2025 Redcedar Adaptation Analysis

2025-03-24

knitr::opts\_chunk$set(echo = TRUE)

## Introduction

Welcome to the analyses of western redcedar trees planted in for the Open Redcedar Adaptation Network

### Research Topic

Dieback of western redcedar is increasinly reported throughout the northwest. Planting trees from southern seed sources may be a solution. The open redcedar adaptation network planted redcedar trees from two seed sources to compare the growth of trees from Oregon seed zones to trees from Washington seed zones.

### Data

Redcedar trees were planted in 2022 and remeasured each spring thereafter. The dataset analyzed here includes two years of remeasurements after planting.

### Research Question

Is there a difference in growth rates between Oregon seed zones and Washington Seed zones.

## Data Wrangling

data <- read\_excel("data/WRC Tree Measurements 2024.xlsx", sheet = "2024 Corrected")

names(data)

## [1] "New Tree Number" "Tree Number" "Site"   
## [4] "Seed Zone" "Date Measured" "Tree Height (cm)"   
## [7] "Tree Diameter (mm)" "Health" "Comment"   
## [10] "Year Measured" "Elk Browse"

str(data)

## tibble [519 × 11] (S3: tbl\_df/tbl/data.frame)  
## $ New Tree Number : num [1:519] NA NA NA NA NA NA NA NA NA NA ...  
## $ Tree Number : num [1:519] 227 228 231 232 234 235 236 237 238 239 ...  
## $ Site : chr [1:519] "Puyallup" "Puyallup" "Puyallup" "Puyallup" ...  
## $ Seed Zone : chr [1:519] "OR" "OR" "OR" "OR" ...  
## $ Date Measured : POSIXct[1:519], format: NA NA ...  
## $ Tree Height (cm) : num [1:519] NA NA NA NA NA NA NA NA NA NA ...  
## $ Tree Diameter (mm): num [1:519] NA NA NA NA NA NA NA NA NA NA ...  
## $ Health : chr [1:519] NA NA NA NA ...  
## $ Comment : chr [1:519] NA NA NA NA ...  
## $ Year Measured : num [1:519] 2024 2024 2024 2024 2024 ...  
## $ Elk Browse : chr [1:519] NA NA NA NA ...

Below we set the site and seed zone variables as factors

data$Site <- as.factor(data$Site)  
data$`Seed Zone` <- as.factor(data$`Seed Zone`)  
#data$`Elk Browse` <- as.factor(data$`Elk Browse`)

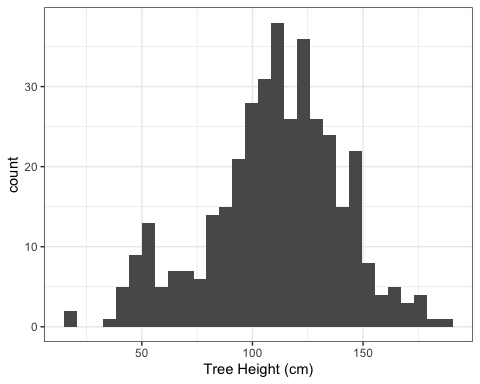
## Data Visualization

Lets look at the distribution of the data

ggplot(data, aes(`Tree Height (cm)`) ) +geom\_histogram() +theme\_bw()

## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.

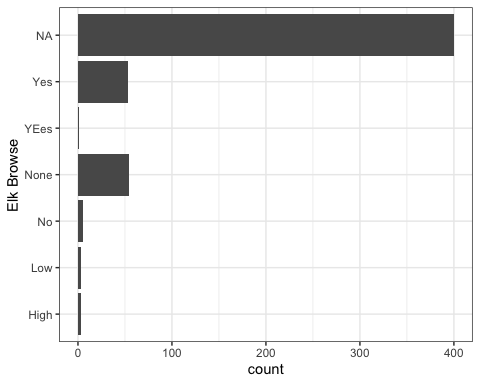
## Warning: Removed 142 rows containing non-finite outside the scale range  
## (`stat\_bin()`).



It looks like there are a lot of trees in the 50cm range. Perhaps that is related to elk browse?

Let try to plot it without any of the trees that had elk browse

ggplot(data,aes(`Elk Browse`))+geom\_bar() +coord\_flip()+theme\_bw()



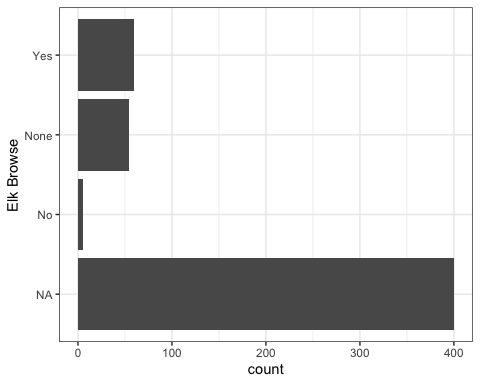
data.browsed <- data %>% mutate(`Elk Browse` = recode(`Elk Browse`, YEes = 'Yes', Low = 'Yes', High = 'Yes' ))

str(data.browsed)

## tibble [519 × 11] (S3: tbl\_df/tbl/data.frame)  
## $ New Tree Number : num [1:519] NA NA NA NA NA NA NA NA NA NA ...  
## $ Tree Number : num [1:519] 227 228 231 232 234 235 236 237 238 239 ...  
## $ Site : Factor w/ 5 levels "Puyallup","Renton",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Seed Zone : Factor w/ 3 levels "OR","Wa","WA": 1 1 1 1 1 1 1 1 1 1 ...  
## $ Date Measured : POSIXct[1:519], format: NA NA ...  
## $ Tree Height (cm) : num [1:519] NA NA NA NA NA NA NA NA NA NA ...  
## $ Tree Diameter (mm): num [1:519] NA NA NA NA NA NA NA NA NA NA ...  
## $ Health : chr [1:519] NA NA NA NA ...  
## $ Comment : chr [1:519] NA NA NA NA ...  
## $ Year Measured : num [1:519] 2024 2024 2024 2024 2024 ...  
## $ Elk Browse : chr [1:519] NA NA NA NA ...

data.browsed$`Elk Browse`[is.na(data.browsed$`Elk Browse`)] <- "NA"

ggplot(data.browsed,aes(`Elk Browse`))+geom\_bar() +coord\_flip()+theme\_bw()

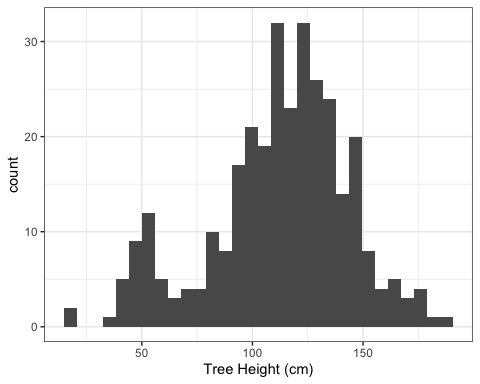


data.browsed.filtered <- data.browsed %>% filter(`Elk Browse` != "Yes")

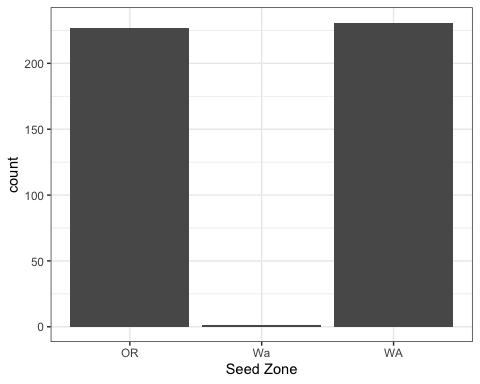
ggplot(data.browsed.filtered, aes(`Tree Height (cm)`) ) +geom\_histogram() +theme\_bw()

## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.

## Warning: Removed 142 rows containing non-finite outside the scale range  
## (`stat\_bin()`).



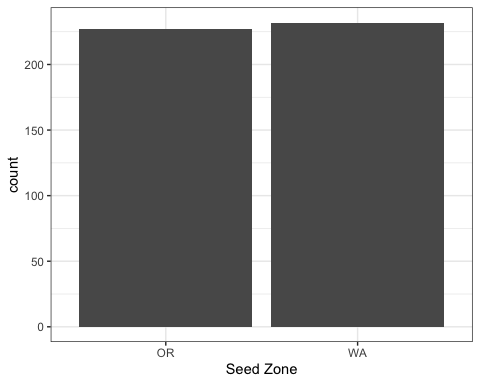
ggplot(data.browsed.filtered,aes(`Seed Zone`)) +geom\_bar() +theme\_bw()



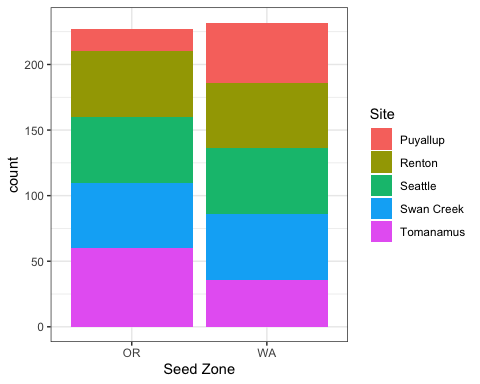
Oh, we have one tree in seed zone labeled ‘Wa’ instead of ‘WA’

data.browsed.filtered$`Seed Zone`[data.browsed.filtered$`Seed Zone`=="Wa"] <- "WA"

ggplot(data.browsed.filtered,aes(`Seed Zone`)) +geom\_bar() +theme\_bw()



ggplot(data.browsed.filtered,aes(`Seed Zone`,fill=Site)) +geom\_bar() +theme\_bw()



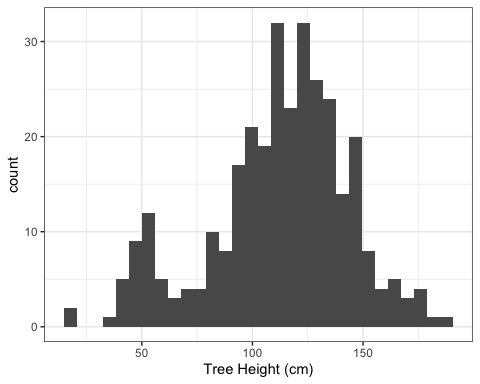
Lets filter out the Puyallup site

data.browsed.site.filtered <- data.browsed.filtered %>% filter(Site!="Puyallup")

ggplot(data.browsed.site.filtered, aes(`Tree Height (cm)`) ) +geom\_histogram() +theme\_bw()

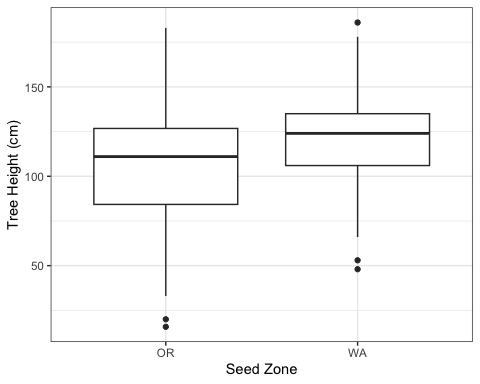
## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.

## Warning: Removed 79 rows containing non-finite outside the scale range  
## (`stat\_bin()`).



ggplot(data.browsed.filtered,aes(`Seed Zone`,`Tree Height (cm)`)) +geom\_boxplot() +theme\_bw()

## Warning: Removed 142 rows containing non-finite outside the scale range  
## (`stat\_boxplot()`).



#ggplot(df, aes(x=Category, y=Mean, fill=Quality)) + geom\_bar(position=position\_dodge(), stat="identity", colour='black') + geom\_errorbar(aes(ymin=Mean-sd, ymax=Mean+sd), width=.2)

## Analyses

Lets compare the standard deviation between groups

library(car)

## Loading required package: carData

##   
## Attaching package: 'car'

## The following object is masked from 'package:dplyr':  
##   
## recode

## The following object is masked from 'package:purrr':  
##   
## some

leveneTest(`Tree Height (cm)` ~ `Seed Zone`, data=data.browsed.site.filtered)

## Levene's Test for Homogeneity of Variance (center = median)  
## Df F value Pr(>F)   
## group 1 16.015 7.84e-05 \*\*\*  
## 315   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

According to the Levene’s Test for Homogenity of Variance, the standard deviation in tree height between Oregon and Washington seed zones is significantly different. reject the null -

“tests the null hypothesis that the population variances are equal (called homogeneity of variance or homoscedasticity). If the resulting p-value of Levene’s test is less than some significance level (typically 0.05), the obtained differences in sample variances are unlikely to have occurred based on random sampling from a population with equal variances. Thus, the null hypothesis of equal variances is rejected and it is concluded that there is a difference between the variances in the population.” - Wikipedia article <https://en.wikipedia.org/wiki/Levene%27s_test>

Basically, there is a difference in variation. Therefore, an anova can’t be used to compare the tree heights in the data.

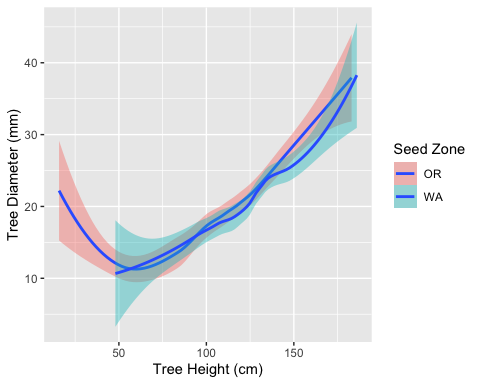
Determining the tests that are appropriate

Depending on the distribution and shape of the data, we may or may not be able to compare groups with anova or t-tests.

ggplot(data.browsed.filtered,aes(`Tree Height (cm)`,`Tree Diameter (mm)`, fill=`Seed Zone`)) +geom\_smooth()

## `geom\_smooth()` using method = 'loess' and formula = 'y ~ x'

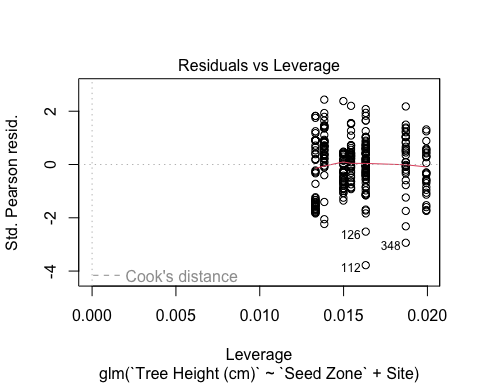
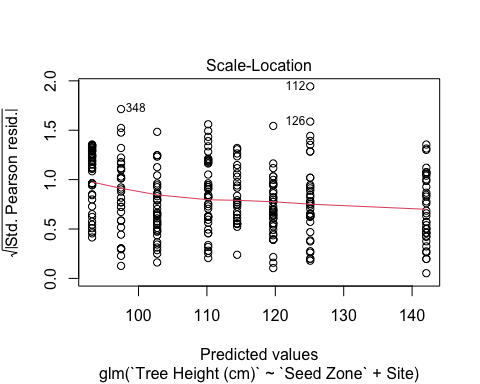
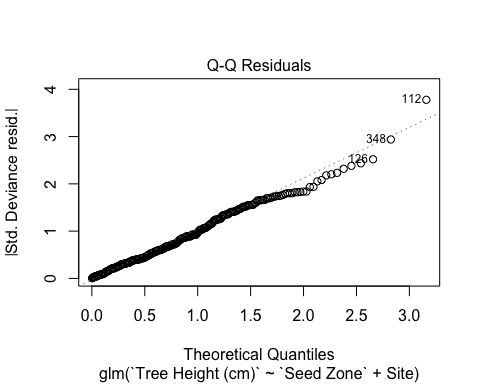
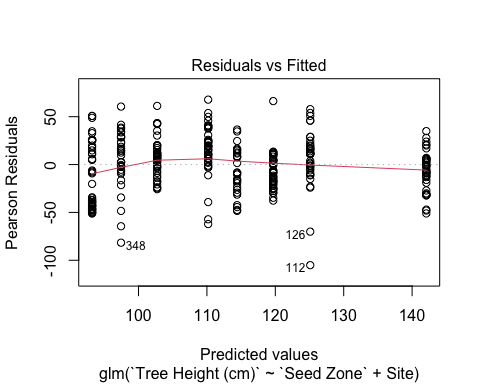
## Warning: Removed 144 rows containing non-finite outside the scale range  
## (`stat\_smooth()`).



model.one <- glm(`Tree Height (cm)` ~ `Seed Zone` + Site, data=data.browsed.site.filtered)  
  
summary(model.one)

##   
## Call:  
## glm(formula = `Tree Height (cm)` ~ `Seed Zone` + Site, data = data.browsed.site.filtered)  
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 93.212 3.237 28.791 < 2e-16 \*\*\*  
## `Seed Zone`WA 16.971 3.155 5.379 1.47e-07 \*\*\*  
## SiteSeattle 31.895 4.307 7.405 1.23e-12 \*\*\*  
## SiteSwan Creek 9.514 4.206 2.262 0.0244 \*   
## SiteTomanamus 4.234 4.571 0.926 0.3551   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for gaussian family taken to be 786.7835)  
##   
## Null deviance: 317466 on 316 degrees of freedom  
## Residual deviance: 245476 on 312 degrees of freedom  
## (79 observations deleted due to missingness)  
## AIC: 3020.3  
##   
## Number of Fisher Scoring iterations: 2

plot(model.one)



## Results