Math 3070 § 1.	**Hospitalization Data:**	Name: Example
Treibergs	**Means and Trimmed Means.**	September 2, 2010

Data File Used in this Analysis:

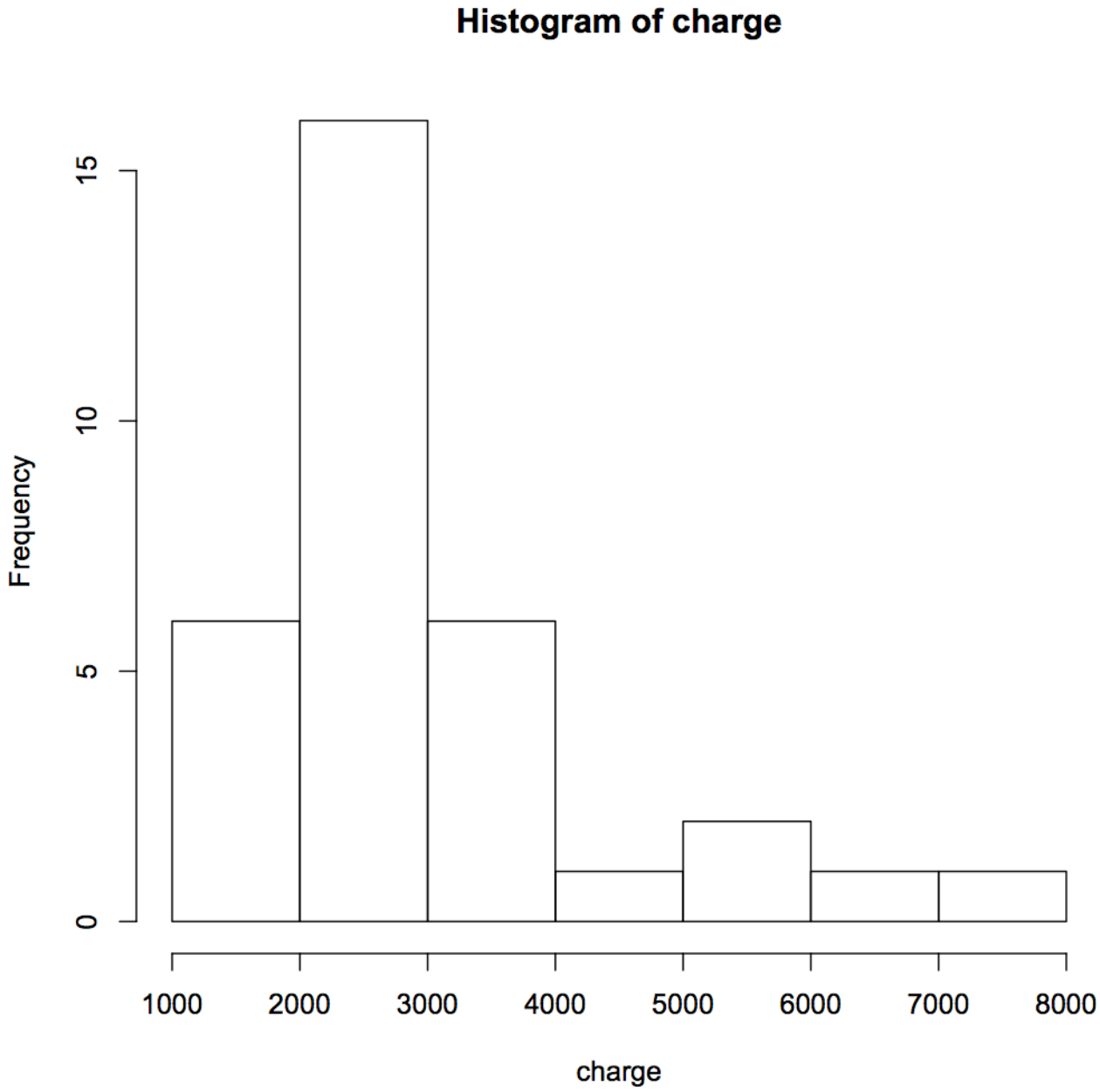
```
# Math 3070 - 1    Hospital Data    Treibergs
#
# From E. W Frees, Estimating densities of functions of observations,
# Journal of the American Statistical Association, 1994. From text by
# W. Rosenkrantz, Probability and Statistics for Science, Engineering and
# Finance, Chapman Hall/CRC, 2009. charge = 1998 total hospital charge in
# dollars for 33 female patients admitted for circulatory disorders.
#
charge
2337
2179
2348
4765
2088
2872
1924
2294
2182
2138
1765
2467
3609
2141
1850
3191
3020
2473
1898
7787
6169
1802
2011
2270
3425
3558
2315
1642
5878
2101
2242
5746
3041
```

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("M3073HospitalData.txt",header=TRUE)
> tt
   charge
1    2337
2    2179
3    2348
4    4765
5    2088
6    2872
7    1924
8    2294
9    2182
10   2138
11   1765
12   2467
13   3609
14   2141
15   1850
16   3191
17   3020
18   2473
19   1898
20   7787
21   6169
22   1802
23   2011
24   2270
25   3425
26   3558
```

```
27   2315
28   1642
29   5878
30   2101
31   2242
32   5746
33   3041
> attach(tt)

> #   List of charges is put into the vector  charge.
> #   Use canned summary
> summary(charge)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   1642    2101    2315    2955    3191    7787

> # Sample Standard Deviation
> sd(charge)
[1] 1481.044

> # Sample variance
> var(charge)
[1] 2193492

> # Plot Histogram
> hist(charge)


> charge
 [1] 2337 2179 2348 4765 2088 2872 1924 2294 2182 2138 1765 2467
[13] 3609 2141 1850 3191 3020 2473 1898 7787 6169 1802 2011 2270
[25] 3425 3558 2315 1642 5878 2101 2242 5746 3041


> # Sort the vector charge
> scharge <- sort(charge);scharge
 [1] 1642 1765 1802 1850 1898 1924 2011 2088 2101 2138 2141 2179
[13] 2182 2242 2270 2294 2315 2337 2348 2467 2473 2872 3020 3041
[25] 3191 3425 3558 3609 4765 5746 5878 6169 7787

> # 2:33 denotes the vector 2,3,4,5,...,31,32.
> # The charges corresponding to these indices are (dropping first and last)
> scharge[2:32]
 [1] 1765 1802 1850 1898 1924 2011 2088 2101 2138 2141 2179 2182
[13] 2242 2270 2294 2315 2337 2348 2467 2473 2872 3020 3041 3191
[25] 3425 3558 3609 4765 5746 5878 6169
```

```
> # The alpha-trimmed mean is gotten by averaging after dropping alpha n terms
> # from front and back of sorted data. Compare hand and canned trimmed mean.
> sum(scharge[2:32])/31;mean(charge,trim=1/33)
[1] 2841.903
[1] 2841.903
> sum(scharge[3:31])/29;mean(charge,trim=2/33)
[1] 2764.310
[1] 2764.310


> # For alpha's that are not fractions of n = number of data we interpolate.
> # For alpha between p/n and (p+1)/n we choose 0 < theta < 1 so that
> # (1 - theta)p/n + \theta (p+1)/n = alpha.
> # Thus theta = (alpha - p/n) / (1/n)
> # For example, if alpha = 0.1,
> theta <- (.1-3/33)/(1/33);theta
[1] 0.3
> (1-theta)*(3/33)+theta*(4/33)
[1] 0.1
> # and the alpha trimmed mean is the interpolation of the p/n and (p+1)/n trimmed means
> xtr1<-sum(scharge[4:30])/27;xtr1;mean(charge,trim=3/33)
[1] 2684.630
[1] 2684.630
> xtr2<-sum(scharge[5:29])/25;xtr2;mean(charge,trim=4/33)
[1] 2595.56
[1] 2595.56
> (1-theta)*xtr1+theta*xtr2;mean(charge,trim=.1)
[1] 2657.909
[1] 2684.630
> # Note that R does not interpolate the same as Devore. R removes the fraction of trim
> # and computes the mean of the remaining middle data. In this case
> # R's mean(charge,trim=0.1) = Devore's mean(charge, trim=3/33)
```


Histogram of charge
Frequency
0
5
10
15
1000
2000
3000
4000
5000
6000
7000
8000
charge