Yields_Report

Import data and libraries

```
library(knitr)
library(plyr)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
##
## The following objects are masked from 'package:plyr':
##
##     arrange, count, desc, failwith, id, mutate, rename, summarise,
##     summarize
##
## The following object is masked from 'package:stats':
##
##     filter
##
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(tidyr)
library(ggvis)
library(ggplot2)
```

```
##
## Attaching package: 'ggplot2'
##
## The following object is masked from 'package:ggvis':
##
##     resolution
```

```
library(gridExtra)
```

```
## Loading required package: grid
```

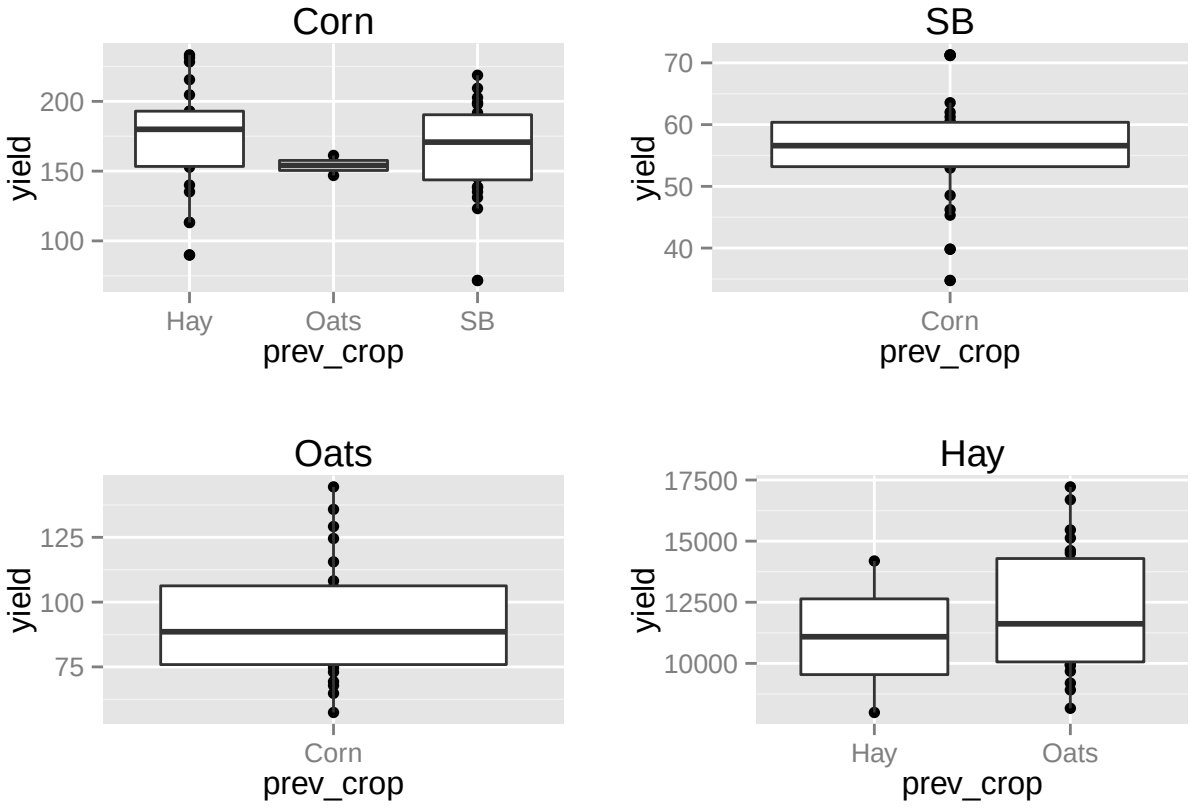
```
library(xtable)
library(reshape2)
dat <- read.csv("/Users/epwalsh/GitHub/PFI/data/PFI_clean.csv")
```

```
yields <- dat %>%
  filter(item_type %in% c("Expense"), farmer == "Thompson") %>%
  spread(item, value) %>%
  select(-c(5))

yields <- yields %>% mutate(year_before = year - 1) %>%
```

```
  left_join(select(yields, 1:4),
            by=c("field_id" = "field_id", "year_before" = "year")) %>%
  select(-c(2,38))

# Get crop yields
value <- dat %>%
  filter(item_type %in% c("Unit Quantity"), farmer == "Thompson") %>%
  select(c(1,3,7))

# Add crop yields to yields df
yields <- yields %>%
  left_join(value, by = c("field_id" = "field_id", "year" = "year")) %>%
  rename(crop = crop.x, prev_crop = crop.y, yield = value) %>%
  select(-3)

# Reorder variables
yields <- yields[,c(2,1,3,36,37,4:35)]
```

Yield by previous crop

Below are boxplots of the yield for each crop based on the crop that was planted on the same field during the previous year. Soy beans and oats always followed corn according to Thompson's rotation system. What is interesting is how the yields for corn and hay vary.



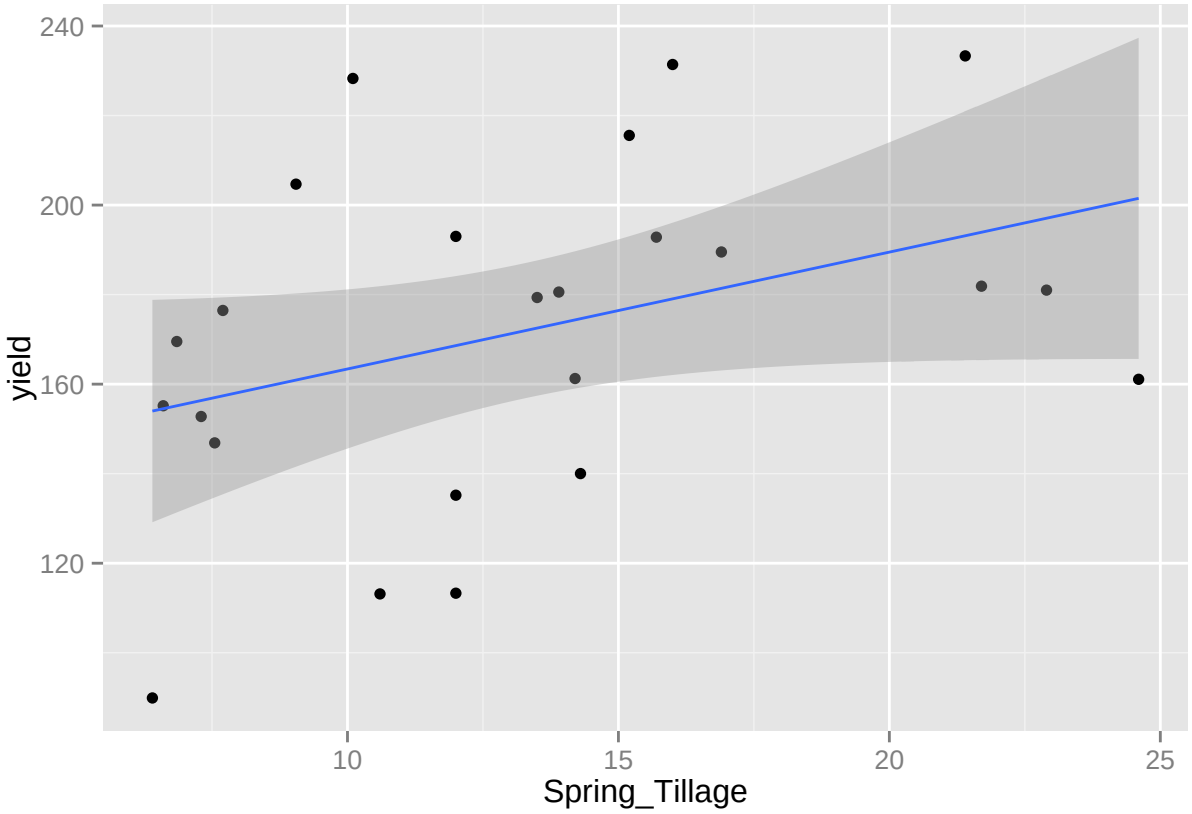
When oats were planted in the previous year, the yield for corn is significantly lower than when hay or soy beans were planted the year before. However there is not much of a difference between the yield of corn when the previous crop was hay versus soy beans. The yield of hay is also slightly higher when the previous crop was oats, as opposed to hay.

A future study could further examine these variations by looking at how the yields of soy beans and oats are affected if the previous crop is something other than corn. Obviously the knowledge that one crop performs signficantly better when a certain crop is planted on the same field the year before could be very valuable information. This could help the farmer develop an optimized rotation strategy.

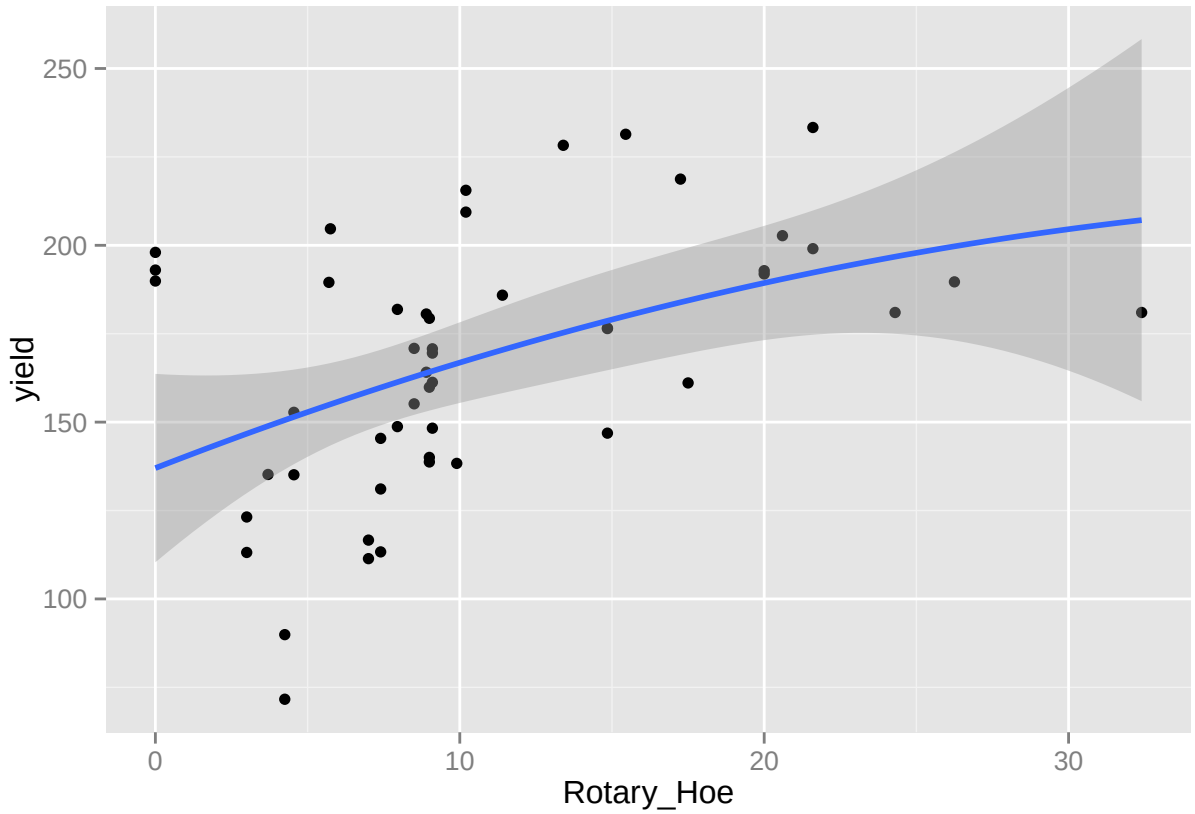
Yield by expense

After examining how each individual expense related to the yield of each crop, we found four interesting relationships: three having to do with corn and one having to do with soy beans.

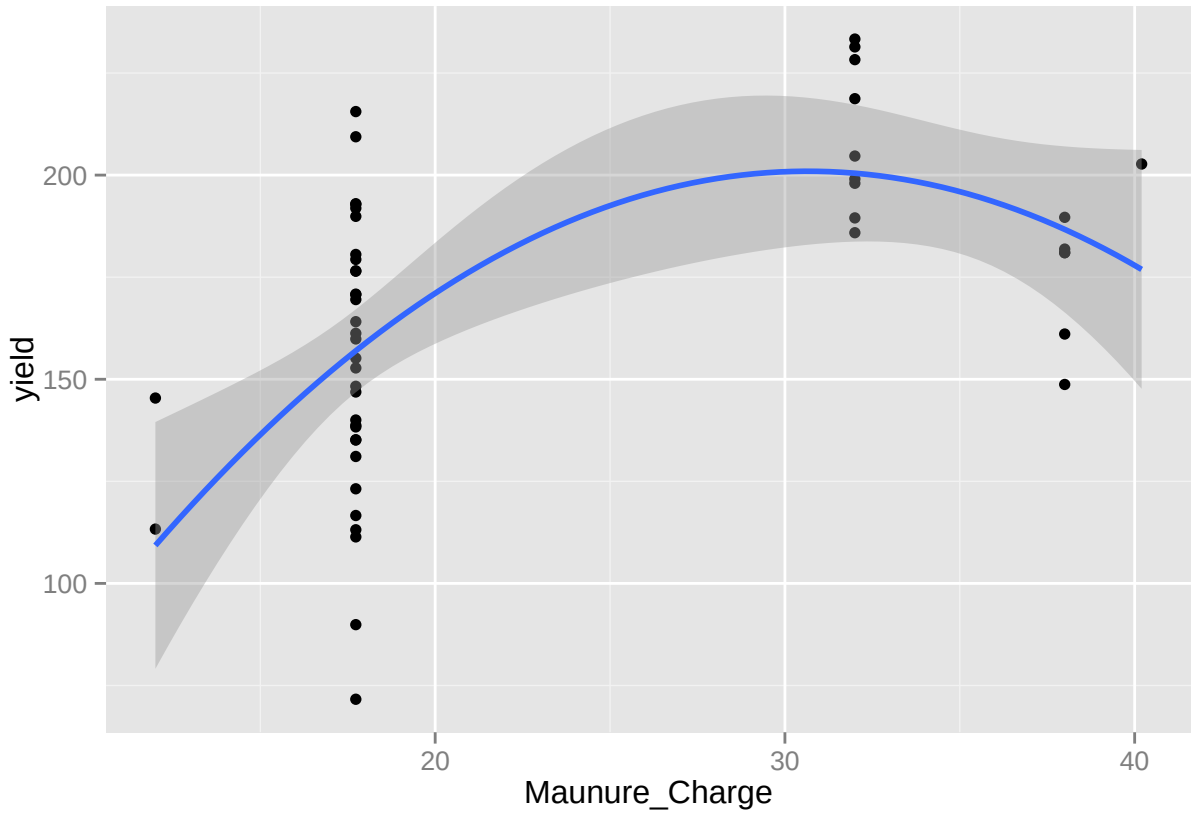
Corn and `Spring_Tillage` have a moderately strong positive relationship. As the money spent on tilling in the spring increases, the yield of corn generally increases. The plot below illustrates this relationship after removing all the years when no money was spent on spring tillage.



The plot below shows the expenses of `Rotary_Hoe` against the yield of corn with a quadratic trend fitted to the data. The trend suggests that the more money spent on `Rotary_Hoe`, the higher the expected yield will be, but the marginal benifit to yield decreases as the expense increases.

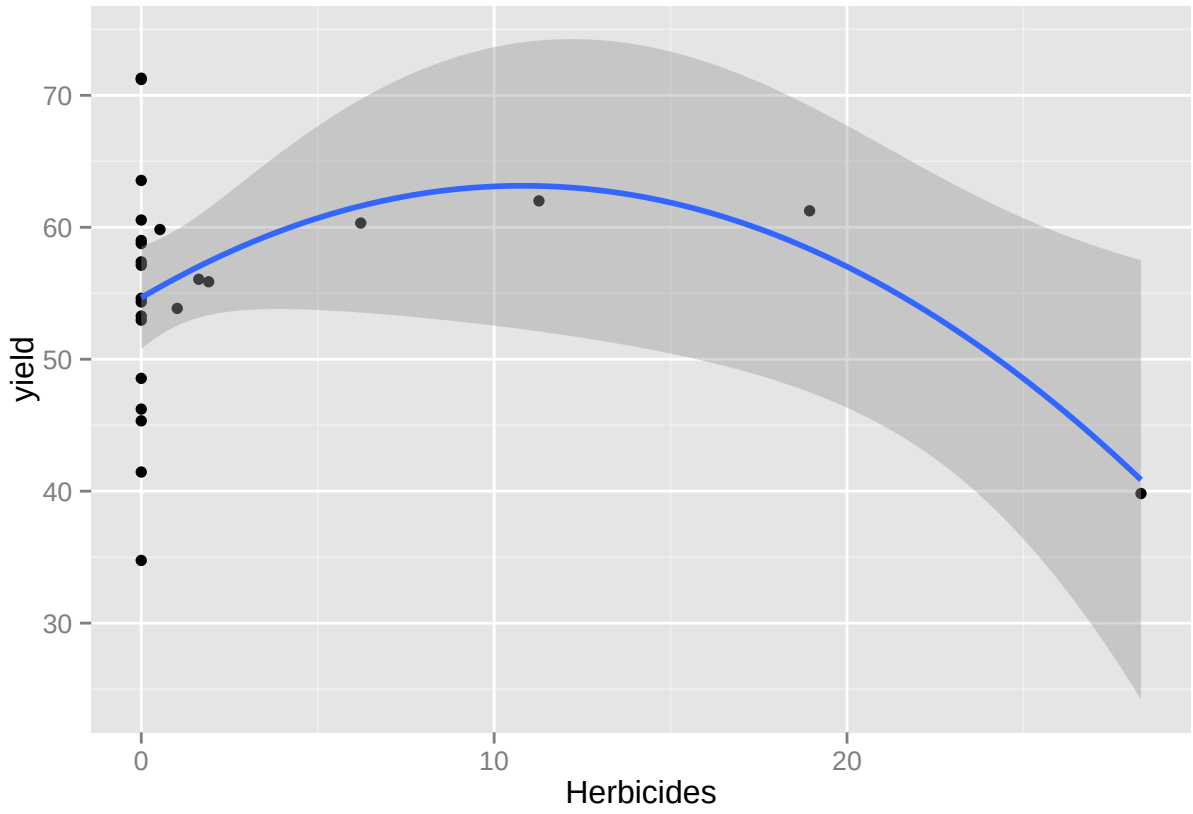


The amount spent on maunure exhibits a concave-down, quadratic relationship with the yield of corn. The expected yield increases as the expense on maunure increases up to a point, after which the expected yield actually decreases.



The amount of money spend on herbicides appears to have a quadratic, concave-down relationship with the yield of soy beans.

```
yields %>% filter(crop == "SB") %>%
ggplot(aes(Herbicides, yield)) + geom_point() +
stat_smooth(method = "lm", formula = y ~ x + I(x^2), size = 1)
```


70
60
50
40
30
yield
0
10
20
Herbicides