N-rate Timing

3/21/23

To do

Distill brainstorming done in the EDA scripts, report.Rmd and Dominics word doc into a singular first draft.

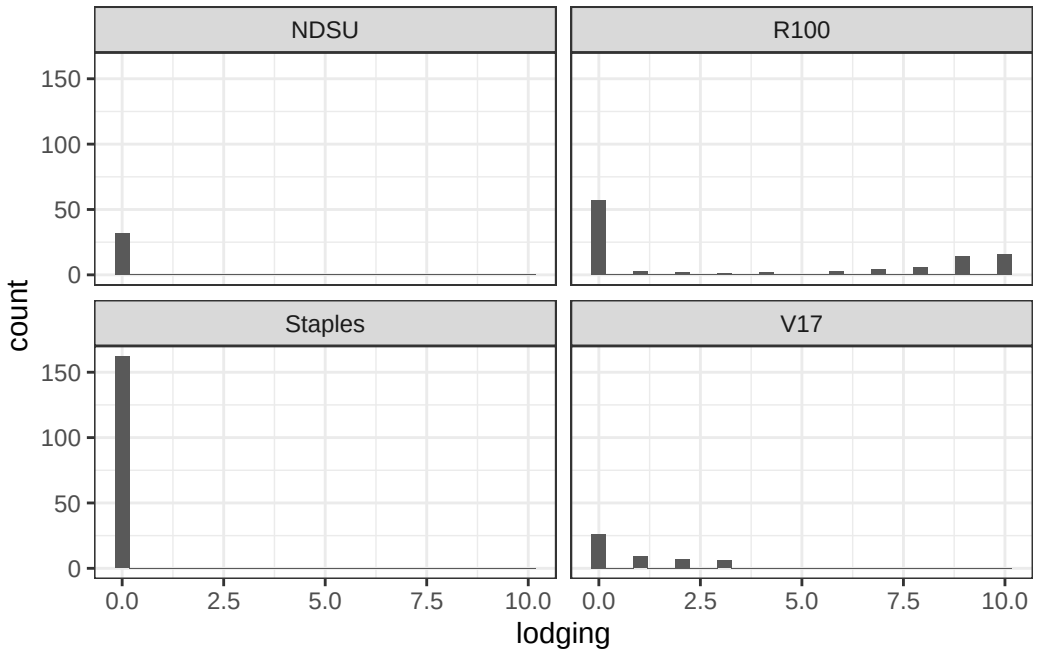
Data

Methods

Lodging

Understanding lodging is not a primary objective of this experiment, this data was only collected as a covariate if there happened to be a lot of lodging.

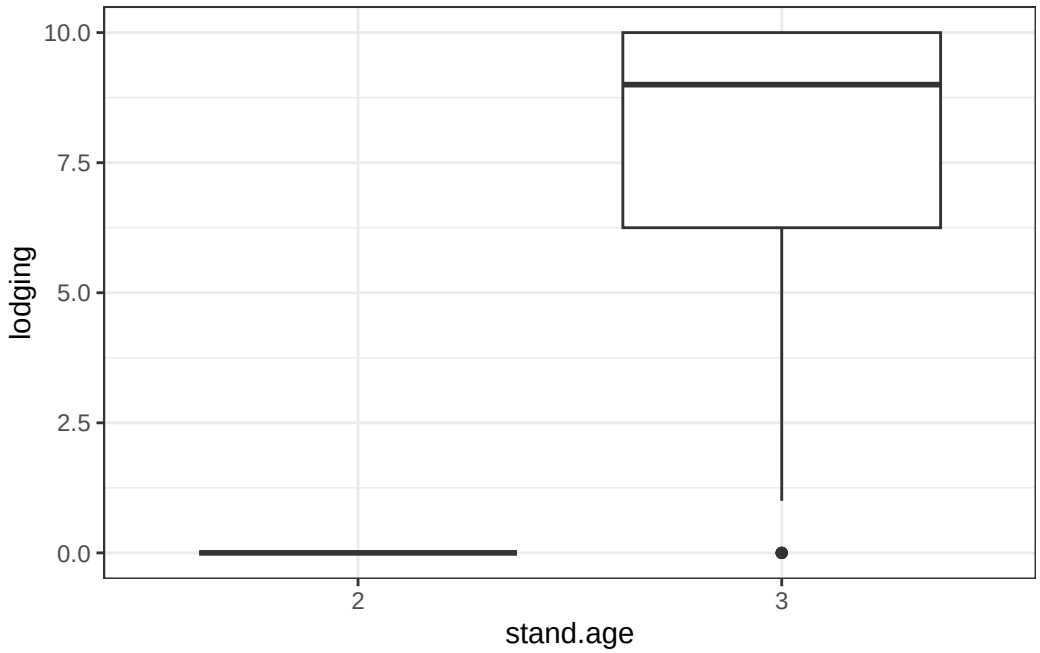
In general, if lodging is above 6, the yield data is questionable.



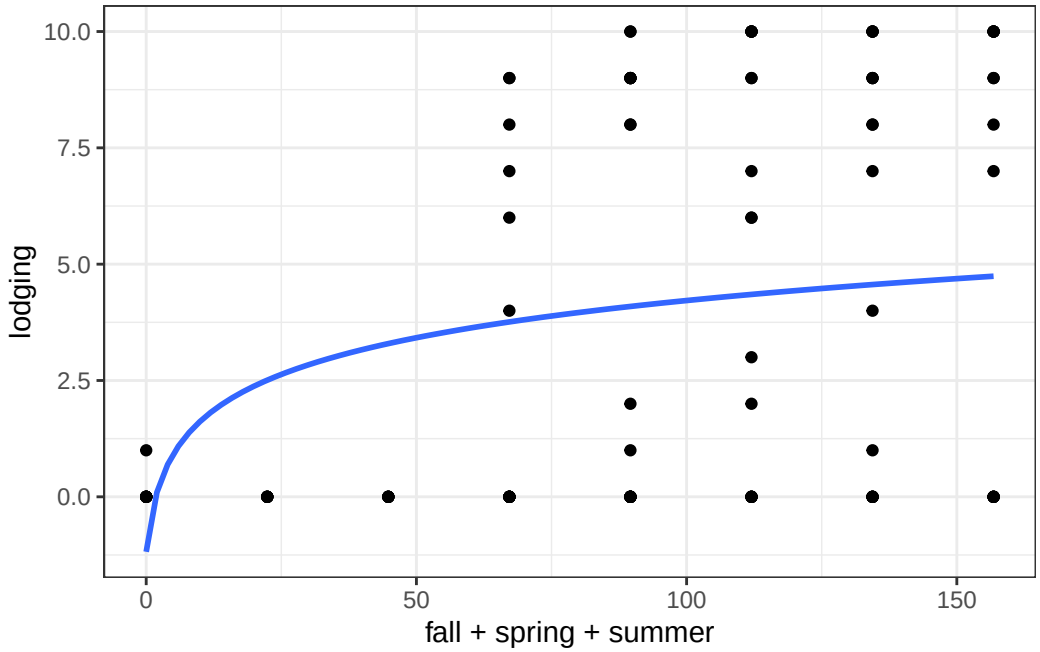
Only R100 and V17 showed lodging, and only R100 had severe lodging to the point where the yield data probably is not very accurate.

R100

R100 only had N applied in the second and third year.

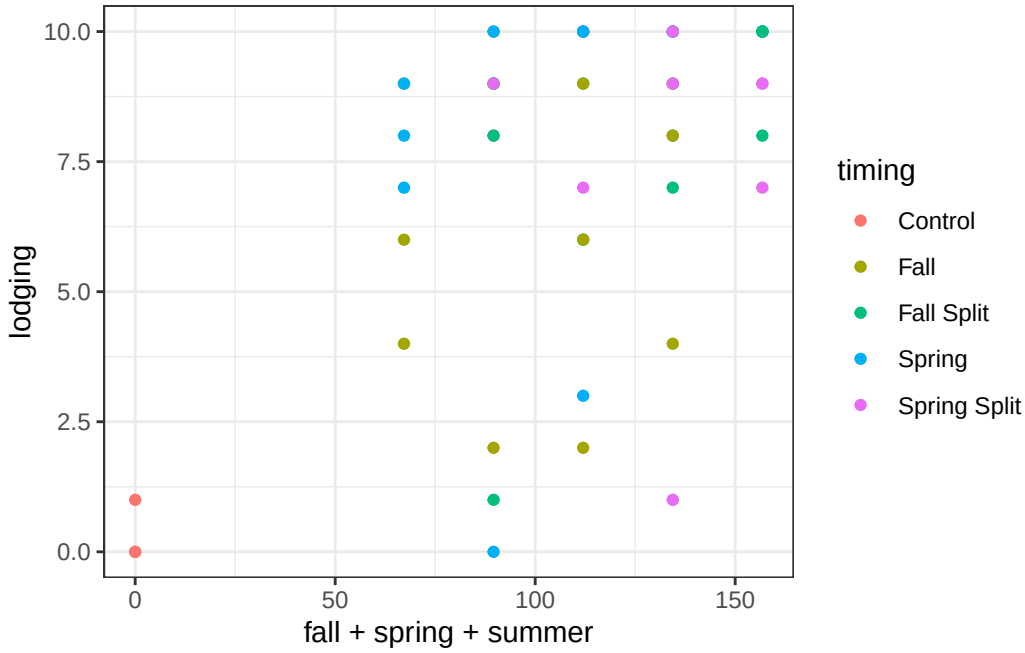


Lodging only occurred in the stands third year of production.

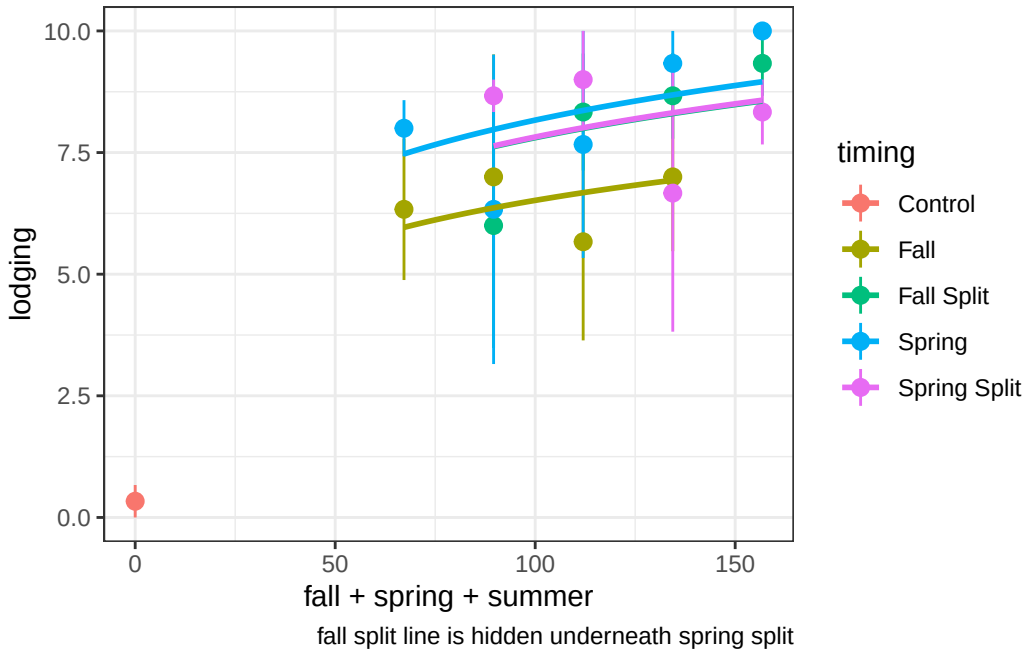


We observe a general increase in lodging as nitrogen rate increases. We fit a logistic curve since lodging cannot be greater than 10 and we expect as we increase N rate more lodging will

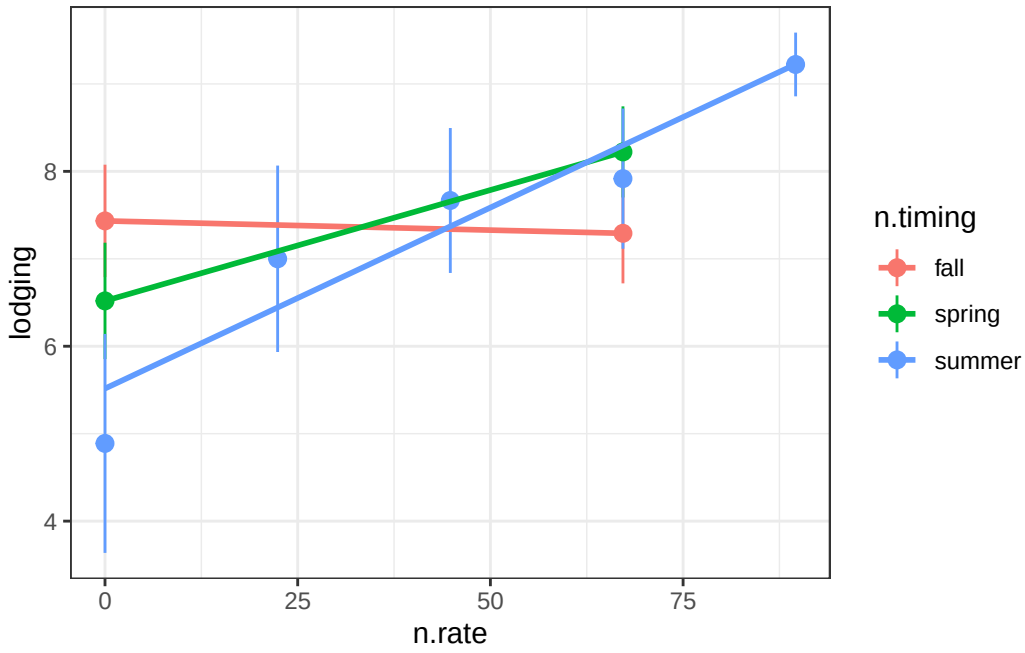
get closer to 10. This curve is obviously not perfect.



Looking at treatment timing with N rate, I see a bit more red dots (fall) with lower lodging, but it's messy. Sprint split has some of the most lodging, but also an outlier at near 150 kg N ha where there was no lodging



Here we are fitting a logistic regression with a y intercept of zero because we assume at 0N there is no lodging (as shown with control plots) and that lodging score will increase as nitrogen rate increases but that lodging will never exceed 10. The takeaway from this figure is that there is no lodging at 0N and that you see less lodging when you apply in fall and more when you apply in spring and summer.



This is a funkier way to look at this data based on how I organized the data and it is confusing so may not be worth trying to interpret.. Every lodging observation has a corresponding amount of N applied in fall, spring and summer. All of the spring plots, for example, also are plots where no N was applied in the fall. So within each of the timings (fall, spring or summer), the entire lodging dataset is represented. This also means the lodging dataset is duplicated 3 times when we do this. I am not sure if this is ok, but this is how I think this can be interpreted.

With more N there is more lodging, but we also see more lodging when the N is applied closer to harvest.

Table 1: Linear regression parameters for lodging response to nitrogen timing

Timing of N	m (lodging ~ m*nrate + b)
fall	-0.002
spring	0.025
summer	0.041

Analysis

When we fit logistic curves, it looked like the lodging was lowest in the control and fall plots. Let's see if this is true.

```
Analysis of Variance Table

Response: lodging
                        Df  Sum Sq Mean Sq  F value Pr(>F)
log(n.total + 1)         1 3109.08 3109.08 428.8900 <2e-16 ***
timing                   5   19.87    3.97   0.5481 0.7388
log(n.total + 1):timing  3   18.84    6.28   0.8664 0.4655
Residuals               45  326.21    7.25
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We cannot reject Ho that timing has an impact on lodging.

```
 timing       emmean    SE df lower.CL upper.CL .group
 Control       0.333 1.559 49    -2.80     3.47  1
 Fall          6.500 0.779 49     4.93     8.07   2
 Fall Split    8.083 0.779 49     6.52     9.65   2
 Spring Split  8.167 0.779 49     6.60     9.73   2
 Spring        8.267 0.697 49     6.87     9.67   2

Confidence level used: 0.95
P value adjustment: tukey method for comparing a family of 5 estimates
significance level used: alpha = 0.05
NOTE: If two or more means share the same grouping symbol,
      then we cannot show them to be different.
      But we also did not show them to be the same.
```

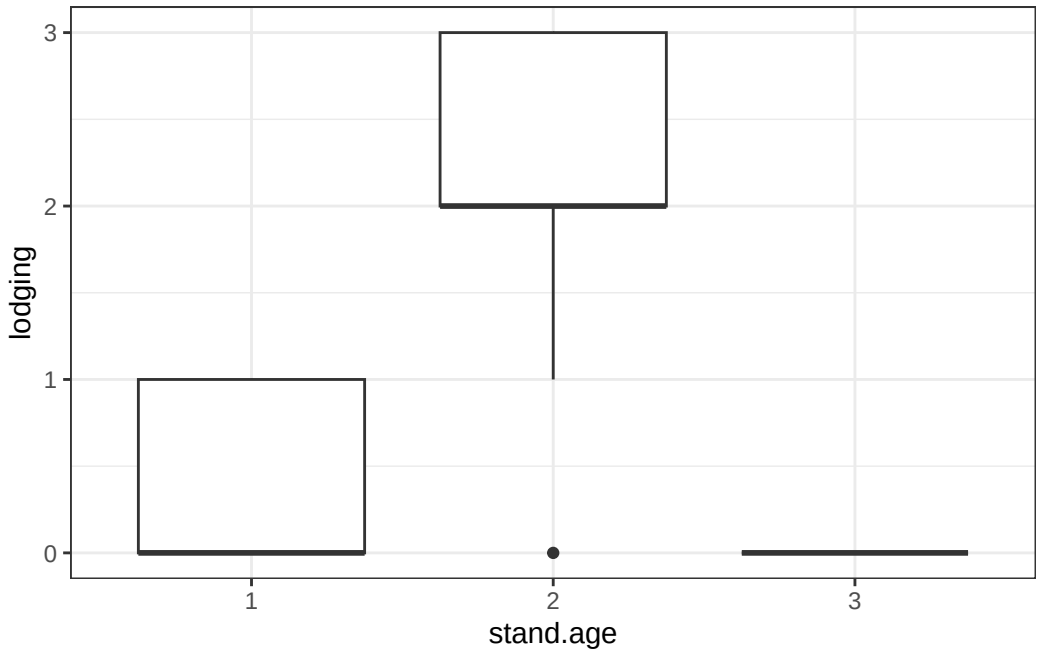
Obviously, if we look just at timing then the control plots have less lodging than the rest of the plots receiving N. Takeaway here is that lodging is correlated with increases in N, but we cannot reject the Ho that timing does not have an impact. It looks like fall applied N had less impact on lodging, but we cannot reject Ho.

Based on analysis, we are better off just fitting a general curve to lodging by nrate, we will need to do more analysis to find the best curve with the most fitting relationship.

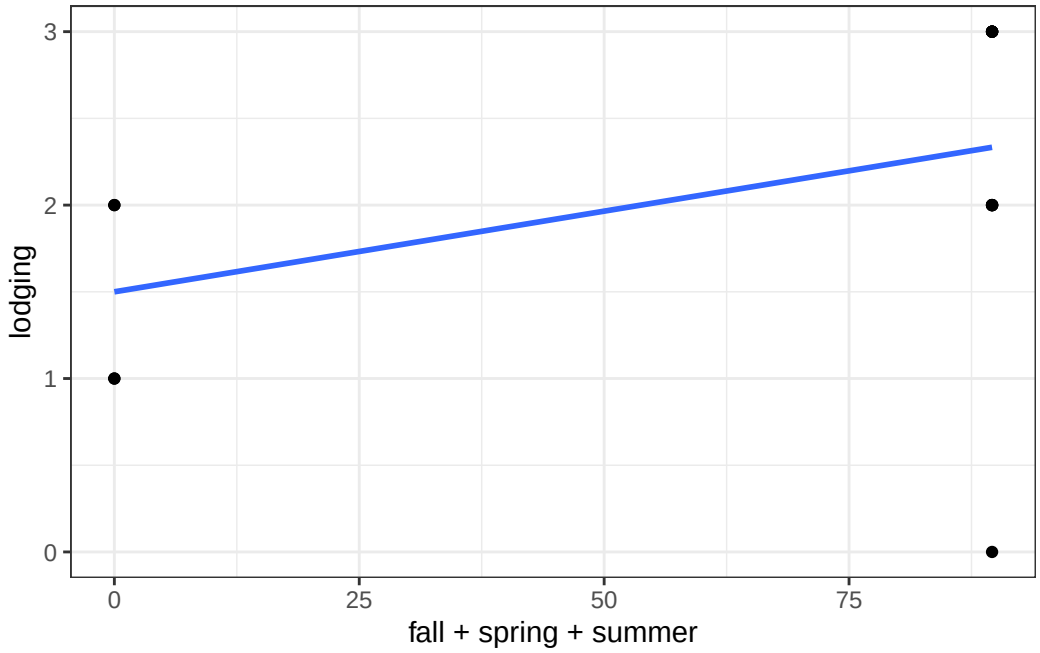
V17

Lodging is below threshold of 6, meaning we think this yield data is still good.

We are just looking to see if similar trends to R100 where applying more N leads to more lodging and where applying N closer to harvest correlates with more lodging

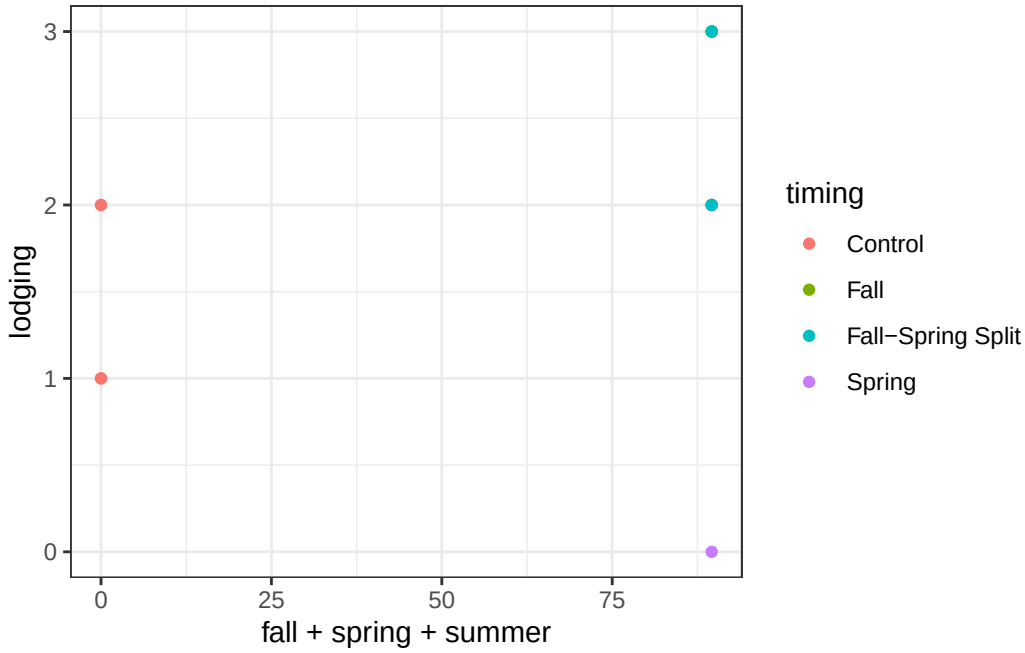


Lodging only occurred in the stands second year of production. This could've just been a harvest timing issue, but harvest was 3Aug so unlikely.



Not a strong relationship between N rate and lodging. There is a good amount of lodging in

control. Nothing to see here.



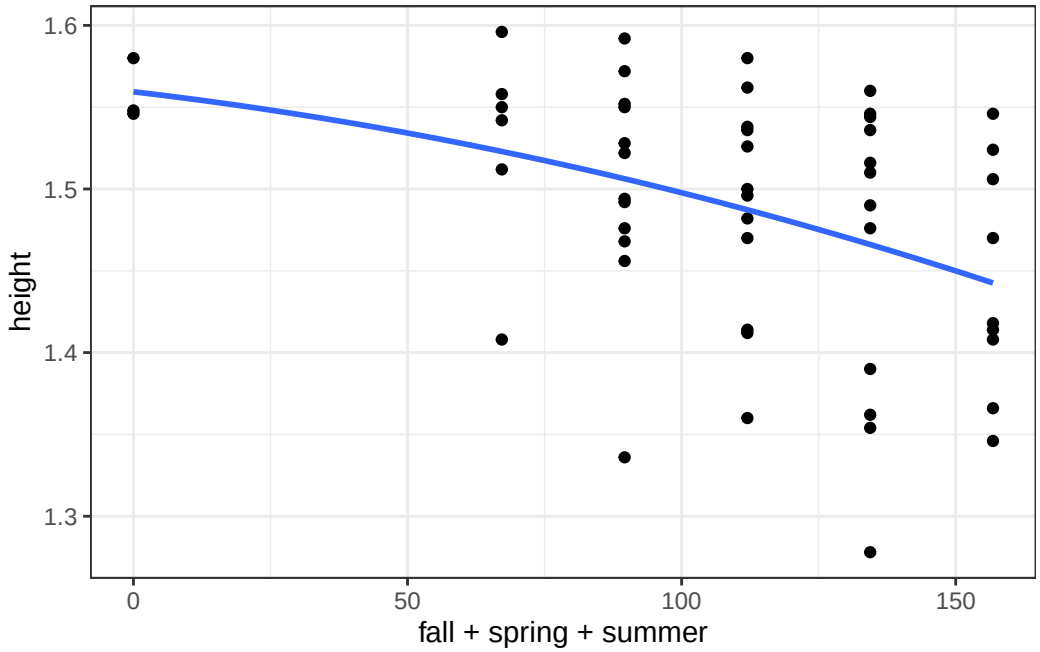
Plant height

Plant height is also not a measurement of primary interest.

We are curious to what extent plant height relates to lodging in R100

We are curious more broadly how plant height was distributed across experiments.

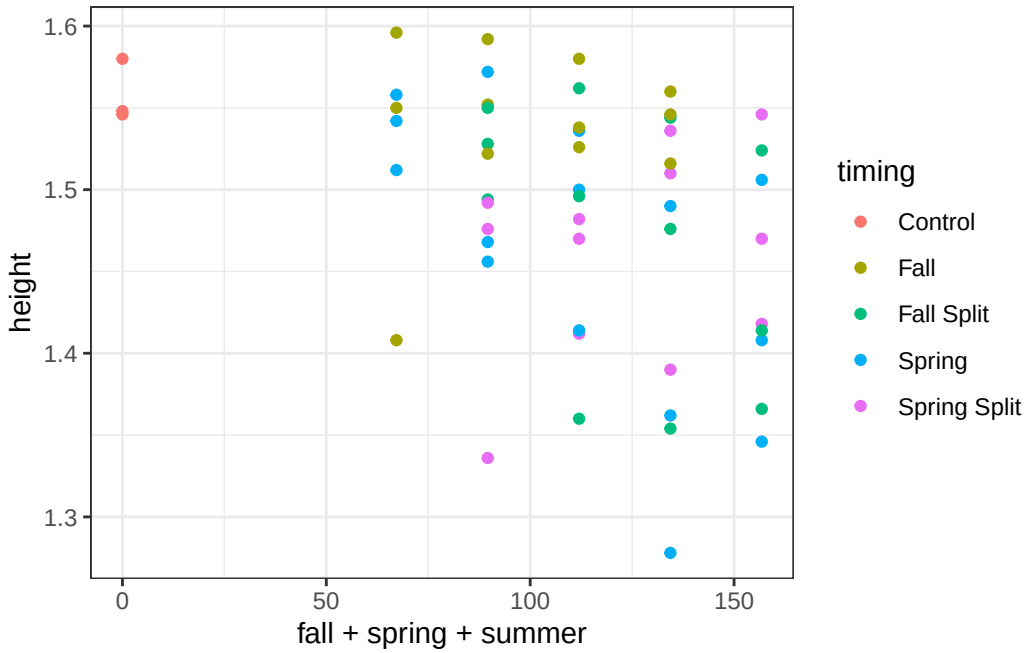
R100 - third year with lodging only



We observe an overall trend of decreasing plant height as N rate increases.

```
Analysis of Variance Table

Response: height
          Df   Sum Sq  Mean Sq F value   Pr(>F)
n.total    1 0.049647 0.049647 10.7931 0.001908 **
timing     4 0.037013 0.009253  2.0116 0.107787
Residuals 48 0.220797 0.004600
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Cannot reject Ho that timing has no effect on the effect of n rate on plant height



We have learned from R100 in it's third stand age that as nrate increases, there is an increase in lodging and a decrease in plant height.

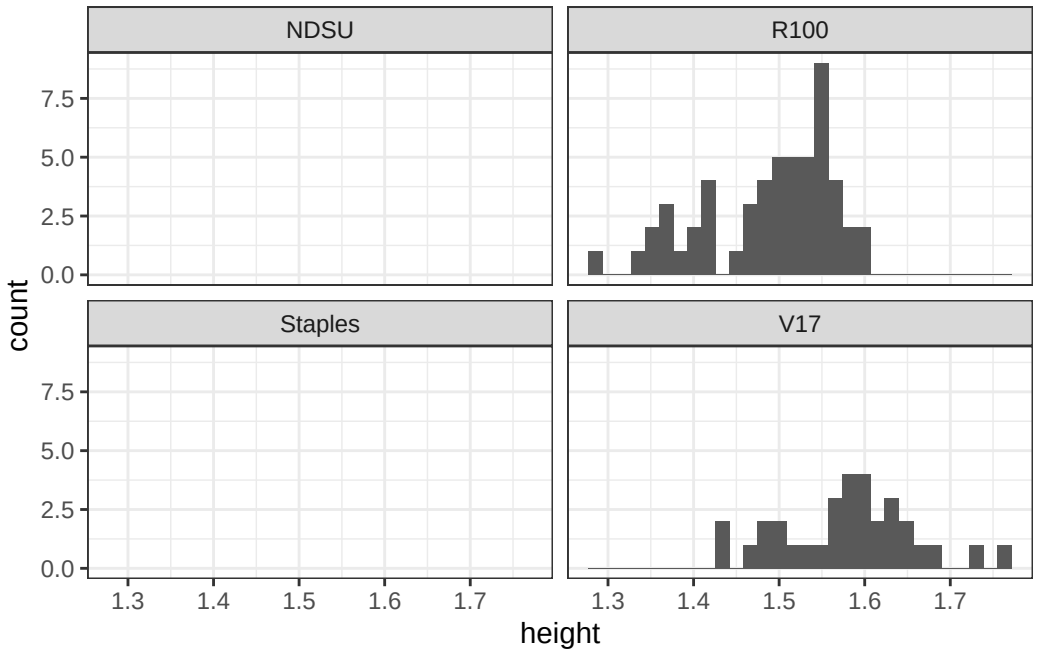
Let's see if they're correlated

```
           lodging     height
lodging  1.0000000 -0.4993655
height  -0.4993655  1.0000000
```

We observe a pearson correlation coefficient of -0.5 between height and lodging. This is considered between a moderate and strong correlation.

Beyond R100

We only have height data for V17 in stand ages 2 and 3 besides the R100 in stand age 3.



alright, we only have two sites with plant height data

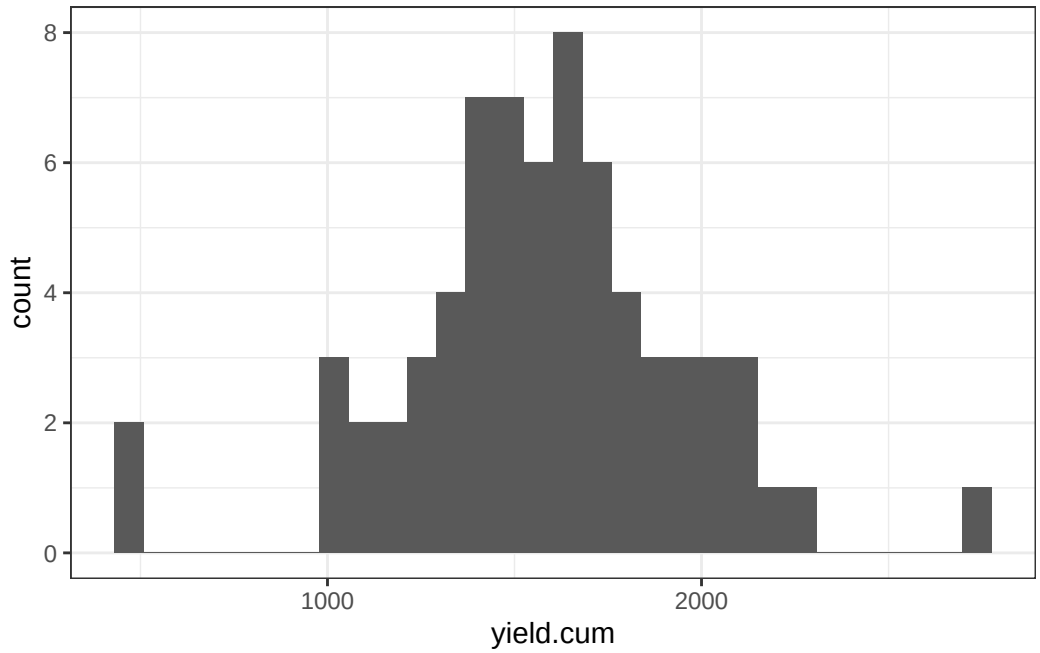
```
Analysis of Variance Table

Response: height
          Df   Sum Sq  Mean Sq F value Pr(>F)
n.total    1 0.000056 0.000056  0.0092  0.924
Residuals 30 0.181829 0.006061
```

We cannot reject Ho that plant height is the same across n rates

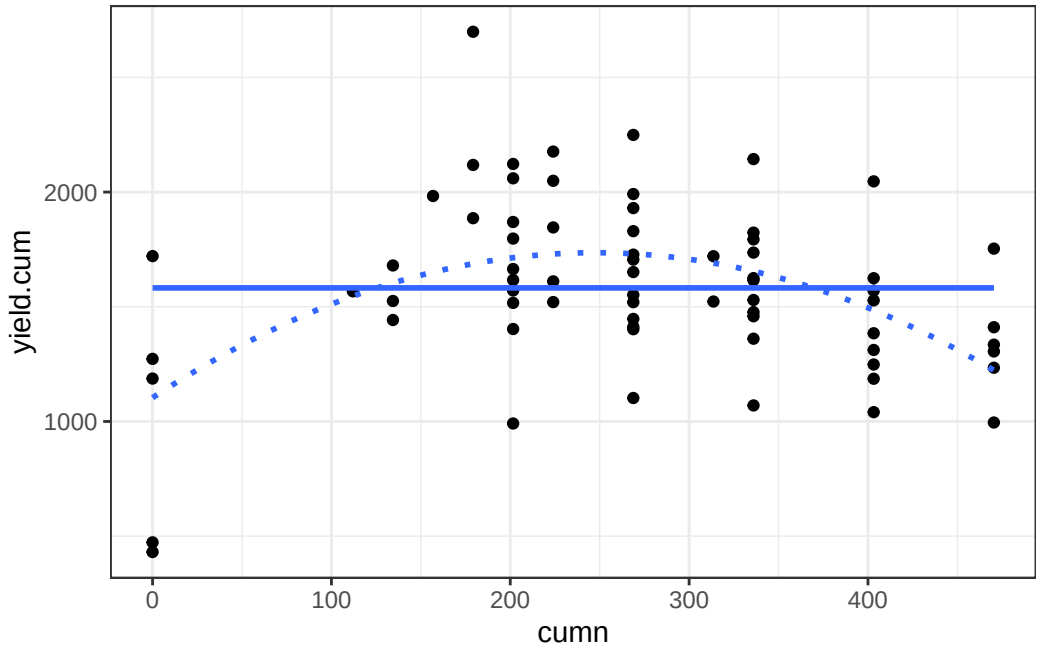
Yield

cumulative

Cumulative yield of kernza stands after 3 years of N fertilizer (V17 and Staples)



quadratic model



```
[1] 1000.713
```

```
[1] 1020.782
```

Ok, so there is an effect of cumulative N when modelled quadratically versus linear

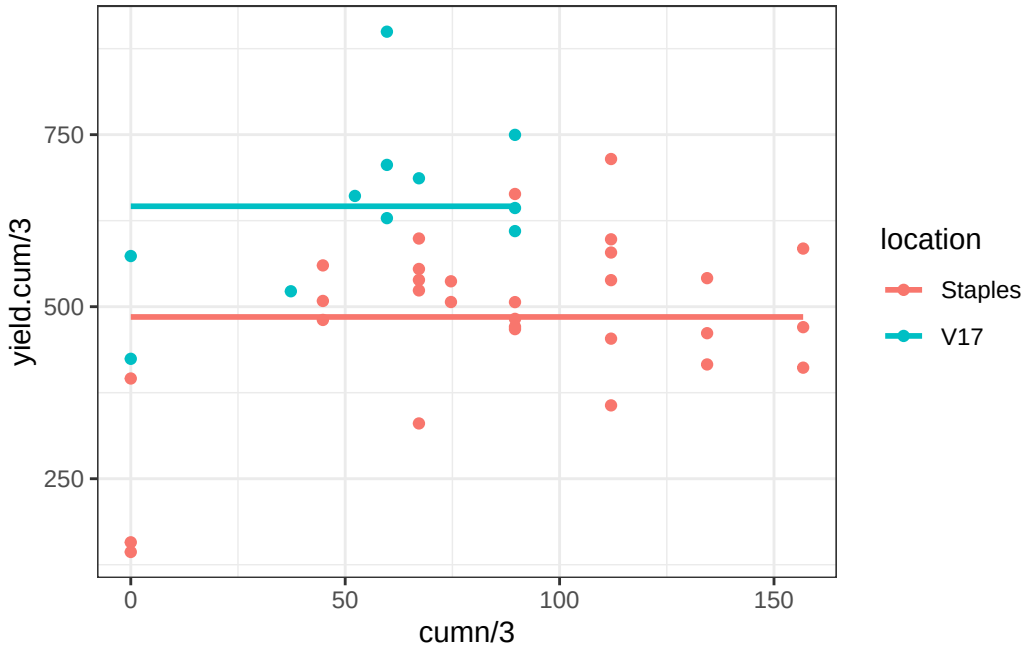
```
Analysis of Variance Table

Response: yield.cum
                       Df  Sum Sq Mean Sq F value    Pr(>F)
poly(cumn, 2)           2 2011131 1005565 12.0239 0.0001701 ***
location                1 2181936 2181936 26.0901 2.065e-05 ***
timing                  2  284394  142197  1.7003 0.2009473
poly(cumn, 2):location  2  114253   57126  0.6831 0.5132778
poly(cumn, 2):timing    3  186149   62050  0.7419 0.5360498
location:timing         2   71804   35902  0.4293 0.6551831
Residuals              28 2341662   83631
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Now that I'm modeling quadratically, we can reject the Ho there is no relationship between nitrogen rate and yield, but we cannot reject Ho that timing doesn't matter

mixed effect model

```
Analysis of Deviance Table (Type II Wald chisquare tests)

Response: yield.cum
                              Chisq Df Pr(>Chisq)
poly(cumn, 2)                0.5868  2     0.7457
location                    27.2054  1  1.829e-07 ***
timing                       4.0312  2     0.1332
poly(cumn, 2):location       1.1353  2     0.5669
poly(cumn, 2):timing         2.8742  3     0.4114
location:timing              0.8907  2     0.6406
poly(cumn, 2):location:timing        0
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The best model I can create (lowest AIC) finds a difference in cumulative yield by location, but not an effect of nrate, timing or an interaction.



TAKEAWAY: We applied N at differing rates and timings over 3 years at two locations. We cannot reject the Ho that the amount of N and the timing of N do not impact the cumulative yield over the 3 years. I would say our data suggests a quadratic or logarithmic relationship at around 60 kg N ha per year does not result in noticeably greater grain yields.