A Course In Business Statistics

4th Edition

Chapter 3

Describing Data Using Numerical Measures

Chapter Goals

After completing this chapter, you should be able to:

- Compute and interpret the mean, median, and mode for a set of data
- Compute the range, variance, and standard deviation and know what these values mean
- Construct and interpret a box and whiskers plot
- Compute and explain the coefficient of variation and z scores
- Use numerical measures along with graphs, charts, and tables to describe data

Descriptive Statistics

- With charts and graphs, we can have an idea about the distribution of the data. How the data is distributed among the values of the variables.
- When we are concerned with describing the characteristics of the frequency distributions, we need some measures.
- What are the characteristics?
  - Where is the center?
  - What is the range?
  - What is the shape of the distribution?

Summary Measures

- Measures of Center and Location
  - Mean, median, mode, geometric mean, midrange
- Other measures of Location
  - Weighted mean, percentiles, quartiles
- Measures of Variation
  - Range, interquartile range, variance and standard deviation, coefficient of variation

Marks of Intr.Stats of Last Year (Make-up Exam)

0	0	60	85	90	60
0	25	75	70	60	90
0	60	74	35	67	85
0	0	72	87	36	0
60	35	30	65	70	0
0	70	20	85	95	70
60	87	65	60	72	89
0	60	86	60	60	90
0	80	60	60	82	0
72	95	95	85	60	
60	70	60	85	30	
60	60	0	75	65	

Sorted Marks of Intr.Stats of Last Year (Make-up Exam)

0	0	60	60	72	87
0	20	60	65	74	87
0	25	60	65	75	89
0	30	60	65	75	90
0	30	60	67	80	90
0	35	60	70	82	90
0	35	60	70	85	95
0	36	60	70	85	95
0	60	60	70	85	95
0	60	60	70	85	
0	60	60	72	85	
0	60	60	72	86	

Construct a Grouped Frequency Distribution to Obtain a Histogram

- STURGES RULE
- Min=0
- Max=95
- N=69
- Interval width = (Max – Min) / k
- k (Sturges) = 1 + 3.3 * log(N)

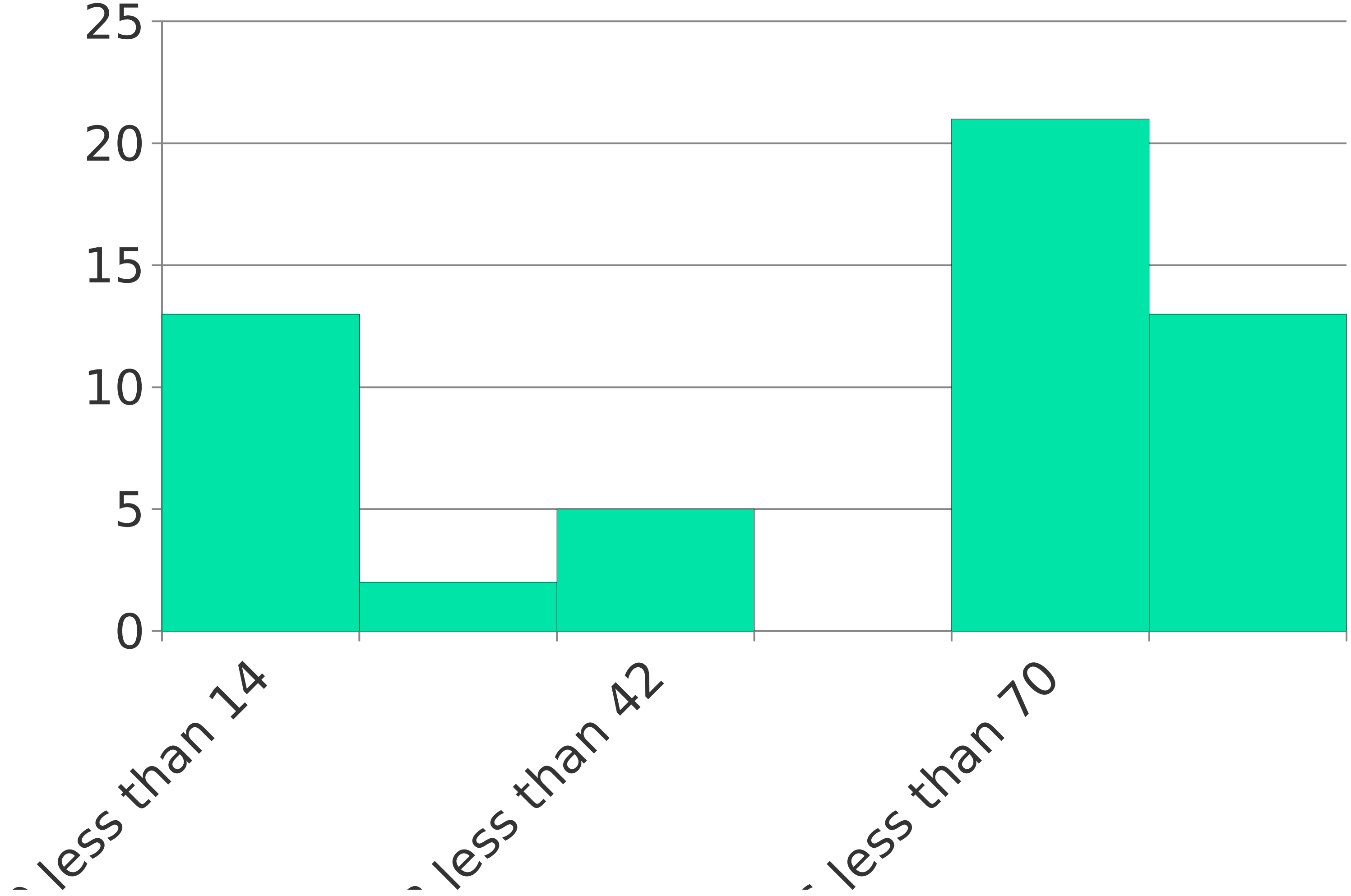
  = 1 + 3.3 * log(69) = 7.068
- Interval width = (95 – 0) / 7 = 13.57
- Take 14 as the interval width

Grouped Frequency Distribution

Marks	Number of Students
0 less than 14	13
14 less than 28	2
28 less than 42	5
42 less than 56	0
56 less than 70	21
70 less than 84	13
84 less than 98	15

Histogram

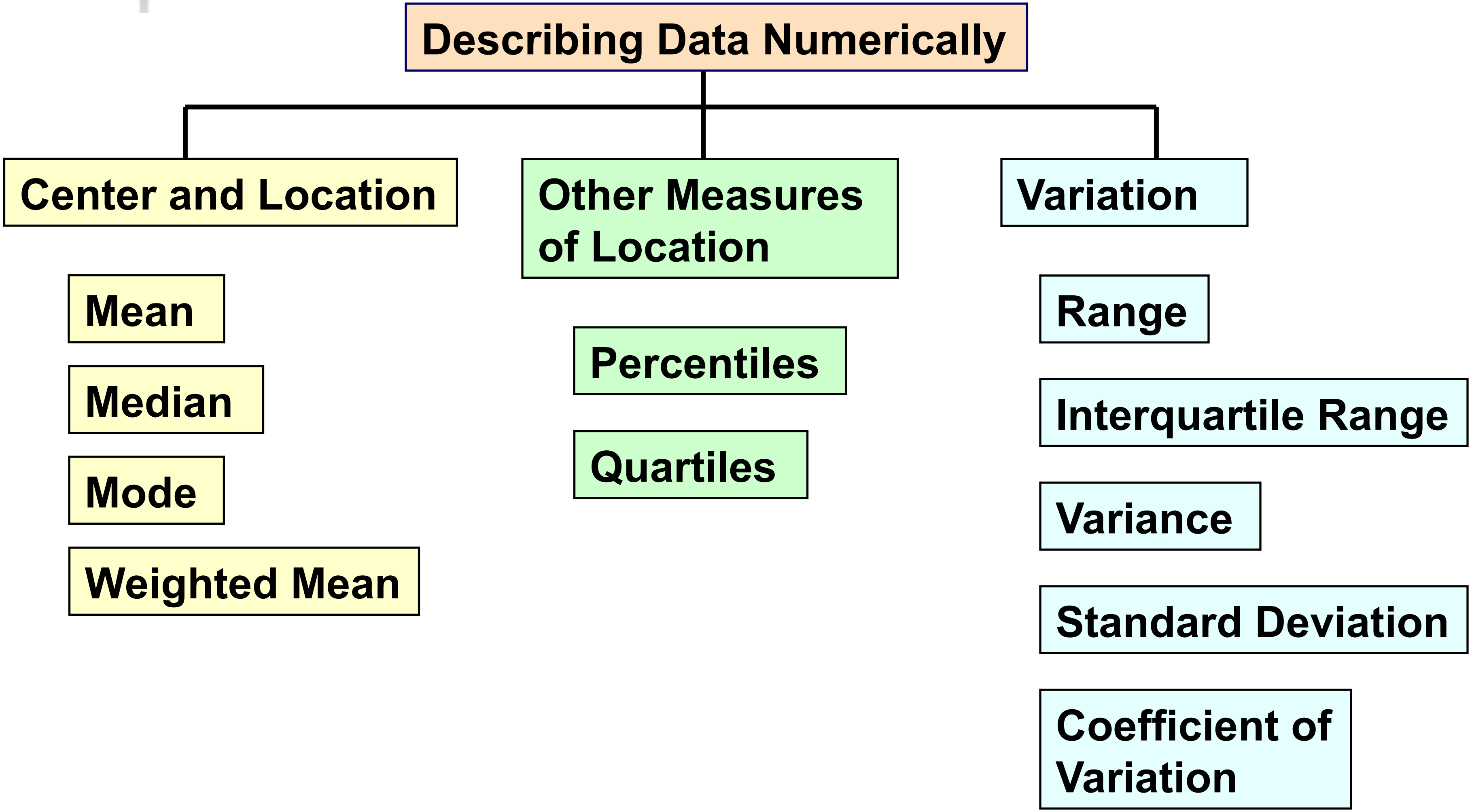
Histogram of Marks of Intr.Stats 2011-2012



No answers to the following questions

- What is the success level of students?
- How did Student A do in the exam?
- What is the success level of the make-up exam compared to final exam?
- How an individual mark is different from all the other values?
- Where the marks are gathered?
- How is the intensity of this gathering?
- All these questions are answered with measures of central tendency, variation and distributional measures.

Summary Measures



Describing Data Numerically

Center and Location

- Mean
- Median
- Mode
- Weighted Mean

Other Measures of Location

- Percentiles
- Quartiles

Variation

- Range
- Interquartile Range
- Variance
- Standard Deviation
- Coefficient of Variation

Measures of Central Tendency

- By «central tendency» we mean the tendency of the observations to centre around a particular value.

- Three measures of central tendency are commonly used in statistical analysis - the mode, the median, and the mean

- Each measure is designed to represent a typical score

Measures of Central Tendency

- the mode, the median, and the mean

- The choice of which measure to use depends on:
  - the shape of the distribution (whether normal or skewed), and
  - the variable’s “level of measurement” (data are nominal, ordinal or interval).
  - the availability of the data (open-end or closed-end data)

From the Point of Measurement Levels

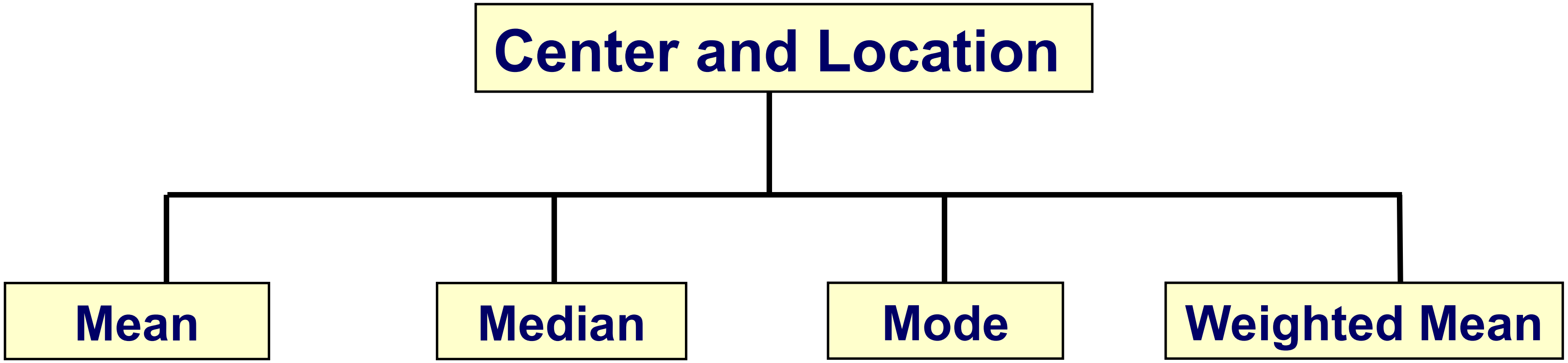
Level of measurement	Measure of central tendency
Nominal	Mode
Ordinal	Median and mode
Interval/Ratio	Mean, median and mode

Some Basic Notations

- X is a single raw score
- Xi is to the *i* th score in a set
- $X\ n$ is the last score in a set
- Set consists of $X\ 1$, $X\ 2$, ….Xn
- $\sum X = X\ 1 + X\ 2 + ... + X\ n$

Measures of Center and Location

Overview



Center and Location

Mean

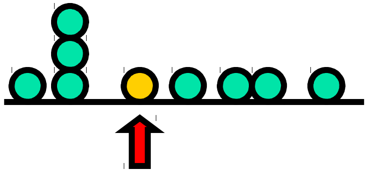
Median

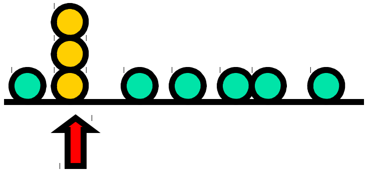
Mode

Weighted Mean

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$$





$$\bar{X}_W = \frac{\sum w_i x_i}{\sum w_i}$$

$$\mu_W = \frac{\sum w_i x_i}{\sum w_i}$$

Mean (Arithmetic Average)

- The Mean is the arithmetic average of data values
  - Sample mean

n = Sample Size

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

  - Population mean

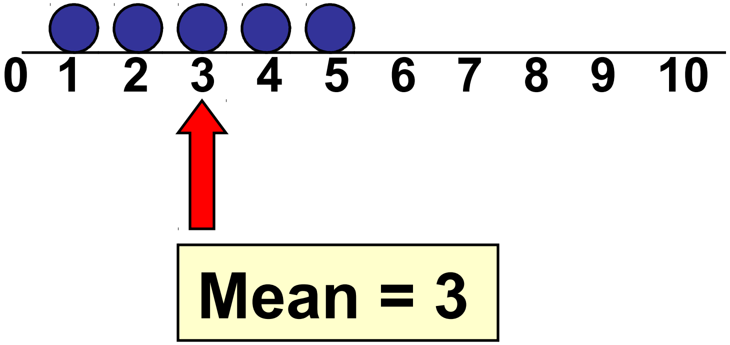
N = Population Size

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} = \frac{x_1 + x_2 + \cdots + x_N}{N}$$

Mean (Arithmetic Average)

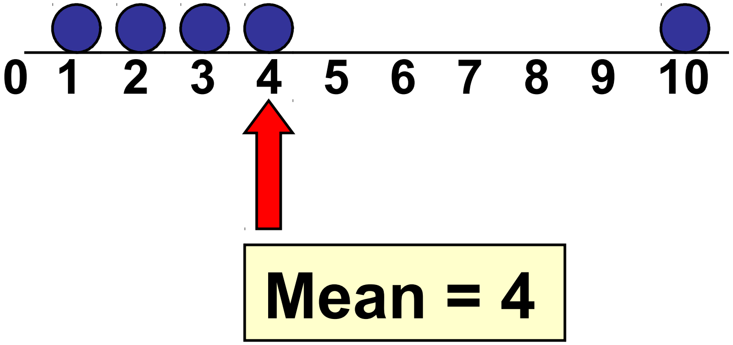
(continued)

- The most common measure of central tendency
- Mean = sum of values divided by the number of values
- Affected by extreme values (outliers)



Mean = 3

$$\frac{1+2+3+4+5}{5}=\frac{15}{5}=3$$



Mean = 4

$$\frac{1+2+3+4+10}{5}=\frac{20}{5}=4$$

Mean (Ungrouped Frequency Distribution)

- Sample mean

$$\bar{X} = \frac{\sum_{i=1}^{k} f_i X_i}{\sum_{i=1}^{k} f_i} = \frac{\sum_{i=1}^{k} f_i X_i}{n}$$

- Population mean

$$\mu = \frac{\sum_{i=1}^{k} f_i X_i}{\sum_{i=1}^{k} f_i} = \frac{\sum_{i=1}^{k} f_i X_i}{N}$$

Mean (Grouped Frequency Distribution)

- Sample mean

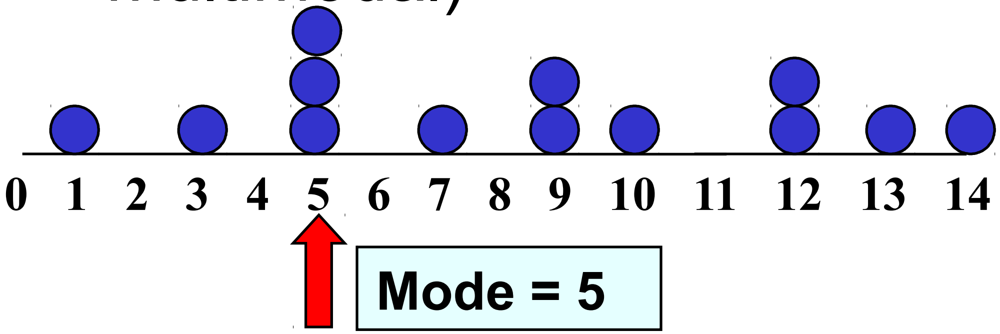
$$\bar{X} \approx \frac{\sum_{i=1}^{k} f_i m_i}{\sum_{i=1}^{k} f_i} = \frac{\sum_{i=1}^{k} f_i m_i}{n}$$

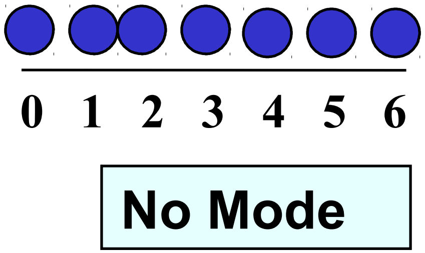
- Population mean

$$\mu \approx \frac{\sum_{i=1}^{k} f_i m_i}{\sum_{i=1}^{k} f_i} = \frac{\sum_{i=1}^{k} f_i m_i}{N}$$

Mode

- Value that occurs most often. The most common observation in a group of scores
- Not affected by extreme values
- Can be obtained for open-end data
- Used for either numerical or categorical data
- There may be no mode
- There may be several modes (unimodal, bimodal, or multimodal)





Mode (Categorical Data)

- If the data is categorical (measured on the nominal scale) then *only* the mode can be calculated.
- The most frequently occurring score (mode) is Vanilla.

Flavor	***f***
Vanilla	28
Chocolate	22
Strawberry	15
Neapolitan	8
Butter Pecan	12
Rocky Road	9
Fudge Ripple	6

Mode (Ungrouped Data)

- The mode can also be calculated with ordinal and higher data, but it often is not appropriate.
  - If other measures can be calculated, the mode would never be the first choice!
- 7, 7, 7, 20, 23, 23, 24, 25, 26 has a mode of 7, but obviously it doesn’t make much sense.

Mode (Grouped Data)

$$Mode \cong l_1 + \frac{\Delta_1}{\Delta_1 + \Delta_2} * i_{mode}$$

To be able to use this formula in a grouped frequency distribution, the 2 **interval widths** before and after the mode interval must be identical.

Median

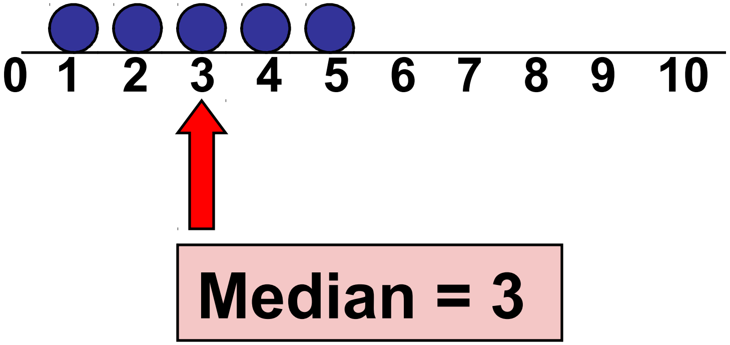
- The number that divides a distribution of scores exactly in half.
- Better than mode because only one score can be median and the median will usually be around where most scores fall.
- If data are perfectly normal, the mode is the median.
- The median is computed when data are ordinal scale or when they are highly skewed.
- Not affected by extreme values

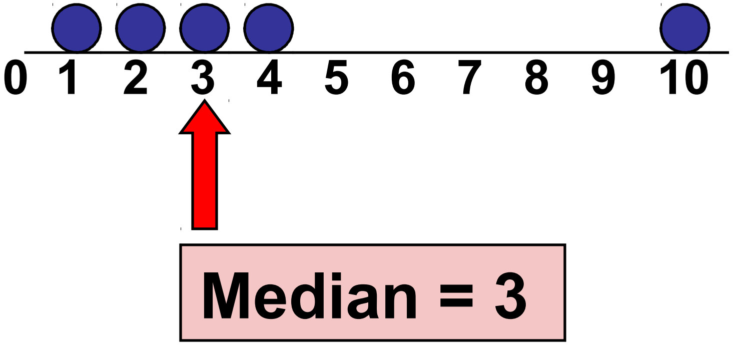
Median (Raw Data and Ungrouped Freq)

- Median is the (N+1)/2 nd value in the data.
- FIRSTLY, arrange all values in rank order from low to high (min to max)
  - If you have an odd number of scores pick the middle score. (7+1)/2 = 4th value is the median:
    - 1 4 6 **8** 12 14 18
    - Median is 8
  - If you have an even number of scores, take the average of the middle two. (8+1)/2 = 4.5 (4th and 5th values)
    - 1 4 6 **7** **8** 12 14 16
    - Median is (7+8)/2 = 7.5

Median (Raw Data and Ungrouped Freq)

- Not affected by extreme values





- In an ordered array, the median is the “middle” number
  - If n or N is odd, the median is the middle number
  - If n or N is even, the median is the average of the two middle numbers

Median (Grouped Freq)

- Median is the N/2 nd value in the data.

$$Med \cong \ l_1 + \frac{\frac{N}{2} - \sum_{i=1}^{med-1} f_i}{f_{med}} * i_m$$

Which measure of location is the “best”?

If data are perfectly normal, then the mean, median and mode are exactly the same.

I would prefer to use the mean whenever possible since it uses information from EVERY score. BUT...

Mean is generally used, unless extreme values (outliers) exist

Then **median** is often used, since the median is not sensitive to extreme values.

Example: Median home prices may be reported for a region – less sensitive to outliers

An example

- Consider these means for weekly candy bar consumption.

X = {7, 8, 6, 7, 7, 6, 8, 7}

$\bar{X}$ = (7+8+6+7+7+6+8+7)/8

$\bar{X}$ = 7

X = {12, 2, 0, 14, 10, 9, 5, 4}

$\bar{X}$ = (12+2+0+14+10+9+5+4)/8

$\bar{X}$ = 7

- On Friday, we will do some practical exercises on mean, mode and median.
- See you on Friday.

- Today we will see the other measures of central tendency such as:
  - Quartiles – Percentiles
  - Weighted Mean
  - Geometric Mean

and we will see when we use these measures.

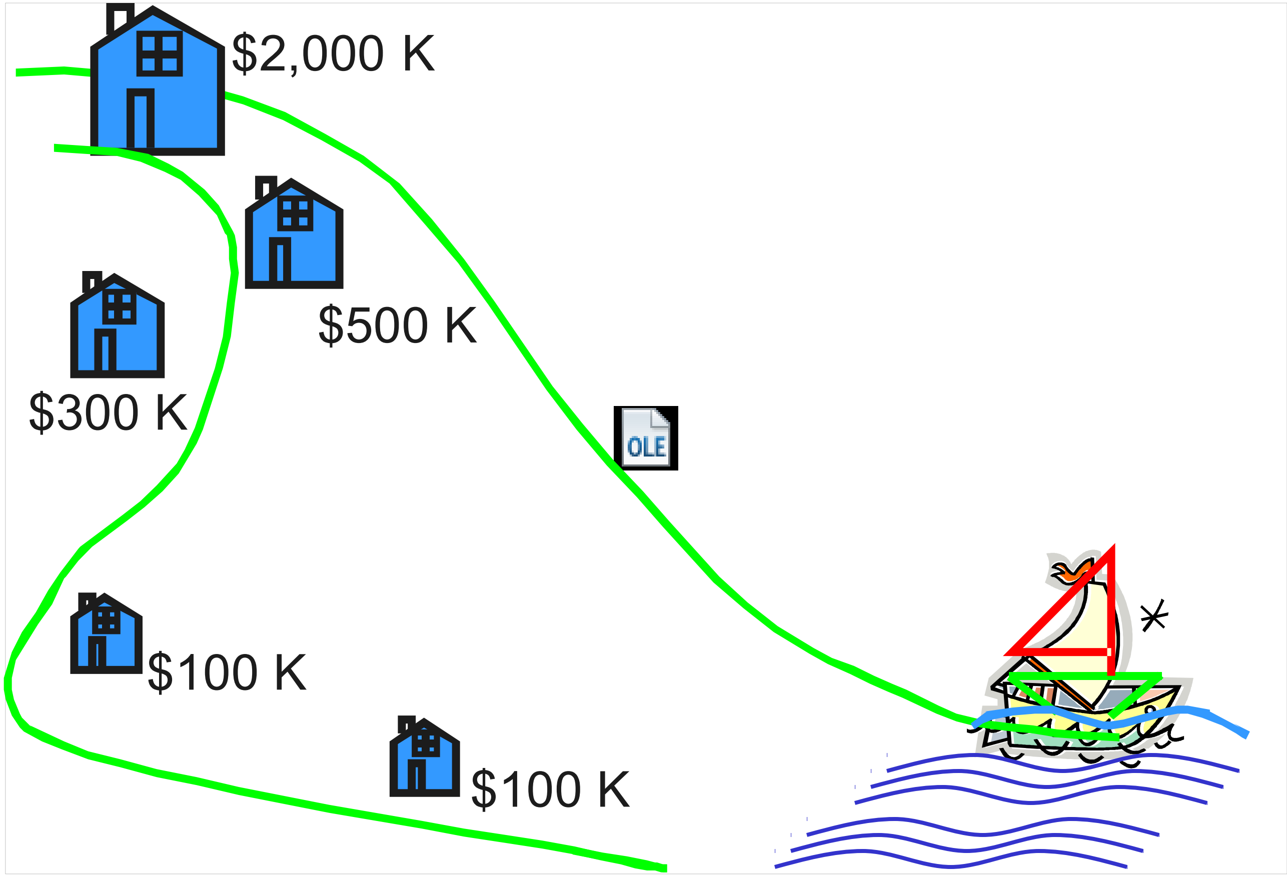
Remember last week we learned about MEAN, MEDIAN and MODE?

Review Example

Five houses on a hill by the beach

House Prices:

$2,000,000
500,000
300,000
100,000
100,000



Summary Statistics

House Prices:

$2,000,000
500,000
300,000
100,000
100,000

Sum **3,000,000**

Mean: ($3,000,000/5)
= **$600,000**

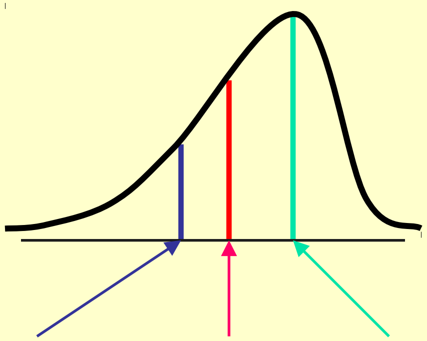
Median: middle value of ranked data
= **$300,000**

Mode: most frequent value
= **$100,000**

Shape of a Distribution

- Describes how data is distributed
- Symmetric or skewed

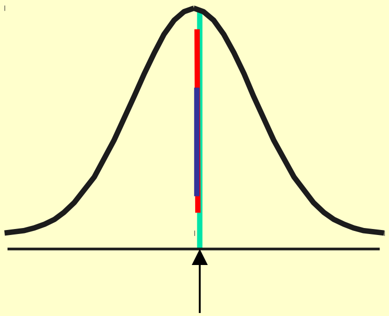
Left-Skewed



Mean < Median < Mode

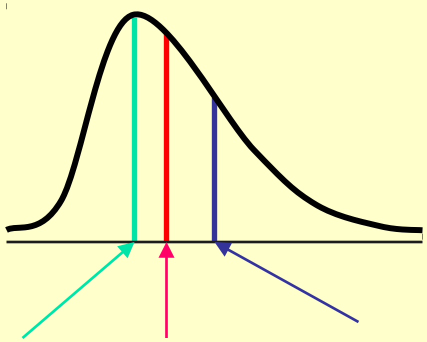
(Longer tail extends to left)

Symmetric



Mean = Median = Mode

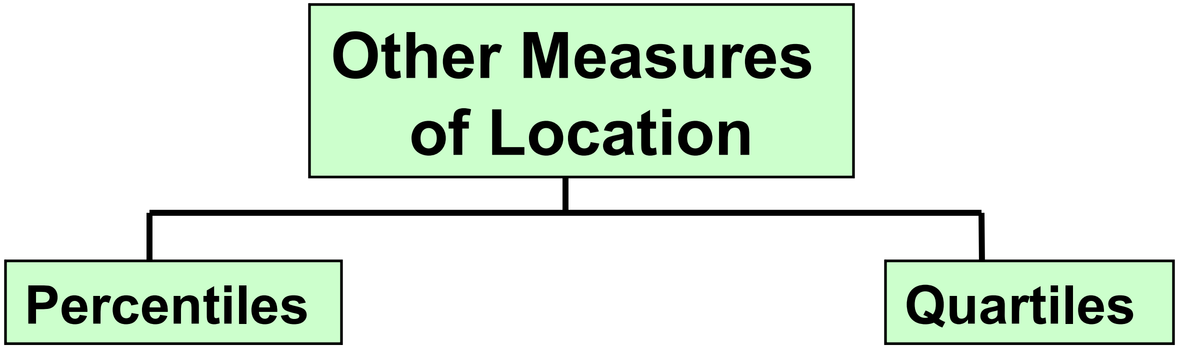
Right-Skewed



Mode < Median < Mean

(Longer tail extends to right)

Other Location Measures



Other Measures of Location

Percentiles

Quartiles

The pth percentile in a data array:

- p% are less than or equal to this value
- (100 – p)% are greater than or equal to this value

(where $0 \le p \le 100$)

- 1st quartile = 25th percentile
- 2nd quartile = 50th percentile
= median
- 3rd quartile = 75th percentile

Percentiles

- For a set of measurements arranged in increasing order, the ***p*th percentile** is a value such that *p* percent of the measurements fall at or below the value, and (100-*p*) percent of the measurements fall at or above the value.

How to calculate the pth percentile

- **Step 1:** Arrange the measurements in increasing order.
- **Step 2:** Calculate the index

$$\boldsymbol{i = \left(\frac{p}{100}\right)(n+1)}$$

- **Step 3: (a)** If *i* is an integer, this integer denotes the position of the *p*th percentile in the ordered arrangement.
- **(b)** If *i* is not an integer, the *p*th percentile is the average of the measurements in positions *i* and *i+1* in the ordered arrangement.

Percentiles

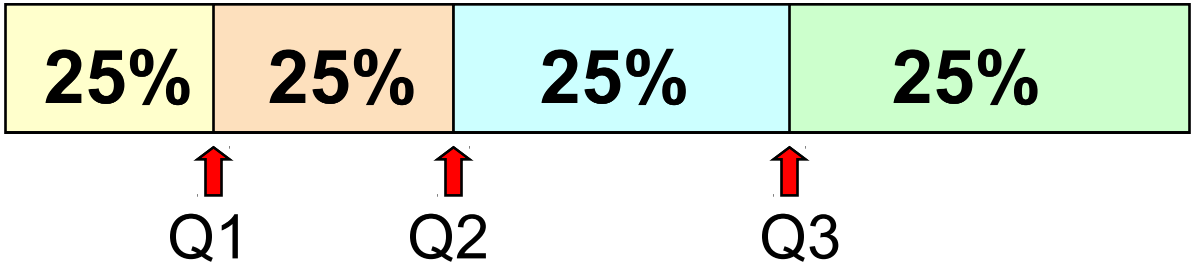
- An economist has randomly selected a sample of n=12 households from a Midwestern city and has determined last year's income for each household as follows:

7,524 11,070 18,211 26,817 36,551
41,286 49,312 57,283 72,814 90,416
135,540 190,250

  - Example: The 50th percentile in an ordered array of 12 values is the value in the 6th and 7th positions: (since 6.5 is not an integer we take the mean of the 6th and the 7th values)

Quartiles

- Quartiles split the ranked data into 4 equal groups



- Example: Find the first quartile

Sample Data in Ordered Array: 11 12 13 16 16 17 18 21 22

(n = 9)

Q1 = 25th percentile, so find the $\frac{25}{100}(9+1) = 2.5$ position

so use the value half way between the 2nd and 3rd values,

so **Q1 = 12.5**

Quartiles in Grouped Data

- N/4

$$Q_1 \cong l_1 + \frac{\frac{N}{4} - \sum_{i=1}^{Q_1-1} f_i}{f_{Q_1}} i_{Q_1}$$

- N/2

$$Q_2 = Med \cong l_1 + \frac{\frac{N}{2} - \sum_{i=1}^{Med-1} f_i}{f_{med}} i_{med}$$

- 3N/4

$$Q_3 \cong l_1 + \frac{\frac{3N}{4} - \sum_{i=1}^{Q_3-1} f_i}{f_{Q_3}} i_{Q_3}$$

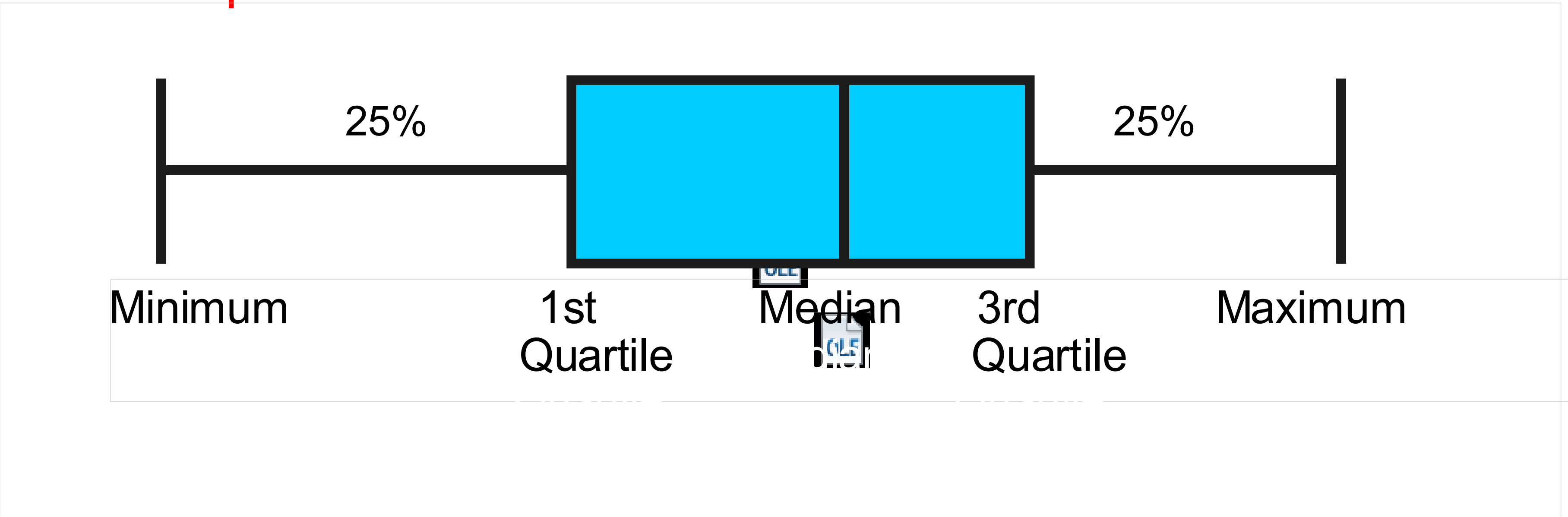
- Last week’s tutorial Example 2:
  - Salary distribution of workers ($)
  - Find the first, second and third quartiles.

Box and Whisker Plot

- A Graphical display of data using 5-number summary:

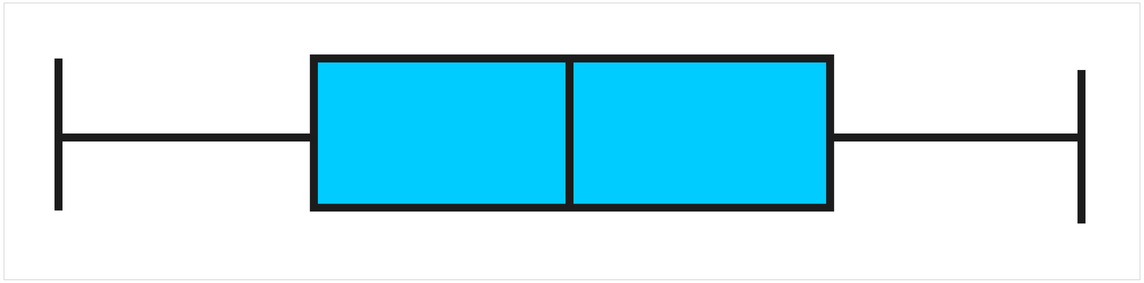
Minimum -- Q1 -- Median -- Q3 -- Maximum

Example:



Shape of Box and Whisker Plots

- The Box and central line are centered between the endpoints if data is symmetric around the median



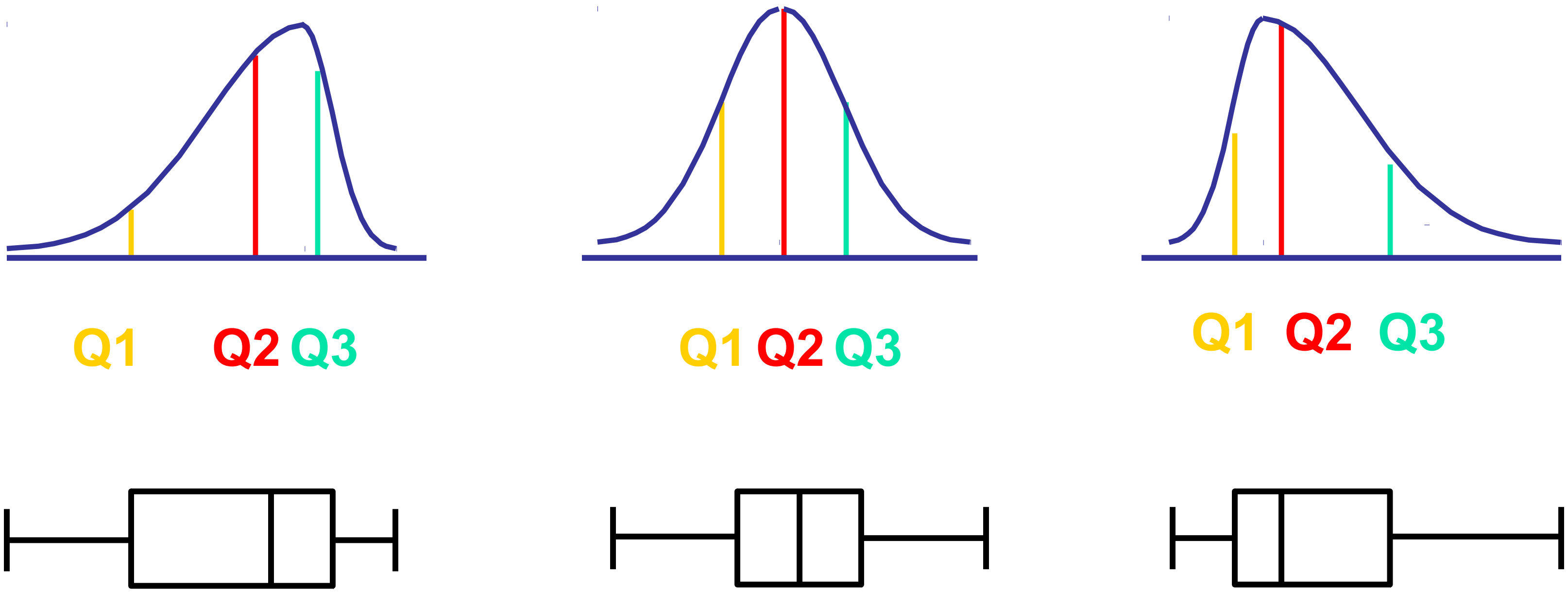
- A Box and Whisker plot can be shown in either vertical or horizontal format

Distribution Shape and Box and Whisker Plot

Left-Skewed

Symmetric

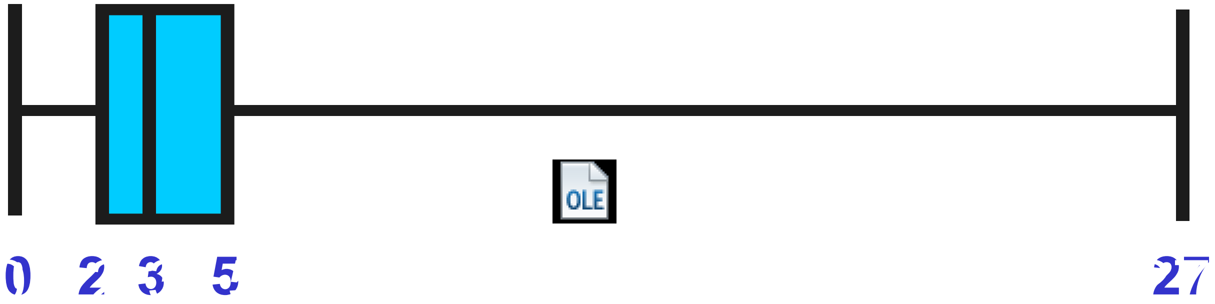
Right-Skewed



Box-and-Whisker Plot Example

- Below is a Box-and-Whisker plot for the following data:

Min		Q1			Q2			Q3		Max
0	2	2	2	3	3	4	5	5	10	27



- This data is very right skewed, as the plot depicts

Other Types of Means

- Weighted Mean
- Geometric Mean
- Harmonic Mean
- Squared Mean

Weighted Mean

- A mean where different measurements are given with different weights based on their importance

Example: AGNO Calculation

Marks	Credits
5	4
6	2
7	3
8	4

Weighted Grade Point Average:

$$\bar{X}_w = \frac{\sum w_i x_i}{\sum w_i}$$

Geometric Mean

- Used when we have a geometric series such as growth rates, population growth rates, compound interest rates...
- In such cases arithmetic mean is not a good measure.

$$G = \sqrt[N]{X_1 X_2 \cdots X_N}$$

Example

- The Standard & Poor’s 500 stock index is a commonly used measure of stock market performance in the United States. In the table below, we give the value of the S&P 500 index on the first day of market trading for each year from 2000 to 2005.

Year	S&P 500 Index
2000	1,455.22
2001	1,283.27
2002	1,154.67
2003	909.03
2004	1,108.48
2005	1,211.92

Source: http://table.finance.yahoo.com.

1) What are the percentage changes (rates of return) for the S&P 500 index for the years from 2000 to 2001 and from 2001 to 2002 and so on?

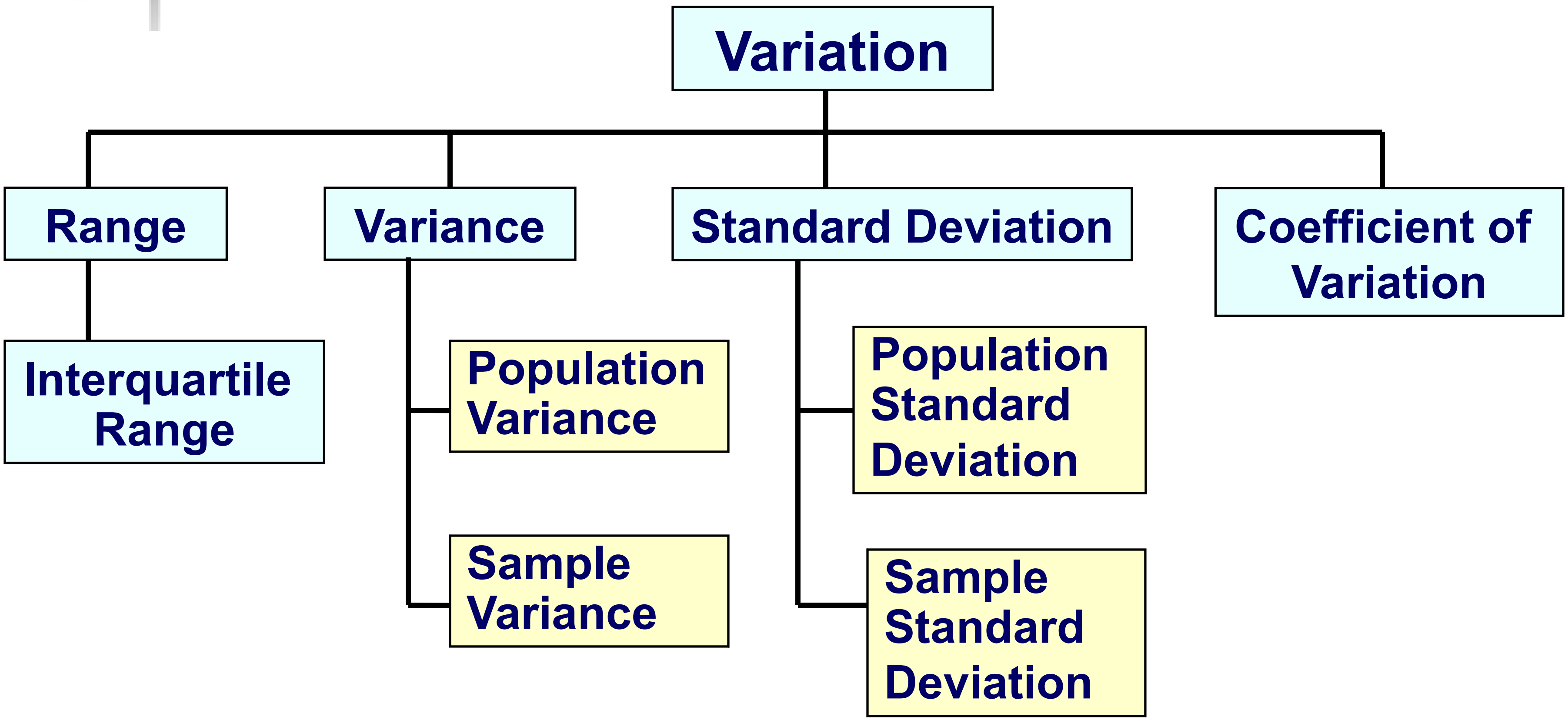
2) Calculate the mean return for the S&P 500 index over the period from 2000 to 2005.

3) Suppose that an investment of $1,000,000 is made in 2000 and that the portfolio performs with returns equal to those of the S&P 500 index. What is the investment portfolio worth in 2006?

- See you on Friday.

- Please solve the problems, at least try, given in the handouts.

Measures of Variation



Variation

Range

Interquartile Range

Variance

Population Variance

Sample Variance

Standard Deviation

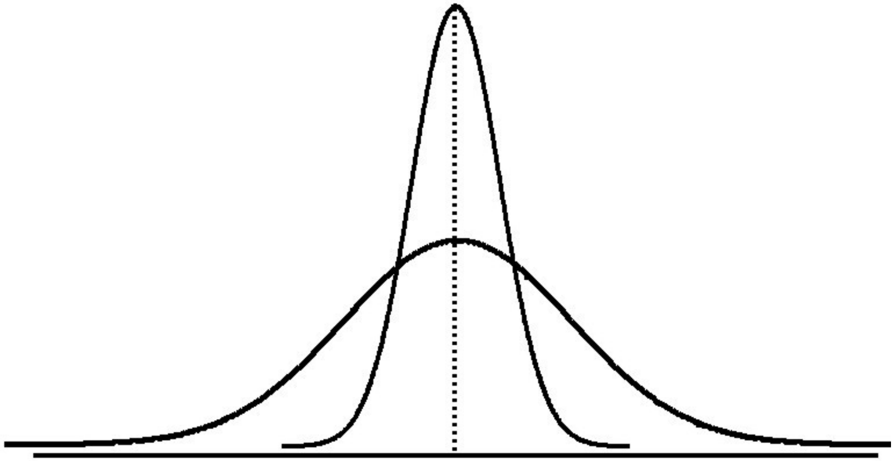
Population Standard Deviation

Sample Standard Deviation

Coefficient of Variation

Variation

Measures of variation give information on the **spread** or **variability** of the data values.



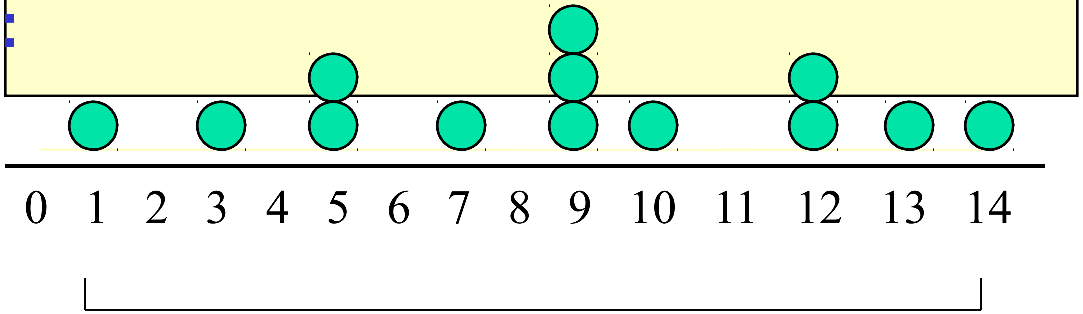
Same center, different variation

Range

- Simplest measure of variation
- Difference between the largest and the smallest observations:

Range = xmaximum – xminimum

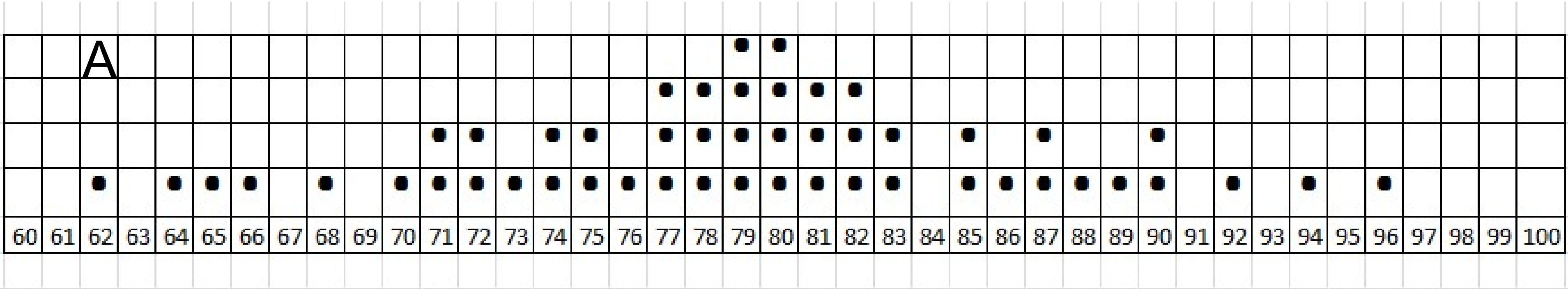
Example:

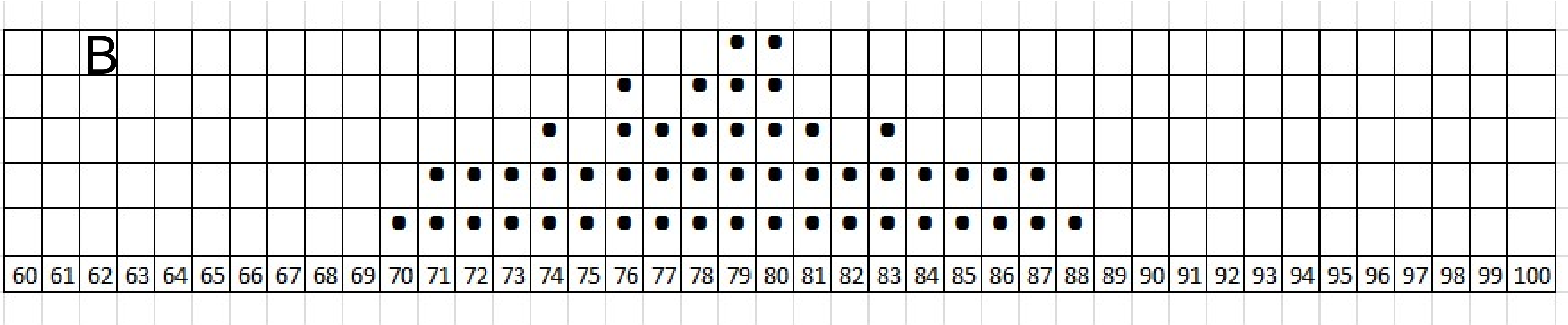


Range = 14 - 1 = 13

Compare the 2 dot-diagrams.

- What seems to you the most striking difference between the two distributions?



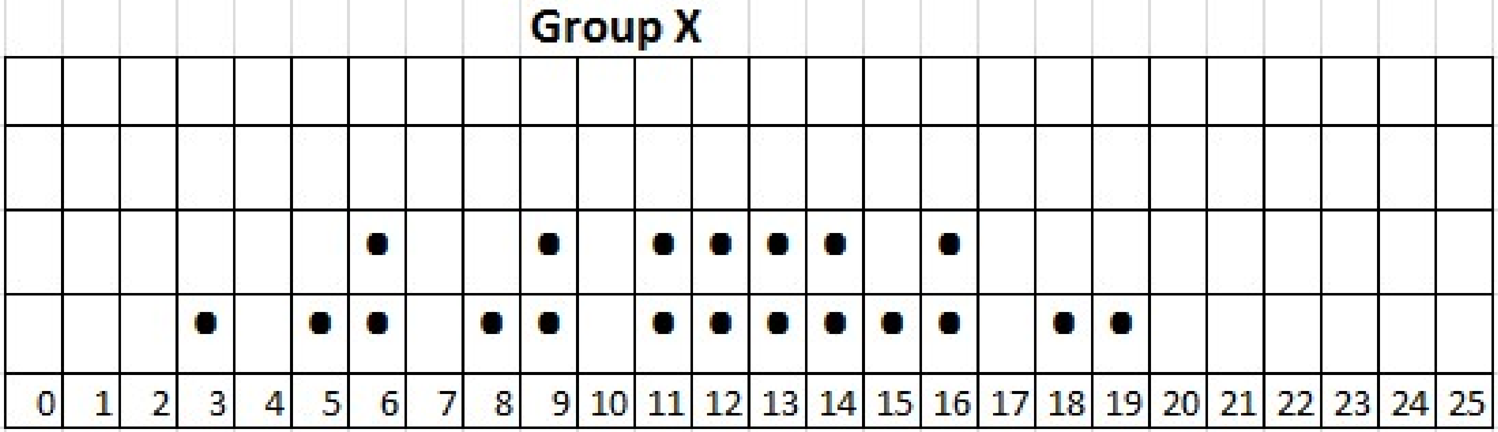


The Range

- 1) The observed values in B are far less spread out than they are in A. Compare the ranges:
  - In A, range = 96-62 = 34
  - In B, range = 88-70 = 18

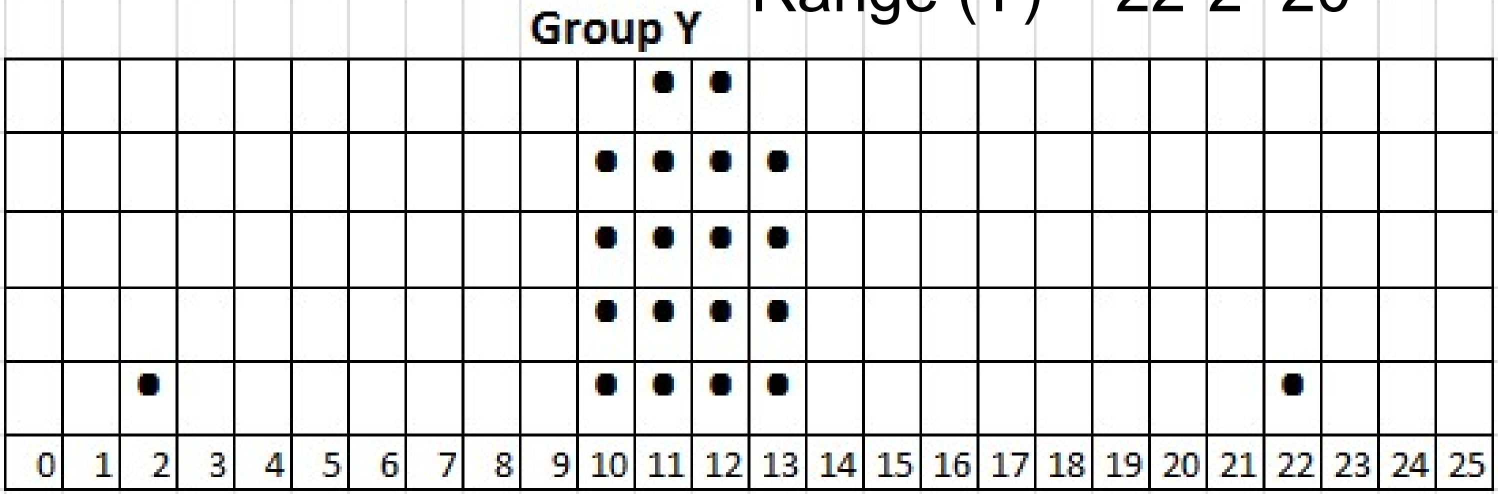
- Range is the simplest measure of dispersion.
- However we cannot trust range much. Why?
- It depends totally on just two values: The smallest and largest. These may be outliers.

Which of the two distributions is the more dispersed? Which range is bigger?



Range (X) = 19-3=16

Range (Y) = 22-2=20



Disadvantages of the Range

- Sensitive to outliers

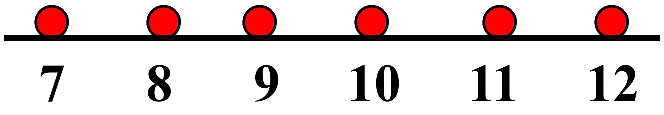
1,1,1,1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,2,3,3,3,3,4,5

Range = 5 - 1 = 4

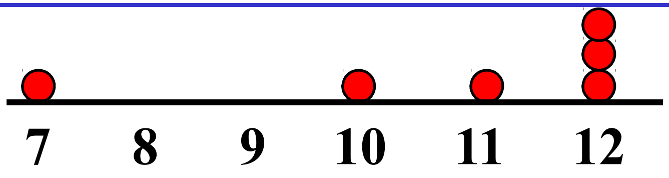
1,1,1,1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,2,3,3,3,3,4,120

Range = 120 - 1 = 119

- Ignores the way in which data are distributed



Range = 12 - 7 = 5



Range = 12 - 7 = 5

- Therefore, to overcome the disadvantages of the range we use a fairer measure of dispersion by taking a kind of «mini range» from nearer the centre of a distribution.

- This is the interquartile range.

Interquartile Range

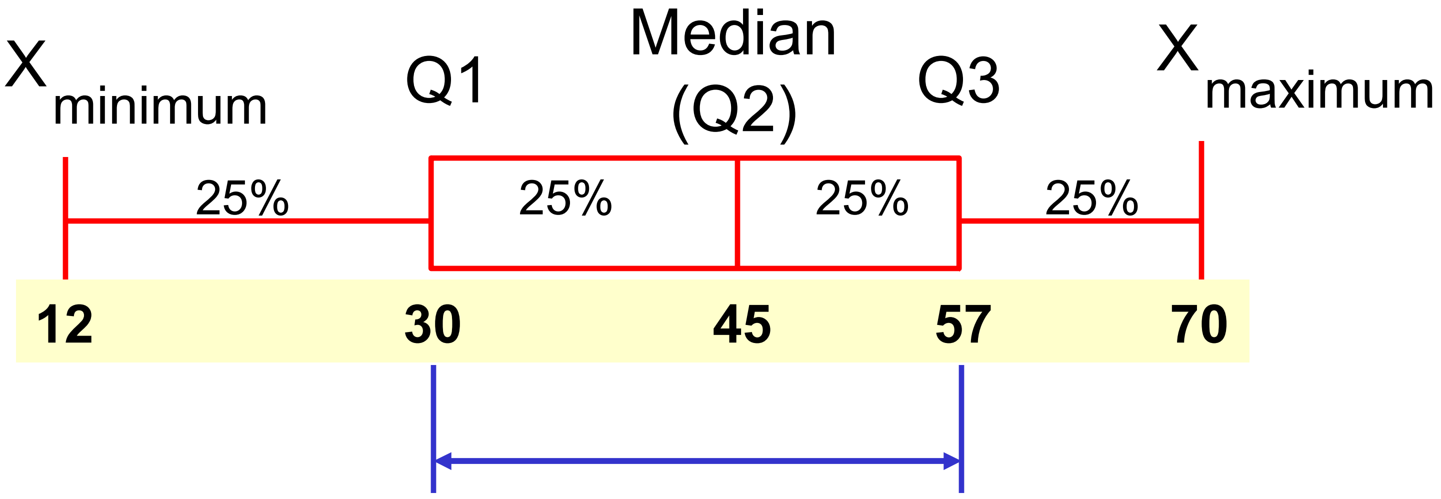
- Can eliminate some outlier problems by using the **interquartile range**

- Eliminate some high-and low-valued observations and calculate the range from the remaining values.

- Interquartile range = 3rd quartile – 1st quartile

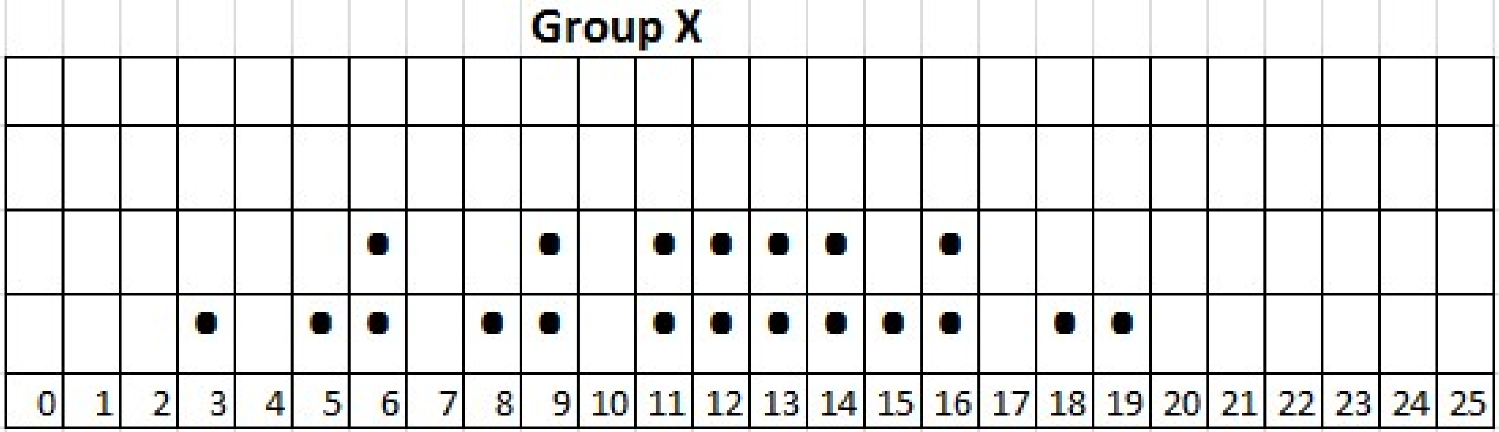
Interquartile Range

Example:



Interquartile range
= 57 – 30 = 27

Interquartile Range Versus Range

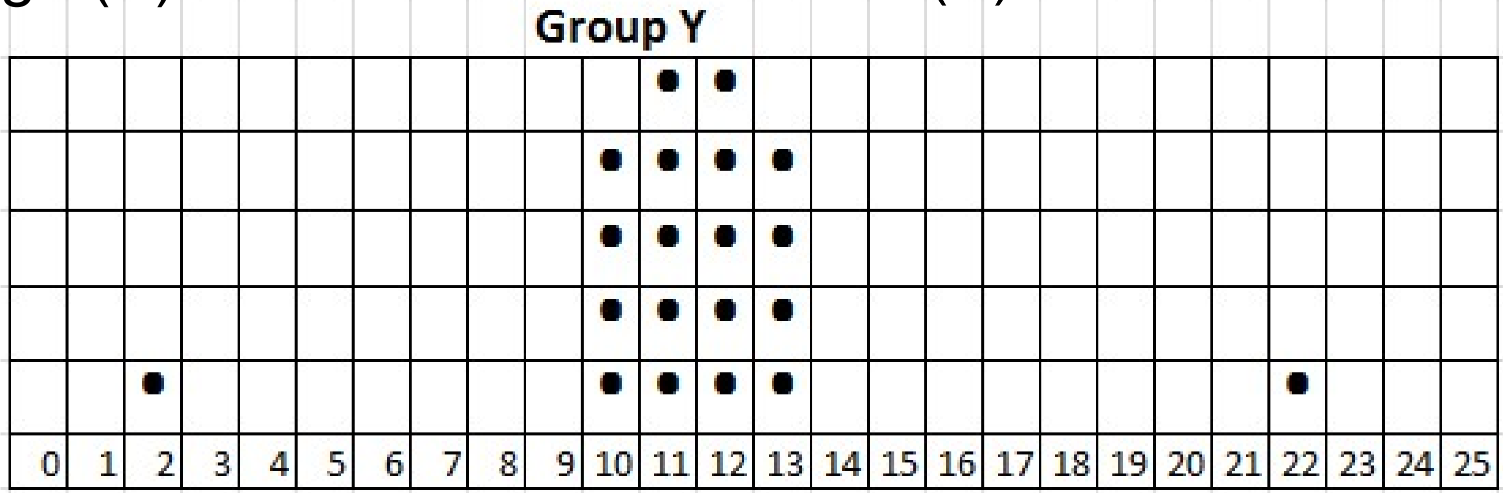


Range (X) = 19-3=16

Range (Y) = 22-2=20

IQR (X) = 14.5-8.5=6

IQR (Y) = 12.5-10.5=2



- You will no doubt agree that the inter-quartile range gives a more reasonable indication of the distribution than does the full range.

- Now we will focus on the measure of dispersion that is more commonly used than any other: the **STANDARD DEVIATION**.

- Like the mean, the standard deviation takes all the observed values into account.

Which of these sets of values would you expect to have the larger standard deviation?

- (a) 6 24 37 49 64 (mean = 36)
- (b) 111 114 117 118 120 (mean=116)

- The values in (a) are more dispersed than those in (b), so we can expect the standard deviation to be larger.
- Let’s see how it works:

- in (b) 111 114 117 118 120 (mean=116)
- Deviation from the mean 116:
- -5 -2 +1 +2 +4
- Now we cannot simply take an average of the deviations. Why?

- They add up to zero! So to overcome this we square each deviation:
- 25 4 1 4 16
- The mean of these squared deviations is called **VARIANCE**

Variance

- Average of squared deviations of values from the mean
  - **Sample** variance:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n - 1}$$

  - **Population** variance:

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}$$

- But!
- Variance is in squared values of our data:
  - Squared marks
  - Squared TL
  - Squared meters (for heigths)
  - and so on....

What would you do in this case to get the measure of dispersion back into the same units as the observed values?

Take the square root? So that is called the **STANDARD DEVIATION**

Standard Deviation

- Most commonly used measure of variation
- Shows variation about the mean
- Has the same units as the original data
  - **Sample** standard deviation:

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

  - **Population** standard deviation:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}}$$

Standard deviation (Ungrouped Frequency Distribution)

- Sample standard deviation

$$s = \sqrt{\frac{\sum_{i=1}^{k} f_i (X_i - \bar{X})^2}{n-1}}$$

- Population standard deviation

$$\sigma = \sqrt{\frac{\sum_{i=1}^{k} f_i (X_i - \bar{X})^2}{N}}$$

Standard deviation (Grouped Frequency Distribution)

- Sample standard deviation

$$s \cong \sqrt{\frac{\sum_{i=1}^{k} f_i (m_i - \bar{X})^2}{n-1}}$$

- Population standard deviation

$$\sigma \cong \sqrt{\frac{\sum_{i=1}^{k} f_i (m_i - \bar{X})^2}{N}}$$

Calculation Example: Sample Standard Deviation

Sample Data (Xi) : **10 12 14 15 17 18 18 24**

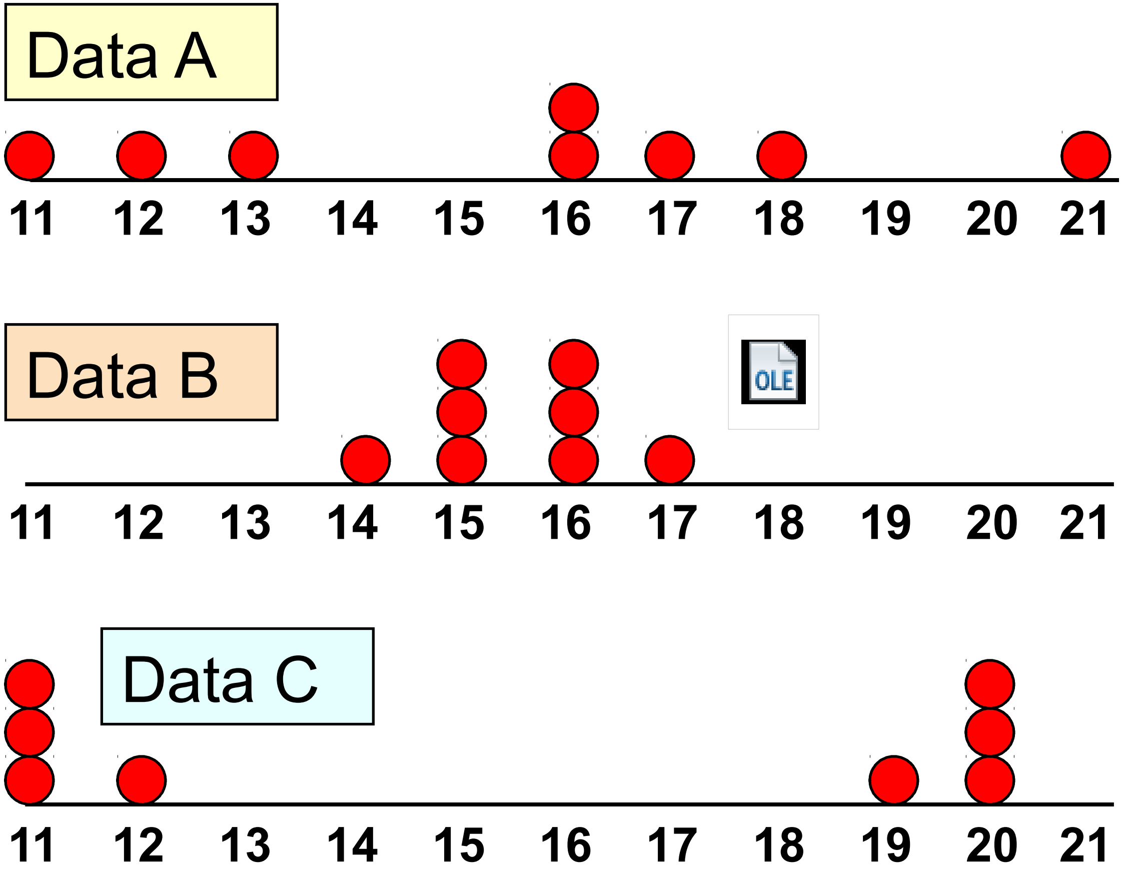
n = 8 Mean = $\bar{x}$ = 16

$$s = \sqrt{\frac{(10-\bar{x})^2 + (12-\bar{x})^2 + (14-\bar{x})^2 + \cdots + (24-\bar{x})^2}{n-1}}$$

$$= \sqrt{\frac{(10-16)^2 + (12-16)^2 + (14-16)^2 + \cdots + (24-16)^2}{8-1}}$$

$$= \sqrt{\frac{126}{7}} = \boxed{4.2426}$$

Comparing Standard Deviations



Mean = 15.5
s = 3.338

Mean = 15.5
s = .9258

Mean = 15.5
s = 4.57

Coefficient of Variation

- Measures relative variation
- Always in percentage (%)
- Shows variation relative to mean
- Is used to compare two or more sets of data measured in different units

Population

$$CV = \left(\frac{\sigma}{\mu} \right) \cdot 100\%$$

Sample

$$CV = \left(\frac{s}{\bar{x}} \right) \cdot 100\%$$

Comparing Coefficient of Variation

- Stock A:
  - Average price last year = $50
  - Standard deviation = $5

$$CV_A = \left(\frac{s}{\bar{x}}\right)\cdot 100\% \quad \frac{\$5}{\$50}\cdot 100\% = 10\%$$

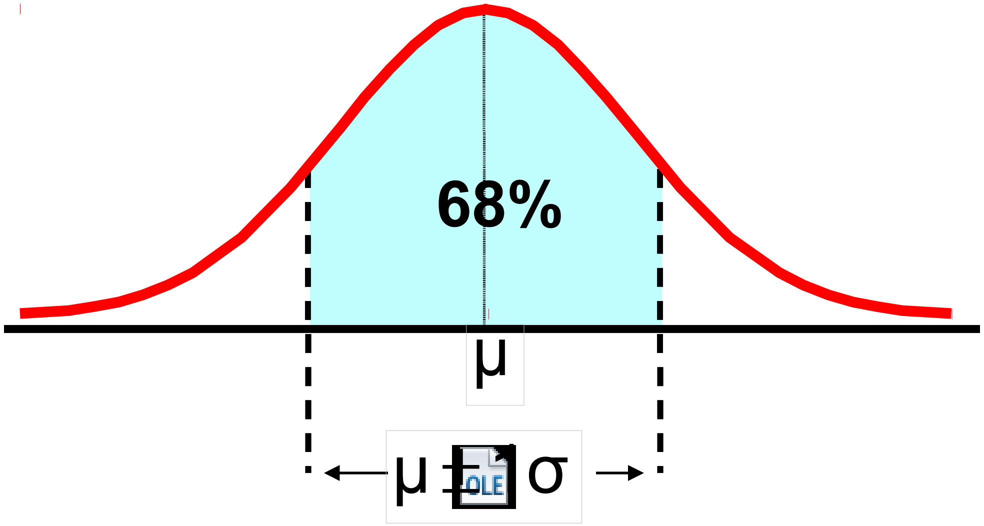
- Stock B:
  - Average price last year = $100
  - Standard deviation = $5

$$CV_B = \left(\frac{s}{\bar{x}}\right)\cdot 100\% \quad \frac{\$5}{\$100}\cdot 100\% = 5\%$$

Both stocks have the same standard deviation, but stock B is less variable relative to its price

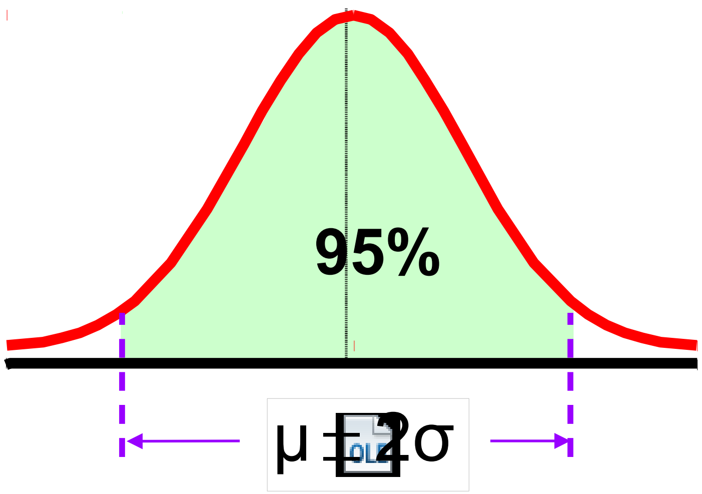
The Empirical Rule

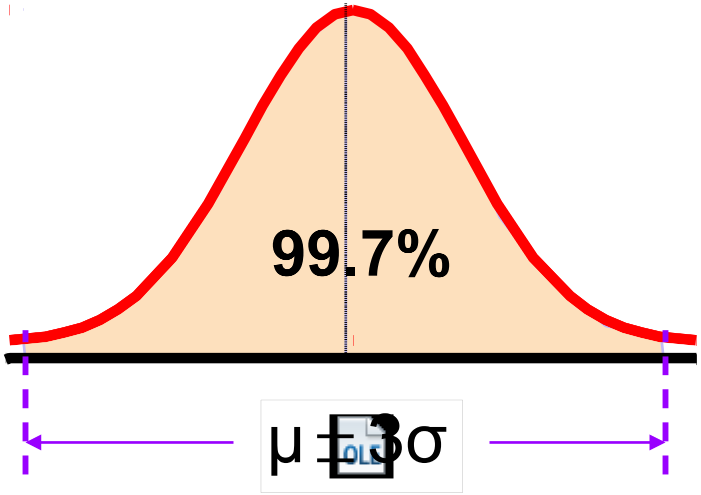
- If the data distribution is bell-shaped, then the interval:
- $\mu \pm 1\sigma$ contains about 68% of the values in the population or the sample



The Empirical Rule

- μ ± 2σ contains about 95% of the values in the population or the sample
- μ ± 3σ contains about 99.7% of the values in the population or the sample





Tchebysheff’s Theorem

- Regardless of how the data are distributed, at least (1 - 1/k2) of the values will fall within k standard deviations of the mean

  - Examples:

At least .... within

(1 - 1/12) = 0% .... k=1 $(\mu \pm 1\sigma)$

(1 - 1/22) = 75% .... k=2 $(\mu \pm 2\sigma)$

(1 - 1/32) = 89% .... k=3 $(\mu \pm 3\sigma)$

Standardized Data Values

- A standardized data value refers to the number of standard deviations a value is from the mean

- Standardized data values are sometimes referred to as z-scores

Standardized Population Values

$$z = \frac{x - \mu}{\sigma}$$

where:

- x = original data value
- μ = population mean
- σ = population standard deviation
- z = standard score

(number of standard deviations x is from μ)

Standardized Sample Values

$$z = \frac{x - \bar{x}}{s}$$

where:

- x = original data value
- $\bar{x}$ = sample mean
- s = sample standard deviation
- z = standard score

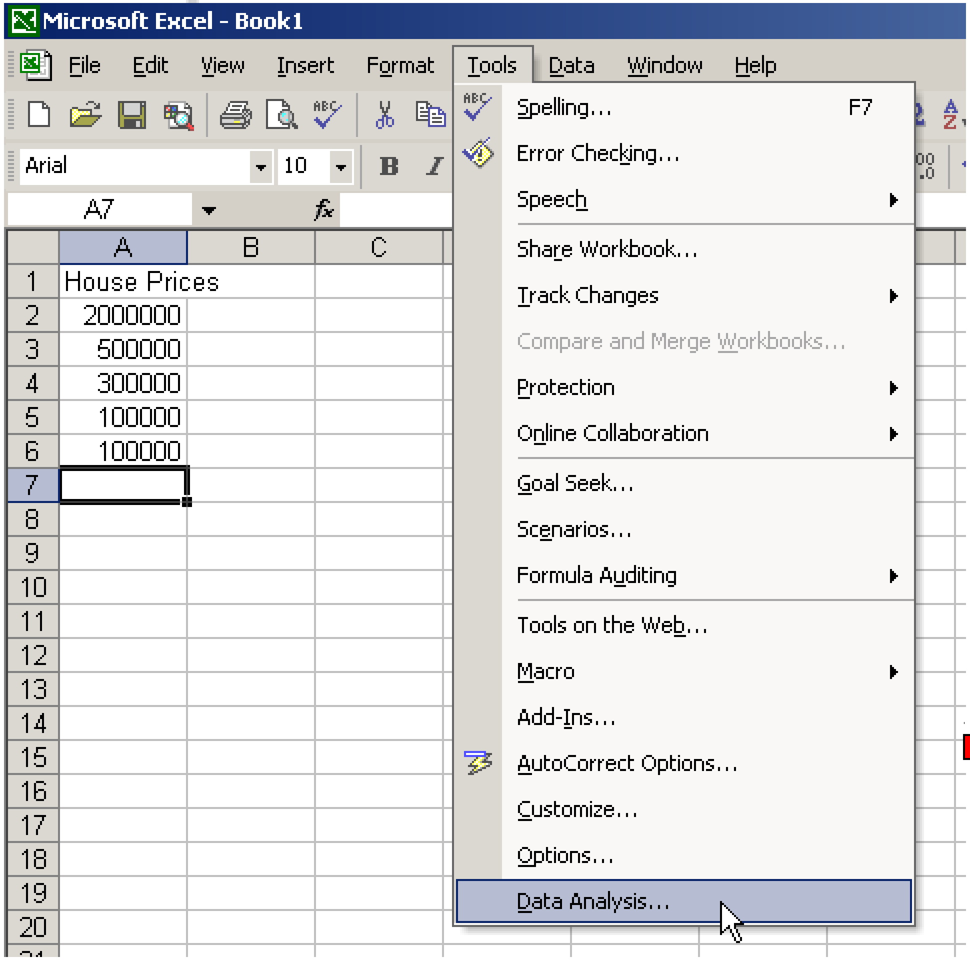
(number of standard deviations x is from μ)

Using Microsoft Excel

- Descriptive Statistics are easy to obtain from Microsoft Excel
  - Use menu choice:

    tools / data analysis / descriptive statistics
  - Enter details in dialog box

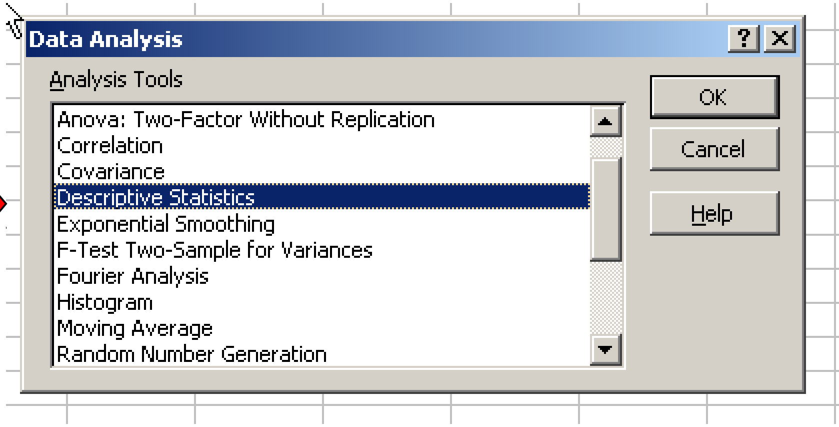
Using Excel



	A	B	C
1	House Prices		
2	2000000		
3	500000		
4	300000		
5	100000		
6	100000		
7			

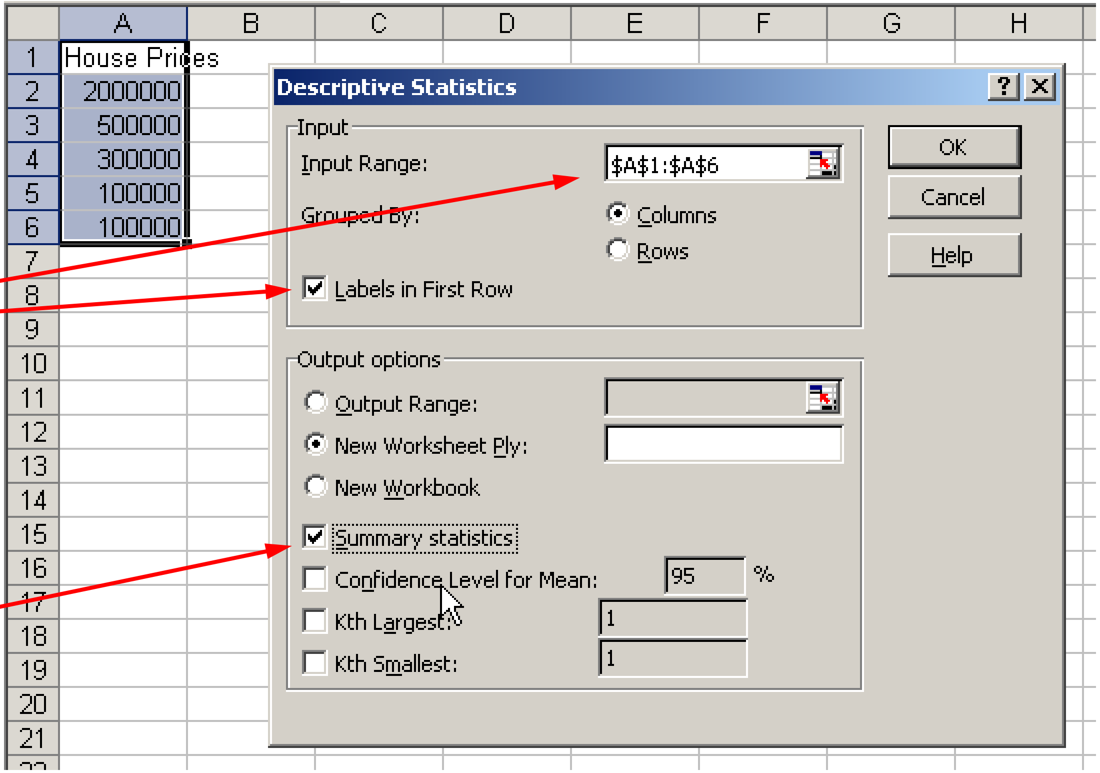
- Use menu choice:

  tools / data analysis / descriptive statistics



Using Excel

(continued)



	A
1	House Prices
2	2000000
3	500000
4	300000
5	100000
6	100000

Descriptive Statistics

Input

Input Range: $A$1:$A$6

Grouped By: ● Columns ○ Rows

☑ Labels in First Row

Output options

○ Output Range:

● New Worksheet Ply:

○ New Workbook

☑ Summary statistics

□ Confidence Level for Mean: 95 %

□ Kth Largest: 1

□ Kth Smallest: 1

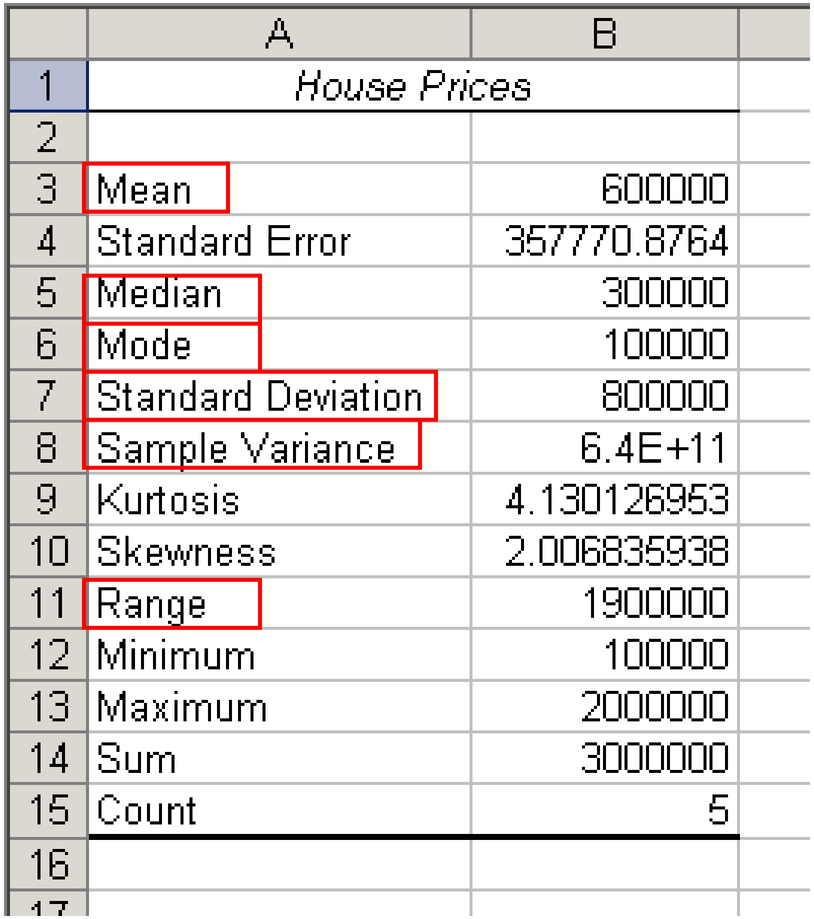
- Enter dialog box details
- Check box for summary statistics
- Click OK

Excel output

Microsoft Excel descriptive statistics output, using the house price data:

House Prices:

$2,000,000
500,000
300,000
100,000
100,000



	A	B
1	*House Prices*	
2		
3	Mean	600000
4	Standard Error	357770.8764
5	Median	300000
6	Mode	100000
7	Standard Deviation	800000
8	Sample Variance	6.4E+11
9	Kurtosis	4.130126953
10	Skewness	2.006835938
11	Range	1900000
12	Minimum	100000
13	Maximum	2000000
14	Sum	3000000
15	Count	5
16		
17		

Chapter Summary

- Described measures of center and location
  - Mean, median, mode, geometric mean, midrange
- Discussed percentiles and quartiles
- Described measure of variation
  - Range, interquartile range, variance, standard deviation, coefficient of variation
- Created Box and Whisker Plots

Chapter Summary

(continued)

- Illustrated distribution shapes
  - Symmetric, skewed
- Discussed Tchebysheff's Theorem
- Calculated standardized data values