Worksheet 4

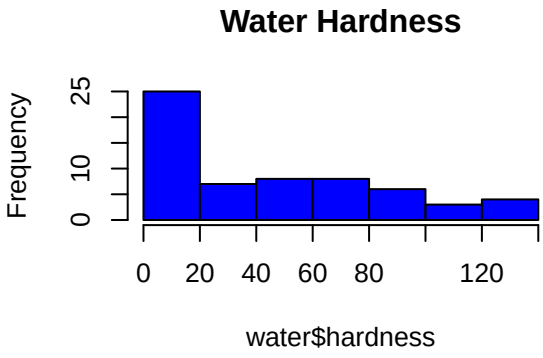
Jarod Wright

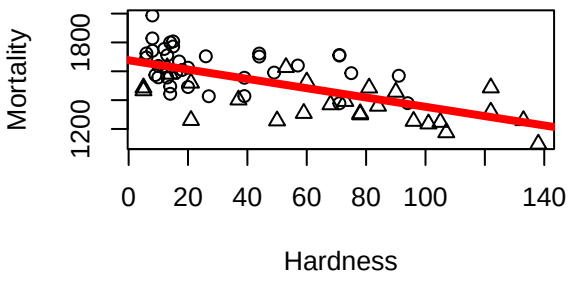
9/13/2021

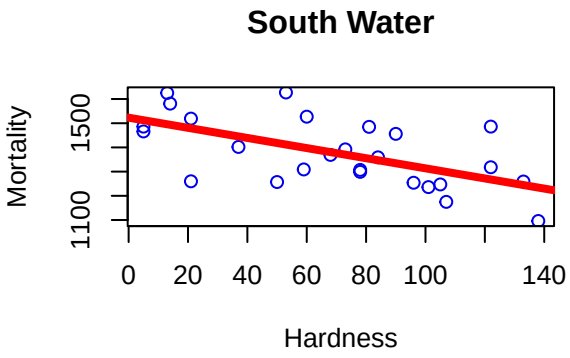
True Difference in means is nonzero: 43.7, 52.55 : Estimation in feet is more precise.

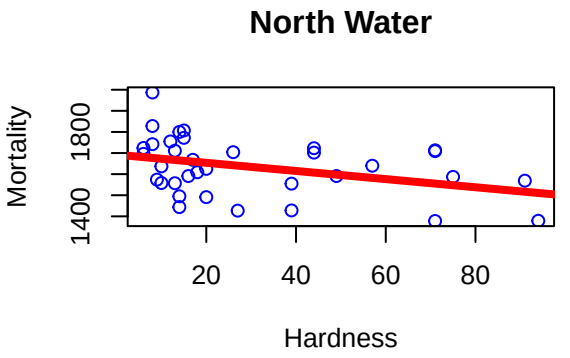
Water Hardness v. Mortality

```
library(HSAUR3)
library(dplyr)
attach(water)

par(mfrow=c(2,2))
water.south <- water %>% filter(location == "South")
water.north <- water %>% filter(location != "South")
hist(water$hardness, main = "Water Hardness", col = "blue")
#boxplot(water$hardness, main = "Water Hardness")

plot(water$hardness,water$mortality, pch=as.numeric(water$location), xlab = "Hardness", ylab = "Mortali
abline(lm(water$mortality~water$hardness), col="red", lwd=4)
plot(water.south$hardness,water.south$mortality, pch=as.numeric(water$location[2]), main = "South Water
abline(lm(water.south$mortality~water.south$hardness), col="red", lwd=4)
plot(water.north$hardness,water.north$mortality, pch=as.numeric(water$location[!2]), main = "North Wate
abline(lm(water.north$mortality~water.north$hardness), col="red", lwd=4)
```










```
# test correlation between mortality and hardness)
cor1 = cor(water.south$mortality, water.south$hardness)
cor.test(water.south$mortality, water.south$hardness)
```

```
##
##  Pearson's product-moment correlation
##
## data:  water.south$mortality and water.south$hardness
## t = -3.6949, df = 24, p-value = 0.001135
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.8023588 -0.2801435
## sample estimates:
##       cor
## -0.6021533
```

```
cor2 = cor(water.north$mortality, water.north$hardness)
cor.test(water.north$mortality, water.north$hardness)
```

```
##
##  Pearson's product-moment correlation
##
## data:  water.north$mortality and water.north$hardness
## t = -2.2778, df = 33, p-value = 0.02934
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.62506516 -0.04030171
## sample estimates:
##       cor
## -0.3685978
```

```r
#correlation for the north is approx. half of correlation for the south
```

```r
data("rearrests")

 x <-matrix(c(sum(rearrests[c(1,3)]),sum(rearrests[1:2]),sum(rearrests[c(2,4)]),sum(rearrests[3:4])), nr
 colnames(x) <-c('Yes','No')
 rownames(x) <-c('adult court','juvenile court')
 rearrests.tab <-as.table(x)

 print(rearrests)
```

```
##              Juvenile court
## Adult court   Rearrest No rearrest
##   Rearrest         158         515
##   No rearrest      290        1134
```

```r
 print(x)
```

```
##                Yes   No
## adult court    673 1424
## juvenile court 448 1649
```

```r
 prop.test(x)
```

```
##
##  2-sample test for equality of proportions with continuity correction
##
## data:  x
## X-squared = 61.088, df = 1, p-value = 5.458e-15
## alternative hypothesis: two.sided
## 95 percent confidence interval:
##  0.0802301 0.1343622
## sample estimates:
##    prop 1    prop 2
## 0.3209347 0.2136385
```

```r
 # Approx 10% higher rearrest rate when tried in adult court
```