ASSignment 4

Harsh

October 13, 2017

Q1. (i)

solution

```r
#install.packages('agridat')
library("agridat")
```

```
## Warning: package 'agridat' was built under R version 3.3.3
```

```r
library("ggplot2")
```

```
## Warning: package 'ggplot2' was built under R version 3.3.3
```

```r
byds=agridat::minnesota.barley.yield

ggplot(data=byds, aes(x=year, y = yield ,colour = gen)) +
  geom_line()+ facet_wrap(~site)
```


q1.(ii)

```r
library("GGally")
```

```
## Warning: package 'GGally' was built under R version 3.3.3
```

```
library(broom)
```

```
## Warning: package 'broom' was built under R version 3.3.3
```
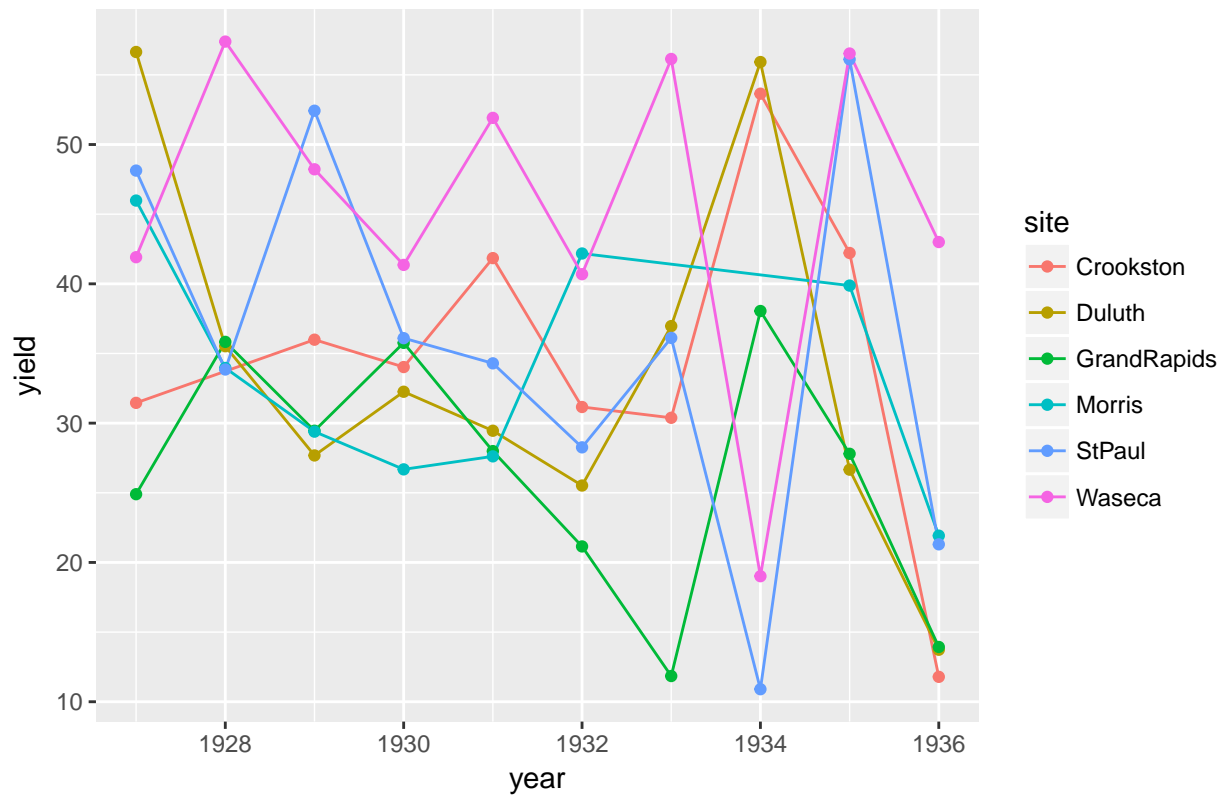
```
library(tidyr)
```

```
## Warning: package 'tidyr' was built under R version 3.3.3
```

```
bydsm=aggregate(yield~year+site,mean,data=byds)
```
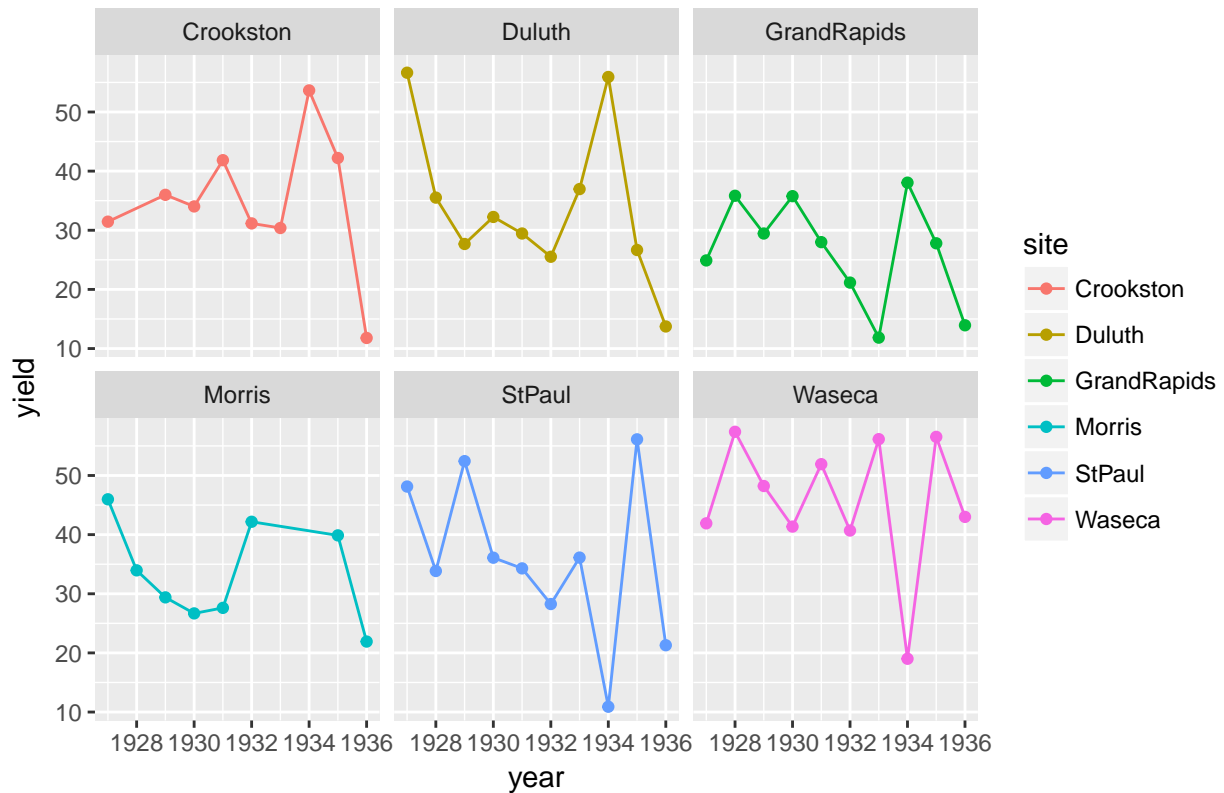
```
ggplot(data=bydsm, aes(x=year, y = yield ,colour = site)) +
  geom_line()+geom_point()+ggtitle("singal figure plot without facet")
```



```
ggplot(data=bydsm, aes(x=year, y = yield ,colour = site)) +
  geom_line()+geom_point()+facet_wrap(~site)+ggtitle("figure plot with facet")
```

figure plot with facet

if we see that above plot its that curves are not equal for years are not eaqual for yield movement but if we see the increase and decrease movement for location which almost equal all location for all yield,

Q2.

(i) Explain why you chose this model and not some other one.

interaction between the varity and variable are process with help of rlm model so selsect rlm().

```
library(MASS)
```

```
## Warning: package 'MASS' was built under R version 3.3.3
```

```
bydsrlm=rlm(yield~gen+year+site,psi = psi.bisquare,data=byds)


myg <- aggregate(yield ~ year + gen, mean, data =byds)
ggplot(myg, aes(x = yield, y = year,color=year)) +geom_point() +facet_wrap(~ gen) +
ggtitle(" the plot of Year Variety Interaction")
```
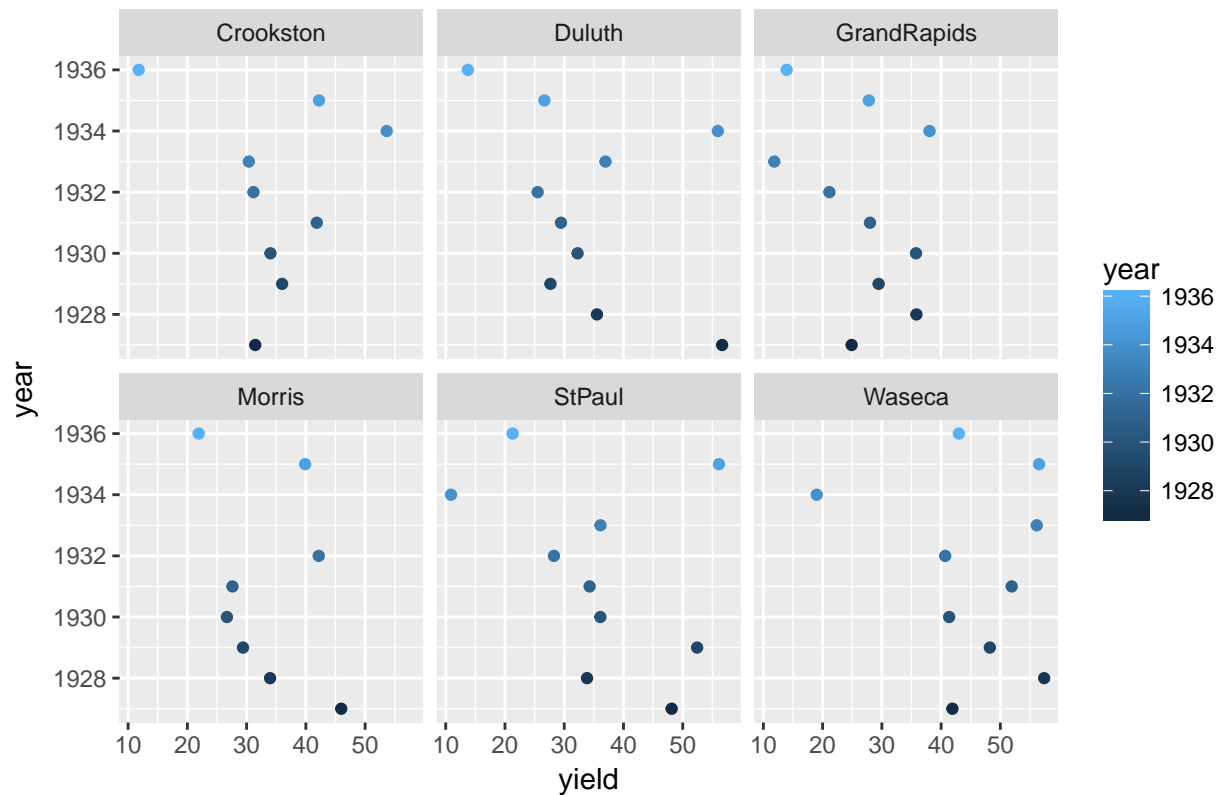
the plot of Year Variety Interaction
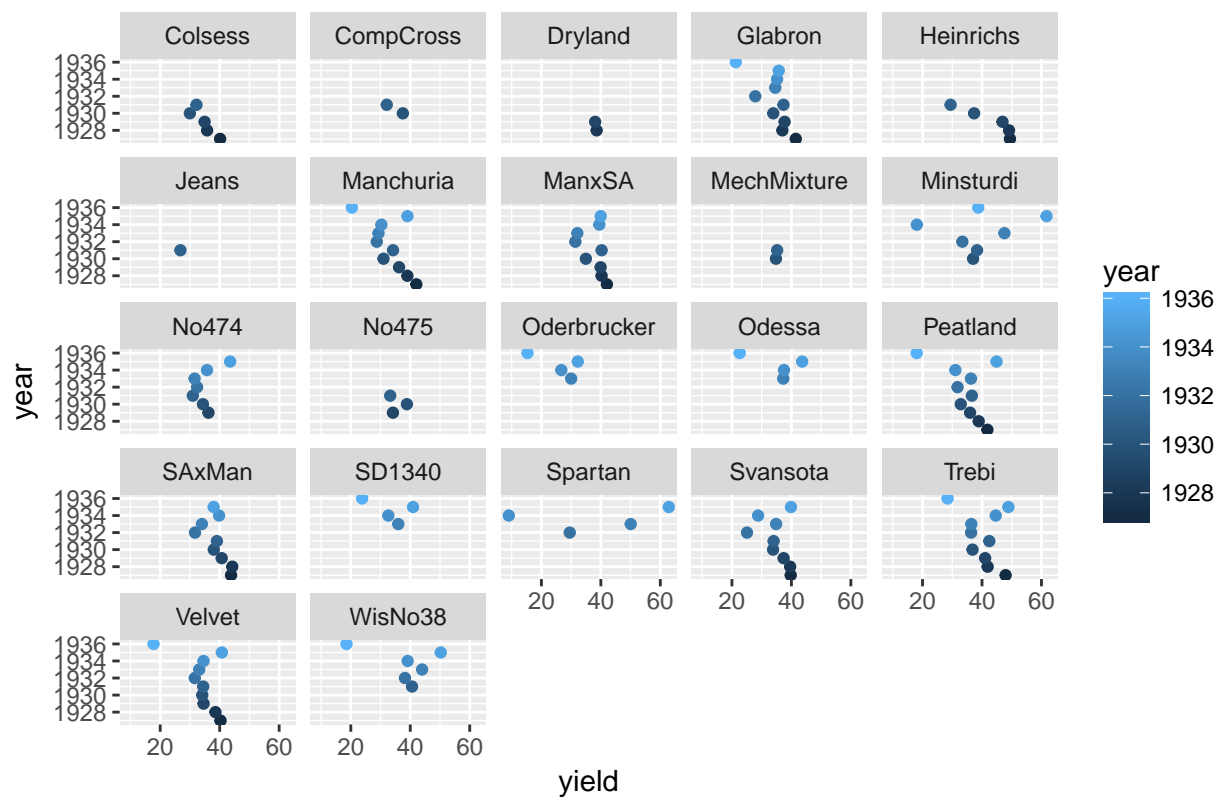

```r
mys <- aggregate(yield ~ year + site, mean, data = byds )
ggplot(mys, aes(x = yield, y = year, color=year)) + geom_point() + facet_wrap(~ site) +
ggtitle("the plot of YearSite Interaction")
```

the plot of YearSite Interaction

to selest rlm () the reason is the minnsota.barely.yieldtoo has outlier to corrected the problem wea are selecting rlm.
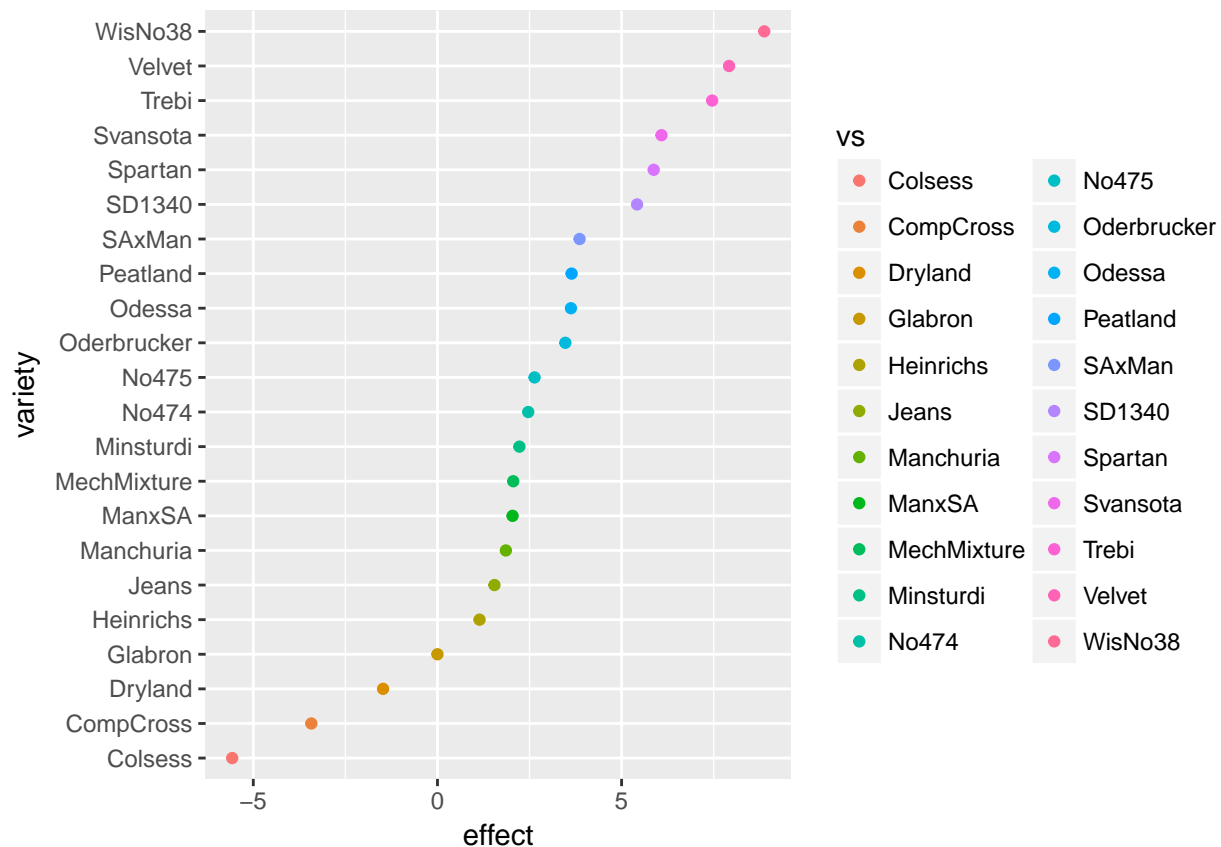
(ii) solutions

```r
library(MASS)
bydsrlm=rlm(yield~gen+year+site,psi = psi.bisquare,data=byds)


bts =dummy.coef(bydsrlm)
sts =sort(bts$site)
ss =factor(names(bts$site),levels =names(bts$site))
ssdf =data.frame(effect =sts,site =ss)
ggplot(data=ssdf,aes(x =effect,y =site,color=ss)) +geom_point()
```
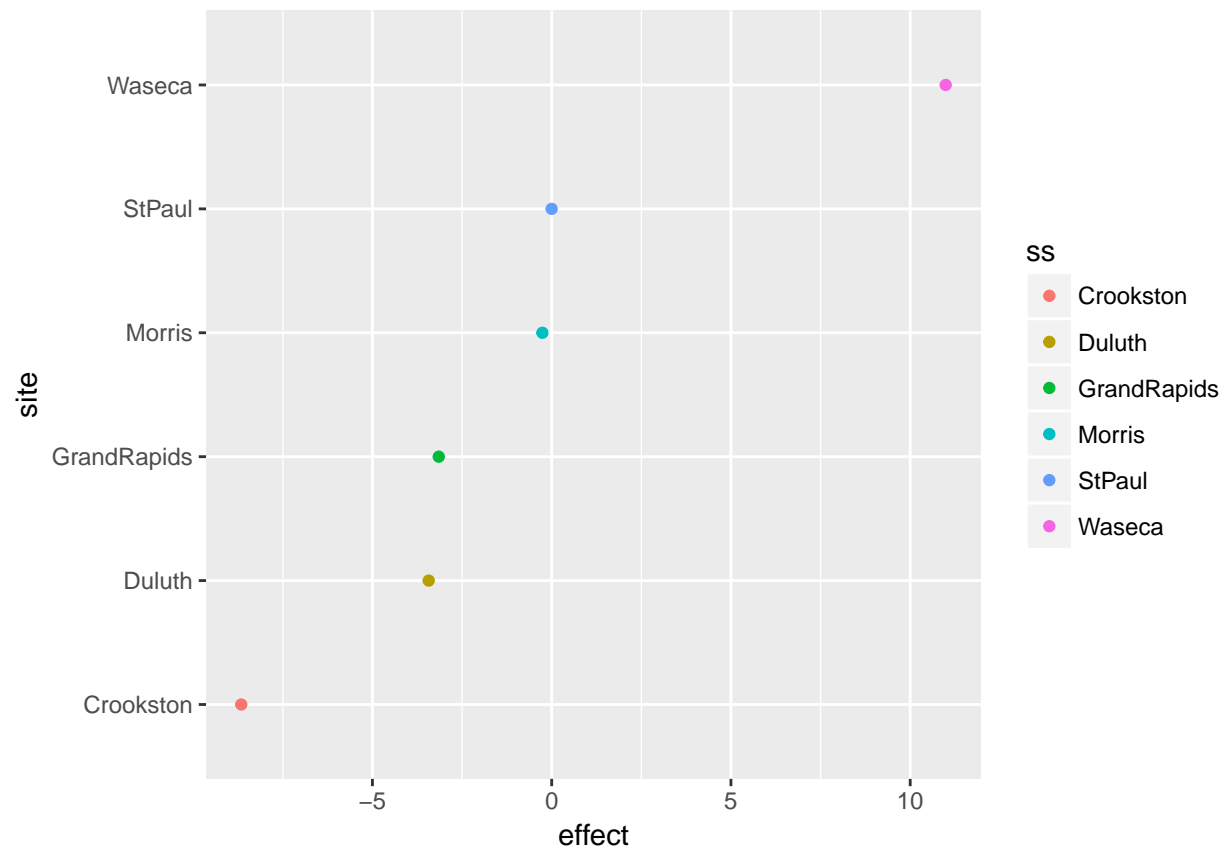
```
vts =sort(bts$gen)

vs =factor(names(bts$gen),levels =names(bts$gen))
vdf =data.frame(effect =vts,variety =vs)
ggplot(data=vdf,aes(x =effect,y =variety,color=vs)) +geom_point()
```

Q3. solution

```
brlmdf =augment(bydsrlm)
```

```
## Warning: Deprecated: please use `purrr::possibly()` instead

## Warning: Deprecated: please use `purrr::possibly()` instead

## Warning: Deprecated: please use `purrr::possibly()` instead

## Warning: Deprecated: please use `purrr::possibly()` instead

## Warning: Deprecated: please use `purrr::possibly()` instead
```
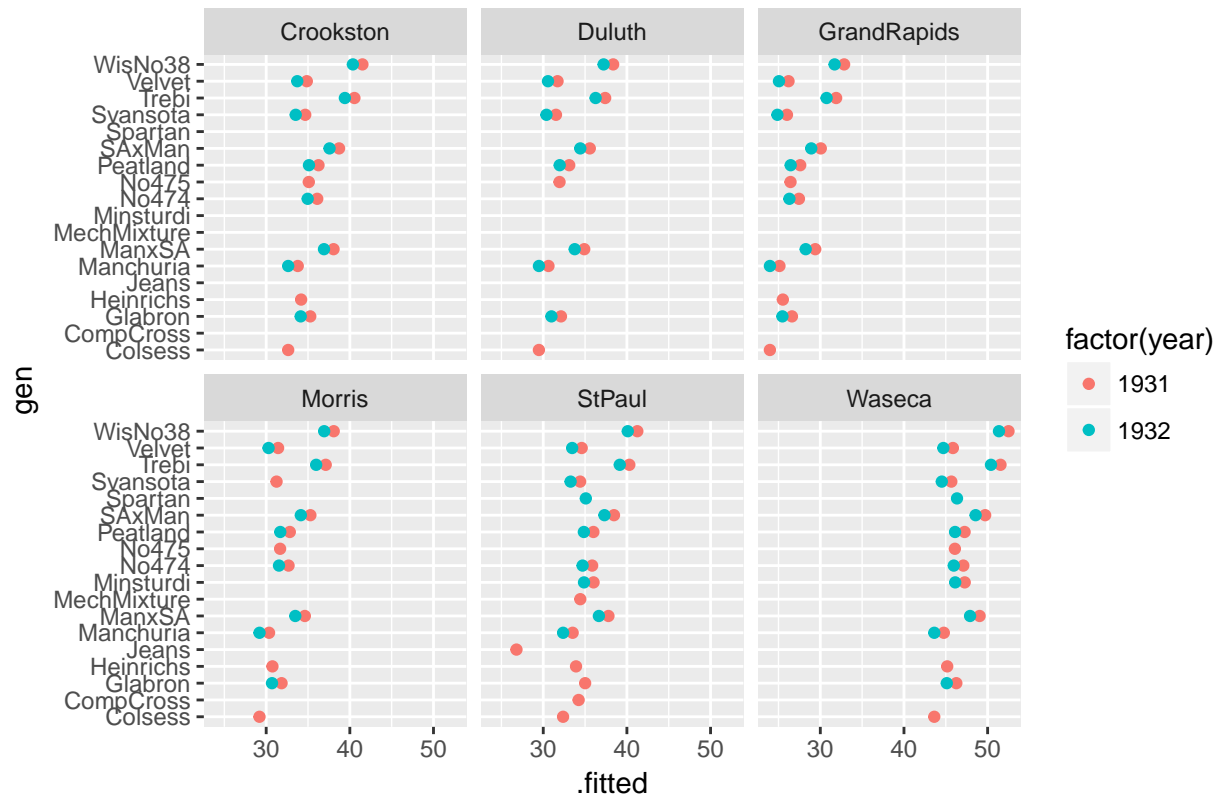
```
baly <- subset(brlmdf, year %in% c(1931,1932))
baly$year = factor(baly$year)

ggplot(baly, aes(y=gen,x=.fitted,color=factor(year))) +
  geom_point() + facet_wrap(~ site) + ggtitle(" Plot of 1931 to 1932 year")
```
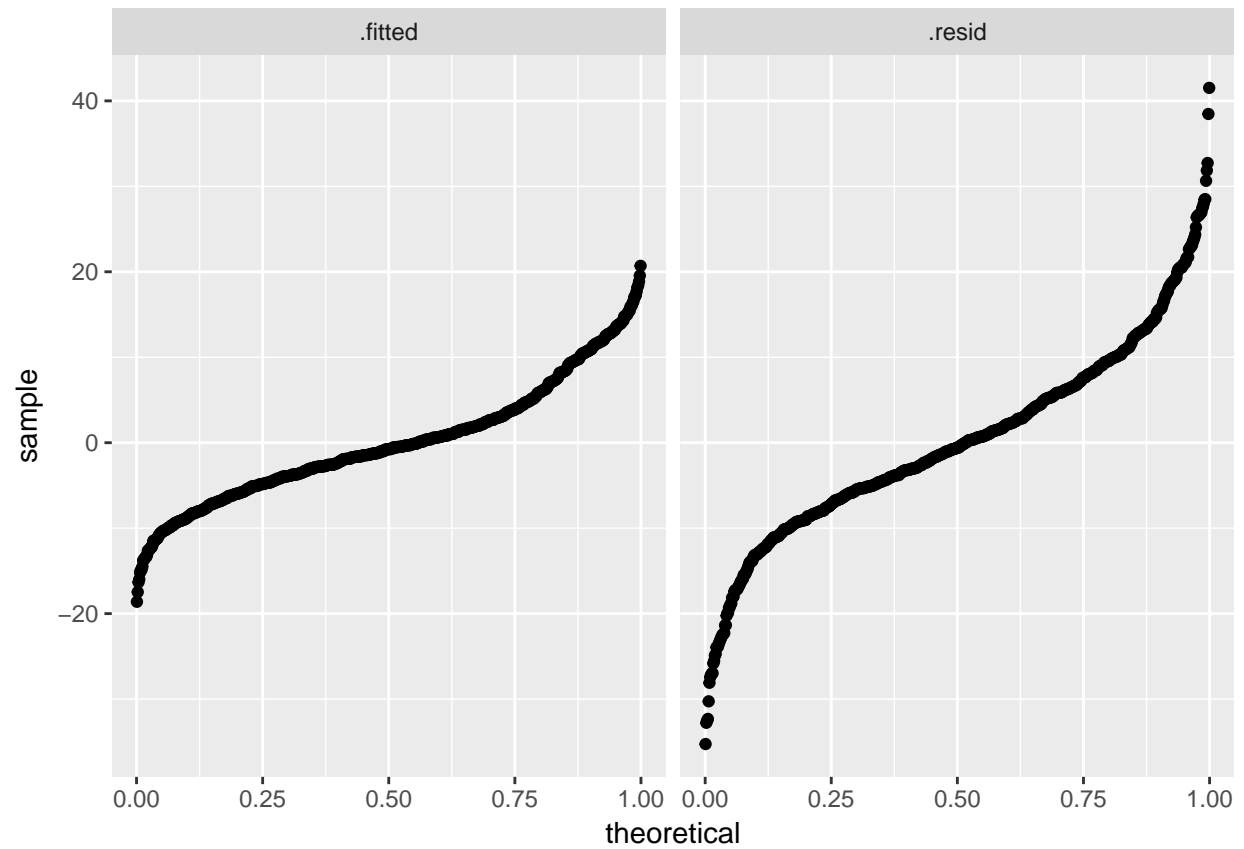
Plot of 1931 to 1932 year


```
brlmdf$.fitted = brlmdf$.fitted -mean(brlmdf$.fitted)
brlml = brlmdf %>%gather(component, value,c(.fitted, .resid))
ggplot(brlml,aes(sample =value)) +stat_qq(distribution ="qunif") +facet_grid(~component)
```

if we see the plot its morries islook like equal and solution which i gave problem 1.2 condition is also so based on this two fact we say there no mistaken data