Solution to first Homework

8. Chapter 1
a. Number observations equal 2 x 2 x 2 = 8
b. This could be called an analytic study because the data would be collected on an existing process. There is no sampling frame.

15. Chapter 1

CRUNCHY — 4 |*3 represents 34

CREAMY — 3*|6 represents 36

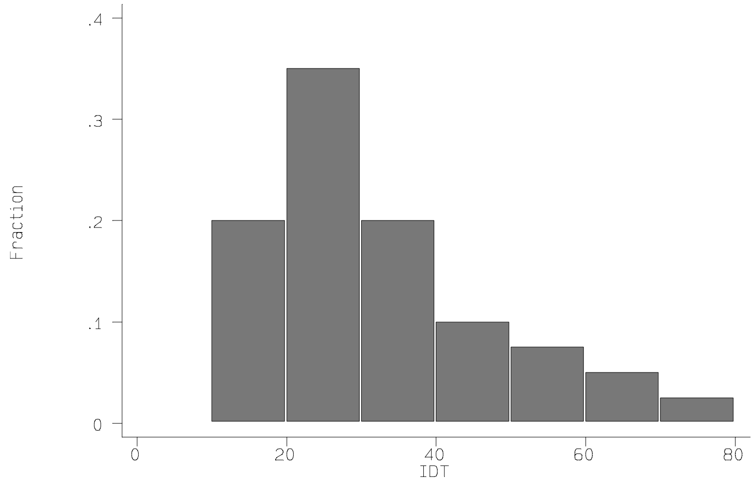
```
        2* | 2
  644| *3 3* | 0069
77220| *4 4* | 00145
 6320| *5 5* | 003666
  222| *6 6* | 258
   55| *7
    0| *8
```

Crunchy received higher scores than Creamy and has three of the highest scores. Creamy received lower scores than Crunchy and has three of the lowest scores. Crunchy is skewed to the right and Creamy scores are skewed to the left. However, both have bulk of the data common in the middle from the 30's to the 60's.

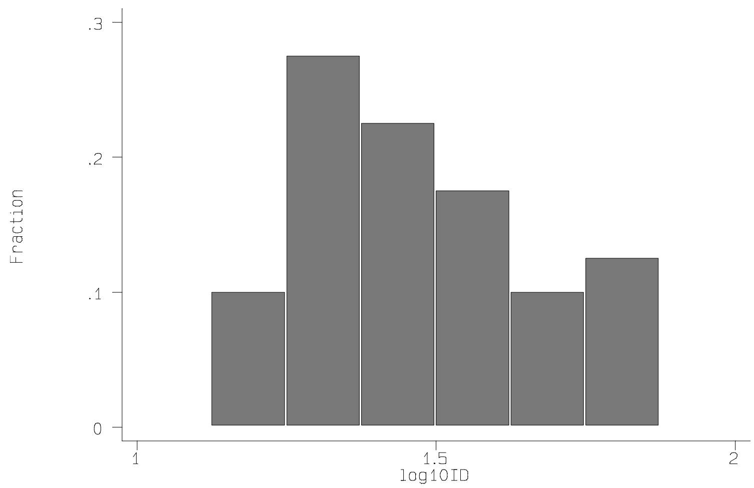
Chapter 1
22. A very large percentage of the data values are greater than 0, which indicates that most, but not all, runners do slow down at the end of the race. The histogram is also positively skewed, which means that some runners slow down a lot compared to the others. A typical value for this data would be in the neighborhood of 200 seconds. The proportion of the runners who ran the last 5 km faster than they did the first 5 km is very small, about 1% or so.

Chapter 1

25. Histogram of original data:



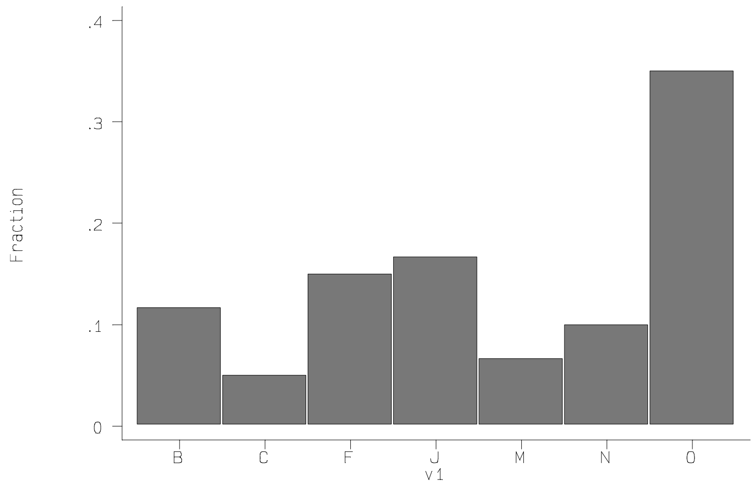
Histogram of transformed data:



The transformation creates a much more symmetric, mound-shaped histogram.

29. **Chapter 1**

Complaint	Frequency	Relative Frequency
B	7	0.1167
C	3	0.0500
F	9	0.1500
J	10	0.1667
M	4	0.0667
N	6	0.1000
O	21	0.3500
Total =	60	1. 0000



33. **Chapter 1**
a. x-bar = 192.57, sample median =189. The mean is larger than the median, but they are still fairly close together.
b. Changing the one value, x-bar = 189.71, sample median =189. The mean is lowered, the median stays the same.
c. trimmed mean when eliminating the smallest and largest observations is 191. Trimming percentage is 1/14 = 7.14 % (trimmed from each tail).
d. For n = 13, Σx = (119.7692) x 13 = 1,557
For n = 14, Σx = 1,557 + 159 = 1,716
x-bar = 122. 5714 or 122.57

.

36. Chapter 1
a. A stem-and leaf display of this data appears below:

```
32 |55          stem: ones in seconds
33 |49          leaf: tenths in seconds
34 |
35 |6699
36 |34469
37 |03345
38 |9
39 |2347
40 |23
41 |
42 |4
```

The display is reasonably symmetric, so the mean and median will be close.

b. The sample mean is = 9638/26 = 370.7. The sample median is
= (369+370)/2 = 369.50.
c. The largest value (currently 424) could be increased by any amount. Doing so will not

change the fact that the middle two observations are 369 and 170, and hence, the median will not change. However, the value x = 424 can not be changed to a number less than 370 (a change of 424-370 = 54) since that *will* lower the values(s) of the two middle observations.
d. Expressed in minutes, the mean is (370.7 sec)/(60 sec) = 6.18 min; the median is 6.16 min.

49. Chapter 1
a. Sum of data = 2.75+ …+ 3.01 = 56.80, Sum of square of data = (2.75)(2.75)+…+(3.01)(3.01) = 197.804
b. Sample variance s-squared = {197.804- (56.8)(56.8)/17}/16 = 8.0252/16 =0.5016
s the sample standard deviation is the square-root of 0.5016 = 0.708

59. Chapter 1
a. ED: median = .4 (the 14th value in the *sorted* list of data). The lower quartile (median of the lower half of the data, including the median, since n is odd) is (.1+.1)/2 = .1. The upper quartile is (2.7+2.8)/2 = 2.75. Therefore, IQR = 2.75 - .1 = 2.65.
Non-ED: median = (1.5+1.7)/2 = 1.6. The lower quartile (median of the lower 25 observations) is .3; the upper quartile (median of the upper half of the data) is 7.9. Therefore, IQR = 7.9 - .3 = 7.6.
b. ED: mild outliers are less than .1 - 1.5(2.65) = -3.875 or greater than 2.75 + 1.5(2.65) = 6.725. Extreme outliers are less than .1 - 3(2.65) = -7.85 or greater than 2.75 + 3(2.65) = 10.7. So, the two largest observations (11.7, 21.0) are extreme outliers and the next two largest values (8.9, 9.2) are mild outliers. There are no outliers at the lower end of the data.
Non-ED: mild outliers are less than .3 - 1.5(7.6) = -11.1 or greater than 7.9 + 1.5(7.6) = 19.3. Note that there are no mild outliers in the data, hence there can not be any extreme outliers either.
c. A comparative boxplot appears below. The outliers in the ED data are clearly visible. There is noticeable positive skewness in both samples; the Non-Ed data has more variability then the Ed data; the typical values of the ED data tend to be smaller than those for the Non-ED data.