Phenology Analysis

Moriah Young

February 19, 2021

COLLABORATORS: Phoebe Zarnetske, Mark Hammond, Pat Bills, Kara Dobson
DATA INPUT: Cleaned phenology data csv from the shared Google drive
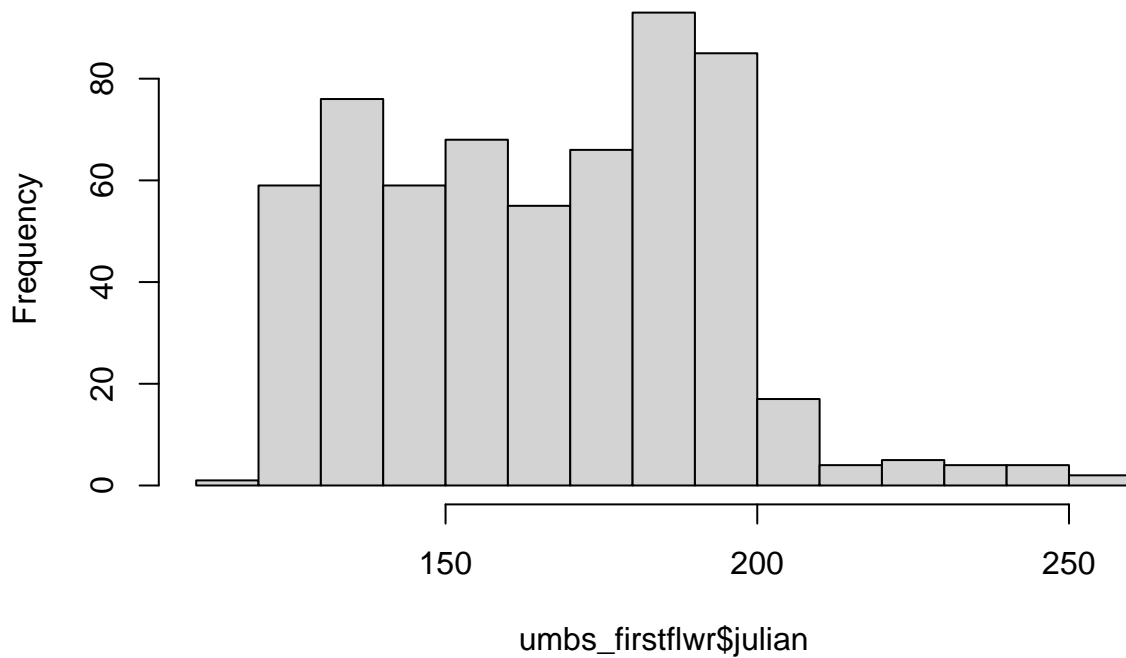DATA OUTPUT: Code and Rmd are in the scripts folder in Github
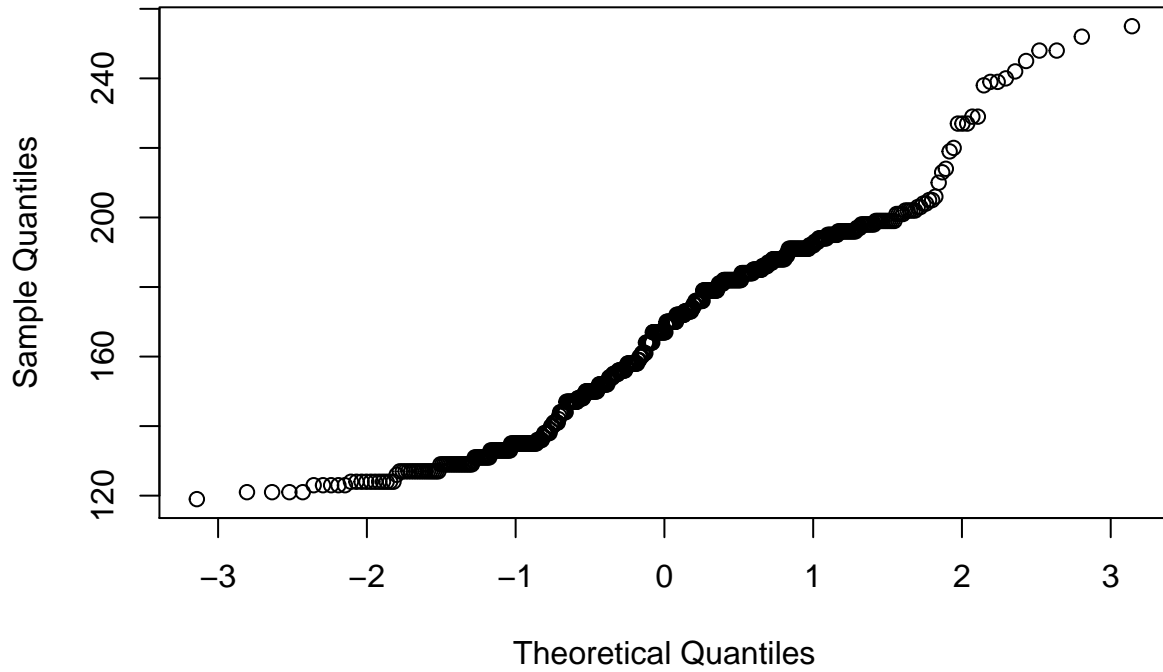PROJECT: warmXtrophic

UMBS Data

First Flower

```
umbs_firstflwr <- subset(FirstFlower_all, site == "umbs")  # pull out umbs only data

# first flower
hist(umbs_firstflwr$julian)  # poisson?
```
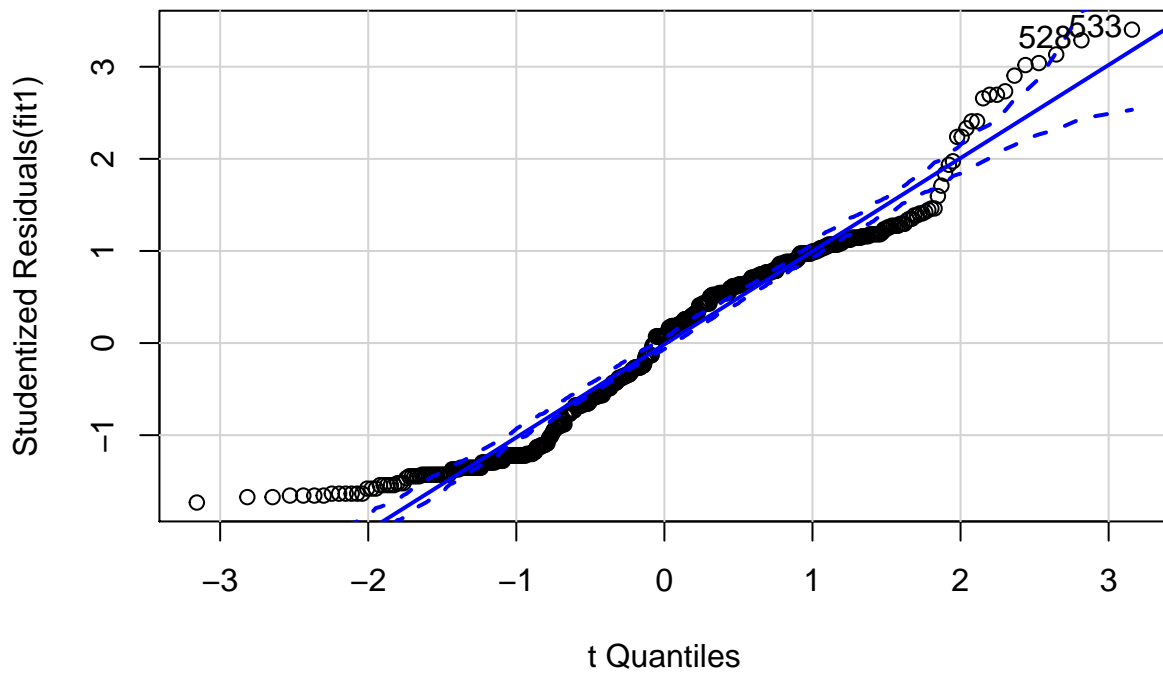

Histogram of umbs_firstflwr$julian

```
qqnorm(umbs_firstflwr$julian)
```

Normal Q–Q Plot


```
fit1 <- lm(julian ~ state, data = umbs_firstflwr)
qqPlot(fit1)
```
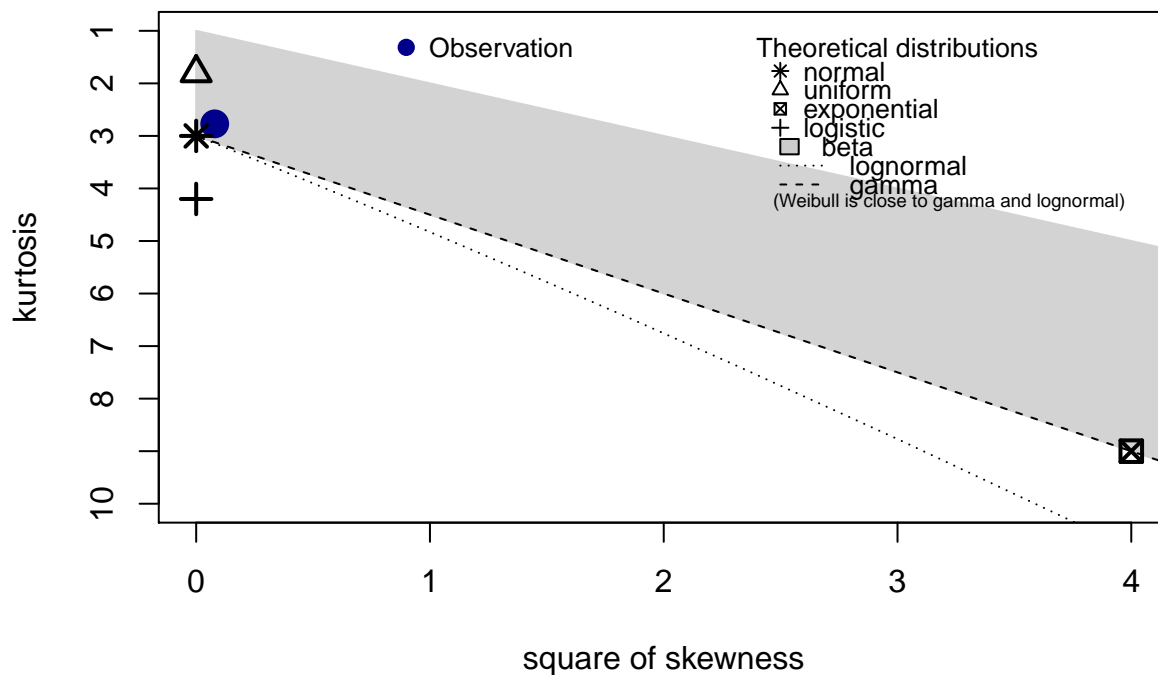


```
## [1] 528 533
```

```
shapiro.test(umbs_firstflwr$julian)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  umbs_firstflwr$julian
## W = 0.96131, p-value = 1.928e-11
```

```
descdist(umbs_firstflwr$julian, discrete = FALSE)  # uniform? normal?
```
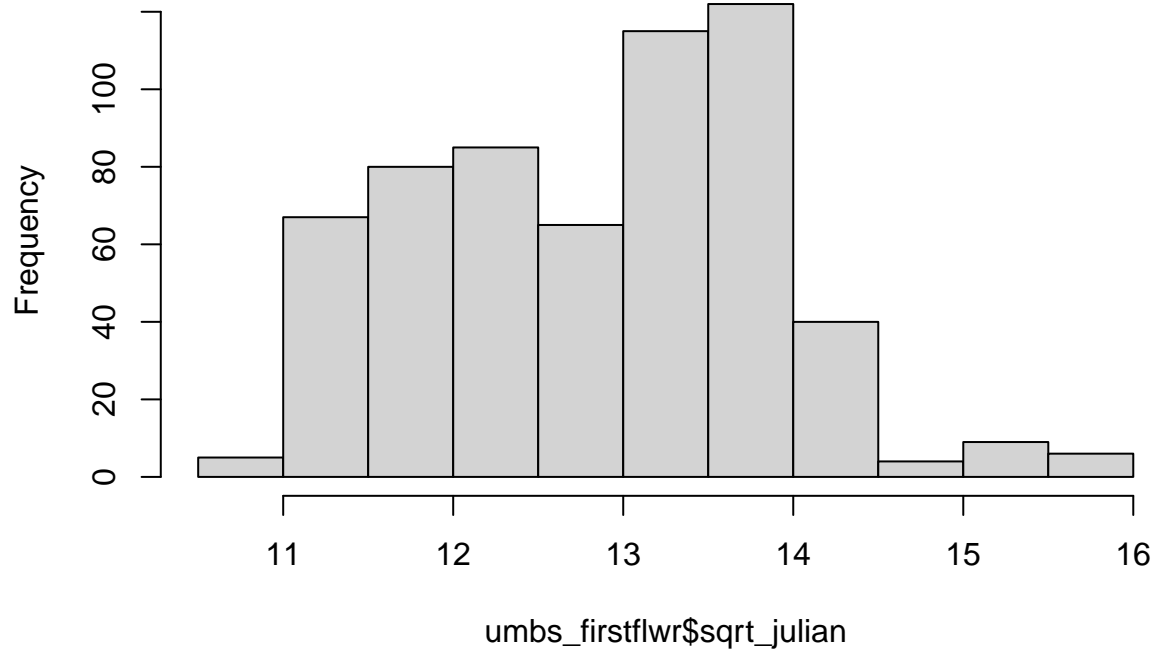

Cullen and Frey graph

```
## summary statistics
## ------
## min:  119    max:   255
## median:  167
## mean:  166.2709
## estimated sd:  26.70761
## estimated skewness:  0.281686
## estimated kurtosis:  2.770431
```

```
skewness(umbs_firstflwr$julian)
```

```
## [1] 0.2809789
```

```
umbs_firstflwr$sqrt_julian <- sqrt(umbs_firstflwr$julian)
hist(umbs_firstflwr$sqrt_julian)
```

Histogram of umbs_firstflwr$sqrt_julian
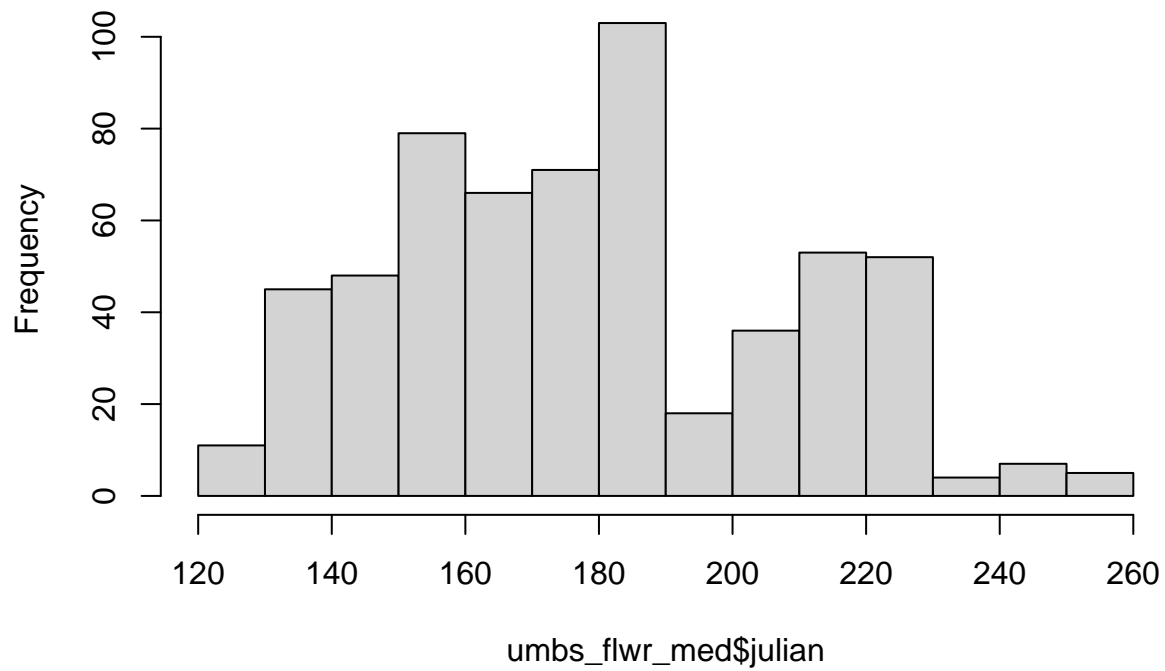

umbs_firstflwr$sqrt_julian

```r
skewness(umbs_firstflwr$sqrt_julian, na.rm = TRUE)
```

```
## [1] 0.1021979
```

Median Flower

```r
umbs_flwr_med <- subset(MedianFlwr_all, site == "umbs")  # pull out umbs only data

hist(umbs_flwr_med$julian)
```

Histogram of umbs_flwr_med$julian


```
qqnorm(umbs_flwr_med$julian)
```

Normal Q−Q Plot
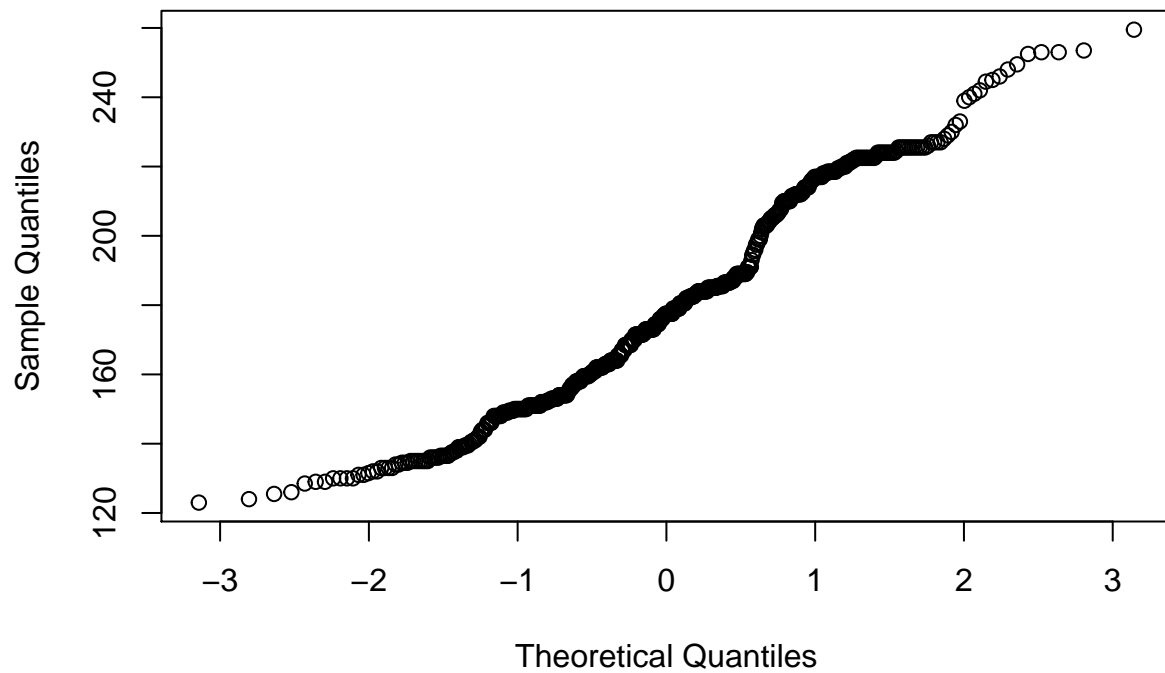
```
fit2 <- lm(julian ~ state, data = umbs_flwr_med)
qqPlot(fit2)
```



```
## [1] 528 533
```
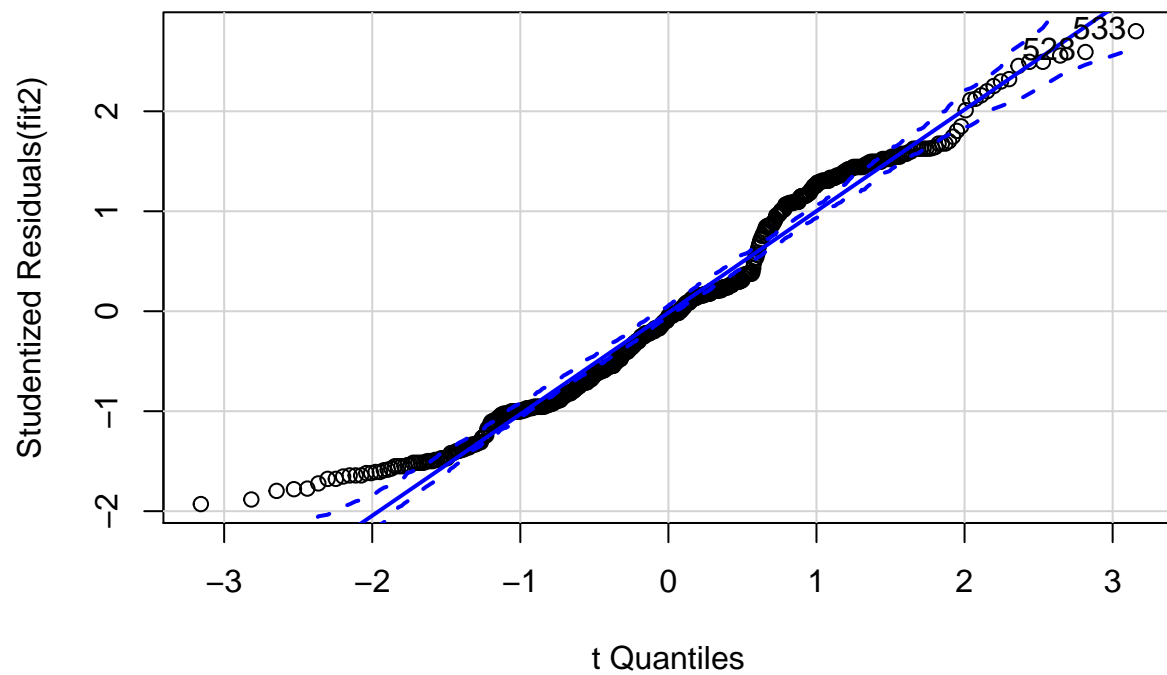
```
shapiro.test(umbs_flwr_med$julian)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  umbs_flwr_med$julian
## W = 0.96918, p-value = 6.893e-10
```

```
descdist(umbs_flwr_med$julian, discrete = FALSE)  # uniform?
```

Cullen and Frey graph
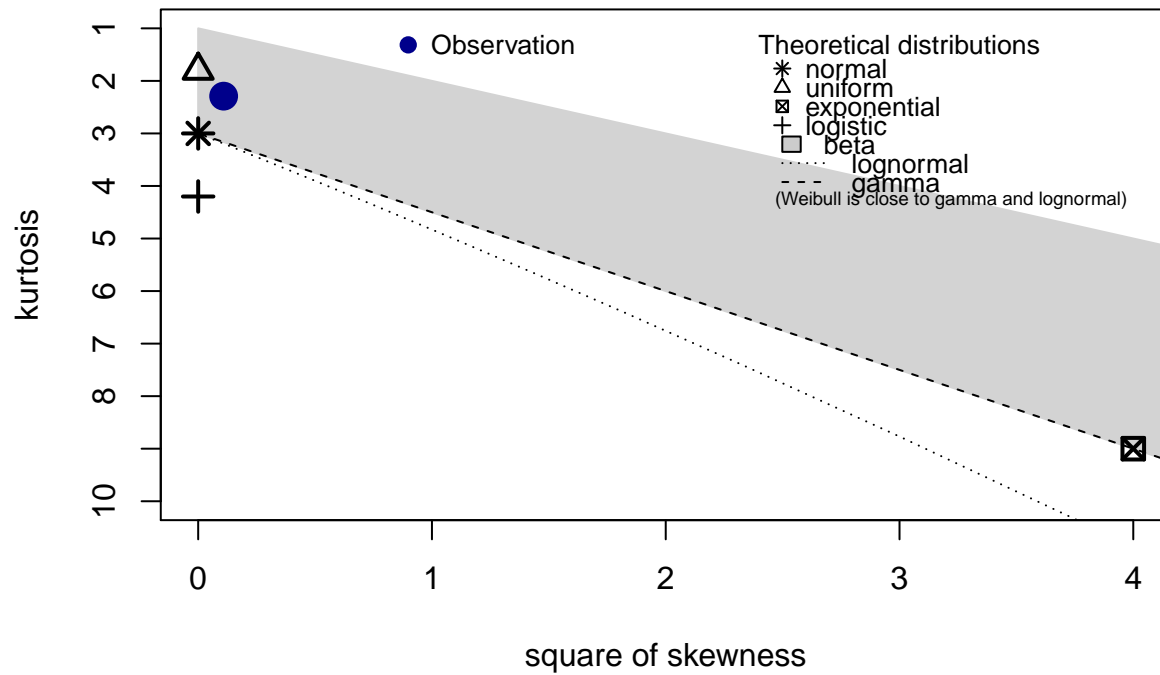

```
## summary statistics
## ------
## min:  123    max:  259.5
## median:  177.5
## mean:  179.1296
## estimated sd:  29.31017
## estimated skewness:  0.3300838
## estimated kurtosis:  2.290566
```

```r
skewness(umbs_flwr_med$julian, na.rm = TRUE)  # slight positive skew
```

```
## [1] 0.3292553
```

```r
# Because of slight positive skew I'm transforming the data by taking the square
# root.
umbs_flwr_med$sqrt_julian <- sqrt(umbs_flwr_med$julian)
hist(umbs_flwr_med$sqrt_julian)
```

Histogram of umbs_flwr_med$sqrt_julian

umbs_flwr_med$sqrt_julian
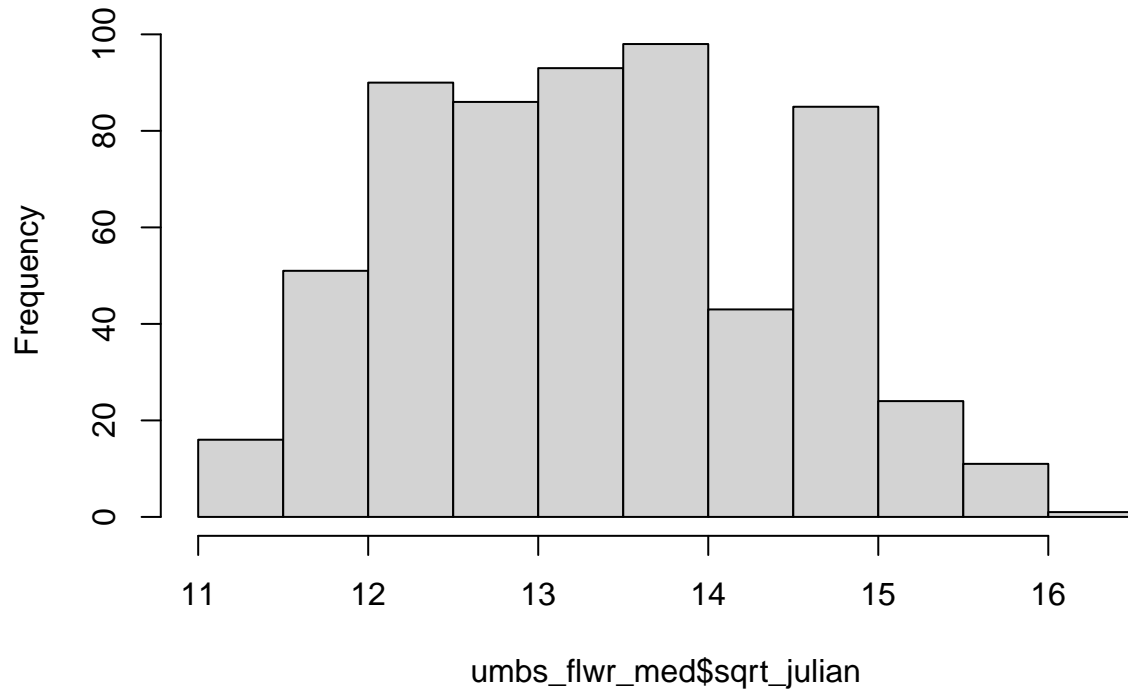
```r
skewness(umbs_flwr_med$sqrt_julian, na.rm = TRUE)
```

```
## [1] 0.187288
```

```r
ggdensity(umbs_flwr_med, x = "sqrt_julian", fill = "lightgray", title = "CONT") +
    stat_overlay_normal_density(color = "red", linetype = "dashed")
```
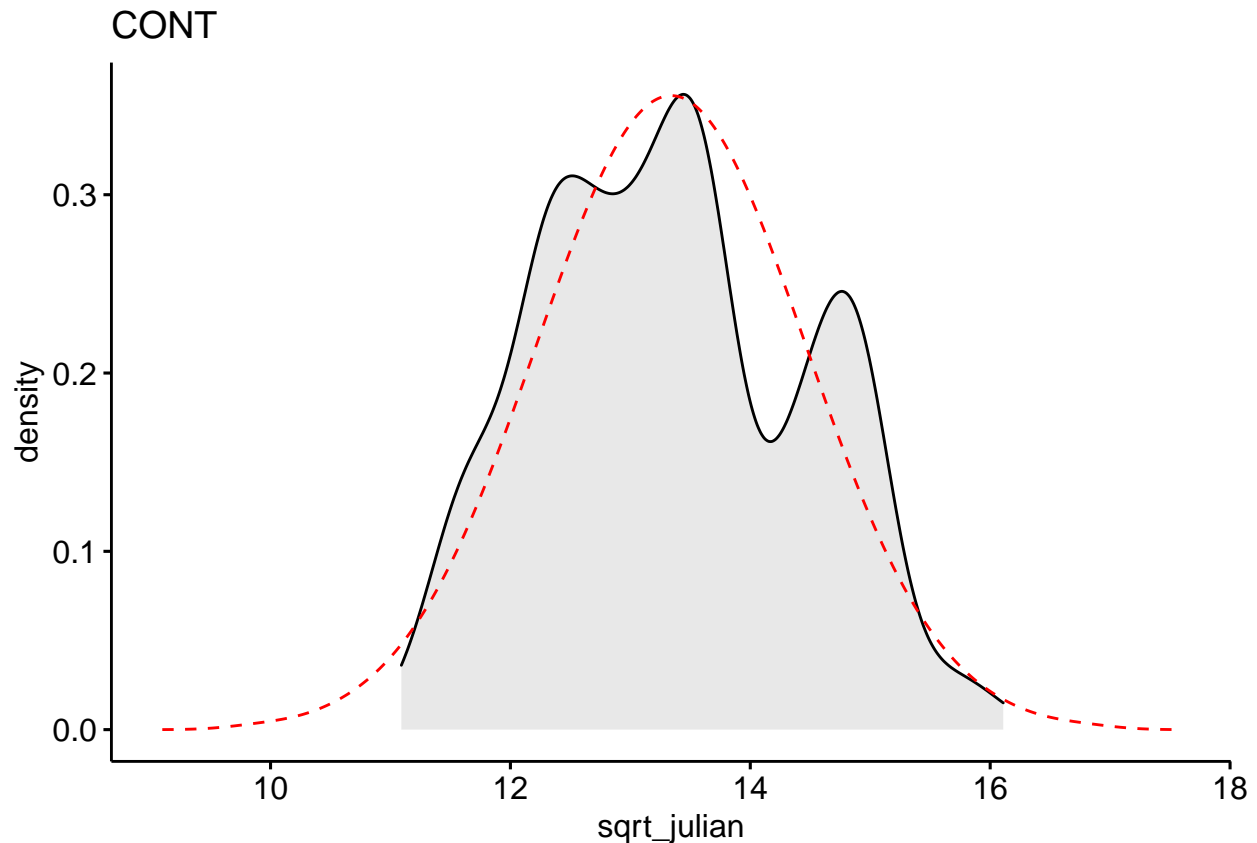
```
# there's that weird dip around 14 on the x-axis
```

Models! What is the effect of warming (state) on the median date of flower?

```
# state, year and insecticide are fixed effects with interaction btw state and
# year - species and plot are random effects

# normal dist
umbsMED_norm <- lmer(sqrt_julian ~ state * year_factor + insecticide + (1 | species) +
    (1 | plot), data = umbs_flwr_med, REML = FALSE)
```

```
## boundary (singular) fit: see ?isSingular
```

```
summary(umbsMED_norm)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: sqrt_julian ~ state * year_factor + insecticide + (1 | species) +
##     (1 | plot)
##    Data: umbs_flwr_med
##
##      AIC      BIC   logLik deviance df.resid
##    556.1    617.6   -264.1    528.1      584
##
## Scaled residuals:
```

```
##      Min      1Q Median      3Q     Max
## -5.9560 -0.5553 -0.0481  0.6076  3.1940
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  species  (Intercept) 1.4509   1.2045
##  plot     (Intercept) 0.0000   0.0000
##  Residual             0.1144   0.3383
## Number of obs: 598, groups:  species, 30; plot, 24
##
## Fixed effects:
##                             Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)                13.284926   0.229672 33.417175  57.843  < 2e-16 ***
## statewarmed                -0.076647   0.070570 569.956854  -1.086  0.27788
## year_factor3                0.185417   0.068576 571.254488   2.704  0.00706 **
## year_factor4                0.271762   0.068068 569.218046   3.993 7.39e-05 ***
## year_factor5                0.481660   0.066097 569.834495   7.287 1.06e-12 ***
## year_factor6                0.377016   0.065514 569.588411   5.755 1.42e-08 ***
## insecticideno_insects      -0.037231   0.028926 569.953425  -1.287  0.19857
## statewarmed:year_factor3    0.009110   0.095666 570.804816   0.095  0.92417
## statewarmed:year_factor4   -0.070302   0.095122 569.440556  -0.739  0.46017
## statewarmed:year_factor5   -0.003494   0.092306 569.861732  -0.038  0.96982
## statewarmed:year_factor6   -0.014094   0.091263 569.429882  -0.154  0.87732
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##            (Intr) sttwrm yr_fc3 yr_fc4 yr_fc5 yr_fc6 insct_ stt:_3 stt:_4
## statewarmed -0.155
## year_factr3 -0.169  0.508
## year_factr4 -0.159  0.511  0.538
## year_factr5 -0.168  0.526  0.551  0.563
## year_factr6 -0.171  0.534  0.560  0.573  0.591
## insctcdn_ns -0.091  0.022  0.014  0.009  0.028  0.029
## sttwrmd:y_3  0.115 -0.732 -0.701 -0.374 -0.382 -0.387 -0.034
## sttwrmd:y_4  0.110 -0.737 -0.372 -0.702 -0.388 -0.390 -0.009  0.545
## sttwrmd:y_5  0.118 -0.759 -0.383 -0.390 -0.702 -0.401 -0.041  0.558  0.566
## sttwrmd:y_6  0.121 -0.771 -0.389 -0.395 -0.407 -0.693 -0.038  0.566  0.570
##            stt:_5
## statewarmed
## year_factr3
## year_factr4
## year_factr5
## year_factr6
## insctcdn_ns
## sttwrmd:y_3
## sttwrmd:y_4
## sttwrmd:y_5
## sttwrmd:y_6  0.587
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see ?isSingular
```

```r
umbsFF_norm <- lmer(sqrt_julian ~ state * year_factor + insecticide + (1 | species) +
    (1 | plot), data = umbs_firstflwr, REML = FALSE)
```

```
## boundary (singular) fit: see ?isSingular
```

```r
summary(umbsFF_norm)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: sqrt_julian ~ state * year_factor + insecticide + (1 | species) +
##     (1 | plot)
##    Data: umbs_firstflwr
##
##      AIC       BIC   logLik deviance df.resid
##    833.4     894.9   -402.7    805.4      584
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.4481 -0.4587 -0.0401  0.4855  3.8916
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  species  (Intercept) 1.3916   1.180
##  plot     (Intercept) 0.0000   0.000
##  Residual             0.1866   0.432
## Number of obs: 598, groups:  species, 30; plot, 24
##
## Fixed effects:
##                            Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)               13.020511   0.231268  35.666865  56.301  < 2e-16 ***
## statewarmed               -0.176740   0.090072 571.144985  -1.962 0.050223 .
## year_factor3              -0.006892   0.087495 573.056632  -0.079 0.937240
## year_factor4               0.290144   0.086898 569.919464   3.339 0.000896 ***
## year_factor5               0.511511   0.084367 570.875961   6.063 2.43e-09 ***
## year_factor6               0.355532   0.083629 570.527942   4.251 2.48e-05 ***
## insecticideno_insects     -0.048498   0.036919 571.163350  -1.314 0.189505
## statewarmed:year_factor3   0.128240   0.122074 572.401094   1.051 0.293927
## statewarmed:year_factor4   0.047980   0.121428 570.304058   0.395 0.692892
## statewarmed:year_factor5   0.102263   0.117818 570.987523   0.868 0.385773
## statewarmed:year_factor6   0.082943   0.116502 570.286883   0.712 0.476790
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) sttwrm yr_fc3 yr_fc4 yr_fc5 yr_fc6 insct_ stt:_3 stt:_4
## statewarmed -0.197
## year_factr3 -0.214  0.509
## year_factr4 -0.202  0.511  0.538
## year_factr5 -0.213  0.526  0.551  0.563
## year_factr6 -0.217  0.534  0.561  0.573  0.591
## insctcdn_ns -0.115  0.022  0.014  0.009  0.028  0.029
## sttwrmd:y_3  0.146 -0.733 -0.701 -0.374 -0.382 -0.387 -0.034
## sttwrmd:y_4  0.140 -0.737 -0.373 -0.702 -0.388 -0.390 -0.009  0.545
## sttwrmd:y_5  0.150 -0.759 -0.383 -0.389 -0.702 -0.401 -0.041  0.559  0.565
## sttwrmd:y_6  0.153 -0.771 -0.389 -0.394 -0.407 -0.694 -0.038  0.566  0.570
##             stt:_5
## statewarmed
```

```
## year_factr3
## year_factr4
## year_factr5
## year_factr6
## insctcdn_ns
## sttwrmd:y_3
## sttwrmd:y_4
## sttwrmd:y_5
## sttwrmd:y_6   0.587
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see ?isSingular
```

```
#### More models! This needs more work, I need to write out what each model is doing

model1 <- lmer(julian ~ state + (1 | species), data = umbs_flwr_med, REML = FALSE)
summary(model1)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: julian ~ state + (1 | species)
##    Data: umbs_flwr_med
##
##      AIC      BIC   logLik deviance df.resid
##   4556.8   4574.4  -2274.4   4548.8      594
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.1600 -0.5713  0.0391  0.6061  2.8542
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  species  (Intercept) 1152.4   33.947
##  Residual              95.4    9.767
## Number of obs: 598, groups:  species, 30
##
## Fixed effects:
##             Estimate Std. Error       df t value Pr(>|t|)
## (Intercept) 184.6390     6.3178  30.2580  29.225  < 2e-16 ***
## statewarmed  -2.4581     0.8266 570.4546  -2.974  0.00306 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr)
## statewarmed -0.069
```

```
model2 <- lmer(julian ~ state * origin + (1 | species) + (1 | year_factor), data = umbs_flwr_med,
    REML = FALSE)
```

```
## fixed-effect model matrix is rank deficient so dropping 1 column / coefficient
```

```
summary(model2)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: julian ~ state * origin + (1 | species) + (1 | year_factor)
##    Data: umbs_flwr_med
##
##      AIC      BIC   logLik deviance df.resid
##   4477.5   4521.4  -2228.8   4457.5      588
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.5869 -0.5575 -0.0647  0.6465  3.2150
##
## Random effects:
##  Groups      Name        Variance Std.Dev.
##  species     (Intercept) 1073.61  32.766
##  year_factor (Intercept)   21.38   4.624
##  Residual                  79.59   8.921
## Number of obs: 598, groups:  species, 30; year_factor, 5
##
## Fixed effects:
##                         Estimate Std. Error     df t value Pr(>|t|)
## (Intercept)               159.38      34.03  32.98   4.683 4.68e-05 ***
## statewarmed                26.00      48.03  32.71   0.541    0.592
## originBoth                -14.39      41.48  32.36  -0.347    0.731
## originExotic               22.17      35.29  32.50   0.628    0.534
## originNative               30.41      35.12  32.58   0.866    0.393
## statewarmed:originExotic  -28.61      48.03  32.73  -0.596    0.555
## statewarmed:originNative  -28.08      48.05  32.77  -0.584    0.563
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) sttwrm orgnBt orgnEx orgnNt sttw:E
## statewarmed -0.706
## originBoth  -0.001 -0.579
## originExotc -0.961  0.680  0.000
## originNativ -0.965  0.684  0.000  0.931
## sttwrmd:rgE  0.705 -1.000  0.579 -0.681 -0.684
## sttwrmd:rgN  0.705 -1.000  0.579 -0.680 -0.684  0.999
## fit warnings:
## fixed-effect model matrix is rank deficient so dropping 1 column / coefficient
```

```
model3 <- lmer(julian ~ state + origin + (1 | species), data = umbs_flwr_med, REML = FALSE)
summary(model3)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: julian ~ state + origin + (1 | species)
##    Data: umbs_flwr_med
##
```

```
##      AIC      BIC   logLik deviance df.resid
##   4561.9   4592.7  -2274.0   4547.9      591
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.1612 -0.5707  0.0387  0.6068  2.8544
##
## Random effects:
##  Groups    Name         Variance Std.Dev.
##  species   (Intercept) 1119.98   33.466
##  Residual                95.39    9.767
## Number of obs: 598, groups:  species, 30
##
## Fixed effects:
##              Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)  172.2261    24.6548  33.1689   6.986 5.33e-08 ***
## statewarmed   -2.4523     0.8266 570.3389  -2.967  0.00314 **
## originBoth     5.4837    34.6898  32.4882   0.158  0.87537
## originExotic   8.6220    26.5177  32.6064   0.325  0.74715
## originNative  18.3418    26.2767  32.7596   0.698  0.49009
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) sttwrm orgnBt orgnEx
## statewarmed -0.017
## originBoth  -0.710 -0.012
## originExotc -0.929 -0.001  0.661
## originNativ -0.938  0.001  0.667  0.872
```

Updated from Kileigh's script

```r
# should plot be a random effect?
moda <- lmer(sqrt_julian ~ state * origin + insecticide + (1 | species) + (1 | plot),
    umbs_flwr_med, REML = FALSE)
```

```
## fixed-effect model matrix is rank deficient so dropping 1 column / coefficient
```

```
## boundary (singular) fit: see ?isSingular
```

```r
mod1 <- lmer(sqrt_julian ~ state * origin + insecticide + (1 | species), umbs_flwr_med,
    REML = FALSE)
```

```
## fixed-effect model matrix is rank deficient so dropping 1 column / coefficient
```

```r
anova(mod1, moda)
```

```
## Data: umbs_flwr_med
## Models:
## mod1: sqrt_julian ~ state * origin + insecticide + (1 | species)
## moda: sqrt_julian ~ state * origin + insecticide + (1 | species) +
## moda:     (1 | plot)
```

```
##      npar    AIC    BIC  logLik deviance Chisq Df Pr(>Chisq)
## mod1   10 657.62 701.56 -318.81   637.62
## moda   11 659.62 707.95 -318.81   637.62     0  1          1
```

```r
AIC(mod1, moda)
```

```
##      df      AIC
## mod1 10 657.6214
## moda 11 659.6214
```

```r
# Do we need insects?
mod2 <- lmer(sqrt_julian ~ state * origin + (1 | species), umbs_flwr_med, REML = FALSE)
```

```
## fixed-effect model matrix is rank deficient so dropping 1 column / coefficient
```

```r
anova(mod2, mod1)
```

```
## Data: umbs_flwr_med
## Models:
## mod2: sqrt_julian ~ state * origin + (1 | species)
## mod1: sqrt_julian ~ state * origin + insecticide + (1 | species)
##      npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## mod2    9 657.08 696.62 -319.54   639.08
## mod1   10 657.62 701.56 -318.81   637.62 1.4592  1     0.2271
```

```r
# Do we need interaction term?
mod3 <- lmer(sqrt_julian ~ state + origin + (1 | species), umbs_flwr_med, REML = FALSE)
anova(mod3, mod2)
```

```
## Data: umbs_flwr_med
## Models:
## mod3: sqrt_julian ~ state + origin + (1 | species)
## mod2: sqrt_julian ~ state * origin + (1 | species)
##      npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## mod3    7 653.49 684.25 -319.75   639.49
## mod2    9 657.08 696.62 -319.54   639.08 0.4101  2     0.8146
```

```r
anova(mod3)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##         Sum Sq Mean Sq NumDF  DenDF F value  Pr(>F)
## state  1.2126 1.21257     1 570.62  8.7183 0.00328 **
## origin 0.0940 0.03133     3  31.58  0.2253 0.87810
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
summary(mod3)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: sqrt_julian ~ state + origin + (1 | species)
##    Data: umbs_flwr_med
##
##      AIC      BIC   logLik deviance df.resid
##    653.5    684.2   -319.7    639.5      591
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.5266 -0.5311  0.0436  0.6151  2.4797
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  species  (Intercept) 1.4679   1.2116
##  Residual             0.1391   0.3729
## Number of obs: 598, groups:  species, 30
##
## Fixed effects:
##              Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)  13.11383    0.89652 33.57863  14.627 3.97e-16 ***
## statewarmed  -0.09319    0.03156 570.61566  -2.953  0.00328 **
## originBoth    0.21597    1.26075 32.81637   0.171  0.86504
## originExotic  0.30173    0.96384 32.94953   0.313  0.75621
## originNative  0.59840    0.95519 33.12020   0.626  0.53530
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) sttwrm orgnBt orgnEx
## statewarmed -0.018
## originBoth  -0.711 -0.013
## originExotc -0.930 -0.001  0.661
## originNativ -0.938  0.001  0.667  0.873
```

```r
confint(mod3, method = "boot", nsim = 999)
```

```
## Computing bootstrap confidence intervals ...


##                   2.5 %      97.5 %
## .sig01        0.8123143  1.41149227
## .sigma        0.3521670  0.39392262
## (Intercept)  11.4197678 14.80310243
## statewarmed  -0.1556652 -0.02741583
## originBoth    -2.0164054  2.56082824
## originExotic  -1.5491234  2.15428315
## originNative  -1.1692723  2.37299371
```

```r
difflsmeans(mod3, test.effs = NULL, ddf = "Satterthwaite")
```

```
## Least Squares Means table:
##
```

```
##       Estimate Std. Error df t value lower upper Pr(>|t|)
##
##   Confidence level: 95%
##   Degrees of freedom method: Satterthwaite
```