Math 3070 § 1. **Aquifer Data:** Name: Example
Treibergs **Summary statistics and Boxplot.** August 5, 2010

Data File Used in this Analysis:

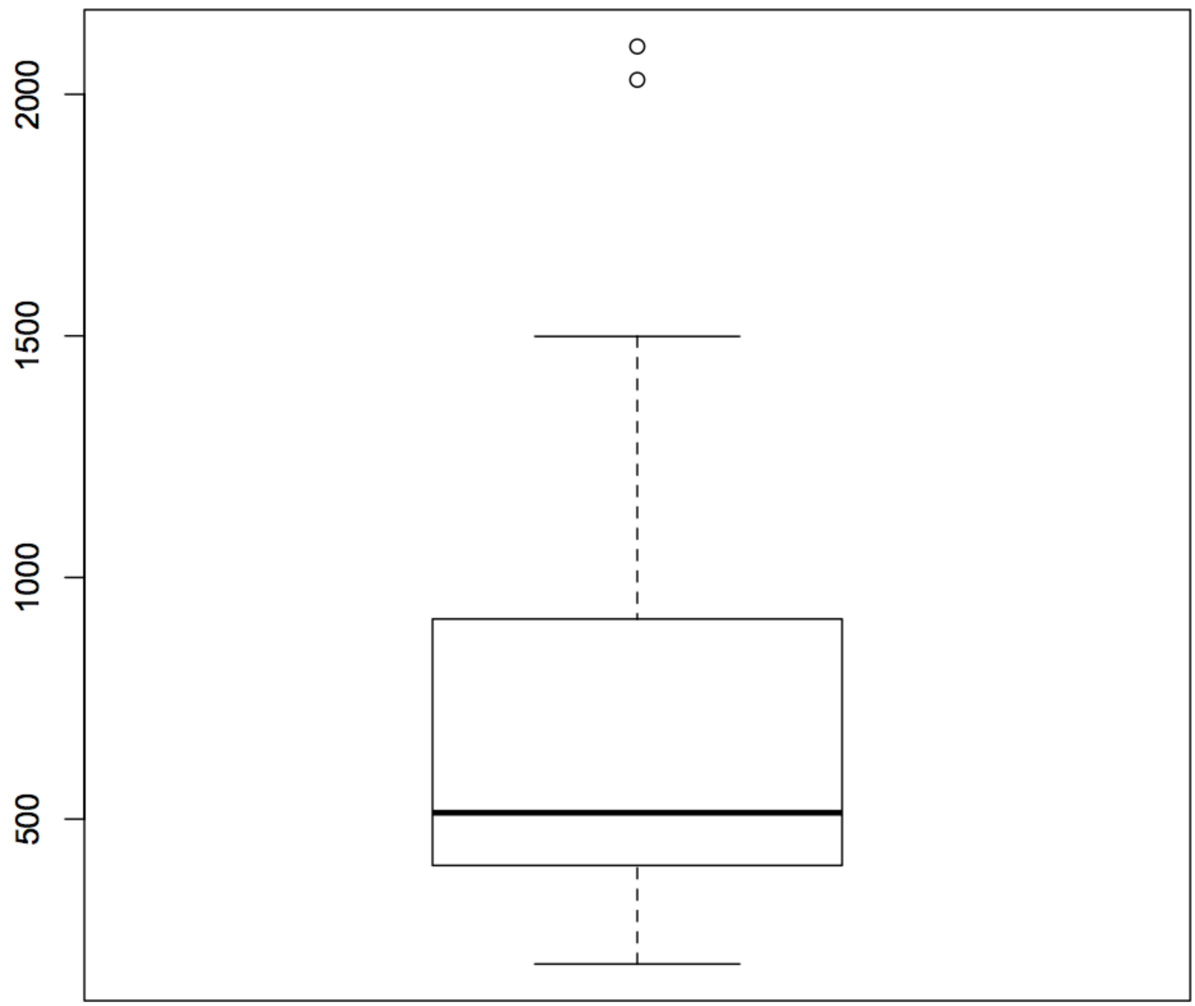
```
# Math 3070 - 1    Aquifer Data          Aug. 5, 2010
# treibergs
#
# From navidi, principles of Statistics for Engineers and Scientists
# Mcgraw hill 2010. taken from an article S Jeen, J. Kim et. al.,
# "Hydrochemical Characteristics of Groundwater in Mid-Western Coastal
# Aquifer System", Geosciences Journal, 2001. here are measurements of
# the electrical conductivity in microsiemens per centimeter of 23
# samples of an aquifer system in Korea.
#
Conductivity
2099
528
2030
1350
1018
384
1499
1265
375
424
789
810
522
513
488
200
215
486
257
557
260
461
500
```

R Session:

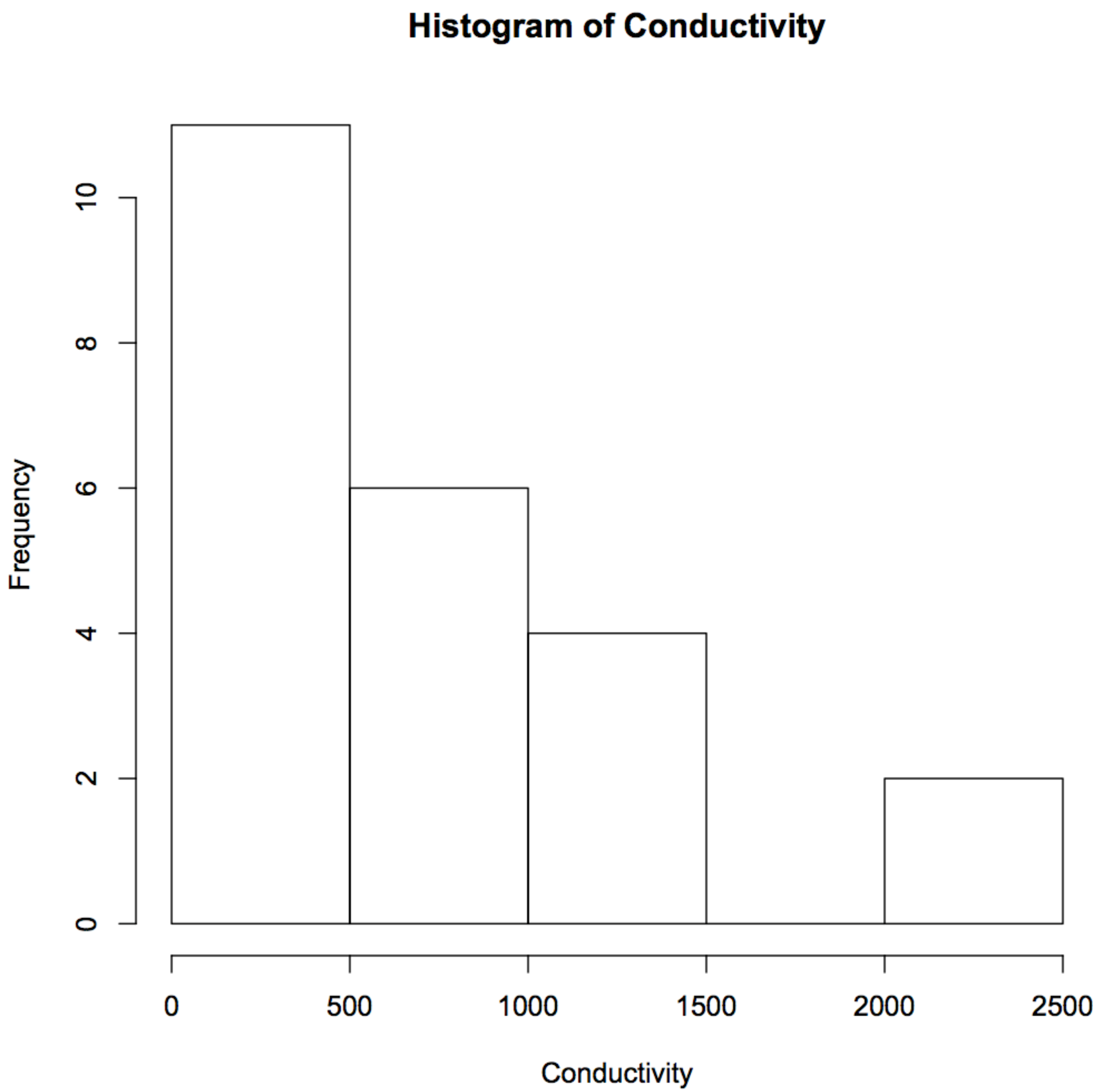
```
R version 2.10.1 (2009-12-14)
Copyright (C) 2009 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale
```

```
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[R.app GUI 1.31 (5537) powerpc-apple-darwin9.8.0]

> tt <- read.table("M3073AquiferData.txt",header=TRUE)
> tt
   Conductivity
1          2099
2           528
3          2030
4          1350
5          1018
6           384
7          1499
8          1265
9           375
10          424
11          789
12          810
13          522
14          513
15          488
16          200
17          215
18          486
19          257
20          557
21          260
22          461
23          500
> attach(tt)

> # Compute summary statistics of Conductivity
> summary(Conductivity)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  200.0   404.0   513.0   740.4   914.0  2099.0
> mean(Conductivity)
[1] 740.4348
> sd(Conductivity)
[1] 549.8366
> sd(Conductivity)^2; var(Conductivity)
[1] 302320.3
[1] 302320.3
```

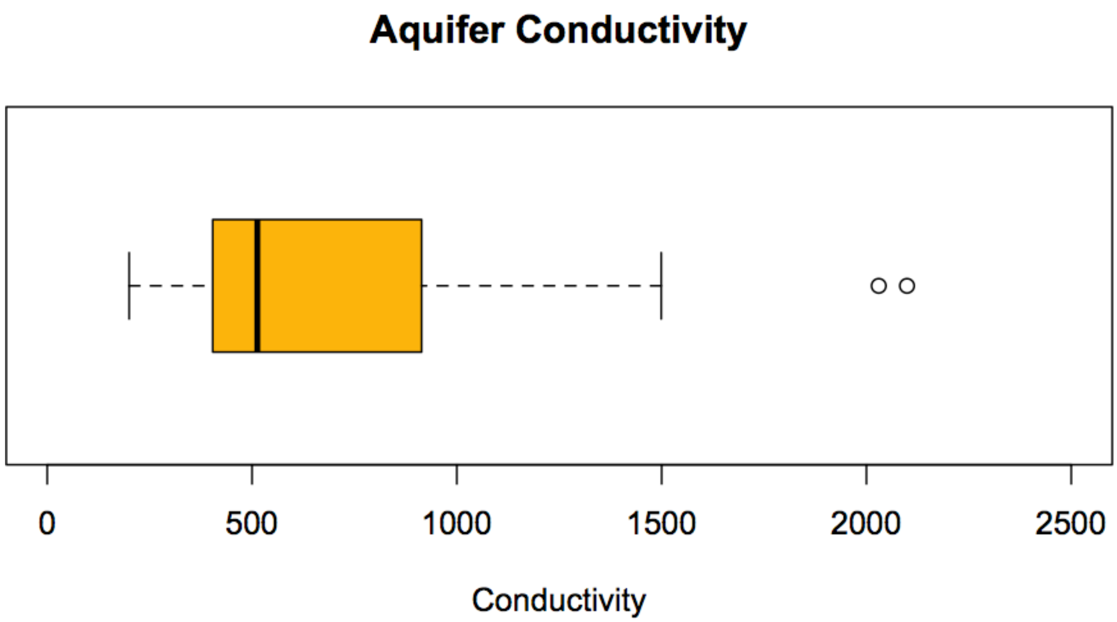
```
> # We view the boxplot and histogram. We set the ranges the same on both
> # and adjust the margins.
> boxplot(Conductivity)
> hist(Conductivity)
> xr <- c(0,2500)
> layout(matrix(1:2,ncol=1))
> par(mar=par("mar")*c(0.8,1,1,1))
> boxplot(Conductivity,ylim=xr,horizontal=TRUE,
  xlab="Conductivity", col="orange",main="Aquifer Conductivity")
> hist(Conductivity,xlim=xr,xlab="",ylab="",main="", axes=FALSE, col="yellow")
> axis(1)
>
>
>  # Compute summary statistics of log(Conductivity)
> lc <- log(Conductivity)
> summary(lc)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  5.298   6.000   6.240   6.380   6.811   7.649
> mean(lc)
[1] 6.380078
> sd(lc)
[1] 0.6748329
> var(lc)
[1] 0.4553994

> # We view the histogram. We set the ranges the same on both and adjust the margins.
> hist(lc)
> lxr <- c(5,8)
> boxplot(lc,ylim=lxr,horizontal=TRUE, xlab="log(Conductivity)",
  col="lightblue",main="Aquifer Conductivity")
> hist(lc,xlim=lxr,xlab="",ylab="",main="", axes=FALSE, col="lightgreen")
> axis(1)
>
```

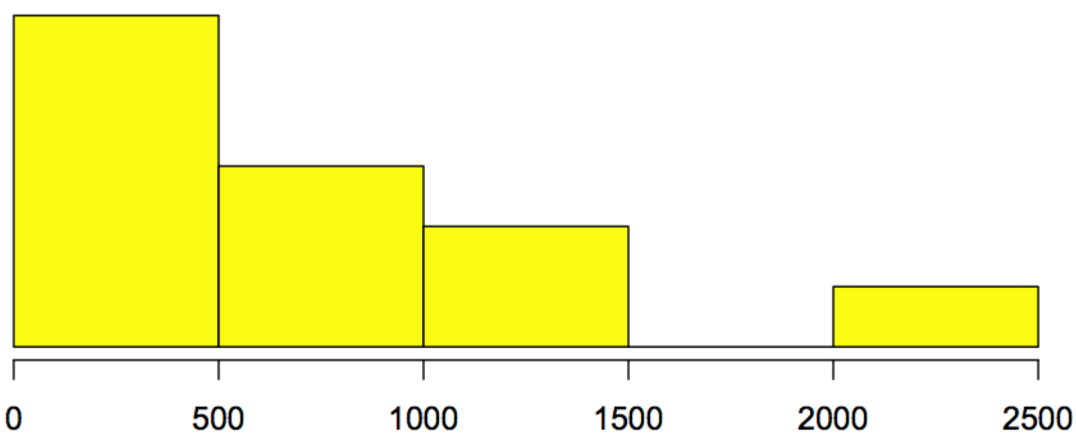


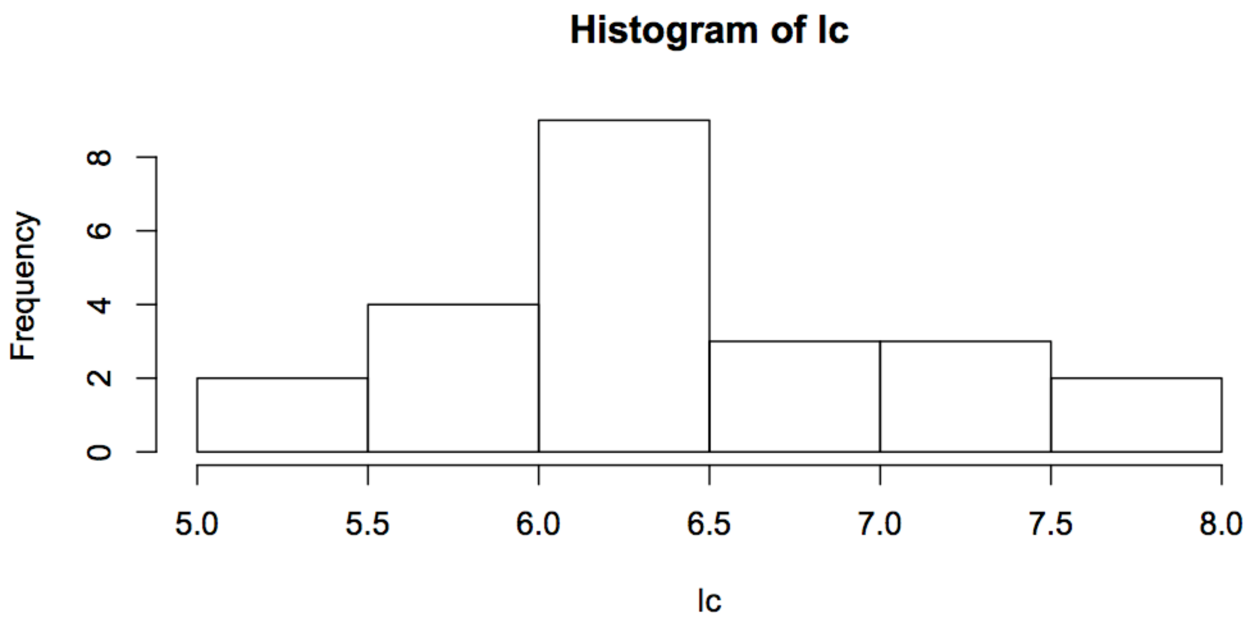
2000
1500
1000
500

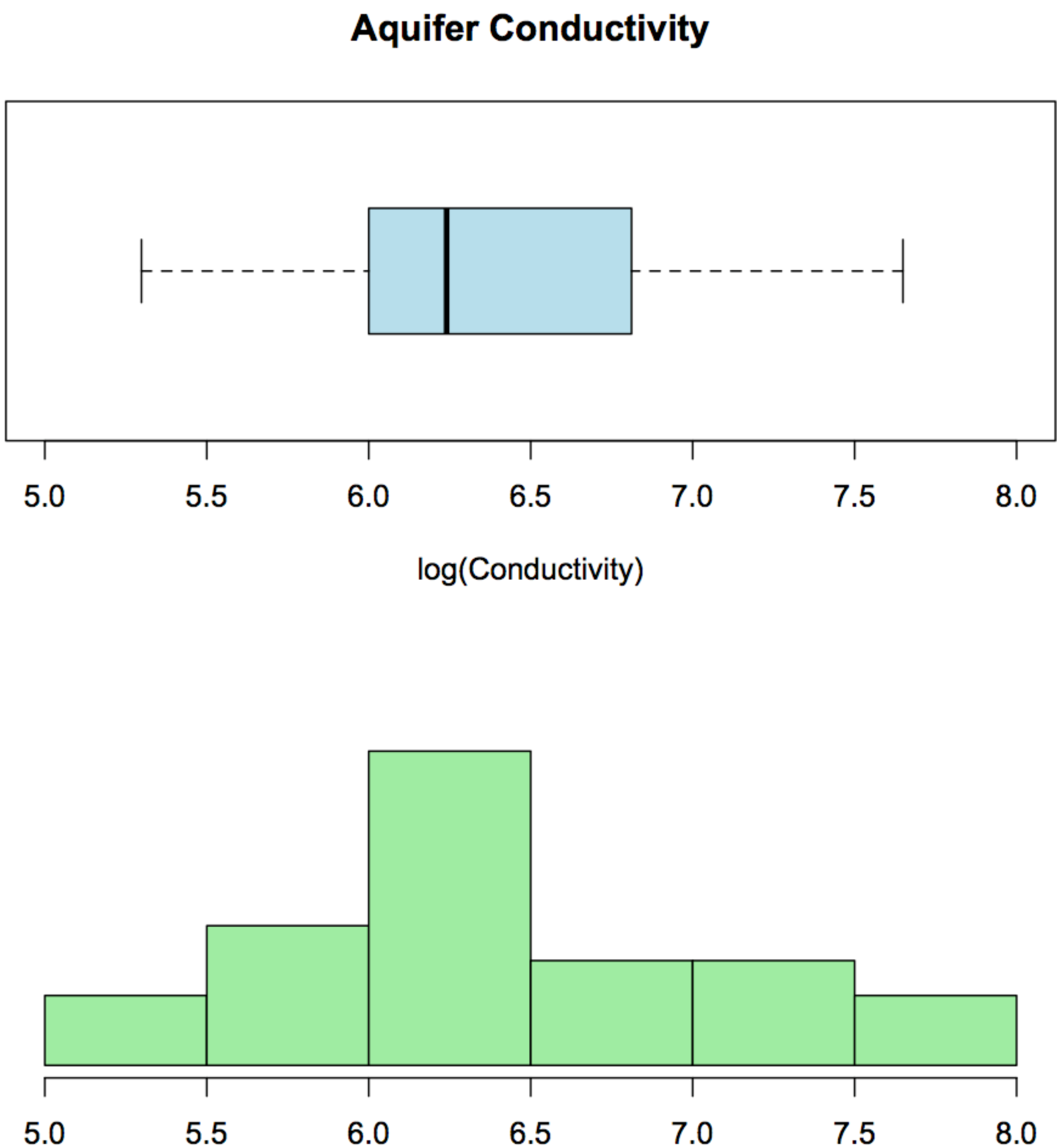


Histogram of Conductivity
Frequency
0
2
4
6
8
10
0
500
1000
1500
2000
2500
Conductivity


Aquifer Conductivity
0
500
1000
1500
2000
2500
Conductivity


0
500
1000
1500
2000
2500


Histogram of lc
Frequency
0
2
4
6
8
5.0
5.5
6.0
6.5
7.0
7.5
8.0
lc


Aquifer Conductivity
5.0
5.5
6.0
6.5
7.0
7.5
8.0
log(Conductivity)
5.0
5.5
6.0
6.5
7.0
7.5
8.0