Methods for Describing Data Continued

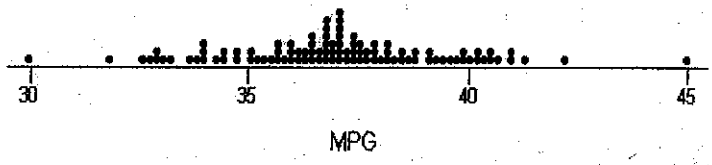
Sorted EPA Mileage Ratings on 100 Cars

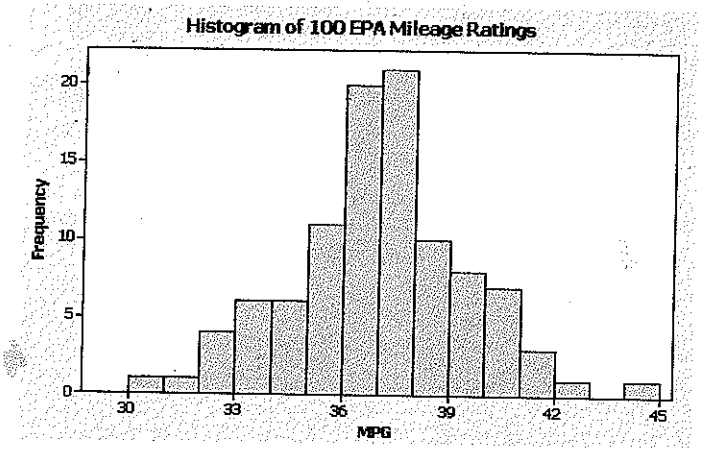
30.0	33.9	35.2	36.1	36.7	37.0	37.4	38.0	39.0	40.2
31.8	33.9	35.3	36.2	36.7	37.0	37.4	38.1	39.0	40.3
32.5	34.0	35.5	36.3	36.7	37.0	37.5	38.2	39.3	40.5
32.7	34.2	35.6	36.3	36.8	37.1	37.6	38.2	39.4	40.5
32.9	34.4	35.6	36.4	36.8	37.1	37.6	38.3	39.5	40.7
32.9	34.5	35.7	36.4	36.8	37.1	37.7	38.4	39.7	41.0
33.1	34.8	35.8	36.5	36.9	37.2	37.7	38.5	39.8	41.0
33.2	34.8	35.9	36.5	36.9	37.2	37.8	38.6	39.9	41.2
33.6	35.0	35.9	36.6	36.9	37.3	37.9	38.7	40.0	42.1
33.8	35.1	36.0	36.6	37.0	37.3	37.9	38.8	40.1	44.9

A Stem and Leaf Display

```
 1     30  0
 2     31  8
 6     32  5799
 12    33  126899
 18    34  024588
 29    35  01235667899
 49    36  01233445566777888999
(21)   37  000011122334456677899
 30    38  0122345678
 20    39  00345789
 12    40  0123557
 5     41  002
 2     42  1
 1     43
 1     44  9
```

A Dotplot of MPG





Descriptive Statistics

Variable	N	Mean	Median	Sample StDev	Minimum	Maximum	Range	Q1	Q2
MPG									

25% of the MPG values are less than _____ ; 50% of the MPG values are less than _____;
75% of the MPG values are less than _____.
About _____% of the MPG values are within one StDev of the mean.
About _____% of the MPG values are within two StDevs of the mean.
About _____% of the MPG values are within three StDevs of the mean.

Summation Notation and Computation of Descriptive Statistics

$$\sum_{i=1}^{n} i = 1 + 2 + 3 + \ldots + n.$$

We read "$\sum_{i=1}^{n} x_i$" as "The sum of the measurements denoted by x_i beginning with x_1 and ending with x_n."

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + x_3 + \ldots + x_n$$

Formula for a Sample Mean: $$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Formula for Sample Variance: $$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$

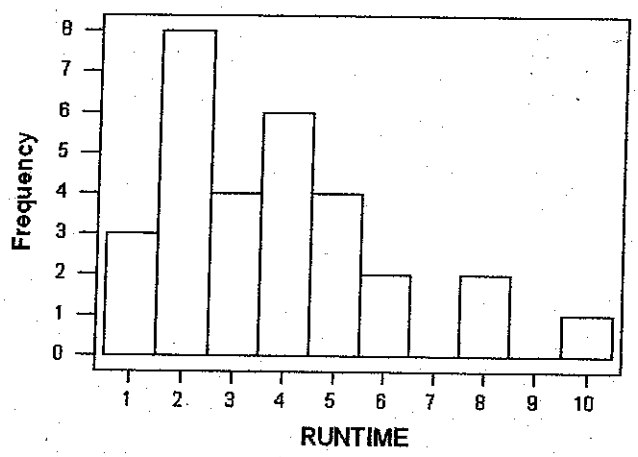
Formula for Sample Standard Deviation: $s = \sqrt{s^2}$

Example of run times (in minutes) of 30 rats running through a maze:

1.97	5.36	9.70	6.06	1.93	7.60
1.74	4.02	1.71	5.63	2.02	2.06
3.77	3.81	1.15	4.25	4.55	3.65
0.60	1.06	8.29	4.44	5.15	3.16
2.75	3.20	2.47	5.21	3.37	1.65

Descriptive Statistics

Variable	N	Mean	Median	StDev
RUNTIME	30	3.744	3.510	2.198



To be within 1 standard deviation of the mean, a run time must be in the interval (_____, _____). _____% of the run times are in that interval.

To be within two standard deviations of the mean, run time must be in the interval (_____, _____). _____% of the run times are in that interval.

To be within three standard deviations of the mean, run time must be in the interval (_____, _____). _____% of the run times are in that interval.