Greenup Analyses

Kara Dobson

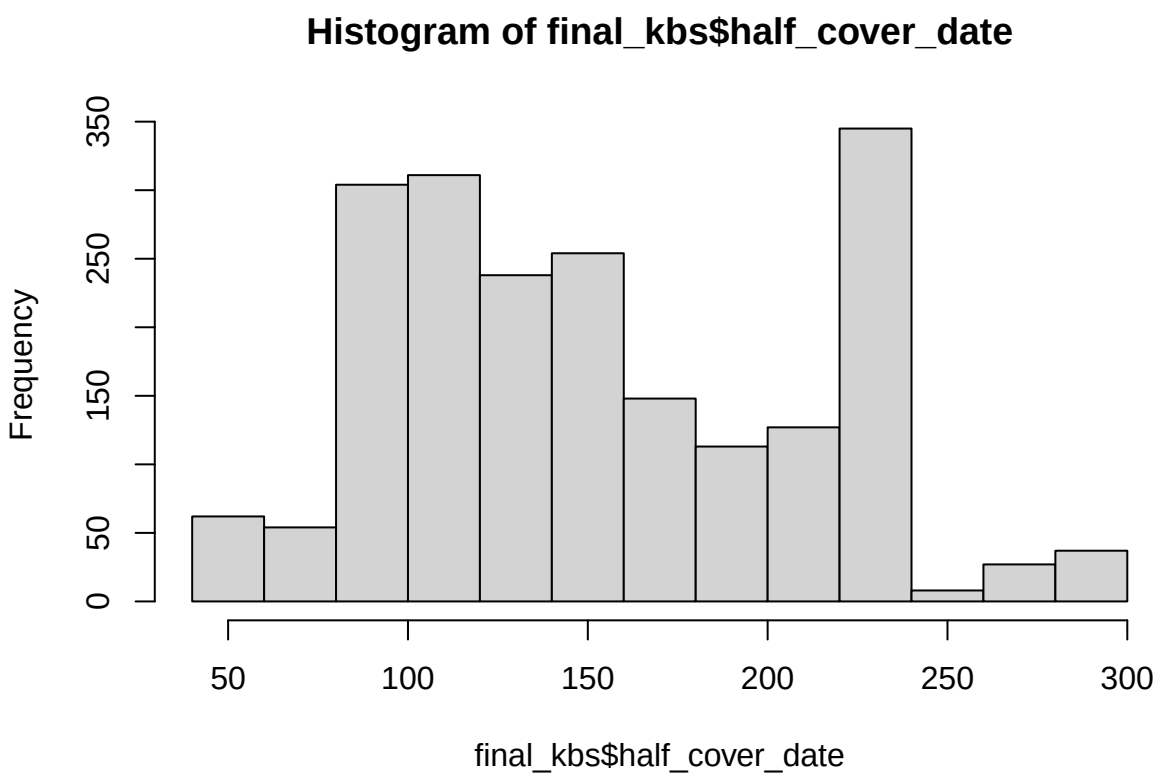
February 18, 2021

COLLABORATORS: Phoebe Zarnetske, Moriah Young, Mark Hammond, Pat Bills
DATA INPUT: Data imported as csv files from shared Google drive L1 plant comp folder
PROJECT: warmXtrophic

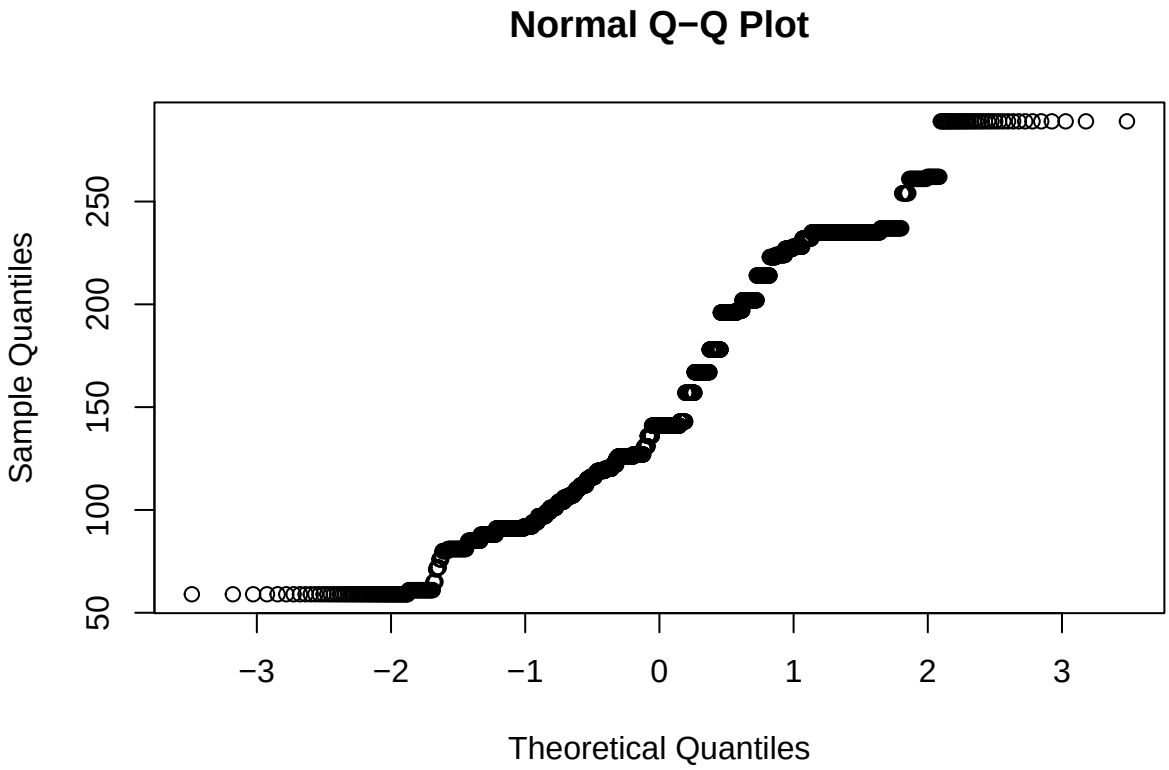
Starting with KBS

First, checking for normality

```
hist(final_kbs$half_cover_date)
```



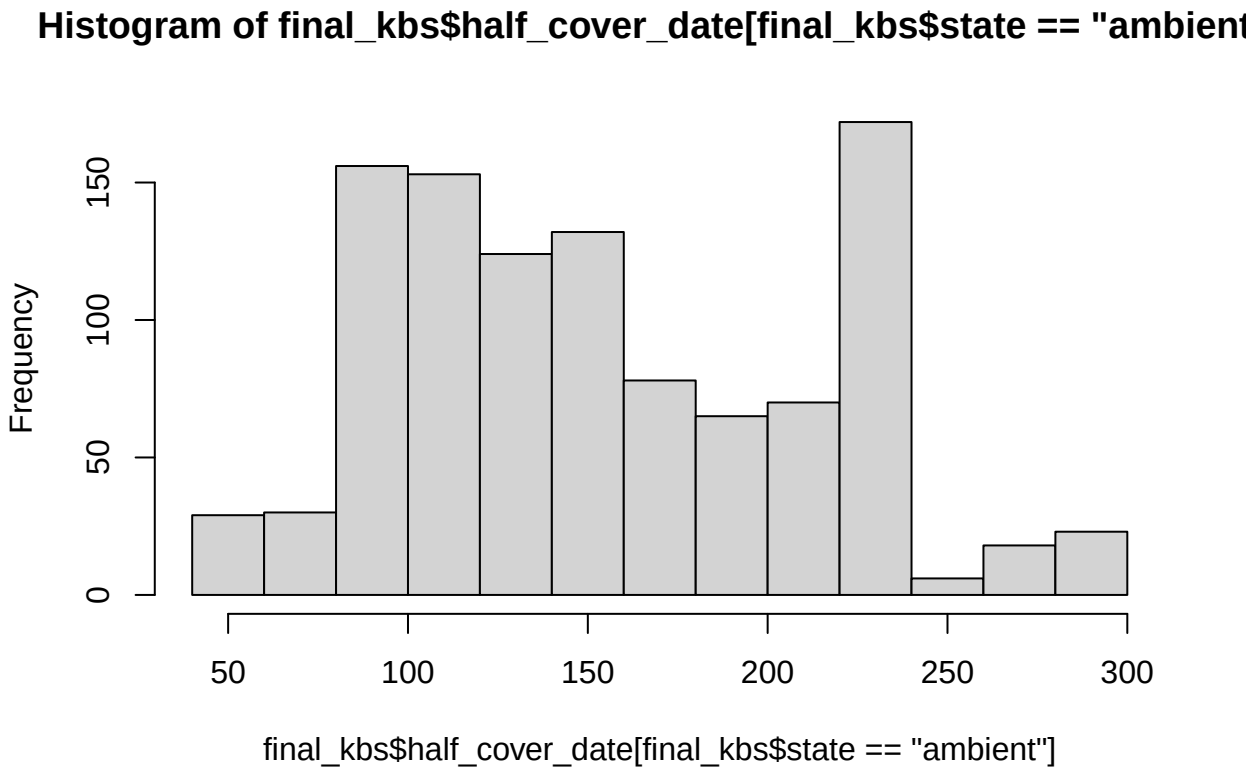

```
qqnorm(final_kbs$half_cover_date)
```



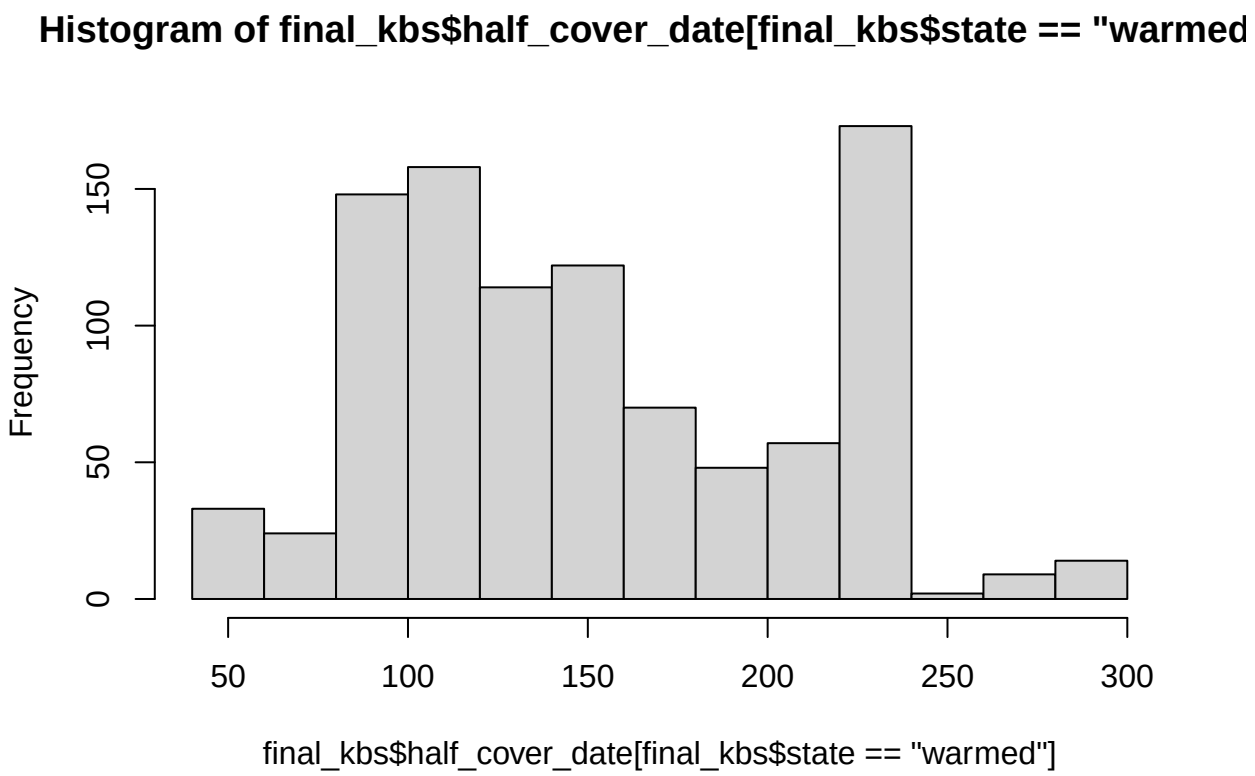

```
shapiro.test(final_kbs$half_cover_date)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  final_kbs$half_cover_date
## W = 0.94273, p-value < 2.2e-16
```

```
hist(final_kbs$half_cover_date[final_kbs$state == "ambient"])
```

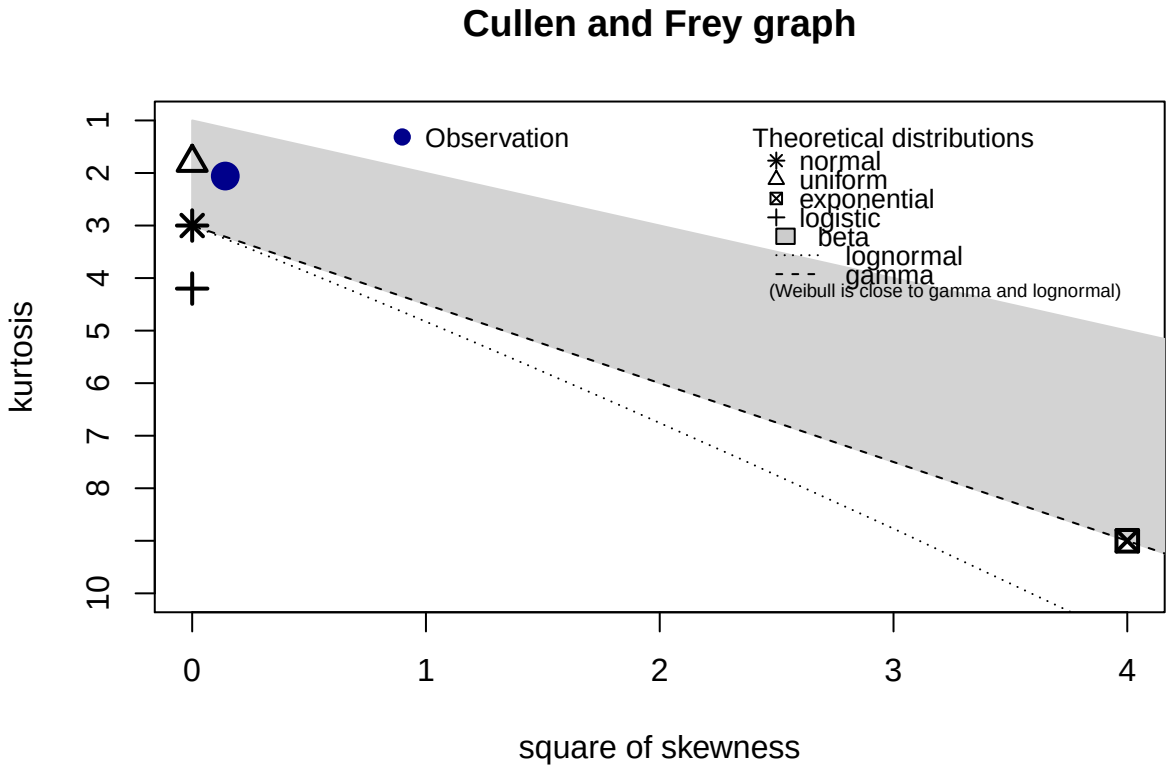



```
hist(final_kbs$half_cover_date[final_kbs$state == "warmed"])
```



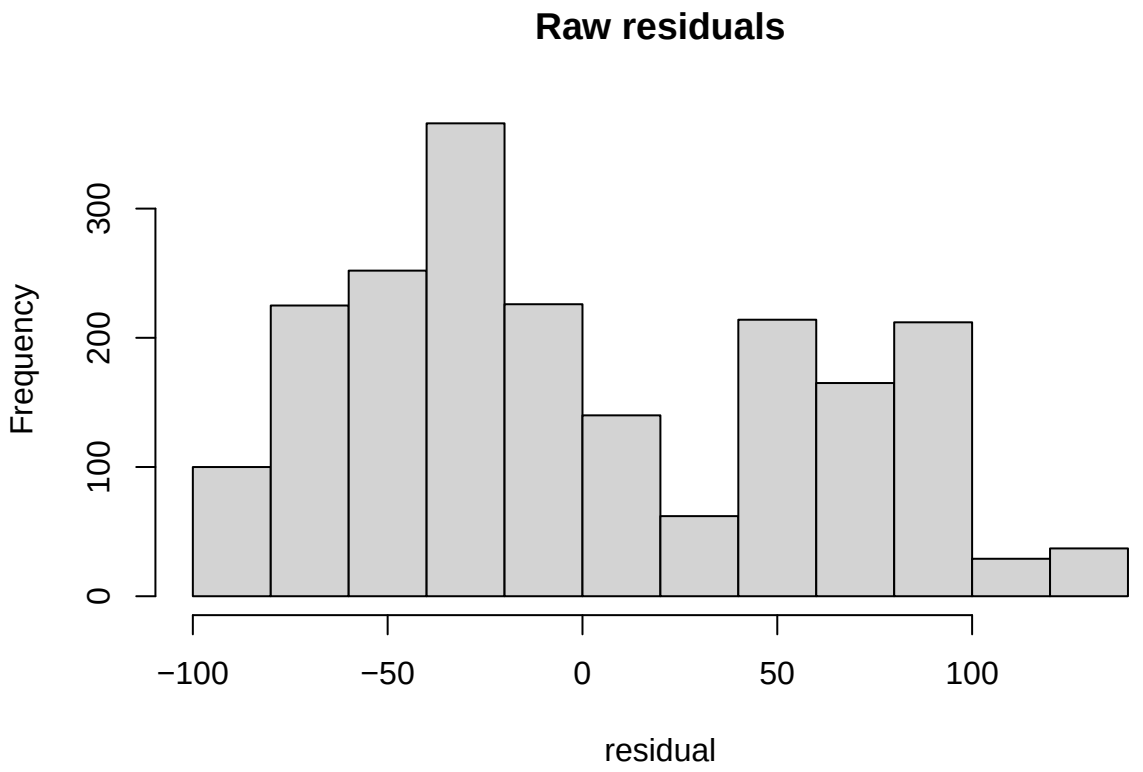
Not normal, and previously attempted transformations don't help

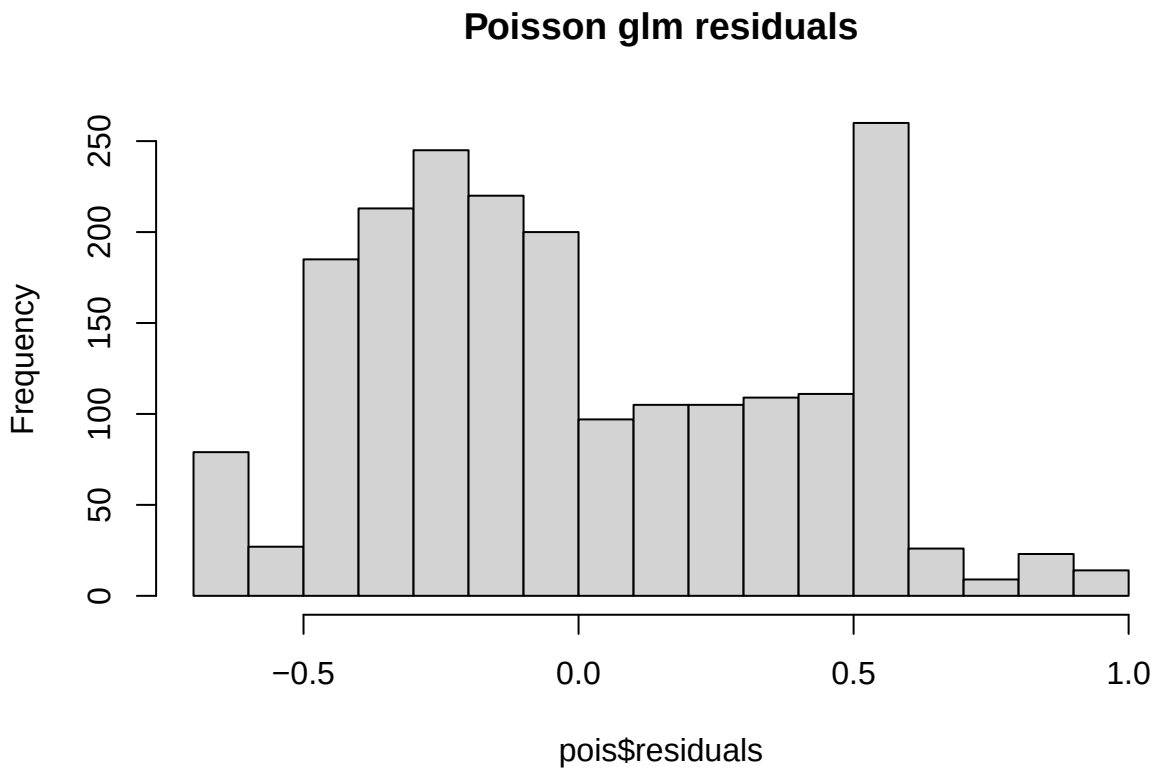
Seeing what other distribution could fit




```
## summary statistics
## ------
## min:  59   max:  289
## median:  141
## mean:  152.8644
## estimated sd:  57.73216
## estimated skewness:  0.3763172
## estimated kurtosis:  2.058193
```

While uniform looks the closest, I'll try poisson





Below I try a few different generalized linear models with poisson distribution:

An interaction between state and year, plus insecticide as a fixed effect and species and plot as random effects

```
moda <- glmer(half_cover_date ~ state*year + insecticide + (1|species) + (1|plot),
              data=final_kbs, family = poisson)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly unider
##  - Rescale variables?
```

```
summary(moda)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: poisson  ( log )
## Formula: half_cover_date ~ state * year + insecticide + (1 | species) +
##     (1 | plot)
##    Data: final_kbs
##
##      AIC      BIC   logLik deviance df.resid
##  42924.6  42963.9 -21455.3  42910.6     2021
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -9.0161 -2.8352 -0.6425  2.3763 14.4656
##
## Random effects:
##  Groups  Name        Variance Std.Dev.
##  species (Intercept) 0.069954 0.26449
##  plot    (Intercept) 0.001848 0.04299
## Number of obs: 2028, groups:  species, 55; plot, 24
##
## Fixed effects:
##                         Estimate Std. Error z value Pr(>|z|)
## (Intercept)             5.020670   0.039631 126.684  < 2e-16 ***
## statewarmed             0.009389   0.019347   0.485   0.6275
## year                    0.003600   0.001520   2.368   0.0179 *
## insecticideno_insects   0.028934   0.017937   1.613   0.1067
## statewarmed:year       -0.009698   0.002116  -4.584 4.57e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) sttwrm year   insct_
## statewarmed -0.243
## year        -0.130  0.249
## insctcdn_ns -0.227 -0.002  0.008
## statwrmd:yr  0.089 -0.375 -0.648  0.006
## convergence code: 0
## Model is nearly unidentifiable: very large eigenvalue
##  - Rescale variables?
```

No interaction between state and year, but with state and insecticide as fixed effects and species and plot as random effects

```
modb <- glmer(half_cover_date ~ state + year + insecticide + (1|species) + (1|plot),
              data=final_kbs, family = poisson)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly unide
##  - Rescale variables?
```

```
summary(modb)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: poisson  ( log )
## Formula: half_cover_date ~ state + year + insecticide + (1 | species) +
##     (1 | plot)
##    Data: final_kbs
##
##      AIC      BIC   logLik deviance df.resid
##  42943.6  42977.3 -21465.8  42931.6     2022
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -9.0053 -2.8448 -0.6575  2.3929 14.2130
##
## Random effects:
##  Groups  Name        Variance Std.Dev.
##  species (Intercept) 0.069954 0.26449
##  plot    (Intercept) 0.001847 0.04297
## Number of obs: 2028, groups:  species, 55; plot, 24
##
## Fixed effects:
##                          Estimate Std. Error z value Pr(>|z|)
## (Intercept)            5.0367372  0.0394723 127.602   <2e-16 ***
## statewarmed           -0.0238775  0.0179292  -1.332    0.183
## year                  -0.0009128  0.0011579  -0.788    0.430
## insecticideno_insects  0.0294351  0.0179303   1.642    0.101
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) sttwrm year
## statewarmed -0.227
## year        -0.096  0.007
## insctcdn_ns -0.228  0.001  0.015
## convergence code: 0
## Model is nearly unidentifiable: very large eigenvalue
##  - Rescale variables?
```

State and insecticide as fixed effects & year, species and plot as random effects

```
modc <- glmer(half_cover_date ~ state + insecticide + (1|year) + (1|species) + (1|plot),
              data=final_kbs, family = poisson)
summary(modc)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: poisson  ( log )
## Formula: half_cover_date ~ state + insecticide + (1 | year) + (1 | species) +
##     (1 | plot)
##    Data: final_kbs
##
##      AIC      BIC   logLik deviance df.resid
##  41436.8  41470.5 -20712.4  41424.8     2022
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -9.3418 -2.8134 -0.5215  2.1783 13.1976
##
## Random effects:
##  Groups  Name        Variance Std.Dev.
##  species (Intercept) 0.068407 0.26155
##  plot    (Intercept) 0.001820 0.04266
##  year    (Intercept) 0.005866 0.07659
## Number of obs: 2028, groups:  species, 55; plot, 24; year, 6
##
## Fixed effects:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)            5.01945    0.04991 100.579   <2e-16 ***
## statewarmed           -0.02683    0.01781  -1.507   0.1319
## insecticideno_insects  0.03093    0.01781   1.737   0.0824 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) sttwrm
## statewarmed -0.178
## insctcdn_ns -0.178  0.001
```

Because no distributions seems to match well, I'll try a Friedman's test

```
#friedman_kbs <- final_kbs %>%
#  friedman_test(half_cover_date ~ state)
```

I get this - Error: Must extract column with a single valid subscript. x Subscript `var` can't be `NA`

Can't figure out what this means

If I include the blocks portion of the formula (from the documentation) I get this error

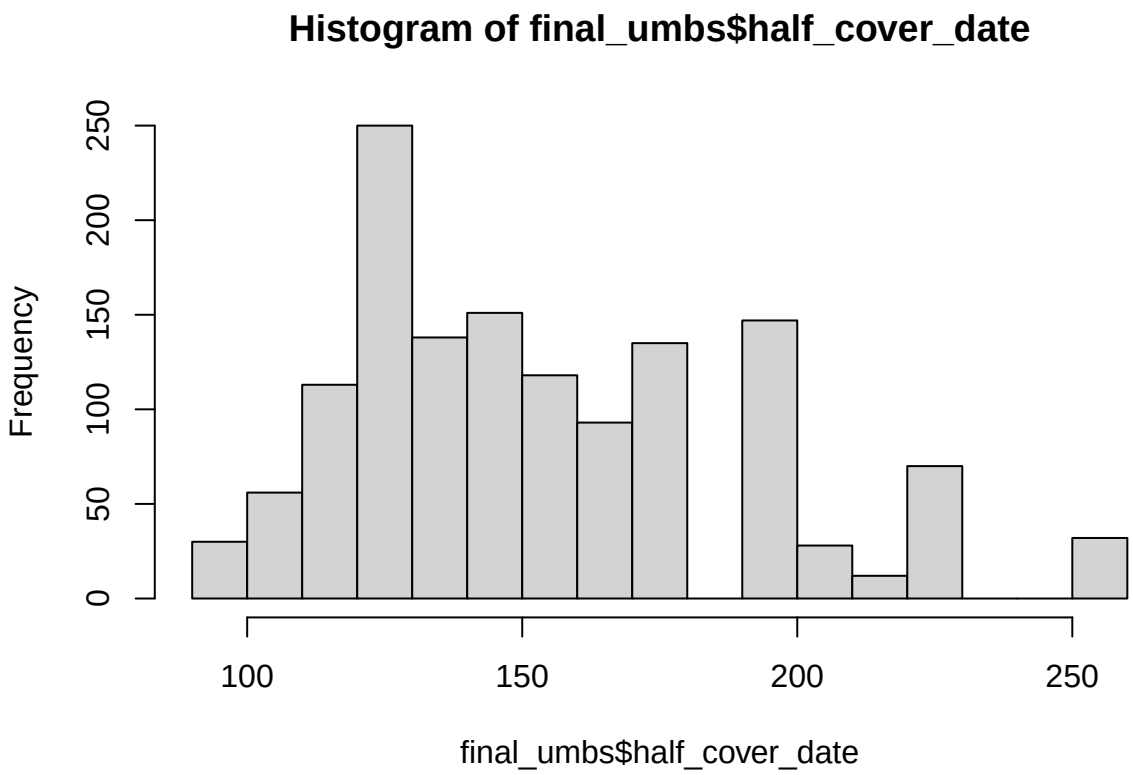
```
#friedman_kbs <- final_kbs %>%
#  friedman_test(half_cover_date ~ state | plot)
```

Error in friedman.test.default(c(141L, 202L, 122L, 101L, 127L, 120L, 197L, : not an unreplicated complete block design

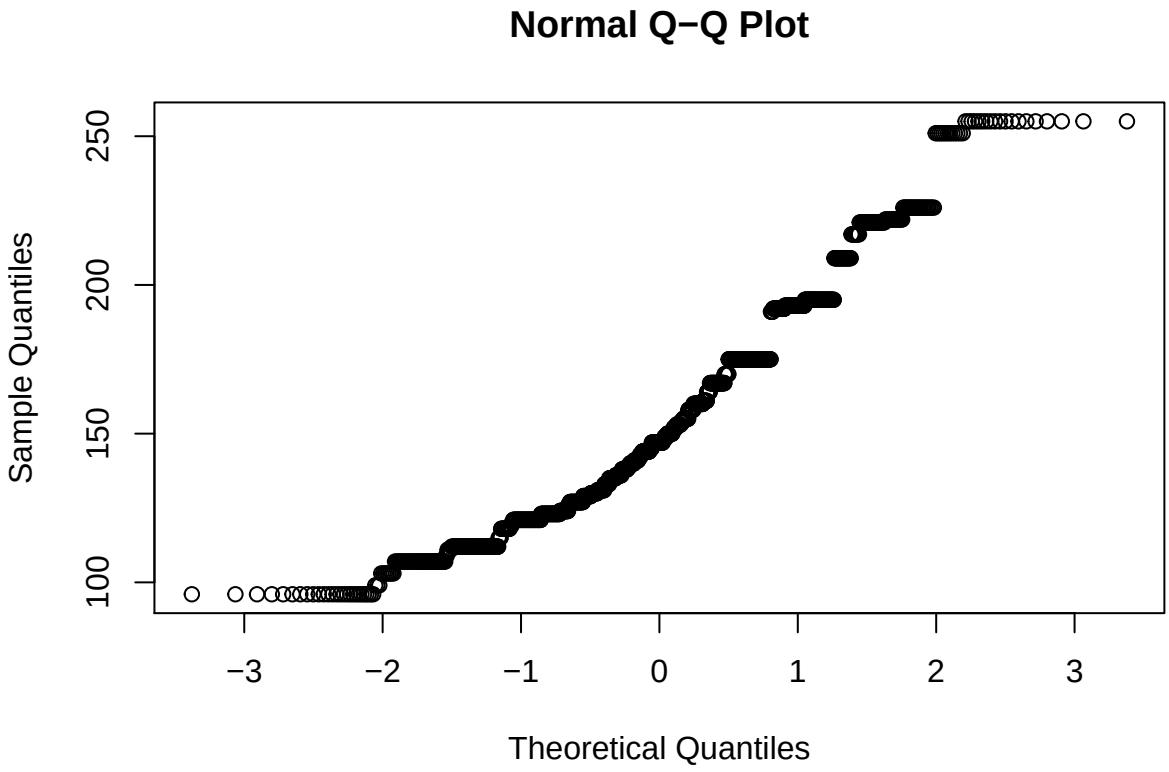
UMBS

Checking for normality

```
hist(final_umbs$half_cover_date)
```



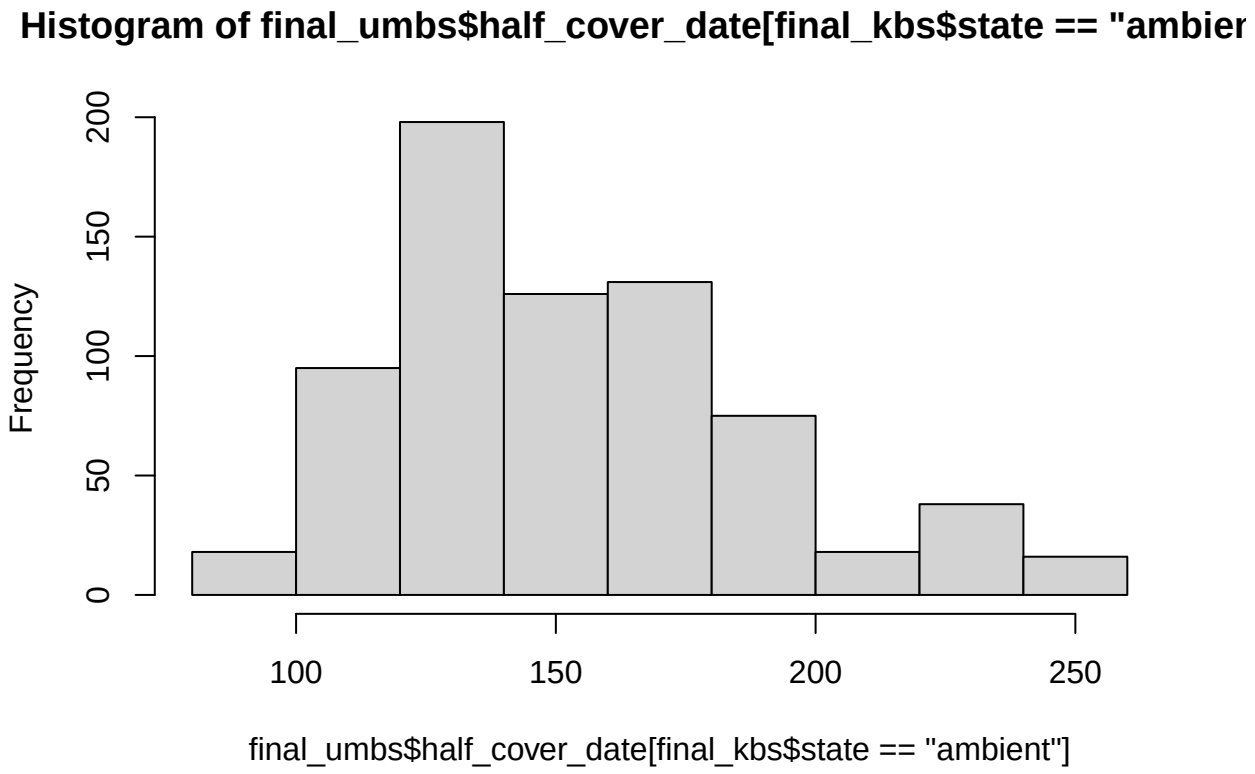

```
qqnorm(final_umbs$half_cover_date)
```



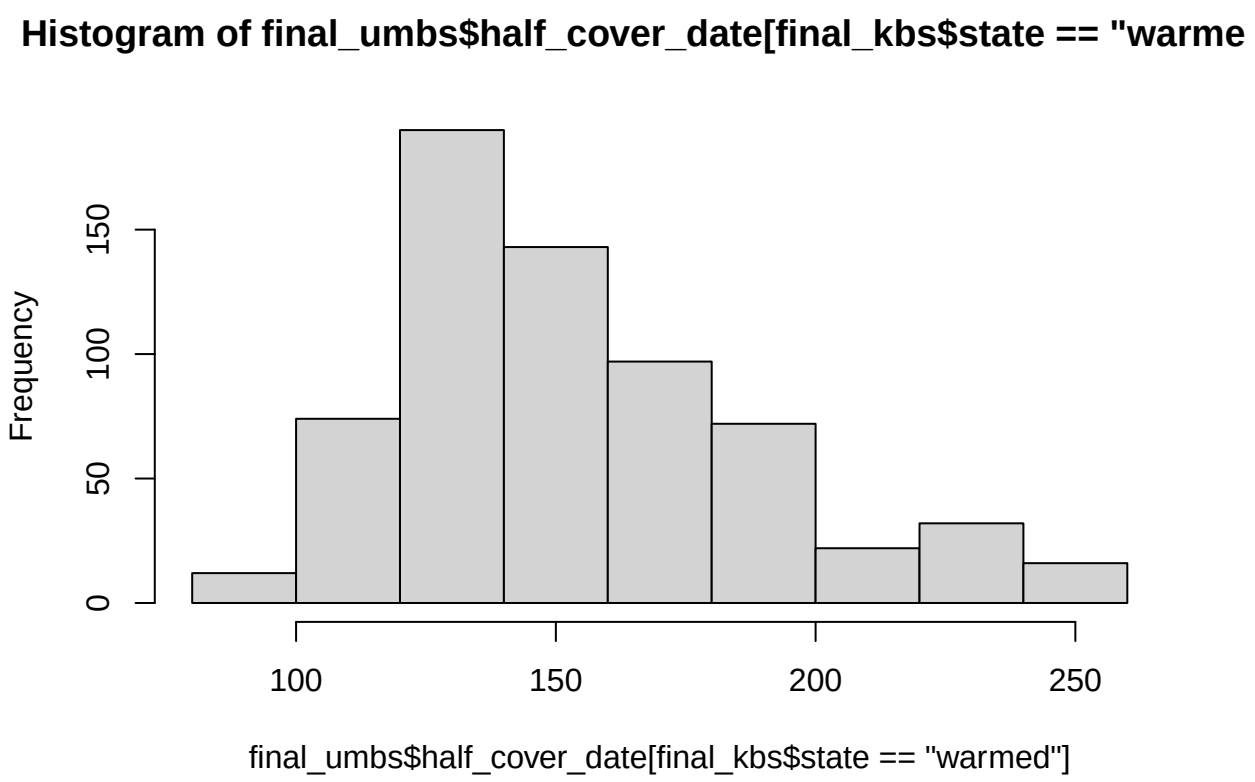

```
shapiro.test(final_umbs$half_cover_date)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  final_umbs$half_cover_date
## W = 0.94754, p-value < 2.2e-16
```

```
hist(final_umbs$half_cover_date[final_kbs$state == "ambient"])
```



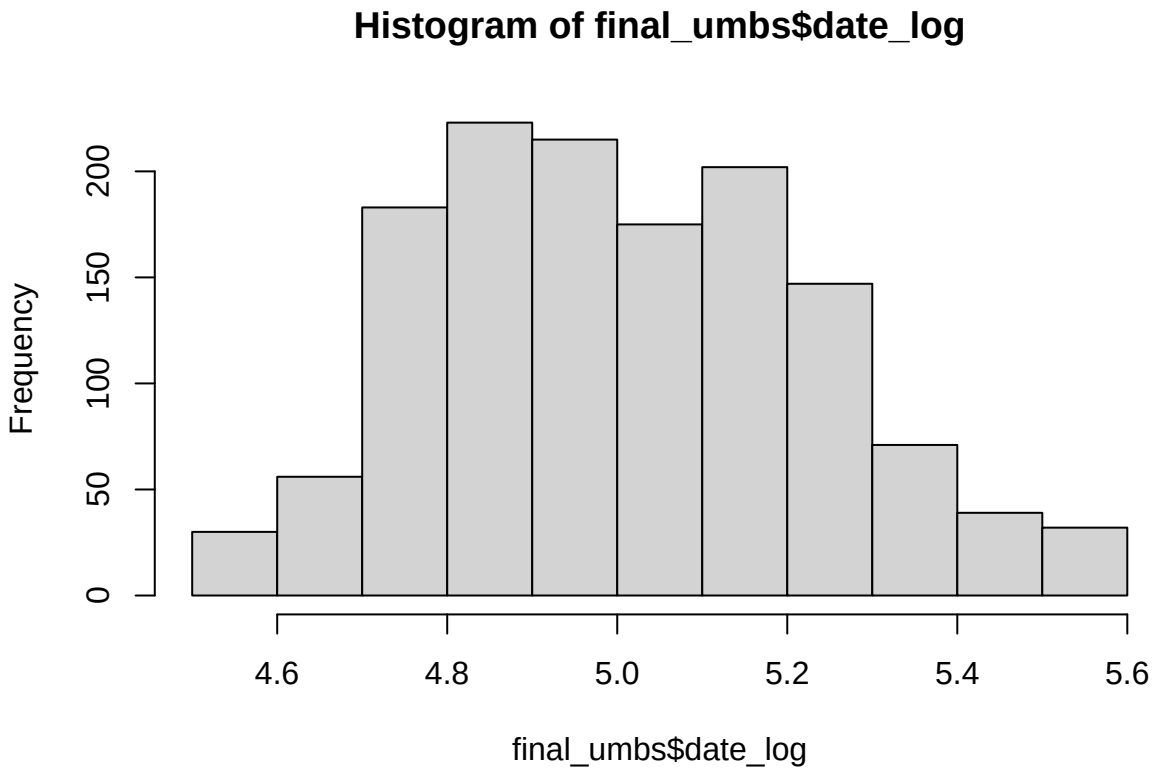

```
hist(final_umbs$half_cover_date[final_kbs$state == "warmed"])
```



These look pretty good

Trying log transformation

```
final_umbs$date_log <- log(final_umbs$half_cover_date)
hist(final_umbs$date_log)
```



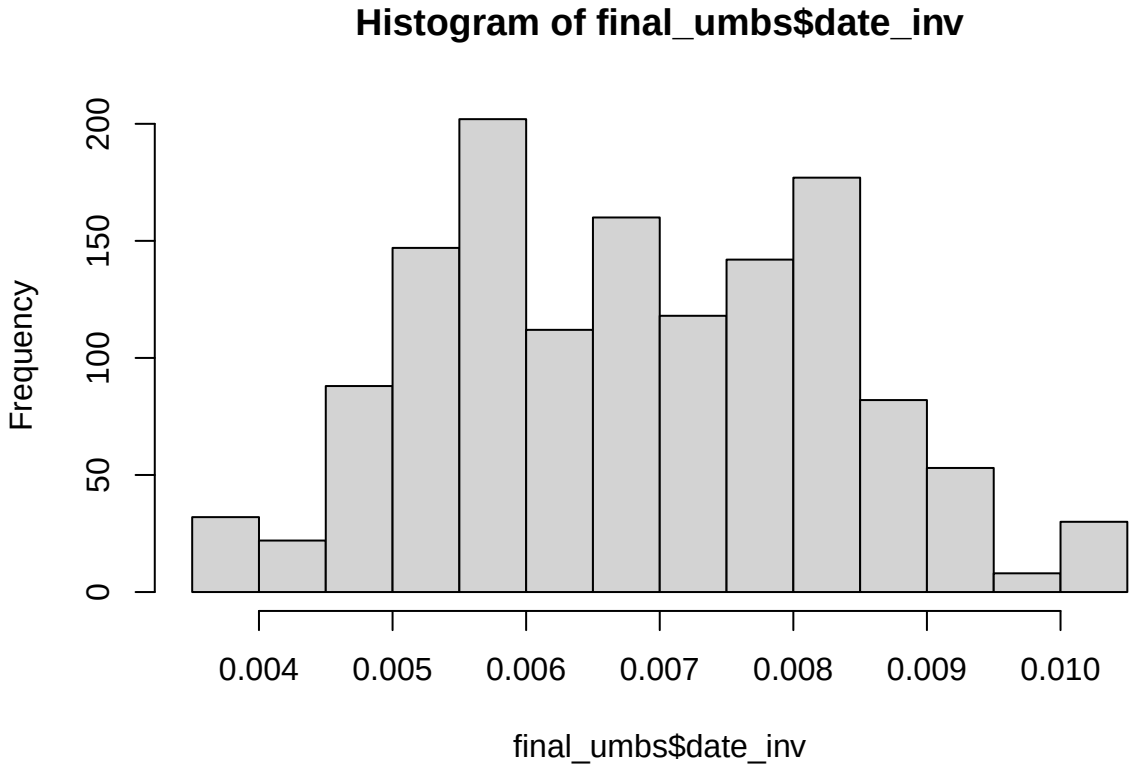

```
shapiro.test(final_umbs$date_log)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  final_umbs$date_log
## W = 0.97728, p-value = 6.765e-14
```

I think this looks good but shapiro-wilk is lower than 0.05

Trying inverse tranformation

```
final_umbs$date_inv <- 1/(final_umbs$half_cover_date)
hist(final_umbs$date_inv)
```




```
shapiro.test(final_umbs$date_inv)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  final_umbs$date_inv
## W = 0.97993, p-value = 6.679e-13
```

This also looks good but is also still low for shapiro-wilk