Section 5.8 Case Study: Words with Friends

Loaded needed packages.

```
library(Stat2Data)
library(mosaic)
```

EXAMPLE 5.25 The value of a blank tile

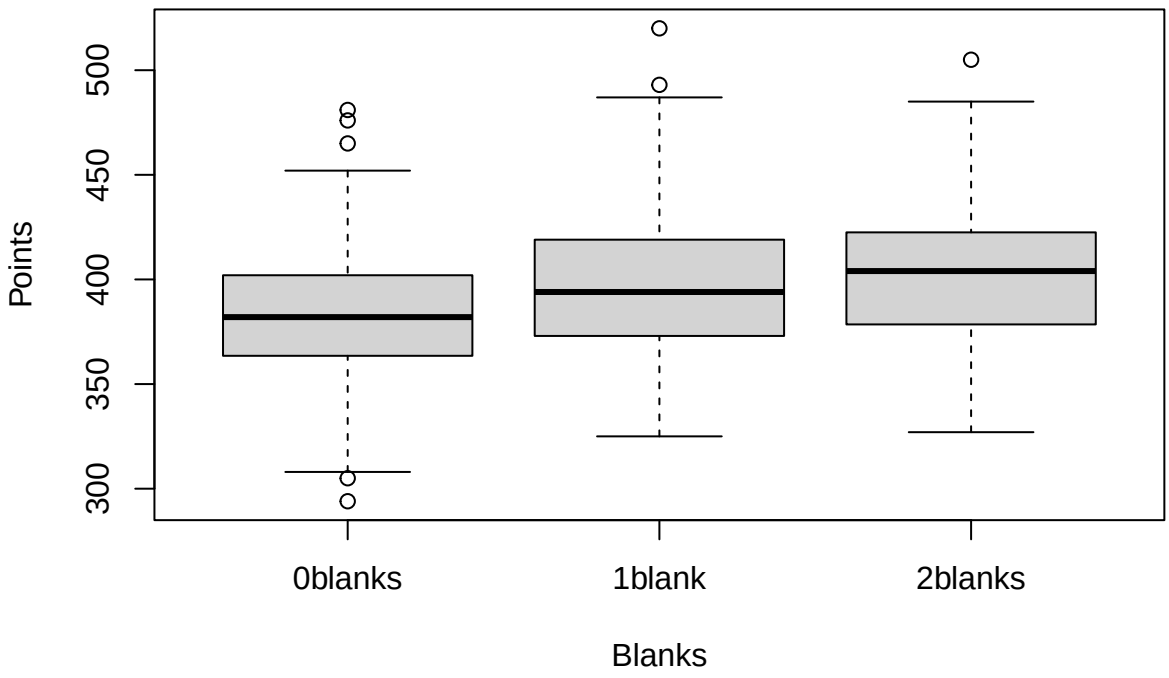
Create a dataframe for **WordsWithFriends** and look at the structure of the data.

```
data("WordsWithFriends")
str(WordsWithFriends)
```

```
## 'data.frame':    444 obs. of  11 variables:
##  $ Points      : int  357 386 434 418 411 373 367 294 383 386 ...
##  $ OppPoints   : int  302 311 364 357 356 261 343 254 331 327 ...
##  $ WinMargin   : int  55 75 70 61 55 112 24 40 52 59 ...
##  $ Start       : Factor w/ 2 levels "first","pass": 1 1 2 1 1 1 1 2 2 1 ...
##  $ Ss          : int  4 2 2 1 3 1 2 5 0 3 ...
##  $ BlanksNumber: int  0 0 0 1 1 2 0 0 1 0 ...
##  $ J           : int  1 0 1 1 0 0 0 0 0 0 ...
##  $ Q           : int  0 0 0 0 1 0 1 1 0 0 ...
##  $ X           : int  1 0 0 1 0 1 1 0 1 1 ...
##  $ Z           : int  1 0 0 1 1 0 1 0 1 0 ...
##  $ Blanks      : Factor w/ 3 levels "0blanks","1blank",..: 1 1 1 2 2 3 1 1 2 1 ...
```

FIGURE 5.28 Parallel boxplots of the words with friends data

```
boxplot(Points~Blanks,data=WordsWithFriends,ylab="Points")
```



Compute summary statistics for Points.

```
favstats(~Points, data=WordsWithFriends)
```

```
##  min  Q1 median     Q3 max     mean       sd   n missing
##  294 371    393 418.25 520 395.5946 34.72605 444       0
```

Compute summary statistics for the groups.

```
WordsStats=favstats(Points~Blanks,data=WordsWithFriends)[c("Blanks","n","mean","sd")]
WordsStats
```

```
##    Blanks   n     mean       sd
## 1 0blanks 111 384.7117 33.46488
## 2  1blank 217 396.9355 34.34487
## 3 2blanks 116 403.5000 34.31960
```

Group standard deviations are very close! No issues with equal variances.

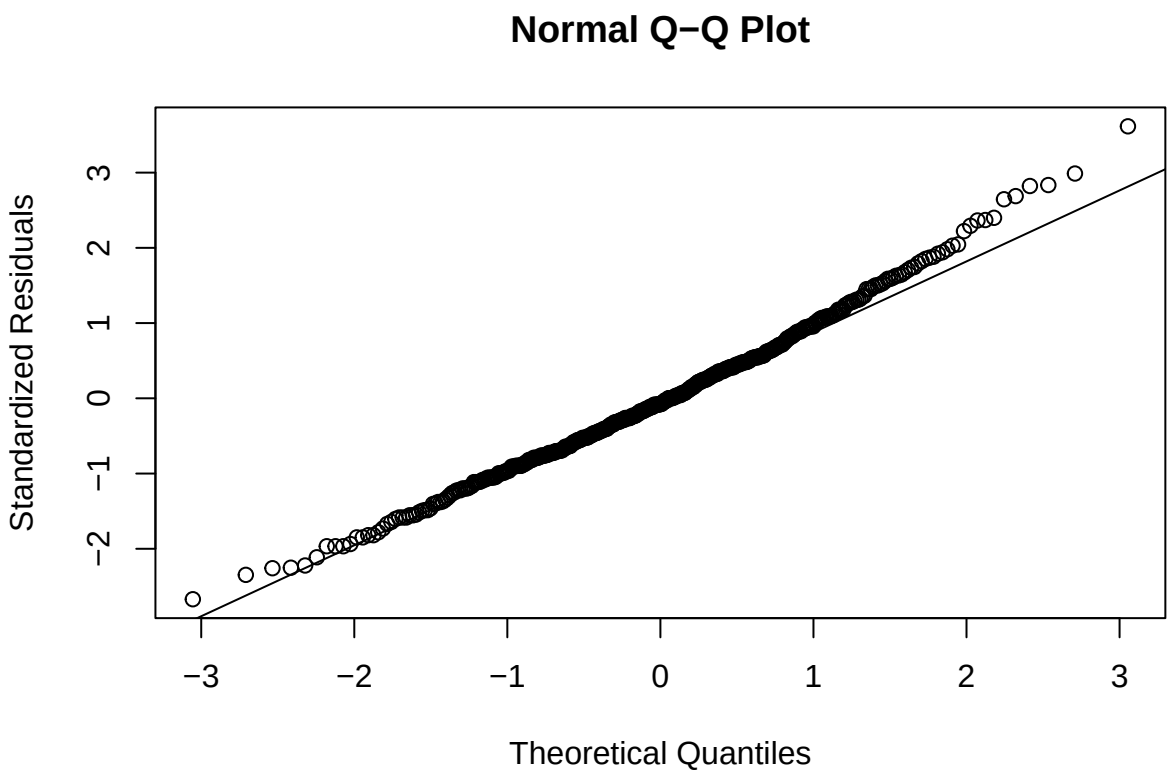
One-way ANOVA for Word Points

```
WordsAnova=aov(Points~Blanks,data=WordsWithFriends)
summary(WordsAnova)
```

```
##              Df Sum Sq Mean Sq F value   Pr(>F)
## Blanks        2  20786   10393   8.927 0.000158 ***
## Residuals   441 513427    1164
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

FIGURE 5.29 Normal quantile plot of residuals from ANOVA fit to words with friends data

```
qqnorm(rstandard(WordsAnova),ylab="Standardized Residuals")
qqline(rstandard(WordsAnova))
```



Extract the MSE and compute CI for the difference between and 1 and 2 blanks

```
tstar=qt(0.975,WordsAnova$df.residual)
MSE=summary(WordsAnova)[[1]]$"Mean Sq"[2]    #gets the MSE from aov summary
sqrt(MSE)
```

```
## [1] 34.12086
```

```
n1=WordsStats$n[2]
n2=WordsStats$n[3]
c(n2,n1)
```

```
## [1] 116 217
```

```
LSD=tstar*sqrt(MSE)*sqrt(1/n1+1/n2)   #Margin of error for the CI is LSD for this pair
round(LSD,2)
```

```
## [1] 7.71
```

```
Diff=WordsStats$mean[3]-WordsStats$mean[2]
round(Diff,1)
```

```
## [1] 6.6
```

```
Lower=Diff-LSD
Upper=Diff+LSD
round(c(Lower,Upper),1)
```

```
## [1] -1.1 14.3
```

Effect size for 2 versus 1 blank

```
EffectSize=Diff/sqrt(MSE)
EffectSize
```

```
## [1] 0.1923901
```

Alternative Solution

Use the asbio package to get the LSD confidence intervals.

```
library(asbio)
```

```
pairw.anova(WordsWithFriends$Points,WordsWithFriends$Blanks,method="lsd")
```

```
##
## 95% LSD confidence intervals
##
##                          LSD      Diff     Lower    Upper  Decision Adj. p-value
## mu0blanks-mu1blank  7.82541 -12.22377 -20.04918 -4.39836 Reject H0      0.00227
## mu0blanks-mu2blanks 8.90398 -18.78829 -27.69227 -9.88431 Reject H0        4e-05
## mu1blank-mu2blanks  7.71303  -6.56452 -14.27754  1.14851    FTR H0      0.09509
```

Note: The signs change when the software computes the interval for mu1-mu2.