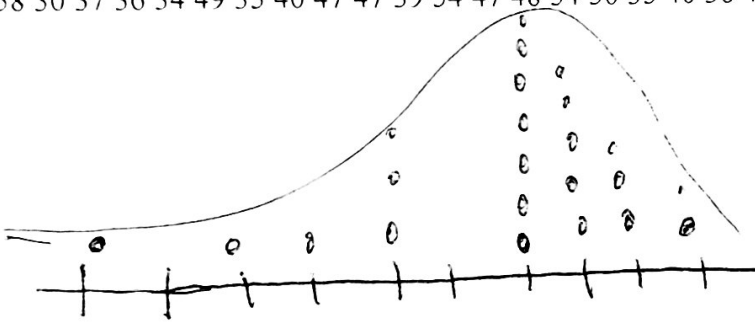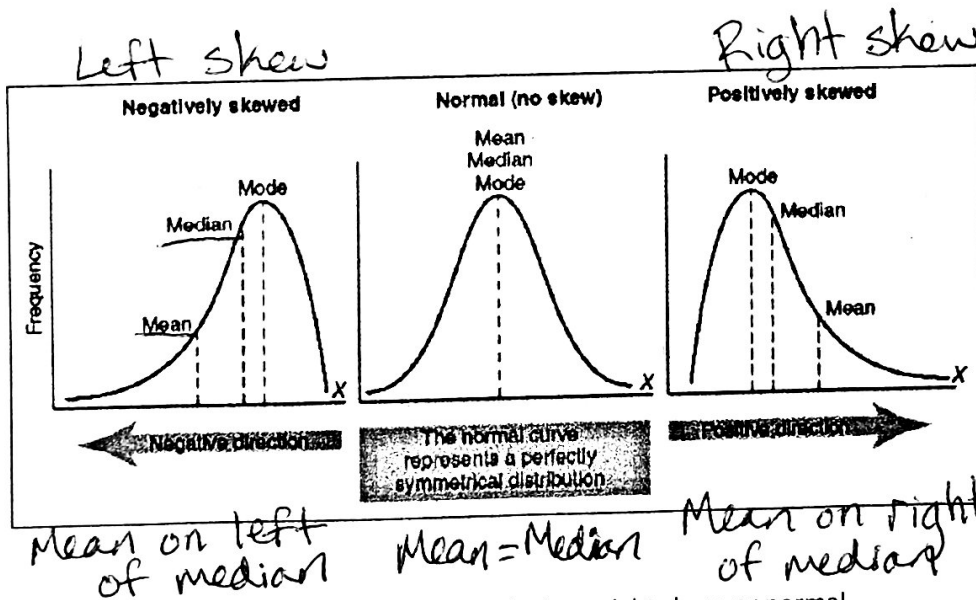# 12.3 Types of Distribution

1) Suppose you take a sample of the ages of people living in your apartment building. Make a dot plot to represent the following data:

58 30 37 36 34 49 35 40 47 47 39 54 47 48 54 50 35 40 38 47 48 34 40 46 49 47 35 48 47 46

Data clustered on right
spread out on left

Left skew

Rather than making a dot plot of your data every time, your measures of center can reveal how your data is distributed.

Left skew   Right skew

Negatively skewed	Normal (no skew)	Positively skewed

Mean on left of median   Mean = Median   Mean on right of median

2) Determine whether the distribution of each data set is left skew, right skew, or normal.

a. 7 12 3 14 17 20 5 3 17 4 13 2 15 9 15 18 16 9 1 6

$\bar{x} = 10.3$ Med = 10.5   Approximately normal

b. 6283 5700 6381 6274 5700 5896 5972 6075 5993 5581

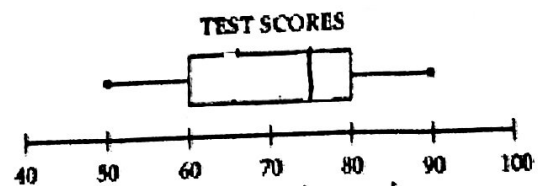$\bar{x} = 5985.5$ Med = 5982.5   Approximately normal (slightly right skew)

3) Your math class took a test. The scores are represented in the box plot below. What is the distribution? Do you feel your class understood the material? Why or why not?

25% failed
50% scored 75-90
25% scored 60-70
About average understanding

TEST SCORES

Slight left skew

Standard Deviation

The variance and standard deviation of a data set measure how much your data is spread out from the mean.

$$\text{Variance: } \sigma^2 = \frac{\sum(x-x)^2}{n} \qquad \text{Standard Deviation: } \sigma = \sqrt[2]{\frac{\sum(x-\overline{x})^2}{n}}$$

Finding Variance and Standard Deviation
*Find the mean, $\overline{x}$, of the n values in a data set
*Find the difference, $x - \overline{x}$, between each value x and the mean.
*Find the average (mean) of these squares. This is the variance:
*Take the square root of the variance. This the standard deviation:

4) Find the mean, variance, and standard deviation of the following:

a) 6.9 8.7 7.6 4.8 9.0

x	$\overline{x}$	$x - \overline{x}$	$(x - \overline{x})^2$

c) 52 63 65 77 80 82

x	$\overline{x}$	$x - \overline{x}$	$(x - \overline{x})^2$

Ain't nobody got time for this.

Finding the Standard Deviation using a Calculator:

Step 1: Use **STAT EDIT** to enter the data in list L1
Step 2: In **STAT CALC** select the **1-Var Stats** options

1 – Var Stats
$\overline{x}$ = mean
σx = standard deviation

5) Find the mean, median, and standard deviation of each.

a. The Dow Jones Industrial average for the first 12 weeks of 1922:

| 1911.31	1956.07	1903.51	1958.22	1910.48	1983.26	2014.59	2023.21	2057.86	2034.98	2087.37	2067.14

$\overline{x} = 1992.33 \qquad \sigma = 61.85$

b. The Dow Jones Industrial average for the first 12 weeks of 2008:

12800.18	12606.3	12099.3	12207.17	12743.19	12182.13	12348.21	12381.02	12266.39	11893.69	11951.09	11972.25

Normal Distribution

In a normal distribution, all of the data is centered around the mean (which is also the median and the mode).

Memorize percentages
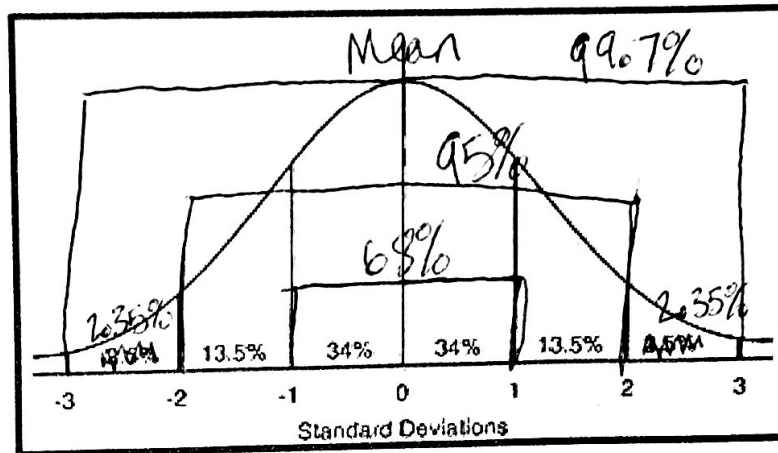
We can use what is called the **empirical rule** to make predictions about the data using the standard deviation. This rule came from the area under the curve for each section of the normal distribution.

- __68__ % of the data falls within __1__ standard deviation of the mean, with __34__ % above the mean and __34__ % below the mean
- __95__ % of the data falls within __2__ standard deviation of the mean, with __47.5__% above the mean and __47.5__% below the mean
- __99.7__% of the data falls within __3__ standard deviation of the mean, with __49.9__% above the mean and __49.9__% below the mean

6) The table displays the number of hurricanes in the Atlantic Ocean from 1992 to 2006.

Year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Number	4	4	3	11	10	3	10	8	8	9	4	7	9	14	5

Source: National Hurricane Center

a. Find the mean and standard deviation of the data.
b. Draw a bell curve to represent the data. How many standard deviations away from the mean does all the data fall?

Standard Deviations Away: $\dfrac{x - \bar{x}}{\sigma}$ $\left(\dfrac{\text{data point} - \text{mean}}{SD} \right)$
from mean

7) Determine how many standard deviations away from the mean the min and max fall. What kind of distribution is it based on that information?

The mean length of Beethoven's nine symphonies is (37 minutes) the standard deviation is 12 minutes. 27 min, 30 min, 47 min, 35 min, 30 min, 40 min, 35 min, (22 min) (65 min)

What % of data is between 25 & 61 min?

81.5%

34%   34%   13.5%

1 13 25 37 49 61 73

Min: $\dfrac{22-37}{12} = -1.33$ SD's away

Slight right skew

Max: $\dfrac{65-37}{12} = 2.10$ SD's away