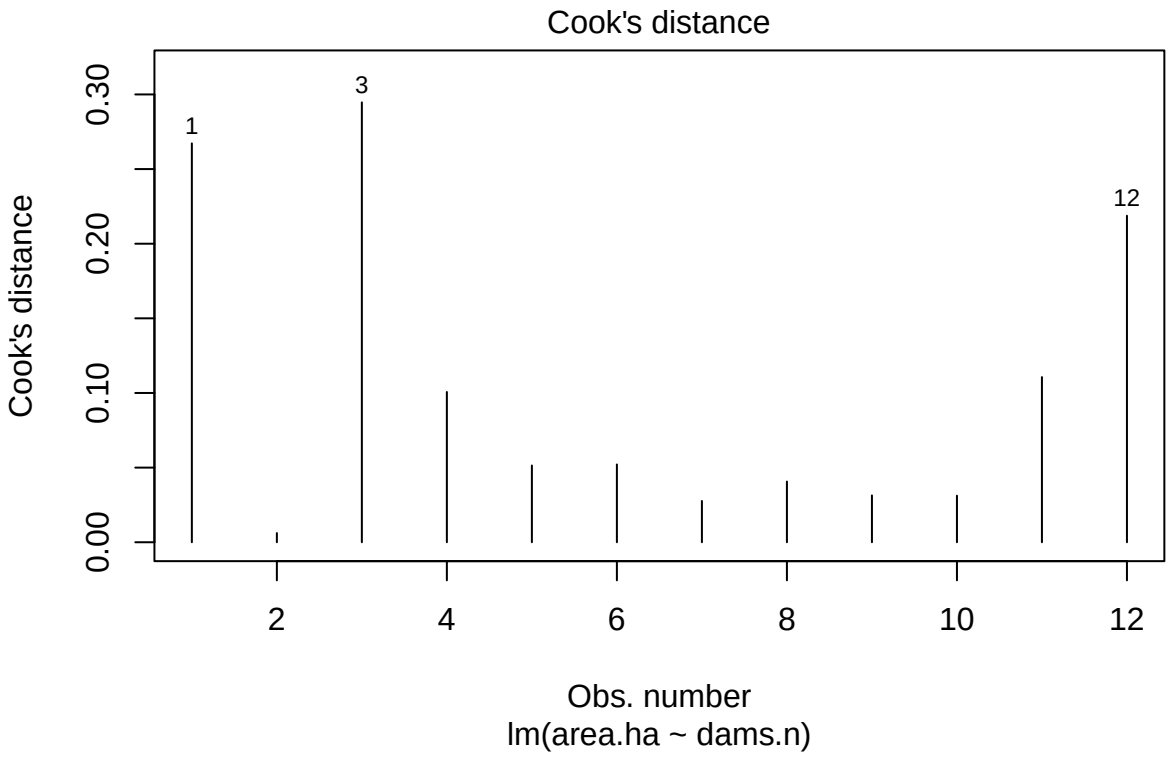
```
## Non-constant Variance Score Test
## Variance formula: ~ fitted.values
## Chisquare = 1.054123, Df = 1, p = 0.30456
```

#Cooks D

```
ggplot(as_tibble(cooks.distance(mod)), aes(value)) + geom_boxplot()
```

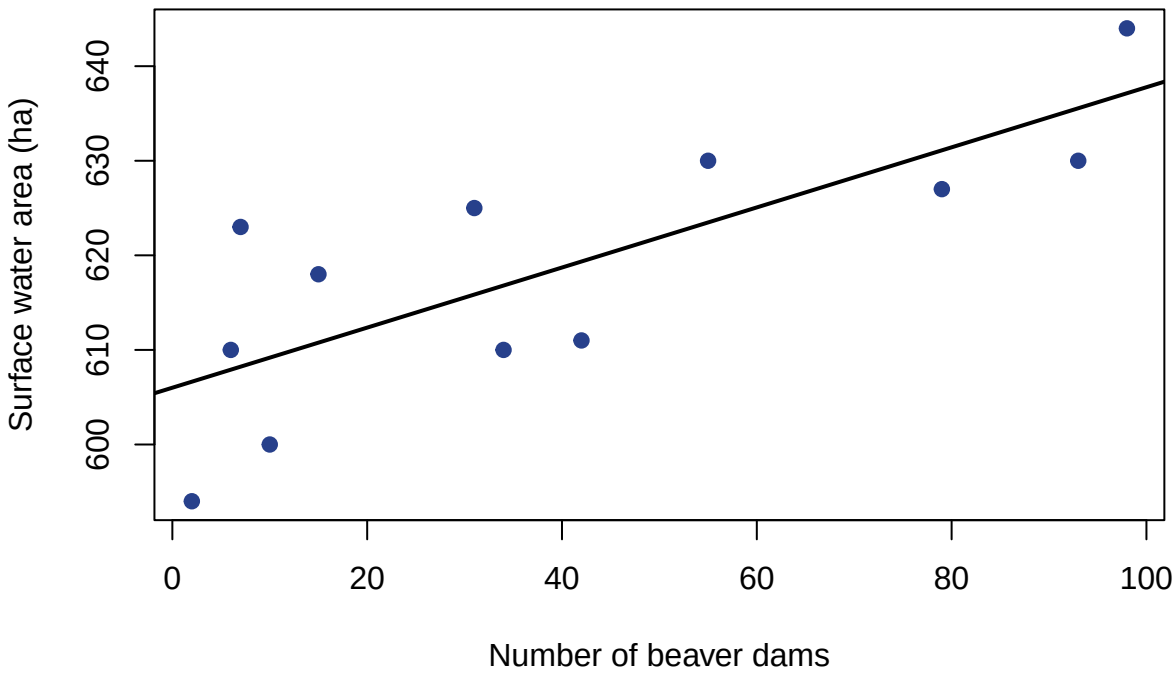



```
plot(mod,which=4)
```



Make plot of beaver dams and surface water

```
plot(area.ha ~ dams.n, data,
     pch = 19,
     col = "royalblue4",
     ylab = "Surface water area (ha)",
     xlab =  "Number of beaver dams")
#add regression line
#make line width thicker
abline(mod, lwd=2)
```



References:

https://bookdown.org/kroppheather/IntEnvData/introduction-to-linear-regression.html https://github.com/kroppheather/Intro_Enviro_Data/blob/main/IntroEnvData/data/tutorial_5/beaver_dam.csv