

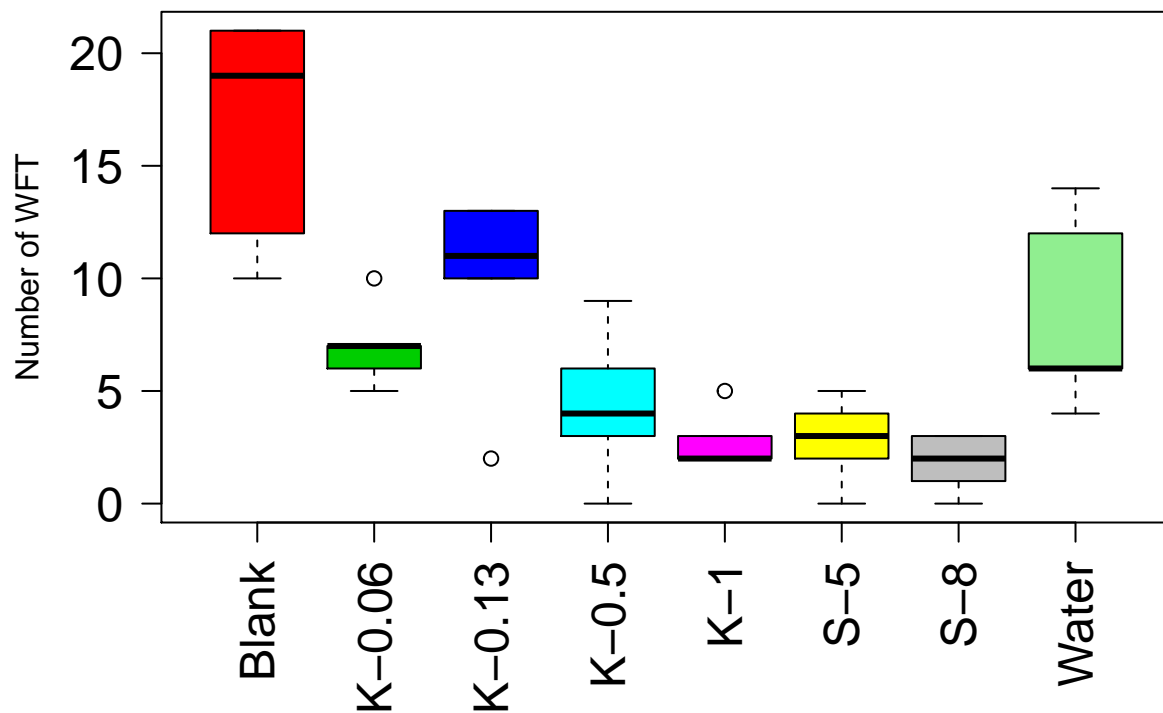
Project1

Congxing Zhu

11/20/2017

```
Tre <- rep(c("Blank", "Water", "K-1", "K-0.5", "K-0.13", "K-0.06", "S-8", "S-5"),
  each = 5)
Rep <- rep(1:5, length = 40)
Num <- rep(20, length = 40)
WFT <- c(12, 19, 21, 21, 10, 14, 4, 12, 6, 6, 2, 2, 3, 2, 5, 0, 6, 3, 9, 4,
  13, 2, 10, 11, 13, 7, 5, 10, 6, 7, 3, 0, 1, 2, 3, 4, 3, 5, 0, 2)
Data1 <- data.frame(Tre = as.factor(Tre), Rep, Num, WFT)
Data1$Num[Data1$WFT > 20] <- 21
```

```
plot(WFT ~ Tre, data = Data1, col = c(2:8, "lightgreen"), cex.axis = 1.5, las = 2,
  xlab = "", ylab = "Number of WFT")
```



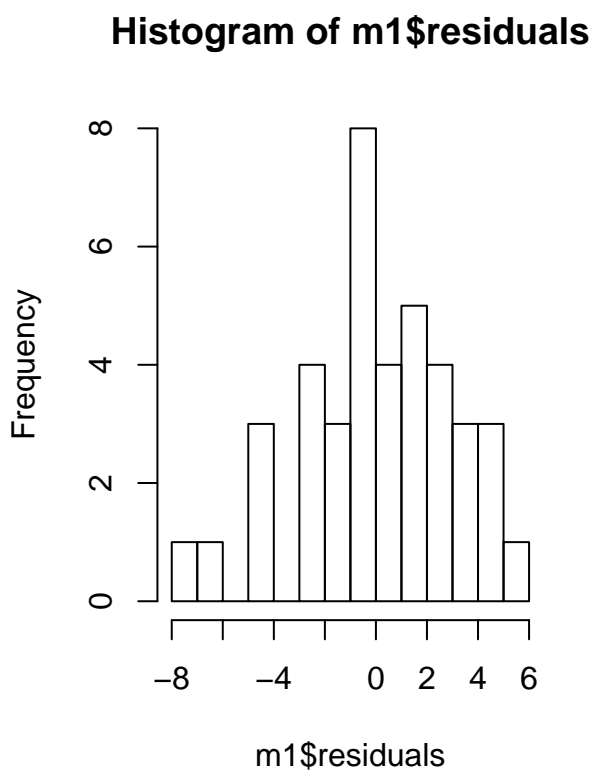
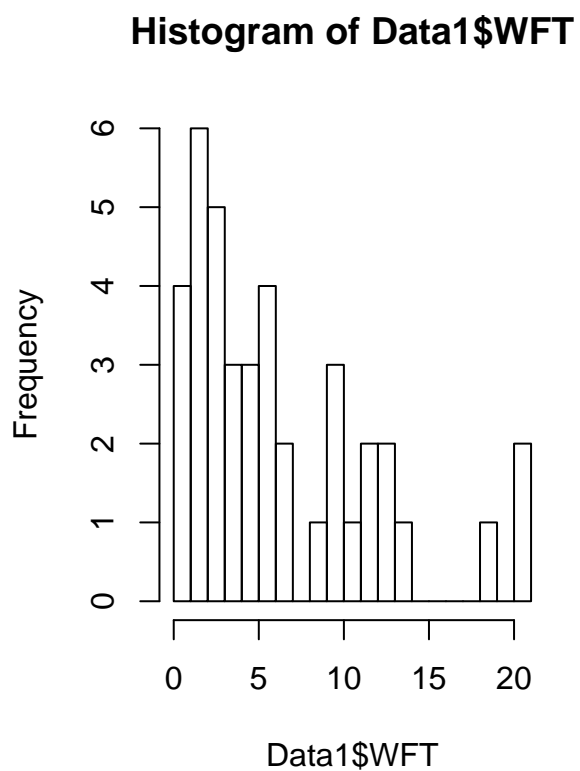
```
tapply(WFT, Tre, mean, data = Data1) # mean of the treatments
```

```
## Blank K-0.06 K-0.13 K-0.5 K-1 S-5 S-8 Water
## 16.6 7.0 9.8 4.4 2.8 2.8 1.8 8.4
```

```
m1 <- lm(WFT ~ Tre - 1, data = Data1) #fit the data into simple linear model
summary(m1)
```

```
##
```

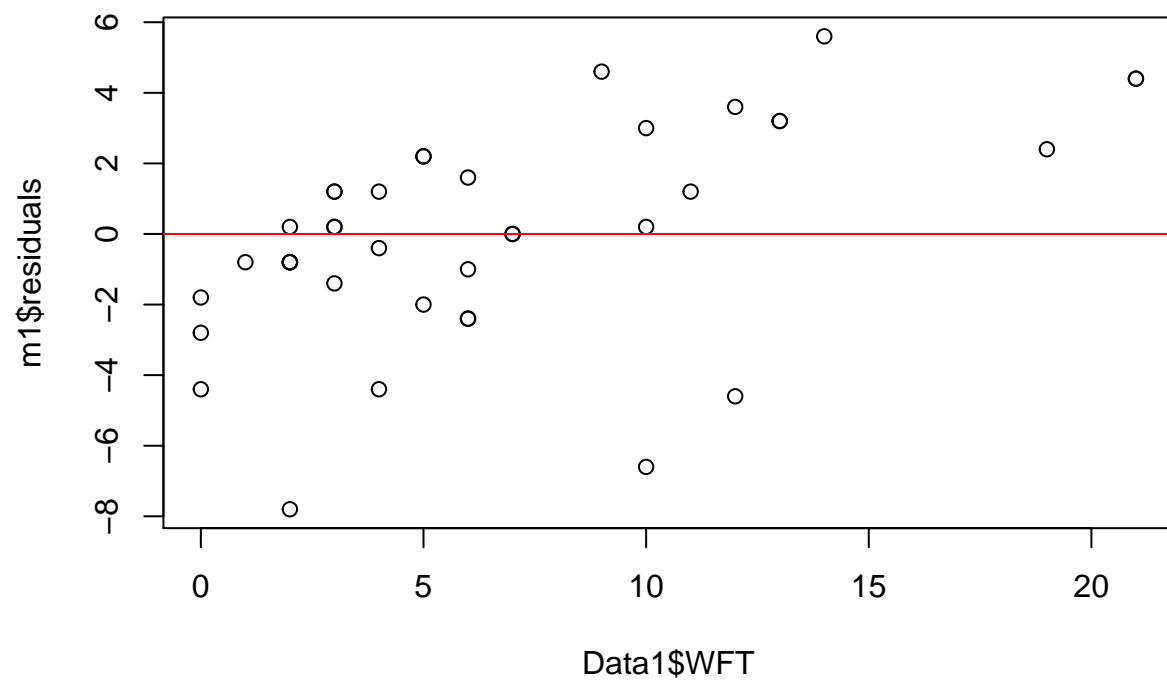
```
## Call:
## lm(formula = WFT ~ Tre - 1, data = Data1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
##    -7.8    -1.5     0.1     2.2     5.6
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## TreBlank       16.600      1.489  11.147 1.49e-12 ***
## TreK-0.06       7.000      1.489   4.701 4.73e-05 ***
## TreK-0.13       9.800      1.489   6.581 2.04e-07 ***
## TreK-0.5        4.400      1.489   2.955 0.00583 **
## TreK-1          2.800      1.489   1.880 0.06920 .
## TreS-5          2.800      1.489   1.880 0.06920 .
## TreS-8          1.800      1.489   1.209 0.23561
## TreWater        8.400      1.489   5.641 3.08e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.33 on 32 degrees of freedom
## Multiple R-squared:  0.8818, Adjusted R-squared:  0.8523
## F-statistic: 29.84 on 8 and 32 DF,  p-value: 9.861e-13
par(mfrow = c(1, 2))
hist(Data1$WFT, breaks = 15) #Checking the assumption of normality
hist(m1$residuals, breaks = 15)
```



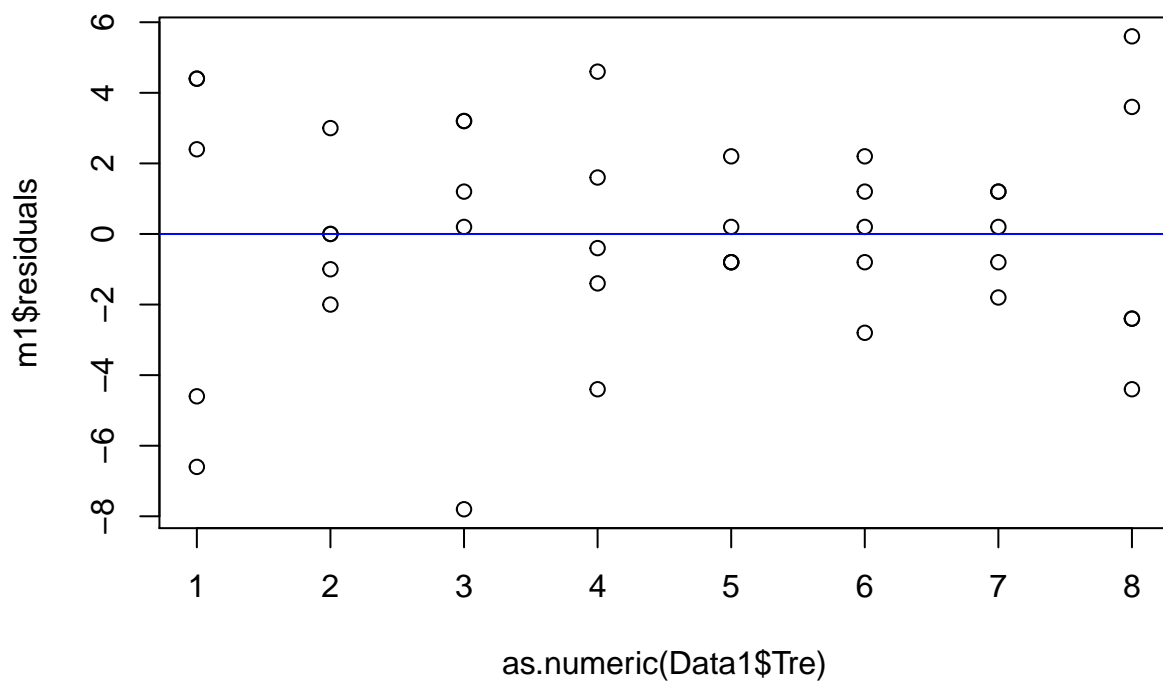
```
shapiro.test(m1$residuals)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  m1$residuals
## W = 0.97475, p-value = 0.5015
```

```
plot(Data1$WFT, m1$residuals) #Checking the assumption of constant variance
abline(h = 0, col = "red")
```



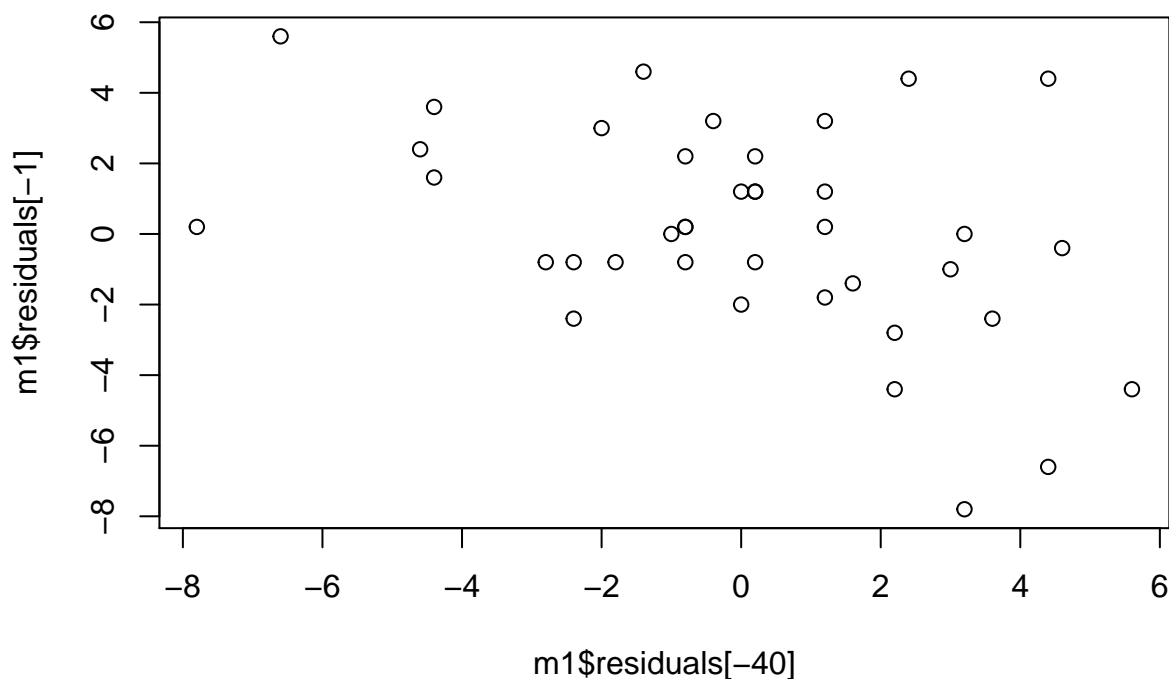
```
plot(as.numeric(Data1$Tre), m1$residuals)
abline(h = 0, col = "blue")
```



```
plot(m1$residuals[-40], m1$residuals[-1]) #Checking the assumption of independence
cor(m1$residuals[-40], m1$residuals[-1])
```

```
## [1] -0.4439954
```

```
library(lmtest)
```



```
dwtest(m1)
```

```
##
## Durbin-Watson test
##
## data: m1
## DW = 2.7976, p-value = 0.941
## alternative hypothesis: true autocorrelation is greater than 0
```

```
anova(m1) #anova to check whether some of the coefficients are not zero
```

```
## Analysis of Variance Table
```

```
##
```

```
## Response: WFT
```

```
##      Df Sum Sq Mean Sq F value    Pr(>F)
## Tre      8 2647.2   330.90  29.844 9.861e-13 ***
## Residuals 32  354.8    11.09
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Differ <- TukeyHSD(x = aov(Data1$WFT ~ Data1$Tre), "Data1$Tre", conf.level = 0.95) #Tukey test to comp
Differ
```

```
## Tukey multiple comparisons of means
```

```
## 95% family-wise confidence level
```

```
##
```

```
## Fit: aov(formula = Data1$WFT ~ Data1$Tre)
```

```
##
```

```
## $`Data1$Tre`
##               diff               lwr               upr               p adj
## K-0.06-Blank -9.600000e+00 -16.4217809 -2.77821906 0.0016403
## K-0.13-Blank -6.800000e+00 -13.6217809  0.02178094 0.0512141
## K-0.5-Blank  -1.220000e+01 -19.0217809 -5.37821906 0.0000497
## K-1-Blank    -1.380000e+01 -20.6217809 -6.97821906 0.0000057
## S-5-Blank    -1.380000e+01 -20.6217809 -6.97821906 0.0000057
## S-8-Blank    -1.480000e+01 -21.6217809 -7.97821906 0.0000015
## Water-Blank  -8.200000e+00 -15.0217809 -1.37821906 0.0098726
## K-0.13-K-0.06 2.800000e+00  -4.0217809  9.62178094 0.8806731
## K-0.5-K-0.06 -2.600000e+00  -9.4217809  4.22178094 0.9152039
## K-1-K-0.06   -4.200000e+00 -11.0217809  2.62178094 0.5010193
## S-5-K-0.06   -4.200000e+00 -11.0217809  2.62178094 0.5010193
## S-8-K-0.06   -5.200000e+00 -12.0217809  1.62178094 0.2445883
## Water-K-0.06  1.400000e+00  -5.4217809  8.22178094 0.9973914
## K-0.5-K-0.13 -5.400000e+00 -12.2217809  1.42178094 0.2062402
## K-1-K-0.13   -7.000000e+00 -13.8217809 -0.17821906 0.0409971
## S-5-K-0.13   -7.000000e+00 -13.8217809 -0.17821906 0.0409971
## S-8-K-0.13   -8.000000e+00 -14.8217809 -1.17821906 0.0126302
## Water-K-0.13 -1.400000e+00  -8.2217809  5.42178094 0.9973914
## K-1-K-0.5    -1.600000e+00  -8.4217809  5.22178094 0.9940651
## S-5-K-0.5    -1.600000e+00  -8.4217809  5.22178094 0.9940651
## S-8-K-0.5    -2.600000e+00  -9.4217809  4.22178094 0.9152039
## Water-K-0.5   4.000000e+00  -2.8217809 10.82178094 0.5611598
## S-5-K-1       7.549517e-15  -6.8217809  6.82178094 1.0000000
## S-8-K-1      -1.000000e+00  -7.8217809  5.82178094 0.9997024
## Water-K-1     5.600000e+00  -1.2217809 12.42178094 0.1725205
## S-8-S-5      -1.000000e+00  -7.8217809  5.82178094 0.9997024
## Water-S-5     5.600000e+00  -1.2217809 12.42178094 0.1725205
## Water-S-8     6.600000e+00  -0.2217809 13.42178094 0.0636517
```

```
library(multcomp)
```

```
cld(summary(glht(m1, linfct = mcp(Tre = "Tukey"))), decreasing = T) #Tukey test to attribute them to d
```

```
## Blank K-0.06 K-0.13 K-0.5 K-1 S-5 S-8 Water
## "a" "bc" "ab" "bc" "c" "c" "c" "bc"
```

```
library(agricolae)
```

```
model <- aov(Data1$WFT ~ Data1$Tre, data = Data1)
```

```
df <- df.residual(model)
```

```
MSerror <- deviance(model)/df
```

```
with(Data1, HSD.test(WFT, Tre, df, MSerror, group = TRUE, console = TRUE, main = "Difference between th
```

```
##
```

```
## Study: Difference between the treatments
```

```
##
```

```
## HSD Test for WFT
```

```
##
```

```
## Mean Square Error: 11.0875
```

```
##
```

```
## Tre, means
```

```
##
```

```
## WFT std r Min Max
```

```
## Blank 16.6 5.224940 5 10 21
```

```
## K-0.06 7.0 1.870829 5 5 10
```

```

## K-0.13  9.8 4.549725 5  2  13
## K-0.5   4.4 3.361547 5  0  9
## K-1     2.8 1.303840 5  2  5
## S-5     2.8 1.923538 5  0  5
## S-8     1.8 1.303840 5  0  3
## Water   8.4 4.335897 5  4  14
##
## Alpha: 0.05 ; DF Error: 32
## Critical Value of Studentized Range: 4.58106
##
## Minimum Significant Difference: 6.821781
##
## Treatments with the same letter are not significantly different.
##
##      WFT groups
## Blank  16.6      a
## K-0.13  9.8      ab
## Water   8.4      bc
## K-0.06  7.0      bc
## K-0.5   4.4      bc
## K-1     2.8      c
## S-5     2.8      c
## S-8     1.8      c

```

```

m2 <- glm(cbind(WFT, Num - WFT) ~ Tre - 1, family = quasibinomial(link = "logit"),
  data = Data1) #Fit the data into a quasibinomial model(overdispersion)
summary(m2)

```

```

##
## Call:
## glm(formula = cbind(WFT, Num - WFT) ~ Tre - 1, family = quasibinomial(link = "logit"),
##      data = Data1)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -3.7537  -0.8334   0.0447   1.3009   2.9424
##
## Coefficients:
##      Estimate Std. Error t value Pr(>|t|)
## TreBlank     1.47440    0.43228   3.411 0.001771 **
## TreK-0.06    -0.61904    0.35636  -1.737 0.091982 .
## TreK-0.13    -0.04001    0.34001  -0.118 0.907073
## TreK-0.5     -1.26567    0.41032  -3.085 0.004180 **
## TreK-1       -1.81529    0.48985  -3.706 0.000794 ***
## TreS-5       -1.81529    0.48985  -3.706 0.000794 ***
## TreS-8       -2.31363    0.59393  -3.895 0.000469 ***
## TreWater     -0.32277    0.34438  -0.937 0.355645
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasibinomial family taken to be 2.889032)
##
##      Null deviance: 388.00  on 40  degrees of freedom
## Residual deviance: 106.28  on 32  degrees of freedom
## AIC: NA

```



```
##
## Number of Fisher Scoring iterations: 5
library(multcomp)
Differ2 <- summary(glht(m2, linfct = mcp(Tre = "Tukey"))) #Tukey test
Differ2

##
## Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: glm(formula = cbind(WFT, Num - WFT) ~ Tre - 1, family = quasibinomial(link = "logit"),
## data = Data1)
##
## Linear Hypotheses:
## Estimate Std. Error z value Pr(>|z|)
## K-0.06 - Blank == 0 -2.0934 0.5602 -3.737 0.00448 **
## K-0.13 - Blank == 0 -1.5144 0.5500 -2.754 0.10379
## K-0.5 - Blank == 0 -2.7401 0.5960 -4.597 < 0.001 ***
## K-1 - Blank == 0 -3.2897 0.6533 -5.035 < 0.001 ***
## S-5 - Blank == 0 -3.2897 0.6533 -5.035 < 0.001 ***
## S-8 - Blank == 0 -3.7880 0.7346 -5.157 < 0.001 ***
## Water - Blank == 0 -1.7972 0.5527 -3.252 0.02422 *
## K-0.13 - K-0.06 == 0 0.5790 0.4925 1.176 0.93697
## K-0.5 - K-0.06 == 0 -0.6466 0.5435 -1.190 0.93300
## K-1 - K-0.06 == 0 -1.1963 0.6058 -1.975 0.49127
## S-5 - K-0.06 == 0 -1.1963 0.6058 -1.975 0.49149
## S-8 - K-0.06 == 0 -1.6946 0.6926 -2.447 0.21368
## Water - K-0.06 == 0 0.2963 0.4956 0.598 0.99887
## K-0.5 - K-0.13 == 0 -1.2257 0.5329 -2.300 0.28698
## K-1 - K-0.13 == 0 -1.7753 0.5963 -2.977 0.05698 .
## S-5 - K-0.13 == 0 -1.7753 0.5963 -2.977 0.05656 .
## S-8 - K-0.13 == 0 -2.2736 0.6844 -3.322 0.01929 *
## Water - K-0.13 == 0 -0.2828 0.4839 -0.584 0.99902
## K-1 - K-0.5 == 0 -0.5496 0.6390 -0.860 0.98899
## S-5 - K-0.5 == 0 -0.5496 0.6390 -0.860 0.98899
## S-8 - K-0.5 == 0 -1.0480 0.7219 -1.452 0.82838
## Water - K-0.5 == 0 0.9429 0.5357 1.760 0.64018
## S-5 - K-1 == 0 0.0000 0.6928 0.000 1.00000
## S-8 - K-1 == 0 -0.4983 0.7699 -0.647 0.99811
## Water - K-1 == 0 1.4925 0.5988 2.493 0.19265
## S-8 - S-5 == 0 -0.4983 0.7699 -0.647 0.99811
## Water - S-5 == 0 1.4925 0.5988 2.493 0.19331
## Water - S-8 == 0 1.9909 0.6865 2.900 0.07028 .
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)
cld(Differ2, decreasing = T)

## Blank K-0.06 K-0.13 K-0.5 K-1 S-5 S-8 Water
## "a" "bc" "ab" "bc" "bc" "bc" "c" "bc"
```

```

library(gamlss)
m3 <- gamlss(cbind(WFT, Num - WFT) ~ Tre - 1, data = Data1, family = BB) #Fit the data into a beta bin

## GAMLSS-RS iteration 1: Global Deviance = 203.4452
## GAMLSS-RS iteration 2: Global Deviance = 189.7678
## GAMLSS-RS iteration 3: Global Deviance = 189.1909
## GAMLSS-RS iteration 4: Global Deviance = 189.1859
## GAMLSS-RS iteration 5: Global Deviance = 189.1859

summary(m3)

## *****
## Family: c("BB", "Beta Binomial")
##
## Call: gamlss(formula = cbind(WFT, Num - WFT) ~ Tre - 1, family = BB,
## data = Data1)
##
## Fitting method: RS()
##
## -----
## Mu link function: logit
## Mu Coefficients:
##      Estimate Std. Error t value Pr(>|t|)
## TreBlank      1.82749    0.48898   3.737 0.000754 ***
## TreK-0.06     -0.57443    0.32644  -1.760 0.088324 .
## TreK-0.13     -0.07011    0.32238  -0.217 0.829250
## TreK-0.5      -1.30862    0.38950  -3.360 0.002082 **
## TreK-1        -1.62913    0.40392  -4.033 0.000333 ***
## TreS-5        -1.74253    0.43050  -4.048 0.000320 ***
## TreS-8        -2.12554    0.47706  -4.455 0.000102 ***
## TreWater     -0.32218    0.32378  -0.995 0.327412
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## -----
## Sigma link function: log
## Sigma Coefficients:
##      Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -2.4170     0.3837  -6.299 5.26e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## -----
## No. of observations in the fit: 40
## Degrees of Freedom for the fit: 9
##      Residual Deg. of Freedom: 31
##                               at cycle: 5
##
## Global Deviance:      189.1859
## AIC:                  207.1859
## SBC:                  222.3858
## *****

library(mgcv)
m4 <- gam(list(WFT ~ Tre - 1, ~Tre), family = gaulss(link = list("identity",

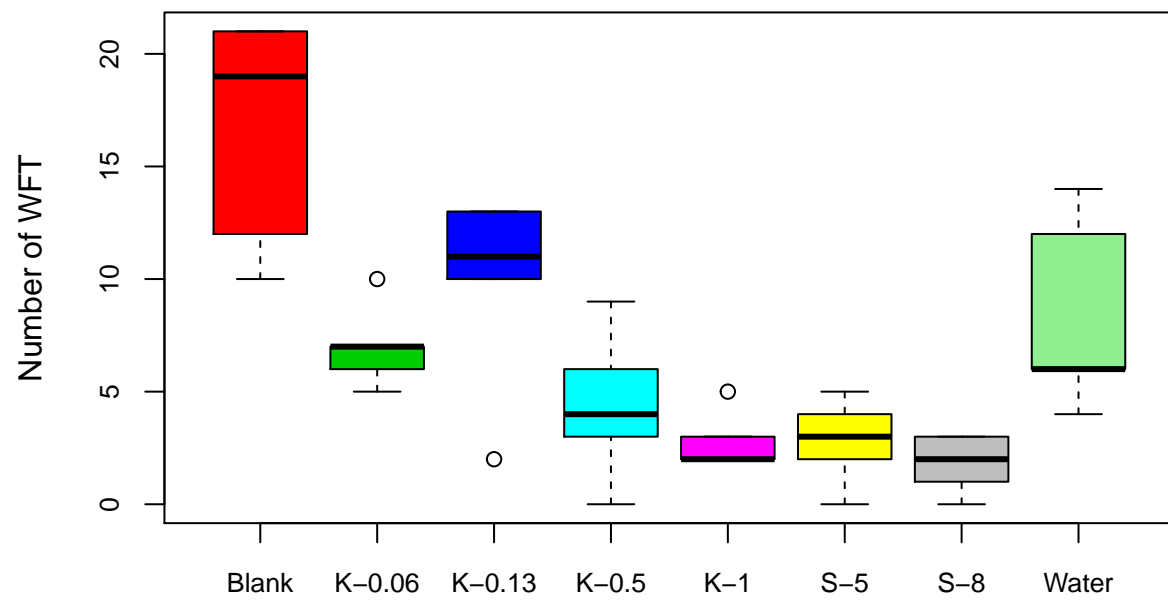
```

```

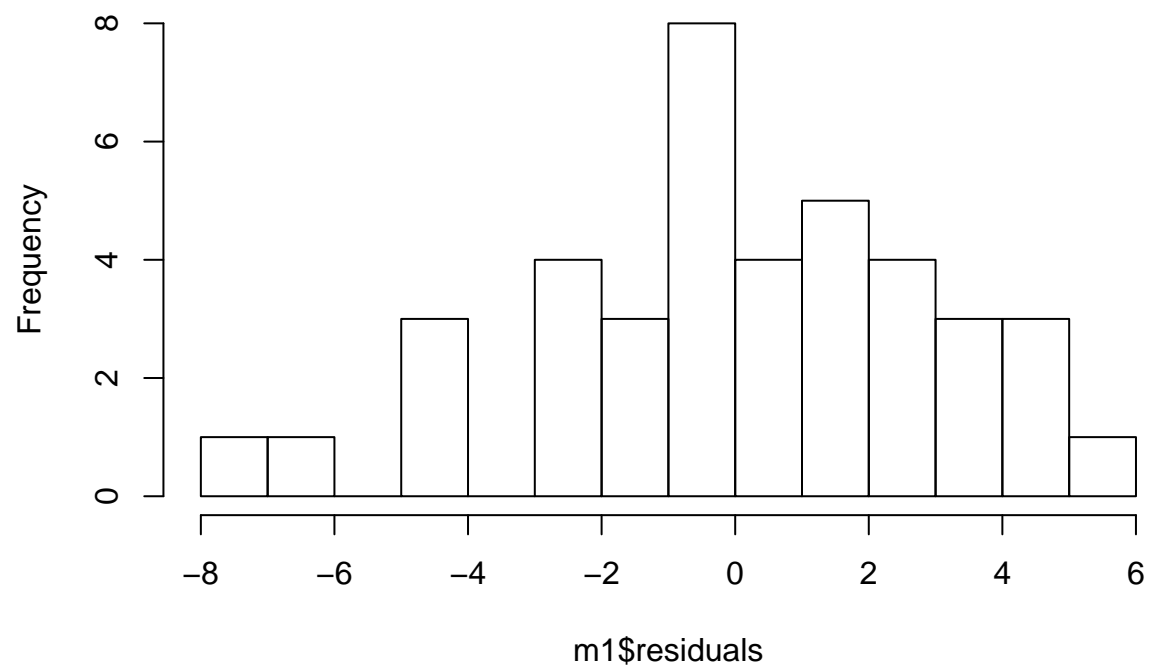
    "logb")), data = Data1) #Another method to deal with outlier talked in class(change the model)
summary(m4)

##
## Family: gaulss
## Link function: identity logb
##
## Formula:
## WFT ~ Tre - 1
## ~Tre
##
## Parametric coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## TreBlank      16.6000    2.0900   7.943 1.98e-15 ***
## TreK-0.06      7.0000    0.7483   9.354 < 2e-16 ***
## TreK-0.13      9.8000    1.8199   5.385 7.25e-08 ***
## TreK-0.5       4.4000    1.3446   3.272 0.001067 **
## TreK-1         2.8000    0.5215   5.369 7.93e-08 ***
## TreS-5         2.8000    0.7694   3.639 0.000274 ***
## TreS-8         1.8000    0.5215   3.451 0.000558 ***
## TreWater       8.4000    1.7344   4.843 1.28e-06 ***
## (Intercept).1  1.5397    0.3169   4.859 1.18e-06 ***
## TreK-0.06.1    -1.0309    0.4490  -2.296 0.021686 *
## TreK-0.13.1    -0.1387    0.4482  -0.309 0.757003
## TreK-0.5.1     -0.4422    0.4484  -0.986 0.324056
## TreK-1.1       -1.3946    0.4496  -3.102 0.001924 **
## TreS-5.1       -1.0030    0.4490  -2.234 0.025498 *
## TreS-8.1       -1.3946    0.4496  -3.102 0.001924 **
## TreWater.1     -0.1870    0.4483  -0.417 0.676637
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Deviance explained = 83.3%
## -REML = 84.729 Scale est. = 1          n = 40
plot(WFT ~ Tre, data = Data1, col = c(2:8, "lightgreen"), cex.axis = 0.8, xlab = "",
     ylab = "Number of WFT") #for project figure

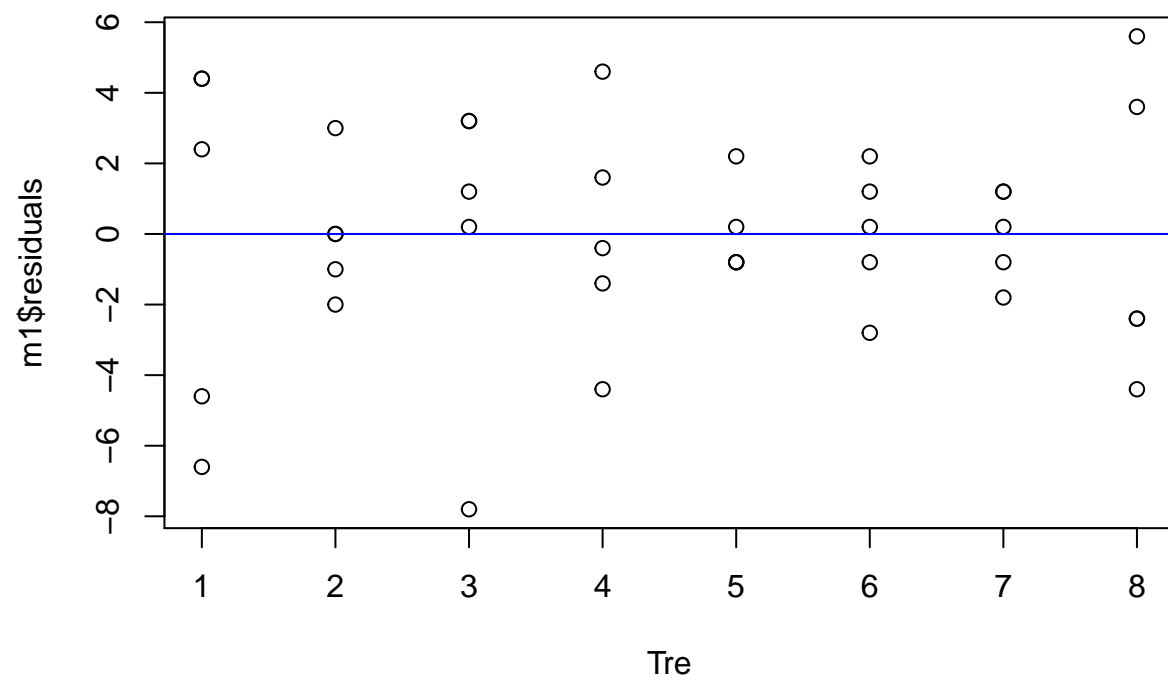
```



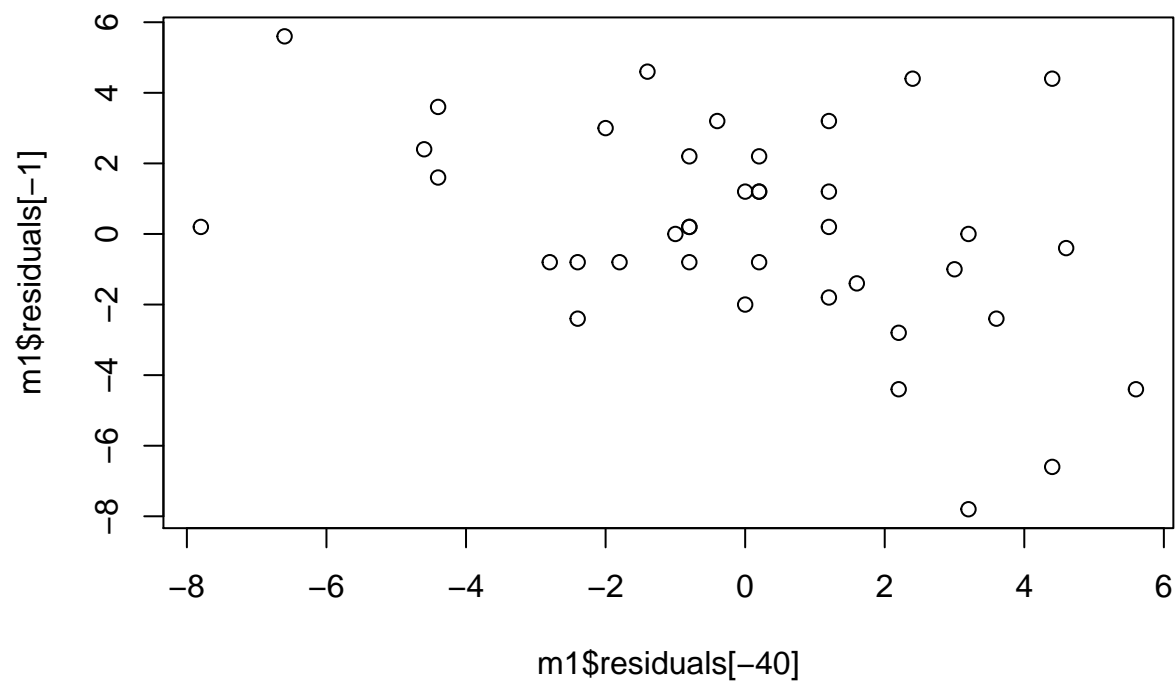
```
hist(m1$residuals, breaks = 15, main = "") #for project figure
```



```
plot(as.numeric(Data1$Tre), m1$residuals, xlab = "Tre") #for project figure  
abline(h = 0, col = "blue") #for project figure
```



```
plot(m1$residuals[-40], m1$residuals[-1]) #for project figure
```



```
plot(WFT ~ as.numeric(Tre), data = Data1, xlab = "", ylab = "Number of WFT",
      ylim = c(-2, 25)) #for project figure
arrows(2, 0, 2.9, 1.9, col = "blue") #for project figure
```

