

Statistical Analyses of Amaranth Characterization Paper

Ali Baturaygil

June 4 , 2021

Packages

Data processing

- I want to remove biomass genotypes, as we don't have categorical data for them (latitudinal groups, species) and some genotypes which did not show successful field emergence.

```
data <- filter(data, !(plot %in% c(205, 510, 526, 610, 712, 810, 1024)), species!="biomass" )>%
  dplyr::select(genotypes, species, country, plot, seed_setting, ft_adjusted, ph_adjusted, longitude)

data_species <- data
summary(data)
```

```
##   genotypes      species      country      plot
## Length:254      Length:254      Length:254      Min.   : 101.0
## Class :character Class :character Class :character 1st Qu.: 313.5
## Mode  :character Mode  :character Mode  :character Median : 526.0
##                                     Mean  : 562.3
##                                     3rd Qu.: 815.8
##                                     Max.   :1028.0
##
##   seed_setting  ft_adjusted  ph_adjusted  longitude
## Min.   :0.0000  Min.   : 50.0  Min.   : 45.2  Min.   : -108.83
## 1st Qu.:0.0000  1st Qu.:102.0  1st Qu.:196.0  1st Qu.: -90.79
## Median :0.0000  Median :110.0  Median :224.1  Median : -77.59
## Mean   :0.1929  Mean   :106.3  Mean   :220.3  Mean   : -79.99
## 3rd Qu.:0.0000  3rd Qu.:113.0  3rd Qu.:249.2  3rd Qu.: -72.96
## Max.   :1.0000  Max.   :160.0  Max.   :359.9  Max.   : -42.88
##                                     NA's   :11
##                                     NA's   :18
##                                     NA's   :28
##   latitude      elevation      prec      temp
## Min.   : -41.035  Min.   : 43  Min.   : 0.50  Min.   : 3.60
## 1st Qu.: -13.636  1st Qu.:1887  1st Qu.: 53.35  1st Qu.:12.92
## Median : -9.000  Median :2650  Median : 64.46  Median :14.83
## Mean   : -3.286  Mean   :2436  Mean   : 72.36  Mean   :15.38
## 3rd Qu.: 14.694  3rd Qu.:3105  3rd Qu.: 78.75  3rd Qu.:17.89
## Max.   : 27.417  Max.   :4518  Max.   :344.00  Max.   :26.92
## NA's   :26      NA's   :28  NA's   :28  NA's   :28
##   sr      GDD10c
## Min.   :11788  Min.   : 416.1
## 1st Qu.:15077  1st Qu.: 884.4
## Median :16339  Median : 967.0
## Mean   :16222  Mean   : 911.6
## 3rd Qu.:17260  3rd Qu.: 984.9
## Max.   :19655  Max.   :1196.2
```

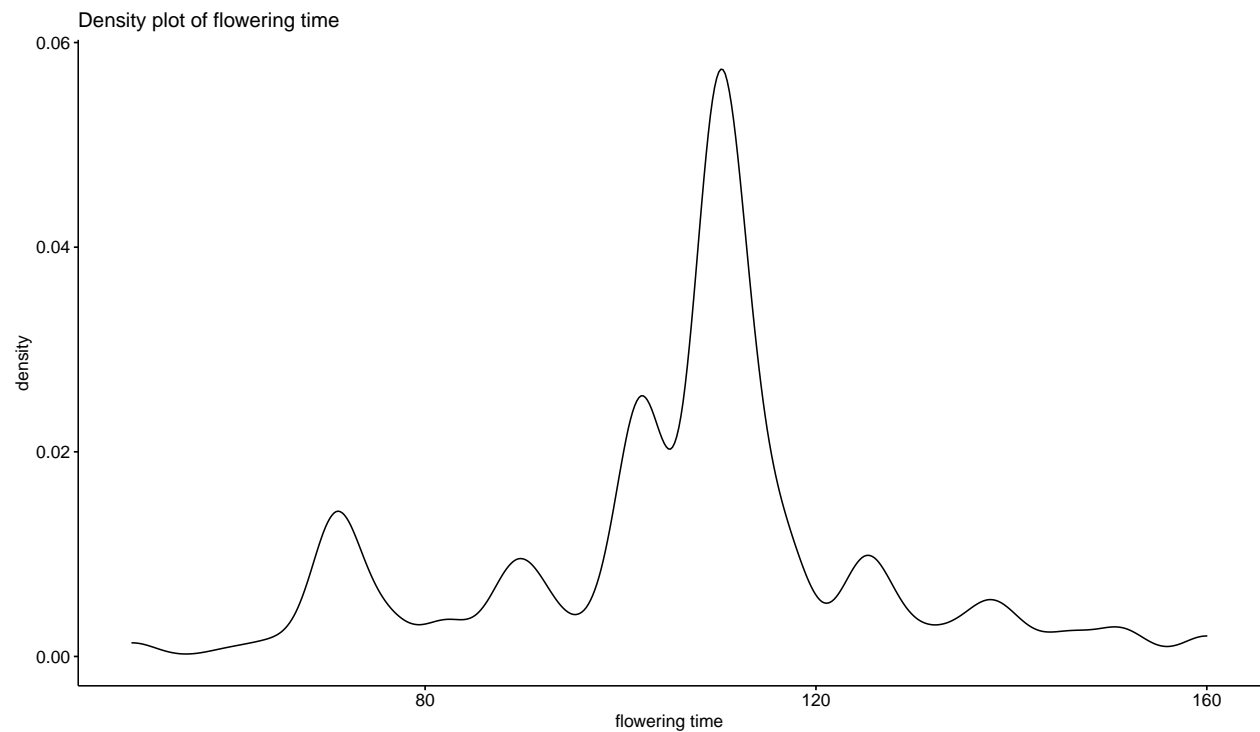
```
## NA's :28      NA's :11
```

1. Comparison of species

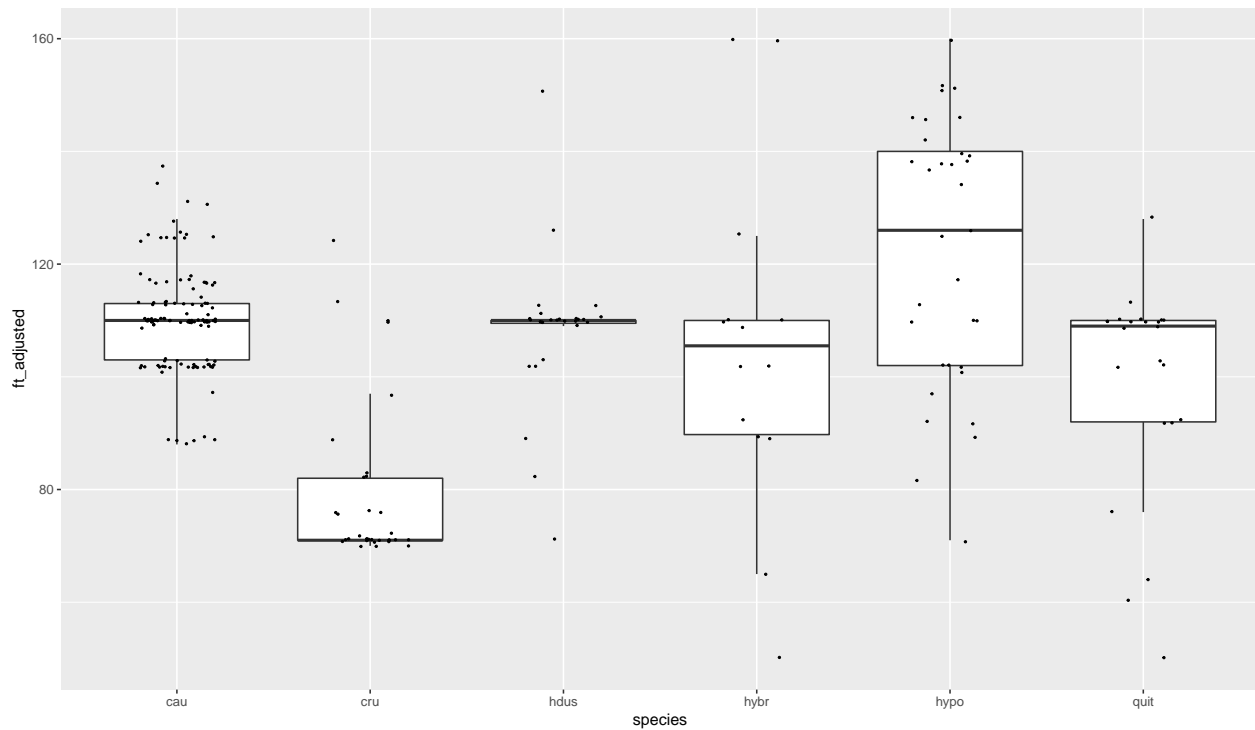
Phenotypic traits

A. Flowering time

```
ggdensity(data_species$ft_adjusted,  
           main = "Density plot of flowering time",  
           xlab = "flowering time")
```



```
ggplot(data = data_species, mapping = aes(x = species, y = ft_adjusted)) +  
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_1 <- glm(ft_adjusted ~ species, data = data_species, family = gaussian)
summary(md_1)
```

```
##
## Call:
## glm(formula = ft_adjusted ~ species, family = gaussian, data = data_species)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -55.214  -8.281  -0.281   6.719  54.786
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   110.281     1.484   74.302 < 2e-16 ***
## speciescru    -30.843     3.170  -9.729 < 2e-16 ***
## specieshdus    -2.021     3.392  -0.596 0.551752
## specieshybr    -5.066     4.488  -1.129 0.260073
## specieshypo    12.053     3.133   3.848 0.000153 ***
## speciesquit   -11.063     3.622  -3.054 0.002515 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 251.1295)
##
##      Null deviance: 94172  on 242  degrees of freedom
## Residual deviance: 59518  on 237  degrees of freedom
## (11 observations deleted due to missingness)
## AIC: 2040.3
##
## Number of Fisher Scoring iterations: 2
```

```

lrtest(md_1)

## Likelihood ratio test
##
## Model 1: ft_adjusted ~ species
## Model 2: ft_adjusted ~ 1
##   #Df LogLik Df Chisq Pr(>Chisq)
## 1    7 -1013.2
## 2    2 -1068.9 -5 111.5 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

lsmeans_species_ft <- LSD.test(md_1, "species", p.adj="none" )
lsmeans_species_ft

## $statistics
##      MSerror Df      Mean      CV
## 251.1295 237 106.2922 14.90897
##
## $parameters
##      test p.adjusted name.t ntr alpha
## Fisher-LSD      none species    6 0.05
##
## $means
##      ft_adjusted      std    r      LCL      UCL Min Max    Q25    Q50 Q75
## cau    110.28070  9.297929 114 107.35677 113.20464  88 137 103.00 110.0 113
## cru    79.43750 14.840958  32  73.91869  84.95631  70 124  71.00  71.0  82
## hdus   108.25926 13.586171  27 102.25114 114.26738  71 151 109.50 110.0 110
## hybr   105.21429 30.209979  14  96.87063 113.55794  50 160  89.75 105.5 110
## hypo   122.33333 23.766661  33 116.89879 127.76788  71 160 102.00 126.0 140
## quit    99.21739 19.284842  23  92.70776 105.72702  50 128  92.00 109.0 110
##
## $comparison
## NULL
##
## $groups
##      ft_adjusted groups
## hypo   122.33333      a
## cau    110.28070      b
## hdus   108.25926      b
## hybr   105.21429     bc
## quit    99.21739      c
## cru    79.43750      d
##
## attr(,"class")
## [1] "group"

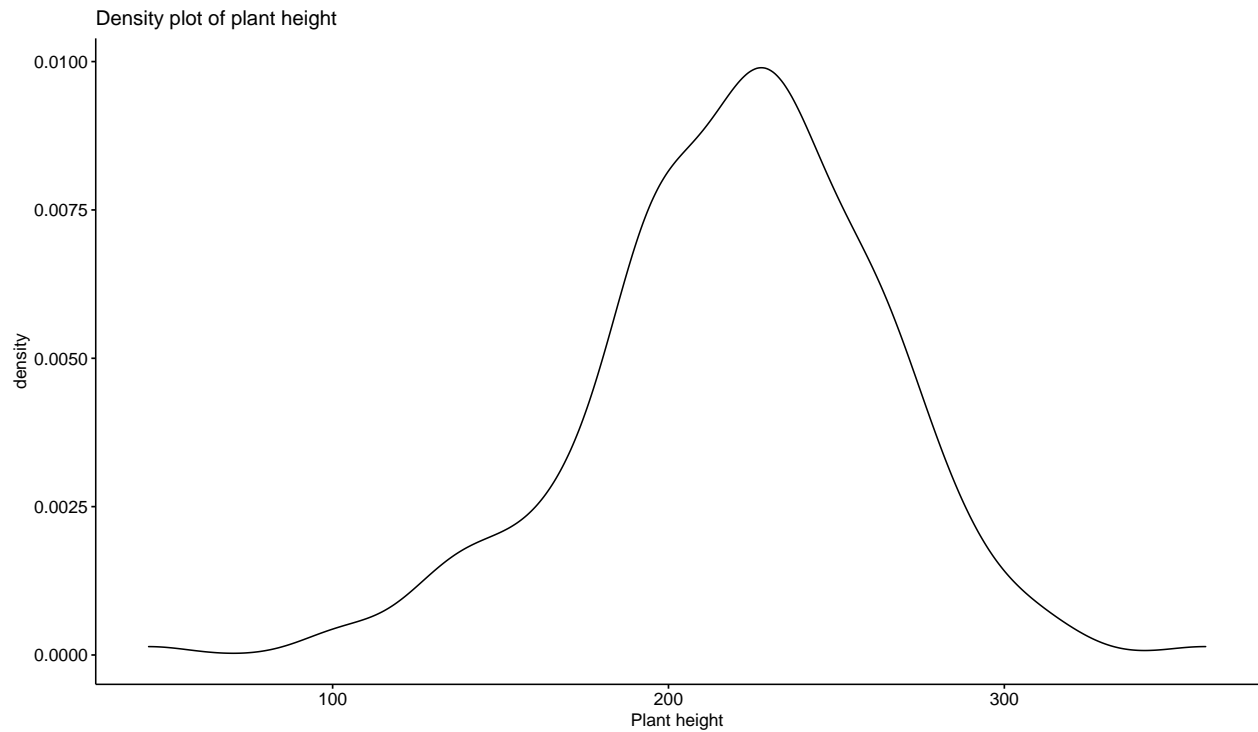
```

B. Plant height

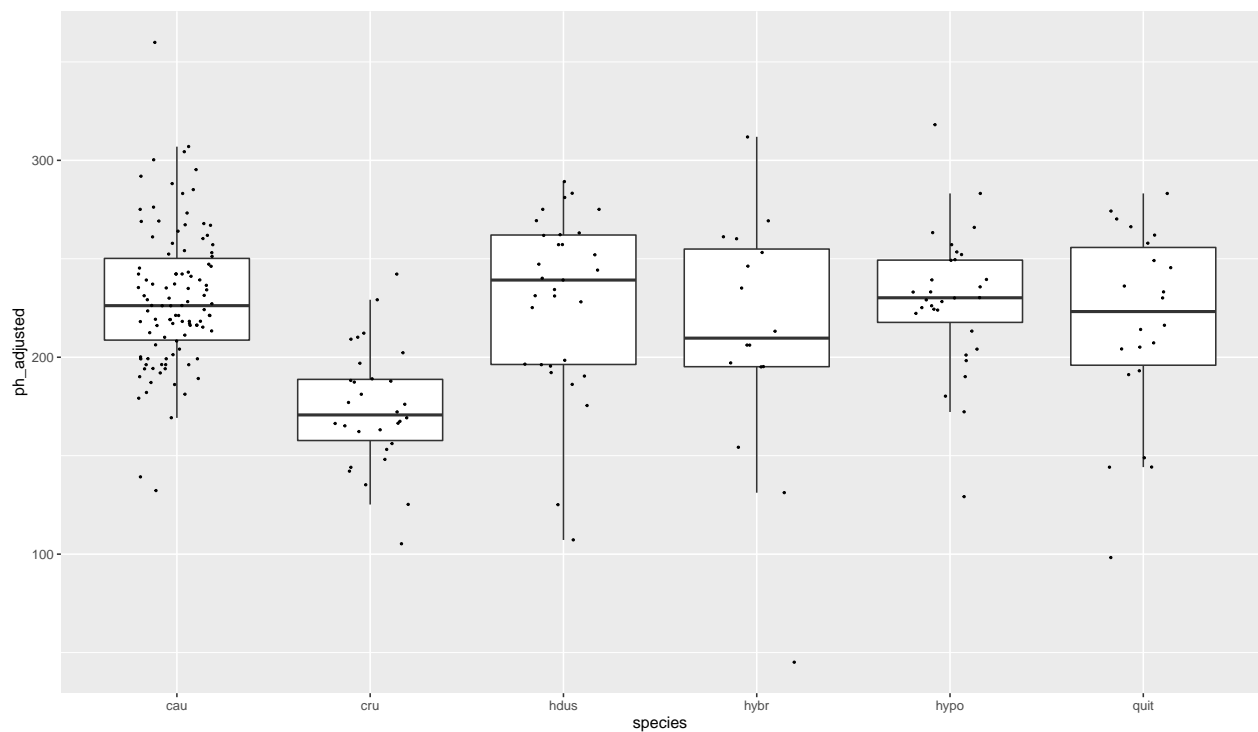
```

ggdensity(data_species$ph_adjusted,
  main = "Density plot of plant height",
  xlab = "Plant height")

```



```
ggplot(data = data_species, mapping = aes(x = species, y = ph_adjusted)) +  
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_2 <-glm(ph_adjusted ~ species , data = data_species, family = gaussian)  
summary(md_2)
```

```
##  
## Call:
```

```
## glm(formula = ph_adjusted ~ species, family = gaussian, data = data_species)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -166.109   -21.357    -0.856    25.053   129.823
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 230.1269     3.8541  59.709 < 2e-16 ***
## speciescru  -55.7602     8.2061  -6.795 9.19e-11 ***
## specieshdus  -0.7253     8.1023  -0.090  0.9288
## specieshybr -18.8175    10.6426  -1.768  0.0784 .
## specieshypo  -1.0559     8.1023  -0.130  0.8964
## speciesquit -13.0860     9.2965  -1.408  0.1606
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 1574.569)
##
##      Null deviance: 442163  on 235  degrees of freedom
## Residual deviance: 362151  on 230  degrees of freedom
## (18 observations deleted due to missingness)
## AIC: 2415
##
## Number of Fisher Scoring iterations: 2
```

```
lrtest(md_2)
```

```
## Likelihood ratio test
##
## Model 1: ph_adjusted ~ species
## Model 2: ph_adjusted ~ 1
##   #Df LogLik Df Chisq Pr(>Chisq)
## 1    7 -1200.5
## 2    2 -1224.1 -5 47.11  5.396e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
lsmeans_species_ph <- LSD.test(md_2, "species", p.adj="none" )
lsmeans_species_ph
```

```
## $statistics
##      MSerror Df      Mean      CV
## 1574.569 230 220.3091 18.01144
##
## $parameters
##      test p.adjusted name.t ntr alpha
## Fisher-LSD      none species 6 0.05
##
## $means
##      ph_adjusted      std      r      LCL      UCL      Min      Max      Q25      Q50
## cau      230.1269 35.25617 106 222.5329 237.7208 132.2 359.95 208.700 226.2
## cru      174.3667 30.13885  30 160.0922 188.6411 105.2 242.20 157.700 170.7
## hdus      229.4016 44.29120  31 215.3593 243.4439 107.2 289.20 196.325 239.2
## hybr      211.3094 63.21781  16 191.7633 230.8555  45.2 311.95 195.200 209.7
```

```
## hypo    229.0710 34.76474 31 215.0286 243.1133 129.2 318.20 217.700 230.2
## quit    217.0409 48.91085 22 200.3719 233.7099 98.2 283.20 195.950 223.2
##          Q75
## cau    250.2000
## cru    188.7625
## hdus   262.0750
## hybr   254.9500
## hypo   249.3250
## quit   255.7625
##
## $comparison
## NULL
##
## $groups
##      ph_adjusted groups
## cau    230.1269      a
## hdus    229.4016      a
## hypo    229.0710      a
## quit    217.0409      a
## hybr    211.3094      a
## cru     174.3667      b
##
## attr("class")
## [1] "group"
```

C. Seed setting

- It is a binomial trait, so I use binomial regression model.

```
data_species$seed_setting <- as.numeric(data_species$seed_setting)
```

```
data_species$seed_setting <- as.integer(data_species$seed_setting)
```

```
md_3 <- glm(seed_setting ~ species, data = data_species, family = binomial(link = "logit"))
summary(md_3)
```

```
##
## Call:
## glm(formula = seed_setting ~ species, family = binomial(link = "logit"),
##      data = data_species)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.06885  -0.45118  -0.00005  -0.00005   2.16119
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -20.57    1653.37  -0.012   0.990
## speciescru     22.58    1653.37   0.014   0.989
## specieshdus    18.33    1653.37   0.011   0.991
## specieshybr    18.62    1653.37   0.011   0.991
## specieshypo    19.79    1653.37   0.012   0.990
## speciesquit    18.67    1653.37   0.011   0.991
##
## (Dispersion parameter for binomial family taken to be 1)
```

```
##
## Null deviance: 249.13 on 253 degrees of freedom
## Residual deviance: 117.78 on 248 degrees of freedom
## AIC: 129.78
##
## Number of Fisher Scoring iterations: 19

lrtest(md_3)

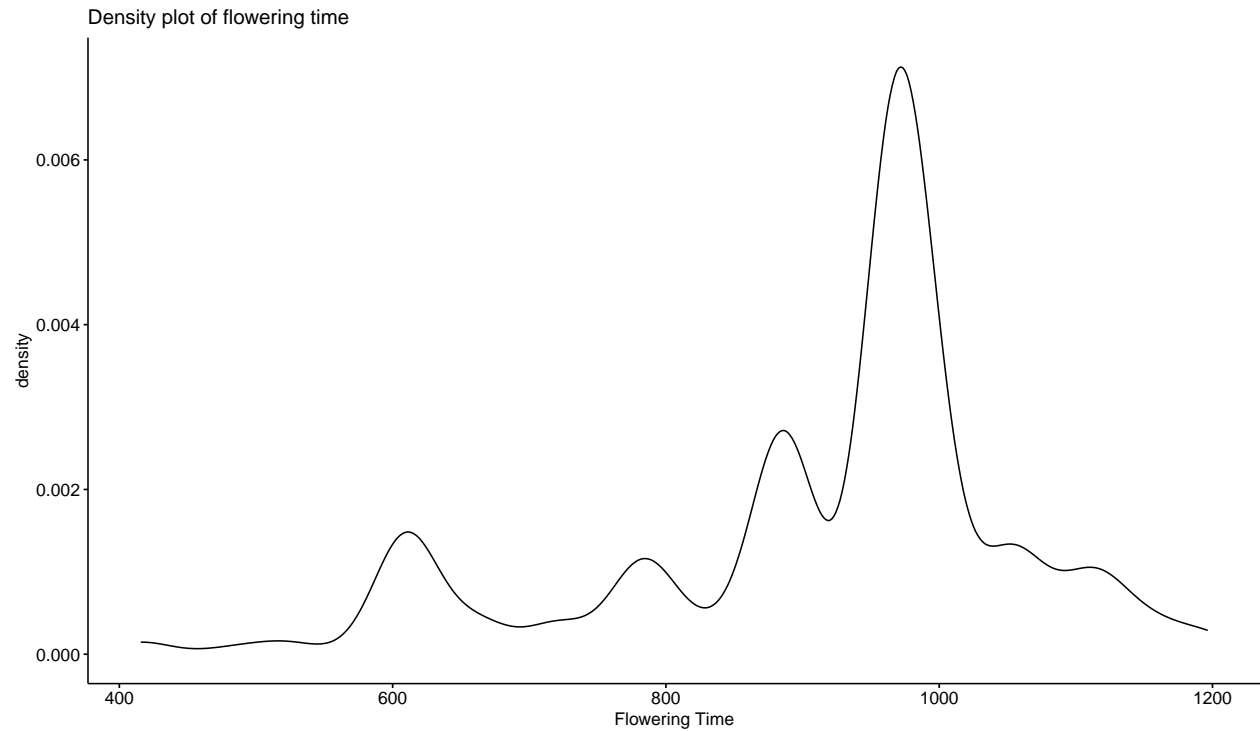
## Likelihood ratio test
##
## Model 1: seed_setting ~ species
## Model 2: seed_setting ~ 1
## #Df LogLik Df Chisq Pr(>Chisq)
## 1 6 -58.892
## 2 1 -124.567 -5 131.35 < 2.2e-16 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

lsmeans_species_ss <- LSD.test(md_3, "species", p.adj="none" )
lsmeans_species_ss

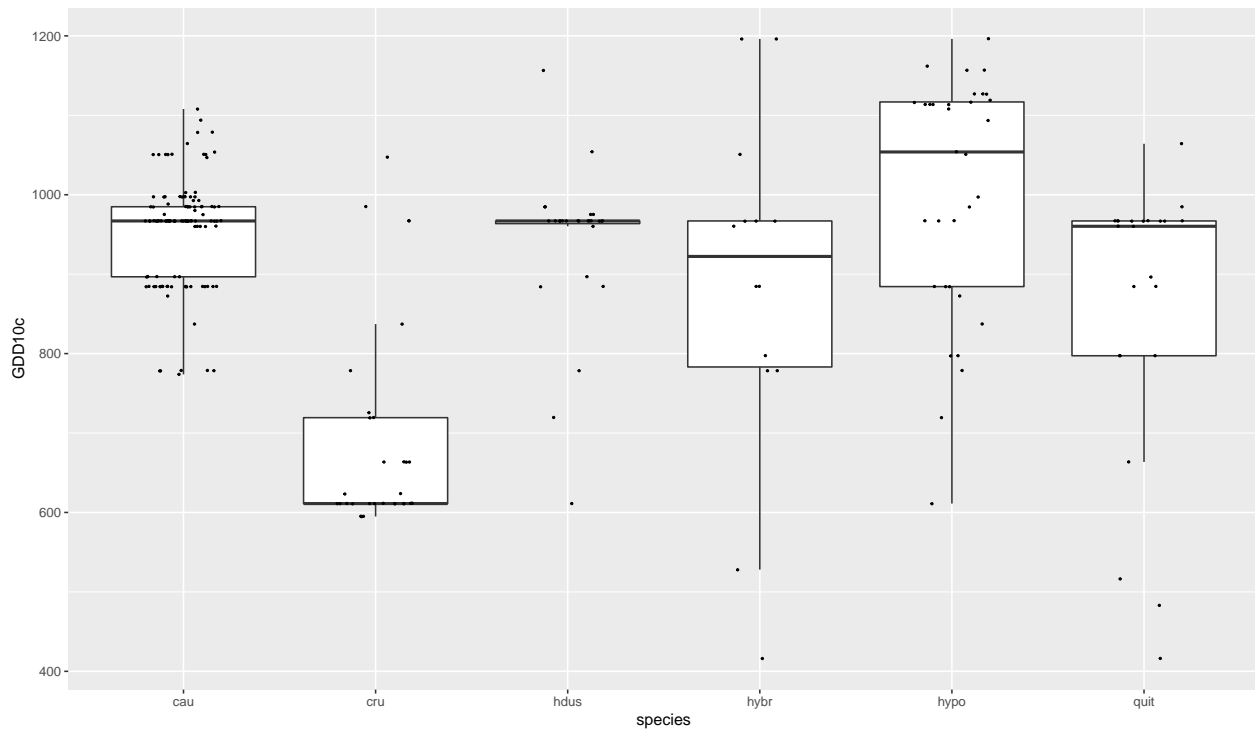
## $statistics
## MSerror Df Mean CV
## 0.4749391 248 0.1929134 357.2371
##
## $parameters
## test p.adjusted name.t ntr alpha
## Fisher-LSD none species 6 0.05
##
## $means
## seed_setting std r LCL UCL Min Max Q25 Q50 Q75
## cau 0.00000000 0.0000000 115 -0.1265735 0.1265735 0 0 0 0 0
## cru 0.88235294 0.3270350 34 0.6495695 1.1151364 0 1 1 1 1
## hdus 0.09677419 0.3005372 31 -0.1470129 0.3405613 0 1 0 0 0
## hybr 0.12500000 0.3415650 16 -0.2143373 0.4643373 0 1 0 0 0
## hypo 0.31428571 0.4710082 35 0.0848518 0.5437196 0 1 0 0 1
## quit 0.13043478 0.3443502 23 -0.1525921 0.4134617 0 1 0 0 0
##
## $comparison
## NULL
##
## $groups
## seed_setting groups
## cru 0.88235294 a
## hypo 0.31428571 b
## quit 0.13043478 bc
## hybr 0.12500000 bc
## hdus 0.09677419 bc
## cau 0.00000000 c
##
## attr(,"class")
## [1] "group"
```


D. Growing degree days (Tbase is 10°C)

```
ggdensity(data_species$GDD10c,  
  main = "Density plot of flowering time",  
  xlab = "Flowering Time")
```



```
ggplot(data = data_species, mapping = aes(x = species, y = GDD10c)) +  
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_4 <-glm(GDD10c ~ species , data = data_species, family = gaussian)
summary(md_4)
```

```
##
## Call:
## glm(formula = GDD10c ~ species, family = gaussian, data = data_species)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -467.61   -70.42    12.23    47.72   359.99
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    954.77     11.17  85.460 < 2e-16 ***
## speciescru   -267.61     23.86 -11.214 < 2e-16 ***
## specieshdus   -13.90     25.53  -0.544 0.586621
## specieshybr   -71.11     33.78  -2.105 0.036352 *
## specieshypo    48.66     23.58   2.064 0.040128 *
## speciesquit   -93.50     27.27  -3.429 0.000714 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 14229.2)
##
##      Null deviance: 5567350  on 242  degrees of freedom
## Residual deviance: 3372322  on 237  degrees of freedom
## (11 observations deleted due to missingness)
## AIC: 3021.4
##
## Number of Fisher Scoring iterations: 2
```

```
lrtest(md_4)
```

```
## Likelihood ratio test
##
## Model 1: GDD10c ~ species
## Model 2: GDD10c ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    7 -1503.7
## 2    2 -1564.6 -5 121.82 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
lsmeans_species_GDD10c <- LSD.test(md_4, "species", p.adj="none" )
lsmeans_species_GDD10c
```

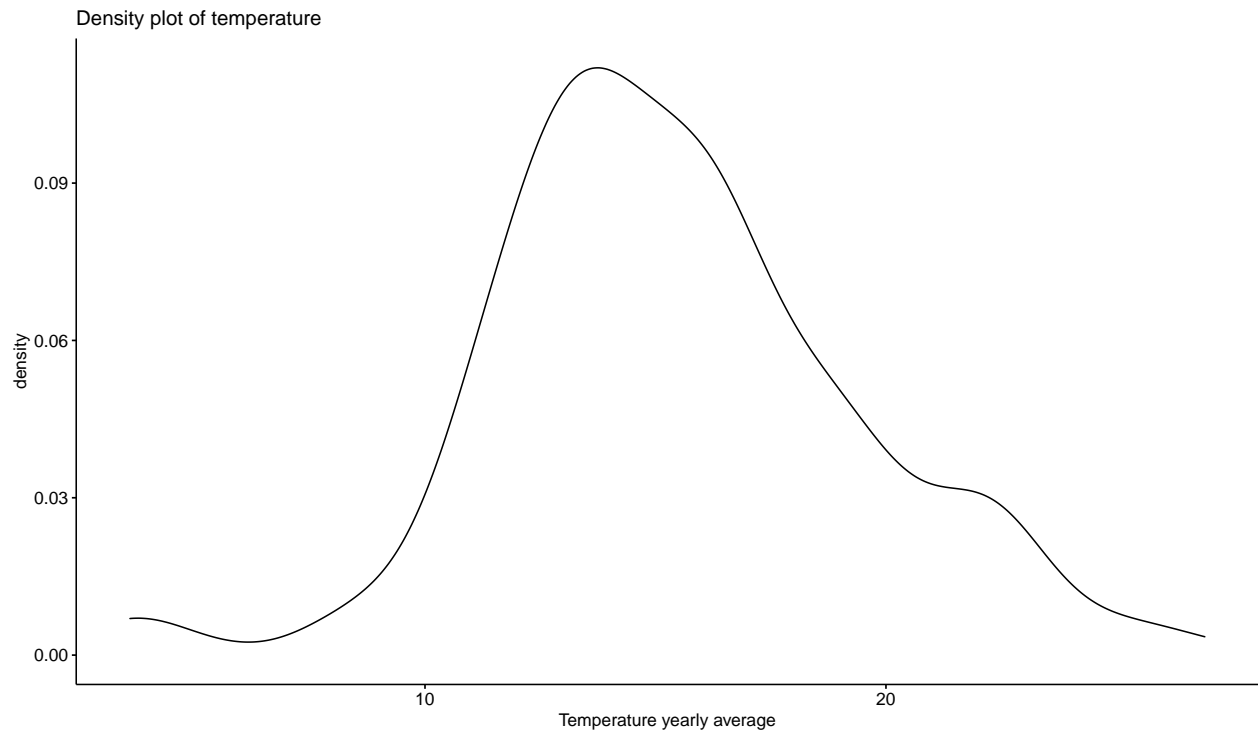
```
## $statistics
##   MSerror Df      Mean      CV
## 14229.2 237 911.6481 13.08468
##
## $parameters
##           test p.adjusted name.t ntr alpha
## Fisher-LSD      none species   6  0.05
##
## $means
##           GDD10c      std    r      LCL      UCL      Min      Max      Q25
## cau   954.7715  68.84005 114 932.7620  976.7810 773.80 1107.95 896.7000
## cru   687.1609 130.24622  32 645.6190  728.7029 594.90 1047.15 611.2000
## hdus   940.8704 101.52242  27 895.6452  986.0955 611.20 1156.65 963.6500
## hybr   883.6643 219.19917  14 820.8588  946.4698 416.05 1196.15 783.1625
## hypo 1003.4348 150.82466  33 962.5272 1044.3425 611.20 1196.15 884.3500
## quit   861.2717 177.14793  23 812.2715  910.2719 416.05 1064.15 797.3000
##           Q50      Q75
## cau   967.000  984.90
## cru   611.200  719.35
## hdus   967.000  967.00
## hybr   922.325  967.00
## hypo 1053.850 1116.80
## quit   960.300  967.00
##
## $comparison
## NULL
##
## $groups
##           GDD10c groups
## hypo 1003.4348      a
## cau   954.7715      b
## hdus   940.8704     bc
## hybr   883.6643     cd
## quit   861.2717      d
## cru   687.1609      e
##
## attr("class")
## [1] "group"
```

2. Climatic & Geographical Variables

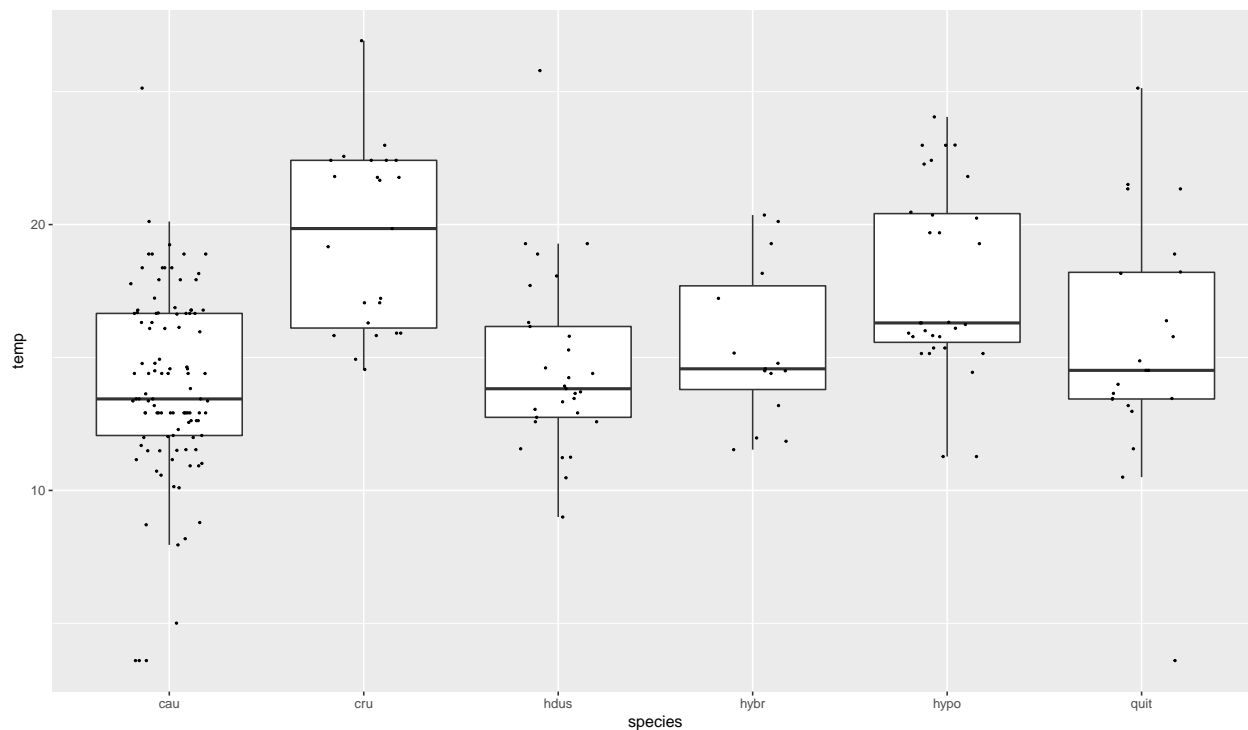
A. Temperature

Distribution of the data.

```
ggdensity(data_species$temp,  
  main = "Density plot of temperature",  
  xlab = "Temperature yearly average")
```



```
ggplot(data = data_species, mapping = aes(x = species, y = temp)) +  
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_5 <-glm(temp ~ species , data = data_species, family = gaussian)
summary(md_5)
```

```
##
## Call:
## glm(formula = temp ~ species, family = gaussian, data = data_species)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -11.8773  -2.0715  -0.6303   2.7193  11.1943
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  13.9391     0.3460  40.286 < 2e-16 ***
## speciescru    5.5732     0.8194   6.801 9.65e-11 ***
## specieshdus   0.7213     0.7465   0.966  0.3350
## specieshybr   1.5026     0.9827   1.529  0.1277
## specieshypo   3.8991     0.7274   5.360 2.09e-07 ***
## speciesquit   1.5382     0.8346   1.843  0.0667 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 12.69022)
##
##      Null deviance: 3607.3  on 225  degrees of freedom
## Residual deviance: 2791.8  on 220  degrees of freedom
## (28 observations deleted due to missingness)
## AIC: 1223.5
##
## Number of Fisher Scoring iterations: 2
```

```
lrtest(md_5)
```

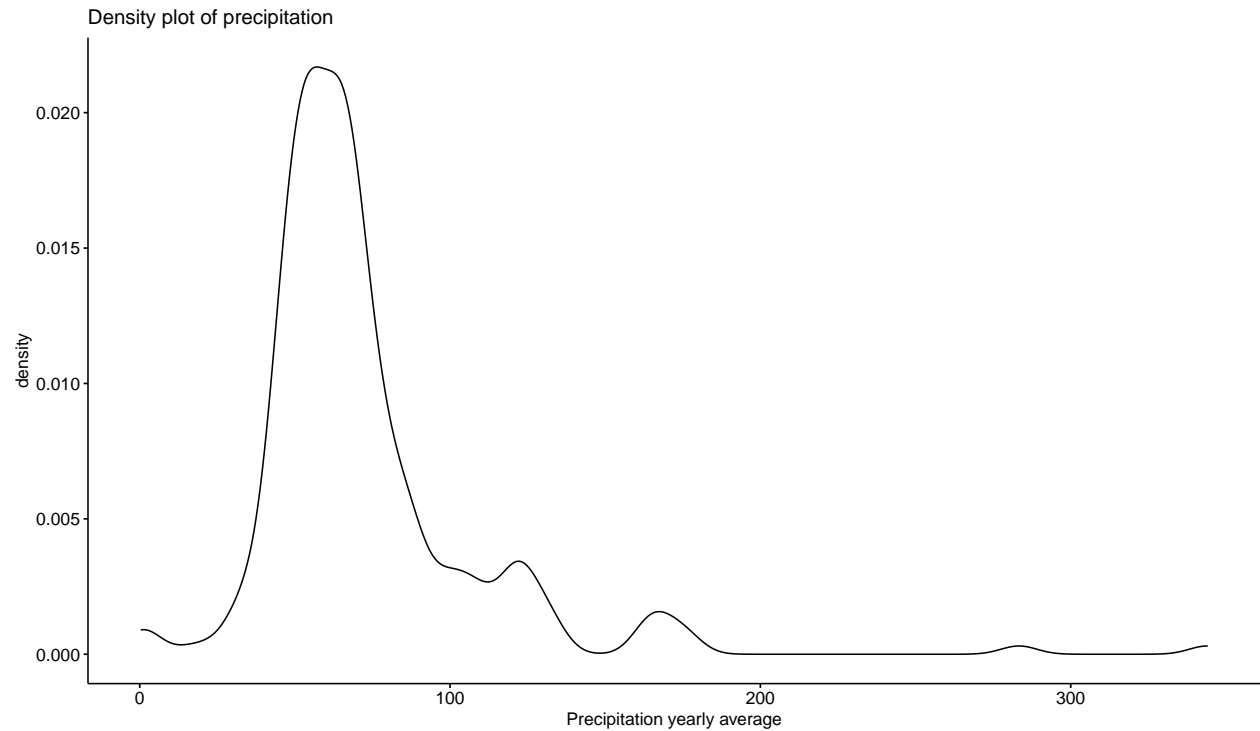
```
## Likelihood ratio test
##
## Model 1: temp ~ species
## Model 2: temp ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    7 -604.75
## 2    2 -633.71 -5 57.913  3.278e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
lsmeans_species_temp <- LSD.test(md_5, "species", p.adj="none" )
lsmeans_species_temp
```

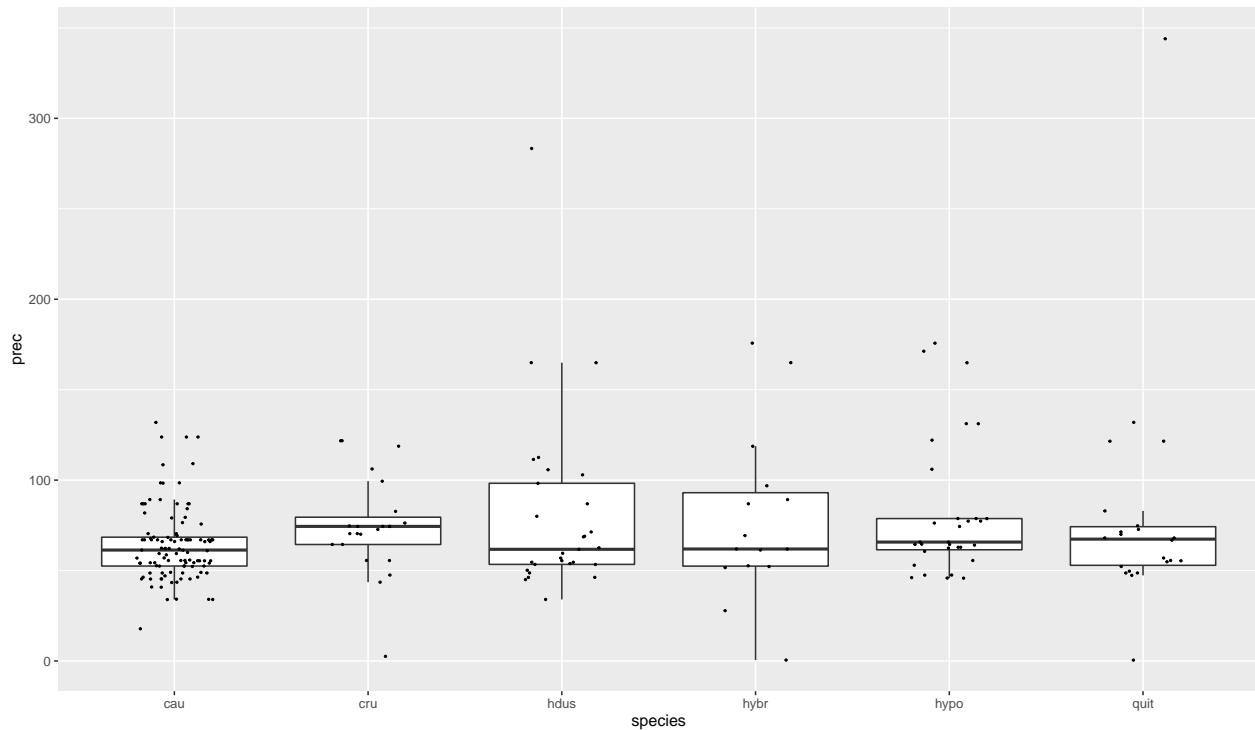
```
## $statistics
##      MSerror Df      Mean      CV
## 12.69022 220 15.38311 23.15743
##
## $parameters
##      test p.adjusted name.t ntr alpha
## Fisher-LSD      none species  6 0.05
##
## $means
##      temp      std    r      LCL      UCL      Min      Max      Q25
## cau 13.93907 3.505962 106 13.25716 14.62098  3.60000 25.13333 12.06667
## cru 19.51232 3.414973  23 18.04841 20.97623 14.55000 26.91667 16.10833
## hdus 14.66034 3.376270  29 13.35664 15.96405  9.00000 25.79167 12.75000
## hybr 15.44167 2.937050  15 13.62894 17.25440 11.53333 20.35833 13.79583
## hypo 17.83817 3.494242  31 16.57722 19.09912 11.27500 24.05000 15.57083
## quit 15.47727 4.569101  22 13.98046 16.97408  3.60000 25.13333 13.43958
##      Q50      Q75
## cau 13.44167 16.65833
## cru 19.85000 22.41667
## hdus 13.82500 16.16667
## hybr 14.57500 17.69583
## hypo 16.30000 20.41250
## quit 14.51667 18.20417
##
## $comparison
## NULL
##
## $groups
##      temp groups
## cru 19.51232    a
## hypo 17.83817    a
## quit 15.47727    b
## hybr 15.44167    b
## hdus 14.66034    b
## cau 13.93907    b
##
## attr(,"class")
## [1] "group"
```

B. Precipitation

```
ggdensity(data_species$prec,  
  main = "Density plot of precipitation",  
  xlab = "Precipitation yearly average")
```



```
ggplot(data = data_species, mapping = aes(x = species, y = prec)) +  
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_6 <-glm(prec ~ species , data = data_species, family = gaussian)
summary(md_6)
```

```
##
## Call:
## glm(formula = prec ~ species, family = gaussian, data = data_species)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -79.652  -19.013   -7.422    5.496  263.848
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    64.359     3.555  18.103  <2e-16 ***
## speciescru     10.271     8.420   1.220  0.2238
## specieshdus    16.888     7.671   2.202  0.0287 *
## specieshybr    13.763    10.098   1.363  0.1743
## specieshypo    17.031     7.474   2.279  0.0236 *
## speciesquit    15.792     8.576   1.842  0.0669 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 1339.809)
##
##      Null deviance: 308314  on 225  degrees of freedom
## Residual deviance: 294758  on 220  degrees of freedom
## (28 observations deleted due to missingness)
## AIC: 2276.5
##
## Number of Fisher Scoring iterations: 2
```



```
lrtest(md_6)
```

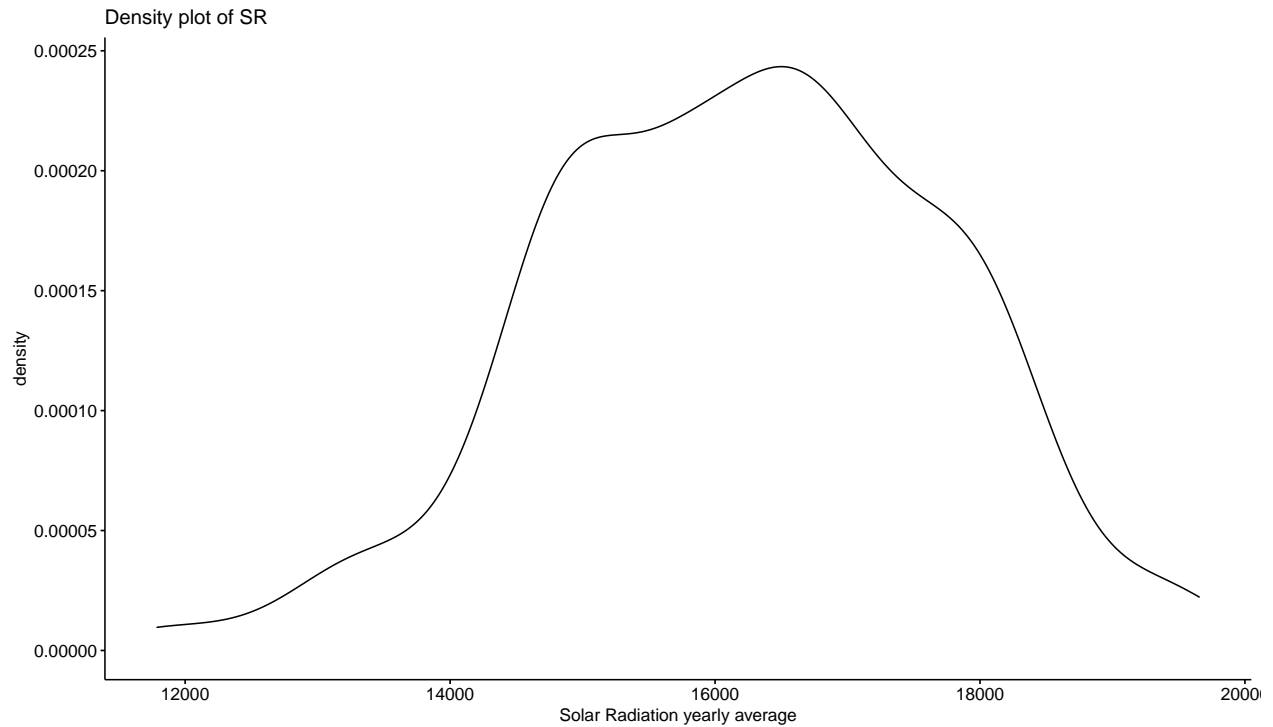
```
## Likelihood ratio test
##
## Model 1: prec ~ species
## Model 2: prec ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    7 -1131.3
## 2    2 -1136.3 -5 10.162    0.07079 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
lsmeans_species_prec <- LSD.test(md_6, "species", p.adj="none" )
lsmeans_species_prec
```

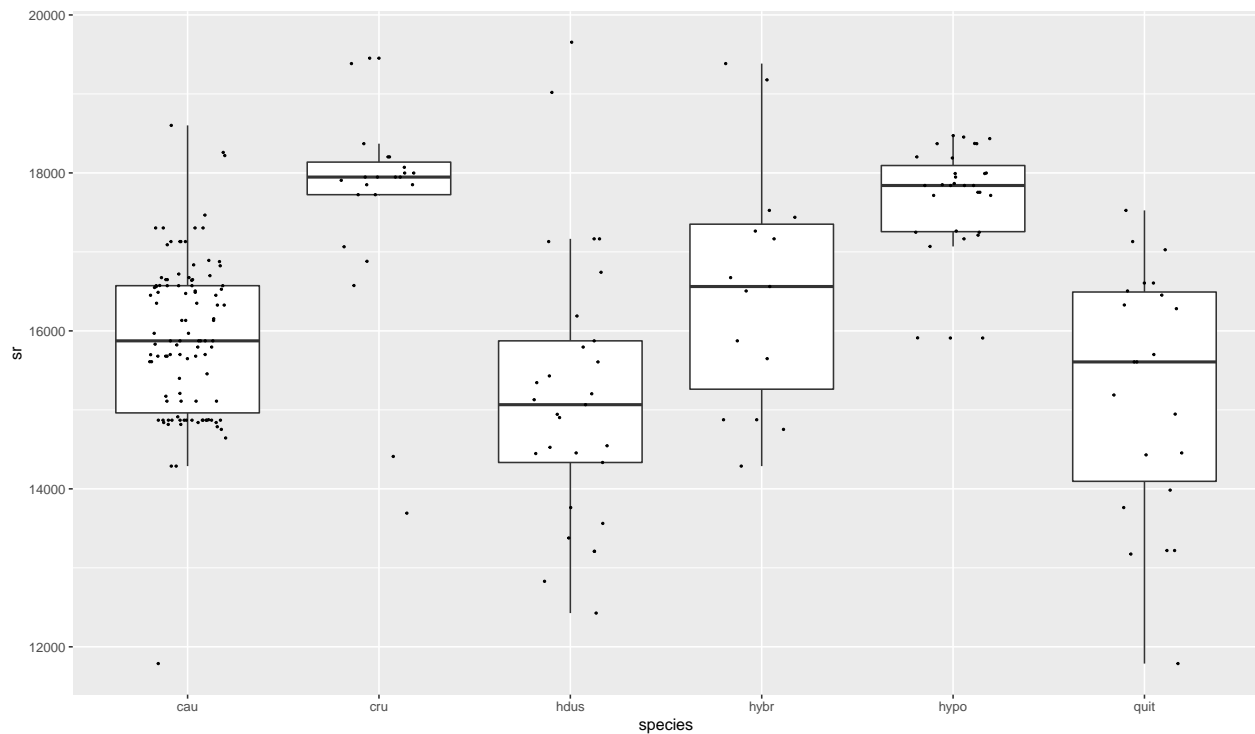
```
## $statistics
##      MSerror Df      Mean      CV
## 1339.809 220 72.35841 50.58625
##
## $parameters
##      test p.adjusted name.t ntr alpha
## Fisher-LSD      none species 6 0.05
##
## $means
##      prec      std  r      LCL      UCL      Min      Max      Q25
## cau 64.35928 20.33200 106 57.35260 71.36596 17.833333 131.9167 52.50000
## cru 74.63043 26.92531 23 59.58858 89.67229 2.583333 121.7500 64.41667
## hdus 81.24713 50.96899 29 67.85140 94.64285 34.000000 283.3333 53.41667
## hybr 78.12222 46.84483 15 59.49622 96.74822 0.500000 175.7500 52.45833
## hypo 81.38978 37.40327 31 68.43338 94.34619 45.833333 175.7500 61.50000
## quit 80.15152 65.45480 22 64.77160 95.53143 0.500000 344.0000 52.89583
##      Q50      Q75
## cau 61.33333 68.45833
## cru 74.41667 79.50000
## hdus 61.75000 98.25000
## hybr 61.91667 93.04167
## hypo 65.75000 78.75000
## quit 67.37500 74.25000
##
## $comparison
## NULL
##
## $groups
##      prec groups
## hypo 81.38978 a
## hdus 81.24713 a
## quit 80.15152 ab
## hybr 78.12222 ab
## cru 74.63043 ab
## cau 64.35928 b
##
## attr(,"class")
## [1] "group"
```

C. Solar Radiation

```
ggdensity(data_species$sr,  
  main = "Density plot of SR",  
  xlab = "Solar Radiation yearly average")
```



```
ggplot(data = data_species, mapping = aes(x = species, y = sr)) +  
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_7 <-glm(sr ~ species , data = data_species, family = gaussian)
summary(md_7)
```

```
##
## Call:
## glm(formula = sr ~ species, family = gaussian, data = data_species)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -4128.6   -803.9    110.9    663.5   4446.4
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  15916.4     118.3  134.599 < 2e-16 ***
## speciescru    1762.0     280.0   6.292 1.68e-09 ***
## specieshdus   -708.0     255.1  -2.775 0.00599 **
## specieshybr    617.5     335.9   1.839 0.06732 .
## specieshypo   1753.1     248.6   7.052 2.25e-11 ***
## speciesquit   -664.5     285.2  -2.330 0.02074 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 1482219)
##
##      Null deviance: 501678361  on 225  degrees of freedom
## Residual deviance: 326088184  on 220  degrees of freedom
## (28 observations deleted due to missingness)
## AIC: 3860.5
##
## Number of Fisher Scoring iterations: 2
```

```
lrtest(md_7)
```

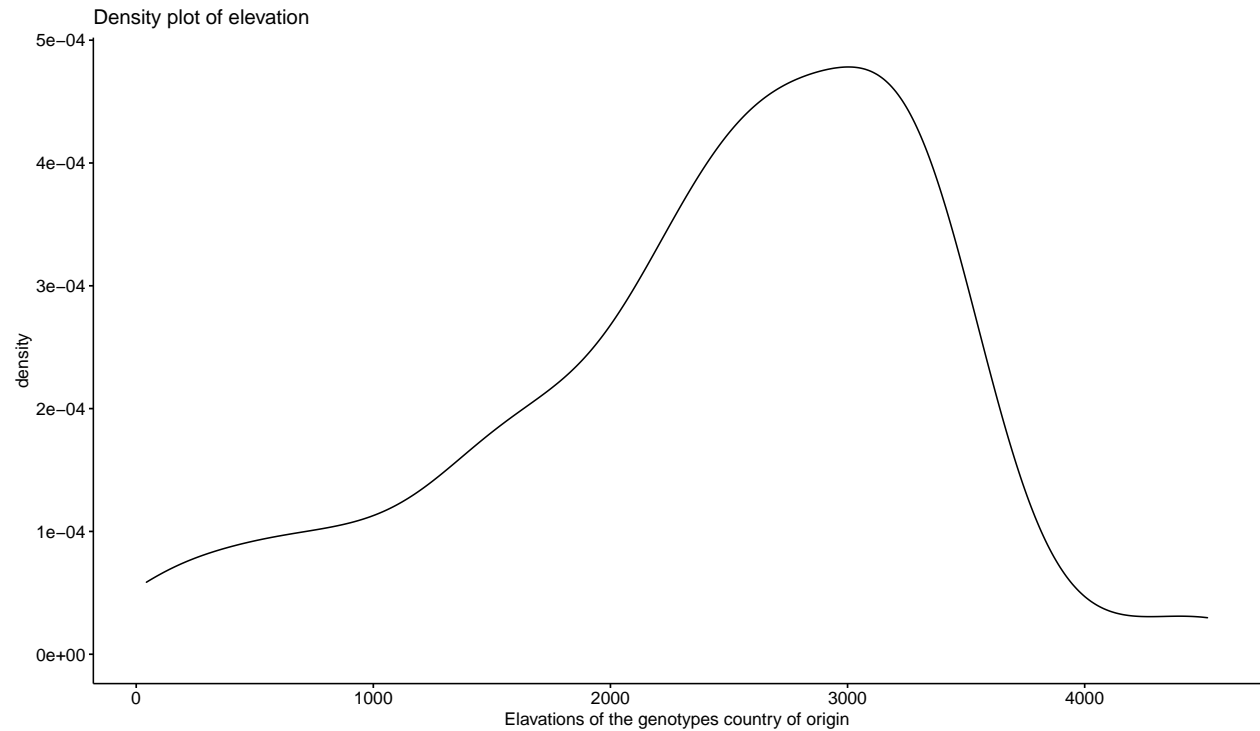
```
## Likelihood ratio test
##
## Model 1: sr ~ species
## Model 2: sr ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    7 -1923.3
## 2    2 -1971.9 -5  97.359  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
lsmeans_species_sr <- LSD.test(md_7, "species", p.adj="none" )
lsmeans_species_sr
```

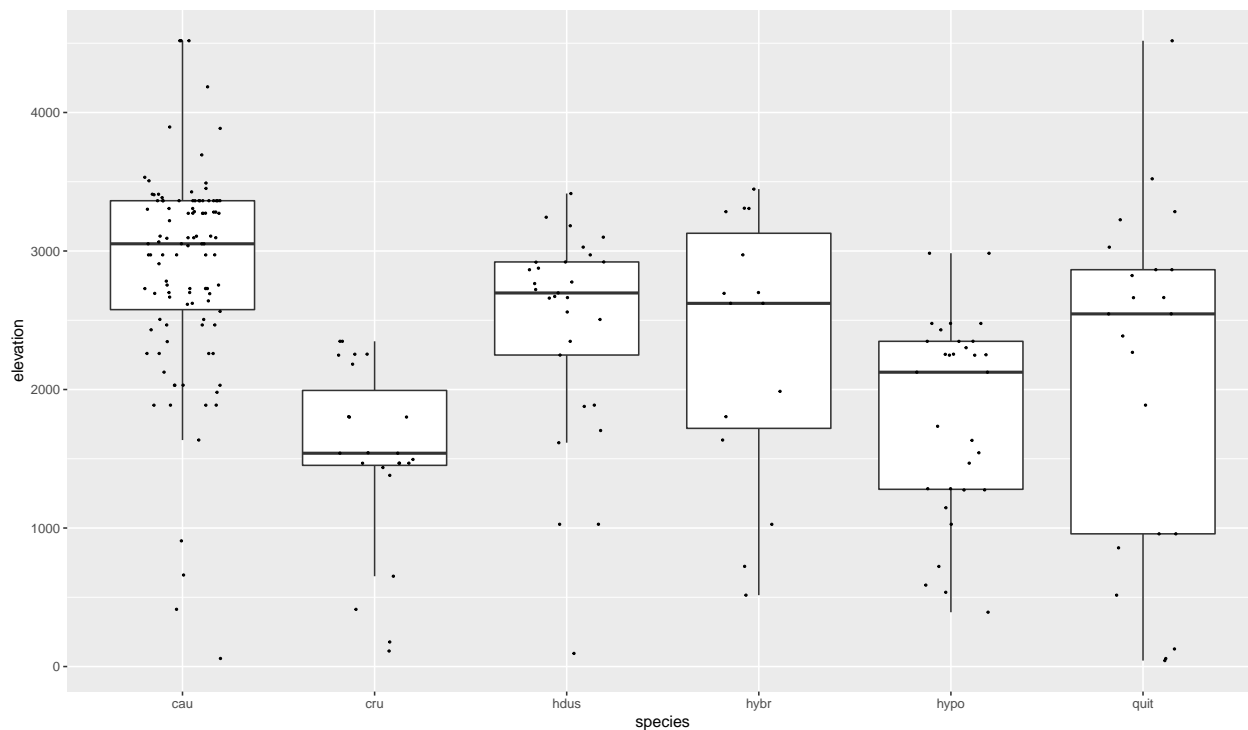
```
## $statistics
##   MSerror Df      Mean      CV
## 1482219 220 16221.64 7.505185
##
## $parameters
##           test p.adjusted name.t ntr alpha
## Fisher-LSD      none species   6 0.05
##
## $means
##           sr      std   r      LCL      UCL      Min      Max      Q25
## cau 15916.41 999.2097 106 15683.36 16149.46 11787.83 18601.75 14961.58
## cru 17678.38 1345.0603  23 17178.08 18178.69 13692.25 19452.83 17724.42
## hdus 15208.39 1716.7614  29 14762.83 15653.94 12427.17 19654.75 14333.83
## hybr 16533.92 1530.9086  15 15914.40 17153.44 14288.00 19384.33 15261.62
## hypo 17669.48  709.5543  31 17238.53 18100.42 15909.75 18472.50 17256.25
## quit 15251.95 1558.5838  22 14740.40 15763.51 11787.83 17525.42 14095.31
##           Q50      Q75
## cau 15873.75 16572.17
## cru 17947.42 18136.88
## hdus 15065.50 15873.75
## hybr 16561.75 17350.92
## hypo 17840.83 18093.79
## quit 15607.00 16492.19
##
## $comparison
## NULL
##
## $groups
##           sr groups
## cru 17678.38      a
## hypo 17669.48      a
## hybr 16533.92      b
## cau 15916.41      b
## quit 15251.95      c
## hdus 15208.39      c
##
## attr(,"class")
## [1] "group"
```

D. Elevation

```
ggdensity(data_species$elevation,  
  main = "Density plot of elevation",  
  xlab = "Elavations of the genotypes country of origin")
```



```
ggplot(data = data_species, mapping = aes(x = species, y = elevation)) +  
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_8 <-glm(elevation ~ species , data = data_species, family = gaussian)
summary(md_8)
```

```
##
## Call:
## glm(formula = elevation ~ species, family = gaussian, data = data_species)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2830.6   -412.6    206.5    474.4   2399.6
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2888.58     78.86  36.628 < 2e-16 ***
## speciescru  -1357.97    186.77  -7.271 6.18e-12 ***
## specieshdus  -430.09    170.15  -2.528  0.0122 *
## specieshybr  -578.71    223.98  -2.584  0.0104 *
## specieshypo -1063.09    165.79  -6.412 8.63e-10 ***
## speciesquit  -770.17    190.22  -4.049 7.14e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 659244.6)
##
##      Null deviance: 199626032  on 225  degrees of freedom
## Residual deviance: 145033821  on 220  degrees of freedom
## (28 observations deleted due to missingness)
## AIC: 3677.4
##
## Number of Fisher Scoring iterations: 2
```

```

lrtest(md_8)

## Likelihood ratio test
##
## Model 1: elevation ~ species
## Model 2: elevation ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    7 -1831.7
## 2    2 -1867.8 -5 72.202  3.565e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

lsmeans_species_elevation <- LSD.test(md_8, "species", p.adj="none" )
lsmeans_species_elevation

## $statistics
##      MSerror Df      Mean      CV
## 659244.6 220 2435.982 33.33106
##
## $parameters
##      test p.adjusted name.t ntr alpha
## Fisher-LSD      none species 6 0.05
##
## $means
##      elevation      std    r      LCL      UCL Min  Max      Q25  Q50      Q75
## cau  2888.575  738.9232 106 2733.153 3043.998  58 4518 2576.75 3052 3363.0
## cru  1530.609  655.7926  23 1196.949 1864.268 112 2348 1452.50 1540 1993.5
## hdus 2458.483  760.1605  29 2161.338 2755.628  95 3415 2249.00 2697 2921.0
## hybr 2309.867  977.1904  15 1896.704 2723.030 515 3447 1719.50 2622 3128.0
## hypo 1825.484  717.9926  31 1538.084 2112.884 392 2984 1279.50 2125 2348.0
## quit 2118.409 1257.8891  22 1777.251 2459.567  43 4518  958.00 2546 2865.0
##
## $comparison
## NULL
##
## $groups
##      elevation groups
## cau  2888.575      a
## hdus 2458.483      b
## hybr 2309.867     bc
## quit 2118.409     bc
## hypo 1825.484     cd
## cru  1530.609      d
##
## attr(,"class")
## [1] "group"

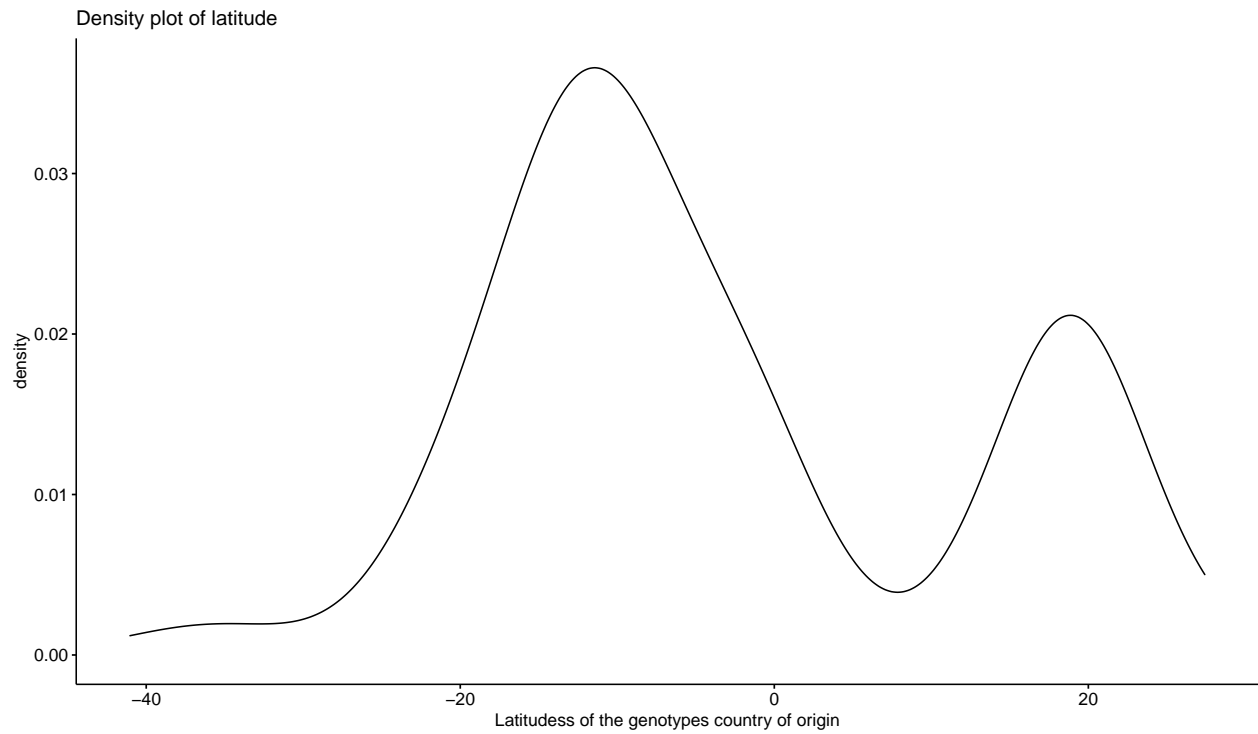
```

E. Latitude

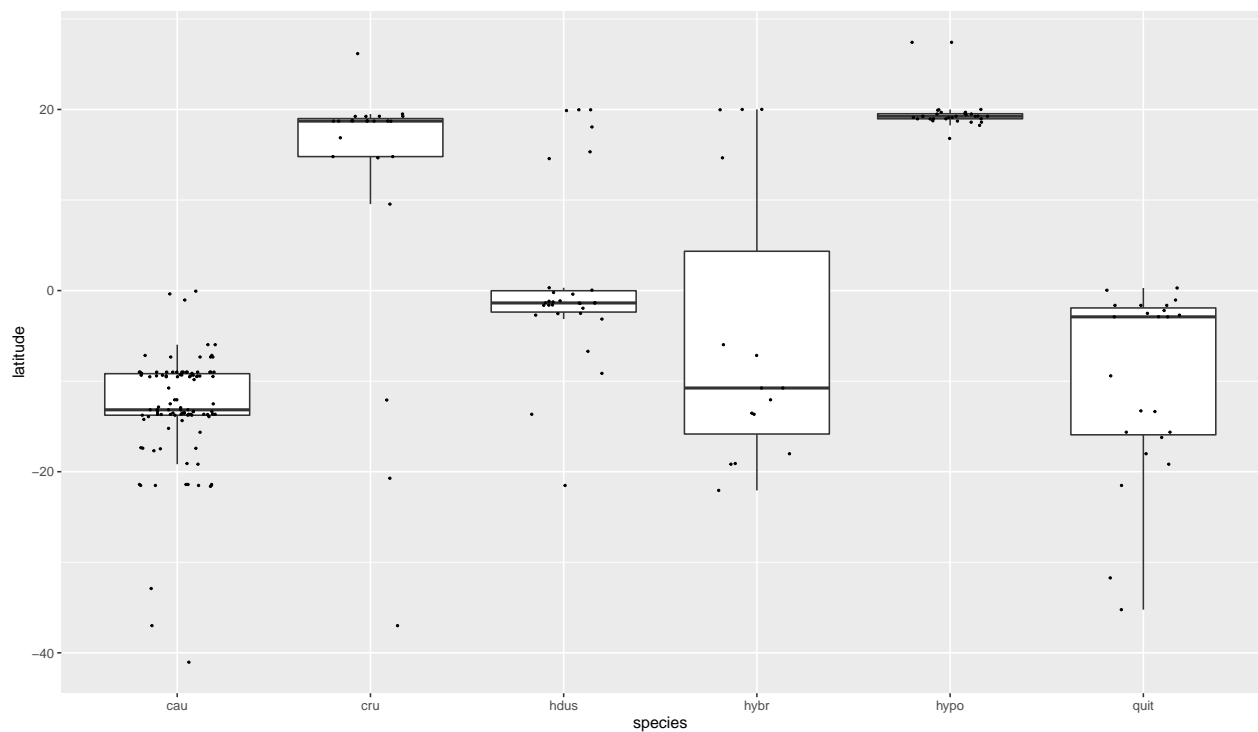
```

ggdensity(data_species$latitude,
  main = "Density plot of latitude",
  xlab = "Latitudess of the genotypes country of origin")

```



```
ggplot(data = data_species, mapping = aes(x = species, y = latitude)) +  
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_9 <-glm(latitude ~ species , data = data_species, family = gaussian)  
summary(md_9)
```

```
##  
## Call:
```



```
## glm(formula = latitude ~ species, family = gaussian, data = data_species)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -49.698  -2.332  -0.409   3.931  25.187
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -12.9305     0.8611 -15.016 < 2e-16 ***
## speciescru   25.6286     2.0393  12.567 < 2e-16 ***
## specieshdus  13.8769     1.8334   7.569 9.97e-13 ***
## specieshybr   7.7557     2.4457   3.171 0.00173 **
## specieshypo  32.5897     1.8103  18.003 < 2e-16 ***
## speciesquit   2.8905     2.0393   1.417 0.15778
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 78.60063)
##
##      Null deviance: 51146  on 227  degrees of freedom
## Residual deviance: 17449  on 222  degrees of freedom
## (26 observations deleted due to missingness)
## AIC: 1650
##
## Number of Fisher Scoring iterations: 2
```

```
lrtest(md_9)
```

```
## Likelihood ratio test
##
## Model 1: latitude ~ species
## Model 2: latitude ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    7 -818.02
## 2    2 -940.61 -5 245.19 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

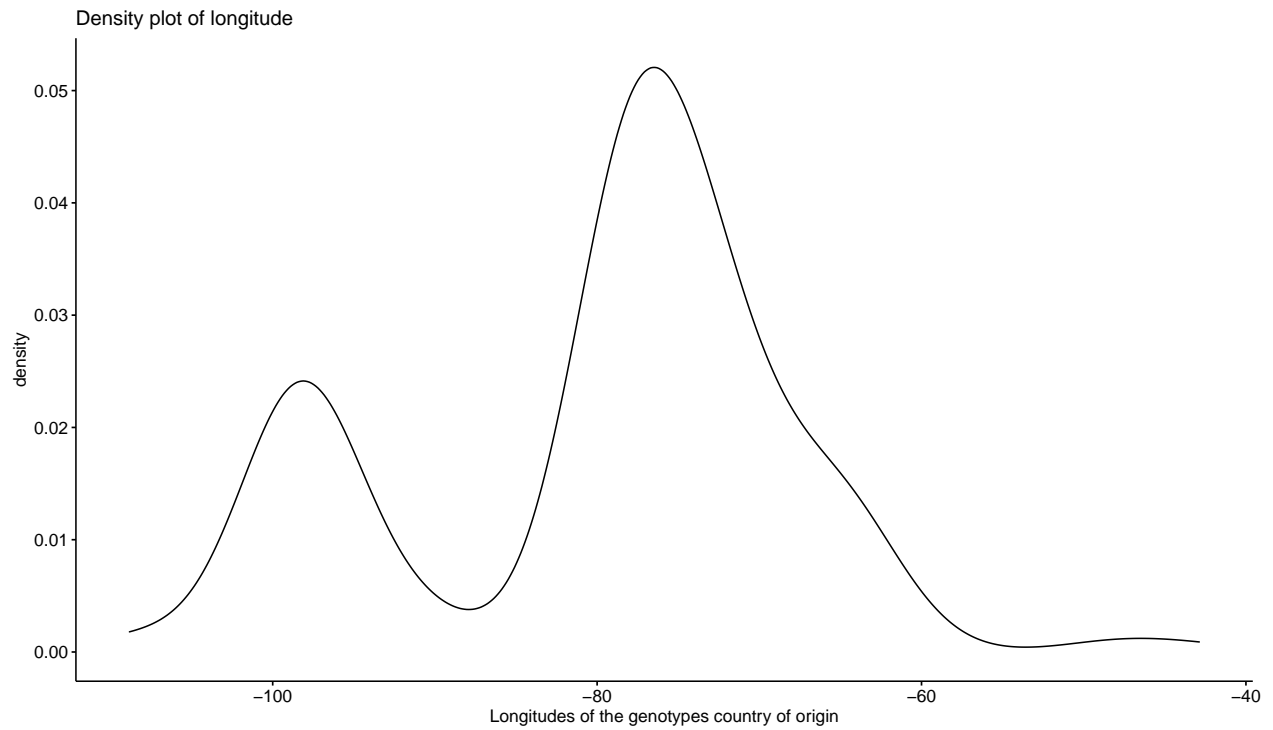
```
lsmeans_species_latitude <- LSD.test(md_9, "species", p.adj="none" )
lsmeans_species_latitude
```

```
## $statistics
##      MSerror Df      Mean      CV
##    78.60063 222 -3.286399 -269.7694
##
## $parameters
##      test p.adjusted name.t ntr alpha
## Fisher-LSD      none species    6 0.05
##
## $means
##      latitude      std    r      LCL      UCL      Min      Max
## cau -12.9305355  6.053224 106 -14.627537 -11.2335339 -41.03454 -0.063325
## cru  12.6980219 15.025957  23  9.054921  16.3411232 -36.99958 26.166667
## hdus  0.9463838  9.778241  30 -2.243498  4.1362652 -21.51777 19.966667
## hybr -5.1748345 15.560844  15 -9.686008 -0.6636609 -22.06622 20.012118
```

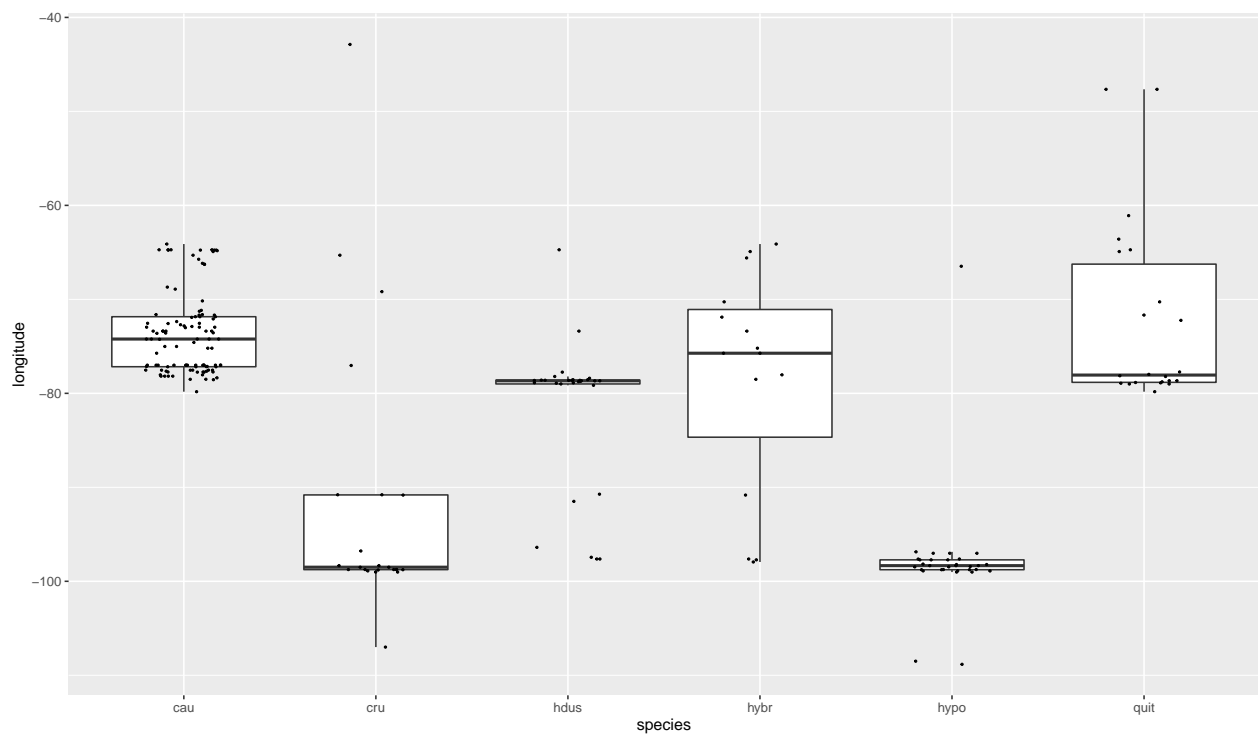
```
## hypo 19.6591845 2.154195 31 16.521175 22.7971945 16.80000 27.416667
## quit -10.0400192 10.297513 23 -13.683121 -6.3969179 -35.23390 0.287972
##          Q25          Q50          Q75
## cau -13.761848 -13.164547 -9.17521950
## cru 14.796239 18.716667 18.99452400
## hdus -2.371932 -1.364866 -0.01530725
## hybr -15.838098 -10.759094 4.34273750
## hypo 18.966667 19.250000 19.52897500
## quit -15.925000 -2.899469 -1.91589850
##
## $comparison
## NULL
##
## $groups
##      latitude groups
## hypo 19.6591845    a
## cru 12.6980219    b
## hdus 0.9463838    c
## hybr -5.1748345    d
## quit -10.0400192   de
## cau -12.9305355    e
##
## attr("class")
## [1] "group"
```

F. Longitude

```
ggdensity(data_species$longitude,
          main = "Density plot of longitude",
          xlab = "Longitudes of the genotypes country of origin")
```



```
ggplot(data = data_species, mapping = aes(x = species, y = longitude)) +
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_10 <-glm(longitude ~ species , data = data_species, family = gaussian)
summary(md_10)
```

```
##
## Call:
```

```
## glm(formula = longitude ~ species, family = gaussian, data = data_species)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -19.447   -4.046   -0.416    2.625   48.972
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -73.6599     0.7673  -95.994 < 2e-16 ***
## speciescru  -18.1875     1.8173  -10.008 < 2e-16 ***
## specieshdus  -7.7483     1.6556   -4.680 5.01e-06 ***
## specieshybr  -4.8382     2.1794   -2.220 0.0274 *
## specieshypo -24.1986     1.6131  -15.001 < 2e-16 ***
## speciesquit  1.5494     1.8509    0.837 0.4034
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 62.414)
##
##      Null deviance: 32568  on 225  degrees of freedom
## Residual deviance: 13731  on 220  degrees of freedom
## (28 observations deleted due to missingness)
## AIC: 1583.5
##
## Number of Fisher Scoring iterations: 2
```

```
lrtest(md_10)
```

```
## Likelihood ratio test
##
## Model 1: longitude ~ species
## Model 2: longitude ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    7 -784.76
## 2    2 -882.35 -5 195.19 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
lsmeans_species_longitude <- LSD.test(md_10, "species", p.adj="none" )
lsmeans_species_longitude
```

```
## $statistics
##      MSerror Df      Mean      CV
##      62.414 220 -79.99466 -9.875977
##
## $parameters
##      test p.adjusted name.t ntr alpha
## Fisher-LSD      none species 6 0.05
##
## $means
##      longitude      std  r      LCL      UCL      Min      Max
## cau  -73.65988  4.356625 106  -75.17216 -72.14760 -79.82561 -64.11433
## cru  -91.84738 14.698078  23  -95.09392 -88.60084 -107.00000 -42.87518
## hdus  -81.40820  7.796249  29  -84.29945 -78.51695  -97.62778 -64.72605
## hybr  -78.49813 11.914513  15  -82.51825 -74.47801  -97.94471 -64.11433
```

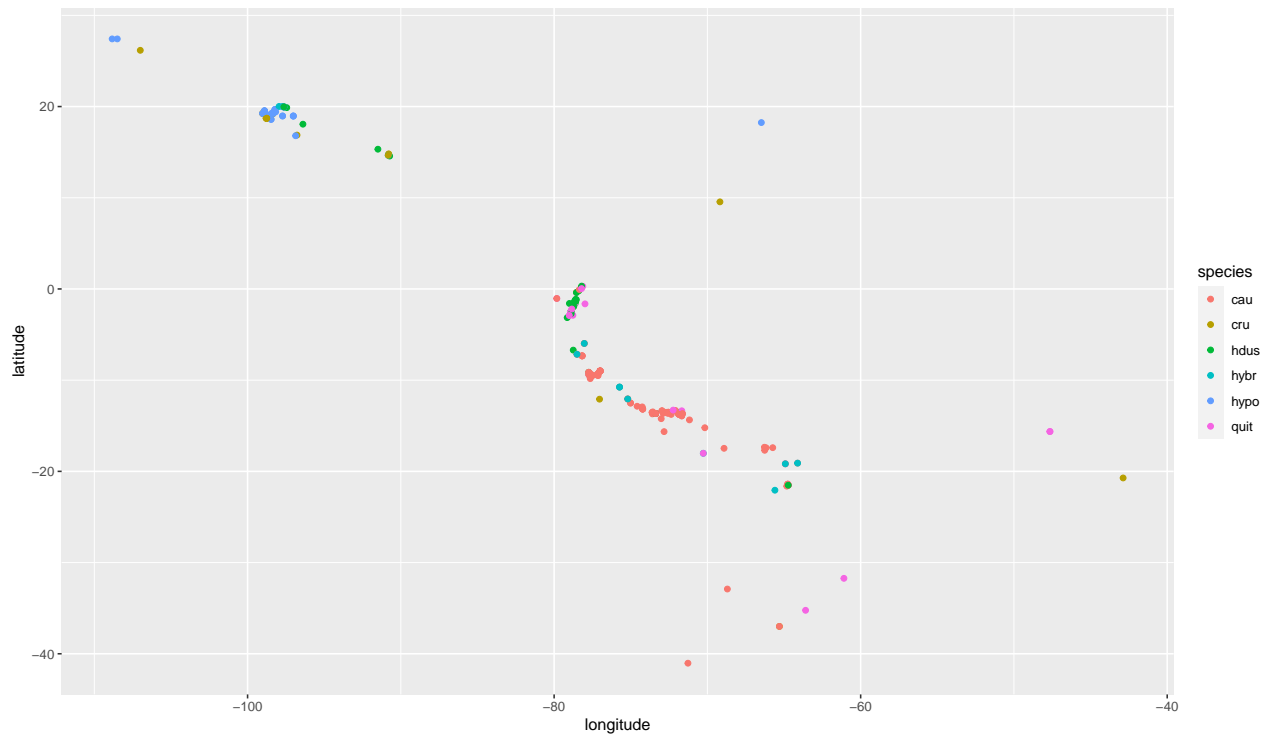
```
## hypo -97.85852 6.412741 31 -100.65495 -95.06209 -108.83333 -66.47890
## quit -72.11049 9.910231 22 -75.43000 -68.79099 -79.82561 -47.65000
##           Q25           Q50           Q75
## cau  -77.16731 -74.22260 -71.84630
## cru  -98.75833 -98.50000 -90.80932
## hdus -79.00128 -78.66724 -78.59322
## hybr -84.66627 -75.73134 -71.07873
## hypo -98.76667 -98.33414 -97.71667
## quit -78.82909 -78.05610 -66.25272
##
## $comparison
## NULL
##
## $groups
##      longitude groups
## quit -72.11049      a
## cau  -73.65988      a
## hybr -78.49813      b
## hdus -81.40820      b
## cru  -91.84738      c
## hypo -97.85852      d
##
## attr("class")
## [1] "group"
```

- It should be determined how to divide the accessions into different latitudinal groups.

```
prelim_plot <- ggplot(data, aes(x = longitude, y = latitude, colour = species)) +
  geom_point()
```

```
prelim_plot
```

```
## Warning: Removed 28 rows containing missing values (geom_point).
```



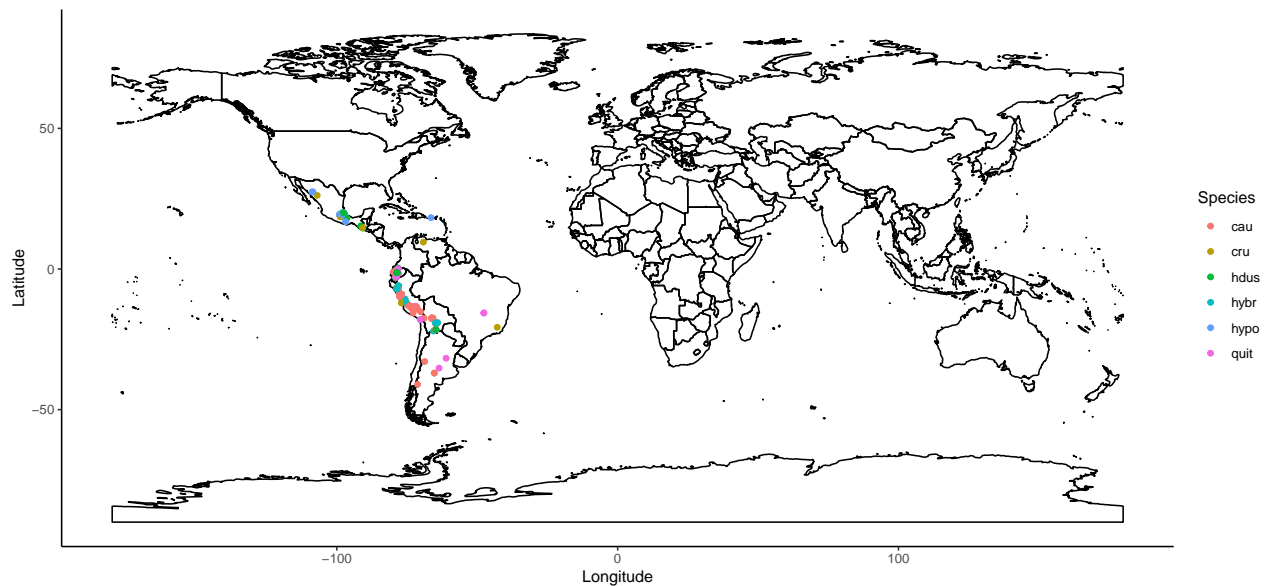
```
world <- getMap(resolution = "low")

with_world <- ggplot() +
  geom_polygon(data = world,
    aes(x = long, y = lat, group = group),
    fill = NA, colour = "black") +
  geom_point(data = data, # Add and plot species data
    aes(x = longitude, y = latitude,
      colour = species)) +
  coord_quickmap() + # Prevents stretching when resizing
  theme_classic() + # Remove ugly grey background
  xlab("Longitude") +
  ylab("Latitude") +
  guides(colour=guide_legend(title="Species"))
```

```
## Regions defined for each Polygons
```

```
with_world
```

```
## Warning: Removed 28 rows containing missing values (geom_point).
```



I need a smaller map, focused on the Latin America.

```
mapa <- borders("world", regions = c("Brazil", "Uruguay", "Argentina", "French Guiana", "Suriname", "Colombia",
  "Bolivia", "Ecuador", "Chile", "Paraguay", "Peru", "Guyana", "Panama",
  "Nicaragua", "Honduras", "El Salvador", "Belize", "Guatemala", "Mexico",
  "Caribe", "Puerto Rico", "Dominican Republic", "Haiti", "Jamaica",
  "Dominica", "Saba", "United States of America"),
  fill = "grey70", colour = "black")
```

Filtering the accessions without missing coordinate values

```
data_complete_coordinates <- filter(data, !(longitude=="NA" | latitude=="NA"))
dim(data_complete_coordinates)
```

```
## [1] 226 14
```

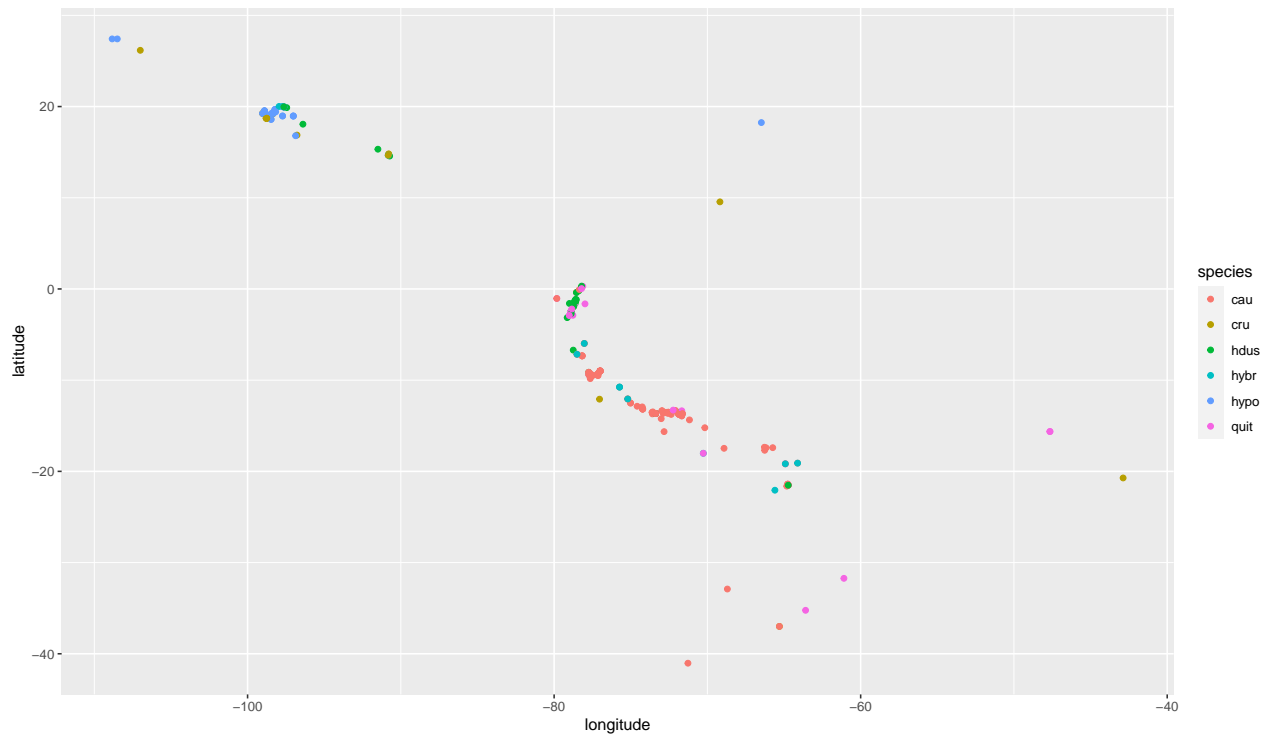
```
#data_complete_coordinates
sum(is.na(data_complete_coordinates$longitude))
```

```
## [1] 0
```

226 genotypes will be shown on the map

First let's look at the coordinates

```
ggplot(data_complete_coordinates, aes(x = longitude, y = latitude, colour = species)) +
  geom_point()
```



```
table(data_complete_coordinates$country)
```

```
##
```

```
## ARG BOL BRA ECU GTM MEX PER PTR VEN
```

```
##   6  21   3  33   6  53 102   1   1
```

```
map1 <- ggplot() + mapa + theme_bw() + xlab("Longitude (decimals)") + ylab("Latitude (decimals)") +
  theme(panel.border = element_blank(), panel.grid.major = element_line(colour = "grey80"), panel
  geom_point(data = data, # Add and plot species data
    aes(x = longitude, y = latitude,
        colour = species)) +
  coord_quickmap() + # Prevents stretching when resizing
  theme_bw() + # Remove ugly grey background
  xlab("Longitude") +
  ylab("Latitude") +
  guides(colour=guide_legend(title="Species"))+
  theme(legend.position="none")
```

```
map2<- ggplot() + mapa + theme_bw() + xlab("Longitude (decimals)") + ylab("Latitude (decimals)") +
  theme(panel.border = element_blank(), panel.grid.major = element_line(colour = "grey80"), panel
  geom_point(data = data, # Add and plot species data
    aes(x = longitude, y = latitude), color="blue") +
  coord_quickmap() + # Prevents stretching when resizing
  #theme_classic() + # Remove ugly grey background
  theme_bw() +
  xlab("Longitude") +
  ylab("Latitude")
```

Filtering the South American accessions.


```
filter(data, country=="VEN")
```

```
##   genotypes species country plot seed_setting ft_adjusted ph_adjusted longitude
## 1 cru_72_n    cru      VEN 508           1          110          NA -69.18047
##   latitude elevation      prec      temp      sr GDD10c
## 1 9.550365      177 99.41667 26.91667 16881      967
```

```
south_american_genotypes <-filter(data_complete_coordinates, latitude < 9.56 )
dim(south_american_genotypes)
```

```
## [1] 166 14
```

```
summary(south_american_genotypes)
```

```
##   genotypes      species      country      plot
## Length:166      Length:166      Length:166      Min.   : 102.0
## Class :character Class :character Class :character 1st Qu.: 319.2
## Mode  :character Mode  :character Mode  :character Median : 602.5
##                                           Mean  : 566.4
##                                           3rd Qu.: 818.5
##                                           Max.   :1028.0
##
##   seed_setting      ft_adjusted      ph_adjusted      longitude
## Min.   :0.00000      Min.   : 50.0      Min.   :107.2      Min.   : -79.83
## 1st Qu.:0.00000      1st Qu.:102.0      1st Qu.:204.2      1st Qu.: -78.03
## Median :0.00000      Median :110.0      Median :231.2      Median : -75.73
## Mean   :0.04217      Mean   :108.3      Mean   :229.6      Mean   : -73.68
## 3rd Qu.:0.00000      3rd Qu.:113.0      3rd Qu.:257.9      3rd Qu.: -71.85
## Max.   :1.00000      Max.   :160.0      Max.   :359.9      Max.   : -42.88
##
##           NA's      :2           NA's      :11
##   latitude      elevation      prec      temp
## Min.   : -41.035      Min.   :  43      Min.   :  0.50      Min.   :  3.60
## 1st Qu.: -13.760      1st Qu.:2466      1st Qu.: 52.25      1st Qu.:12.58
## Median : -10.759      Median :2870      Median : 61.17      Median :13.77
## Mean   : -11.309      Mean   :2694      Mean   : 65.52      Mean   :14.26
## 3rd Qu.:  -7.205      3rd Qu.:3283      3rd Qu.: 70.31      3rd Qu.:16.65
## Max.   :   9.550      Max.   :4518      Max.   :344.00      Max.   :26.92
##
##           sr      GDD10c
## Min.   :11788      Min.   : 416.1
## 1st Qu.:14869      1st Qu.: 884.4
## Median :15701      Median : 967.0
## Mean   :15639      Mean   : 937.8
## 3rd Qu.:16570      3rd Qu.: 984.9
## Max.   :19178      Max.   :1196.2
##
##           NA's      :2
```

Dividing South American accessions into two groups using the median value.

```
mean(south_american_genotypes$latitude)
```

```
## [1] -11.30898
```

```
sa_1 <-filter(south_american_genotypes, latitude >= median(south_american_genotypes$latitude))
sa_2 <-filter(south_american_genotypes, latitude < median(south_american_genotypes$latitude))
```

```
dim(sa_1)
```

```
## [1] 84 14
dim(sa_2)

## [1] 82 14
range(sa_1$latitude)

## [1] -10.759094  9.550365
range(sa_2$latitude)

## [1] -41.03454 -12.06110
  • Central American group
ca <- filter(data_complete_coordinates, latitude > 9.56)
dim(ca)

## [1] 60 14
summary(ca)

##      genotypes      species      country      plot
## Length:60      Length:60      Length:60      Min.   : 106.0
## Class :character Class :character Class :character 1st Qu.: 323.0
## Mode  :character Mode  :character Mode  :character Median : 523.5
##                                     Mean   : 573.1
##                                     3rd Qu.: 813.5
##                                     Max.   :1027.0
##
##      seed_setting      ft_adjusted      ph_adjusted      longitude
## Min.   :0.0000      Min.   : 70.0      Min.   :129.2      Min.   : -108.83
## 1st Qu.:0.0000      1st Qu.: 76.0      1st Qu.:176.2      1st Qu.: -98.77
## Median :0.0000      Median :102.0      Median :212.2      Median : -98.33
## Mean   :0.4333      Mean   :107.2      Mean   :208.7      Mean   : -97.45
## 3rd Qu.:1.0000      3rd Qu.:138.0      3rd Qu.:233.2      3rd Qu.: -97.63
## Max.   :1.0000      Max.   :160.0      Max.   :318.2      Max.   : -66.48
##                                     NA's   :8      NA's   :5
##      latitude      elevation      prec      temp
## Min.   :14.57      Min.   : 95      Min.   : 45.83      Min.   :11.28
## 1st Qu.:18.72      1st Qu.:1356      1st Qu.: 64.33      1st Qu.:15.89
## Median :19.12      Median :1719      Median : 74.42      Median :17.47
## Mean   :19.07      Mean   :1722      Mean   : 91.29      Mean   :18.49
## 3rd Qu.:19.51      3rd Qu.:2254      3rd Qu.:113.25      3rd Qu.:21.78
## Max.   :27.42      Max.   :2984      Max.   :283.33      Max.   :25.79
##
##      sr      GDD10c
## Min.   :15910      Min.   : 594.9
## 1st Qu.:17264      1st Qu.: 663.6
## Median :17858      Median : 884.4
## Mean   :17833      Mean   : 888.9
## 3rd Qu.:18203      3rd Qu.:1113.5
## Max.   :19655      Max.   :1196.2
##                                     NA's   :8
```

Let's show the three different latitudinal groups (ca, sa_1 and sa_2) on the map

```

ca$latgroup <- "ca"
sa_1$latgroup <- "sa1"
sa_2$latgroup <- "sa2"
accessions_grouped_by_latitude <- rbind(ca, sa_1, sa_2)
data_gbl <- accessions_grouped_by_latitude
#dim(data2)
data_gbl

```

	genotypes	species	country	plot	seed_setting	ft_adjusted	ph_adjusted
## 1	cru_78_y	cru	MEX	106	1	70	187.95
## 2	hypo_157_y	hypo	MEX	109	0	137	223.95
## 3	hypo_158_y	hypo	MEX	112	1	101	265.95
## 4	cru_100_y	cru	MEX	116	1	70	176.95
## 5	hypo_174	hypo	MEX	128	0	140	235.70
## 6	hypo_146_y	hypo	MEX	201	0	146	230.20
## 7	hypo_148_y	hypo	MEX	202	0	138	NA
## 8	hypo_170_n	hypo	MEX	214	0	160	198.20
## 9	hybr_118_n	hybr	MEX	217	0	160	197.20
## 10	cru_89_y	cru	MEX	220	0	NA	229.20
## 11	hypo_159_y	hypo	MEX	225	0	138	226.20
## 12	cru_101_y	cru	MEX	305	1	76	202.20
## 13	hypo_153_y	hypo	MEX	306	1	117	318.20
## 14	hypo_141_y	hypo	MEX	316	0	134	233.20
## 15	cru_77_y	cru	MEX	317	1	71	153.20
## 16	cru_76_y	cru	MEX	325	1	71	156.20
## 17	hypo_165_n	hypo	MEX	401	0	151	239.20
## 18	hypo_168_n	hypo	MEX	408	0	151	NA
## 19	cru_91_y	cru	GTM	412	1	82	NA
## 20	cru_93_y	cru	MEX	419	1	89	209.20
## 21	hdus_124_y	hdus	GTM	420	1	89	231.20
## 22	hybr_119_y	hybr	MEX	426	0	102	195.20
## 23	cru_92_y	cru	MEX	503	0	113	212.20
## 24	hypo_160_y	hypo	MEX	504	1	92	180.20
## 25	hypo_166_n	hypo	MEX	506	0	138	225.20
## 26	hypo_143_y	hypo	MEX	512	0	125	213.20
## 27	hypo_167_n	hypo	MEX	513	1	92	229.20
## 28	hypo_149_n	hypo	PTR	515	1	82	129.20
## 29	hdus_137_n	hdus	MEX	516	1	71	228.20
## 30	cru_70_y	cru	MEX	522	1	71	142.20
## 31	hypo_145_y	hypo	MEX	525	0	102	172.20
## 32	hypo_144_y	hypo	MEX	527	0	113	190.20
## 33	hypo_150_n	hypo	MEX	612	0	110	NA
## 34	hypo_156_y	hypo	MEX	614	0	110	222.20
## 35	hypo_142_y	hypo	MEX	619	0	146	283.20
## 36	cru_68_y	cru	MEX	620	1	71	162.20
## 37	cru_98_y	cru	MEX	624	1	71	165.20
## 38	hypo_169_n	hypo	MEX	626	1	97	228.20
## 39	cru_74_y	cru	GTM	627	1	76	NA
## 40	hypo_172_n	hypo	MEX	714	0	NA	249.20
## 41	hypo_164_n	hypo	MEX	715	0	142	257.20
## 42	hybr_109_y	hybr	MEX	725	0	NA	154.20
## 43	hdus_130_y	hdus	MEX	728	1	110	186.20
## 44	hybr_121_n	hybr	GTM	801	0	NA	206.20
## 45	hypo_152_n	hypo	MEX	812	0	138	263.20

## 46	cru_94_y	cru	MEX	818	1	71	169.20
## 47	cru_82_n	cru	MEX	822	1	71	163.20
## 48	cru_79_y	cru	MEX	904	1	76	181.20
## 49	hypo_154_y	hypo	MEX	905	1	102	252.20
## 50	hypo_155_n	hypo	MEX	909	0	146	233.20
## 51	cru_86_n	cru	MEX	1001	1	72	167.45
## 52	hypo_147_y	hypo	MEX	1006	0	126	249.45
## 53	hdus_136_y	hdus	MEX	1010	0	NA	195.45
## 54	hypo_151_y	hypo	MEX	1012	0	152	253.45
## 55	hypo_140_y	hypo	MEX	1014	0	NA	224.45
## 56	hdus_132_n	hdus	GTM	1016	0	NA	234.45
## 57	cru_88_y	cru	GTM	1017	1	83	166.45
## 58	cru_95_y	cru	MEX	1019	0	72	166.45
## 59	hypo_162_y	hypo	MEX	1021	0	139	239.45
## 60	hdus_135_y	hdus	MEX	1027	0	NA	175.45
## 61	cau_339_y	cau	PER	102	0	124	234.95
## 62	cau_281_y	cau	PER	110	0	88	229.95
## 63	cau_27_y	cau	PER	114	0	109	266.95
## 64	quit_178_y	quit	ECU	117	0	109	257.95
## 65	hdus_298_y	hdus	ECU	119	0	109	251.95
## 66	cau_340_y	cau	PER	121	0	116	191.95
## 67	cau_354_y	cau	PER	123	0	137	257.95
## 68	quit_190_n	quit	ECU	126	0	109	261.95
## 69	hdus_125_y	hdus	ECU	127	0	102	261.95
## 70	hdus_122_y	hdus	PER	204	0	113	125.20
## 71	cau_316_y	cau	PER	206	0	110	194.20
## 72	cau_345_y	cau	PER	207	0	110	226.20
## 73	cau_51_y	cau	PER	209	0	89	196.20
## 74	hdus_123_y	hdus	PER	210	0	110	247.20
## 75	quit_280_y	quit	ECU	211	0	110	214.20
## 76	hdus_293_y	hdus	ECU	224	0	110	196.20
## 77	hybr_111_n	hybr	PER	226	0	102	260.20
## 78	cau_344_y	cau	PER	227	0	102	226.20
## 79	cau_330_y	cau	PER	304	0	110	283.20
## 80	cau_31_y	cau	PER	307	0	102	199.20
## 81	hdus_297_y	hdus	ECU	311	0	110	257.20
## 82	quit_180_n	quit	ECU	319	0	110	249.20
## 83	cau_43_y	cau	PER	320	0	113	196.20
## 84	cau_42_n	cau	PER	321	0	113	242.20
## 85	hdus_134_n	hdus	ECU	322	0	102	263.20
## 86	cau_56_y	cau	PER	323	0	110	269.20
## 87	hdus_131_y	hdus	ECU	324	0	82	240.20
## 88	hdus_138_y	hdus	ECU	328	0	110	275.20
## 89	hdus_294_y	hdus	ECU	415	0	110	275.20
## 90	hdus_302_y	hdus	ECU	417	0	113	231.20
## 91	quit_304_y	quit	ECU	418	0	110	230.20
## 92	cau_285_y	cau	PER	421	0	89	221.20
## 93	cau_366_y	cau	PER	423	0	113	NA
## 94	hybr_113_n	hybr	PER	425	0	125	235.20
## 95	quit_184_y	quit	ECU	501	0	110	191.20
## 96	cau_30_y	cau	PER	502	0	113	243.20
## 97	cru_72_n	cru	VEN	508	1	110	NA
## 98	cau_22_n	cau	PER	509	0	113	NA
## 99	quit_182_n	quit	PER	511	0	102	144.20

## 100	cau_346_y	cau	PER	518	0	110	211.20
## 101	cau_343_y	cau	PER	519	0	89	169.20
## 102	cau_341_y	cau	PER	523	0	102	179.20
## 103	hdus_129_y	hdus	ECU	602	0	110	239.20
## 104	cau_50_y	cau	PER	606	0	102	NA
## 105	quit_188_n	quit	ECU	607	0	113	193.20
## 106	cau_360_y	cau	PER	608	0	102	139.20
## 107	cau_23_n	cau	PER	613	0	102	242.20
## 108	cau_13_y	cau	PER	615	0	89	260.20
## 109	hdus_296_y	hdus	ECU	616	0	110	257.20
## 110	cau_19_y	cau	PER	618	0	110	NA
## 111	cau_283_y	cau	PER	622	0	110	253.20
## 112	cau_47_y	cau	ECU	628	0	110	246.20
## 113	cau_323_y	cau	PER	706	0	117	237.20
## 114	cau_28_y	cau	PER	707	0	110	216.20
## 115	cau_351_y	cau	PER	726	0	102	199.20
## 116	cau_325_y	cau	PER	727	0	131	132.20
## 117	cau_20_y	cau	PER	804	0	110	267.20
## 118	hdus_288_y	hdus	ECU	807	0	110	262.20
## 119	hdus_128_y	hdus	ECU	809	0	110	281.20
## 120	quit_290_y	quit	ECU	815	0	110	266.20
## 121	hybr_289_y	hybr	PER	819	0	89	253.20
## 122	cau_58_y	cau	PER	823	0	102	216.20
## 123	hdus_295_y	hdus	ECU	827	0	110	225.20
## 124	cau_329_y	cau	PER	828	0	117	219.20
## 125	cau_327_y	cau	PER	906	0	102	186.20
## 126	cau_342_y	cau	PER	907	0	110	221.20
## 127	cau_24_n	cau	PER	908	0	110	257.20
## 128	quit_179_n	quit	ECU	910	0	110	274.20
## 129	hdus_292_y	hdus	ECU	911	0	110	289.20
## 130	quit_194_n	quit	ECU	914	0	110	270.20
## 131	cau_326_y	cau	PER	915	0	134	242.20
## 132	cau_11_n	cau	PER	918	0	117	NA
## 133	cau_60_y	cau	PER	921	0	125	199.20
## 134	quit_193_n	quit	ECU	923	0	110	236.20
## 135	hdus_276_y	hdus	ECU	924	0	NA	192.20
## 136	hybr_110_n	hybr	PER	926	0	110	195.20
## 137	cau_359_y	cau	PER	927	0	102	228.20
## 138	cau_16_y	cau	PER	1004	0	103	304.45
## 139	hdus_300_y	hdus	ECU	1007	0	111	196.45
## 140	quit_191_n	quit	ECU	1008	0	103	245.45
## 141	cau_57_y	cau	ECU	1009	0	103	231.45
## 142	cau_29_n	cau	PER	1013	0	111	213.45
## 143	hdus_301_y	hdus	ECU	1020	0	126	190.45
## 144	hdus_299_y	hdus	ECU	1028	0	111	269.45
## 145	cau_314_y	cau	PER	103	0	101	306.95
## 146	cau_32_n	cau	PER	108	0	109	291.95
## 147	cau_59_y	cau	PER	111	0	112	267.95
## 148	hybr_116_y	hybr	PER	113	0	109	311.95
## 149	cau_353_n	cau	PER	115	0	109	359.95
## 150	quit_187_y	quit	BRA	118	1	64	148.95
## 151	cau_349_y	cau	PER	120	0	116	261.95
## 152	cau_17_y	cau	BOL	122	0	109	268.95
## 153	cau_315_y	cau	PER	124	0	109	263.95

## 154	cau_338_n	cau	BOL	208	0	117	247.20
## 155	cau_38_n	cau	PER	212	0	113	194.20
## 156	cau_361_y	cau	PER	221	0	NA	199.20
## 157	cau_34_y	cau	PER	228	0	110	210.20
## 158	cau_64_y	cau	BOL	301	0	103	300.20
## 159	quit_279_y	quit	PER	302	0	92	283.20
## 160	cau_286_y	cau	PER	303	0	113	218.20
## 161	cau_46_y	cau	PER	310	0	110	NA
## 162	hybr_107_n	hybr	PER	312	0	92	246.20
## 163	cau_368_y	cau	PER	313	0	110	204.20
## 164	cau_347_y	cau	PER	318	0	117	234.20
## 165	hybr_115_y	hybr	BOL	326	0	110	269.20
## 166	cau_54_n	cau	PER	327	0	110	295.20
## 167	cau_33_y	cau	ARG	402	0	125	285.20
## 168	cau_319_y	cau	PER	403	0	110	239.20
## 169	cau_18_y	cau	PER	406	0	125	217.20
## 170	cru_99_y	cru	BRA	407	1	97	242.20
## 171	cau_335_y	cau	PER	409	0	102	NA
## 172	cau_318_y	cau	PER	410	0	102	NA
## 173	cau_367_y	cau	PER	411	0	131	229.20
## 174	cau_337_y	cau	PER	413	0	102	231.20
## 175	cau_332_y	cau	BOL	414	0	125	239.20
## 176	cau_324_y	cau	PER	416	0	102	261.20
## 177	cau_310_y	cau	BOL	422	0	110	276.20
## 178	cau_62_y	cau	PER	427	0	113	190.20
## 179	cau_49_y	cau	BOL	505	0	110	219.20
## 180	cau_350_y	cau	PER	507	0	110	217.20
## 181	cau_308_y	cau	BOL	514	0	110	221.20
## 182	cru_85_n	cru	ARG	517	1	71	144.20
## 183	cau_307_y	cau	BOL	521	0	102	196.20
## 184	cau_362_y	cau	PER	601	0	110	200.20
## 185	cau_363_y	cau	PER	603	0	125	201.20
## 186	cau_352_y	cau	PER	604	0	102	235.20
## 187	quit_192_y	quit	BOL	611	0	92	NA
## 188	cau_14_y	cau	PER	617	0	102	275.20
## 189	cau_336_y	cau	PER	621	0	97	215.20
## 190	cau_317_y	cau	PER	625	0	89	189.20
## 191	cau_48_y	cau	PER	702	0	110	206.20
## 192	cau_369_y	cau	PER	704	0	113	NA
## 193	cau_312_y	cau	PER	705	0	113	227.20
## 194	hdus_127_y	hdus	PER	709	0	151	107.20
## 195	cau_284_y	cau	PER	710	0	110	194.20
## 196	quit_196_y	quit	PER	716	0	110	207.20
## 197	hybr_114_n	hybr	PER	717	1	50	131.20
## 198	cru_67_y	cru	PER	718	1	110	210.20
## 199	cau_53_y	cau	PER	719	0	110	226.20
## 200	cau_35_n	cau	BOL	720	0	110	226.20
## 201	quit_181_n	quit	ARG	722	0	76	204.20
## 202	cau_321_y	cau	BOL	723	0	117	182.20
## 203	cau_37_n	cau	BOL	724	0	117	216.20
## 204	cau_309_y	cau	BOL	802	0	102	288.20
## 205	cau_36_n	cau	PER	805	0	110	254.20
## 206	cau_356_y	cau	PER	806	0	128	181.20
## 207	cau_305_y	cau	BOL	816	0	110	273.20

## 208	quit_186_n	quit	PER	817	1	50	144.20
## 209	cau_15_y	cau	ARG	824	0	125	219.20
## 210	quit_185_y	quit	BRA	825	0	128	216.20
## 211	cau_331_y	cau	PER	826	0	113	187.20
## 212	hybr_120_y	hybr	PER	901	0	160	261.20
## 213	cau_313_n	cau	PER	902	0	125	245.20
## 214	cau_311_y	cau	PER	903	0	110	241.20
## 215	quit_195_n	quit	ARG	912	0	92	205.20
## 216	hybr_117_y	hybr	BOL	913	0	89	213.20
## 217	cau_21_y	cau	ARG	920	0	102	218.20
## 218	quit_183_n	quit	BOL	922	0	102	233.20
## 219	hybr_303_y	hybr	BOL	925	0	110	206.20
## 220	cau_25_y	cau	PER	1002	0	126	252.45
## 221	cau_358_y	cau	BOL	1003	0	114	236.45
## 222	cau_41_n	cau	PER	1005	0	111	235.45
## 223	cau_12_n	cau	BOL	1018	0	118	242.45
## 224	cau_355_n	cau	PER	1023	0	103	212.45
## 225	cau_322_y	cau	PER	1025	0	118	223.45
## 226	hdus_139_y	hdus	BOL	1026	0	103	198.45
##	longitude	latitude	elevation	prec	temp	sr	GDD10c
## 1	-98.50000	18.750000	1540	70.416667	21.775000	17724.42	594.90
## 2	-97.71667	20.001389	723	175.750000	20.358333	17263.58	1107.95
## 3	-98.75000	18.750000	1543	76.250000	21.808334	17850.33	872.40
## 4	-99.01667	19.250000	2348	55.583333	15.916667	18202.75	594.90
## 5	-98.21667	19.683333	2984	77.333333	11.275000	17715.00	1116.80
## 6	-98.33333	19.250000	2348	65.750000	15.358334	17992.25	1126.95
## 7	-98.76667	19.116667	2477	78.750000	15.150000	17840.83	1113.45
## 8	-98.21667	19.683333	2984	77.333333	11.275000	17715.00	1196.15
## 9	-97.94471	20.012118	1987	96.833333	15.166667	17438.25	1196.15
## 10	-96.77312	16.874277	1495	70.083333	21.666667	17064.92	NA
## 11	-97.01667	18.966667	1275	131.166667	19.691667	15909.75	1113.45
## 12	-98.33414	19.239048	2255	64.416667	15.825000	17998.58	663.60
## 13	-98.75000	18.716667	1468	74.416667	22.416667	17947.42	997.45
## 14	-97.71667	18.966667	2125	45.833333	15.783333	17248.92	1093.65
## 15	-99.01667	19.250000	2348	55.583333	15.916667	18202.75	611.20
## 16	-98.33414	19.239048	2255	64.416667	15.825000	17998.58	611.20
## 17	-98.76667	19.116667	2477	78.750000	15.150000	17840.83	1156.65
## 18	-97.71667	18.966667	2125	45.833333	15.783333	17248.92	1156.65
## 19	-90.79379	14.796239	1801	121.750000	17.058333	19452.83	719.35
## 20	-98.89401	19.498709	2248	47.500000	16.300000	18370.25	778.45
## 21	-90.73043	14.571722	1616	102.916667	18.066667	19654.75	778.45
## 22	-97.71667	20.001389	723	175.750000	20.358333	17263.58	884.35
## 23	-107.00000	26.166667	2183	82.750000	14.550000	18071.00	984.90
## 24	-108.83333	27.416667	392	60.583333	24.050000	18454.00	797.30
## 25	-98.33414	19.239048	2255	64.416667	15.825000	17998.58	1113.45
## 26	-99.01667	19.250000	2302	52.916667	16.100000	18189.00	1050.70
## 27	-98.89821	19.559241	2254	46.083333	16.241667	18372.83	797.30
## 28	-66.47890	18.244193	536	171.250000	22.275000	18434.33	719.35
## 29	-97.44444	19.880556	1704	105.750000	16.166667	16742.17	611.20
## 30	-98.75000	18.716667	1468	74.416667	22.416667	17947.42	611.20
## 31	-108.50000	27.416667	588	64.083333	22.991667	18472.50	884.35
## 32	-97.01667	18.950000	1146	122.083333	20.241667	15911.58	984.90
## 33	-98.16667	19.416667	2431	62.416667	14.441667	17865.58	967.00
## 34	-97.63333	19.916667	1734	106.000000	16.325000	17069.00	967.00

## 35	-99.01667	19.250000	2348	55.583333	15.916667	18202.75	1126.95
## 36	-98.76667	18.700000	1380	72.750000	22.983333	17906.42	611.20
## 37	-98.75000	18.750000	1543	76.250000	21.808334	17850.33	611.20
## 38	-98.89401	19.498709	2248	47.500000	16.300000	18370.25	837.15
## 39	-90.79379	14.796239	1801	121.750000	17.058333	19452.83	663.60
## 40	-98.46468	18.594558	1284	62.916667	22.983333	17754.83	NA
## 41	-97.01667	18.966667	1275	131.166667	19.691667	15909.75	1118.95
## 42	-97.62778	19.966667	1027	164.916667	19.283334	17165.17	NA
## 43	-96.38867	18.066202	95	283.333333	25.791667	16188.33	967.00
## 44	-90.82485	14.659463	1804	118.750000	17.225000	19384.33	NA
## 45	-98.89401	19.498709	2248	47.500000	16.300000	18370.25	1113.45
## 46	-98.79761	18.724242	1437	74.666667	22.566667	17851.00	611.20
## 47	-98.75000	18.716667	1468	74.416667	22.416667	17947.42	611.20
## 48	-98.50000	18.750000	1540	70.416667	21.775000	17724.42	663.60
## 49	-98.46468	18.594558	1284	62.916667	22.983333	17754.83	884.35
## 50	-98.76667	19.116667	2477	78.750000	15.150000	17840.83	1126.95
## 51	-98.75000	18.716667	1468	74.416667	22.416667	17947.42	623.55
## 52	-98.33333	19.250000	2348	65.750000	15.358334	17992.25	1053.85
## 53	-97.62778	19.966667	1027	164.916667	19.283334	17165.17	NA
## 54	-97.62778	19.966667	1027	164.916667	19.283334	17165.17	1161.85
## 55	-96.86667	16.800000	1633	64.583333	20.466666	17211.83	NA
## 56	-91.49589	15.321985	1878	111.416667	17.708333	19019.33	NA
## 57	-90.82485	14.659463	1804	118.750000	17.225000	19384.33	725.80
## 58	-98.75000	18.716667	1468	74.416667	22.416667	17947.42	623.55
## 59	-98.42435	19.270982	2251	64.500000	16.008333	17840.75	1115.95
## 60	-97.62778	19.966667	1027	164.916667	19.283334	17165.17	NA
## 61	-77.00000	-9.000000	3363	67.000000	12.916667	14869.42	1047.15
## 62	-77.13901	-9.486735	3273	70.416667	12.066667	14815.58	773.80
## 63	-77.72563	-9.416307	4518	55.583333	3.600000	15700.67	960.30
## 64	-78.65881	-1.628928	2865	48.666667	13.458333	15607.00	960.30
## 65	-78.66724	-1.414770	3415	69.000000	9.000000	12427.17	960.30
## 66	-77.53007	-9.514881	3052	48.666667	13.441667	15679.75	992.90
## 67	-75.73134	-10.759094	2622	61.916667	14.500000	14874.92	1107.95
## 68	-78.88653	-2.710103	2664	71.333333	13.650000	13762.92	960.30
## 69	-78.39511	-0.211229	2348	80.000000	15.800000	15344.75	884.35
## 70	-77.74361	-9.138330	2506	34.000000	16.316667	15795.17	984.90
## 71	-77.74406	-9.139853	2506	34.000000	16.316667	15795.17	967.00
## 72	-77.17131	-9.348316	3107	66.000000	13.366667	14839.75	967.00
## 73	-78.17221	-7.334439	2260	54.333333	16.783333	15110.33	778.45
## 74	-78.75203	-6.706250	2722	112.500000	13.825000	14545.33	967.00
## 75	-77.98376	-1.633505	857	344.000000	21.508333	13174.42	967.00
## 76	-78.87554	-2.545267	3100	62.583333	11.233333	14333.83	967.00
## 77	-75.73134	-10.759094	2622	61.916667	14.500000	14874.92	884.35
## 78	-77.15529	-9.325917	3301	66.083333	12.025000	14644.08	884.35
## 79	-77.53007	-9.514881	3052	48.666667	13.441667	15679.75	967.00
## 80	-77.00000	-9.000000	3363	67.000000	12.916667	14869.42	884.35
## 81	-78.62929	-1.607675	2697	45.000000	14.241667	15429.67	967.00
## 82	-78.77416	-2.886787	2268	74.750000	15.783333	13983.83	967.00
## 83	-78.54736	-0.369085	3092	108.500000	11.691667	14912.00	984.90
## 84	-78.03183	-5.973988	2700	89.250000	14.575000	14288.00	984.90
## 85	-79.14679	-3.143396	2249	53.333333	15.283333	14446.92	884.35
## 86	-77.00000	-9.000000	3363	67.000000	12.916667	14869.42	967.00
## 87	-78.54652	-0.395852	2919	98.250000	12.750000	14943.50	719.35
## 88	-78.65881	-1.628928	2865	48.666667	13.458333	15607.00	967.00

## 89	-78.92419	-2.514139	3028	56.916667	11.566667	14454.67	967.00
## 90	-78.60702	-1.375740	2921	54.583333	12.583333	13208.92	984.90
## 91	-78.92419	-2.514139	3028	56.916667	11.566667	14454.67	967.00
## 92	-77.53588	-9.432803	3272	49.000000	11.991667	15609.92	778.45
## 93	-78.17221	-7.334439	2260	54.333333	16.783333	15110.33	984.90
## 94	-75.73134	-10.759094	2622	61.916667	14.500000	14874.92	1050.70
## 95	-78.12844	0.287972	2663	70.000000	13.991667	14429.75	967.00
## 96	-78.50768	-7.162494	2694	69.333333	14.783333	14752.58	984.90
## 97	-69.18047	9.550365	177	99.416667	26.916667	16881.00	967.00
## 98	-77.00000	-9.000000	3363	67.000000	12.916667	14869.42	984.90
## 99	-77.72563	-9.416307	4518	55.583333	3.600000	15700.67	884.35
## 100	-77.00000	-9.000000	3363	67.000000	12.916667	14869.42	967.00
## 101	-78.17221	-7.334439	2260	54.333333	16.783333	15110.33	778.45
## 102	-77.00000	-9.000000	3363	67.000000	12.916667	14869.42	884.35
## 103	-78.59322	-1.134405	2672	46.250000	13.708333	14525.75	967.00
## 104	-77.13901	-9.486735	3273	70.416667	12.066667	14815.58	884.35
## 105	-78.65881	-1.628928	2865	48.666667	13.458333	15607.00	984.90
## 106	-77.00000	-9.000000	3363	67.000000	12.916667	14869.42	884.35
## 107	-77.64205	-9.832345	3385	45.416667	10.141667	15969.25	884.35
## 108	-77.00000	-9.000000	3363	67.000000	12.916667	14869.42	778.45
## 109	-78.72185	-1.945311	3182	50.166667	11.250000	15203.67	967.00
## 110	-77.64655	-9.281319	2667	43.416667	15.966667	16824.08	967.00
## 111	-77.00000	-9.000000	3363	67.000000	12.916667	14869.42	967.00
## 112	-78.35147	-0.063325	2125	60.000000	16.875000	15455.83	967.00
## 113	-78.03183	-5.973988	2700	89.250000	14.575000	14288.00	997.45
## 114	-77.00000	-9.000000	3363	67.000000	12.916667	14869.42	967.00
## 115	-77.00000	-9.000000	3363	67.000000	12.916667	14869.42	884.35
## 116	-77.00000	-9.000000	3363	67.000000	12.916667	14869.42	1078.50
## 117	-78.50163	-7.157655	2692	68.500000	14.775000	14786.33	967.00
## 118	-78.61550	-1.353993	2876	53.833333	12.916667	13562.25	967.00
## 119	-78.54952	-1.177857	2776	53.416667	13.333333	13377.83	967.00
## 120	-79.00689	-2.899469	2546	68.000000	14.516667	13219.25	967.00
## 121	-78.03183	-5.973988	2700	89.250000	14.575000	14288.00	778.45
## 122	-77.00000	-9.000000	3363	67.000000	12.916667	14869.42	884.35
## 123	-78.20595	0.313251	2560	68.666667	14.608333	14902.33	967.00
## 124	-77.00000	-9.000000	3363	67.000000	12.916667	14869.42	997.45
## 125	-77.53588	-9.432803	3272	49.000000	11.991667	15609.92	884.35
## 126	-77.17131	-9.348316	3107	66.000000	13.366667	14839.75	967.00
## 127	-77.74361	-9.138330	2506	34.000000	16.316667	15795.17	967.00
## 128	-78.20761	0.041156	2823	66.750000	13.433333	14946.33	967.00
## 129	-78.88653	-2.710103	2664	71.333333	13.650000	13762.92	967.00
## 130	-79.00689	-2.899469	2546	68.000000	14.516667	13219.25	967.00
## 131	-78.17221	-7.334439	2260	54.333333	16.783333	15110.33	1093.65
## 132	-77.72563	-9.416307	4518	55.583333	3.600000	15700.67	997.45
## 133	-77.53007	-9.514881	3052	48.666667	13.441667	15679.75	1050.70
## 134	-79.82561	-1.043134	58	131.916667	25.133334	11787.83	967.00
## 135	-79.00128	-1.588997	2765	59.583333	13.050000	15128.42	NA
## 136	-78.50768	-7.162494	2694	69.333333	14.783333	14752.58	967.00
## 137	-77.17131	-9.348316	3107	66.000000	13.366667	14839.75	884.35
## 138	-77.72563	-9.416307	4518	55.583333	3.600000	15700.67	896.70
## 139	-78.60702	-1.375740	2921	54.583333	12.583333	13208.92	975.25
## 140	-78.84740	-2.198292	2386	49.666667	14.875000	15188.17	896.70
## 141	-79.82561	-1.043134	58	131.916667	25.133334	11787.83	896.70
## 142	-77.53007	-9.514881	3052	48.666667	13.441667	15679.75	975.25

## 143	-78.62263	-1.253131	2660	46.166667	13.925000	15065.50	1053.85
## 144	-78.66869	-1.347782	3244	61.750000	10.475000	12829.25	975.25
## 145	-73.58656	-13.697981	3219	79.416667	12.558333	16153.75	872.40
## 146	-74.22260	-13.164547	2729	45.333333	16.658333	16572.17	960.30
## 147	-74.25025	-12.931580	2640	40.833333	16.691667	16699.08	980.30
## 148	-71.89052	-13.535464	3309	51.666667	11.850000	16561.75	960.30
## 149	-72.54998	-13.516412	2754	98.500000	16.091667	16451.00	960.30
## 150	-47.65000	-15.633333	958	121.500000	21.341667	16605.83	516.35
## 151	-74.58266	-12.860836	3452	67.750000	11.508333	16548.92	992.90
## 152	-64.75883	-21.409904	2031	52.500000	18.375000	17303.33	960.30
## 153	-73.56196	-13.478882	3038	81.833333	13.833333	15172.67	960.30
## 154	-66.16782	-17.413549	2564	57.000000	17.233333	16719.67	997.45
## 155	-75.19646	-12.061098	3307	61.333333	11.533333	16674.42	984.90
## 156	-64.75883	-21.409904	2031	52.500000	18.375000	17303.33	NA
## 157	-71.91762	-13.668616	3406	61.000000	11.016667	15969.25	967.00
## 158	-64.11433	-19.097700	1635	52.666667	20.116667	15648.33	896.70
## 159	-71.67067	-13.356775	3521	47.333333	10.500000	16281.33	797.30
## 160	-72.36342	-13.756276	3507	76.500000	10.575000	15698.75	984.90
## 161	-70.16667	-15.216667	3895	58.666667	8.791667	17090.42	967.00
## 162	-73.37850	-13.658872	2972	86.916667	14.400000	15873.75	797.30
## 163	-71.84630	-13.761848	3096	54.083333	12.625000	16649.25	967.00
## 164	-72.69996	-13.525102	2782	98.333333	16.133333	16877.25	997.45
## 165	-65.59431	-22.066221	3447	27.833333	11.975000	19178.08	967.00
## 166	-74.22260	-13.164547	2729	45.333333	16.658333	16572.17	967.00
## 167	-68.69652	-32.900585	661	17.833333	16.633333	18259.42	1050.70
## 168	-72.54704	-13.510070	2754	98.500000	16.091667	16451.00	967.00
## 169	-74.22260	-13.164547	2729	45.333333	16.658333	16572.17	1050.70
## 170	-42.87518	-20.722481	652	106.166667	19.850000	14410.17	837.15
## 171	-72.96320	-13.366720	2466	123.833333	17.925000	16132.50	884.35
## 172	-73.37850	-13.658872	2972	86.916667	14.400000	15873.75	884.35
## 173	-72.82476	-15.643097	4185	34.166667	5.008333	18601.75	1078.50
## 174	-72.96320	-13.366720	2466	123.833333	17.925000	16132.50	884.35
## 175	-66.26056	-17.688695	2431	47.083333	17.775000	17466.08	1050.70
## 176	-71.62327	-13.682642	3272	46.333333	11.491667	16326.58	884.35
## 177	-66.27980	-17.354352	2615	62.416667	16.675000	16561.25	967.00
## 178	-75.01390	-12.508010	3410	62.250000	11.158333	16350.08	984.90
## 179	-64.75841	-21.409969	2031	52.500000	18.375000	17303.33	967.00
## 180	-73.37850	-13.658872	2972	86.916667	14.400000	15873.75	967.00
## 181	-68.91265	-17.471669	3885	34.083333	8.708333	18219.42	967.00
## 182	-65.30244	-36.999577	413	43.583333	14.933333	16574.33	611.20
## 183	-64.75841	-21.409969	2031	52.500000	18.375000	17303.33	884.35
## 184	-75.19646	-12.061098	3307	61.333333	11.533333	16674.42	967.00
## 185	-75.01390	-12.508010	3410	62.250000	11.158333	16350.08	1050.70
## 186	-71.67599	-13.918236	3281	59.416667	10.925000	16488.25	884.35
## 187	-64.91465	-19.170805	3284	52.250000	13.191667	16505.33	797.30
## 188	-72.08357	-13.321357	2908	40.916667	14.641667	16473.83	884.35
## 189	-72.96320	-13.366720	2466	123.833333	17.925000	16132.50	837.15
## 190	-73.37850	-13.658872	2972	86.916667	14.400000	15873.75	778.45
## 191	-71.62327	-13.682642	3272	46.333333	11.491667	16326.58	967.00
## 192	-72.88791	-13.633957	2346	109.083333	19.241667	15822.00	984.90
## 193	-72.57468	-13.643851	3066	84.166667	13.633333	16639.92	984.90
## 194	-73.37850	-13.658872	2972	86.916667	14.400000	15873.75	1156.65
## 195	-73.37850	-13.658872	2972	86.916667	14.400000	15873.75	967.00
## 196	-72.23330	-13.283330	3225	54.833333	12.975000	16327.75	967.00

## 197	-70.26695	-18.017324	515	0.500000	18.166667	17525.42	416.05
## 198	-77.03616	-12.084401	112	2.583333	19.166667	13692.25	967.00
## 199	-71.84630	-13.761848	3096	54.083333	12.625000	16649.25	967.00
## 200	-64.72605	-21.517769	1887	55.416667	18.891667	17130.75	967.00
## 201	-63.59285	-35.233903	127	72.750000	16.383333	16452.75	663.60
## 202	-64.72605	-21.517769	1887	55.416667	18.891667	17130.75	997.45
## 203	-64.72605	-21.517769	1887	55.416667	18.891667	17130.75	997.45
## 204	-65.73575	-17.402299	3427	56.916667	12.291667	16527.00	884.35
## 205	-71.67599	-13.918236	3281	59.416667	10.925000	16488.25	967.00
## 206	-71.16999	-14.355805	3694	62.166667	7.950000	16892.42	1064.15
## 207	-64.91465	-19.170805	3284	52.250000	13.191667	16505.33	967.00
## 208	-70.26695	-18.017324	515	0.500000	18.166667	17525.42	416.05
## 209	-71.26836	-41.034544	907	68.333333	8.183333	15208.50	1050.70
## 210	-47.65000	-15.633333	958	121.500000	21.341667	16605.83	1064.15
## 211	-71.84630	-13.761848	3096	54.083333	12.625000	16649.25	984.90
## 212	-75.19646	-12.061098	3307	61.333333	11.533333	16674.42	1196.15
## 213	-73.37850	-13.658872	2972	86.916667	14.400000	15873.75	1050.70
## 214	-73.01293	-14.223123	3532	75.750000	10.100000	16835.92	967.00
## 215	-61.09163	-31.727270	43	83.000000	18.216667	17026.92	797.30
## 216	-64.11433	-19.097700	1635	52.666667	20.116667	15648.33	778.45
## 217	-65.30244	-36.999577	413	43.583333	14.933333	16574.33	884.35
## 218	-64.72605	-21.517769	1887	55.416667	18.891667	17130.75	884.35
## 219	-64.91465	-19.170805	3284	52.250000	13.191667	16505.33	967.00
## 220	-73.61395	-13.533680	3491	79.083333	10.725000	15398.92	1053.85
## 221	-64.82037	-21.622475	1980	55.833333	18.158333	15831.83	988.35
## 222	-74.22260	-13.164547	2729	45.333333	16.658333	16572.17	975.25
## 223	-64.72605	-21.517769	1887	55.416667	18.891667	17130.75	1002.95
## 224	-74.22260	-13.164547	2729	45.333333	16.658333	16572.17	896.70
## 225	-73.37850	-13.658872	2972	86.916667	14.400000	15873.75	1002.95
## 226	-64.72605	-21.517769	1887	55.416667	18.891667	17130.75	896.70
##	latgroup						
## 1	ca						
## 2	ca						
## 3	ca						
## 4	ca						
## 5	ca						
## 6	ca						
## 7	ca						
## 8	ca						
## 9	ca						
## 10	ca						
## 11	ca						
## 12	ca						
## 13	ca						
## 14	ca						
## 15	ca						
## 16	ca						
## 17	ca						
## 18	ca						
## 19	ca						
## 20	ca						
## 21	ca						
## 22	ca						
## 23	ca						

## 24	ca
## 25	ca
## 26	ca
## 27	ca
## 28	ca
## 29	ca
## 30	ca
## 31	ca
## 32	ca
## 33	ca
## 34	ca
## 35	ca
## 36	ca
## 37	ca
## 38	ca
## 39	ca
## 40	ca
## 41	ca
## 42	ca
## 43	ca
## 44	ca
## 45	ca
## 46	ca
## 47	ca
## 48	ca
## 49	ca
## 50	ca
## 51	ca
## 52	ca
## 53	ca
## 54	ca
## 55	ca
## 56	ca
## 57	ca
## 58	ca
## 59	ca
## 60	ca
## 61	sa1
## 62	sa1
## 63	sa1
## 64	sa1
## 65	sa1
## 66	sa1
## 67	sa1
## 68	sa1
## 69	sa1
## 70	sa1
## 71	sa1
## 72	sa1
## 73	sa1
## 74	sa1
## 75	sa1
## 76	sa1
## 77	sa1

## 78	sa1
## 79	sa1
## 80	sa1
## 81	sa1
## 82	sa1
## 83	sa1
## 84	sa1
## 85	sa1
## 86	sa1
## 87	sa1
## 88	sa1
## 89	sa1
## 90	sa1
## 91	sa1
## 92	sa1
## 93	sa1
## 94	sa1
## 95	sa1
## 96	sa1
## 97	sa1
## 98	sa1
## 99	sa1
## 100	sa1
## 101	sa1
## 102	sa1
## 103	sa1
## 104	sa1
## 105	sa1
## 106	sa1
## 107	sa1
## 108	sa1
## 109	sa1
## 110	sa1
## 111	sa1
## 112	sa1
## 113	sa1
## 114	sa1
## 115	sa1
## 116	sa1
## 117	sa1
## 118	sa1
## 119	sa1
## 120	sa1
## 121	sa1
## 122	sa1
## 123	sa1
## 124	sa1
## 125	sa1
## 126	sa1
## 127	sa1
## 128	sa1
## 129	sa1
## 130	sa1
## 131	sa1

## 132	sa1
## 133	sa1
## 134	sa1
## 135	sa1
## 136	sa1
## 137	sa1
## 138	sa1
## 139	sa1
## 140	sa1
## 141	sa1
## 142	sa1
## 143	sa1
## 144	sa1
## 145	sa2
## 146	sa2
## 147	sa2
## 148	sa2
## 149	sa2
## 150	sa2
## 151	sa2
## 152	sa2
## 153	sa2
## 154	sa2
## 155	sa2
## 156	sa2
## 157	sa2
## 158	sa2
## 159	sa2
## 160	sa2
## 161	sa2
## 162	sa2
## 163	sa2
## 164	sa2
## 165	sa2
## 166	sa2
## 167	sa2
## 168	sa2
## 169	sa2
## 170	sa2
## 171	sa2
## 172	sa2
## 173	sa2
## 174	sa2
## 175	sa2
## 176	sa2
## 177	sa2
## 178	sa2
## 179	sa2
## 180	sa2
## 181	sa2
## 182	sa2
## 183	sa2
## 184	sa2
## 185	sa2

```
## 186      sa2
## 187      sa2
## 188      sa2
## 189      sa2
## 190      sa2
## 191      sa2
## 192      sa2
## 193      sa2
## 194      sa2
## 195      sa2
## 196      sa2
## 197      sa2
## 198      sa2
## 199      sa2
## 200      sa2
## 201      sa2
## 202      sa2
## 203      sa2
## 204      sa2
## 205      sa2
## 206      sa2
## 207      sa2
## 208      sa2
## 209      sa2
## 210      sa2
## 211      sa2
## 212      sa2
## 213      sa2
## 214      sa2
## 215      sa2
## 216      sa2
## 217      sa2
## 218      sa2
## 219      sa2
## 220      sa2
## 221      sa2
## 222      sa2
## 223      sa2
## 224      sa2
## 225      sa2
## 226      sa2
```

Latitudinal grouping on the map

```
map3 <- ggplot() + mapa + theme_bw() +
  theme(panel.border = element_blank(), panel.grid.major = element_line(colour = "grey80"), panel
  geom_point(data = data_gbl, # Add and plot species data
    aes(x = longitude, y = latitude,
        colour = latgroup)) +
  coord_quickmap() + # Prevents stretching when resizing
  theme_bw() + # Remove ugly grey background
  xlab("Longitude") +
  ylab("Latitude") +
  theme(legend.position="none")
  #guides(colour=guide_legend(title="Accessions grouped by latitude"))
```

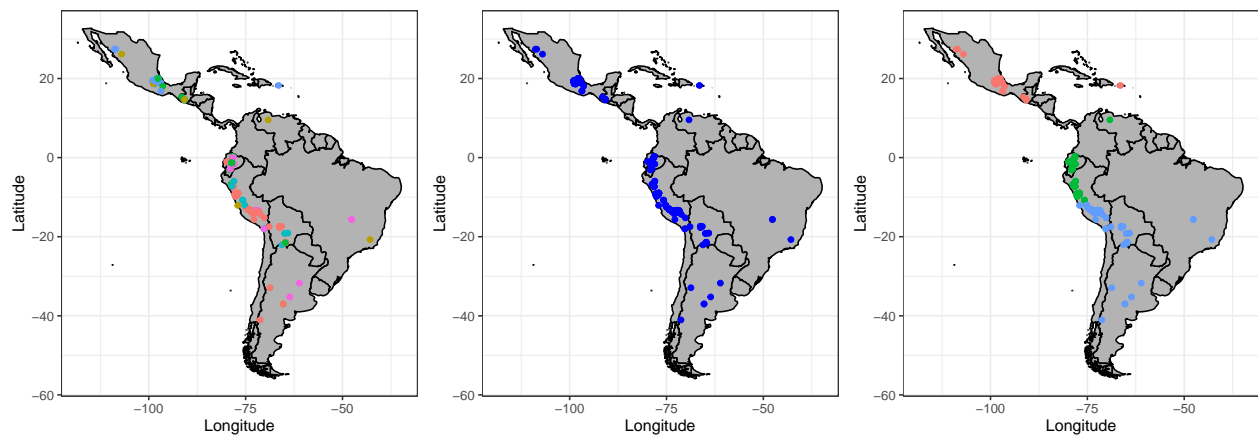
```
#map3
```

Three maps together. First is showing the accessions by species, second shows all accessions with the same color. Third shows our latitudinal grouping.

```
maps <- grid.arrange(map1, map2, map3, ncol = 3)
```

```
## Warning: Removed 28 rows containing missing values (geom_point).
```

```
## Warning: Removed 28 rows containing missing values (geom_point).
```



```
maps
```

```
## TableGrob (1 x 3) "arrange": 3 grobs
##   z      cells  name      grob
## 1 1 (1-1,1-1) arrange gtable[layout]
## 2 2 (1-1,2-2) arrange gtable[layout]
## 3 3 (1-1,3-3) arrange gtable[layout]
```

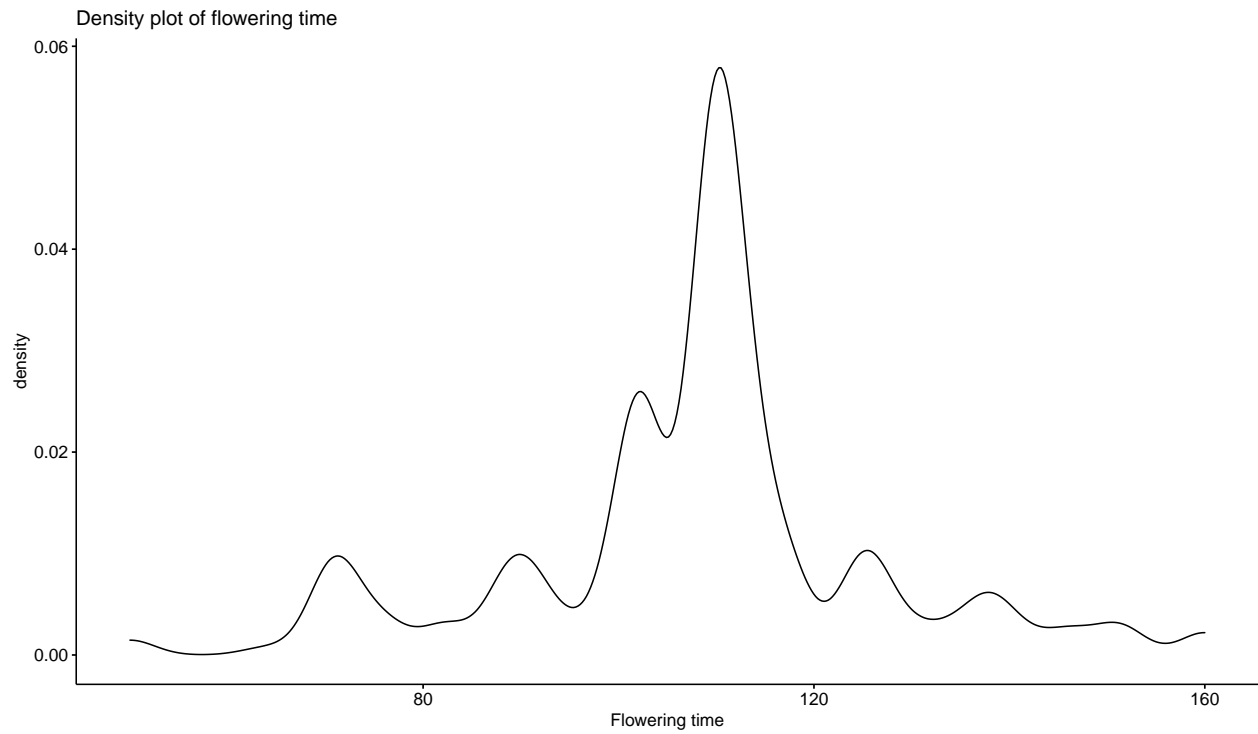
2. Comparison of latitudinal groups

Phenotypic traits

A. Flowering time

```
ggdensity(data_gbl$ft_adjusted,
           main = "Density plot of flowering time",
           xlab = "Flowering time")
```

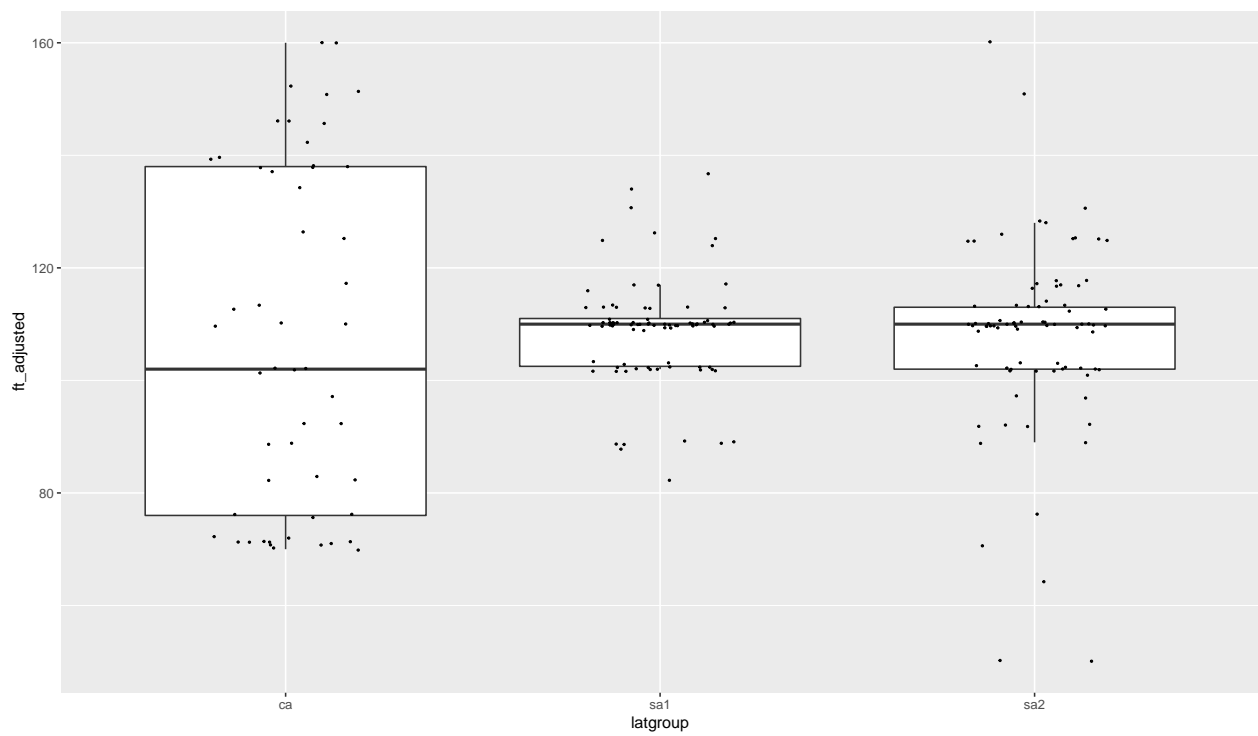
```
## Warning: Removed 10 rows containing non-finite values (stat_density).
```

```
ggplot(data = data_gbl, mapping = aes(x = latgroup, y = ft_adjusted)) +  
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```

Warning: Removed 10 rows containing non-finite values (stat_boxplot).

Warning: Removed 10 rows containing missing values (geom_point).



```
md_11 <-glm(ft_adjusted ~ latgroup , data = data_gbl, family = gaussian)
summary(md_11)

##
## Call:
## glm(formula = ft_adjusted ~ latgroup, family = gaussian, data = data_gbl)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -57.963  -6.723   1.277   5.037  52.827
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 107.1731     2.6365  40.649  <2e-16 ***
## latgroupsa1   1.5498     3.3625   0.461   0.645
## latgroupsa2   0.7899     3.3784   0.234   0.815
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 361.4693)
##
##      Null deviance: 77071  on 215  degrees of freedom
## Residual deviance: 76993  on 213  degrees of freedom
## (10 observations deleted due to missingness)
## AIC: 1890.2
##
## Number of Fisher Scoring iterations: 2
lrtest(md_11)

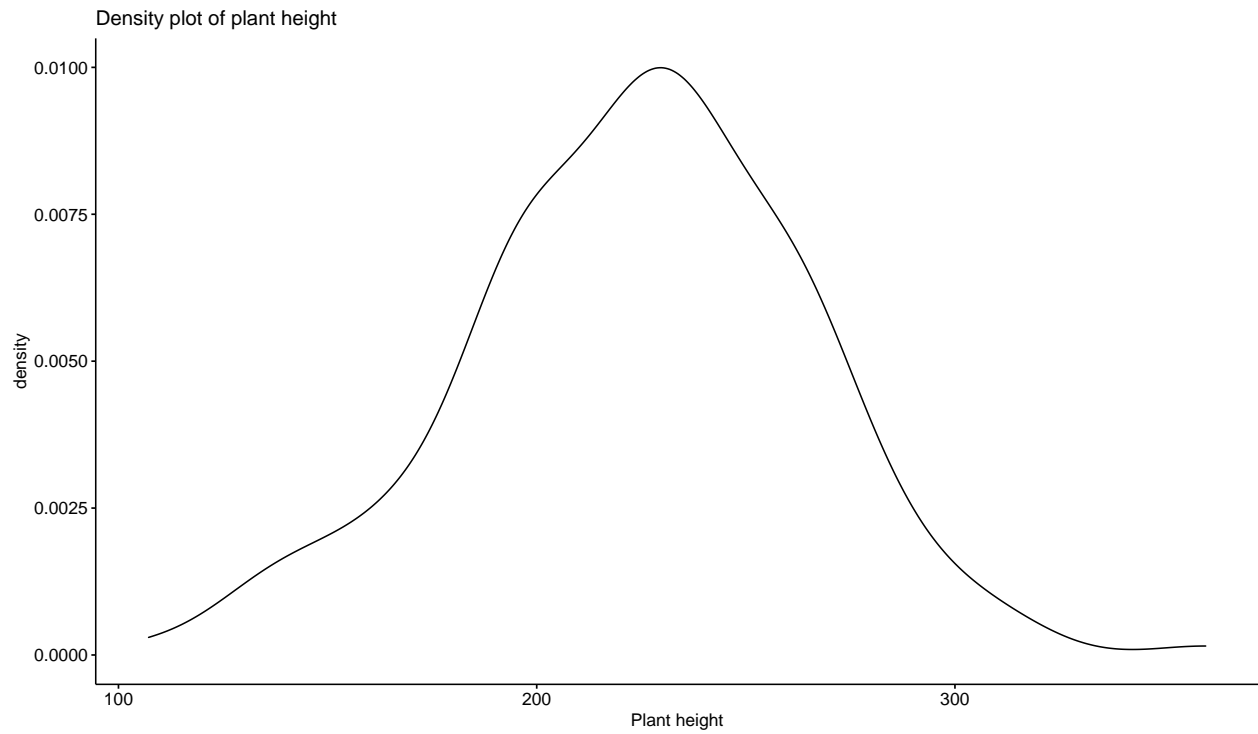
## Likelihood ratio test
##
## Model 1: ft_adjusted ~ latgroup
## Model 2: ft_adjusted ~ 1
##   #Df  LogLik Df  Chisq Pr(>Chisq)
## 1    4 -941.12
## 2    2 -941.23 -2  0.2191    0.8962
#lsmeans_latgroup_ft <- LSD.test(md_11, "latgroup", p.adj="none" )
#lsmeans_latgroup_ft
```

No significant difference

B. Plant height

```
ggdensity(data_gbl$ph_adjusted,
           main = "Density plot of plant height",
           xlab = "Plant height")

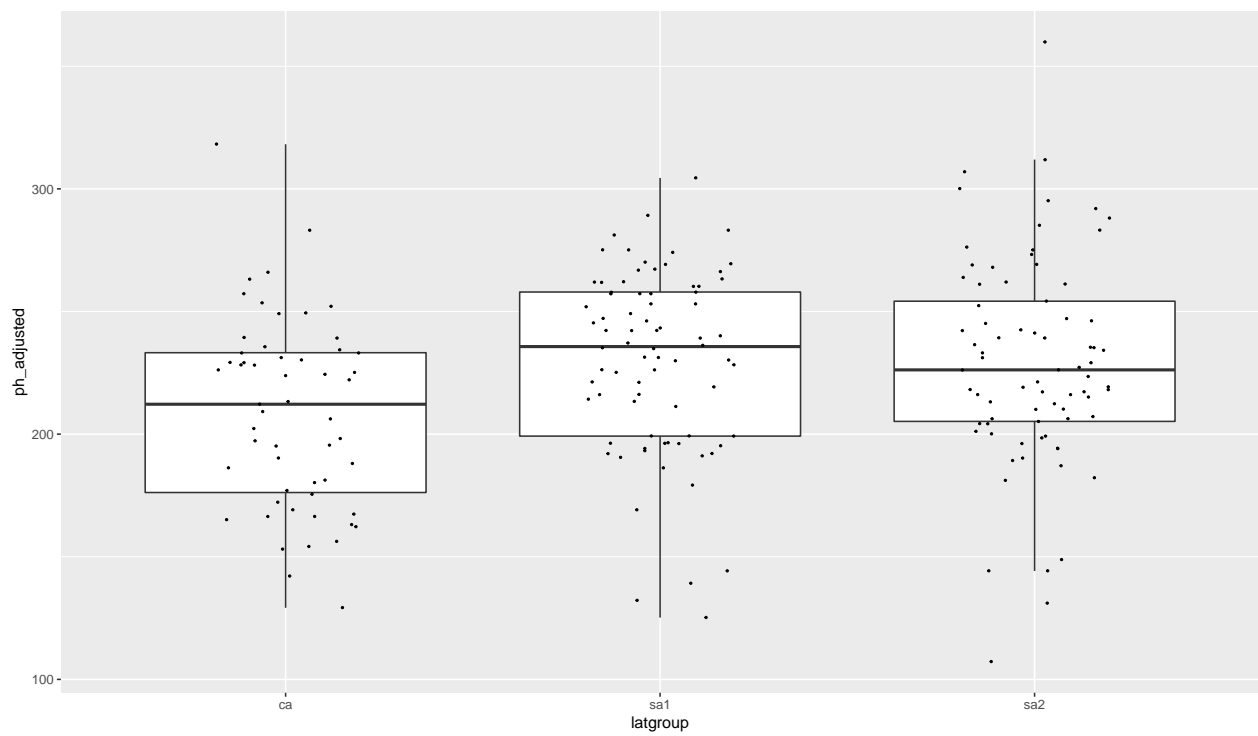
## Warning: Removed 16 rows containing non-finite values (stat_density).
```



```
ggplot(data = data_gbl, mapping = aes(x = latgroup, y = ph_adjusted)) +  
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```

Warning: Removed 16 rows containing non-finite values (stat_boxplot).

Warning: Removed 16 rows containing missing values (geom_point).



```
md_12 <-glm(ph_adjusted ~ latgroup , data = data_gbl, family="gaussian")
summary(md_12)
```

```
##
## Call:
## glm(formula = ph_adjusted ~ latgroup, family = "gaussian", data = data_gbl)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -121.838   -28.355    1.119    26.994   130.912
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  208.727      5.350  39.018 < 2e-16 ***
## latgroupsa1   21.479      6.985   3.075  0.00239 **
## latgroupsa2   20.310      7.004   2.900  0.00414 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 1573.951)
##
##      Null deviance: 343591  on 209  degrees of freedom
## Residual deviance: 325808  on 207  degrees of freedom
## (16 observations deleted due to missingness)
## AIC: 2146.8
##
## Number of Fisher Scoring iterations: 2
```

```
lrtest(md_12)
```

```
## Likelihood ratio test
##
## Model 1: ph_adjusted ~ latgroup
## Model 2: ph_adjusted ~ 1
##   #Df  LogLik Df  Chisq Pr(>Chisq)
## 1    4 -1069.4
## 2    2 -1075.0 -2  11.16   0.003772 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
lsmeans_latgroup_ph <- LSD.test(md_12, "latgroup", p.adj="none" )
lsmeans_latgroup_ph
```

```
## $statistics
##      MSerror Df      Mean      CV
##  1573.951 207  224.1524 17.69914
##
## $parameters
##      test p.adjusted name.t ntr alpha
## Fisher-LSD      none latgroup  3 0.05
##
## $means
##      ph_adjusted      std r      LCL      UCL  Min  Max  Q25  Q50  Q75
## ca      208.7273 38.75286 55 198.1808 219.2738 129.2 318.20 176.2 212.2 233.20
## sa1      230.2064 37.08467 78 221.3503 239.0625 125.2 304.45 199.2 235.7 257.95
```

```
## sa2      229.0377 42.73780 77 220.1242 237.9511 107.2 359.95 205.2 226.2 254.20
##
## $comparison
## NULL
##
## $groups
##      ph_adjusted groups
## sa1      230.2064      a
## sa2      229.0377      a
## ca       208.7273      b
##
## attr(,"class")
## [1] "group"
```

C. Seed setting

```
data_gbl$seed_setting <- as.numeric(data_gbl$seed_setting)
md_13 <- glm(seed_setting ~ latgroup , data = data_gbl, family = binomial(link = "logit"))
summary(md_13)
```

```
##
## Call:
## glm(formula = seed_setting ~ latgroup, family = binomial(link = "logit"),
##      data = data_gbl)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.0658  -0.3898  -0.1548  -0.1548   2.9769
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.2683     0.2605  -1.030   0.303
## latgroupsa1  -4.1506     1.0392  -3.994 6.50e-05 ***
## latgroupsa2  -2.2707     0.4977  -4.562 5.06e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 187.91  on 225  degrees of freedom
## Residual deviance: 135.89  on 223  degrees of freedom
## AIC: 141.89
##
## Number of Fisher Scoring iterations: 7
```

```
lrtest(md_13)
```

```
## Likelihood ratio test
##
## Model 1: seed_setting ~ latgroup
## Model 2: seed_setting ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    3 -67.943
## 2    1 -93.957 -2 52.027   5.04e-12 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

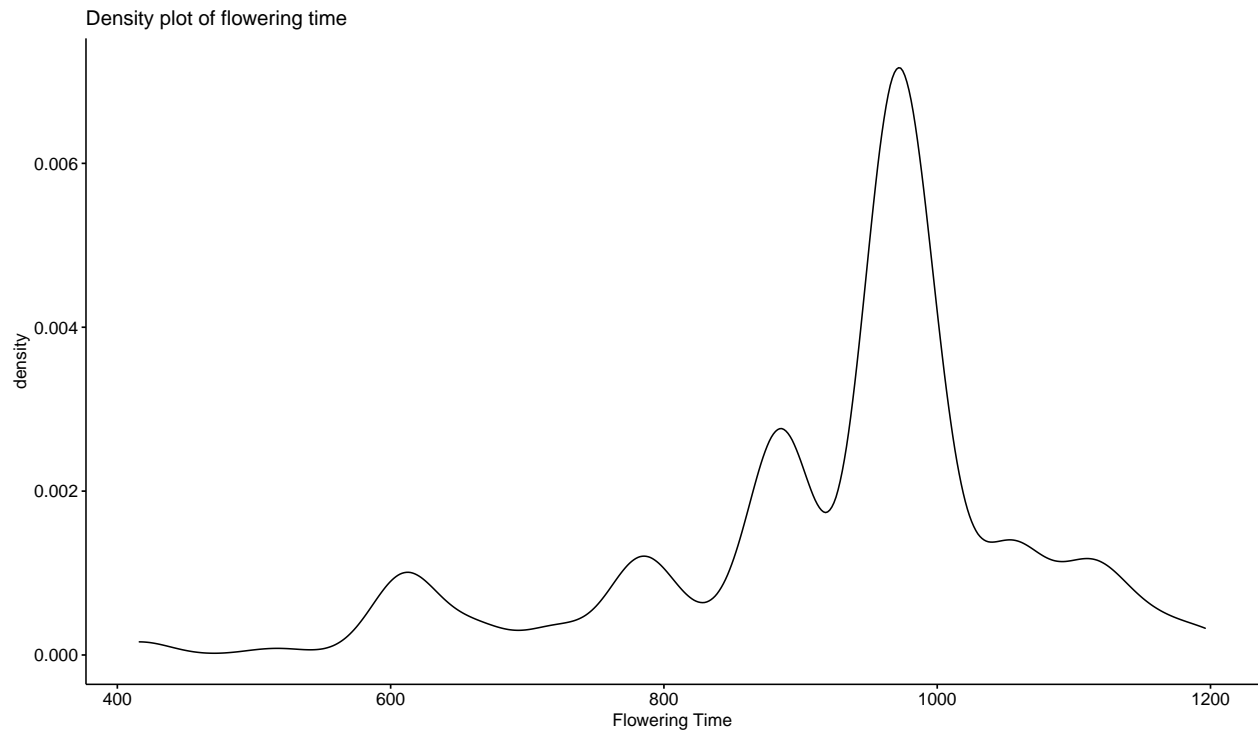
lsmeans_latgroup_ss <- LSD.test(md_13, "latgroup", p.adj="none" )
lsmeans_latgroup_ss

## $statistics
##      MSerror Df      Mean      CV
##    0.6093582 223 0.1460177 534.6023
##
## $parameters
##      test p.adjusted name.t ntr alpha
## Fisher-LSD      none latgroup  3  0.05
##
## $means
##      seed_setting      std r      LCL      UCL Min Max Q25 Q50 Q75
## ca      0.43333333 0.4997174 60  0.23473657 0.6319301  0  1  0  0  1
## sa1      0.01190476 0.1091089 84 -0.15594014 0.1797497  0  1  0  0  0
## sa2      0.07317073 0.2620191 82 -0.09670872 0.2430502  0  1  0  0  0
##
## $comparison
## NULL
##
## $groups
##      seed_setting groups
## ca      0.43333333      a
## sa2      0.07317073      b
## sa1      0.01190476      b
##
## attr(,"class")
## [1] "group"
```

D. Growing degree days (Tbase is 10°C)

```
ggdensity(data_gbl$GDD10c,
          main = "Density plot of flowering time",
          xlab = "Flowering Time")
```

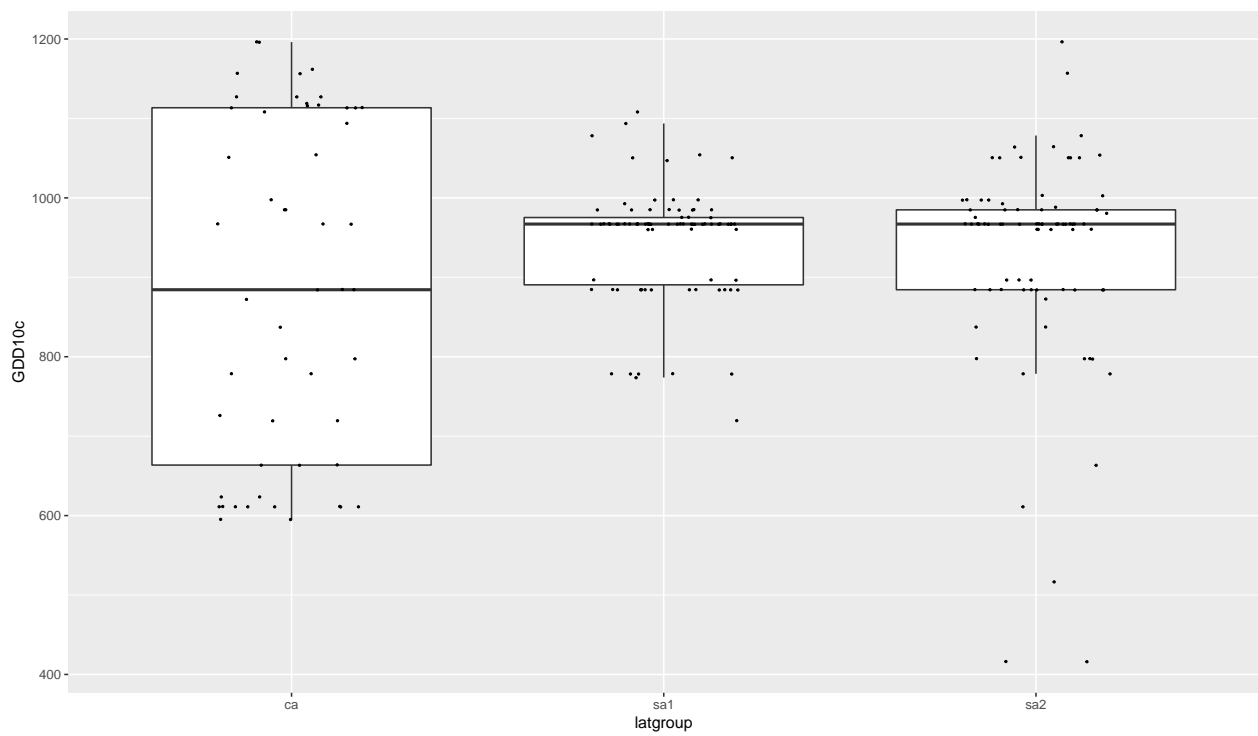
```
## Warning: Removed 10 rows containing non-finite values (stat_density).
```



```
ggplot(data = data_gbl, mapping = aes(x = latgroup, y = GDD10c)) +  
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```

Warning: Removed 10 rows containing non-finite values (stat_boxplot).

Warning: Removed 10 rows containing missing values (geom_point).



```

md_14 <-glm(GDD10c ~ latgroup , data = data_gbl, family = gaussian)
summary(md_14)

##
## Call:
## glm(formula = GDD10c ~ latgroup, family = gaussian, data = data_gbl)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -513.80   -61.21    21.44    55.05   307.22
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   888.93      19.42  45.772  <2e-16 ***
## latgroupsa1    56.63      24.77   2.287  0.0232 *
## latgroupsa2    40.92      24.89   1.644  0.1016
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 19612.81)
##
##      Null deviance: 4281953  on 215  degrees of freedom
## Residual deviance: 4177529  on 213  degrees of freedom
## (10 observations deleted due to missingness)
## AIC: 2752.9
##
## Number of Fisher Scoring iterations: 2
lrtest(md_14)

## Likelihood ratio test
##
## Model 1: GDD10c ~ latgroup
## Model 2: GDD10c ~ 1
##   #Df  LogLik Df  Chisq Pr(>Chisq)
## 1    4 -1372.5
## 2    2 -1375.1 -2  5.3329    0.0695 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
lsmeans_countries <- LSD.test(md_14, "latgroup", p.adj="none" )
lsmeans_countries

## $statistics
##      MSerror Df      Mean      CV
## 19612.81 213 926.0329 15.12319
##
## $parameters
##      test p.adjusted name.t ntr alpha
## Fisher-LSD      none latgroup 3 0.05
##
## $means
##      GDD10c      std  r      LCL      UCL      Min      Max      Q25      Q50
## ca 888.9250 214.10100 52 850.6433 927.2067 594.90 1196.15 663.600 884.35
## sa1 945.5578 72.53505 83 915.2571 975.8586 719.35 1107.95 890.525 967.00

```

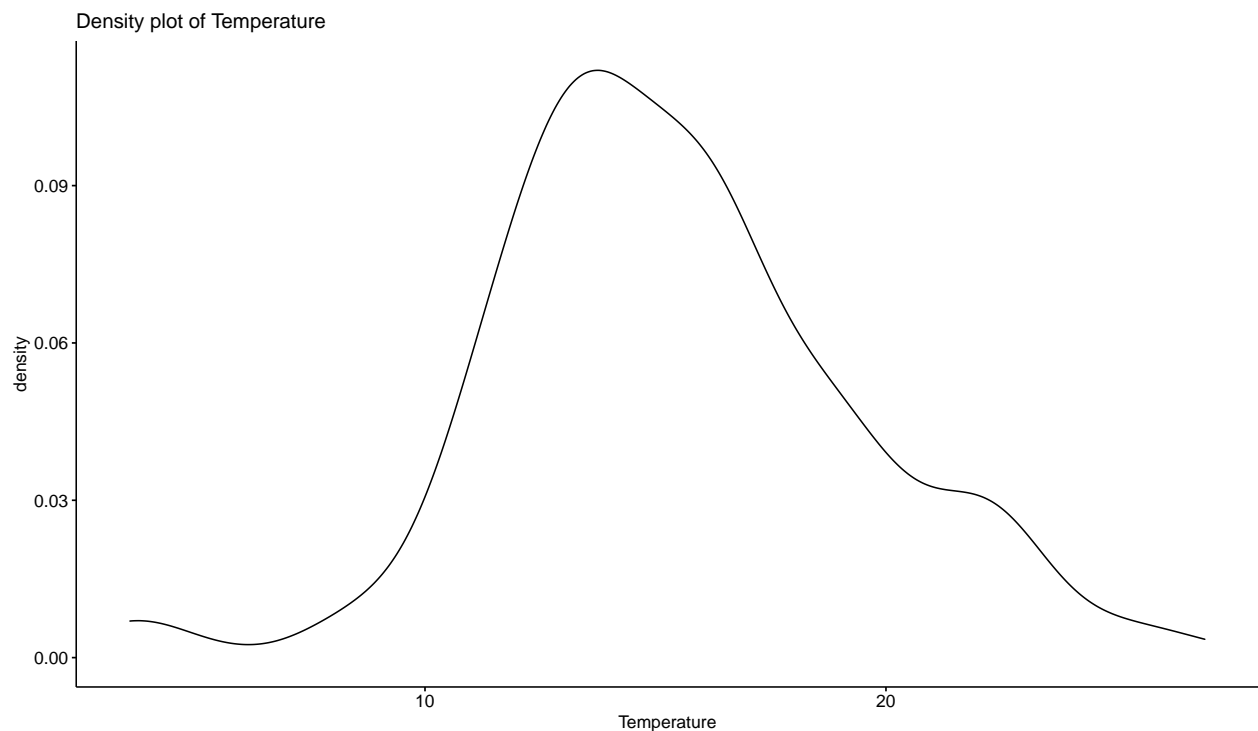


```
## sa2 929.8481 132.67904 81 899.1756 960.5207 416.05 1196.15 884.350 967.00
##      Q75
## ca  1113.45
## sa1  975.25
## sa2  984.90
##
## $comparison
## NULL
##
## $groups
##      GDD10c groups
## sa1 945.5578      a
## sa2 929.8481      ab
## ca  888.9250      b
##
## attr("class")
## [1] "group"
```

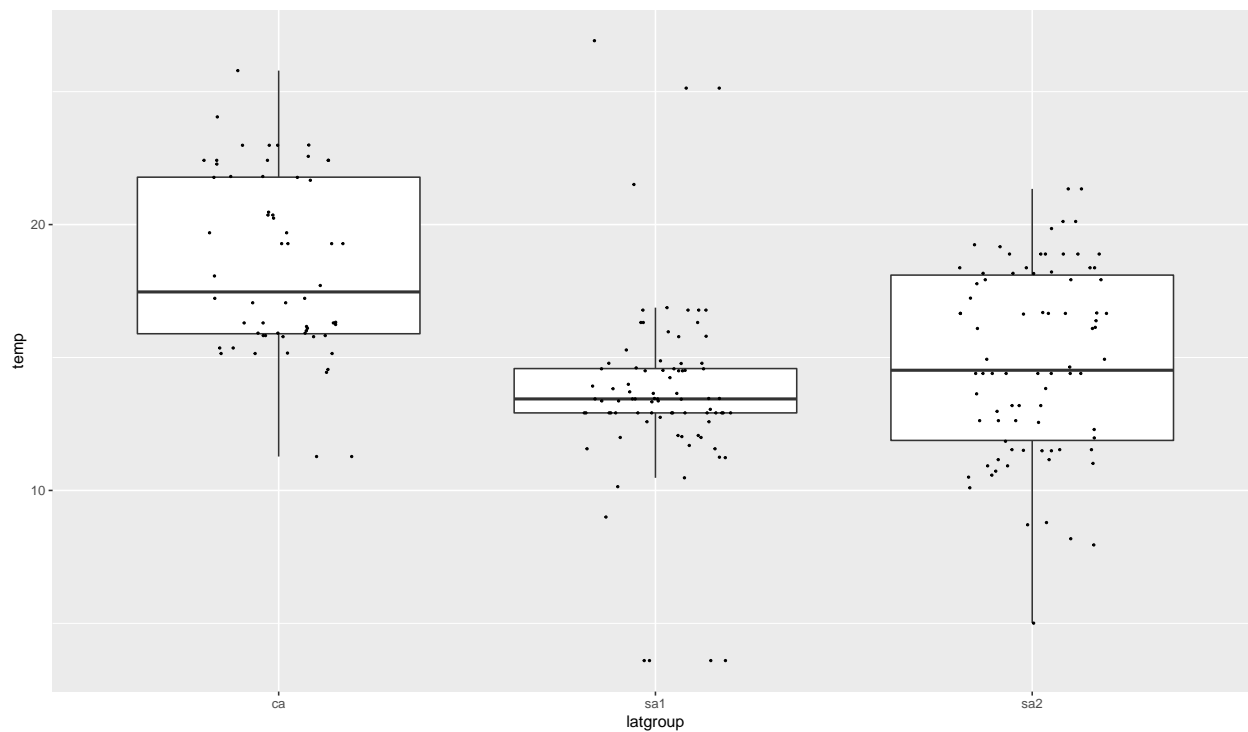
Climatic & Geographical Variables

A. Temperature

```
ggdensity(data_gbl$temp,
  main = "Density plot of Temperature",
  xlab = "Temperature")
```



```
ggplot(data = data_gbl, mapping = aes(x = latgroup, y = temp)) +
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_15 <- glm(temp ~ latgroup, data = data_gbl, family = gaussian)
summary(md_15)
```

```
##
## Call:
## glm(formula = temp ~ latgroup, family = gaussian, data = data_gbl)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -10.0756  -2.2360  -0.2756   2.2494  13.2411
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  18.4872    0.4543  40.691 < 2e-16 ***
## latgroupsa1  -4.8116    0.5949  -8.089 3.85e-14 ***
## latgroupsa2  -3.6262    0.5979  -6.065 5.58e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 12.38515)
##
##      Null deviance: 3607.3  on 225  degrees of freedom
## Residual deviance: 2761.9  on 223  degrees of freedom
## AIC: 1215.1
##
## Number of Fisher Scoring iterations: 2
```

```
lrtest(md_15)
```

```
## Likelihood ratio test
##
## Model 1: temp ~ latgroup
```

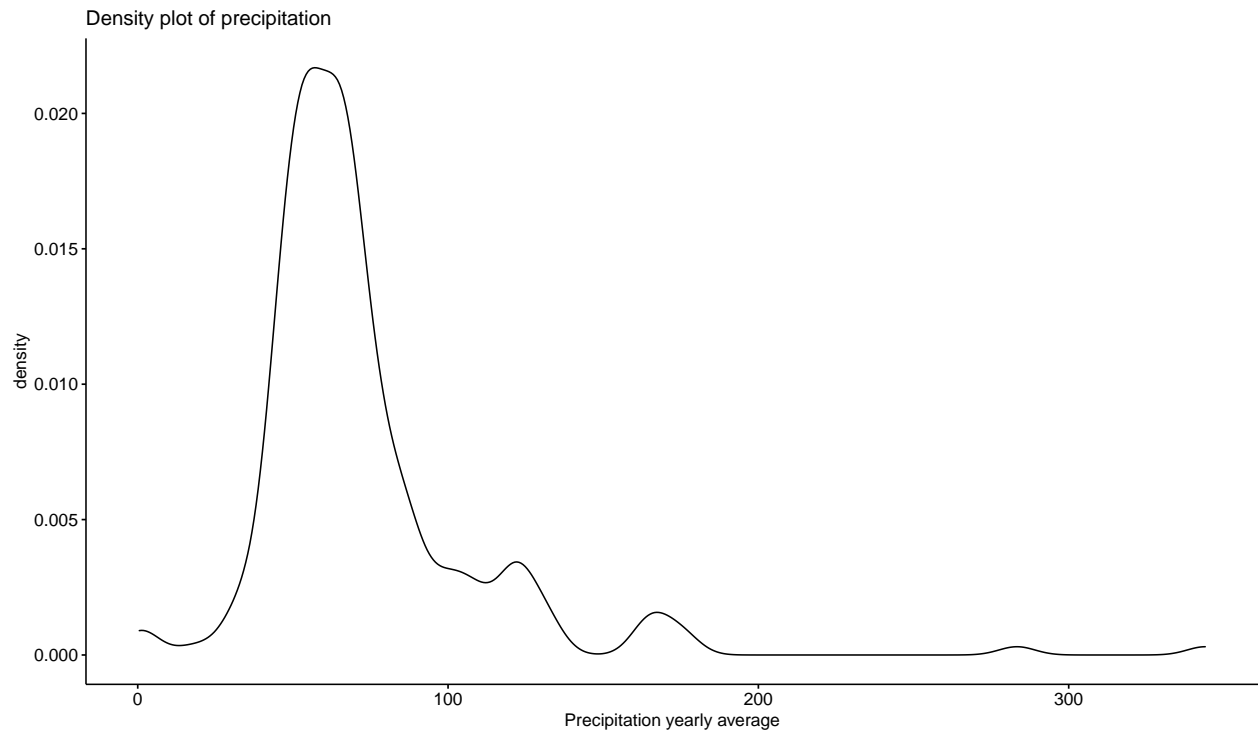
```
## Model 2: temp ~ 1
##   #Df LogLik Df   Chisq Pr(>Chisq)
## 1    4 -603.53
## 2    2 -633.71 -2 60.351  7.851e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

lsmeans_countries_temp <- LSD.test(md_15, "latgroup", p.adj="none" )
lsmeans_countries_temp

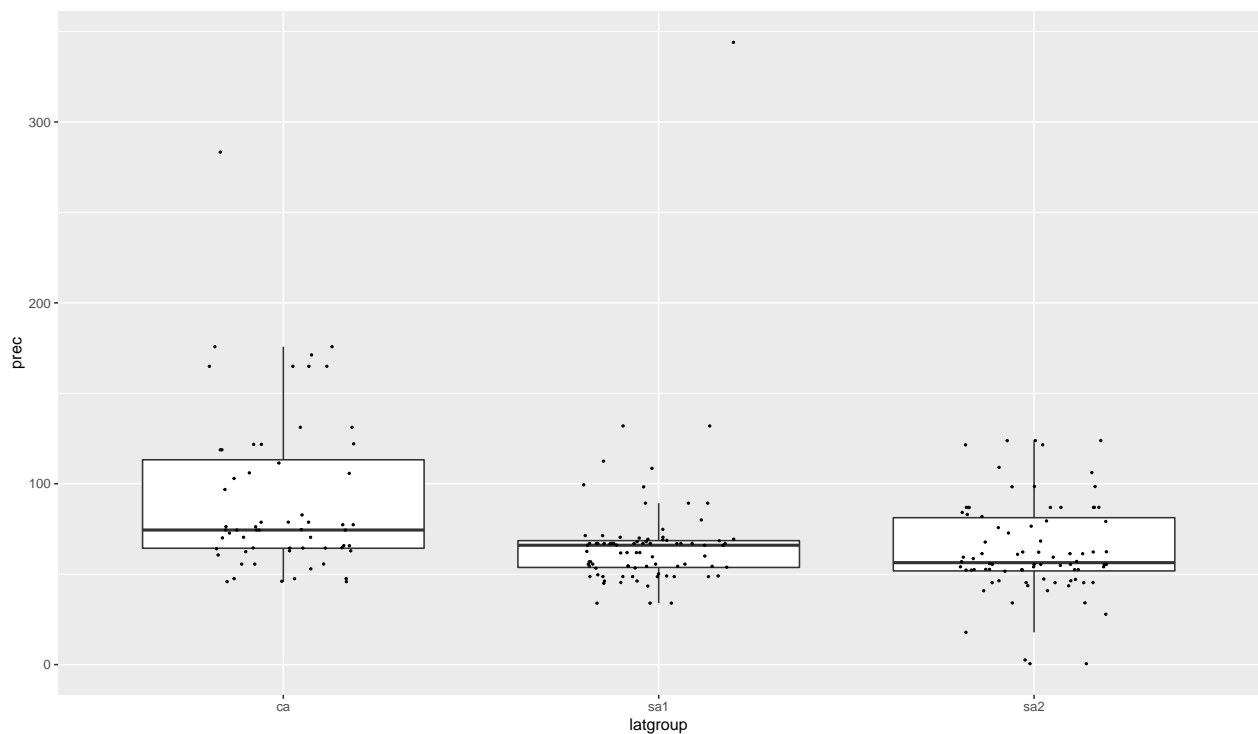
## $statistics
##      MSerror Df      Mean      CV
##    12.38515 223 15.38311 22.87739
##
## $parameters
##      test p.adjusted name.t ntr alpha
## Fisher-LSD      none latgroup  3 0.05
##
## $means
##      temp      std  r      LCL      UCL      Min      Max      Q25      Q50
## ca  18.48722 3.323327 60 17.59189 19.38256 11.275000 25.79167 15.89375 17.46667
## sa1 13.67560 3.640788 84 12.91890 14.43229  3.600000 26.91667 12.91667 13.44167
## sa2 14.86098 3.531290 82 14.09511 15.62685  5.008333 21.34167 11.88125 14.52083
##      Q75
## ca  21.78333
## sa1 14.58333
## sa2 18.10000
##
## $comparison
## NULL
##
## $groups
##      temp groups
## ca  18.48722    a
## sa2 14.86098    b
## sa1 13.67560    c
##
## attr(,"class")
## [1] "group"
```

B. Precipitation

```
ggdensity(data_gbl$prec,
          main = "Density plot of precipitation",
          xlab = "Precipitation yearly average")
```



```
ggplot(data = data_gbl, mapping = aes(x = latgroup, y = prec)) +
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_16 <-glm(prec ~ latgroup , data = data_gbl, family = gaussian)
summary(md_16)
```

```
##
## Call:
```

```
## glm(formula = prec ~ latgroup, family = gaussian, data = data_gbl)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -62.961  -18.128   -8.045    6.807   276.478
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   91.289      4.561  20.015 < 2e-16 ***
## latgroupsa1  -23.767      5.972   -3.980 9.33e-05 ***
## latgroupsa2  -27.828      6.002   -4.636 6.04e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 1248.233)
##
##      Null deviance: 308314  on 225  degrees of freedom
## Residual deviance: 278356  on 223  degrees of freedom
## AIC: 2257.6
##
## Number of Fisher Scoring iterations: 2
```

```
lrtest(md_16)
```

```
## Likelihood ratio test
##
## Model 1: prec ~ latgroup
## Model 2: prec ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    4 -1124.8
## 2    2 -1136.3 -2 23.101  9.631e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

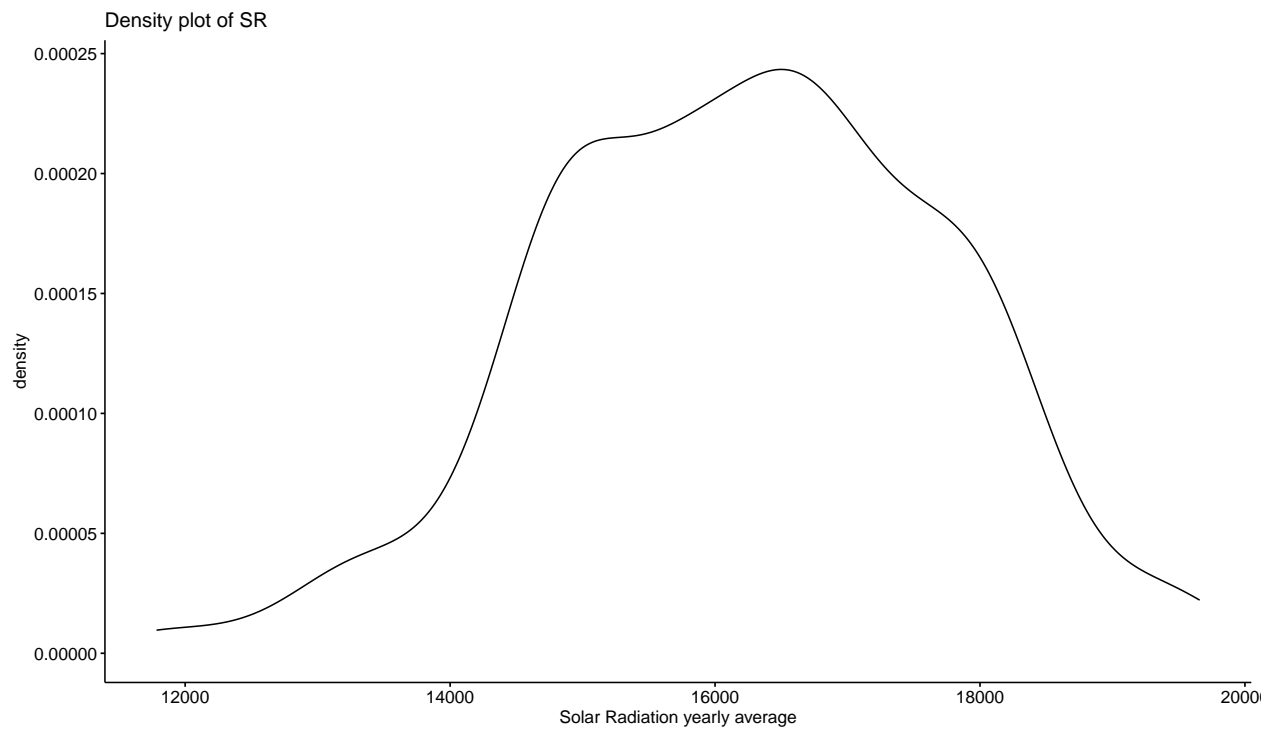
```
lsmeans_latgroup_prec <- LSD.test(md_16, "latgroup", p.adj="none" )
lsmeans_latgroup_prec
```

```
## $statistics
##      MSerror Df      Mean      CV
## 1248.233 223 72.35841 48.82686
##
## $parameters
##      test p.adjusted name.t ntr alpha
## Fisher-LSD      none latgroup 3 0.05
##
## $means
##      prec      std r      LCL      UCL      Min      Max      Q25      Q50
## ca  91.28889 44.85762 60 82.30046 100.27731 45.83333 283.3333 64.33333 74.41667
## sa1 67.52183 35.47180 84 59.92522 75.11843 34.00000 344.0000 53.72917 66.00000
## sa2 63.46138 26.10546 82 55.77269 71.15007 0.50000 123.8333 51.81250 56.37500
##      Q75
## ca 113.25000
## sa1 68.54167
## sa2 81.22917
##
```

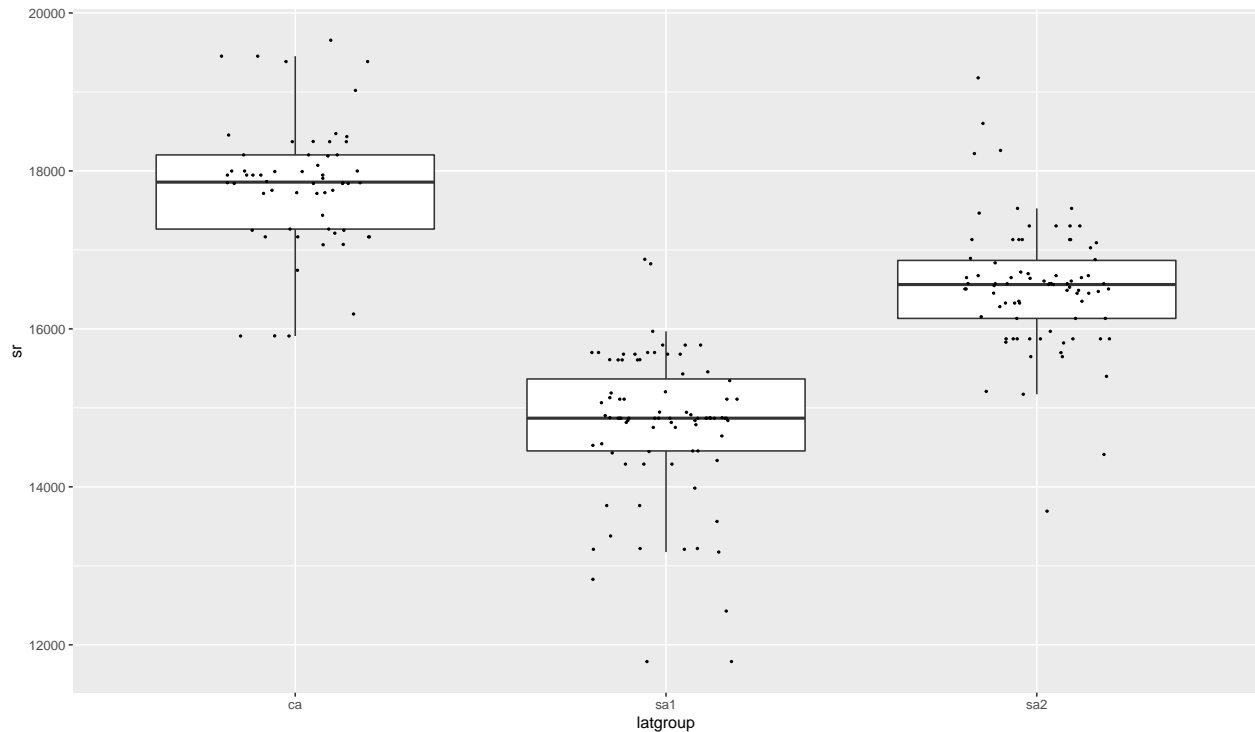
```
## $comparison
## NULL
##
## $groups
##      prec groups
## ca  91.28889    a
## sa1 67.52183    b
## sa2 63.46138    b
##
## attr(,"class")
## [1] "group"
```

C. Solar Radiation

```
ggdensity(data_gbl$sr,
  main = "Density plot of SR",
  xlab = "Solar Radiation yearly average")
```



```
ggplot(data = data_gbl, mapping = aes(x = latgroup, y = sr)) +
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_17 <-glm(sr ~ latgroup , data = data_gbl, family="gaussian")
summary(md_17)
```

```
##
## Call:
## glm(formula = sr ~ latgroup, family = "gaussian", data = data_gbl)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2984.00  -389.61    70.62   369.67  2650.38
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  17833.1      109.9  162.254  <2e-16 ***
## latgroupsa1  -3061.3      143.9  -21.273  <2e-16 ***
## latgroupsa2  -1305.4      144.6   -9.025  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 724791.6)
##
##      Null deviance: 501678361  on 225  degrees of freedom
## Residual deviance: 161628538  on 223  degrees of freedom
## AIC: 3695.9
##
## Number of Fisher Scoring iterations: 2
```

```
lrtest(md_17)
```

```
## Likelihood ratio test
##
```

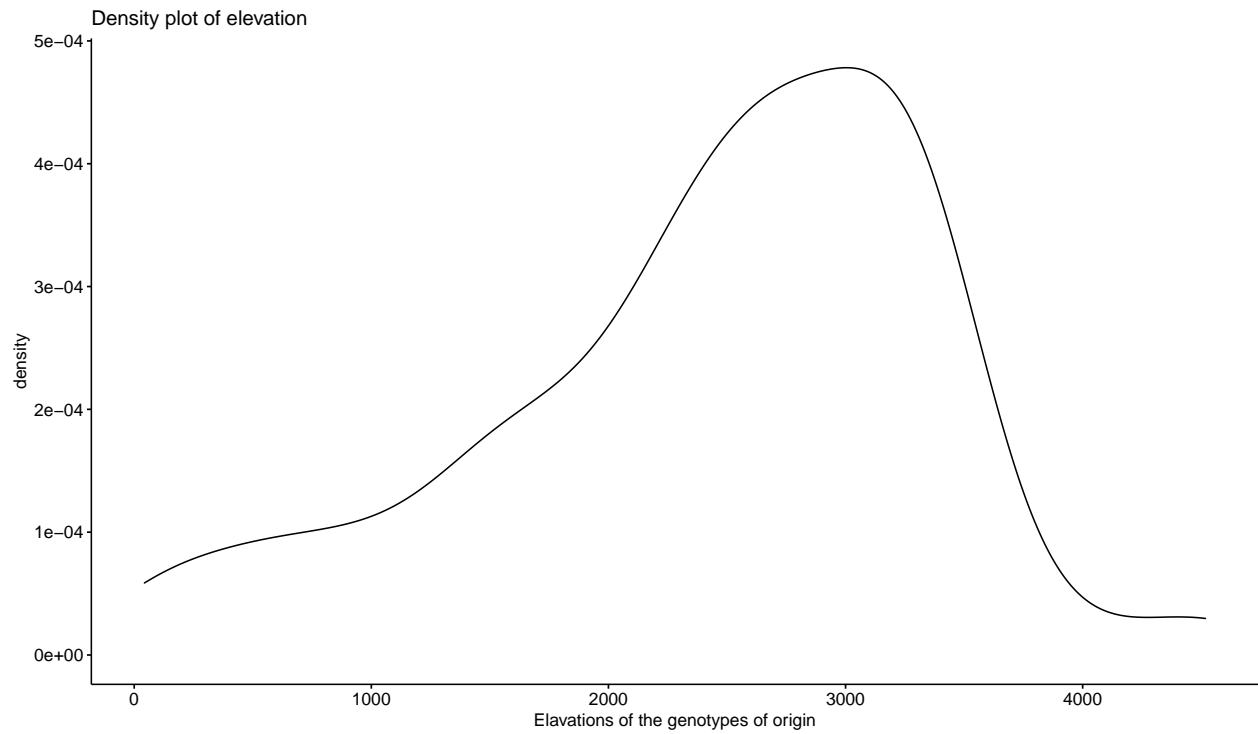
```
## Model 1: sr ~ latgroup
## Model 2: sr ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    4 -1844.0
## 2    2 -1971.9 -2 255.98 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

lsmeans_latgroup_sr <- LSD.test(md_17, "latgroup", p.adj="none" )
lsmeans_latgroup_sr

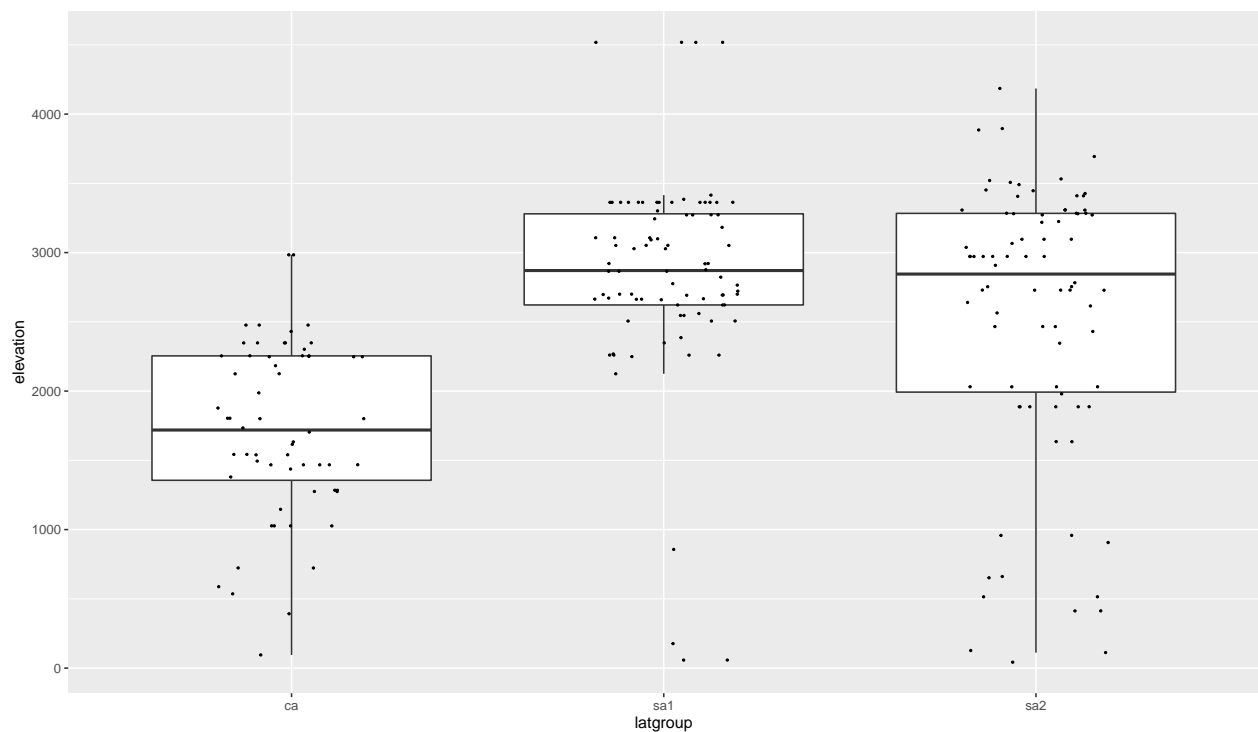
## $statistics
##      MSerror Df      Mean      CV
## 724791.6 223 16221.64 5.248217
##
## $parameters
##      test p.adjusted name.t ntr alpha
## Fisher-LSD      none latgroup 3 0.05
##
## $means
##      sr      std r      LCL      UCL      Min      Max      Q25      Q50
## ca 17833.08 800.1958 60 17616.49 18049.68 15909.75 19654.75 17263.58 17858.29
## sa1 14771.83 936.9841 84 14588.78 14954.89 11787.83 16881.00 14454.67 14869.42
## sa2 16527.71 793.3449 82 16342.43 16712.98 13692.25 19178.08 16132.50 16561.50
##      Q75
## ca 18202.75
## sa1 15365.98
## sa2 16866.92
##
## $comparison
## NULL
##
## $groups
##      sr groups
## ca 17833.08 a
## sa2 16527.71 b
## sa1 14771.83 c
##
## attr(,"class")
## [1] "group"
```

D. Elevation

```
ggdensity(data_gbl$elevation,
          main = "Density plot of elevation",
          xlab = "Elavations of the genotypes of origin")
```

```
ggplot(data = data_gbl, mapping = aes(x = latgroup, y = elevation)) +
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_18 <-glm(elevation ~ latgroup , data =data_gbl, family = "gaussian")
summary(md_18)
```

```
##
## Call:
```

```
## glm(formula = elevation ~ latgroup, family = "gaussian", data = data_gbl)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2791.23   -332.41    93.44   522.84  1668.77
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   1722.1      107.2   16.057 < 2e-16 ***
## latgroupsa1   1127.1      140.4    8.026 5.71e-14 ***
## latgroupsa2    812.9      141.1    5.759 2.78e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 690146.7)
##
##      Null deviance: 199626032  on 225  degrees of freedom
## Residual deviance: 153902717  on 223  degrees of freedom
## AIC: 3684.8
##
## Number of Fisher Scoring iterations: 2

lrtest(md_18)

## Likelihood ratio test
##
## Model 1: elevation ~ latgroup
## Model 2: elevation ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    4 -1838.4
## 2    2 -1867.8 -2 58.788  1.715e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

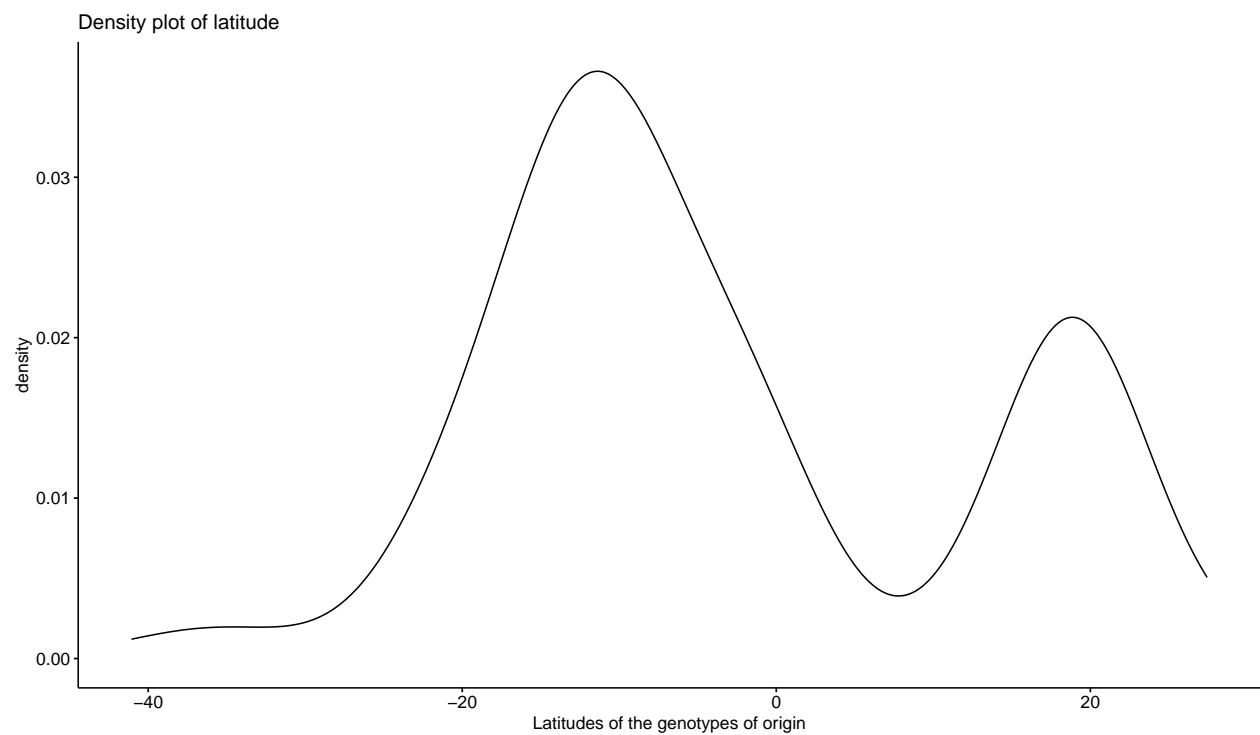
lsmeans_species_elevation <- LSD.test(md_18, "latgroup", p.adj="none" )
lsmeans_species_elevation

## $statistics
##      MSerror Df      Mean      CV
## 690146.7 223 2435.982 34.10331
##
## $parameters
##      test p.adjusted  name.t ntr alpha
## Fisher-LSD      none latgroup  3 0.05
##
## $means
##      elevation      std r      LCL      UCL Min  Max      Q25      Q50      Q75
## ca  1722.133  630.7224 60 1510.781 1933.485  95 2984 1356.00 1719.0 2254.25
## sa1  2849.226  761.5455 84 2670.601 3027.851  58 4518 2622.00 2870.5 3280.00
## sa2  2534.988 1007.9677 82 2354.197 2715.778  43 4185 1992.75 2845.0 3283.25
##
## $comparison
## NULL
##
## $groups
```

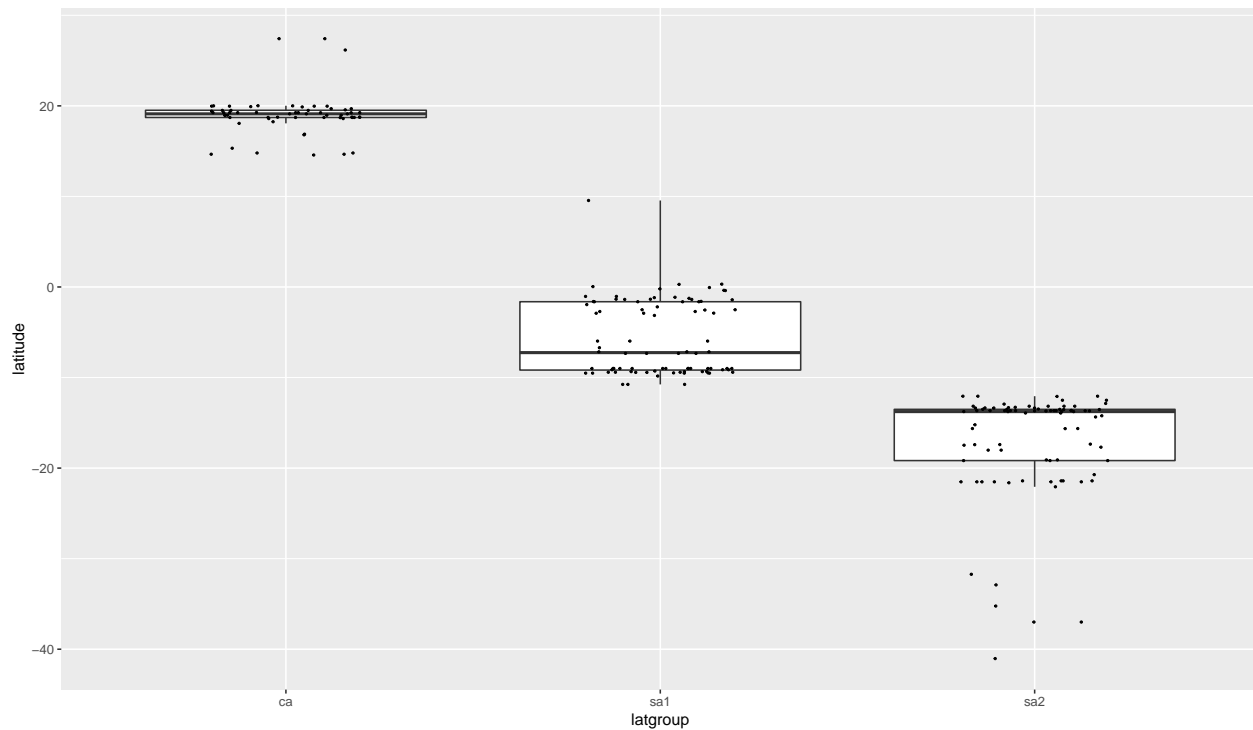
```
##      elevation groups
## sa1  2849.226      a
## sa2  2534.988      b
## ca   1722.133      c
##
## attr(,"class")
## [1] "group"
```

E. Latitude

```
ggdensity(data_gbl$latitude,
  main = "Density plot of latitude",
  xlab = "Latitudes of the genotypes of origin")
```



```
ggplot(data = data_gbl, mapping = aes(x = latgroup, y = latitude)) +
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_19 <- glm(latitude ~ latgroup, data = data_gbl, family = gaussian)
summary(md_19)
```

```
##
## Call:
## glm(formula = latitude ~ latgroup, family = gaussian, data = data_gbl)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -23.9120  -3.3662   0.1697   3.4637  15.1842
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  19.0693    0.6002   31.77  <2e-16 ***
## latgroupsa1 -24.7031    0.7859  -31.43  <2e-16 ***
## latgroupsa2 -36.1919    0.7899  -45.82  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 21.61663)
##
##      Null deviance: 50967.7  on 225  degrees of freedom
## Residual deviance:  4820.5  on 223  degrees of freedom
## AIC: 1340.9
##
## Number of Fisher Scoring iterations: 2
```

```
lrtest(md_19)
```

```
## Likelihood ratio test
##
## Model 1: latitude ~ latgroup
```

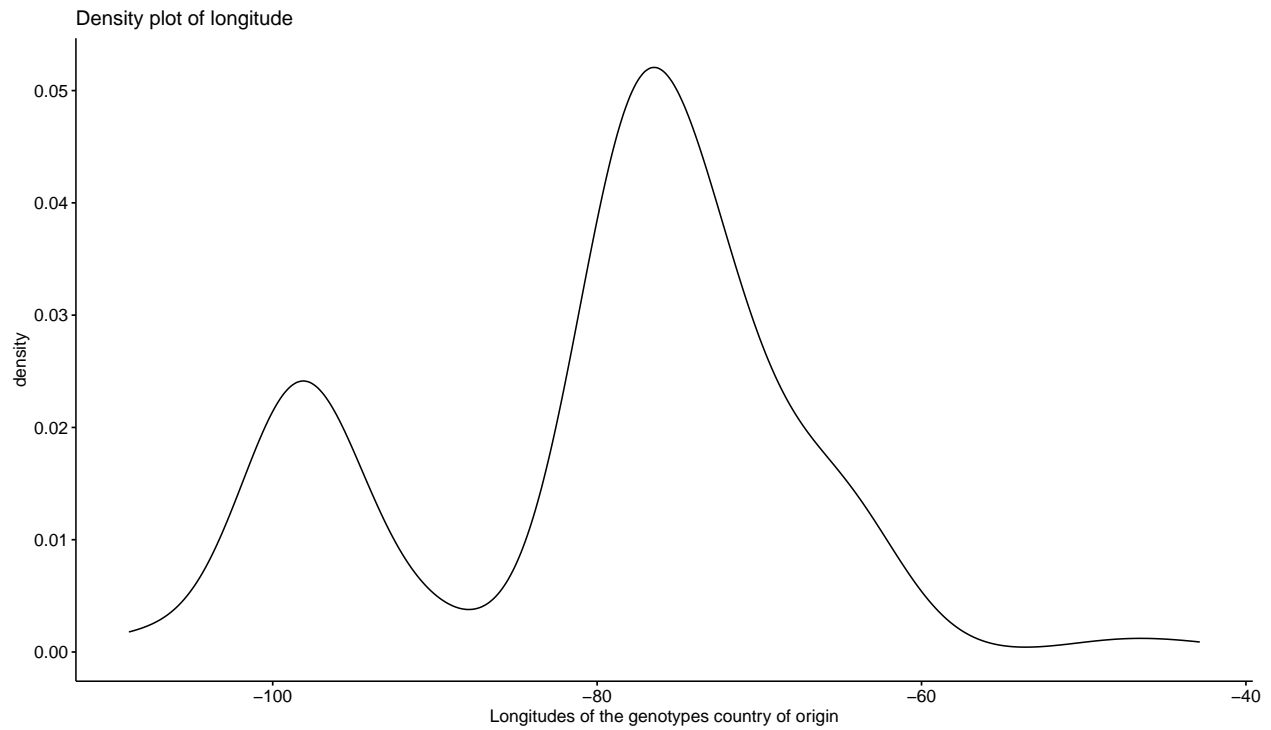
```
## Model 2: latitude ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    4 -666.47
## 2    2 -932.96 -2 532.98 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

lsmeans_latgroup_latitude <- LSD.test(md_19, "latgroup", p.adj="none" )
lsmeans_latgroup_latitude

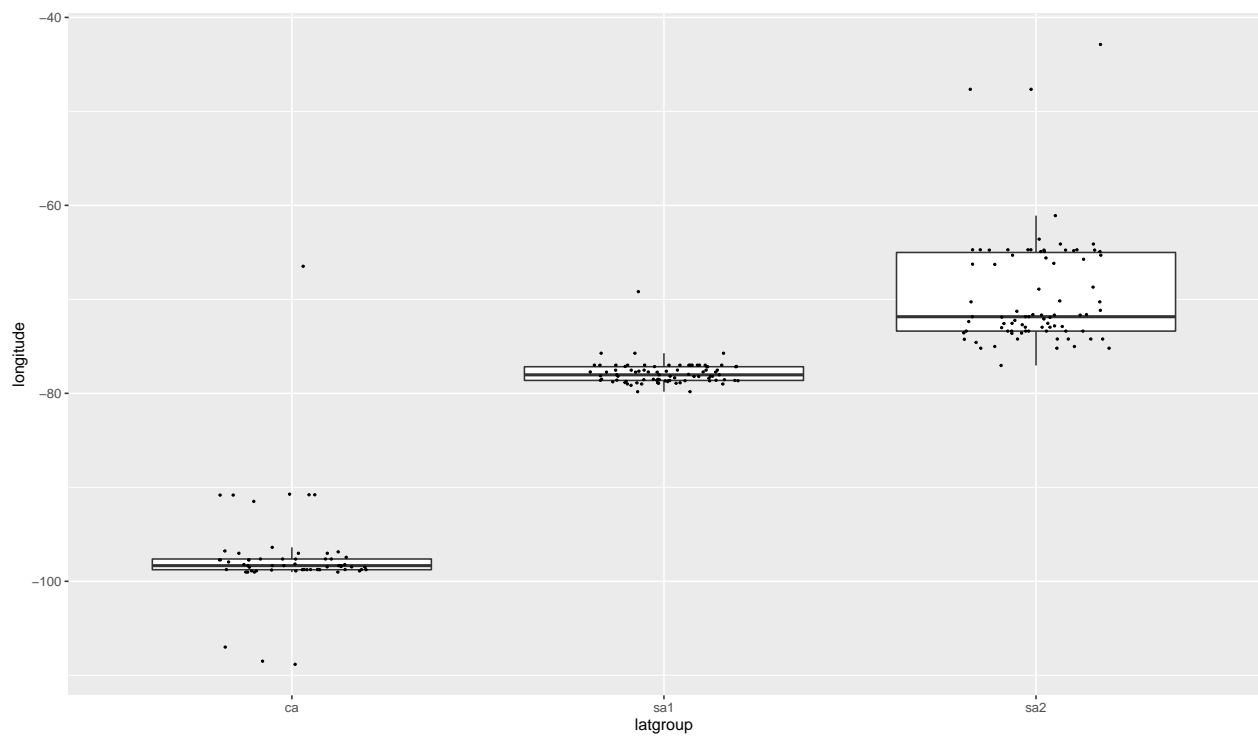
## $statistics
##      MSerror Df      Mean      CV
##    21.61663 223 -3.243948 -143.3244
##
## $parameters
##      test p.adjusted name.t ntr alpha
## Fisher-LSD      none latgroup  3 0.05
##
## $means
##      latitude      std r      LCL      UCL      Min      Max      Q25
## ca  19.069313 2.336447 60  17.886462  20.252163  14.57172  27.416667  18.71667
## sa1  -5.633817 4.098526 84  -6.633508  -4.634126 -10.75909   9.550365  -9.17522
## sa2 -17.122566 6.190596 82 -18.134375 -16.110757 -41.03454 -12.061098 -19.17081
##      Q50      Q75
## ca  19.116667 19.513842
## sa1  -7.248466 -1.628928
## sa2 -13.761848 -13.511656
##
## $comparison
## NULL
##
## $groups
##      latitude groups
## ca  19.069313      a
## sa1  -5.633817      b
## sa2 -17.122566      c
##
## attr(,"class")
## [1] "group"
```

F. Longitude

```
ggdensity(data_gbl$longitude,
          main = "Density plot of longitude",
          xlab = "Longitudes of the genotypes country of origin")
```



```
ggplot(data = data_gbl, mapping = aes(x = latgroup, y = longitude)) +
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_20 <-glm(longitude ~ latgroup , data = data_gbl, family = gaussian)
summary(md_20)
```

```
##
## Call:
```

```

## glm(formula = longitude ~ latgroup, family = gaussian, data = data_gbl)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -11.3810   -1.5644   -0.7279    0.7087   30.9734
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -97.4523     0.5939 -164.09  <2e-16 ***
## latgroupsa1  19.6006     0.7776  25.21  <2e-16 ***
## latgroupsa2  28.0364     0.7815  35.87  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 21.16251)
##
##      Null deviance: 32567.7  on 225  degrees of freedom
## Residual deviance:  4719.2  on 223  degrees of freedom
## AIC: 1336.1
##
## Number of Fisher Scoring iterations: 2

```

```

lrtest(md_20)

```

```

## Likelihood ratio test
##
## Model 1: longitude ~ latgroup
## Model 2: longitude ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    4 -664.07
## 2    2 -882.35 -2 436.56 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

lsmeans_latgroup_longitude <- LSD.test(md_20, "latgroup", p.adj="none" )
lsmeans_latgroup_longitude

```

```

## $statistics
##      MSerror Df      Mean      CV
## 21.16251 223 -79.99466 -5.750726
##
## $parameters
##      test p.adjusted name.t ntr alpha
## Fisher-LSD      none latgroup  3 0.05
##
## $means
##      longitude      std r      LCL      UCL      Min      Max      Q25
## ca -97.45231 5.226773 60 -98.62267 -96.28195 -108.83333 -66.47890 -98.76667
## sa1 -77.85174 1.279016 84 -78.84087 -76.86260 -79.82561 -69.18047 -78.62430
## sa2 -69.41595 6.056964 82 -70.41707 -68.41482 -77.03616 -42.87518 -73.37850
##      Q50      Q75
## ca -98.33414 -97.62778
## sa1 -78.03183 -77.16731
## sa2 -71.84630 -65.01160
##

```

```
## $comparison
## NULL
##
## $groups
##      longitude groups
## sa2 -69.41595      a
## sa1 -77.85174      b
## ca  -97.45231      c
##
## attr("class")
## [1] "group"
```

3. Correlation analysis

```
library(PerformanceAnalytics)
```

```
## Loading required package: xts
##
## Attaching package: 'xts'
## The following objects are masked from 'package:dplyr':
##
##      first, last
##
## Attaching package: 'PerformanceAnalytics'
## The following objects are masked from 'package:agricolae':
##
##      kurtosis, skewness
## The following object is masked from 'package:graphics':
##
##      legend
```

```
dim(data)
```

```
## [1] 254  14
```

I perform correlation analyses using two methods and four different datasets. The methods are Pearson's and Spearman's correlations. If seed setting is involved, Spearman's correlation should be considered. Otherwise, Pearson's.

Datasets are: + Dataset I: All accessions (n=254) + Dataset II: Only within South American accessions (n=181) + Dataset III: Only within Central American group (n=67) + Dataset IV: Only the hypochondriacus accessions within Central American Group (n=32)

Dataset I: All accessions

```
names(data)
```

```
## [1] "genotypes" "species" "country" "plot" "seed_setting"
## [6] "ft_adjusted" "ph_adjusted" "longitude" "latitude" "elevation"
## [11] "prec" "temp" "sr" "GDD10c"
```



```
correlation_all_accessions <- dplyr::select(data, seed_setting, ft_adjusted, ph_adjusted, longitude, latitude, elevation, prec)
correlation_all_accessions$longitude <- abs(correlation_all_accessions$longitude)
correlation_all_accessions$latitude <- abs(correlation_all_accessions$latitude)
head(correlation_all_accessions)
```

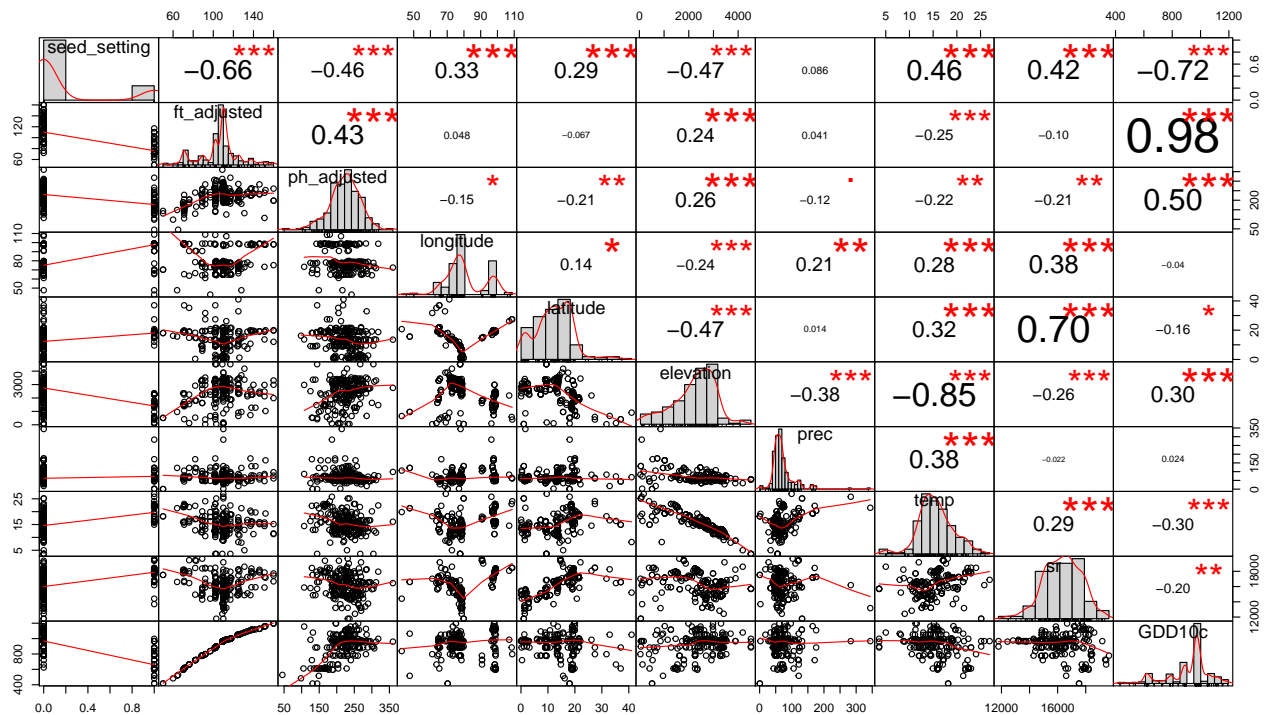
```
##   seed_setting ft_adjusted ph_adjusted longitude latitude elevation    prec
## 1           1           70      188.95         NA         NA         NA     NA
## 2           0           124      234.95  77.00000  9.00000     3363 67.00000
## 3           0           101      306.95  73.58656 13.69798     3219 79.41667
## 4           0           124      196.95         NA         NA         NA     NA
## 5           1           70      187.95  98.50000 18.75000     1540 70.41667
## 6           0           109      291.95  74.22260 13.16455     2729 45.33333

##      temp      sr  GDD10c
## 1      NA      NA  594.90
## 2 12.91667 14869.42 1047.15
## 3 12.55833 16153.75  872.40
## 4      NA      NA 1047.15
## 5 21.77500 17724.42  594.90
## 6 16.65833 16572.17  960.30
```

```
dim(correlation_all_accessions)
```

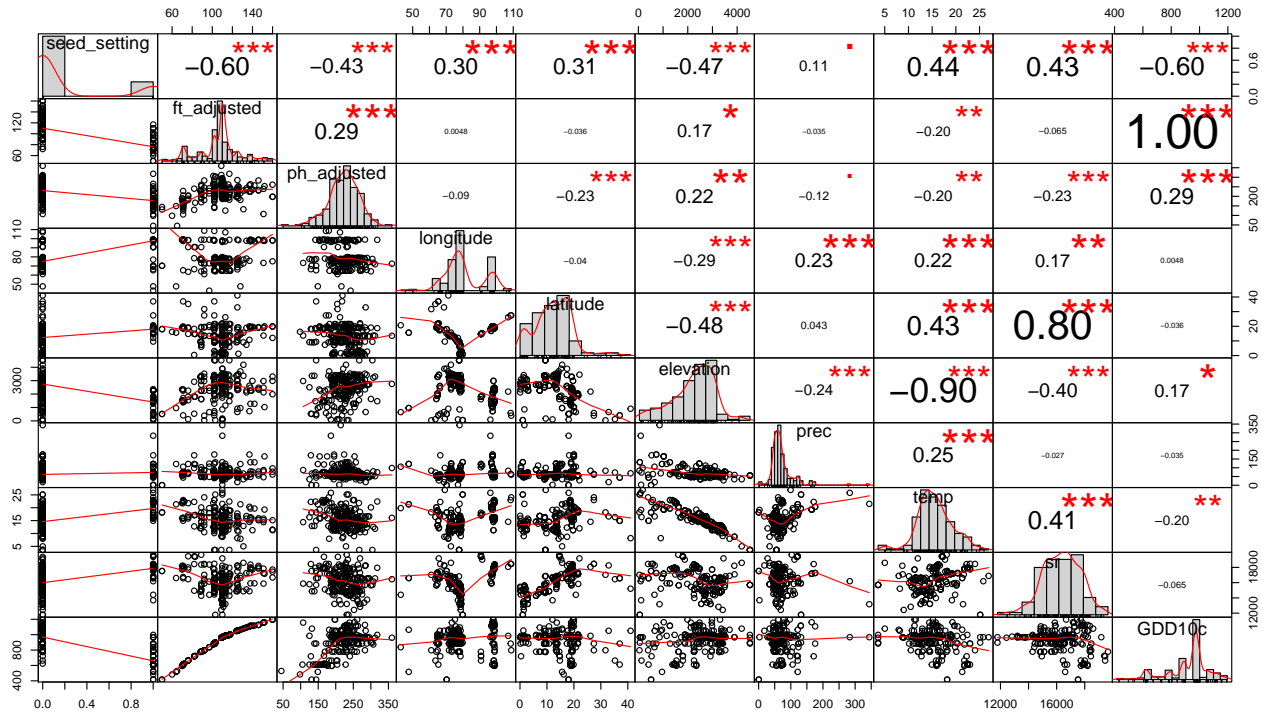
```
## [1] 254  10
```

```
chart.Correlation(
  correlation_all_accessions,
  histogram = TRUE,
  method = c("pearson"))
```



```
chart.Correlation(
  correlation_all_accessions,
  histogram = TRUE,
```

```
method = c("spearman"))
```



Dataset II: South American Accessions

Filtering the South American countries

```
table(data$country)
```

```
##
## ARG BOL BRA CHL ECU GER GTM MEX PER PTR USA VEN
## 7 28 5 1 35 1 6 60 104 1 5 1
correlation_south_american <- filter(data, !(country %in% c("GER", "GTM", "MEX", "USA", "PTR")))
dim(correlation_south_american)
```

```
## [1] 181 14
```

```
table(correlation_south_american$country)
```

```
##
## ARG BOL BRA CHL ECU PER VEN
## 7 28 5 1 35 104 1
```

```
correlation_south_american$longitude <- abs(correlation_south_american$longitude)
correlation_south_american$latitude <- abs(correlation_south_american$latitude)
head(correlation_south_american)
```

```
## genotypes species country plot seed_setting ft_adjusted ph_adjusted longitude
## 1 cau_339_y cau PER 102 0 124 234.95 77.00000
## 2 cau_314_y cau PER 103 0 101 306.95 73.58656
## 3 cru_73_n cru PER 105 0 124 196.95 NA
## 4 cau_32_n cau PER 108 0 109 291.95 74.22260
## 5 cau_281_y cau PER 110 0 88 229.95 77.13901
```

```
## 6  cau_59_y      cau      PER  111          0          112      267.95  74.25025
##   latitude elevation      prec      temp      sr  GDD10c
## 1  9.000000      3363 67.00000 12.91667 14869.42 1047.15
## 2 13.697981      3219 79.41667 12.55833 16153.75  872.40
## 3      NA      NA      NA      NA      NA 1047.15
## 4 13.164547      2729 45.33333 16.65833 16572.17  960.30
## 5  9.486735      3273 70.41667 12.06667 14815.58  773.80
## 6 12.931580      2640 40.83333 16.69167 16699.08  980.30
```

```
dim(correlation_south_american)
```

```
## [1] 181  14
```

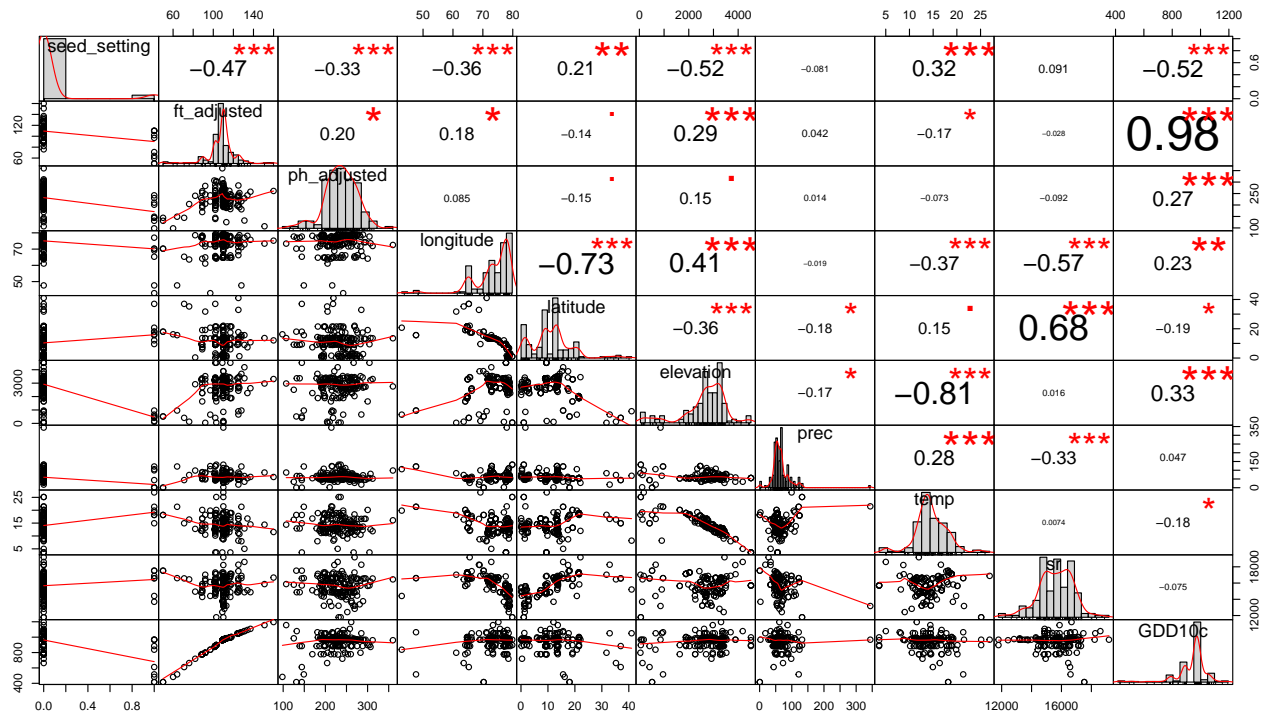
```
correlation_south_american <- dplyr::select(correlation_south_american, seed_setting, ft_adjusted, ph_a
```

```
str(correlation_south_american)
```

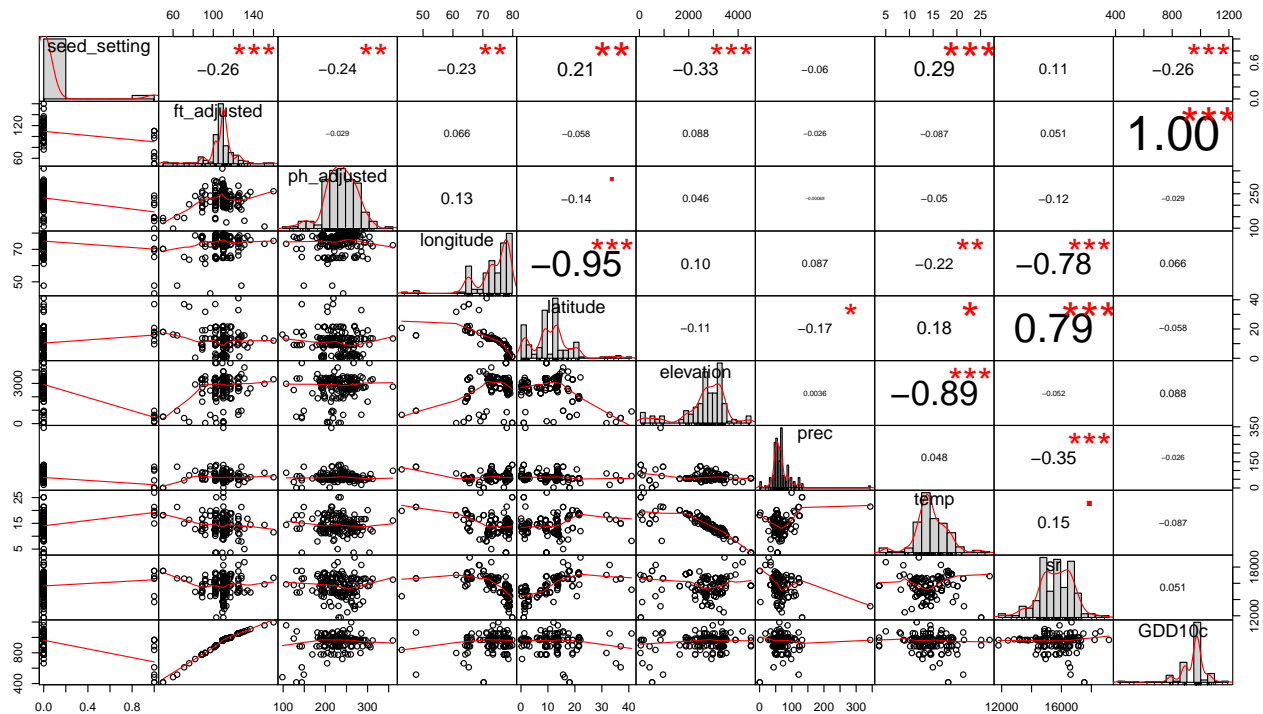
```
## 'data.frame':  181 obs. of  10 variables:
## $ seed_setting: int  0 0 0 0 0 0 0 0 0 0 ...
## $ ft_adjusted : int  124 101 124 109 88 112 109 109 109 109 ...
## $ ph_adjusted : num  235 307 197 292 230 ...
## $ longitude   : num  77 73.6 NA 74.2 77.1 ...
## $ latitude    : num  9 13.7 NA 13.16 9.49 ...
## $ elevation   : int  3363 3219 NA 2729 3273 2640 3309 4518 2754 2865 ...
## $ prec        : num  67 79.4 NA 45.3 70.4 ...
## $ temp        : num  12.9 12.6 NA 16.7 12.1 ...
## $ sr          : num  14869 16154 NA 16572 14816 ...
## $ GDD10c      : num  1047 872 1047 960 774 ...
```

```
library(PerformanceAnalytics)
```

```
chart.Correlation(
  correlation_south_american,
  histogram = TRUE,
  method = c("pearson"))
```



```
chart.Correlation(
  correlation_south_american,
  histogram = TRUE,
  method = c("spearman"))
```



Dataset III: Central American Accessions

Filtering the South American accessions.

```
table(data$country)
```

```
##
## ARG BOL BRA CHL ECU GER GTM MEX PER PTR USA VEN
##   7  28   5   1  35   1   6  60 104   1   5   1
```

```
correlation_central_american <- filter(data, (country %in% c("GTM", "MEX", "PTR")))
dim(correlation_central_american)
```

```
## [1] 67 14
```

```
table(correlation_central_american$country)
```

```
##
## GTM MEX PTR
##   6  60   1
```

```
correlation_central_american$longitude <- abs(correlation_central_american$longitude)
correlation_central_american$latitude <- abs(correlation_central_american$latitude)
head(correlation_central_american)
```

```
##   genotypes species country plot seed_setting ft_adjusted ph_adjusted
## 1   cru_83_n   cru    MEX  101           1           70      188.95
## 2   cru_78_y   cru    MEX  106           1           70      187.95
## 3 hypo_157_y  hypo    MEX  109           0          137      223.95
## 4 hypo_158_y  hypo    MEX  112           1          101      265.95
## 5   cru_100_y   cru    MEX  116           1           70      176.95
## 6   hypo_174  hypo    MEX  128           0          140      235.70
##   longitude latitude elevation      prec      temp      sr  GDD10c
## 1         NA      NA         NA         NA         NA      NA  594.90
## 2  98.50000 18.75000      1540  70.41667  21.77500 17724.42  594.90
## 3  97.71667 20.00139       723 175.75000  20.35833 17263.58 1107.95
## 4  98.75000 18.75000      1543  76.25000  21.80833 17850.33  872.40
## 5  99.01667 19.25000      2348  55.58333  15.91667 18202.75  594.90
## 6  98.21667 19.68333      2984  77.33333  11.27500 17715.00 1116.80
```

```
dim(correlation_central_american)
```

```
## [1] 67 14
```

```
correlation_central_american <- dplyr::select(correlation_central_american, seed_setting, ft_adjusted, ph_adjusted, longitude, latitude, elevation, prec, temp, sr, GDD10c)
```

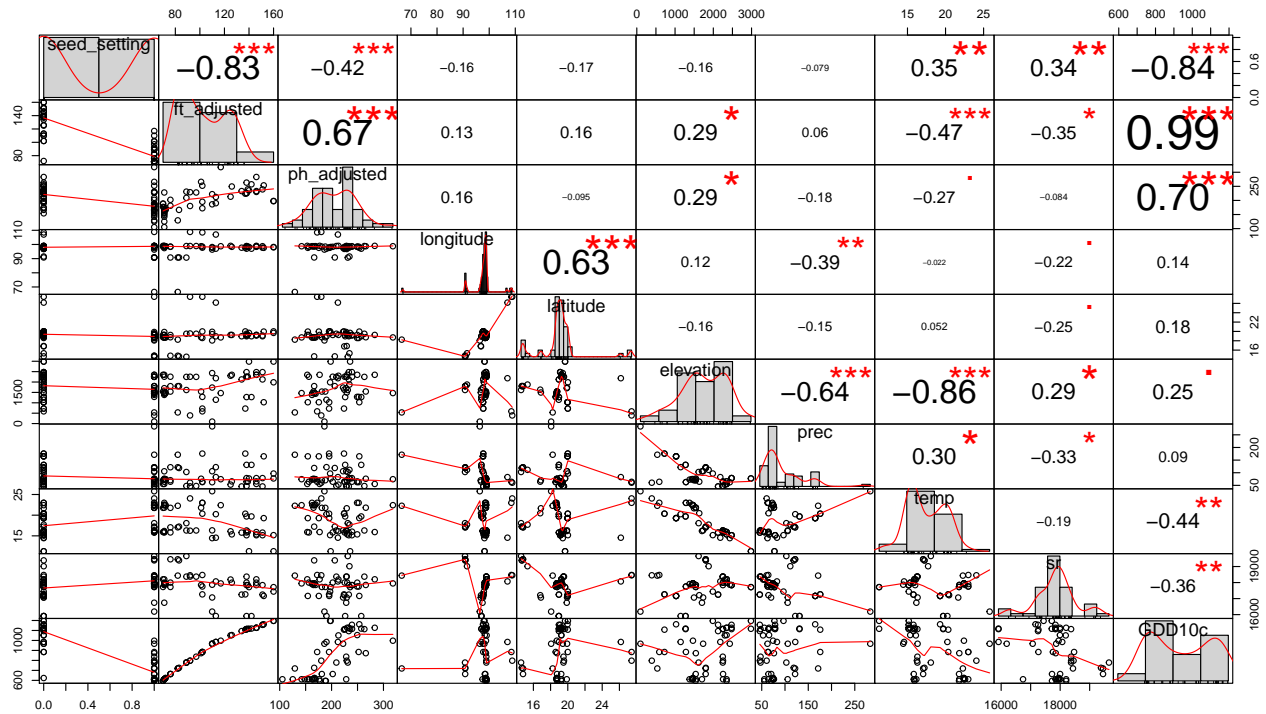
```
str(correlation_central_american)
```

```
## 'data.frame':   67 obs. of  10 variables:
## $ seed_setting: int  1 1 0 1 1 0 0 0 0 1 ...
## $ ft_adjusted : int  70 70 137 101 70 140 146 138 160 71 ...
## $ ph_adjusted : num  189 188 224 266 177 ...
## $ longitude   : num  NA 98.5 97.7 98.8 99 ...
## $ latitude    : num  NA 18.8 20 18.8 19.2 ...
## $ elevation   : int  NA 1540 723 1543 2348 2984 2348 2477 2984 NA ...
## $ prec        : num  NA 70.4 175.8 76.2 55.6 ...
## $ temp        : num  NA 21.8 20.4 21.8 15.9 ...
## $ sr          : num  NA 17724 17264 17850 18203 ...
```

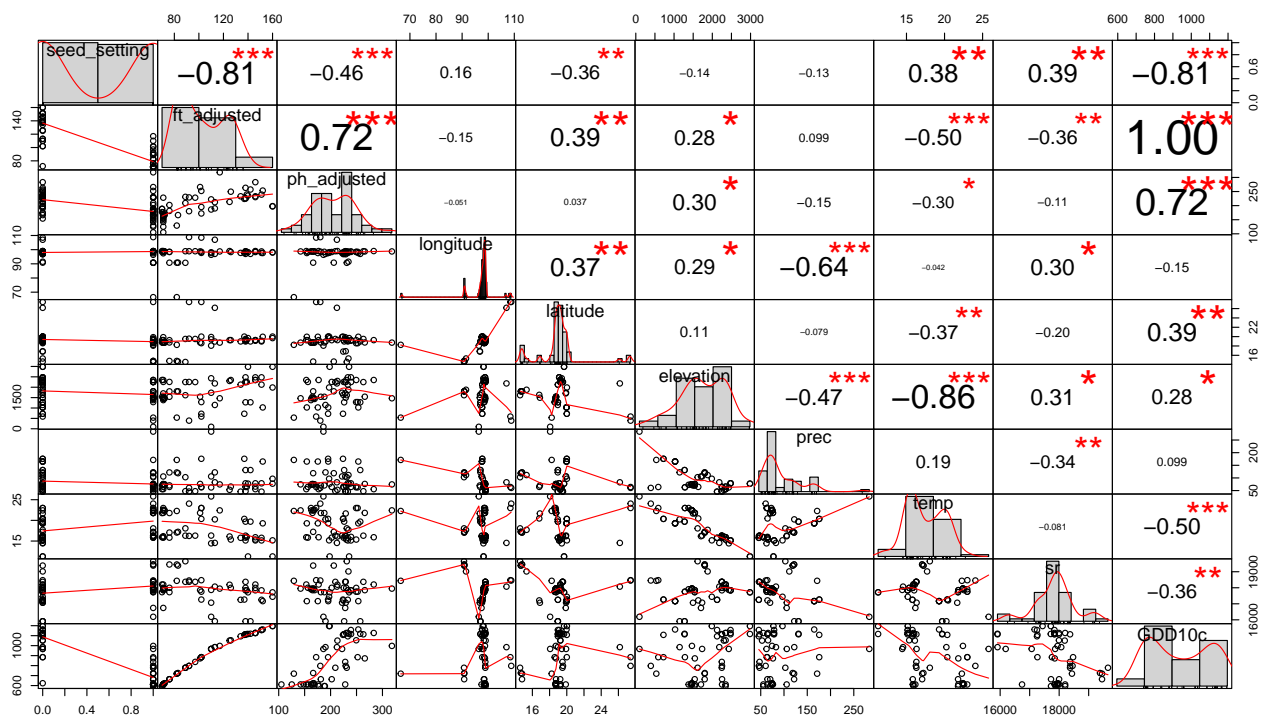
```
## $ GDD10c : num 595 595 1108 872 595 ...
```

```
library(PerformanceAnalytics)
```

```
chart.Correlation(
  correlation_central_american,
  histogram = TRUE,
  method = c("pearson"))
```



```
chart.Correlation(
  correlation_central_american,
  histogram = TRUE,
  method = c("spearman"))
```



Dataset IV: Central American *A. hypochondriacus* Accessions

Filtering *A. hypochondriacus* accessions from Central America.

```
correlation_central_american_hypo <- filter(data, (country %in% c("GTM", "MEX", "PTR"))) %>%
  filter(species=="hypo")
dim(correlation_central_american_hypo)
```

```
## [1] 32 14
```

```
table(correlation_central_american_hypo$country)
```

```
##
## MEX PTR
## 31 1
```

```
correlation_central_american_hypo$longitude <- abs(correlation_central_american_hypo$longitude)
correlation_central_american_hypo$latitude <- abs(correlation_central_american_hypo$latitude)
head(correlation_central_american_hypo)
```

```
##   genotypes species country plot seed_setting ft_adjusted ph_adjusted
## 1 hypo_157_y    hypo     MEX  109           0         137      223.95
## 2 hypo_158_y    hypo     MEX  112           1         101      265.95
## 3 hypo_174     hypo     MEX  128           0         140      235.70
## 4 hypo_146_y    hypo     MEX  201           0         146      230.20
## 5 hypo_148_y    hypo     MEX  202           0         138         NA
## 6 hypo_170_n    hypo     MEX  214           0         160      198.20
##   longitude latitude elevation      prec      temp      sr  GDD10c
## 1  97.71667 20.00139      723 175.75000 20.35833 17263.58 1107.95
## 2  98.75000 18.75000     1543  76.25000 21.80833 17850.33  872.40
## 3  98.21667 19.68333     2984  77.33333 11.27500 17715.00 1116.80
## 4  98.33333 19.25000     2348  65.75000 15.35833 17992.25 1126.95
```

```
## 5  98.76667 19.11667      2477  78.75000 15.15000 17840.83 1113.45
## 6  98.21667 19.68333      2984  77.33333 11.27500 17715.00 1196.15
```

```
dim(correlation_central_american_hypo)
```

```
## [1] 32 14
```

```
correlation_central_american_hypo <- dplyr::select(correlation_central_american_hypo, seed_setting, ft_
```

```
str(correlation_central_american_hypo)
```

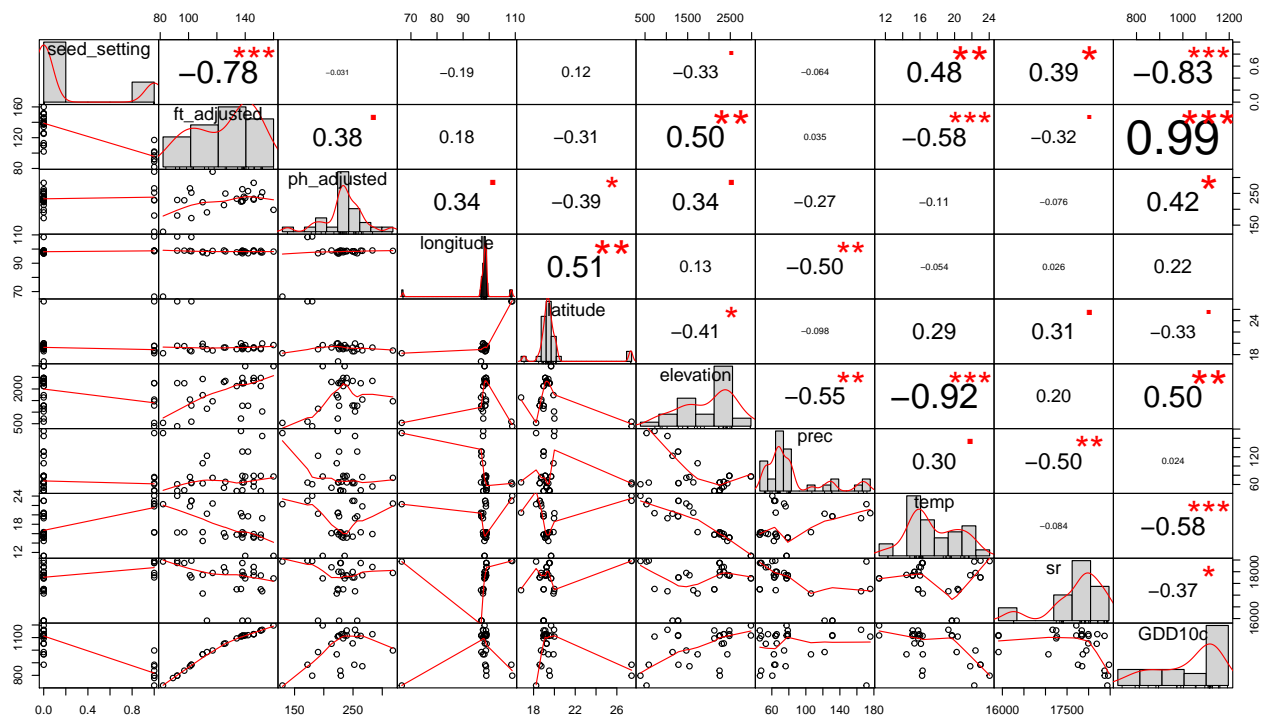
```
## 'data.frame':  32 obs. of  10 variables:
## $ seed_setting: int  0 1 0 0 0 0 0 1 1 0 ...
## $ ft_adjusted : int  137 101 140 146 138 160 138 117 89 134 ...
## $ ph_adjusted : num  224 266 236 230 NA ...
## $ longitude   : num  97.7 98.8 98.2 98.3 98.8 ...
## $ latitude    : num  20 18.8 19.7 19.2 19.1 ...
## $ elevation   : int  723 1543 2984 2348 2477 2984 1275 1468 NA 2125 ...
## $ prec        : num  175.8 76.2 77.3 65.8 78.8 ...
## $ temp        : num  20.4 21.8 11.3 15.4 15.2 ...
## $ sr          : num  17264 17850 17715 17992 17841 ...
## $ GDD10c      : num  1108 872 1117 1127 1113 ...
```

```
dim(correlation_central_american_hypo)
```

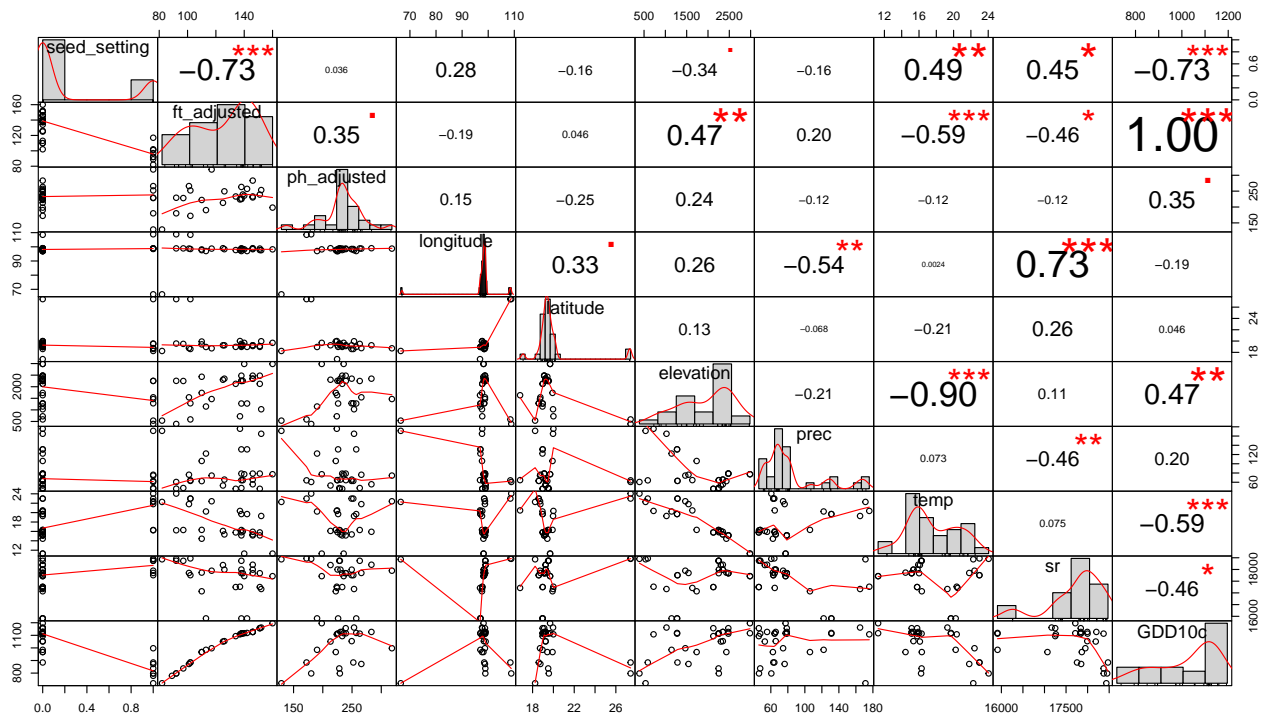
```
## [1] 32 10
```

```
library(PerformanceAnalytics)
```

```
chart.Correlation(
  correlation_central_american_hypo,
  histogram = TRUE,
  method = c("pearson"))
```




```
chart.Correlation(
  correlation_central_american_hypo,
  histogram = TRUE,
  method = c("spearman"))
```



4. Focus on Central American Accessions

We clearly see that Central America is the center of flowering time and seed setting.

```
central_american <- filter(data, (country %in% c("GTM", "MEX", "PTR")))
table(central_american$species)
```

```
##
```

```
## cru hdus hybr hypo
```

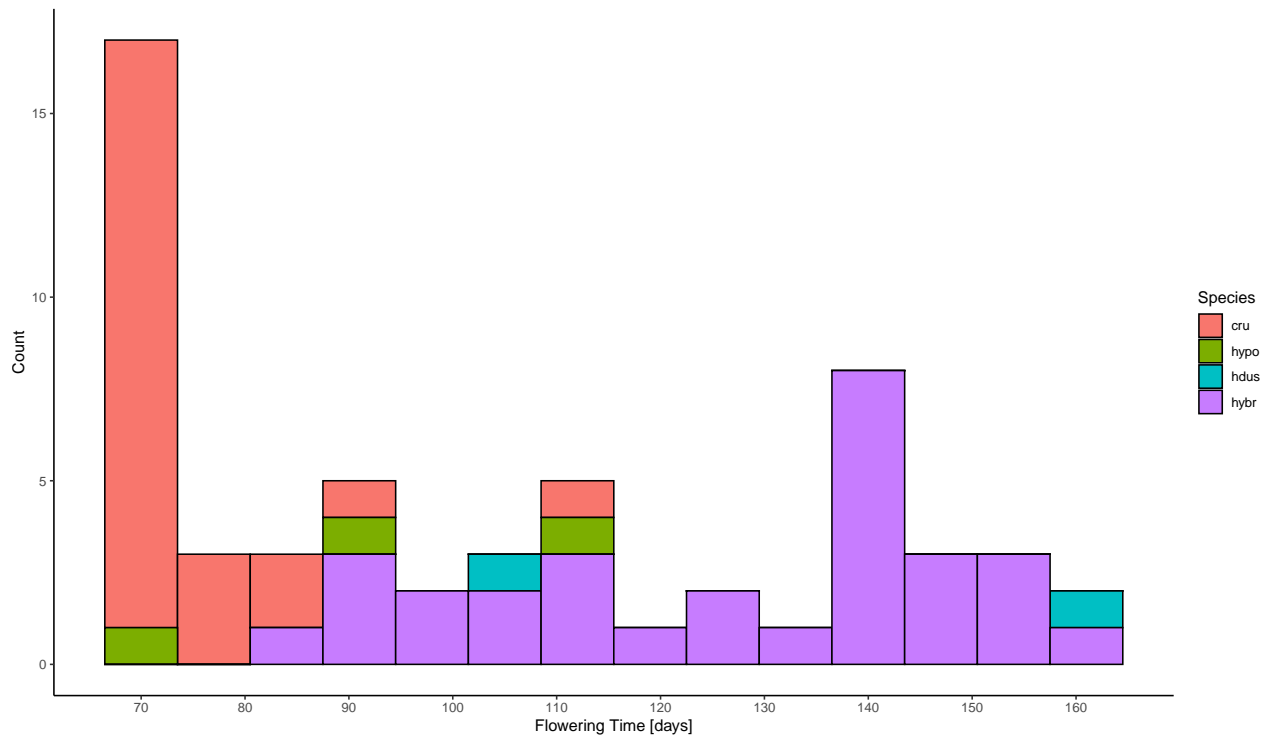
```
## 25 6 4 32
```

```
central_american_plot <- ggplot(central_american, aes(x=ft_adjusted, fill=factor(species)), stat="count") +
  geom_histogram(color="black", binwidth = 7) +
  scale_x_continuous(breaks = scales::pretty_breaks(n = 10)) +
  #scale_y_continuous(expand = expand_scale(mult = 0, add = 0.1)) +
  labs(y="Count", x="Flowering Time [days]") +
```

```
theme(axis.text.x = element_text(colour = "grey20", size = 4.5, hjust = 0.5, vjust = 0.5),
      axis.text.y = element_text(colour = "grey20", size = 4.5),
      text = element_text(size = 5), legend.position=c(0.9, 0.9),
      panel.spacing = unit(1.0, "lines"),
      legend.title = element_text(size=5, face="bold"),
      legend.key.size = unit(4, "mm"),
      legend.text=element_text(size=5)) +
  scale_fill_discrete(name = "Species", labels=c("cru", "hypo", "hdus", "hybr")) +
```

```
scale_color_discrete(name = "Species", labels=c("cru","hypo", "hdus", "hybr"))+
theme_classic()
```

```
central_american_plot
```



```
#ggsave(ca_grain_plot, device = "jpeg", path = NULL, scale = 1, width = 7.5, height = 6.3, units = "cm")
```

There is a clear flowering time difference between the species. Also, *A. cruentus* shows a narrow distribution and *A. hypochondriacus* shows very wide distribution.

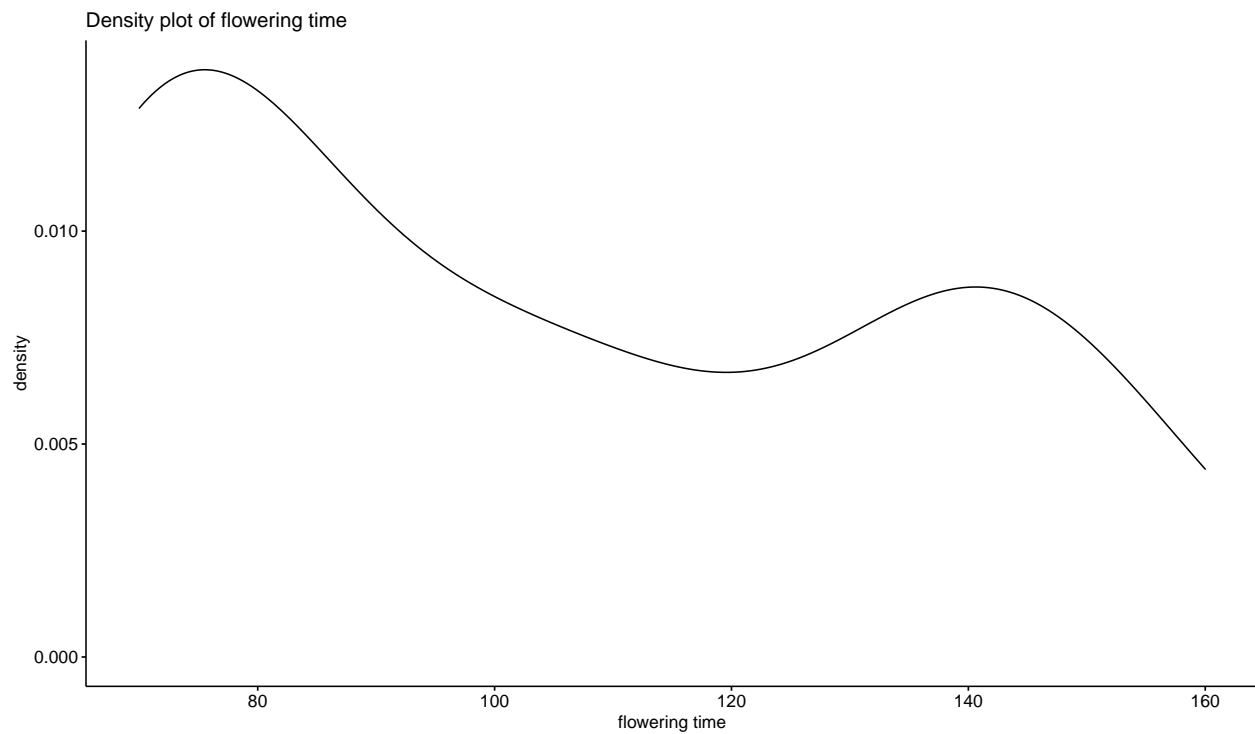
I want to compare the seed setting and non-setting accessions of Central America in terms of phenotype traits and environmental variables.

```
central_american$seed_setting <- as.factor(central_american$seed_setting)
```

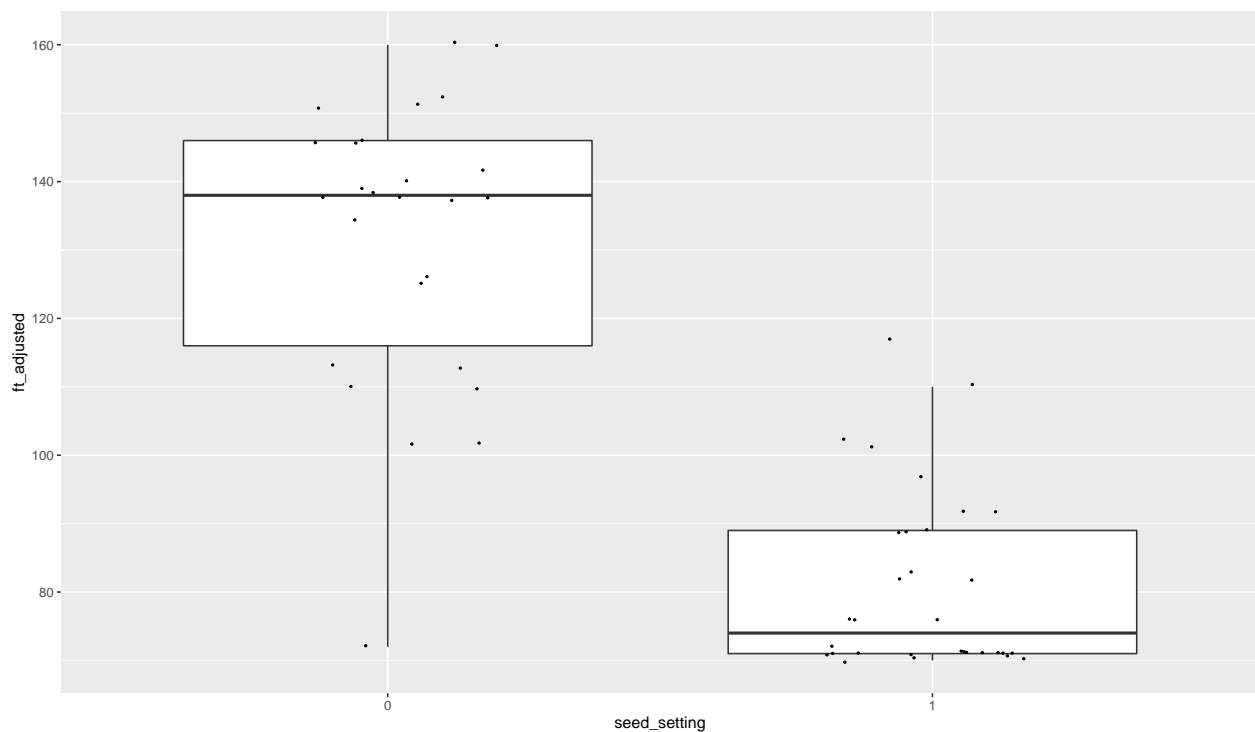
Phenotypic traits

A. Flowering time

```
ggdensity(central_american$ft_adjusted,
  main = "Density plot of flowering time",
  xlab = "flowering time")
```



```
ggplot(data = central_american, mapping = aes(x = seed_setting, y = ft_adjusted)) +  
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_21 <-glm(ft_adjusted ~ seed_setting , data = central_american, family = gaussian)  
summary(md_21)
```

```
##  
## Call:
```

```
## glm(formula = ft_adjusted ~ seed_setting, family = gaussian,
##      data = central_american)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -59.885  -9.844   1.156  10.896  36.156
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    131.885      3.339   39.50 < 2e-16 ***
## seed_setting1  -51.041      4.495  -11.36 3.75e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 289.837)
##
##      Null deviance: 53602  on 57  degrees of freedom
## Residual deviance: 16231  on 56  degrees of freedom
## (9 observations deleted due to missingness)
## AIC: 497.38
##
## Number of Fisher Scoring iterations: 2
```

```
lrtest(md_21)
```

```
## Likelihood ratio test
##
## Model 1: ft_adjusted ~ seed_setting
## Model 2: ft_adjusted ~ 1
##   #Df LogLik Df Chisq Pr(>Chisq)
## 1    3 -245.69
## 2    2 -280.34 -1 69.29 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

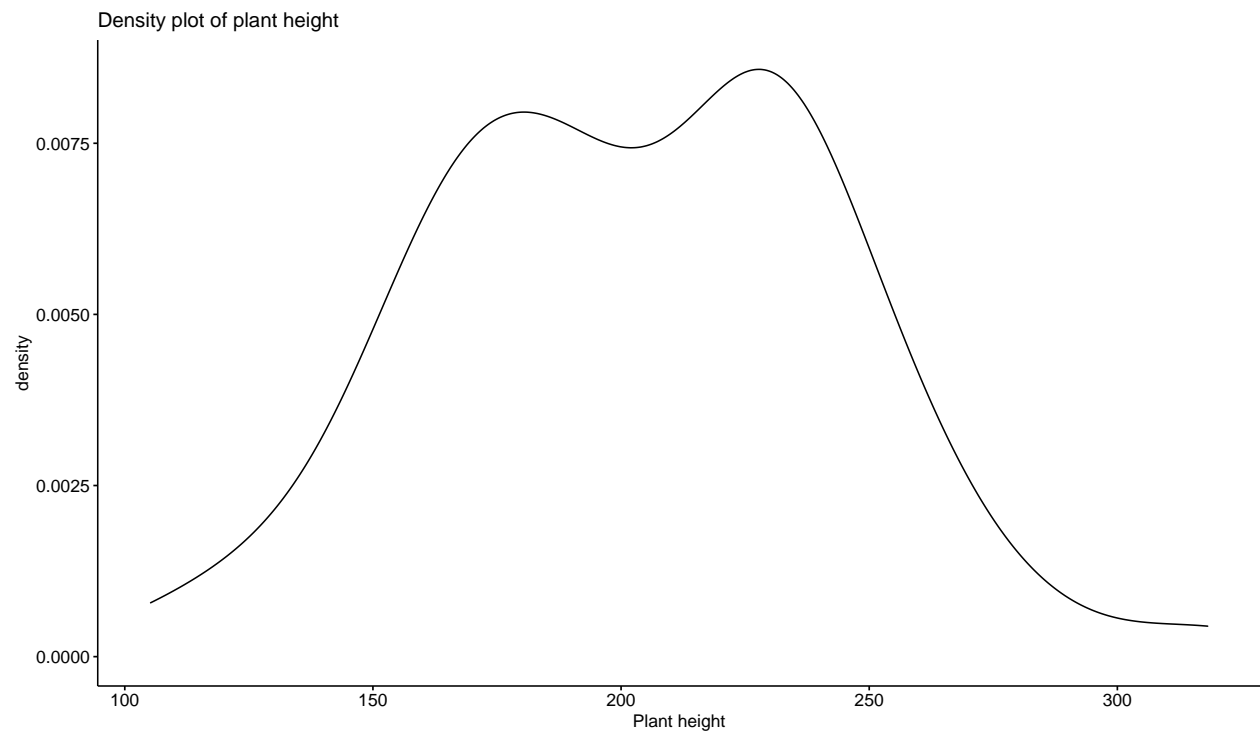
```
lsmeans_ss_ft <- LSD.test(md_21, "seed_setting", p.adj="none" )
lsmeans_ss_ft
```

```
## $statistics
##      MSerror Df      Mean      CV
##    289.837 56 103.7241 16.41334
##
## $parameters
##      test p.adjusted      name.t ntr alpha
## Fisher-LSD      none seed_setting  2 0.05
##
## $means
##   ft_adjusted      std r      LCL      UCL Min Max Q25 Q50 Q75
## 0    131.88462 20.82177 26 125.19620 138.57303  72 160 116 138 146
## 1     80.84375 13.18873 32  74.81489  86.87261  70 117  71  74  89
##
## $comparison
## NULL
##
## $groups
```

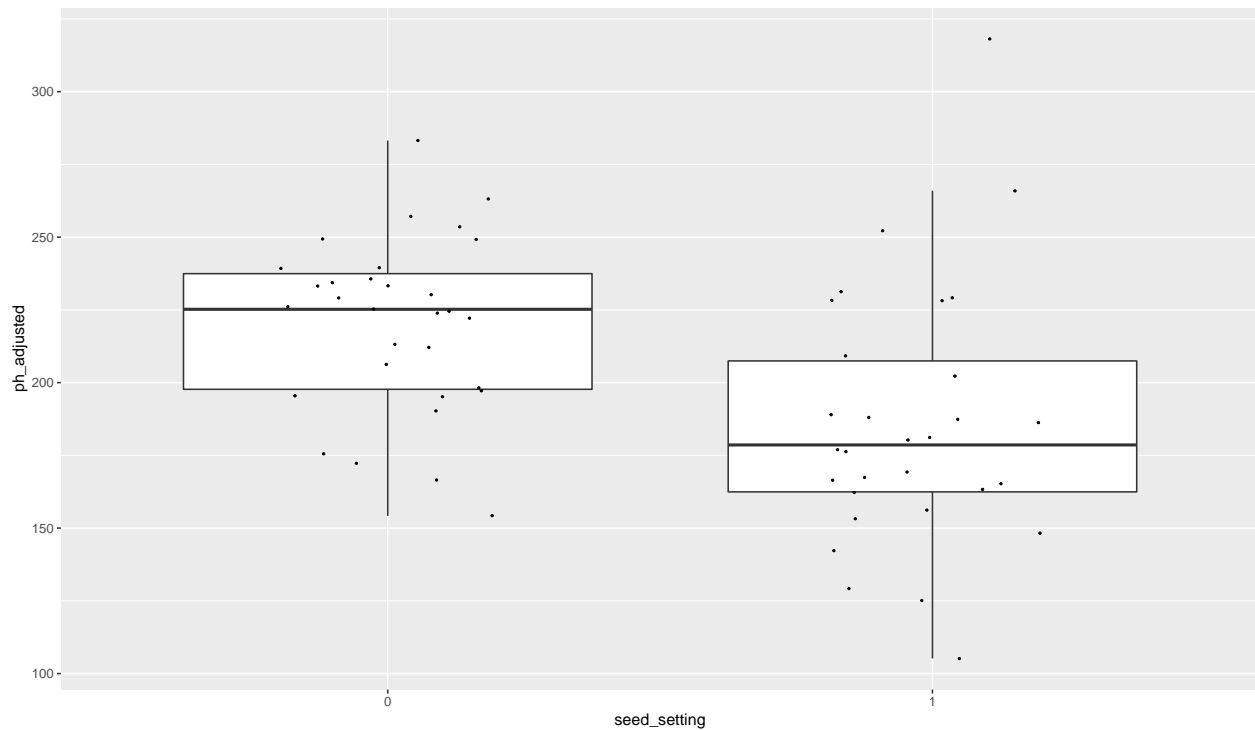
```
##   ft_adjusted groups
## 0   131.88462     a
## 1    80.84375     b
##
## attr(,"class")
## [1] "group"
```

B. Plant height

```
ggdensity(central_american$ph_adjusted,
  main = "Density plot of plant height",
  xlab = "Plant height")
```



```
ggplot(data = central_american, mapping = aes(x = seed_setting, y = ph_adjusted)) +
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_22 <-glm(ph_adjusted ~ seed_setting , data = central_american, family = gaussian)
summary(md_22)
```

```
##
## Call:
## glm(formula = ph_adjusted ~ seed_setting, family = gaussian,
##      data = central_american)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -80.892  -23.073   1.358   18.927  132.108
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    220.273     6.811   32.339 < 2e-16 ***
## seed_setting1  -34.181     9.713   -3.519 0.000841 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 1438.19)
##
##      Null deviance: 102666  on 60  degrees of freedom
## Residual deviance:  84853  on 59  degrees of freedom
## (6 observations deleted due to missingness)
## AIC: 620.62
##
## Number of Fisher Scoring iterations: 2
```

```
lrtest(md_22)
```

```
## Likelihood ratio test
##
```

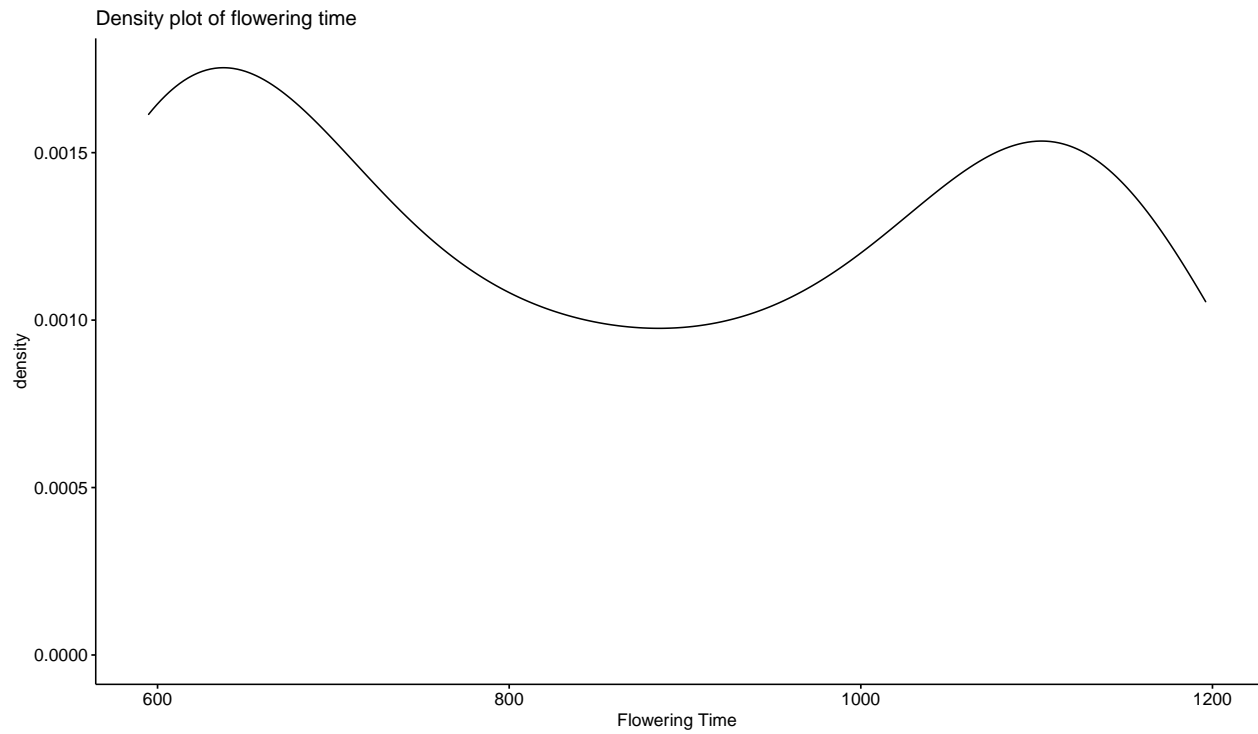
```
## Model 1: ph_adjusted ~ seed_setting
## Model 2: ph_adjusted ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    3 -307.31
## 2    2 -313.12 -1 11.624  0.0006511 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

lsmeans_ss_ph <- LSD.test(md_22, "seed_setting", p.adj="none" )
lsmeans_ss_ph

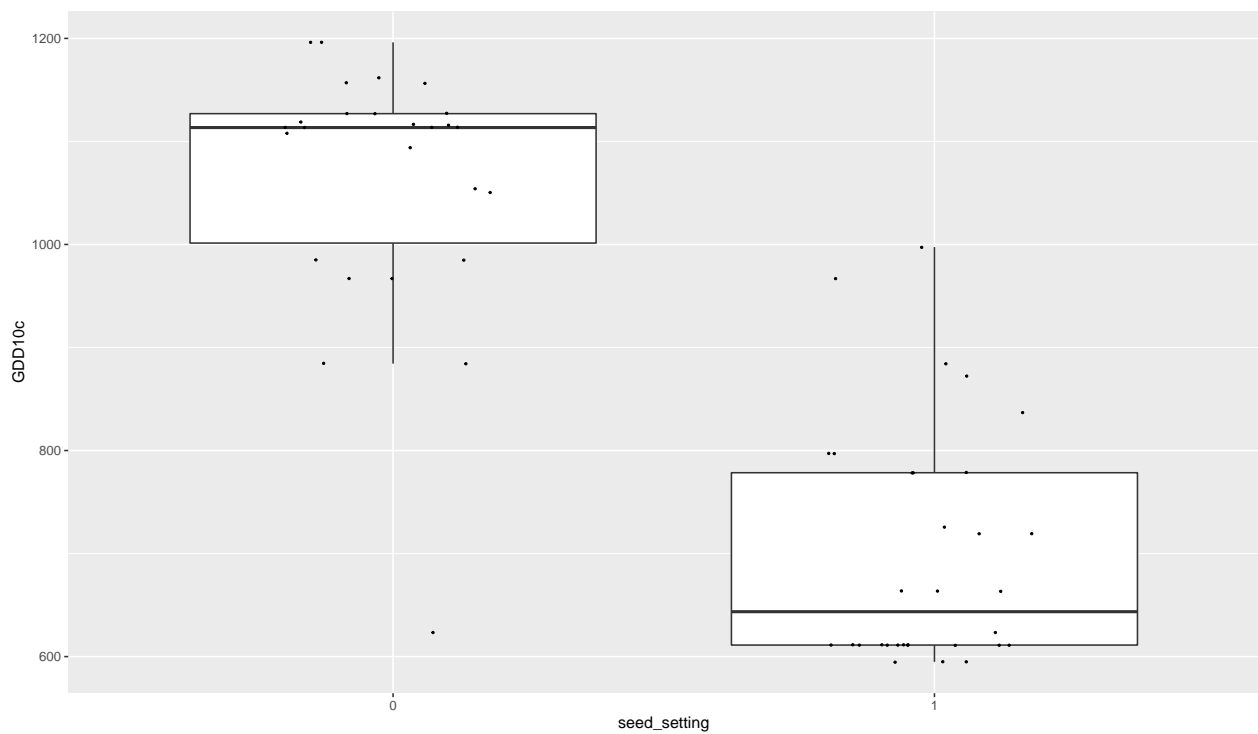
## $statistics
##   MSerror Df      Mean      CV
##   1438.19 59 203.4623 18.63907
##
## $parameters
##           test p.adjusted      name.t ntr alpha
##   Fisher-LSD      none seed_setting   2  0.05
##
## $means
##   ph_adjusted      std r      LCL      UCL   Min   Max   Q25   Q50   Q75
## 0    220.2726 29.88375 31 206.6433 233.9019 154.2 283.2 197.70 225.200 237.45
## 1    186.0917 44.74529 30 172.2371 199.9463 105.2 318.2 162.45 178.575 207.45
##
## $comparison
## NULL
##
## $groups
##   ph_adjusted groups
## 0    220.2726      a
## 1    186.0917      b
##
## attr(,"class")
## [1] "group"
```

C. Growing degree days (Tbase is 10°C)

```
ggdensity(central_american$GDD10c,
           main = "Density plot of flowering time",
           xlab = "Flowering Time")
```



```
ggplot(data = central_american, mapping = aes(x = seed_setting, y = GDD10c)) +  
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_23 <-glm(GDD10c ~ seed_setting , data = central_american, family = gaussian)  
summary(md_23)
```

```
##  
## Call:
```



```
## glm(formula = GDD10c ~ seed_setting, family = gaussian, data = central_american)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -440.14  -88.37   19.78   74.97  297.88
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1063.69      23.49   45.28  <2e-16 ***
## seed_setting1  -364.12      31.63  -11.51  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 14347.88)
##
##      Null deviance: 2705388  on 57  degrees of freedom
## Residual deviance:  803481  on 56  degrees of freedom
## (9 observations deleted due to missingness)
## AIC: 723.7
##
## Number of Fisher Scoring iterations: 2
```

```
lrtest(md_23)
```

```
## Likelihood ratio test
##
## Model 1: GDD10c ~ seed_setting
## Model 2: GDD10c ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    3 -358.85
## 2    2 -394.06 -1 70.415  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
lsmeans_ss_GDD10c <- LSD.test(md_23, "seed_setting", p.adj="none" )
lsmeans_ss_GDD10c
```

```
## $statistics
##      MSerror Df      Mean      CV
##  14347.88 56 862.7974 13.88305
##
## $parameters
##      test p.adjusted      name.t ntr alpha
## Fisher-LSD      none seed_setting  2  0.05
##
## $means
##      GDD10c      std r      LCL      UCL      Min      Max      Q25      Q50
## 0 1063.6923 124.0660 26 1016.6336 1110.7510 623.55 1196.15 1001.35 1113.450
## 1  699.5703 116.2133 32  657.1521  741.9885 594.90  997.45  611.20  643.575
##      Q75
## 0 1126.95
## 1  778.45
##
## $comparison
## NULL
```

```
##
## $groups
##      GDD10c groups
## 0 1063.6923      a
## 1  699.5703      b
##
## attr(,"class")
## [1] "group"
table(central_american$seed_setting)
```

```
##
##  0  1
## 34 33
table(central_american$seed_setting)
```

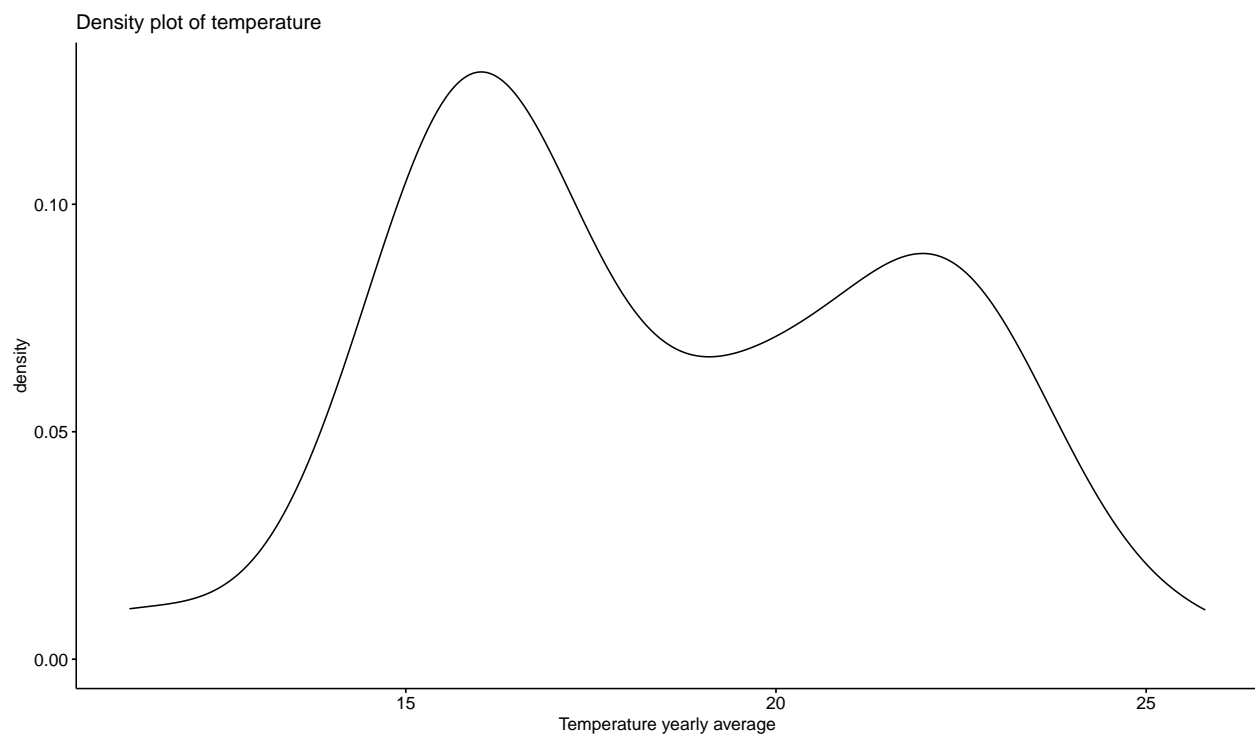
```
##
##  0  1
## 34 33
```

Climatic & Geographical Variables

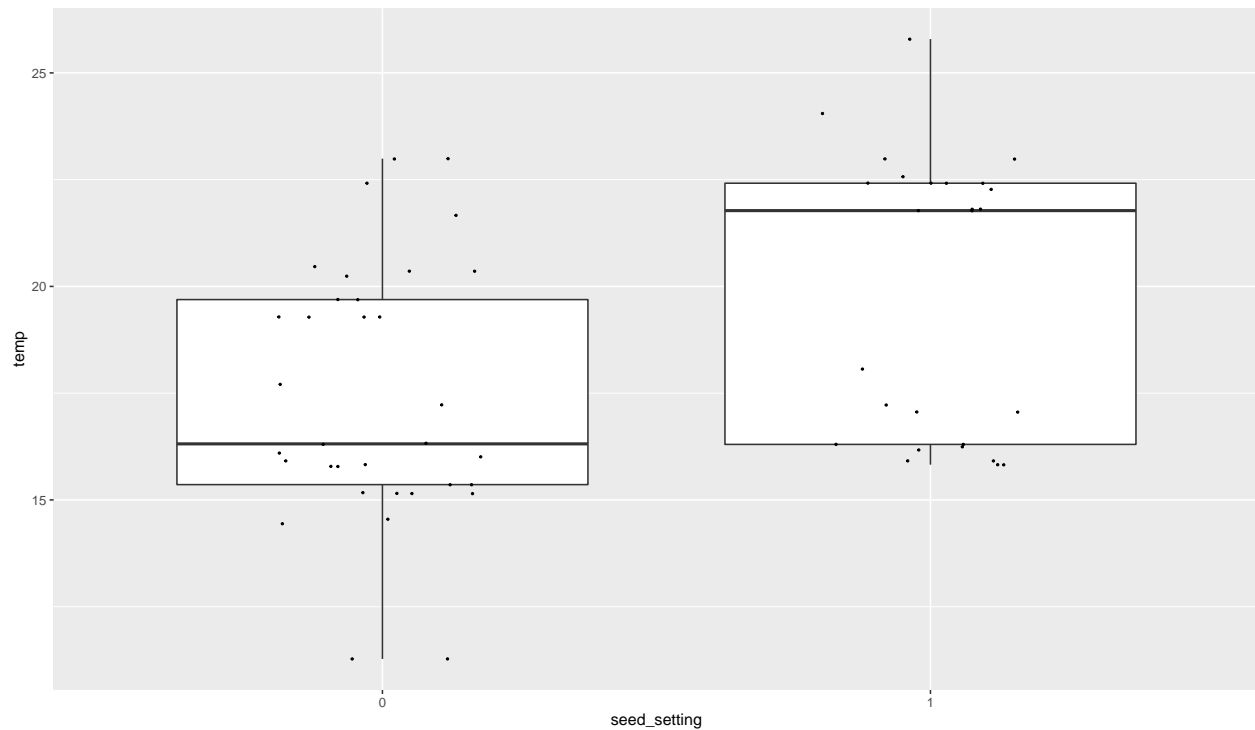
A. Temperature

Distribution of the data.

```
ggdensity(central_american$temp,
  main = "Density plot of temperature",
  xlab = "Temperature yearly average")
```



```
ggplot(data = central_american, mapping = aes(x = seed_setting, y = temp)) +
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_24 <-glm(temp ~ seed_setting , data = central_american, family = gaussian)
summary(md_24)
```

```
##
## Call:
## glm(formula = temp ~ seed_setting, family = gaussian, data = central_american)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -6.1912  -2.3865  -0.6912   2.5942   5.9692
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   17.4662    0.5375  32.492 < 2e-16 ***
## seed_setting1    2.3563    0.8166   2.885  0.00548 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 9.824597)
##
##      Null deviance: 651.63  on 59  degrees of freedom
## Residual deviance: 569.83  on 58  degrees of freedom
## (7 observations deleted due to missingness)
## AIC: 311.33
##
## Number of Fisher Scoring iterations: 2
```

```

lrtest(md_24)

## Likelihood ratio test
##
## Model 1: temp ~ seed_setting
## Model 2: temp ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    3 -152.67
## 2    2 -156.69 -1  8.0483   0.004555 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

lsmeans_ss_temp <- LSD.test(md_24, "seed_setting", p.adj="none" )
lsmeans_ss_temp

## $statistics
##      MSerror Df      Mean      CV
##    9.824597 58 18.48722 16.95453
##
## $parameters
##      test p.adjusted      name.t ntr alpha
## Fisher-LSD      none seed_setting  2  0.05
##
## $means
##      temp      std  r      LCL      UCL      Min      Max      Q25      Q50
## 0 17.46618 3.024049 34 16.39016 18.54220 11.275 22.99167 15.35833 16.3125
## 1 19.82244 3.274421 26 18.59196 21.05291 15.825 25.79167 16.30000 21.7750
##      Q75
## 0 19.69167
## 1 22.41667
##
## $comparison
## NULL
##
## $groups
##      temp groups
## 1 19.82244      a
## 0 17.46618      b
##
## attr("class")
## [1] "group"

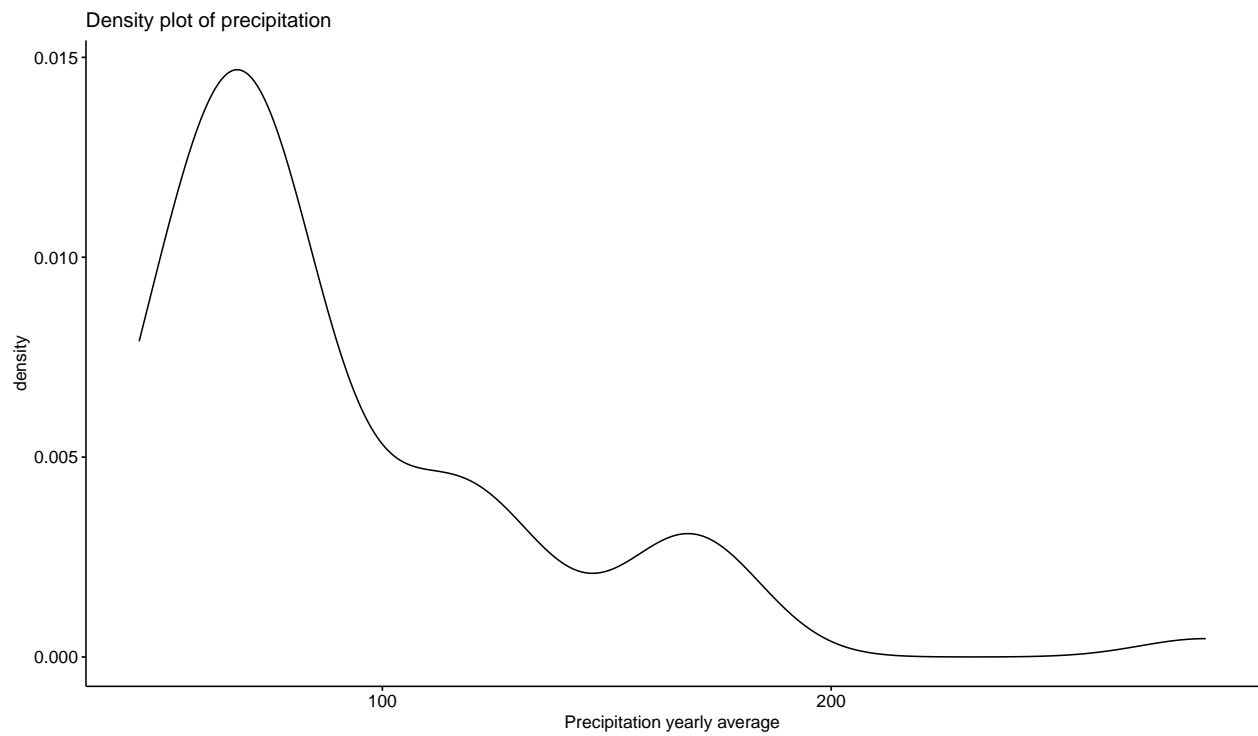
```

B. Precipitation

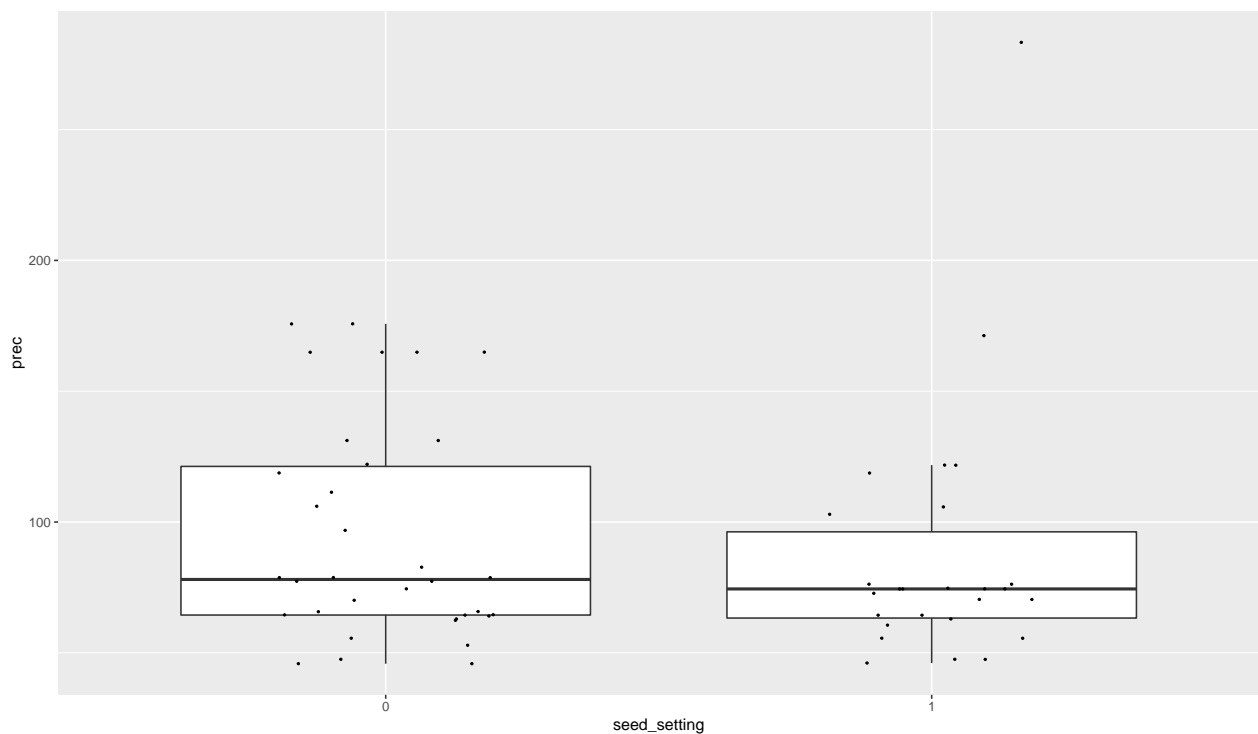
```

ggdensity(central_american$prec,
  main = "Density plot of precipitation",
  xlab = "Precipitation yearly average")

```



```
ggplot(data = central_american, mapping = aes(x = seed_setting, y = prec)) +
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_25 <-glm(prec ~ seed_setting , data = central_american, family = gaussian)
summary(md_25)
```

```
##
## Call:
```

```
## glm(formula = prec ~ seed_setting, family = gaussian, data = central_american)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -48.54  -29.82  -15.63   19.97  196.08
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    94.377      7.735  12.202  <2e-16 ***
## seed_setting1  -7.127     11.750  -0.607    0.546
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 2033.995)
##
##      Null deviance: 118720  on 59  degrees of freedom
## Residual deviance: 117972  on 58  degrees of freedom
## (7 observations deleted due to missingness)
## AIC: 631.3
##
## Number of Fisher Scoring iterations: 2
```

```
lrtest(md_25)
```

```
## Likelihood ratio test
##
## Model 1: prec ~ seed_setting
## Model 2: prec ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    3 -312.65
## 2    2 -312.84 -1  0.3795    0.5379
```

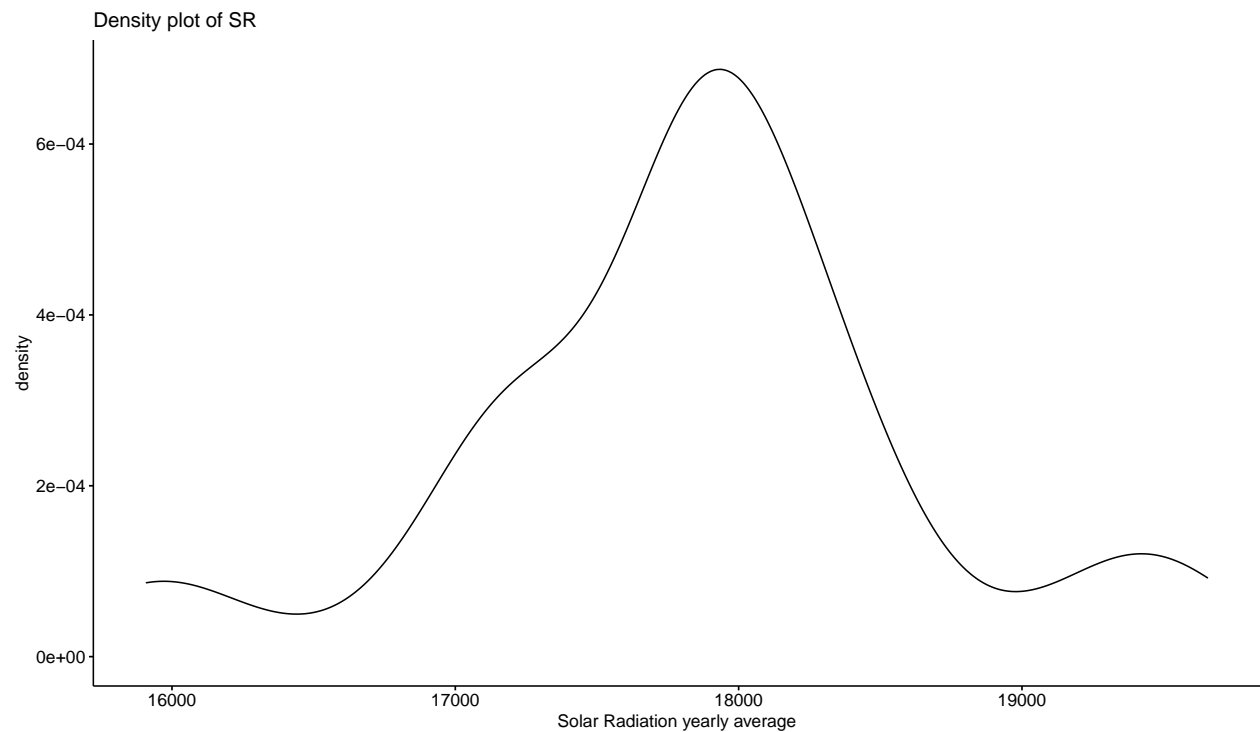
```
lsmeans_ss_prec <- LSD.test(md_25, "seed_setting", p.adj="none" )
lsmeans_ss_prec
```

```
## $statistics
##      MSerror Df      Mean      CV
## 2033.995 58 91.28889 49.40342
##
## $parameters
##      test p.adjusted      name.t ntr alpha
## Fisher-LSD      none seed_setting  2  0.05
##
## $means
##      prec      std r      LCL      UCL      Min      Max      Q25      Q50
## 0 94.37745 41.72935 34 78.89505 109.8599 45.83333 175.7500 64.43750 78.04167
## 1 87.25000 49.19656 26 69.54520 104.9548 46.08333 283.3333 63.29167 74.41667
##      Q75
## 0 121.25
## 1  96.25
##
## $comparison
## NULL
##
## $groups
```

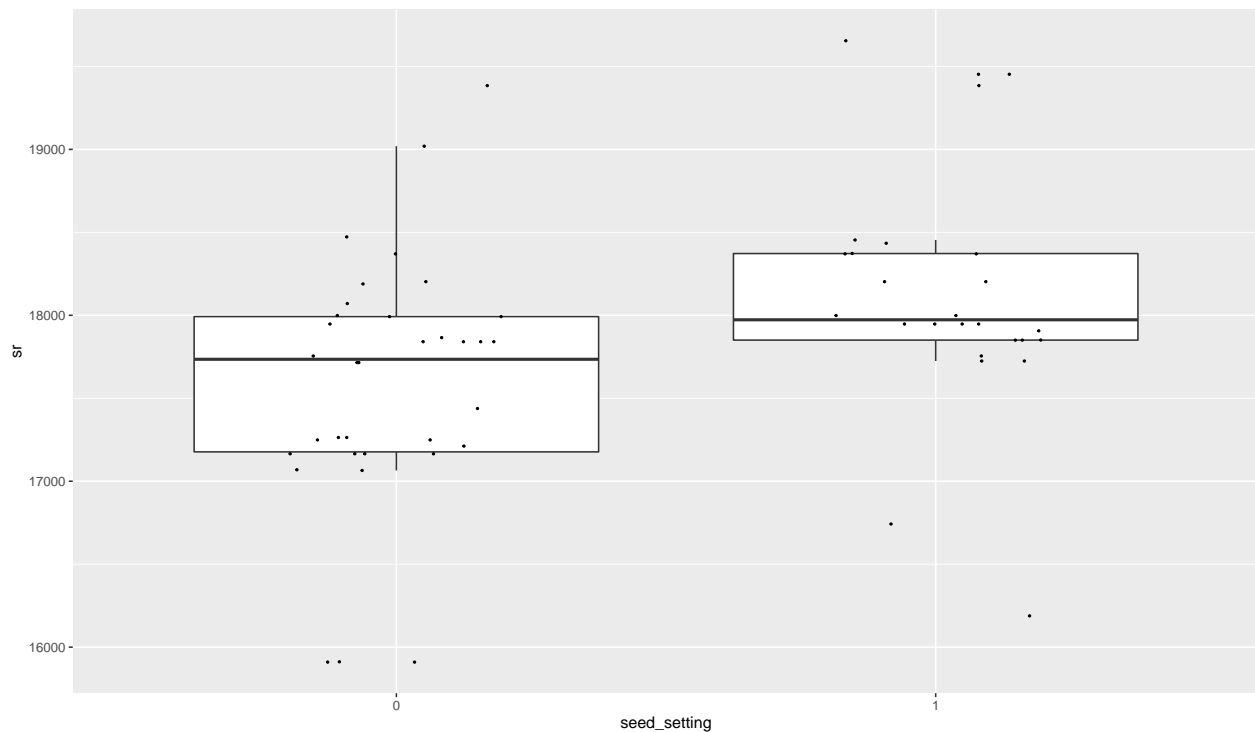
```
##      prec groups
## 0 94.37745      a
## 1 87.25000      a
##
## attr(,"class")
## [1] "group"
```

C. Solar Radiation

```
ggdensity(central_american$sr,
  main = "Density plot of SR",
  xlab = "Solar Radiation yearly average")
```



```
ggplot(data = central_american, mapping = aes(x = seed_setting, y = sr)) +
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_26 <-glm(sr ~ seed_setting , data = central_american, family = gaussian)
summary(md_26)
```

```
##
## Call:
## glm(formula = sr ~ seed_setting, family = gaussian, data = central_american)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1955.17  -356.06   -42.83    320.80   1788.62
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   17595.7     130.1  135.290 < 2e-16 ***
## seed_setting1    547.8     197.6   2.773  0.00747 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 575127.1)
##
##      Null deviance: 37778486  on 59  degrees of freedom
## Residual deviance: 33357370  on 58  degrees of freedom
## (7 observations deleted due to missingness)
## AIC: 969.98
##
## Number of Fisher Scoring iterations: 2
```

```
lrtest(md_26)
```

```
## Likelihood ratio test
##
## Model 1: sr ~ seed_setting
```



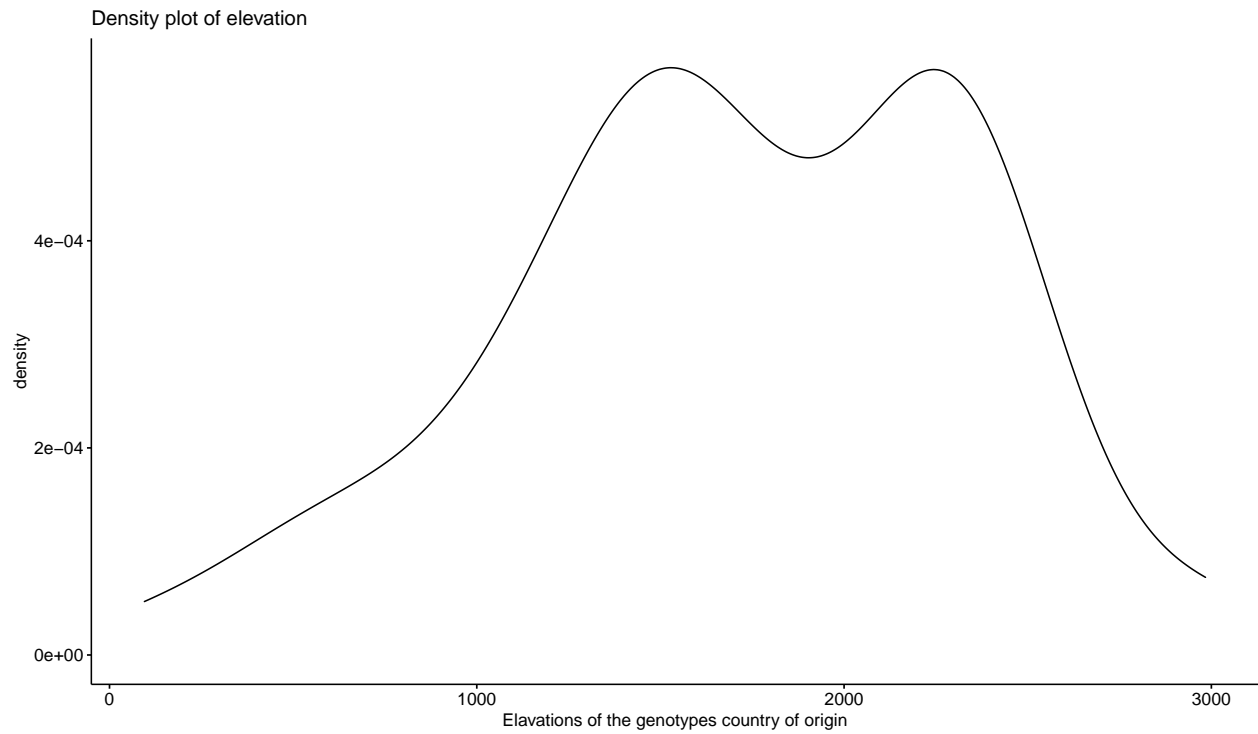
```
## Model 2: sr ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    3 -481.99
## 2    2 -485.72 -1  7.4677   0.006282 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

lsmeans_ss_sr <- LSD.test(md_26, "seed_setting", p.adj="none" )
lsmeans_ss_sr

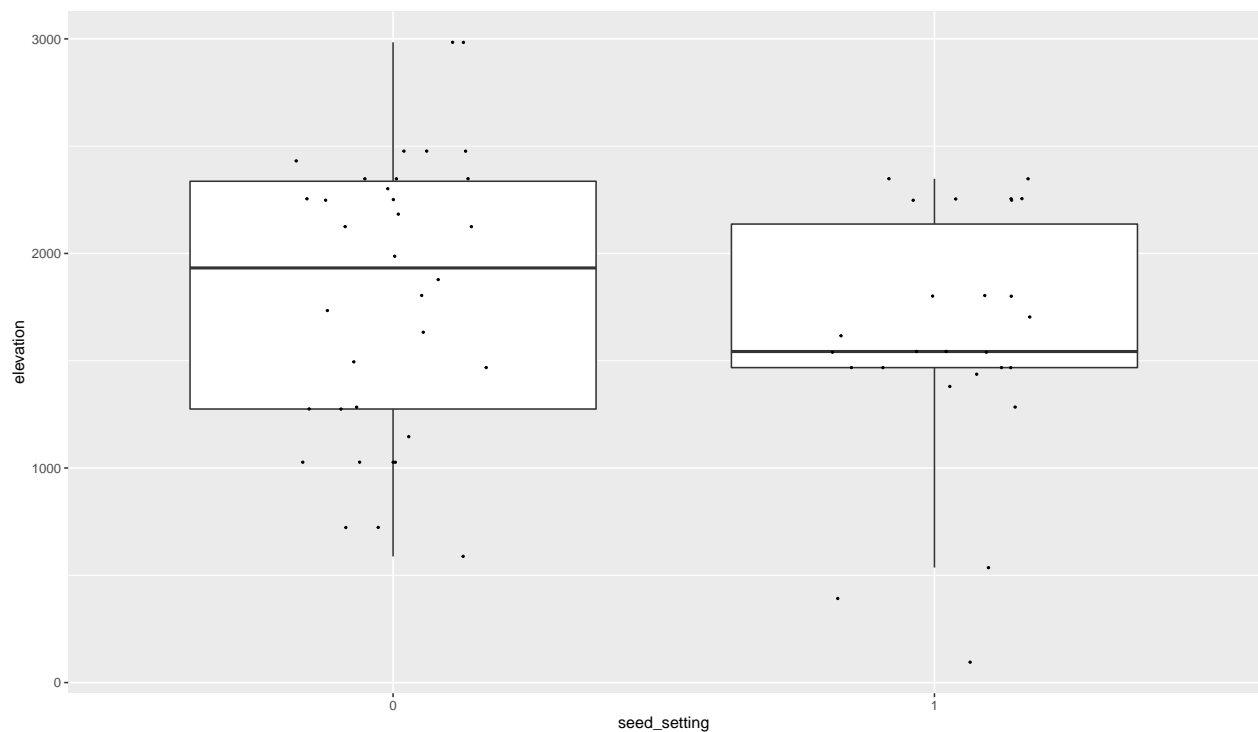
## $statistics
##      MSerror Df      Mean      CV
## 575127.1 58 17833.08 4.252609
##
## $parameters
##      test p.adjusted      name.t ntr alpha
## Fisher-LSD      none seed_setting  2  0.05
##
## $means
##      sr      std  r      LCL      UCL      Min      Max      Q25      Q50
## 0 17595.71 758.0575 34 17335.37 17856.05 15909.75 19384.33 17176.83 17734.92
## 1 18143.50 758.7854 26 17845.79 18441.21 16188.33 19654.75 17850.50 17973.00
##      Q75
## 0 17992.25
## 1 18372.19
##
## $comparison
## NULL
##
## $groups
##      sr groups
## 1 18143.50    a
## 0 17595.71    b
##
## attr(,"class")
## [1] "group"
```

D. Elevation

```
ggdensity(central_american$elevation,
           main = "Density plot of elevation",
           xlab = "Elavations of the genotypes country of origin")
```



```
ggplot(data = central_american, mapping = aes(x = seed_setting, y = elevation)) +  
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_27 <-glm(elevation ~ seed_setting , data = central_american, family = gaussian)  
summary(md_27)
```

```
##  
## Call:
```

```
## glm(formula = elevation ~ seed_setting, family = gaussian, data = central_american)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1514.38   -329.13     1.13    539.65   1175.65
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1808.4      107.7   16.786 <2e-16 ***
## seed_setting1   -199.0      163.7   -1.216   0.229
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 394613.1)
##
##      Null deviance: 23470831  on 59  degrees of freedom
## Residual deviance: 22887562  on 58  degrees of freedom
## (7 observations deleted due to missingness)
## AIC: 947.38
##
## Number of Fisher Scoring iterations: 2
```

```
lrtest(md_27)
```

```
## Likelihood ratio test
##
## Model 1: elevation ~ seed_setting
## Model 2: elevation ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    3 -470.69
## 2    2 -471.44 -1 1.5099      0.2192
```

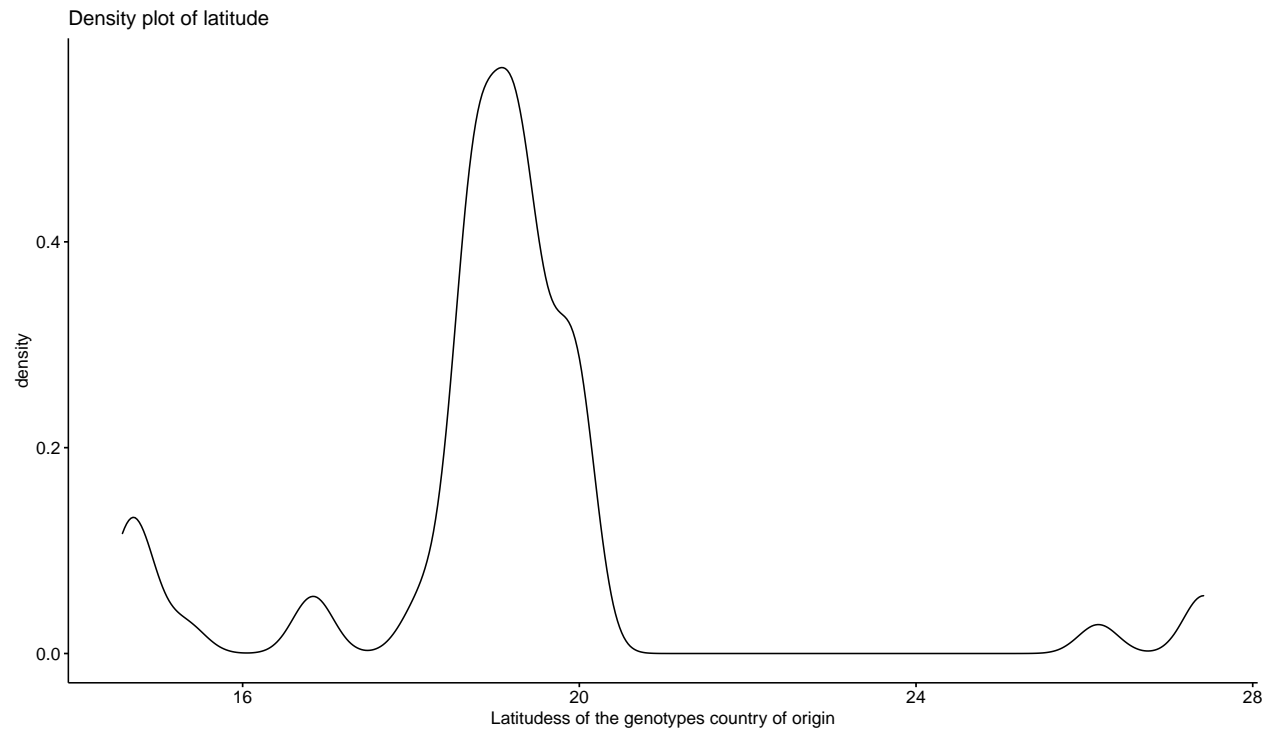
```
lsmeans_ss_elevation <- LSD.test(md_27, "seed_setting", p.adj="none" )
lsmeans_ss_elevation
```

```
## $statistics
##      MSerror Df      Mean      CV
##   394613.1 58 1722.133 36.47699
##
## $parameters
##      test p.adjusted      name.t ntr alpha
## Fisher-LSD      none seed_setting  2  0.05
##
## $means
##   elevation      std r      LCL      UCL Min Max  Q25  Q50  Q75
## 0  1808.353 660.5135 34 1592.703 2024.003 588 2984 1275 1932.5 2336.5
## 1  1609.385 582.7653 26 1362.780 1855.990  95 2348 1468 1543.0 2137.0
##
## $comparison
## NULL
##
## $groups
##   elevation groups
## 0  1808.353      a
## 1  1609.385      a
```

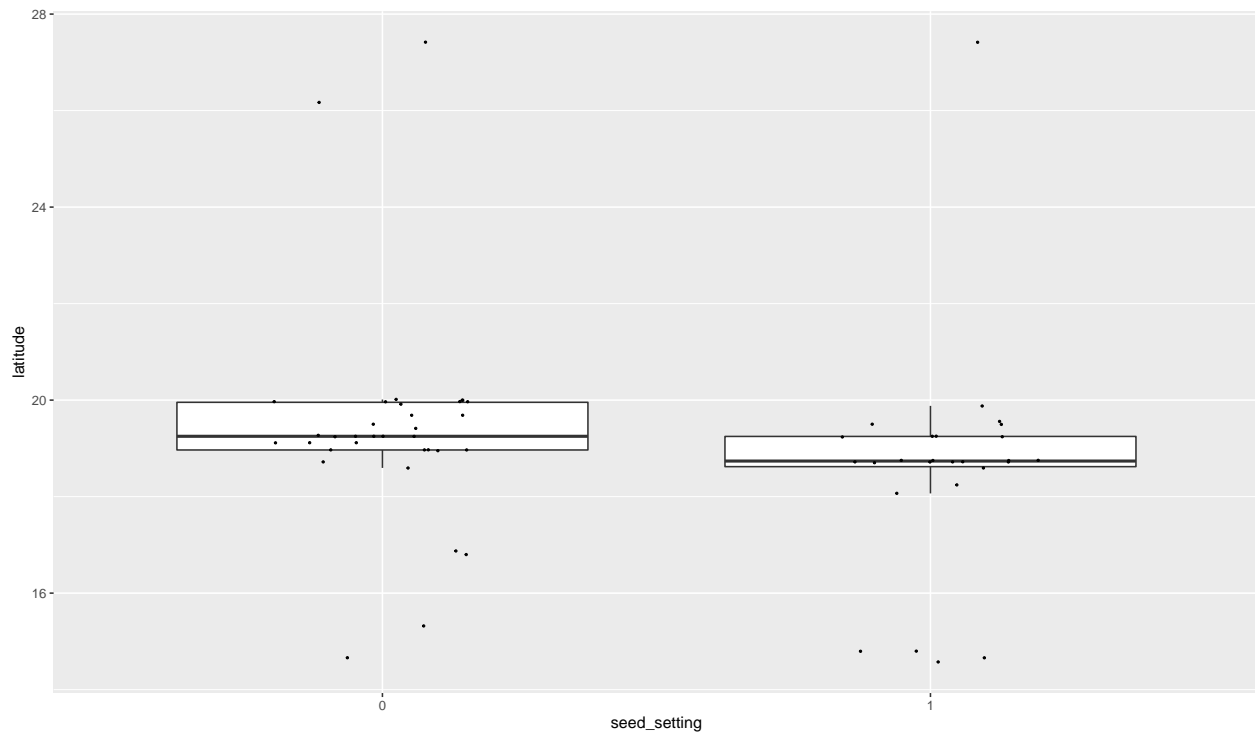
```
##
## attr("class")
## [1] "group"
```

E. Latitude

```
ggsdensity(central_american$latitude,
  main = "Density plot of latitude",
  xlab = "Latitudess of the genotypes country of origin")
```



```
ggplot(data = central_american, mapping = aes(x = seed_setting, y = latitude)) +
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_28 <-glm(latitude ~ seed_setting , data = central_american, family = gaussian)
summary(md_28)
```

```
##
## Call:
## glm(formula = latitude ~ seed_setting, family = gaussian, data = central_american)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -4.7613  -0.4541   0.0987   0.5459   8.8070
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    19.4208     0.3980  48.795  <2e-16 ***
## seed_setting1  -0.8111     0.6046  -1.342   0.185
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 5.385976)
##
##      Null deviance: 322.08  on 59  degrees of freedom
## Residual deviance: 312.39  on 58  degrees of freedom
## (7 observations deleted due to missingness)
## AIC: 275.27
##
## Number of Fisher Scoring iterations: 2
```

```
lrtest(md_28)
```

```
## Likelihood ratio test
##
## Model 1: latitude ~ seed_setting
```

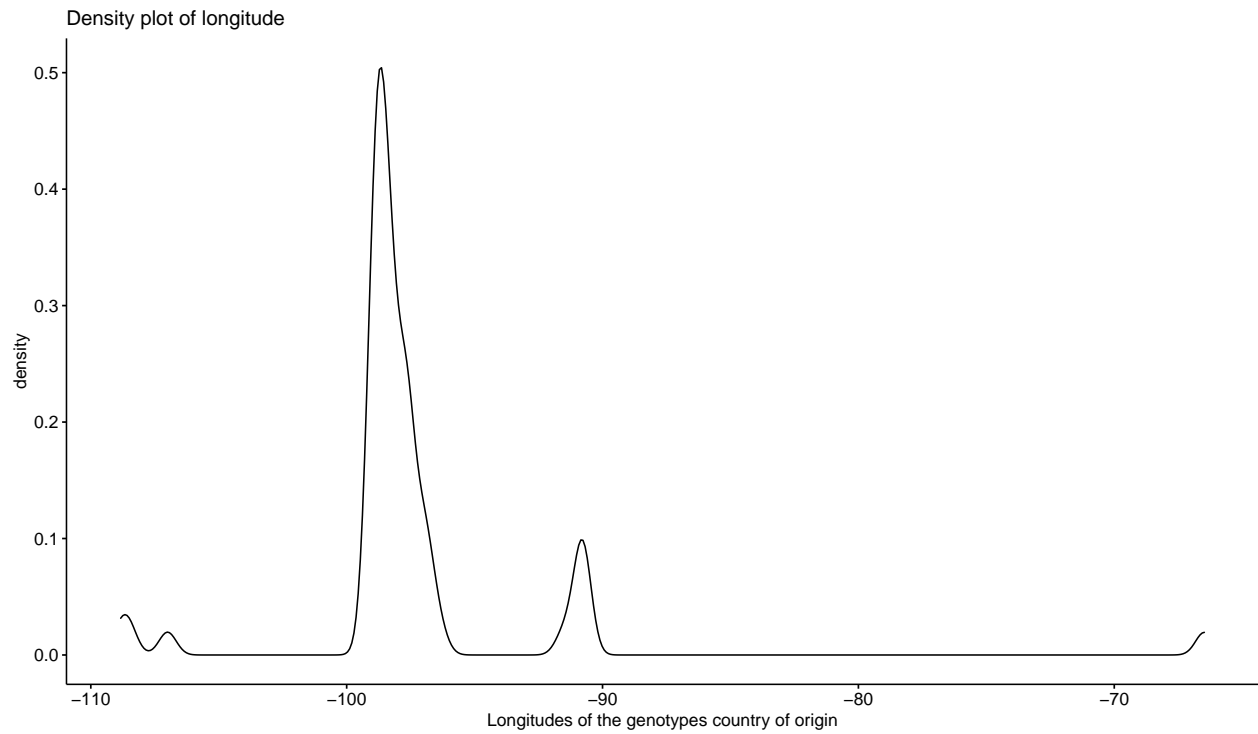
```
## Model 2: latitude ~ 1
##   #Df LogLik Df   Chisq Pr(>Chisq)
## 1    3 -134.63
## 2    2 -135.55 -1 1.8335      0.1757

lsmeans_ss_latitude <- LSD.test(md_28, "seed_setting", p.adj="none" )
lsmeans_ss_latitude

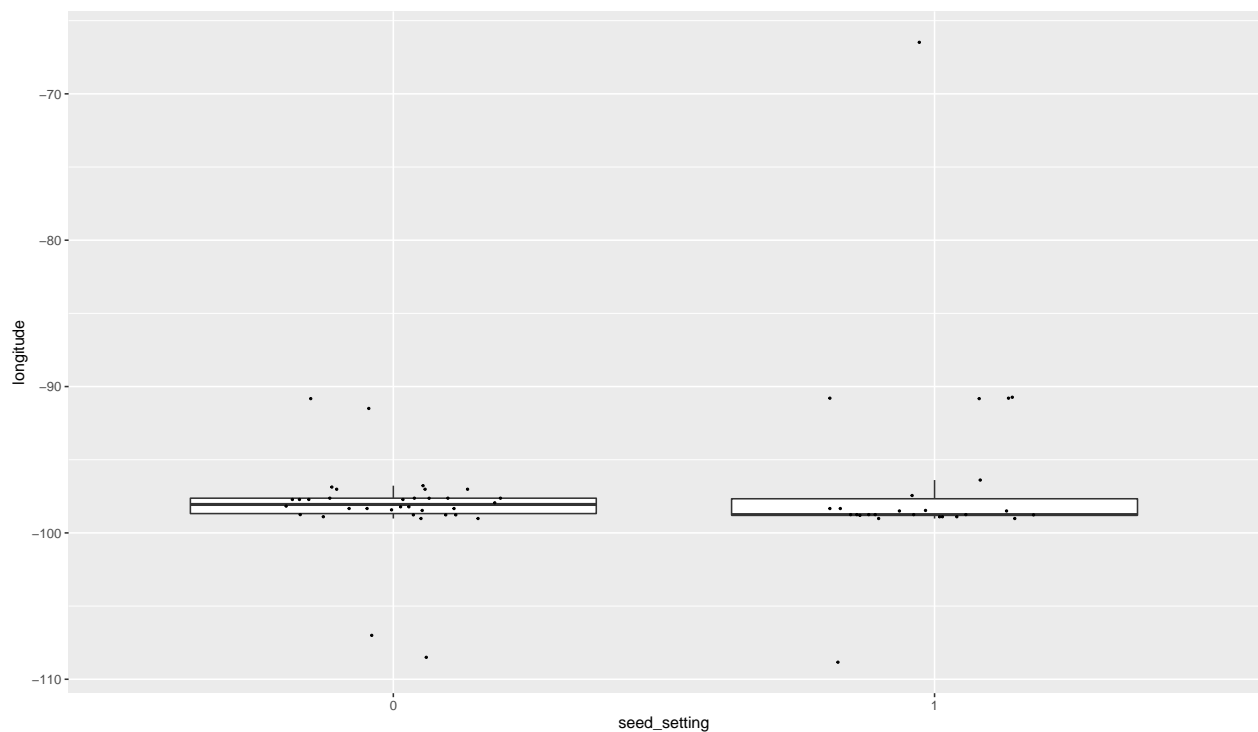
## $statistics
##   MSerror Df      Mean      CV
## 5.385976 58 19.06931 12.17018
##
## $parameters
##      test p.adjusted      name.t ntr alpha
## Fisher-LSD      none seed_setting  2 0.05
##
## $means
##   latitude      std r      LCL      UCL      Min      Max      Q25      Q50
## 0 19.42080 2.253246 34 18.62410 20.21750 14.65946 27.41667 18.96667 19.25000
## 1 18.60967 2.407004 26 17.69861 19.52074 14.57172 27.41667 18.62092 18.73712
##      Q75
## 0 19.95417
## 1 19.24726
##
## $comparison
## NULL
##
## $groups
##   latitude groups
## 0 19.42080      a
## 1 18.60967      a
##
## attr(,"class")
## [1] "group"
```

F. Longitude

```
ggdensity(central_american$longitude,
           main = "Density plot of longitude",
           xlab = "Longitudes of the genotypes country of origin")
```



```
ggplot(data = central_american, mapping = aes(x = seed_setting, y = longitude)) +  
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_29 <-glm(longitude ~ seed_setting , data = central_american, family = gaussian)  
summary(md_29)
```

```
##  
## Call:
```

```
## glm(formula = longitude ~ seed_setting, family = gaussian, data = central_american)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -12.3254  -2.2421  -0.2700   0.5467  30.0290
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -98.1745     0.8925 -109.996 <2e-16 ***
## seed_setting1   1.6666     1.3558   1.229   0.224
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 27.08458)
##
##      Null deviance: 1611.8  on 59  degrees of freedom
## Residual deviance: 1570.9  on 58  degrees of freedom
## (7 observations deleted due to missingness)
## AIC: 372.18
##
## Number of Fisher Scoring iterations: 2
```

```
lrtest(md_29)
```

```
## Likelihood ratio test
##
## Model 1: longitude ~ seed_setting
## Model 2: longitude ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    3 -183.09
## 2    2 -183.86 -1 1.5431    0.2142
```

```
lsmeans_ss_longitude <- LSD.test(md_29, "seed_setting", p.adj="none" )
lsmeans_ss_longitude
```

```
## $statistics
##      MSerror Df      Mean      CV
## 27.08458 58 -97.45231 -5.34034
##
## $parameters
##      test p.adjusted      name.t ntr alpha
## Fisher-LSD      none seed_setting 2 0.05
##
## $means
## longitude      std r      LCL      UCL      Min      Max      Q25
## 0 -98.17452 2.998726 34 -99.96110 -96.38793 -108.5000 -90.82485 -98.67867
## 1 -96.50789 7.139070 26 -98.55093 -94.46484 -108.8333 -66.47890 -98.78987
##      Q50      Q75
## 0 -98.05569 -97.62778
## 1 -98.75000 -97.66687
##
## $comparison
## NULL
##
## $groups
```



```
## longitude groups
## 1 -96.50789      a
## 0 -98.17452      a
##
## attr(,"class")
## [1] "group"
```

5. Focus on Central American *A. hypochondriacus* Accessions

Central America originated *A. hypochondriacus* accessions are the most diverse in flowering time and seed setting. Therefore, I would like to focus on that.

```
central_american_hypo <- filter(data, (country %in% c("GTM", "MEX", "PTR")))%>%
  filter(species=="hypo")

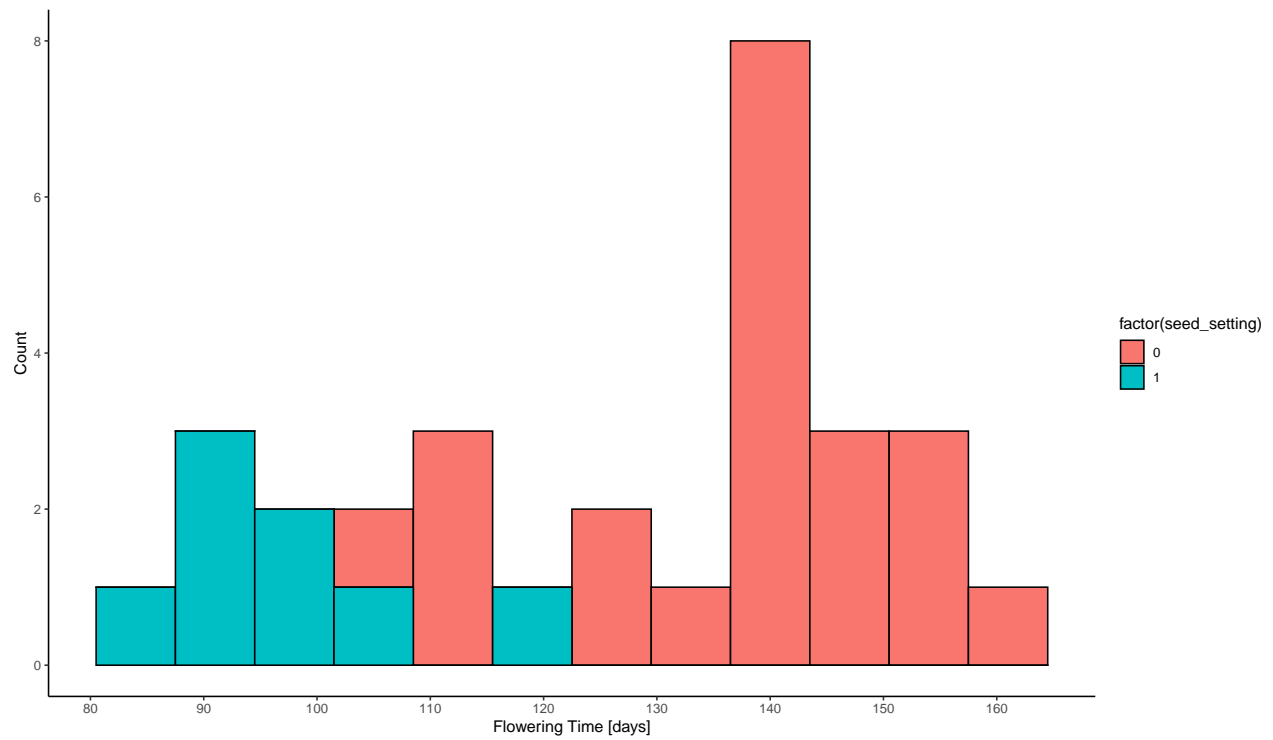
central_american_hypo$seed_setting <- as.factor(central_american_hypo$seed_setting)

central_american_hypo_plot <- ggplot(central_american_hypo, aes(x=ft_adjusted, fill=factor(seed_setting))) +
  geom_histogram(color="black", binwidth = 7) +
  scale_x_continuous(breaks = scales::pretty_breaks(n = 10) ) +
  #scale_y_continuous(expand = expand_scale(mult = 0, add = 0.1)) +
  labs( y="Count", x="Flowering Time [days]") +

  theme(axis.text.x = element_text(colour = "grey20", size = 4.5 , hjust = 0.5, vjust = 0.5),
        axis.text.y = element_text(colour = "grey20", size = 4.5),
        text = element_text(size = 5), legend.position=c(0.9, 0.9),
        panel.spacing = unit(1.0, "lines"),
        legend.title = element_text(size=5, face="bold"),
        legend.key.size = unit(4, "mm"),
        legend.text=element_text(size=5)) +
  #scale_fill_discrete(name = "Species", labels=c("cru", "hypo")) +
  scale_color_discrete(name = "seed_setting", labels=c("0", "1")) +
  theme_classic()

central_american_hypo_plot

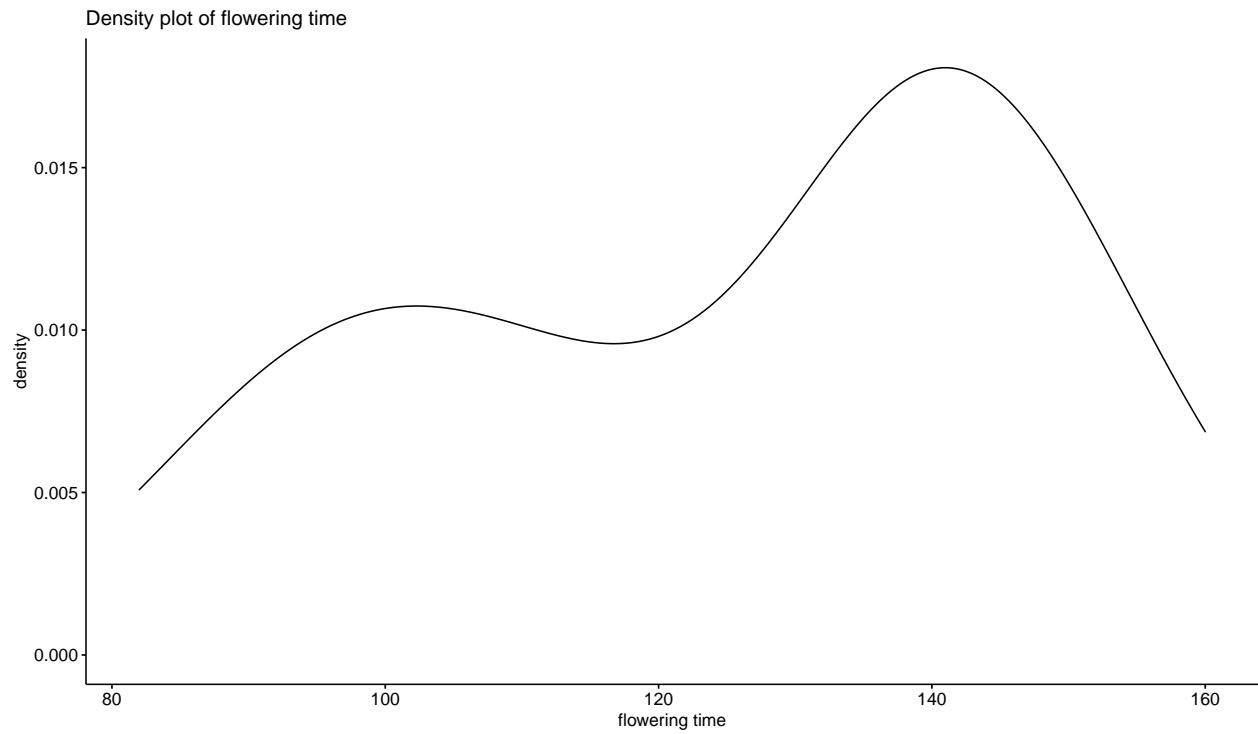
## Warning: Removed 2 rows containing non-finite values (stat_bin).
```



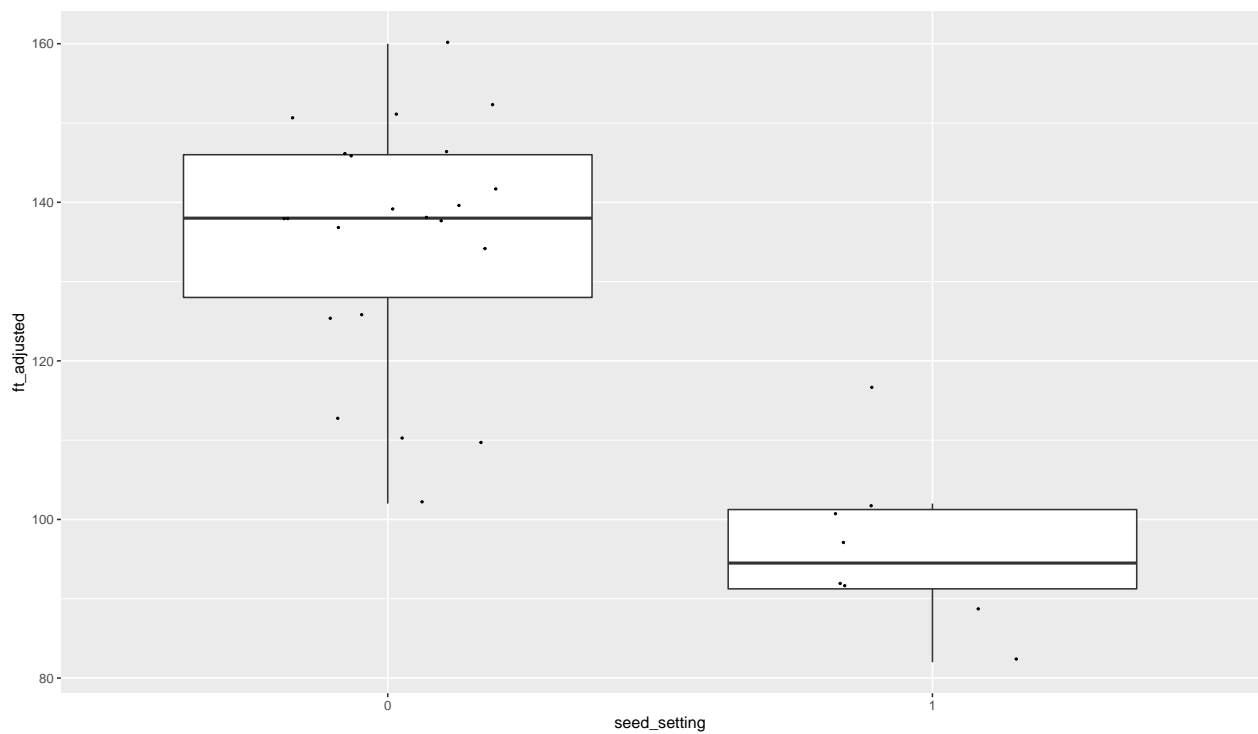
Phenotypic traits

A. Flowering time

```
ggdensity(central_american_hypo$ft_adjusted,  
  main = "Density plot of flowering time",  
  xlab = "flowering time")
```



```
ggplot(data = central_american_hypo, mapping = aes(x = seed_setting, y = ft_adjusted)) +
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_30 <-glm(ft_adjusted ~ seed_setting , data = central_american_hypo, family = gaussian)
summary(md_30)
```

```
##
## Call:
```

```
## glm(formula = ft_adjusted ~ seed_setting, family = gaussian,
##      data = central_american_hypo)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -33.545   -6.750    2.455    9.455   24.455
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    135.545      3.041  44.577 < 2e-16 ***
## seed_setting1  -39.045      5.888  -6.631 3.41e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 203.4091)
##
##      Null deviance: 14639.5  on 29  degrees of freedom
## Residual deviance:  5695.5  on 28  degrees of freedom
## (2 observations deleted due to missingness)
## AIC: 248.52
##
## Number of Fisher Scoring iterations: 2
```

```
lrtest(md_30)
```

```
## Likelihood ratio test
##
## Model 1: ft_adjusted ~ seed_setting
## Model 2: ft_adjusted ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    3 -121.26
## 2    2 -135.42 -1 28.322  1.027e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

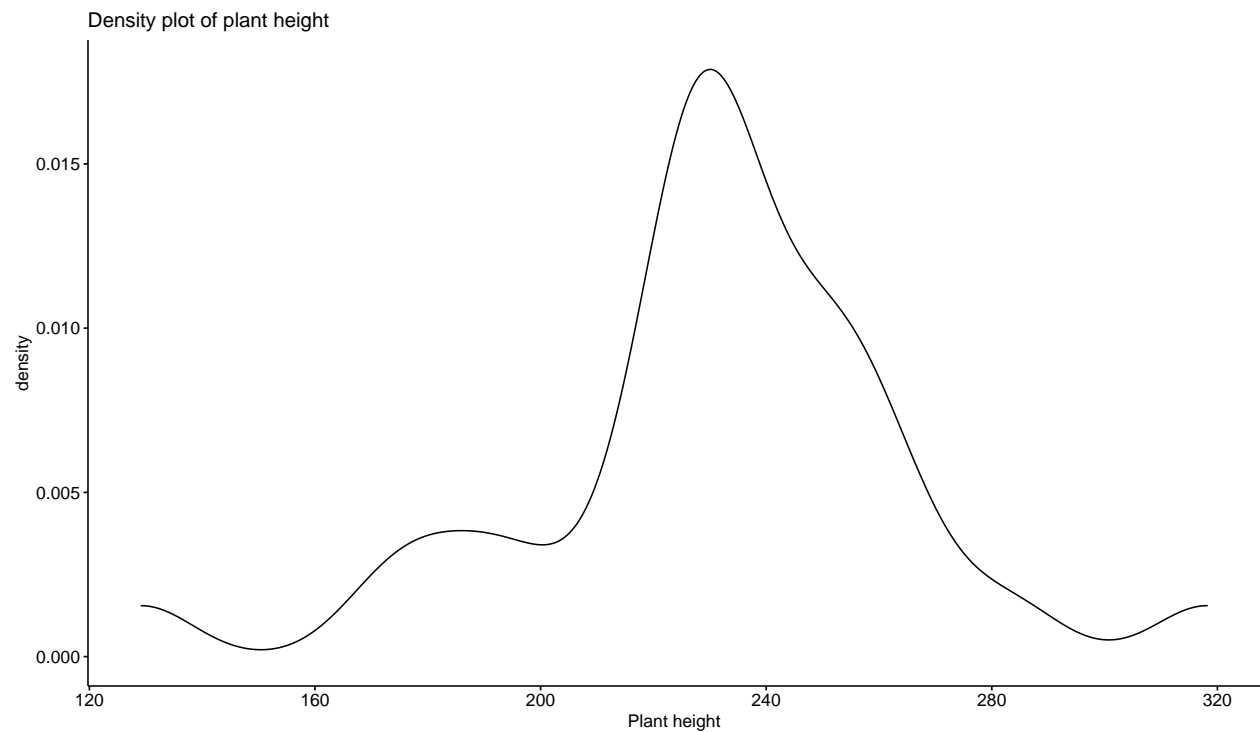
```
lsmeans_ss_ft <- LSD.test(md_30, "seed_setting", p.adj="none" )
lsmeans_ss_ft
```

```
## $statistics
##      MSError Df      Mean      CV
##    203.4091 28 125.1333 11.39757
##
## $parameters
##      test p.adjusted      name.t ntr alpha
## Fisher-LSD      none seed_setting  2 0.05
##
## $means
##   ft_adjusted      std r      LCL      UCL Min Max      Q25      Q50      Q75
## 0    135.5455 15.30243 22 129.31686 141.774 102 160 128.00 138.0 146.00
## 1     96.5000 10.54243  8  86.17104 106.829  82 117  91.25  94.5 101.25
##
## $comparison
## NULL
##
## $groups
```

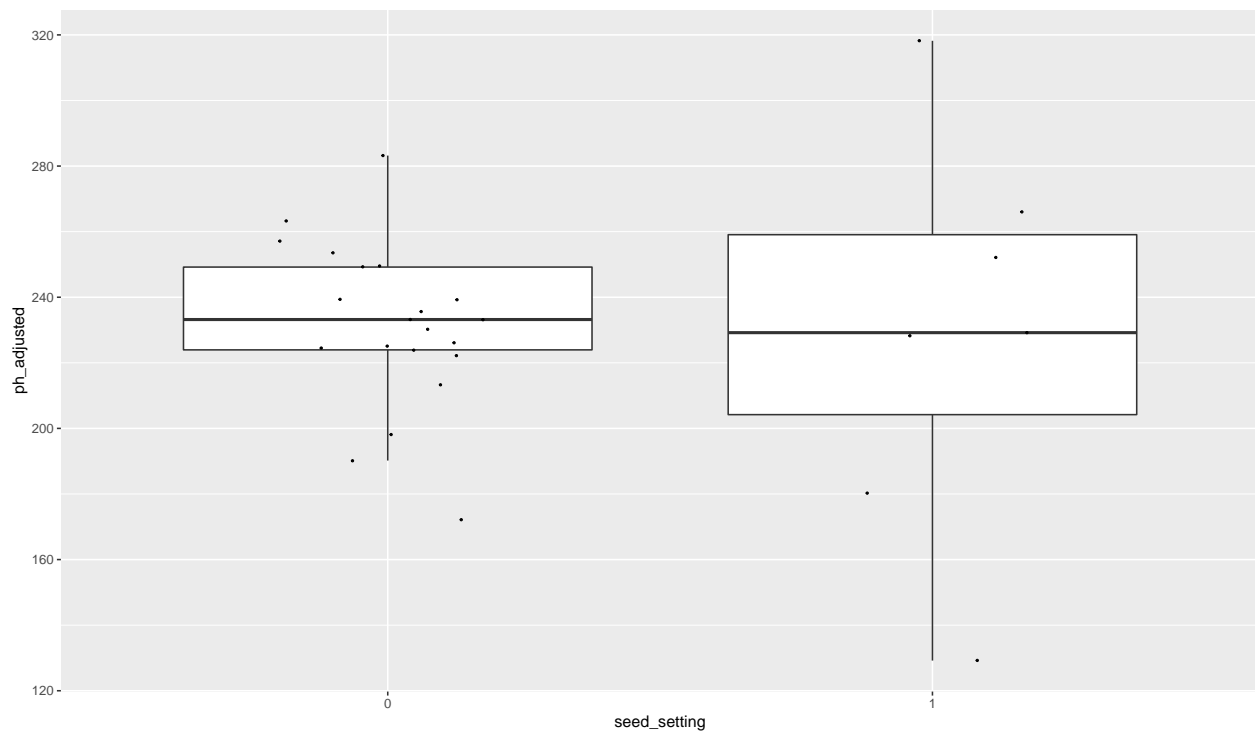
```
##   ft_adjusted groups
## 0    135.5455    a
## 1    96.5000    b
##
## attr(,"class")
## [1] "group"
```

B. Plant height

```
ggdensity(central_american_hypo$ph_adjusted,
  main = "Density plot of plant height",
  xlab = "Plant height")
```



```
ggplot(data = central_american_hypo, mapping = aes(x = seed_setting, y = ph_adjusted)) +
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_31 <-glm(ph_adjusted ~ seed_setting , data = central_american_hypo, family = gaussian)
summary(md_31)
```

```
##
## Call:
## glm(formula = ph_adjusted ~ seed_setting, family = gaussian,
##      data = central_american_hypo)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -99.821   -8.033    0.917   18.905   89.179
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    231.545     7.976   29.031  <2e-16 ***
## seed_setting1    -2.524    15.951   -0.158    0.876
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 1335.861)
##
##      Null deviance: 34766  on 27  degrees of freedom
## Residual deviance: 34732  on 26  degrees of freedom
## (4 observations deleted due to missingness)
## AIC: 284.91
##
## Number of Fisher Scoring iterations: 2
```

```
lrtest(md_31)
```

```
## Likelihood ratio test
##
```

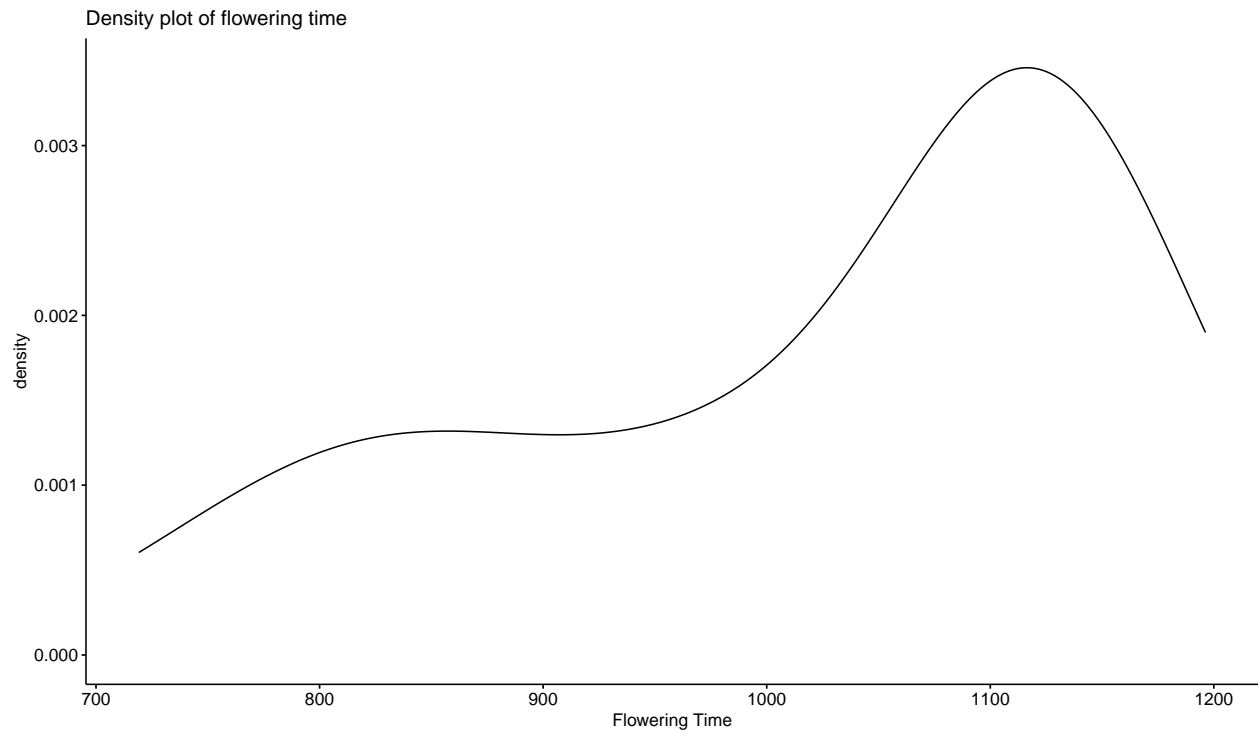
```
## Model 1: ph_adjusted ~ seed_setting
## Model 2: ph_adjusted ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    3 -139.46
## 2    2 -139.47 -1 0.0269      0.8696

lsmeans_ss_ph <- LSD.test(md_31, "seed_setting", p.adj="none" )
lsmeans_ss_ph

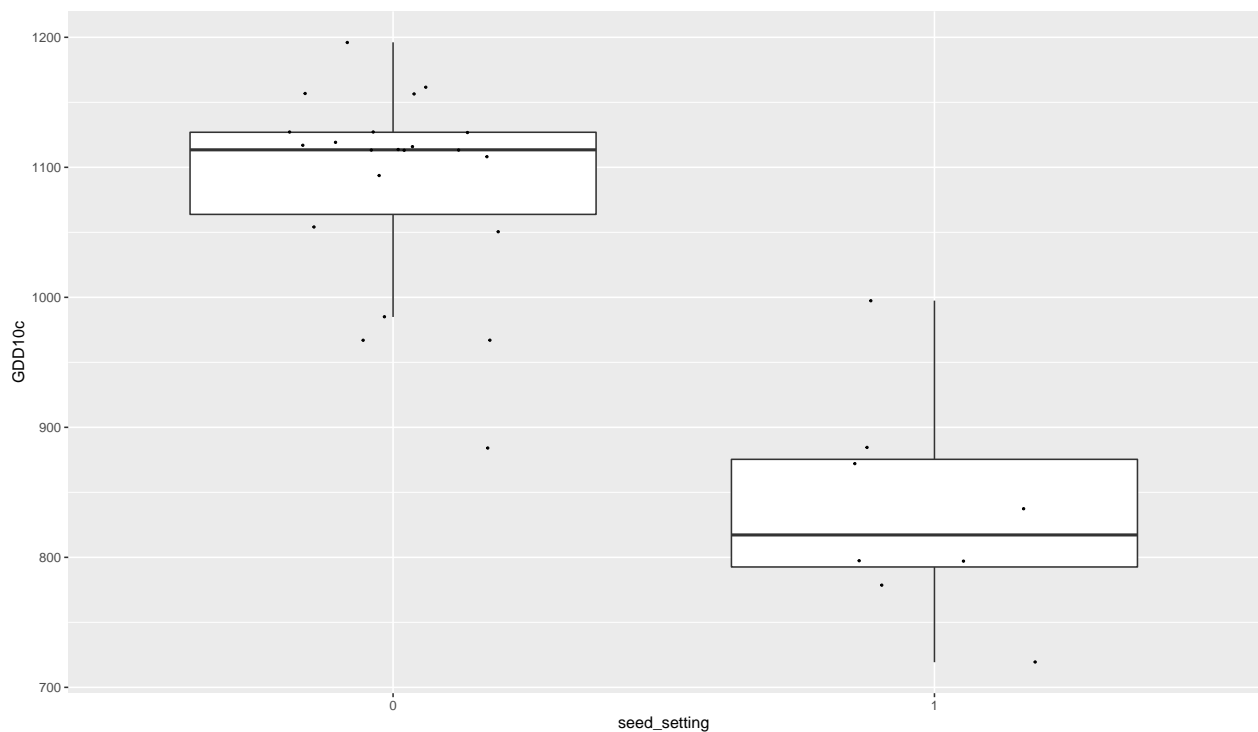
## $statistics
##      MSerror Df      Mean      CV
## 1335.861 26 230.9143 15.82814
##
## $parameters
##      test p.adjusted      name.t ntr alpha
## Fisher-LSD      none seed_setting  2 0.05
##
## $means
##   ph_adjusted      std r      LCL      UCL  Min  Max   Q25   Q50   Q75
## 0    231.5452 25.03004 21 215.1509 247.9396 172.2 283.2 223.95 233.2 249.200
## 1    229.0214 60.83083  7 200.6255 257.4173 129.2 318.2 204.20 229.2 259.075
##
## $comparison
## NULL
##
## $groups
##   ph_adjusted groups
## 0    231.5452      a
## 1    229.0214      a
##
## attr(,"class")
## [1] "group"
```

C. Growing degree days (Tbase is 10°C)

```
ggdensity(central_american_hypo$GDD10c,
  main = "Density plot of flowering time",
  xlab = "Flowering Time")
```



```
ggplot(data = central_american_hypo, mapping = aes(x = seed_setting, y = GDD10c)) +  
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_32 <-glm(GDD10c ~ seed_setting , data = central_american_hypo, family = gaussian)  
summary(md_32)
```

```
##  
## Call:
```



```

## glm(formula = GDD10c ~ seed_setting, family = gaussian, data = central_american_hypo)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -205.06   -38.17    24.04    37.54   161.98
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1089.41      16.63  65.528 < 2e-16 ***
## seed_setting1  -253.94      32.19  -7.888 1.36e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 6080.627)
##
##      Null deviance: 548580  on 29  degrees of freedom
## Residual deviance: 170258  on 28  degrees of freedom
## (2 observations deleted due to missingness)
## AIC: 350.45
##
## Number of Fisher Scoring iterations: 2

```

```

lrtest(md_32)

```

```

## Likelihood ratio test
##
## Model 1: GDD10c ~ seed_setting
## Model 2: GDD10c ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    3 -172.23
## 2    2 -189.78 -1 35.101  3.131e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

lsmeans_ss_GDD10c <- LSD.test(md_32, "seed_setting", p.adj="none" )
lsmeans_ss_GDD10c

```

```

## $statistics
##      MSerror Df      Mean      CV
##    6080.627 28 1021.693 7.632268
##
## $parameters
##      test p.adjusted      name.t ntr alpha
## Fisher-LSD      none seed_setting  2  0.05
##
## $means
##      GDD10c      std r      LCL      UCL      Min      Max      Q25      Q50
## 0 1089.4114 75.78639 22 1055.3565 1123.4662 884.35 1196.15 1063.8000 1113.450
## 1  835.4688 84.21270  8  778.9952  891.9423 719.35  997.45  792.5875  817.225
##      Q75
## 0 1126.9500
## 1  875.3875
##
## $comparison
## NULL

```

```
##
## $groups
##      GDD10c groups
## 0 1089.4114      a
## 1  835.4688      b
##
## attr(,"class")
## [1] "group"
table(central_american_hypo$seed_setting)
```

```
##
##  0  1
## 24  8
table(central_american_hypo$seed_setting)
```

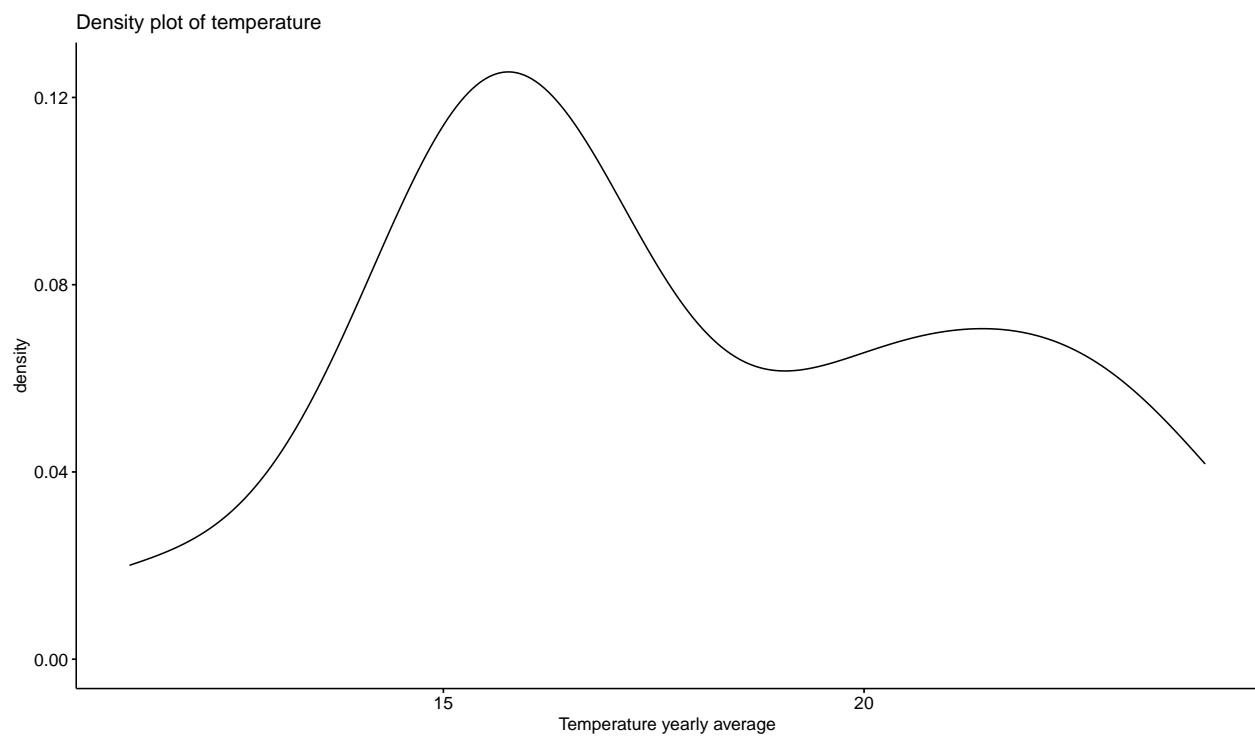
```
##
##  0  1
## 24  8
```

Climatic & Geographical Variables

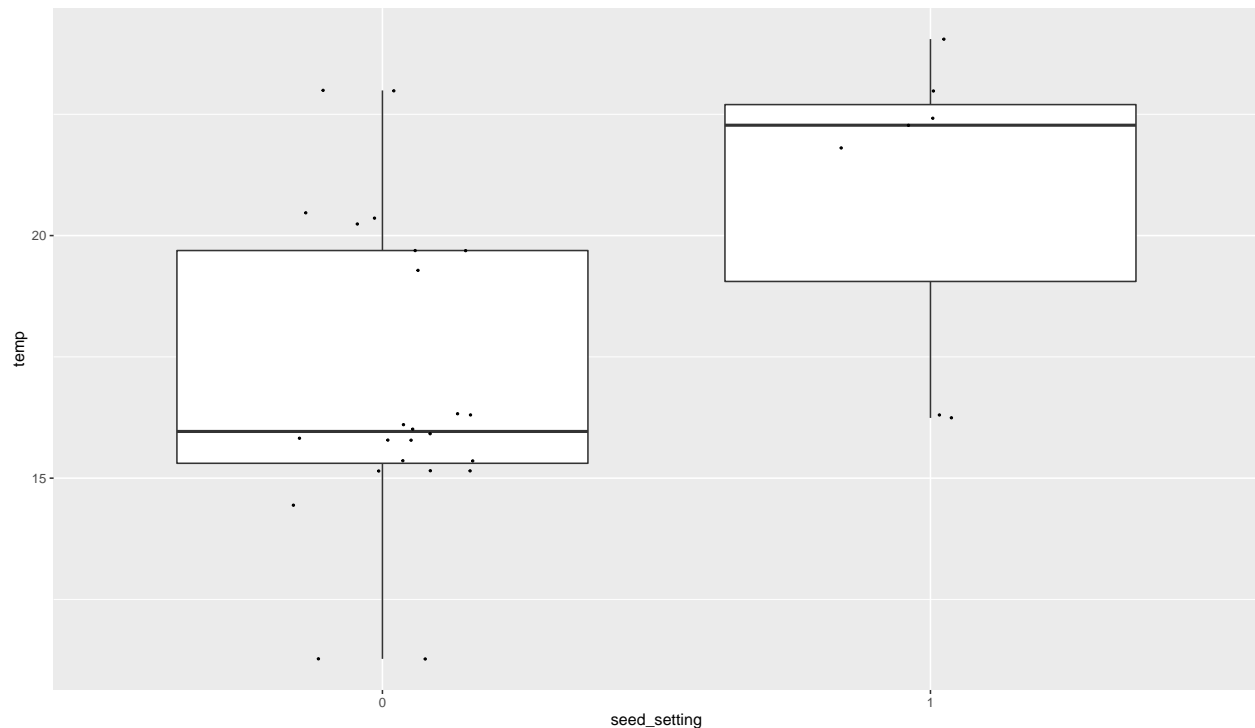
A. Temperature

Distribution of the data.

```
ggdensity(central_american_hypo$temp,
  main = "Density plot of temperature",
  xlab = "Temperature yearly average")
```



```
ggplot(data = central_american_hypo, mapping = aes(x = seed_setting, y = temp)) +
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_33 <-glm(temp ~ seed_setting , data = central_american_hypo, family = gaussian)
summary(md_33)
```

```
##
## Call:
## glm(formula = temp ~ seed_setting, family = gaussian, data = central_american_hypo)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -5.6795  -1.7003  -0.8545   2.5330   6.0372
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    16.954     0.638  26.575  <2e-16 ***
## seed_setting1     3.913     1.343   2.915  0.0068 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 9.76891)
##
##      Null deviance: 366.29  on 30  degrees of freedom
## Residual deviance: 283.30  on 29  degrees of freedom
## (1 observation deleted due to missingness)
## AIC: 162.56
##
## Number of Fisher Scoring iterations: 2
```

```

lrtest(md_33)

## Likelihood ratio test
##
## Model 1: temp ~ seed_setting
## Model 2: temp ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    3 -78.281
## 2    2 -82.263 -1  7.9648   0.00477 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

lsmeans_ss_temp <- LSD.test(md_33, "seed_setting", p.adj="none" )
lsmeans_ss_temp

## $statistics
##      MSerror Df      Mean      CV
##    9.76891 29 17.83817 17.52156
##
## $parameters
##      test p.adjusted      name.t ntr alpha
## Fisher-LSD      none seed_setting  2  0.05
##
## $means
##      temp      std  r      LCL      UCL      Min      Max      Q25      Q50
## 0 16.95451 3.100959 24 15.64967 18.25936 11.27500 22.99167 15.30625 15.9625
## 1 20.86786 3.217960  7 18.45175 23.28396 16.24167 24.05000 19.05417 22.2750
##      Q75
## 0 19.69167
## 1 22.70000
##
## $comparison
## NULL
##
## $groups
##      temp groups
## 1 20.86786      a
## 0 16.95451      b
##
## attr("class")
## [1] "group"

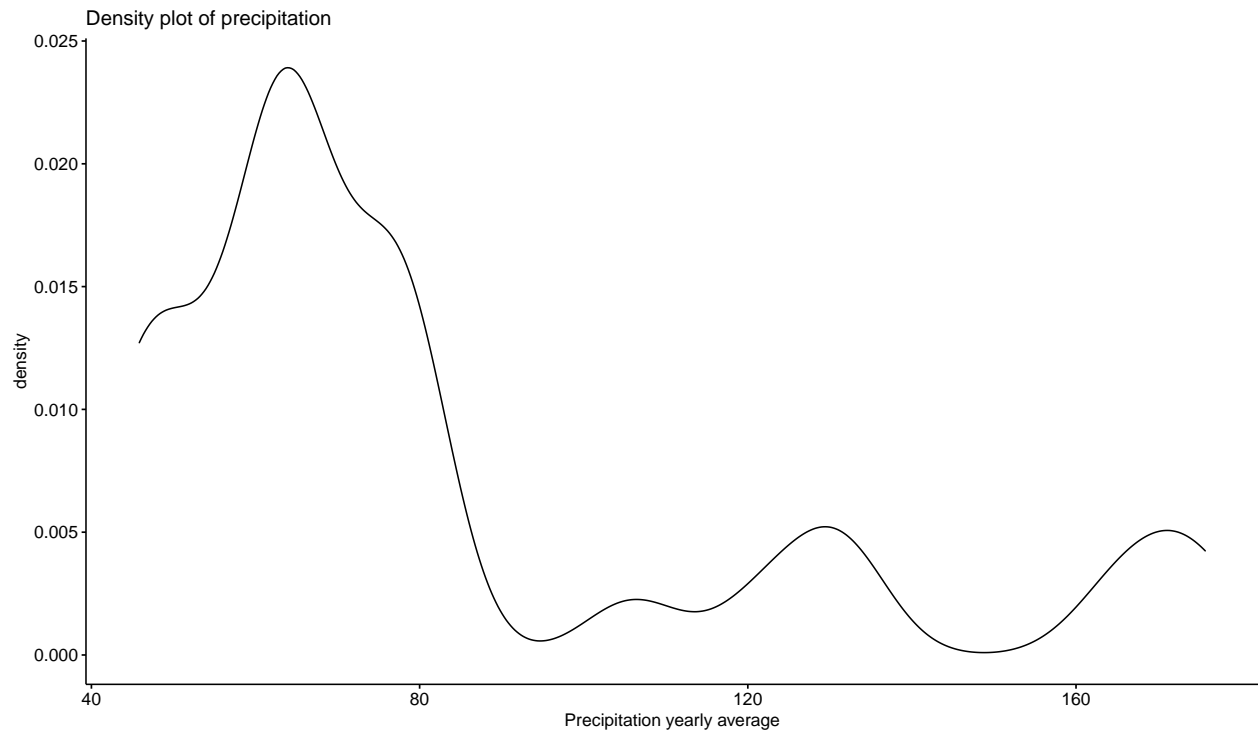
```

B. Precipitation

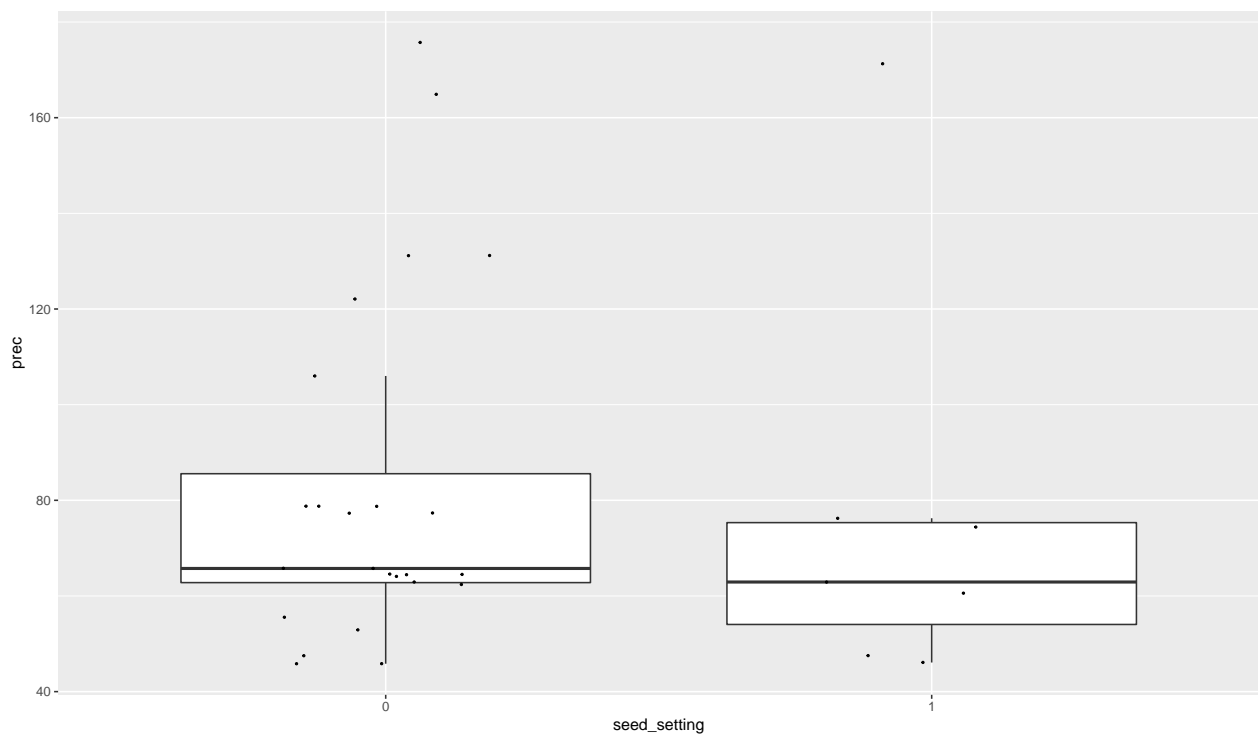
```

ggdensity(central_american_hypo$prec,
  main = "Density plot of precipitation",
  xlab = "Precipitation yearly average")

```



```
ggplot(data = central_american_hypo, mapping = aes(x = seed_setting, y = prec)) +
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_34 <-glm(prec ~ seed_setting , data = central_american_hypo, family = gaussian)
summary(md_34)
```

```
##
## Call:
```

```
## glm(formula = prec ~ seed_setting, family = gaussian, data = central_american_hypo)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -36.837  -20.003  -16.417   -1.667   94.250
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      82.670       7.749  10.668 1.49e-11 ***
## seed_setting1    -5.670      16.308  -0.348   0.731
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 1441.238)
##
##      Null deviance: 41970  on 30  degrees of freedom
## Residual deviance: 41796  on 29  degrees of freedom
## (1 observation deleted due to missingness)
## AIC: 317.38
##
## Number of Fisher Scoring iterations: 2
```

```
lrtest(md_34)
```

```
## Likelihood ratio test
##
## Model 1: prec ~ seed_setting
## Model 2: prec ~ 1
##   #Df LogLik Df Chisq Pr(>Chisq)
## 1    3 -155.69
## 2    2 -155.75 -1 0.129      0.7195
```

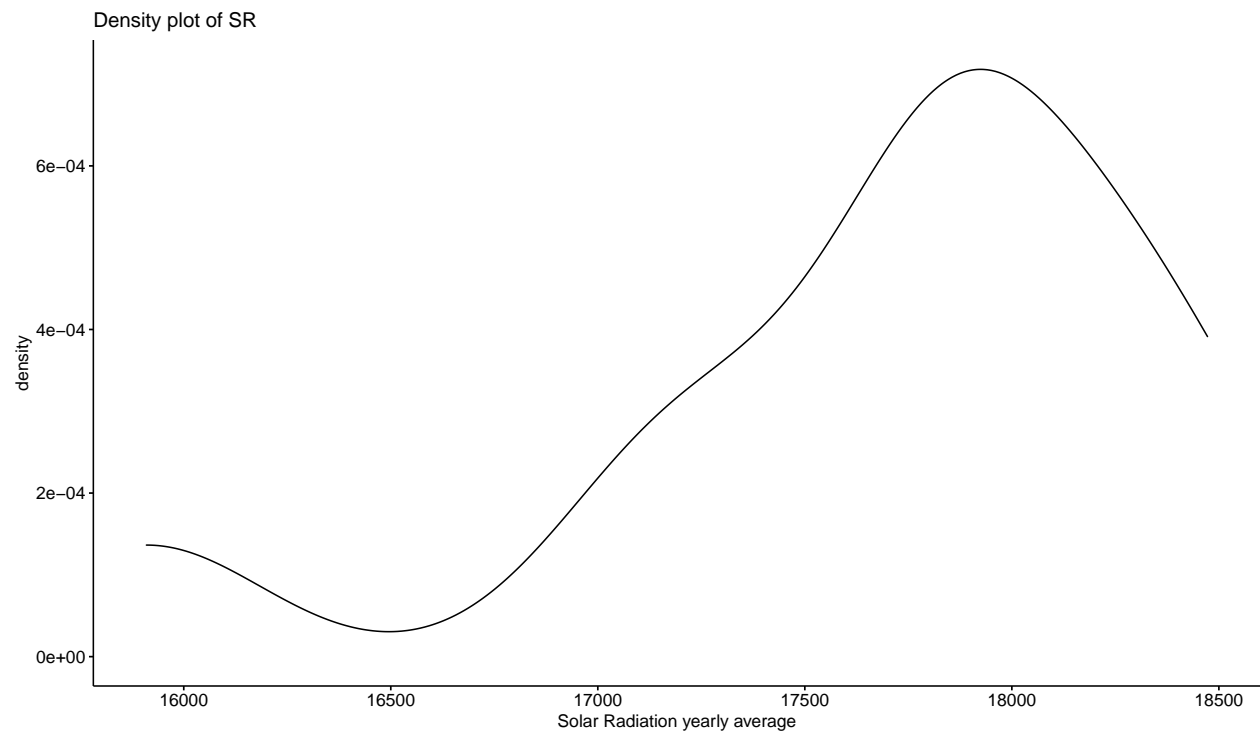
```
lsmeans_ss_prec <- LSD.test(md_34, "seed_setting", p.adj="none" )
lsmeans_ss_prec
```

```
## $statistics
##      MSerror Df      Mean      CV
## 1441.238 29 81.38978 46.64423
##
## $parameters
##      test p.adjusted      name.t ntr alpha
## Fisher-LSD      none seed_setting  2 0.05
##
## $means
##      prec      std r      LCL      UCL      Min      Max      Q25      Q50
## 0 82.67014 36.48207 24 66.82105 98.51923 45.83333 175.75 62.79167 65.75000
## 1 77.00000 43.17452  7 47.65319 106.34681 46.08333 171.25 54.04167 62.91667
##      Q75
## 0 85.56250
## 1 75.33333
##
## $comparison
## NULL
##
## $groups
```

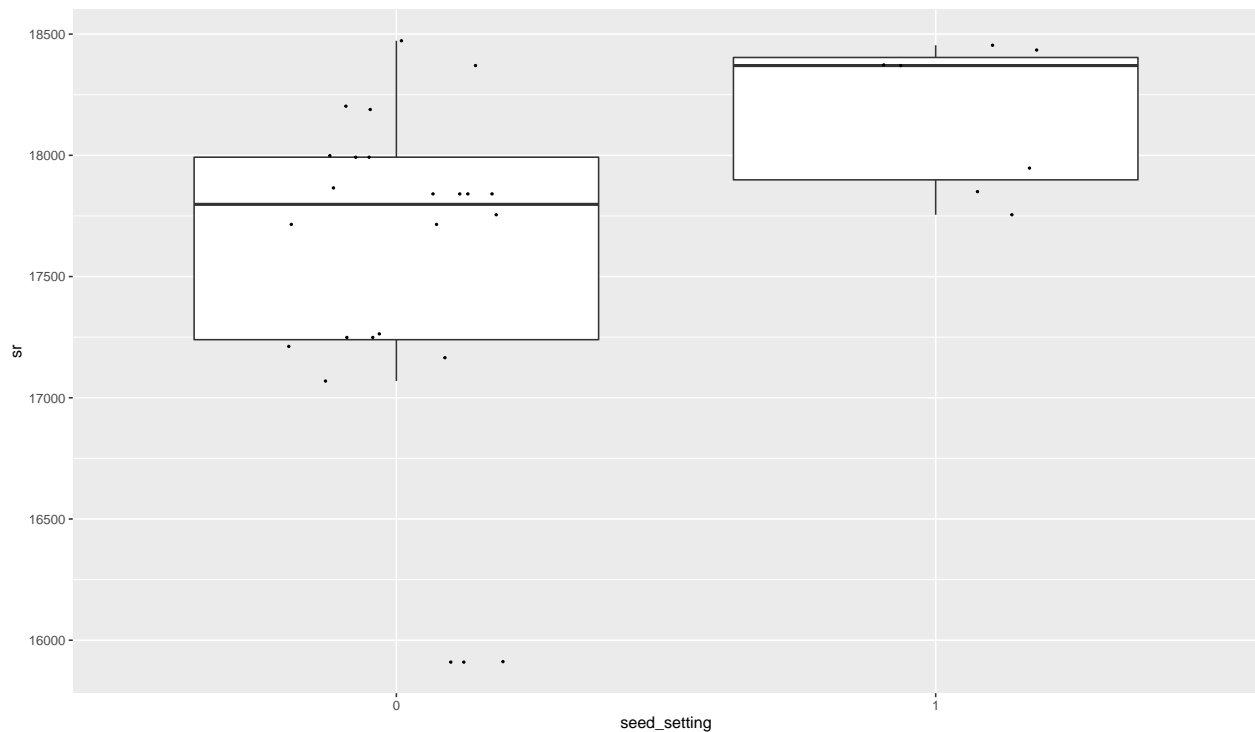
```
##      prec groups
## 0 82.67014      a
## 1 77.00000      a
##
## attr("class")
## [1] "group"
```

C. Solar Radiation

```
ggdensity(central_american_hypo$sr,
  main = "Density plot of SR",
  xlab = "Solar Radiation yearly average")
```



```
ggplot(data = central_american_hypo, mapping = aes(x = seed_setting, y = sr)) +
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_35 <-glm(sr ~ seed_setting , data = central_american_hypo, family = gaussian)
summary(md_35)
```

```
##
## Call:
## glm(formula = sr ~ seed_setting, family = gaussian, data = central_american_hypo)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1614.0   -293.4    203.7    329.5    948.8
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   17523.7     135.9 128.984  <2e-16 ***
## seed_setting1    645.4     285.9   2.257   0.0317 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 442986.6)
##
##      Null deviance: 15104019  on 30  degrees of freedom
## Residual deviance: 12846612  on 29  degrees of freedom
## (1 observation deleted due to missingness)
## AIC: 494.95
##
## Number of Fisher Scoring iterations: 2
```

```
lrtest(md_35)
```

```
## Likelihood ratio test
##
## Model 1: sr ~ seed_setting
```



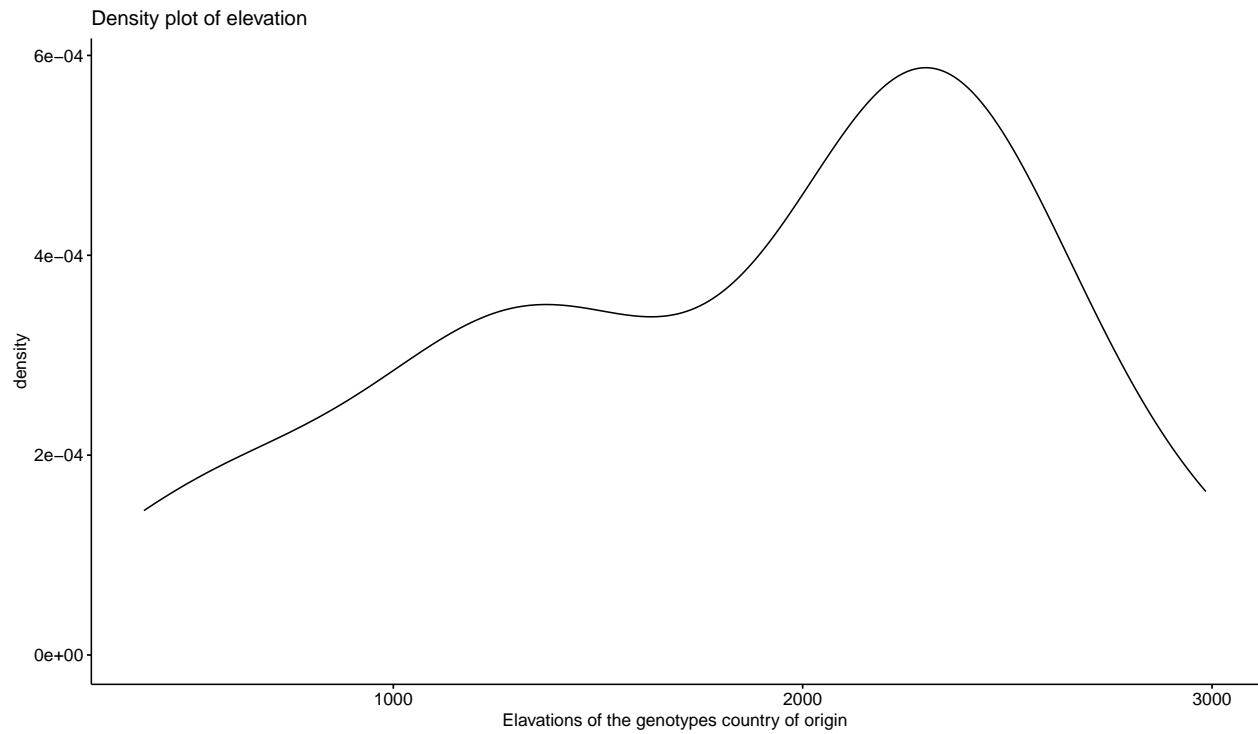
```
## Model 2: sr ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    3 -244.47
## 2    2 -246.98 -1 5.0183    0.02508 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

lsmeans_ss_sr <- LSD.test(md_35, "seed_setting", p.adj="none" )
lsmeans_ss_sr

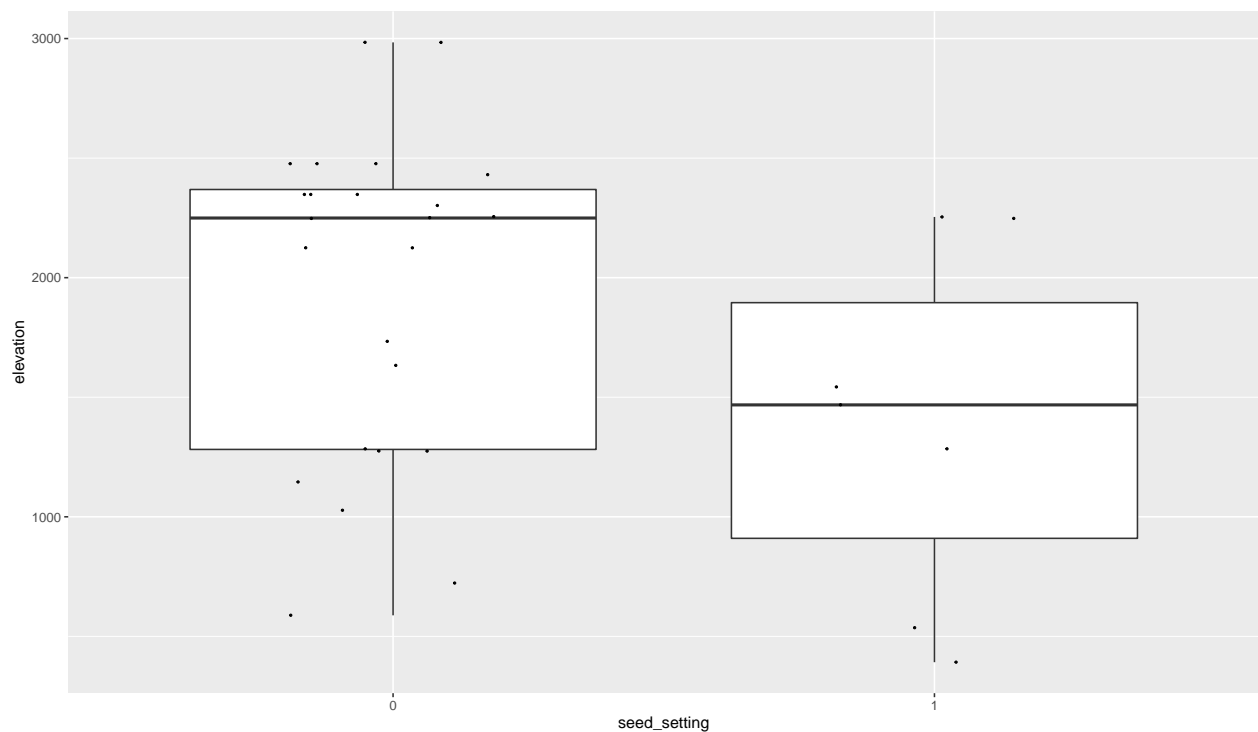
## $statistics
##      MSerror Df      Mean      CV
## 442986.6 29 17669.48 3.766792
##
## $parameters
##      test p.adjusted      name.t ntr alpha
## Fisher-LSD      none seed_setting  2 0.05
##
## $means
##      sr      std r      LCL      UCL      Min      Max      Q25      Q50
## 0 17523.74 731.0134 24 17245.88 17801.60 15909.75 18472.5 17239.65 17797.79
## 1 18169.14 304.3731  7 17654.64 18683.65 17754.83 18454.0 17898.88 18370.25
##      Q75
## 0 17992.25
## 1 18403.58
##
## $comparison
## NULL
##
## $groups
##      sr groups
## 1 18169.14    a
## 0 17523.74    b
##
## attr(,"class")
## [1] "group"
```

D. Elevation

```
ggdensity(central_american_hypo$elevation,
           main = "Density plot of elevation",
           xlab = "Elavations of the genotypes country of origin")
```



```
ggplot(data = central_american_hypo, mapping = aes(x = seed_setting, y = elevation)) +  
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_36 <-glm(elevation ~ seed_setting , data = central_american_hypo, family = gaussian)  
summary(md_36)
```

```
##  
## Call:
```

```

## glm(formula = elevation ~ seed_setting, family = gaussian, data = central_american_hypo)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1364.7   -673.2    172.3    436.8   1031.3
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1952.7      140.5  13.895 2.38e-14 ***
## seed_setting1   -563.4      295.7   -1.905  0.0667 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 473967.4)
##
##      Null deviance: 15465402  on 30  degrees of freedom
## Residual deviance: 13745054  on 29  degrees of freedom
## (1 observation deleted due to missingness)
## AIC: 497.04
##
## Number of Fisher Scoring iterations: 2

```

```

lrtest(md_36)

```

```

## Likelihood ratio test
##
## Model 1: elevation ~ seed_setting
## Model 2: elevation ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    3 -245.52
## 2    2 -247.35 -1  3.6557    0.05588 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

lsmeans_ss_elevation <- LSD.test(md_36, "seed_setting", p.adj="none" )
lsmeans_ss_elevation

```

```

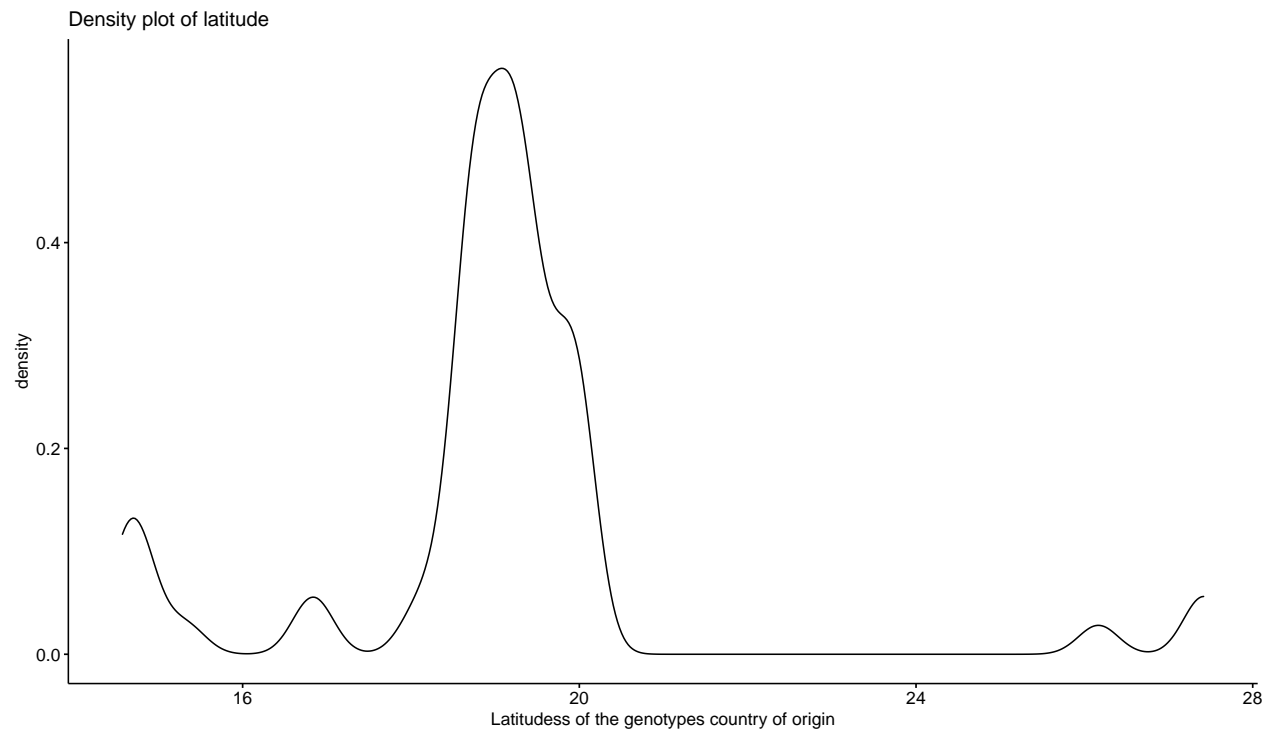
## $statistics
##      MSerror Df      Mean      CV
##  473967.4 29 1825.484 37.71345
##
## $parameters
##      test p.adjusted      name.t ntr alpha
## Fisher-LSD      none seed_setting  2  0.05
##
## $means
##      elevation      std r      LCL      UCL Min Max      Q25      Q50      Q75
## 0  1952.708 675.5464 24 1665.292 2240.124 588 2984 1281.75 2249.5 2368.75
## 1  1389.286 735.8335  7  857.095 1921.476 392 2254  910.00 1468.0 1895.50
##
## $comparison
## NULL
##
## $groups
##      elevation groups

```

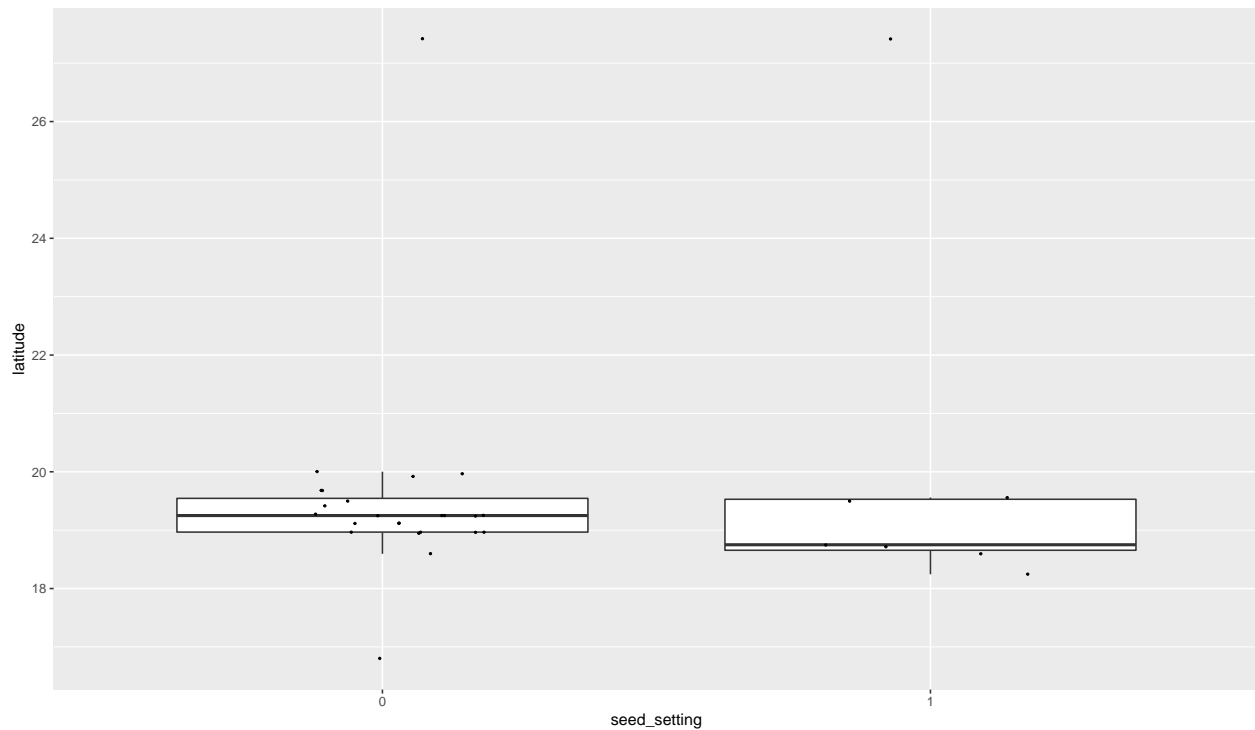
```
## 0 1952.708      a
## 1 1389.286      a
##
## attr(,"class")
## [1] "group"
```

E. Latitude

```
ggsdensity(central_american$latitude,
  main = "Density plot of latitude",
  xlab = "Latitudess of the genotypes country of origin")
```



```
ggplot(data = central_american_hypo, mapping = aes(x = seed_setting, y = latitude)) +
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_37 <-glm(latitude ~ seed_setting , data = central_american_hypo, family = gaussian)
summary(md_37)
```

```
##
## Call:
## glm(formula = latitude ~ seed_setting, family = gaussian, data = central_american_hypo)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.7273  -0.5689  -0.4106  -0.0696   7.8894
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    19.5273     0.4443  43.955  <2e-16 ***
## seed_setting1     0.5842     0.9349   0.625   0.537
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 4.736805)
##
##      Null deviance: 139.22  on 30  degrees of freedom
## Residual deviance: 137.37  on 29  degrees of freedom
## (1 observation deleted due to missingness)
## AIC: 140.12
##
## Number of Fisher Scoring iterations: 2
lrtest(md_37)
```

```
## Likelihood ratio test
##
## Model 1: latitude ~ seed_setting
```

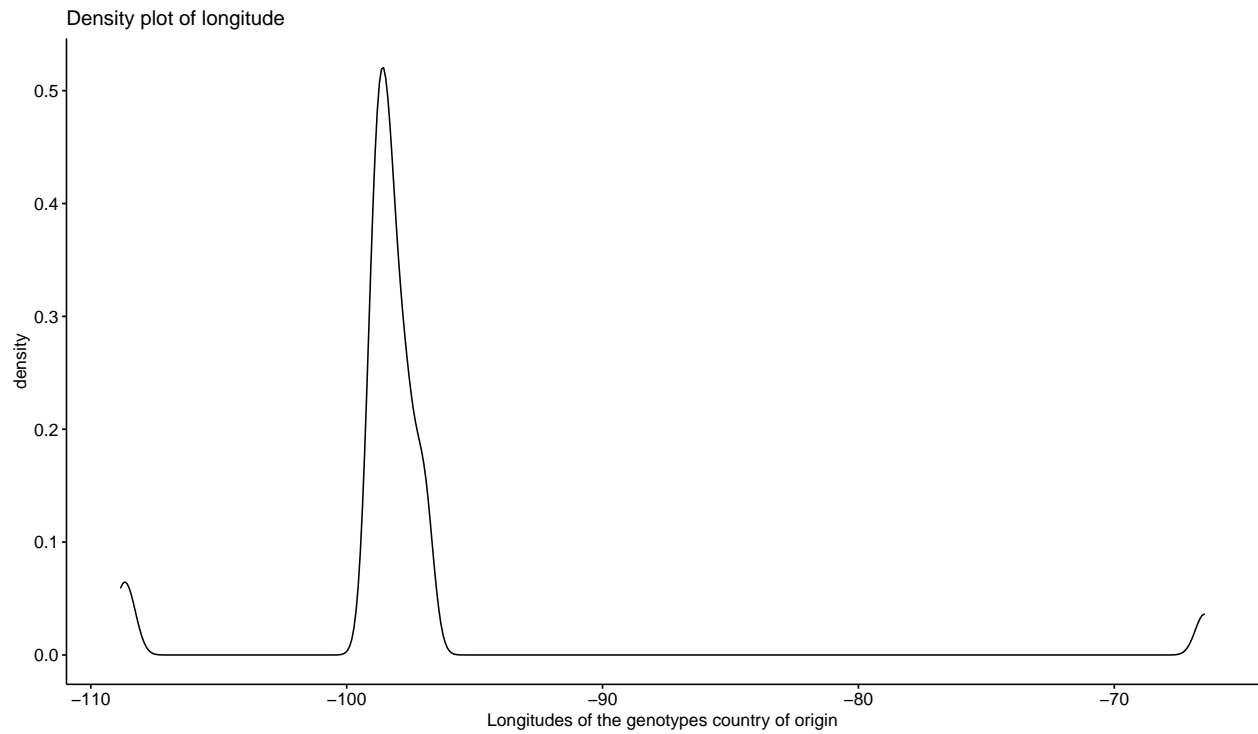
```
## Model 2: latitude ~ 1
##   #Df LogLik Df   Chisq Pr(>Chisq)
## 1    3 -67.062
## 2    2 -67.269 -1 0.4145      0.5197

lsmeans_ss_latitude <- LSD.test(md_37, "seed_setting", p.adj="none" )
lsmeans_ss_latitude

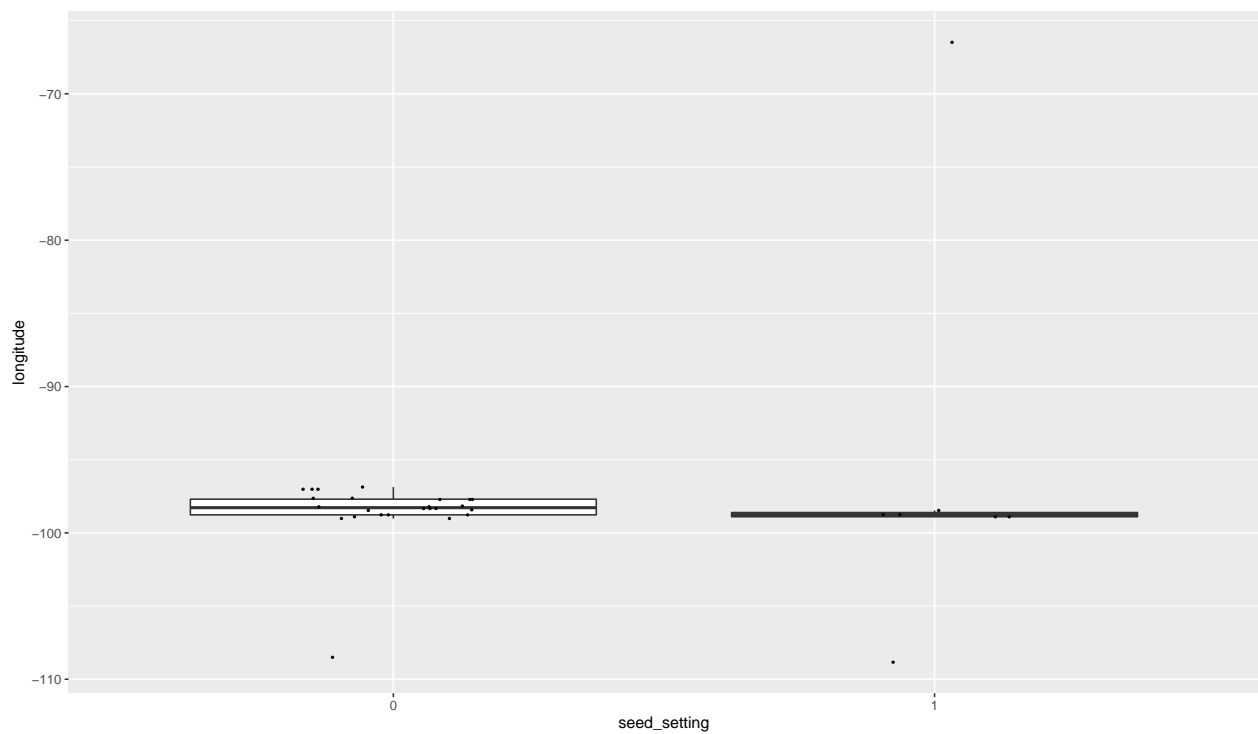
## $statistics
##   MSerror Df      Mean      CV
## 4.736805 29 19.65918 11.07075
##
## $parameters
##      test p.adjusted      name.t ntr alpha
## Fisher-LSD      none seed_setting  2 0.05
##
## $means
##   latitude      std r      LCL      UCL      Min      Max      Q25      Q50
## 0 19.52728 1.790484 24 18.61867 20.43589 16.80000 27.41667 18.96667 19.25
## 1 20.11143 3.256613  7 18.42901 21.79386 18.24419 27.41667 18.65561 18.75
##      Q75
## 0 19.54487
## 1 19.52898
##
## $comparison
## NULL
##
## $groups
##   latitude groups
## 1 20.11143      a
## 0 19.52728      a
##
## attr(,"class")
## [1] "group"
```

F. Longitude

```
ggdensity(central_american_hypo$longitude,
           main = "Density plot of longitude",
           xlab = "Longitudes of the genotypes country of origin")
```



```
ggplot(data = central_american_hypo, mapping = aes(x = seed_setting, y = longitude)) +
  geom_boxplot(outlier.size = -1)+geom_jitter(position=position_jitter(0.2), size=0.4)
```



```
md_38 <-glm(longitude ~ seed_setting , data = central_american_hypo, family = gaussian)
summary(md_38)
```

```
##
## Call:
```

```

## glm(formula = longitude ~ seed_setting, family = gaussian, data = central_american_hypo)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -13.2520   -0.4940    0.1886    0.8060   29.1024
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -98.523      1.306  -75.45  <2e-16 ***
## seed_setting1    2.941      2.748    1.07   0.293
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 40.92449)
##
##      Null deviance: 1233.7  on 30  degrees of freedom
## Residual deviance: 1186.8  on 29  degrees of freedom
## (1 observation deleted due to missingness)
## AIC: 206.97
##
## Number of Fisher Scoring iterations: 2

```

```

lrtest(md_38)

```

```

## Likelihood ratio test
##
## Model 1: longitude ~ seed_setting
## Model 2: longitude ~ 1
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1    3 -100.48
## 2    2 -101.09 -1 1.2011    0.2731

```

```

lsmeans_ss_longitude <- LSD.test(md_38, "seed_setting", p.adj="none" )
lsmeans_ss_longitude

```

```

## $statistics
##      MSerror Df      Mean      CV
## 40.92449 29 -97.85852 -6.537218
##
## $parameters
##      test p.adjusted      name.t ntr alpha
## Fisher-LSD      none seed_setting  2 0.05
##
## $means
## longitude      std r      LCL      UCL      Min      Max      Q25
## 0 -98.52271  2.225085 24 -101.1934 -95.85199 -108.5000 -96.86667 -98.76667
## 1 -95.58131 13.372465  7 -100.5265 -90.63610 -108.8333 -66.47890 -98.89611
##      Q50      Q75
## 0 -98.275 -97.69583
## 1 -98.750 -98.60734
##
## $comparison
## NULL
##
## $groups

```



```
##    longitude groups
## 1 -95.58131      a
## 0 -98.52271      a
##
## attr(,"class")
## [1] "group"
```