2

Juan Jimenez

2022-04-12

```
#########################ask yourself if the h2 suppose to be constant or change?
cc <- read.csv("DATA/ppj220030-sup-0002-tables1.csv")
table(cc$date)
```

```
##
##  1-Sep 12-Aug 14-Aug 16-Aug 20-Aug 22-Aug 23-Aug 26-Aug  3-Sep 30-Aug  5-Sep
##    878    878    878    878    878    878    878    877    878    878    878
##  6-Jul
##    878
```

```
### adding replication information
cc$Rep <- "Rep2"
cc[cc$Row< 3000,] $Rep <- "Rep1"

###Checking our data
head(cc)
```

```
##    Row   Row.Numbers Genotype   Pedigree Treatment   date Canopy_Coverage  Rep
## 1 1001 1001 and 1002     A641 Ames 19311  Nitrogen  6-Jul       29.747379 Rep1
## 2 1001 1001 and 1002     A641 Ames 19311  Nitrogen 23-Aug       15.537977 Rep1
## 3 1001 1001 and 1002     A641 Ames 19311  Nitrogen  1-Sep       15.189726 Rep1
## 4 1001 1001 and 1002     A641 Ames 19311  Nitrogen 22-Aug       19.037283 Rep1
## 5 1001 1001 and 1002     A641 Ames 19311  Nitrogen 12-Aug        9.793073 Rep1
## 6 1001 1001 and 1002     A641 Ames 19311  Nitrogen 26-Aug       21.378869 Rep1
```

```
tail(cc)
```

```
##        Row   Row.Numbers Genotype  Pedigree Treatment   date Canopy_Coverage
## 10530 4583 4583 and 4584    NC262 PI 531085  Nitrogen 23-Aug        17.90676
## 10531 4583 4583 and 4584    NC262 PI 531085  Nitrogen 22-Aug        18.60264
## 10532 4583 4583 and 4584    NC262 PI 531085  Nitrogen 30-Aug        15.17244
## 10533 4583 4583 and 4584    NC262 PI 531085  Nitrogen  5-Sep         9.91415
## 10534 4583 4583 and 4584    NC262 PI 531085  Nitrogen  1-Sep        15.37183
## 10535 4583 4583 and 4584    NC262 PI 531085  Nitrogen 12-Aug        20.12580
##        Rep
## 10530 Rep2
## 10531 Rep2
## 10532 Rep2
## 10533 Rep2
## 10534 Rep2
## 10535 Rep2
```

```
dim(cc)
```

```
## [1] 10535     8
```

```
table(cc$Treatment)
```

```
##
##    Nitrogen No Nitrogen
##        5279        5256
```

```
table(cc$Rep)
```

```
##
## Rep1 Rep2
## 5364 5171
```

```
table(cc$date)
```

```
##
##  1-Sep 12-Aug 14-Aug 16-Aug 20-Aug 22-Aug 23-Aug 26-Aug  3-Sep 30-Aug  5-Sep
##    878    878    878    878    878    878    878    877    878    878    878
##  6-Jul
##    878
```

```
#### sub-setting for date 6 Jul
d1 <- subset(cc, date %in% "6-Jul")
head(d1)
```

```
##     Row   Row.Numbers Genotype   Pedigree Treatment  date Canopy_Coverage  Rep
## 1  1001 1001 and 1002     A641 Ames 19311  Nitrogen 6-Jul       29.747379 Rep1
## 20 1003 1003 and 1004      B76  PI 550483  Nitrogen 6-Jul       47.451452 Rep1
## 33 1005 1005 and 1006     A654  PI 587141  Nitrogen 6-Jul       43.296481 Rep1
## 45 1007 1007 and 1008     A239 Ames 23405  Nitrogen 6-Jul       36.716503 Rep1
## 60 1009 1009 and 1010     A682  PI 587143  Nitrogen 6-Jul       37.797254 Rep1
## 70 1011 1011 and 1012     C123 Ames 19313  Nitrogen 6-Jul        8.849936 Rep1
```

```
tail(d1)
```

```
##        Row   Row.Numbers Genotype   Pedigree Treatment  date Canopy_Coverage
## 10472 4569 4569 and 4570    CO255 Ames 27107  Nitrogen 6-Jul        34.59631
## 10482 4573 4573 and 4574     N192  PI 550566  Nitrogen 6-Jul        37.25802
## 10489 4575 4575 and 4576    NC312 Ames 27152  Nitrogen 6-Jul        29.30559
## 10507 4577 4577 and 4578    NC292 Ames 27142  Nitrogen 6-Jul        27.58182
## 10516 4579 4579 and 4580     SD44  PI 524969  Nitrogen 6-Jul        36.25121
## 10526 4583 4583 and 4584    NC262  PI 531085  Nitrogen 6-Jul        12.03667
##        Rep
## 10472 Rep2
## 10482 Rep2
## 10489 Rep2
## 10507 Rep2
## 10516 Rep2
## 10526 Rep2
```

```
### Defining the model to get multiple linear regression for date 6 Jul
fit1 <- lm(Canopy_Coverage ~ Genotype + Treatment + Genotype:Treatment + Rep,
          data=d1)

#generating ANOVA table
summary(aov(fit1))
```

```
##                      Df Sum Sq Mean Sq F value  Pr(>F)
## Genotype            232  63718     275   8.667 < 2e-16 ***
## Treatment             1   7892    7892 249.031 < 2e-16 ***
## Rep                   1   1604    1604  50.615 4.9e-12 ***
## Genotype:Treatment  224   6987      31   0.984   0.549
## Residuals           419  13278      32
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#As the p-value is less than the significance level 0.05, we can conclude that there are significant di

##Calculating Vprogeny
out1 <- summary(aov(fit1))[[1]]
vprogeny1 <- (out1[1,3] - out1[4,3])/(2*2)


#Calculating H2
h2_1 = vprogeny1/((vprogeny1 + out1[4,3]/(2*2)))

##H2 is very high 0.89. This means that canopy coverage is highly influenced by genetic effects more th

####For in the date 12 Aug
d2 <- subset(cc, date %in% "12-Aug")
head(d2)
```

```
##     Row   Row.Numbers Genotype   Pedigree Treatment   date Canopy_Coverage  Rep
## 5  1001 1001 and 1002     A641 Ames 19311  Nitrogen 12-Aug        9.793073 Rep1
## 24 1003 1003 and 1004      B76  PI 550483  Nitrogen 12-Aug       72.860293 Rep1
## 35 1005 1005 and 1006     A654  PI 587141  Nitrogen 12-Aug       49.738835 Rep1
## 39 1007 1007 and 1008     A239 Ames 23405  Nitrogen 12-Aug       70.890916 Rep1
## 50 1009 1009 and 1010     A682  PI 587143  Nitrogen 12-Aug       71.139594 Rep1
## 66 1011 1011 and 1012     C123 Ames 19313  Nitrogen 12-Aug        5.203955 Rep1
```

```
tail(d2)
```

```
##        Row   Row.Numbers Genotype   Pedigree Treatment   date Canopy_Coverage
## 10468 4569 4569 and 4570    CO255 Ames 27107  Nitrogen 12-Aug        31.74641
## 10484 4573 4573 and 4574     N192  PI 550566  Nitrogen 12-Aug        35.34276
## 10499 4575 4575 and 4576    NC312 Ames 27152  Nitrogen 12-Aug        69.51793
## 10502 4577 4577 and 4578    NC292 Ames 27142  Nitrogen 12-Aug        25.29349
## 10519 4579 4579 and 4580     SD44  PI 524969  Nitrogen 12-Aug        70.22684
## 10535 4583 4583 and 4584    NC262  PI 531085  Nitrogen 12-Aug        20.12580
##        Rep
## 10468 Rep2
## 10484 Rep2
```

```
## 10499 Rep2
## 10502 Rep2
## 10519 Rep2
## 10535 Rep2
```

```
### Defining the model to get multiple linear regression for date 6 Jul
fit2 <- lm(Canopy_Coverage ~ Genotype + Treatment + Genotype:Treatment + Rep,
          data=d2)

#generating ANOVA table
summary(aov(fit2))
```

```
##                      Df Sum Sq Mean Sq F value   Pr(>F)
## Genotype            232 336570  1450.7  16.618  < 2e-16 ***
## Treatment             1   1581  1580.8  18.108 2.58e-05 ***
## Rep                   1   2096  2096.2  24.012 1.37e-06 ***
## Genotype:Treatment  224  20498    91.5   1.048    0.339
## Residuals           419  36578    87.3
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#As the p-value is less than the significance level 0.05, we can conclude that there are significant di

##Calculating Vprogeny
out2 <- summary(aov(fit2))[[1]]
vprogeny2 <- (out2[1,3] - out2[4,3])/(2*2)


#Calculating H2
h2_2 = vprogeny2/((vprogeny2 + out2[4,3]/(2*2)))

##H2 is very high 0.94. This means that canopy coverage is highly influenced by genetic effects more th

####For in the date 14 Aug
d3 <- subset(cc, date %in% "14-Aug")
head(d3)
```

```
##     Row   Row.Numbers Genotype   Pedigree Treatment   date Canopy_Coverage  Rep
## 8  1001 1001 and 1002     A641 Ames 19311  Nitrogen 14-Aug        8.330186 Rep1
## 18 1003 1003 and 1004      B76  PI 550483  Nitrogen 14-Aug       68.743223 Rep1
## 30 1005 1005 and 1006     A654  PI 587141  Nitrogen 14-Aug       48.227840 Rep1
## 40 1007 1007 and 1008     A239 Ames 23405  Nitrogen 14-Aug       69.157072 Rep1
## 52 1009 1009 and 1010     A682  PI 587143  Nitrogen 14-Aug       60.686364 Rep1
## 71 1011 1011 and 1012     C123 Ames 19313  Nitrogen 14-Aug        4.383795 Rep1
```

```
tail(d3)
```

```
##        Row   Row.Numbers Genotype   Pedigree Treatment   date Canopy_Coverage
## 10474 4569 4569 and 4570    CO255 Ames 27107  Nitrogen 14-Aug        20.99052
## 10487 4573 4573 and 4574     N192  PI 550566  Nitrogen 14-Aug        22.83648
## 10490 4575 4575 and 4576    NC312 Ames 27152  Nitrogen 14-Aug        55.09773
## 10501 4577 4577 and 4578    NC292 Ames 27142  Nitrogen 14-Aug        18.50045
```

```
## 10517 4579 4579 and 4580     SD44  PI 524969  Nitrogen 14-Aug        62.04407
## 10528 4583 4583 and 4584    NC262  PI 531085  Nitrogen 14-Aug        13.25547
##        Rep
## 10474 Rep2
## 10487 Rep2
## 10490 Rep2
## 10501 Rep2
## 10517 Rep2
## 10528 Rep2
```

```
### Defining the model to get multiple linear regression for date 6 Jul
fit3 <- lm(Canopy_Coverage ~ Genotype + Treatment + Genotype:Treatment + Rep,
          data=d3)

#generating ANOVA table
summary(aov(fit3))
```

```
##                      Df Sum Sq Mean Sq F value   Pr(>F)
## Genotype            232 378248    1630  15.958  < 2e-16 ***
## Treatment             1      0       0   0.001   0.9766
## Rep                   1   4371    4371  42.779 1.79e-10 ***
## Genotype:Treatment  224  26921     120   1.176   0.0794 .
## Residuals           419  42807     102
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#As the p-value is less than the significance level 0.05, we can conclude that there are significant di

##Calculating Vprogeny
out3 <- summary(aov(fit3))[[1]]
vprogeny3 <- (out3[1,3] - out3[4,3])/(2*2)


#Calculating H2
h2_3 = vprogeny3/((vprogeny3 + out3[4,3]/(2*2)))

##H2 is very high 0.93 This means that canopy coverage is highly influenced by genetic effects more tha


####For in the date 16 Aug
d4 <- subset(cc, date %in% "16-Aug")
head(d4)
```

```
##     Row   Row.Numbers Genotype   Pedigree Treatment   date Canopy_Coverage  Rep
## 12 1001 1001 and 1002     A641 Ames 19311  Nitrogen 16-Aug       11.573201 Rep1
## 15 1003 1003 and 1004      B76  PI 550483  Nitrogen 16-Aug       71.856793 Rep1
## 28 1005 1005 and 1006     A654  PI 587141  Nitrogen 16-Aug       44.140153 Rep1
## 43 1007 1007 and 1008     A239 Ames 23405  Nitrogen 16-Aug       64.226668 Rep1
## 58 1009 1009 and 1010     A682  PI 587143  Nitrogen 16-Aug       59.013240 Rep1
## 68 1011 1011 and 1012     C123 Ames 19313  Nitrogen 16-Aug        4.432205 Rep1
```

```
tail(d4)
```

```
##        Row   Row.Numbers Genotype   Pedigree Treatment   date Canopy_Coverage
## 10467 4569 4569 and 4570    CO255 Ames 27107  Nitrogen 16-Aug        17.18911
## 10485 4573 4573 and 4574     N192  PI 550566  Nitrogen 16-Aug        21.17202
## 10488 4575 4575 and 4576    NC312 Ames 27152  Nitrogen 16-Aug        57.69139
## 10508 4577 4577 and 4578    NC292 Ames 27142  Nitrogen 16-Aug        19.02845
## 10514 4579 4579 and 4580     SD44  PI 524969  Nitrogen 16-Aug        57.30935
## 10529 4583 4583 and 4584    NC262  PI 531085  Nitrogen 16-Aug        11.09715
##        Rep
## 10467 Rep2
## 10485 Rep2
## 10488 Rep2
## 10508 Rep2
## 10514 Rep2
## 10529 Rep2
```

```
### Defining the model to get multiple linear regression for date 6 Jul
fit4 <- lm(Canopy_Coverage ~ Genotype + Treatment + Genotype:Treatment + Rep,
          data=d4)

#generating ANOVA table
summary(aov(fit4))
```

```
##                     Df Sum Sq Mean Sq F value  Pr(>F)
## Genotype           232 345614    1490  17.863 < 2e-16 ***
## Treatment            1    755     755   9.053 0.00278 **
## Rep                  1   8709    8709 104.430 < 2e-16 ***
## Genotype:Treatment 224  25079     112   1.342 0.00524 **
## Residuals          419  34944      83
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#As the p-value is less than the significance level 0.05, we can conclude that there are significant di

##Calculating Vprogeny
out4 <- summary(aov(fit4))[[1]]
vprogeny4 <- (out4[1,3] - out4[4,3])/(2*2)

#Calculating H2
h2_4 = vprogeny4/((vprogeny4 + out4[4,3]/(2*2)))

##H2 is very high 0.92 This means that canopy coverage is highly influenced by genetic effects more tha

####For in the date 20 Aug
d5 <- subset(cc, date %in% "20-Aug")
head(d5)
```

```
##     Row   Row.Numbers Genotype   Pedigree Treatment   date Canopy_Coverage  Rep
## 9  1001 1001 and 1002     A641 Ames 19311  Nitrogen 20-Aug       10.179700 Rep1
## 14 1003 1003 and 1004      B76  PI 550483  Nitrogen 20-Aug       74.095592 Rep1
## 27 1005 1005 and 1006     A654  PI 587141  Nitrogen 20-Aug       54.376656 Rep1
```

```
## 41 1007 1007 and 1008     A239 Ames 23405  Nitrogen 20-Aug       73.009362 Rep1
## 51 1009 1009 and 1010     A682  PI 587143  Nitrogen 20-Aug       68.290128 Rep1
## 65 1011 1011 and 1012     C123 Ames 19313  Nitrogen 20-Aug        6.406647 Rep1
```

```
tail(d5)
```

```
##        Row   Row.Numbers Genotype   Pedigree Treatment   date Canopy_Coverage
## 10464 4569 4569 and 4570    CO255 Ames 27107  Nitrogen 20-Aug        15.60959
## 10478 4573 4573 and 4574     N192  PI 550566  Nitrogen 20-Aug        29.47295
## 10492 4575 4575 and 4576    NC312 Ames 27152  Nitrogen 20-Aug        64.22311
## 10503 4577 4577 and 4578    NC292 Ames 27142  Nitrogen 20-Aug        23.73602
## 10513 4579 4579 and 4580     SD44  PI 524969  Nitrogen 20-Aug        64.44475
## 10524 4583 4583 and 4584    NC262  PI 531085  Nitrogen 20-Aug        16.01849
##        Rep
## 10464 Rep2
## 10478 Rep2
## 10492 Rep2
## 10503 Rep2
## 10513 Rep2
## 10524 Rep2
```

```
### Defining the model to get multiple linear regression for date 6 Jul
fit5 <- lm(Canopy_Coverage ~ Genotype + Treatment + Genotype:Treatment + Rep,
          data=d5)

#generating ANOVA table
summary(aov(fit5))
```

```
##                      Df Sum Sq Mean Sq F value   Pr(>F)
## Genotype            232 356944  1538.6  17.443  < 2e-16 ***
## Treatment             1   1180  1180.1  13.380 0.000287 ***
## Rep                   1    777   776.9   8.808 0.003171 **
## Genotype:Treatment  224  26990   120.5   1.366 0.003343 **
## Residuals           419  36957    88.2
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#As the p-value is less than the significance level 0.05, we can conclude that there are significant di

##Calculating Vprogeny
out5 <- summary(aov(fit5))[[1]]
vprogeny5 <- (out5[1,3] - out5[4,3])/(2*2)

#Calculating H2
h2_5 = vprogeny5/((vprogeny5 + out5[4,3]/(2*2)))

##H2 is very high 0.92 This means that canopy coverage is highly influenced by genetic effects more tha


####For in the date 22 Aug
d6 <- subset(cc, date %in% "22-Aug")
head(d6)
```

```
##     Row   Row.Numbers Genotype   Pedigree Treatment   date Canopy_Coverage  Rep
## 4  1001 1001 and 1002     A641 Ames 19311  Nitrogen 22-Aug        19.03728 Rep1
## 16 1003 1003 and 1004      B76  PI 550483  Nitrogen 22-Aug        75.43454 Rep1
## 29 1005 1005 and 1006     A654  PI 587141  Nitrogen 22-Aug        53.39257 Rep1
## 38 1007 1007 and 1008     A239 Ames 23405  Nitrogen 22-Aug        69.77542 Rep1
## 56 1009 1009 and 1010     A682  PI 587143  Nitrogen 22-Aug        62.87168 Rep1
## 67 1011 1011 and 1012     C123 Ames 19313  Nitrogen 22-Aug        11.63524 Rep1
```

```
tail(d6)
```

```
##        Row   Row.Numbers Genotype   Pedigree Treatment   date Canopy_Coverage
## 10470 4569 4569 and 4570    CO255 Ames 27107  Nitrogen 22-Aug        18.43629
## 10479 4573 4573 and 4574     N192  PI 550566  Nitrogen 22-Aug        32.23088
## 10493 4575 4575 and 4576    NC312 Ames 27152  Nitrogen 22-Aug        68.20305
## 10505 4577 4577 and 4578    NC292 Ames 27142  Nitrogen 22-Aug        25.38502
## 10521 4579 4579 and 4580     SD44  PI 524969  Nitrogen 22-Aug        65.04087
## 10531 4583 4583 and 4584    NC262  PI 531085  Nitrogen 22-Aug        18.60264
##        Rep
## 10470 Rep2
## 10479 Rep2
## 10493 Rep2
## 10505 Rep2
## 10521 Rep2
## 10531 Rep2
```

```
### Defining the model to get multiple linear regression for date 6 Jul
fit6 <- lm(Canopy_Coverage ~ Genotype + Treatment + Genotype:Treatment + Rep,
          data=d6)

#generating ANOVA table
summary(aov(fit6))
```

```
##                      Df Sum Sq Mean Sq F value   Pr(>F)
## Genotype            232 319387  1376.7  18.021  < 2e-16 ***
## Treatment             1   2643  2643.2  34.599 8.29e-09 ***
## Rep                   1    676   676.0   8.848   0.0031 **
## Genotype:Treatment  224  21771    97.2   1.272   0.0183 *
## Residuals           419  32009    76.4
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#As the p-value is less than the significance level 0.05, we can conclude that there are significant di

##Calculating Vprogeny
out6 <- summary(aov(fit6))[[1]]
vprogeny6 <- (out6[1,3] - out6[4,3])/(2*2)

#Calculating H2
h2_6 = vprogeny6/((vprogeny6 + out6[4,3]/(2*2)))

##H2 is very high 0.93 This means that canopy coverage is highly influenced by genetic effects more tha
```

```
####For in the date 23 Aug
d7 <- subset(cc, date %in% "23-Aug")
head(d7)
```

```
##     Row   Row.Numbers Genotype   Pedigree Treatment   date Canopy_Coverage  Rep
## 2  1001 1001 and 1002     A641 Ames 19311  Nitrogen 23-Aug        15.53798 Rep1
## 13 1003 1003 and 1004      B76  PI 550483  Nitrogen 23-Aug        75.10952 Rep1
## 32 1005 1005 and 1006     A654  PI 587141  Nitrogen 23-Aug        52.16748 Rep1
## 48 1007 1007 and 1008     A239 Ames 23405  Nitrogen 23-Aug        71.81781 Rep1
## 55 1009 1009 and 1010     A682  PI 587143  Nitrogen 23-Aug        74.07453 Rep1
## 69 1011 1011 and 1012     C123 Ames 19313  Nitrogen 23-Aug        10.83411 Rep1
```

```
tail(d7)
```

```
##        Row   Row.Numbers Genotype   Pedigree Treatment   date Canopy_Coverage
## 10473 4569 4569 and 4570    CO255 Ames 27107  Nitrogen 23-Aug        20.26747
## 10476 4573 4573 and 4574     N192  PI 550566  Nitrogen 23-Aug        31.52568
## 10498 4575 4575 and 4576    NC312 Ames 27152  Nitrogen 23-Aug        65.63632
## 10511 4577 4577 and 4578    NC292 Ames 27142  Nitrogen 23-Aug        26.28921
## 10518 4579 4579 and 4580     SD44  PI 524969  Nitrogen 23-Aug        67.77127
## 10530 4583 4583 and 4584    NC262  PI 531085  Nitrogen 23-Aug        17.90676
##        Rep
## 10473 Rep2
## 10476 Rep2
## 10498 Rep2
## 10511 Rep2
## 10518 Rep2
## 10530 Rep2
```

```
### Defining the model to get multiple linear regression for date 6 Jul
fit7 <- lm(Canopy_Coverage ~ Genotype + Treatment + Genotype:Treatment + Rep,
          data=d7)

#generating ANOVA table
summary(aov(fit7))
```

```
##                      Df Sum Sq Mean Sq F value   Pr(>F)
## Genotype            232 329716    1421  16.308  < 2e-16 ***
## Treatment             1   3164    3164  36.311 3.69e-09 ***
## Rep                   1   3270    3270  37.522 2.08e-09 ***
## Genotype:Treatment  224  23591     105   1.209   0.0502 .
## Residuals           419  36514      87
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#As the p-value is less than the significance level 0.05, we can conclude that there are significant di

##Calculating Vprogeny
out7 <- summary(aov(fit7))[[1]]
vprogeny7 <- (out7[1,3] - out7[4,3])/(2*2)
```

```r
#Calculating H2
h2_7 = vprogeny7/((vprogeny7 + out7[4,3]/(2*2)))

##H2 is very high 0.93 This means that canopy coverage is highly influenced by genetic effects more tha

####For in the date 26 Aug
d8 <- subset(cc, date %in% "26-Aug")
head(d8)
```

```
##     Row   Row.Numbers Genotype   Pedigree Treatment   date Canopy_Coverage  Rep
## 6  1001 1001 and 1002     A641 Ames 19311  Nitrogen 26-Aug        21.37887 Rep1
## 23 1003 1003 and 1004      B76  PI 550483  Nitrogen 26-Aug        72.06732 Rep1
## 25 1005 1005 and 1006     A654  PI 587141  Nitrogen 26-Aug        55.76904 Rep1
## 37 1007 1007 and 1008     A239 Ames 23405  Nitrogen 26-Aug        74.37770 Rep1
## 57 1009 1009 and 1010     A682  PI 587143  Nitrogen 26-Aug        70.56410 Rep1
## 64 1011 1011 and 1012     C123 Ames 19313  Nitrogen 26-Aug        16.30953 Rep1
```

```r
tail(d8)
```

```
##        Row   Row.Numbers Genotype   Pedigree Treatment   date Canopy_Coverage
## 10471 4569 4569 and 4570    CO255 Ames 27107  Nitrogen 26-Aug        15.53898
## 10483 4573 4573 and 4574     N192  PI 550566  Nitrogen 26-Aug        36.37043
## 10491 4575 4575 and 4576    NC312 Ames 27152  Nitrogen 26-Aug        76.05566
## 10506 4577 4577 and 4578    NC292 Ames 27142  Nitrogen 26-Aug        32.15958
## 10520 4579 4579 and 4580     SD44  PI 524969  Nitrogen 26-Aug        73.31089
## 10527 4583 4583 and 4584    NC262  PI 531085  Nitrogen 26-Aug        21.15264
##        Rep
## 10471 Rep2
## 10483 Rep2
## 10491 Rep2
## 10506 Rep2
## 10520 Rep2
## 10527 Rep2
```

```r
### Defining the model to get multiple linear regression for date 6 Jul
fit8 <- lm(Canopy_Coverage ~ Genotype + Treatment + Genotype:Treatment + Rep,
          data=d8)

#generating ANOVA table
summary(aov(fit8))
```

```
##                      Df Sum Sq Mean Sq F value   Pr(>F)
## Genotype            232 314911    1357  15.905  < 2e-16 ***
## Treatment             1   4076    4076  47.756 1.81e-11 ***
## Rep                   1   1273    1273  14.919  0.00013 ***
## Genotype:Treatment  224  23682     106   1.239  0.03166 *
## Residuals           418  35674      85
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
#As the p-value is less than the significance level 0.05, we can conclude that there are significant di

##Calculating Vprogeny
out8 <- summary(aov(fit8))[[1]]
vprogeny8 <- (out8[1,3] - out8[4,3])/(2*2)

#Calculating H2
h2_8 = vprogeny8/((vprogeny8 + out8[4,3]/(2*2)))

##H2 is very high 0.92 This means that canopy coverage is highly influenced by genetic effects more tha

####For in the date 30 Aug
d9 <- subset(cc, date %in% "30-Aug")
head(d9)
```

```
##     Row   Row.Numbers Genotype   Pedigree Treatment   date Canopy_Coverage  Rep
## 10 1001 1001 and 1002     A641 Ames 19311  Nitrogen 30-Aug        9.945542 Rep1
## 22 1003 1003 and 1004      B76  PI 550483  Nitrogen 30-Aug       75.113953 Rep1
## 31 1005 1005 and 1006     A654  PI 587141  Nitrogen 30-Aug       50.838188 Rep1
## 47 1007 1007 and 1008     A239 Ames 23405  Nitrogen 30-Aug       79.089834 Rep1
## 54 1009 1009 and 1010     A682  PI 587143  Nitrogen 30-Aug       90.333319 Rep1
## 62 1011 1011 and 1012     C123 Ames 19313  Nitrogen 30-Aug       10.285431 Rep1
```

```r
tail(d9)
```

```
##        Row   Row.Numbers Genotype   Pedigree Treatment   date Canopy_Coverage
## 10469 4569 4569 and 4570    CO255 Ames 27107  Nitrogen 30-Aug        5.547609
## 10477 4573 4573 and 4574     N192  PI 550566  Nitrogen 30-Aug       26.156860
## 10497 4575 4575 and 4576    NC312 Ames 27152  Nitrogen 30-Aug       63.367000
## 10500 4577 4577 and 4578    NC292 Ames 27142  Nitrogen 30-Aug       23.513413
## 10512 4579 4579 and 4580     SD44  PI 524969  Nitrogen 30-Aug       62.546048
## 10532 4583 4583 and 4584    NC262  PI 531085  Nitrogen 30-Aug       15.172438
##        Rep
## 10469 Rep2
## 10477 Rep2
## 10497 Rep2
## 10500 Rep2
## 10512 Rep2
## 10532 Rep2
```

```r
### Defining the model to get multiple linear regression for date 6 Jul
fit9 <- lm(Canopy_Coverage ~ Genotype + Treatment + Genotype:Treatment + Rep,
          data=d9)

#generating ANOVA table
summary(aov(fit9))
```

```
##                      Df Sum Sq Mean Sq F value  Pr(>F)
## Genotype            232 364874    1573  14.113 < 2e-16 ***
## Treatment             1    937     937   8.404 0.00394 **
## Rep                   1   9057    9057  81.272 < 2e-16 ***
## Genotype:Treatment 224  31197     139   1.250 0.02650 *
```

```
## Residuals          419  46692     111
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#As the p-value is less than the significance level 0.05, we can conclude that there are significant di

##Calculating Vprogeny
out9 <- summary(aov(fit9))[[1]]
vprogeny9 <- (out9[1,3] - out9[4,3])/(2*2)

#Calculating H2
h2_9 = vprogeny9/((vprogeny9 + out9[4,3]/(2*2)))

##H2 is very high 0.91 This means that canopy coverage is highly influenced by genetic effects more tha

####For in the date 1 Sep
d10 <- subset(cc, date %in% "1-Sep")
head(d10)
```

```
##     Row   Row.Numbers Genotype   Pedigree Treatment  date Canopy_Coverage  Rep
## 3  1001 1001 and 1002     A641 Ames 19311  Nitrogen 1-Sep       15.189726 Rep1
## 17 1003 1003 and 1004      B76  PI 550483  Nitrogen 1-Sep       74.382934 Rep1
## 26 1005 1005 and 1006     A654  PI 587141  Nitrogen 1-Sep       48.696939 Rep1
## 44 1007 1007 and 1008     A239 Ames 23405  Nitrogen 1-Sep       71.570800 Rep1
## 59 1009 1009 and 1010     A682  PI 587143  Nitrogen 1-Sep       76.387863 Rep1
## 63 1011 1011 and 1012     C123 Ames 19313  Nitrogen 1-Sep        7.711603 Rep1
```

```
tail(d10)
```

```
##        Row   Row.Numbers Genotype   Pedigree Treatment  date Canopy_Coverage
## 10466 4569 4569 and 4570    CO255 Ames 27107  Nitrogen 1-Sep        6.362997
## 10480 4573 4573 and 4574     N192  PI 550566  Nitrogen 1-Sep       25.053398
## 10495 4575 4575 and 4576    NC312 Ames 27152  Nitrogen 1-Sep       67.485713
## 10510 4577 4577 and 4578    NC292 Ames 27142  Nitrogen 1-Sep       21.563958
## 10523 4579 4579 and 4580     SD44  PI 524969  Nitrogen 1-Sep       70.742389
## 10534 4583 4583 and 4584    NC262  PI 531085  Nitrogen 1-Sep       15.371834
##        Rep
## 10466 Rep2
## 10480 Rep2
## 10495 Rep2
## 10510 Rep2
## 10523 Rep2
## 10534 Rep2
```

```
### Defining the model to get multiple linear regression for date 6 Jul
fit10 <- lm(Canopy_Coverage ~ Genotype + Treatment + Genotype:Treatment + Rep,
          data=d10)

#generating ANOVA table
summary(aov(fit10))
```

```
##                     Df Sum Sq Mean Sq F value   Pr(>F)
```

```
## Genotype           232 441587    1903  14.332  < 2e-16 ***
## Treatment            1   2971    2971  22.372 3.07e-06 ***
## Rep                  1   3309    3309  24.919 8.79e-07 ***
## Genotype:Treatment 224  32509     145   1.093     0.22
## Residuals          419  55645     133
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#As the p-value is less than the significance level 0.05, we can conclude that there are significant di

##Calculating Vprogeny
out10 <- summary(aov(fit10))[[1]]
vprogeny10 <- (out10[1,3] - out10[4,3])/(2*2)

#Calculating H2
h2_10 = vprogeny10/((vprogeny10 + out10[4,3]/(2*2)))

##H2 is very high 0.93 This means that canopy coverage is highly influenced by genetic effects more tha

####For in the date 3 Sep
d11 <- subset(cc, date %in% "3-Sep")
head(d11)
```

```
##     Row   Row.Numbers Genotype   Pedigree Treatment  date Canopy_Coverage  Rep
## 7  1001 1001 and 1002     A641 Ames 19311  Nitrogen 3-Sep        13.53673 Rep1
## 19 1003 1003 and 1004      B76  PI 550483  Nitrogen 3-Sep        61.73542 Rep1
## 36 1005 1005 and 1006     A654  PI 587141  Nitrogen 3-Sep        41.31357 Rep1
## 42 1007 1007 and 1008     A239 Ames 23405  Nitrogen 3-Sep        51.48036 Rep1
## 53 1009 1009 and 1010     A682  PI 587143  Nitrogen 3-Sep        49.44327 Rep1
## 72 1011 1011 and 1012     C123 Ames 19313  Nitrogen 3-Sep        10.07274 Rep1
```

```
tail(d11)
```

```
##        Row   Row.Numbers Genotype   Pedigree Treatment  date Canopy_Coverage
## 10465 4569 4569 and 4570    CO255 Ames 27107  Nitrogen 3-Sep        2.541246
## 10481 4573 4573 and 4574     N192  PI 550566  Nitrogen 3-Sep       20.475784
## 10494 4575 4575 and 4576    NC312 Ames 27152  Nitrogen 3-Sep       50.602426
## 10504 4577 4577 and 4578    NC292 Ames 27142  Nitrogen 3-Sep       21.216328
## 10522 4579 4579 and 4580     SD44  PI 524969  Nitrogen 3-Sep       55.747890
## 10525 4583 4583 and 4584    NC262  PI 531085  Nitrogen 3-Sep       14.445922
##        Rep
## 10465 Rep2
## 10481 Rep2
## 10494 Rep2
## 10504 Rep2
## 10522 Rep2
## 10525 Rep2
```

```
### Defining the model to get multiple linear regression for date 6 Jul
fit11 <- lm(Canopy_Coverage ~ Genotype + Treatment + Genotype:Treatment + Rep,
          data=d11)

#generating ANOVA table
summary(aov(fit11))
```

```
##                      Df Sum Sq Mean Sq F value  Pr(>F)
## Genotype            232 260177    1121  15.232 < 2e-16 ***
## Treatment             1    801     801  10.882 0.00105 **
## Rep                   1   6455    6455  87.678 < 2e-16 ***
## Genotype:Treatment  224  20558      92   1.247 0.02790 *
## Residuals           419  30848      74
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#As the p-value is less than the significance level 0.05, we can conclude that there are significant di

##Calculating Vprogeny
out11 <- summary(aov(fit11))[[1]]
vprogeny11 <- (out11[1,3] - out11[4,3])/(2*2)

#Calculating H2
h2_11 = vprogeny11/((vprogeny11 + out11[4,3]/(2*2)))

##H2 is very high 0.92 This means that canopy coverage is highly influenced by genetic effects more tha

####For in the date 5 Sep
d12 <- subset(cc, date %in% "5-Sep")
head(d12)
```

```
##     Row   Row.Numbers Genotype   Pedigree Treatment  date Canopy_Coverage  Rep
## 11 1001 1001 and 1002     A641 Ames 19311  Nitrogen 5-Sep        8.183619 Rep1
## 21 1003 1003 and 1004      B76  PI 550483  Nitrogen 5-Sep       67.182663 Rep1
## 34 1005 1005 and 1006     A654  PI 587141  Nitrogen 5-Sep       37.853154 Rep1
## 46 1007 1007 and 1008     A239 Ames 23405  Nitrogen 5-Sep       61.086586 Rep1
## 49 1009 1009 and 1010     A682  PI 587143  Nitrogen 5-Sep       61.464023 Rep1
## 61 1011 1011 and 1012     C123 Ames 19313  Nitrogen 5-Sep       10.114035 Rep1
```

```
tail(d12)
```

```
##        Row   Row.Numbers Genotype   Pedigree Treatment  date Canopy_Coverage
## 10475 4569 4569 and 4570    CO255 Ames 27107  Nitrogen 5-Sep        3.814582
## 10486 4573 4573 and 4574     N192  PI 550566  Nitrogen 5-Sep       15.824509
## 10496 4575 4575 and 4576    NC312 Ames 27152  Nitrogen 5-Sep       50.597114
## 10509 4577 4577 and 4578    NC292 Ames 27142  Nitrogen 5-Sep       16.482220
## 10515 4579 4579 and 4580     SD44  PI 524969  Nitrogen 5-Sep       51.163657
## 10533 4583 4583 and 4584    NC262  PI 531085  Nitrogen 5-Sep        9.914150
##        Rep
## 10475 Rep2
## 10486 Rep2
## 10496 Rep2
## 10509 Rep2
## 10515 Rep2
## 10533 Rep2
```

```
### Defining the model to get multiple linear regression for date 6 Jul
fit12 <- lm(Canopy_Coverage ~ Genotype + Treatment + Genotype:Treatment + Rep,
          data=d12)
```

```r
#generating ANOVA table
summary(aov(fit12))
```

```
##                      Df Sum Sq Mean Sq F value   Pr(>F)
## Genotype            232 343152    1479  17.036  < 2e-16 ***
## Treatment             1    369     369   4.251   0.0399 *
## Rep                   1   5681    5681  65.430 6.56e-15 ***
## Genotype:Treatment  224  21211      95   1.091   0.2250
## Residuals           419  36379      87
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
#As the p-value is less than the significance level 0.05, we can conclude that there are significant di

##Calculating Vprogeny
out12 <- summary(aov(fit12))[[1]]
vprogeny12 <- (out12[1,3] - out12[4,3])/(2*2)

#Calculating H2
h2_12 = vprogeny12/((vprogeny12 + out12[4,3]/(2*2)))

##H2 is very high 0.94 This means that canopy coverage is highly influenced by genetic effects more tha
```

```r
#Creating a table with dates and h2 values
datest=table(cc$date)
datest
```

```
##
##  1-Sep 12-Aug 14-Aug 16-Aug 20-Aug 22-Aug 23-Aug 26-Aug  3-Sep 30-Aug  5-Sep
##    878    878    878    878    878    878    878    877    878    878    878
##  6-Jul
##    878
```

```r
typeof(datest)
```

```
## [1] "integer"
```

```r
mytable = data.frame(date=c("06-Jul","12-Aug","14-Aug","16-Aug","20-Aug","22-Aug","23-Aug","26-Aug","30
                     h2=c(h2_1,h2_2,h2_3,h2_4,h2_5,h2_6,h2_7,h2_8,h2_9,h2_10,h2_11,h2_12))

mytable[order(mytable$date),]
```

```
##      date        h2
## 10 01-Sep 0.9237515
## 11 03-Sep 0.9181626
## 12 05-Sep 0.9359811
## 1  06-Jul 0.8864210
## 2  12-Aug 0.9369237
## 3  14-Aug 0.9262842
## 4  16-Aug 0.9248455
```

```
## 5  20-Aug 0.9216848
## 6  22-Aug 0.9294018
## 7  23-Aug 0.9258939
## 8  26-Aug 0.9221132
## 9  30-Aug 0.9114451
```

```
mytable
```

```
##      date        h2
## 1  06-Jul 0.8864210
## 2  12-Aug 0.9369237
## 3  14-Aug 0.9262842
## 4  16-Aug 0.9248455
## 5  20-Aug 0.9216848
## 6  22-Aug 0.9294018
## 7  23-Aug 0.9258939
## 8  26-Aug 0.9221132
## 9  30-Aug 0.9114451
## 10 01-Sep 0.9237515
## 11 03-Sep 0.9181626
## 12 05-Sep 0.9359811
```

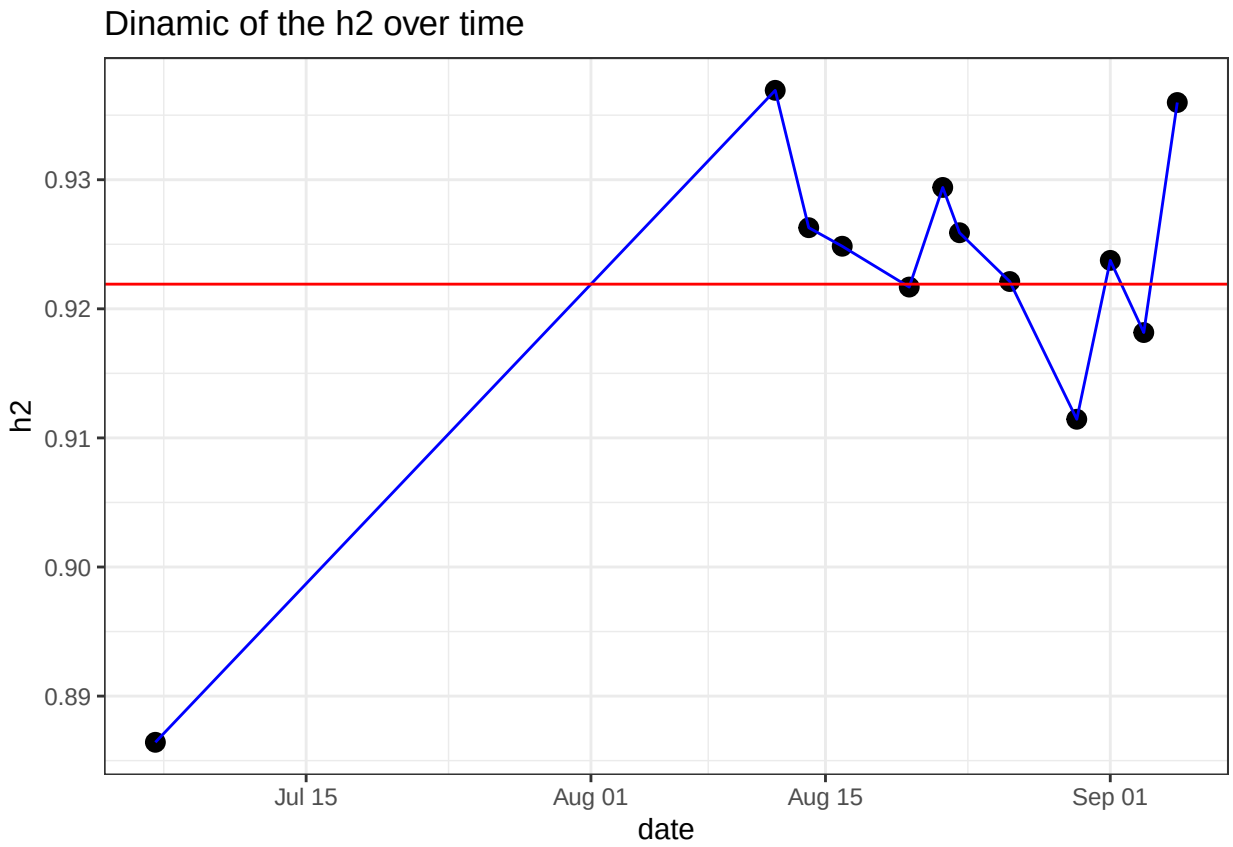
```
class(mytable)
```

```
## [1] "data.frame"
```

```
mytable$date=as.Date(mytable$date, "%d-%b")

#Visualize the data
library(ggplot2)

ggplot(mytable,aes(date,h2,color="h2 dinamic"))+
  geom_point(colour="black", size=3)+ geom_line(colour="blue")+ labs(title="Dinamic of the h2 over time
  theme_bw()+geom_hline(yintercept=mean(mytable$h2),color="red")
```




```
#My Hypothesis is that according with plant physiology, the more older the crop, the more canopy covera

#According to the graph, the h2 value is increasing over time. We can see that the lowest value is in t
#We can see that except for the first data point, the h2 is almost constant and close to the h2 mean ov

#The heritability can change over time because the variance of the genetic values can change due to env

#At the first dates, maybe the plant is in a growing stage in which do not express a lot the genes but

#According to the graph, I reject the null hypothesis. The h2 dynamic over time in the graph is showing
```