river iron

Chris Parrish

January 20, 2016

river iron

reference:
- Cannon, et al., chapter 6, ex. 6.2-6.3, 6.5
- Cannon, et al., Student R Manual, chapter 6
- St. Regis River, NY (Wikipedia)

Import the river data on an array of chemicals.

```
data <- read.table("RiverElements.csv", header=TRUE, sep=",")
head(data[ , c(1, 2, 11)], 4)
```

```
##         River Site  Fe
## 1      Grasse    1 944
## 2      Grasse    2 525
## 3      Grasse    3 327
## 4 Oswegatchie    1 860
```

```
dim(data)
```

```
## [1] 12 27
```

Import the river data on iron. This data is in "long" format.

```
data <- read.table("RiverIron.csv", header=TRUE, sep=",")
head(data, 4)
```

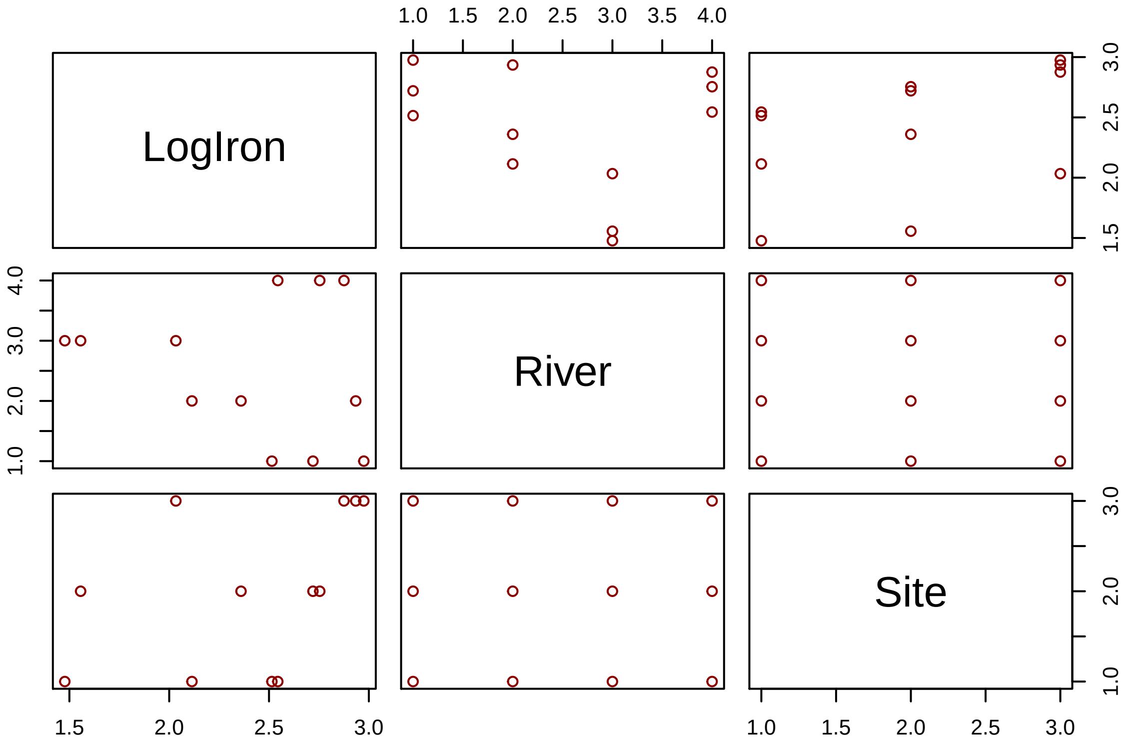
```
##         River       Site Iron LogIron
## 1      Grasse   Upstream  944  2.9750
## 2      Grasse  MidStream  525  2.7202
## 3      Grasse DownStream  327  2.5145
## 4 Oswegatchie   Upstream  860  2.9345
```

```
dim(data)
```

```
## [1] 12  4
```

Scatterplot matrix.

```
pairs(~ LogIron + River + Site, data=data,
      col="darkred")
```



Convert the data to "wide" format.

Compare with Table 6.6, page 284

```
rivers.table <- with(data,
                     tapply(LogIron, list(Site, River), mean))
rivers.table
```

```
##            Grasse Oswegatchie Raquette St. Regis
## DownStream 2.5145      2.1139   1.4771    2.5441
## MidStream  2.7202      2.3598   1.5563    2.7543
## Upstream   2.9750      2.9345   2.0334    2.8756
```

```
row.means <- with(data, tapply(LogIron, Site, mean))
col.means <- with(data, tapply(LogIron, River, mean))
rivers.table <- rbind(cbind(rivers.table, row.means),
                      c(col.means, mean(data$LogIron)))
rownames(rivers.table)[4] <- "Mean"
colnames(rivers.table)[5] <- "Mean"
rivers.table <- rivers.table[c(3, 2, 1, 4), ]  #  permute the rows
rivers.table
```

```
##              Grasse Oswegatchie Raquette St. Regis     Mean
## Upstream   2.975000      2.9345 2.033400  2.875600 2.704625
## MidStream  2.720200      2.3598 1.556300  2.754300 2.347650
## DownStream 2.514500      2.1139 1.477100  2.544100 2.162400
## Mean       2.736567      2.4694 1.688933  2.724667 2.404892
```

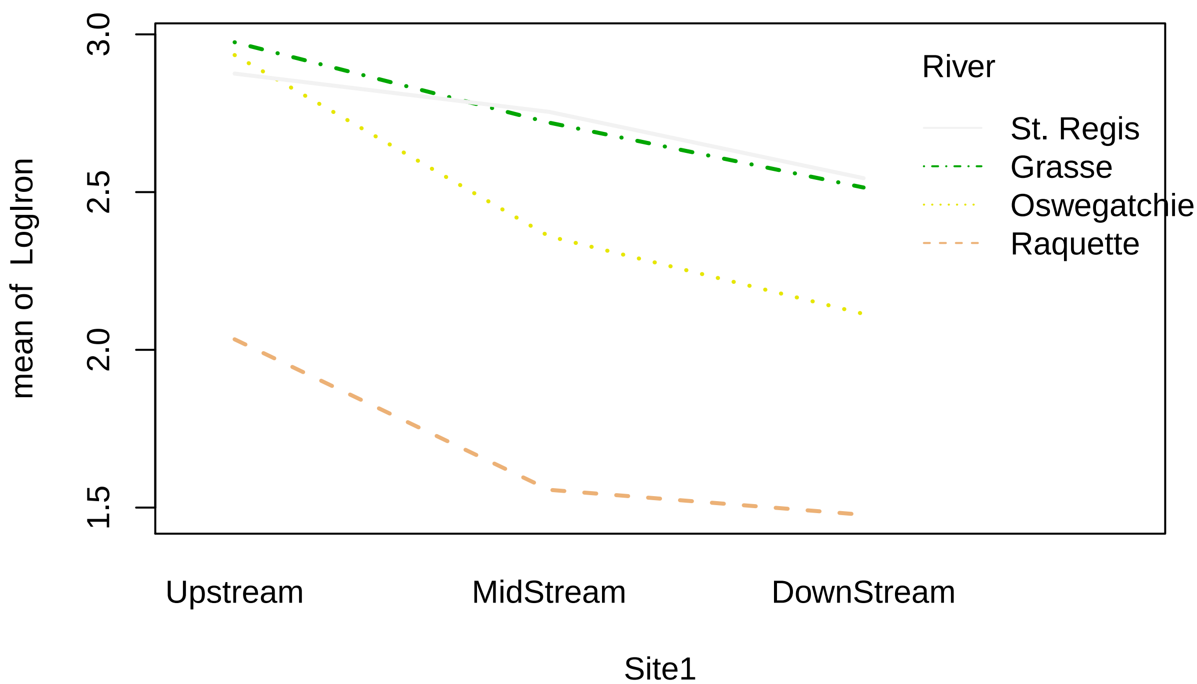
Two-way additive ANOVA.

```
rivers.aov <- aov(LogIron ~ River + Site, data=data)
options(show.signif.stars=FALSE)
summary(rivers.aov)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## River        3 2.1871  0.7290   48.15 0.000137
## Site         2 0.6077  0.3038   20.07 0.002200
## Residuals    6 0.0909  0.0151
```

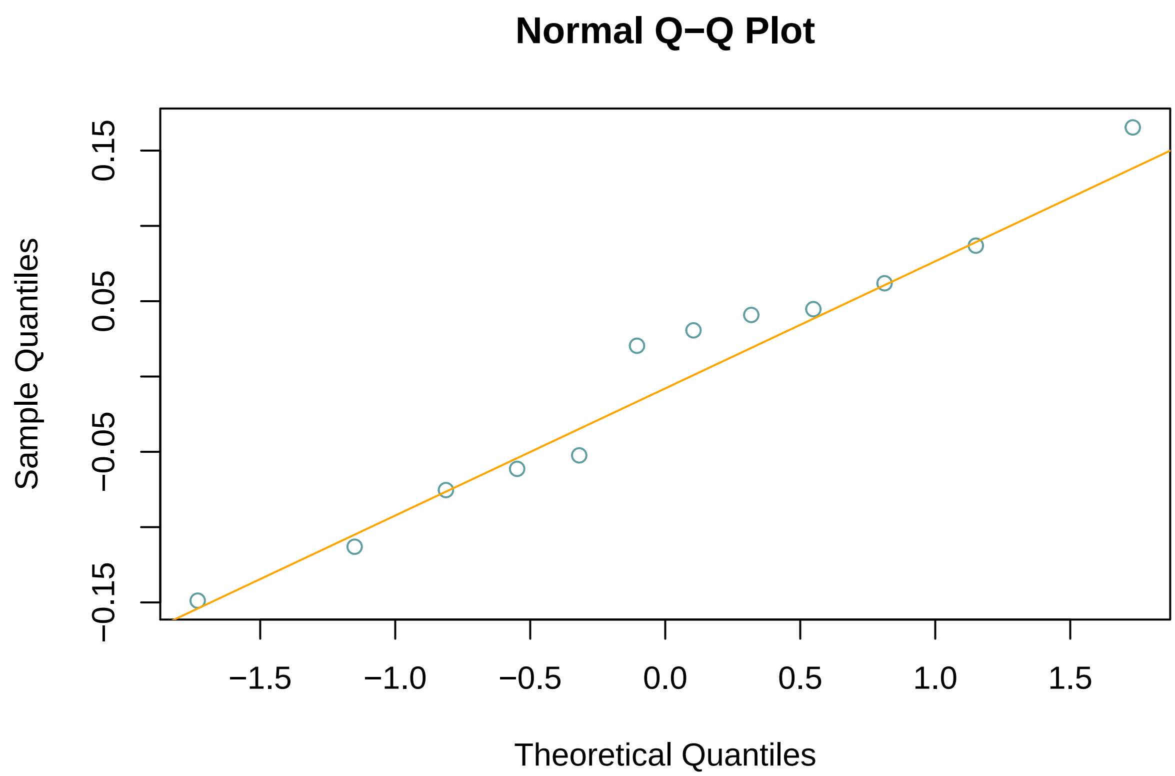
Interaction plot.

```
Site1 <- data$Site[12:1]                 # invert the order
levels(Site1) <- levels(data$Site)[3:1]  # invert the order
with(data,
     interaction.plot(Site1, River, LogIron,
                      col=terrain.colors(4), lwd=2))
```

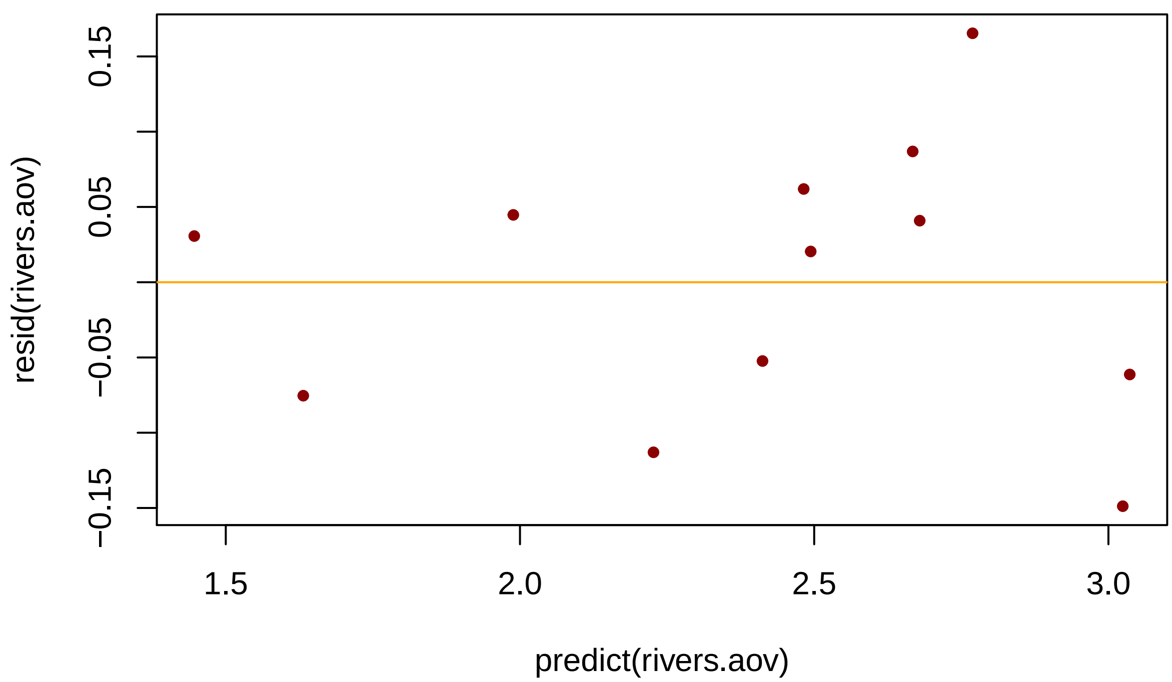


Residuals.

```
qqnorm(resid(rivers.aov), col="cadetblue")
qqline(resid(rivers.aov), col="orange")
```




```
plot(predict(rivers.aov), resid(rivers.aov),
     pch=20, col="darkred")
abline(h=0, col="orange")
```



Fisher's LSD

$$\bar{y}_i - \bar{y}_j \pm \sqrt{MSE\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

```
alpha <- 0.05
df <- 6
t.star <- qt(1 - alpha/2, df=df)
t.star
```

```
## [1] 2.446912
```

```
mse <- 0.0151  #  MSE from ANOVA
sites <- t.star * sqrt(mse * 2 / 4)
rivers <- t.star * sqrt(mse * 2 / 3)
LSD.vals <- rbind(sites, rivers)
colnames(LSD.vals) <- "LSD"
LSD.vals
```

```
##              LSD
## sites  0.2126140
## rivers 0.2455055
```

Testing for differences of means using Fisher's LSD: sites

Which pairs of sites have significantly different means?

```
sites <- rivers.table[4, 1:4]
sites
```

```
##      Grasse Oswegatchie    Raquette   St. Regis
##    2.736567    2.469400    1.688933    2.724667
```

```
differences.of.means <- outer(sites, t(sites), FUN = "-")
differences.of.means
```

```
## , , Grasse
##
##                   [,1]
## Grasse       0.0000000
## Oswegatchie -0.2671667
## Raquette    -1.0476333
## St. Regis   -0.0119000
##
## , , Oswegatchie
##
##                   [,1]
## Grasse       0.2671667
## Oswegatchie  0.0000000
## Raquette    -0.7804667
## St. Regis    0.2552667
##
```

```
## , , Raquette
##
##                   [,1]
## Grasse      1.0476333
## Oswegatchie 0.7804667
## Raquette    0.0000000
## St. Regis   1.0357333
##
## , , St. Regis
##
##                    [,1]
## Grasse       0.0119000
## Oswegatchie -0.2552667
## Raquette    -1.0357333
## St. Regis    0.0000000
```

Testing for differences of means using Fisher's LSD: rivers

Which pairs of rivers have significantly different means?

```
rivers <- rivers.table[1:3, 5]
rivers
```

```
##   Upstream  MidStream DownStream
##   2.704625   2.347650   2.162400
```

```
differences.of.means <- outer(rivers, t(rivers), FUN = "-")
differences.of.means
```

```
## , , Upstream
##
##                  [,1]
## Upstream    0.000000
## MidStream  -0.356975
## DownStream -0.542225
##
## , , MidStream
##
##                  [,1]
## Upstream    0.356975
## MidStream   0.000000
## DownStream -0.185250
##
## , , DownStream
##
##                [,1]
## Upstream   0.542225
## MidStream  0.185250
## DownStream 0.000000
```

Effect plots: without interaction.

```r
rivers.lm1 <- lm(LogIron ~ River + Site, data=data)
library(alr4)
plot(allEffects(rivers.lm1))
```