

Assignment 6

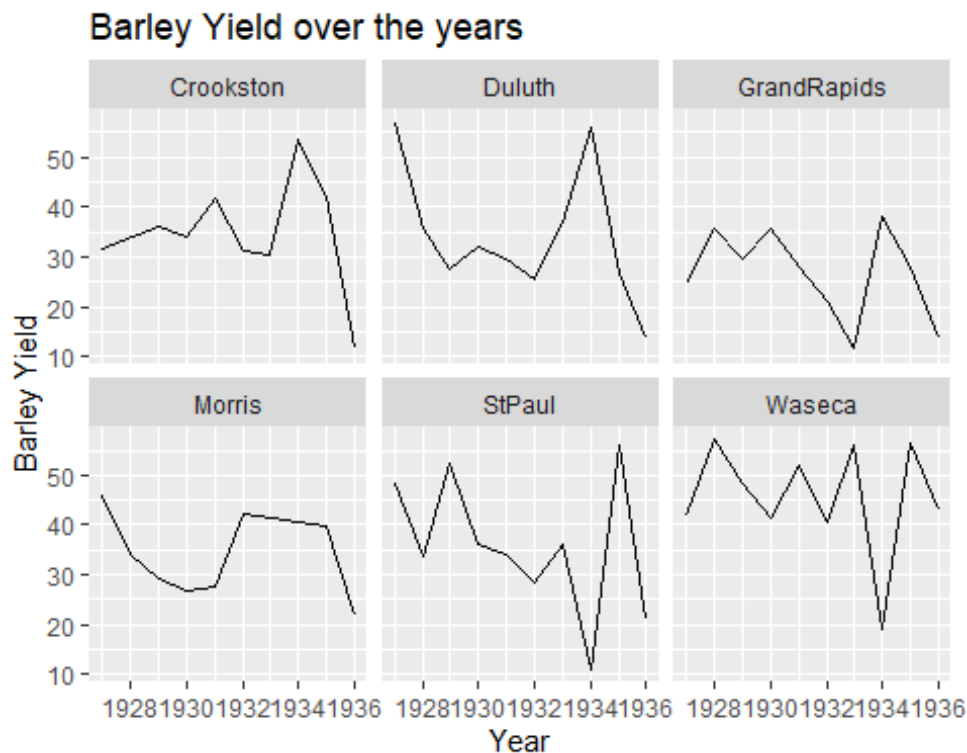
Ojaas Hampiholi

4/15/2020

```
barleyData <- as.data.frame(read.delim(file = "minnesota.barley.yield.txt", header = TRUE, sep = " "))
```

Question 1)

```
barleyAgg <- barleyData %>%  
  group_by(year, site) %>%  
  summarise(yield=mean(yield))  
  
ggplot(barleyAgg, aes(x = barleyAgg$year, y = barleyAgg$yield)) +  
  geom_line() +  
  facet_wrap(~ barleyAgg$site) +  
  ggtitle("Barley Yield over the years") +  
  xlab("Year") +  
  ylab("Barley Yield")
```



The total yield has a decreasing trend from 1927 to 1936 for all the locations except Waseca. The total yield for Waseca for 1936 is almost as same as 1927. The yield for each locations keeps fluctuating over the years.

Question 2)

```
barleyRLM <- rlm(yield ~ gen + factor(year) + site,  
psi = psi.bisquare, data = barleyData)  
summary(barleyRLM)
```

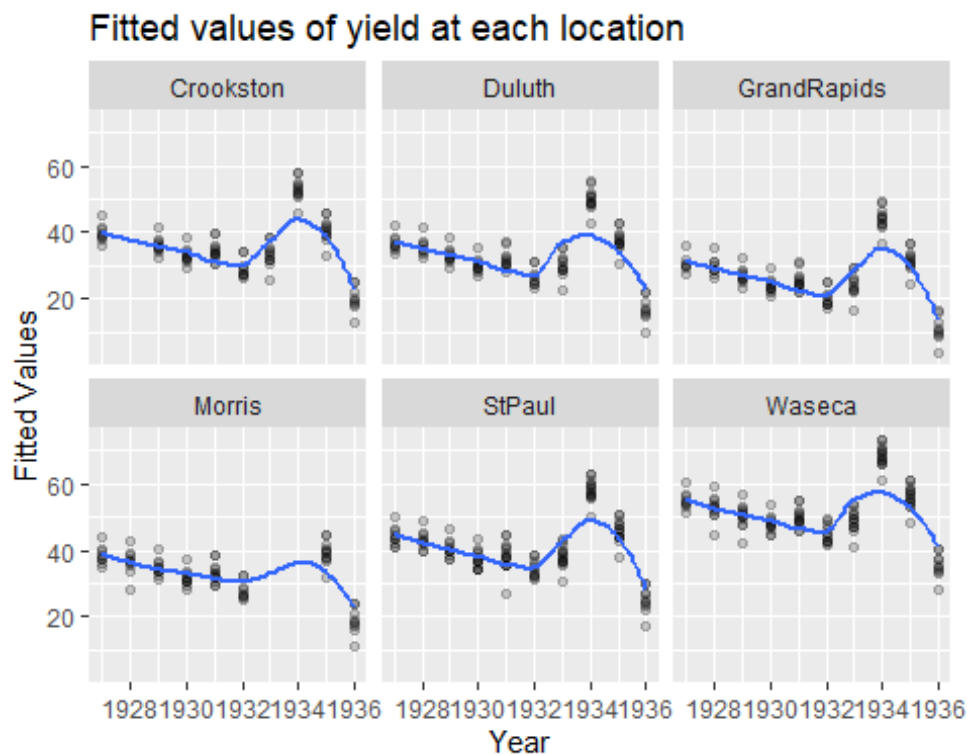
```
##  
## Call: rlm(formula = yield ~ gen + factor(year) + site, data = barleyData,  
##      psi = psi.bisquare)  
## Residuals:  
##      Min      1Q   Median      3Q      Max  
## -62.8500  -6.2416  -0.5303   4.9337  29.2314  
##  
## Coefficients:  
##              Value      Std. Error t value  
## (Intercept)    36.1974      2.0831    17.3764  
## genCompCross   -0.0787      6.4180    -0.0123  
## genDryland     -5.5412      5.3402    -1.0376  
## genGlabron      2.7955      2.0235     1.3815  
## genHeinrichs   -0.0242      3.0078    -0.0081  
## genJeans       -8.6872      8.9091    -0.9751  
## genManchuria    1.4855      2.0177     0.7363  
## genManxSA       4.7102      2.0494     2.2984  
## genMechMixture  0.1213      6.4180     0.0189  
## genMinsturdi    6.0172      3.0061     2.0016  
## genNo474        2.6279      2.1983     1.1954  
## genNo475        3.8993      2.7946     1.3953  
## genOderbrucker -3.4700      2.6003    -1.3344  
## genOdessa       6.0069      2.6003     2.3101  
## genPeatland     3.5442      2.0188     1.7556  
## genSAxMan       5.2776      2.0581     2.5643  
## genSD1340       4.1919      2.6693     1.5704  
## genSpartan      7.2709      3.5915     2.0244  
## genSvansota     2.6863      2.1314     1.2603  
## genTrebi        9.0159      2.0188     4.4661  
## genVelvet       2.2694      2.0188     1.1242  
## genWisNo38      9.1429      2.3003     3.9747  
## factor(year)1928 -1.0656      1.7419    -0.6118  
## factor(year)1929 -3.7258      1.6092    -2.3153  
## factor(year)1930 -6.6967      1.6029    -4.1778  
## factor(year)1931 -5.4798      1.5710    -3.4880  
## factor(year)1932 -11.3348      1.6624    -6.8184  
## factor(year)1933 -6.9980      1.6800    -4.1653  
## factor(year)1934 12.8885      1.6679     7.7273  
## factor(year)1935  0.4452      1.6064     0.2772  
## factor(year)1936 -20.1774      1.7596   -11.4672  
## siteDuluth      -2.9289      1.2176    -2.4055  
## siteGrandRapids -8.9979      1.2219    -7.3639  
## siteMorris      -1.3522      1.3120    -1.0307  
## siteStPaul       4.7695      1.1990     3.9778  
## siteWaseca      15.2932      1.1890    12.8627
```

```
##  
## Residual standard error: 8.446 on 611 degrees of freedom
```

LM tends to fit the data even if the data is a gross outlier, but RLM penalises the outliers in order to follow a better trend in the data. Hence I have used RLM here with yield as response variable and gen, year and site as explanatory variable.

Question 3)

```
library(broom)  
barleyData$fit = fitted.values(barleyRLM)  
barleyData$res = residuals(barleyRLM)  
  
ggplot(barleyData, aes(x=year,y=fit))+geom_point(alpha = 0.2)+geom_smooth(se=  
FALSE)+  
  facet_wrap(~site)+  
  ggtitle("Fitted values of yield at each location") +  
  xlab("Year") +  
  ylab("Fitted Values")  
  
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



From the graph, we can see that fitted graph is same over every location and year. We cannot see any unusual change in Morris for the 1931-1932 year. Average yield in 1932 is lower than average yield in 1931 for all the locations including Morris. Thus, Morris 1931-1932 should have been a natural variation and not a mistake in the recordings.