

Final Project

Ian Baker, Loughlin Claus, Zack Schieberl

12/7/2019

Pledge

I pledge my honor that I have abided by the Stevens Honor System - Ian Baker, Loughlin Claus, Zack Schieberl

11.53

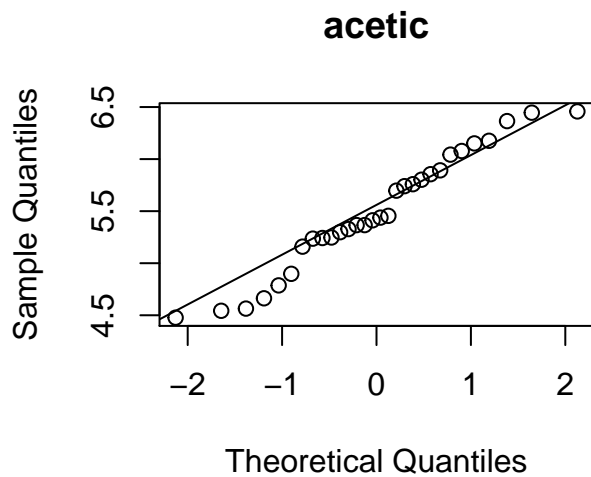
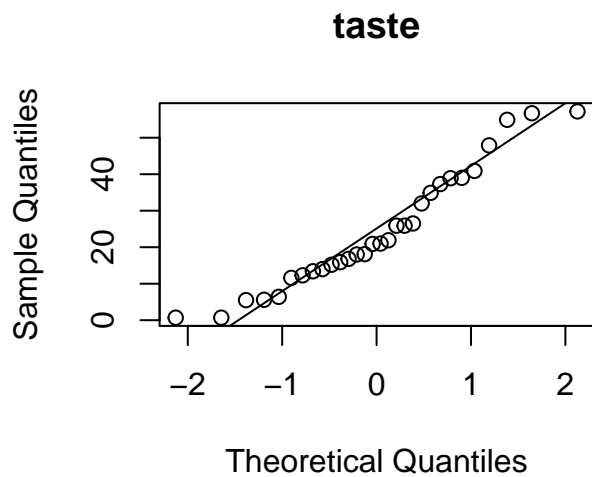
```
cheese <- as.matrix(read.csv2("cheese.csv", header = TRUE, sep = ","))

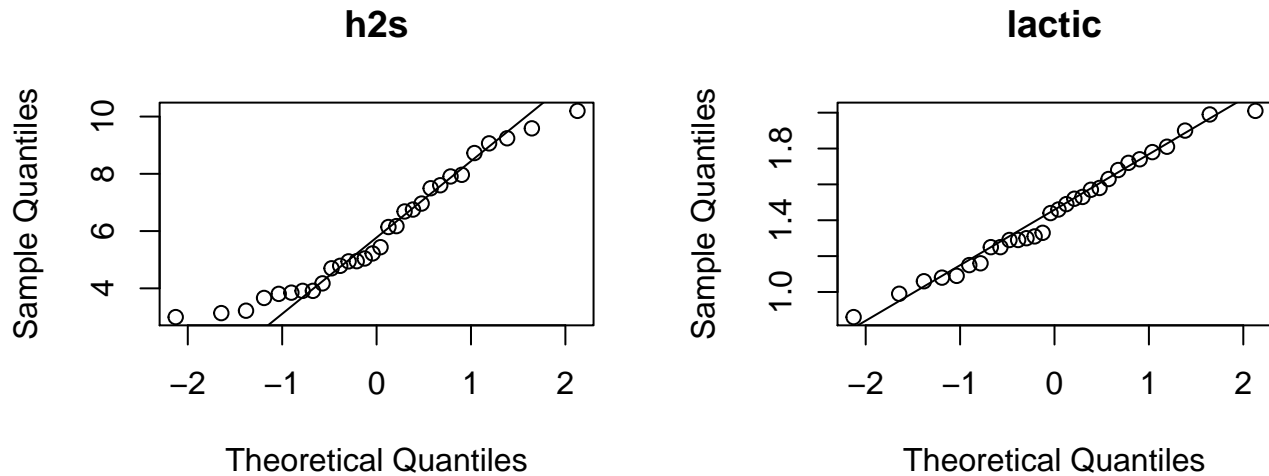
cheeseCols <- colnames(cheese)
for (col in cheeseCols) {
  cur <- as.numeric(cheese[, col])
  # mean, median, sd, iqr
  out <- c(paste("Type:", col), paste("Mean:", round(mean(cur), 2)),
          paste("Median:", round(median(cur), 2)), paste("SD:", round(sd(cur), 2)),
          paste("IQR:", round(IQR(cur), 2)))
  print(format(out, justify = "left", trim = TRUE))
  # stemplot
  stem(cur)
  # normal quantile plot
  qqnorm(cur, main = col)
  qqline(cur)
}
```

```
## [1] "Type: taste " "Mean: 24.53 " "Median: 20.95" "SD: 16.26 "
## [5] "IQR: 23.15 "
##
## The decimal point is 1 digit(s) to the right of the |
##
## 0 | 11666
## 1 | 223456788
## 2 | 112667
## 3 | 25799
## 4 | 18
## 5 | 577
##
## [1] "Type: acetic" "Mean: 5.5 " "Median: 5.42" "SD: 0.57 " "IQR: 0.65 "
##
## The decimal point is 1 digit(s) to the left of the |
##
## 44 | 846
## 46 | 69
## 48 | 0
## 50 | 6
## 52 | 4450377
## 54 | 146
## 56 | 046
## 58 | 069
## 60 | 4858
## 62 | 7
## 64 | 56
```

```
## [1] "Type: h2s    " "Mean: 5.94  " "Median: 5.33" "SD: 2.13    " "IQR: 3.6    "
##
## The decimal point is at the |
##
## 2 |
## 3 | 01278999
## 4 | 27899
## 5 | 024
## 6 | 1278
## 7 | 0569
## 8 | 07
## 9 | 126
## 10 | 2

## [1] "Type: lactic" "Mean: 1.44  " "Median: 1.45" "SD: 0.3    " "IQR: 0.42    "
##
## The decimal point is 1 digit(s) to the left of the |
##
## 8 | 69
## 10 | 68956
## 12 | 5599013
## 14 | 4692378
## 16 | 38248
## 18 | 109
## 20 | 1
```



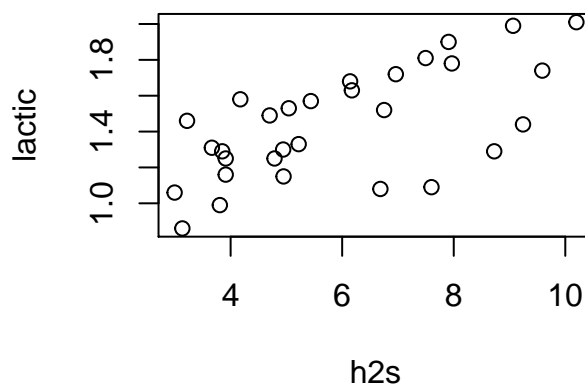
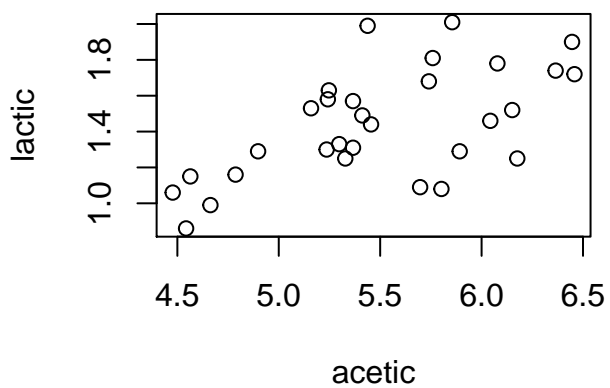
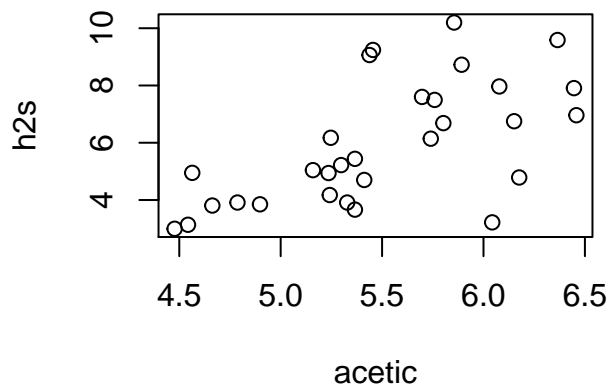
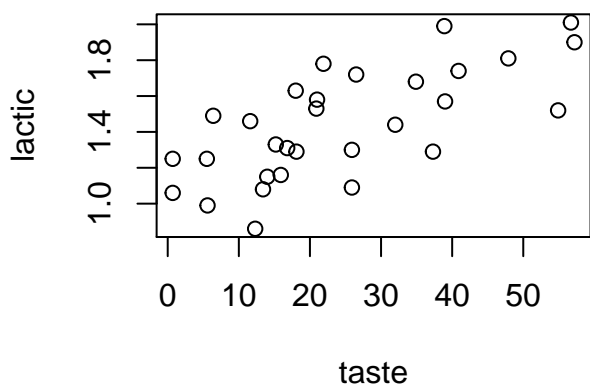
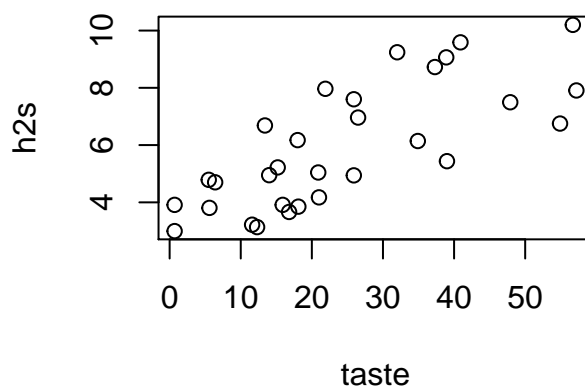
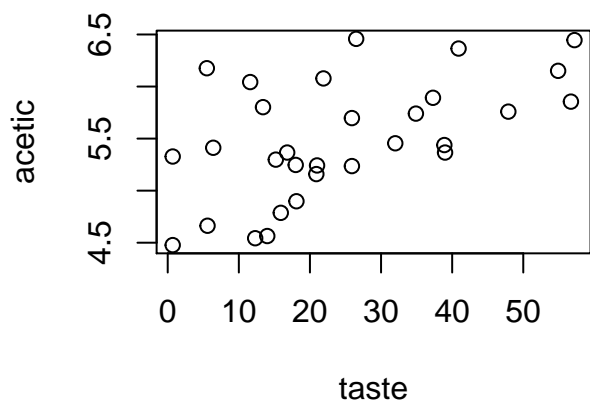


While H2S and Taste have some right skew, and Acetic has two peaks, the data all appears to be relatively normal. There are no outliers in the data.

11.54

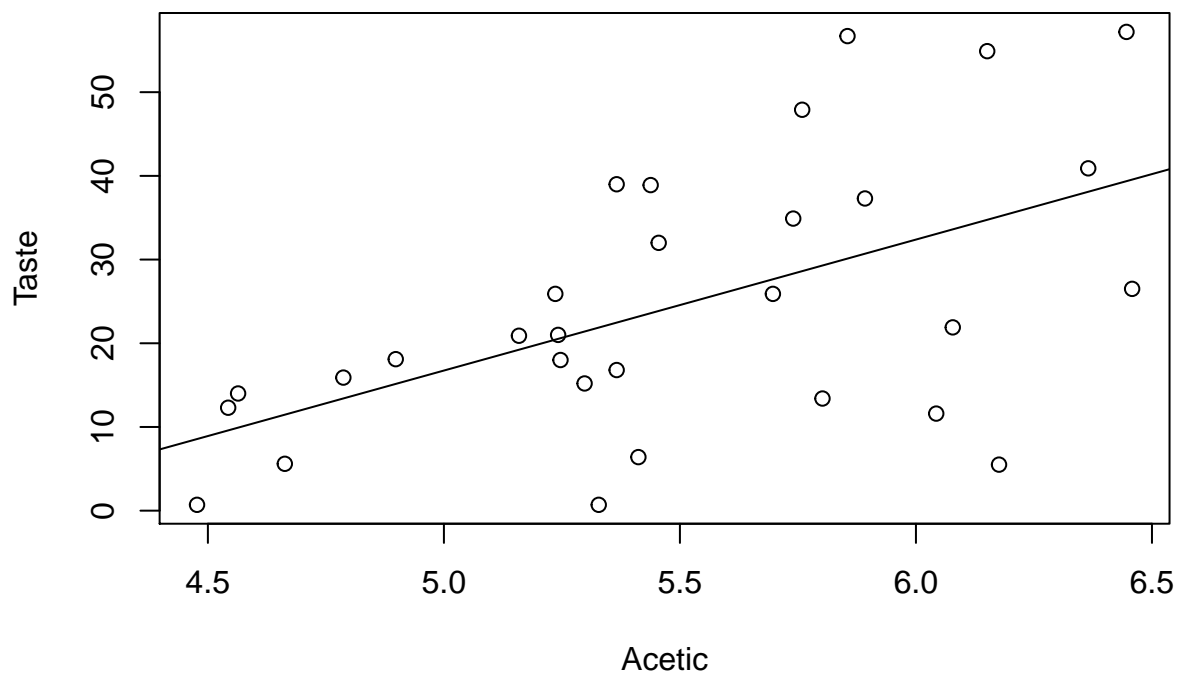
```
for (col in cheeseCols) {
  colIdx <- grep(col, cheeseCols)
  col1Data <- as.numeric(cheese[, col])
  for (col2 in cheeseCols) {
    if (colIdx < grep(col2, cheeseCols)) {
      col2Data <- as.numeric(cheese[, col2])
      plot(col1Data, col2Data, xlab = col, ylab = col2)
      correl <- cor.test(col1Data, col2Data)
      cat("Correlation between", col, "and", col2, "is:", correl$estimate,
          "with a p-value of", correl$p.value, "\n")
    }
  }
}
```

```
## Correlation between taste and acetic is: 0.5495393 with a p-value of 0.001658192
## Correlation between taste and h2s is: 0.7557523 with a p-value of 1.373783e-06
## Correlation between taste and lactic is: 0.7042362 with a p-value of 1.405117e-05
## Correlation between acetic and h2s is: 0.6179559 with a p-value of 0.0002739173
## Correlation between acetic and lactic is: 0.6037826 with a p-value of 0.0004113657
## Correlation between h2s and lactic is: 0.6448123 with a p-value of 0.0001198401
```

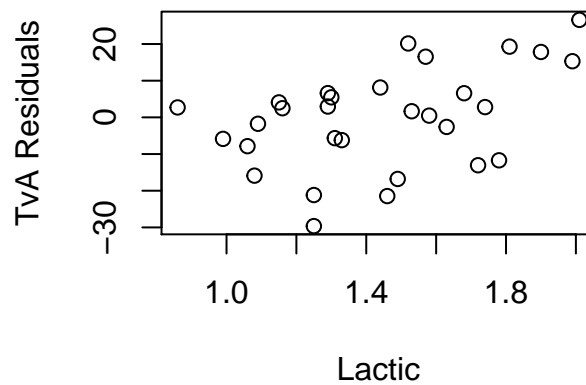
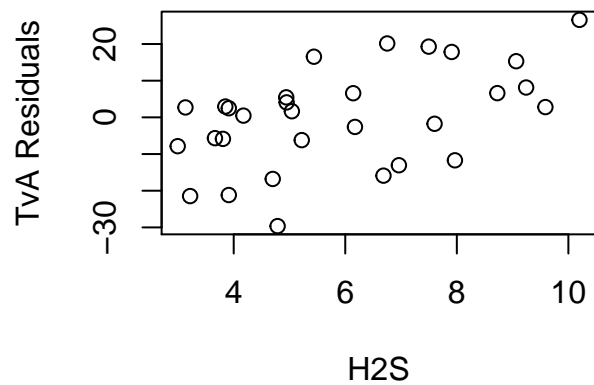


11.55

```
tasteCol <- as.numeric(cheese[, "taste"])
aceticCol <- as.numeric(cheese[, "acetic"])
tasteVsAcetic <- lm(tasteCol ~ aceticCol, data.frame(cheese))
plot(aceticCol, tasteCol, xlab = "Acetic", ylab = "Taste")
abline(tasteVsAcetic)
```



```
tVsAResiduals <- residuals(tasteVsAcetic)
plot(cheese[, "h2s"], tVsAResiduals, xlab = "H2S", ylab = "TvA Residuals")
plot(cheese[, "lactic"], tVsAResiduals, xlab = "Lactic", ylab = "TvA Residuals")
```



The residuals both have a normal distribution and seem to be positively associated with Lactic and H2S.