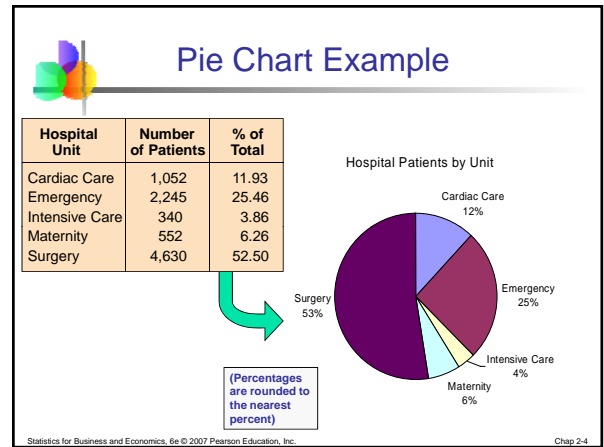
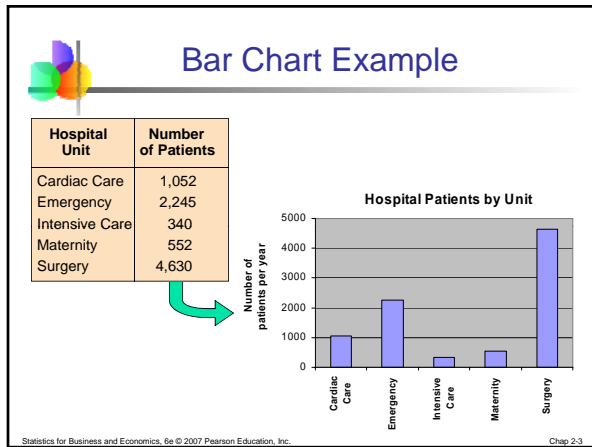
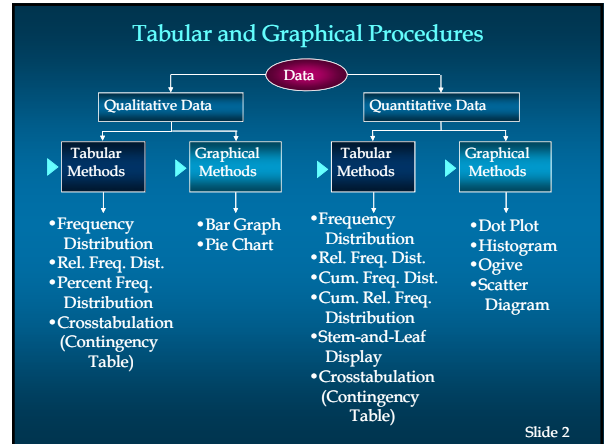


Graphical Presentations

Presented by:
Mahendra AN

Sources:
<http://business.clayton.edu/arjomand/business/stat%20presentations/bus33101.html>
http://www.cliffsnotes.com/study_guide/
 Anderson, Sweeney, Williams, *Statistics for Business and Economics*, 10 e, Thomson, 2008



Stem-and-Leaf Diagram

- A simple way to see distribution details in a data set

METHOD: Separate the sorted data series into leading digits (the **stem**) and the trailing digits (the **leaves**)

Statistics for Business and Economics, 6e © 2007 Pearson Education, Inc. Chap 2-5

Stem-and-Leaf Display

- A **stem-and-leaf display** shows both the **rank order** and **shape of the distribution** of the data.
- It is **similar to a histogram** on its side, but it has the **advantage** of showing the actual data values.
- The **first digits** of each data item are arranged to the left of a vertical line.
- To the right of the vertical line we record the **next digit** for each item **in rank order**--when the leaf unit is not shown, it is assumed to equal 1.
- Each line in the display is referred to as a **stem**.
- Each digit on a stem is a **leaf**.

Slide 6

Example

Data in ordered array:
21, 24, 24, 26, 27, 27, 30, 32, 38, 41

- Here, use the 10's digit for the stem unit:
 - 21 is shown as
 - 38 is shown as

Stem	Leaf
2	1
3	8

Statistics for Business and Economics, 6e © 2007 Pearson Education, Inc. Chap 2-7

Example (continued)

Data in ordered array:
21, 24, 24, 26, 27, 27, 30, 32, 38, 41

- Completed stem-and-leaf diagram:

Stem	Leaves
2	1 4 4 6 7 7
3	0 2 8
4	1

Statistics for Business and Economics, 6e © 2007 Pearson Education, Inc. Chap 2-8

Stretched Stem-and-Leaf Display

- If we believe the original stem-and-leaf display has condensed the data too much, we can **stretch the display** by using **two stems** for each leading digit(s).
- Whenever a stem value is stated twice, the first value corresponds to leaf values of 0 - 4, and the second value corresponds to leaf values of 5 - 9.

Slide 9

Stretched Stem-and-Leaf Display

5	2
5	7
6	2 2 2 2
6	5 6 7 8 8 8 9 9 9
7	1 1 2 2 3 4 4
7	5 5 5 6 7 8 9 9 9
8	0 0 2 3
8	5 8 9
9	1 3
9	7 7 7 8 9
10	1 4
10	5 5 9

✓ The first value corresponds to leaf values of 0 - 4, and the second value corresponds to leaf values of 5 - 9

Slide 10

Leaf Units

- A single digit is used to define each leaf.
- In the preceding example, the leaf unit was 1.
- Leaf units may be 100, 10, 1, 0.1, and so on.
- Where the leaf unit is not shown, it is assumed to equal 1.

Slide 11

Using other stem units

- Using the 100's digit as the stem:
 - Round off the 10's digit to form the leaves

Stem	Leaf
6	1
7	8
...	
12	2

 - 613 would become
 - 776 would become
 - ...
 - 1224 becomes

Statistics for Business and Economics, 6e © 2007 Pearson Education, Inc. Chap 2-12

Using other stem units (continued)

- Using the 100's digit as the stem:
 - The completed stem-and-leaf display:

Stem	Leaves
6	1 3 6
7	2 2 5 8
8	3 4 6 6 9 9
9	1 3 3 6 8
10	3 5 6
11	4 7
12	2

Statistics for Business and Economics, 6e © 2007 Pearson Education, Inc. Chap 2-13

Example: Leaf Unit = 0.1

If we have data with values such as

▶ 8.6 11.7 9.4 9.1 10.2 11.0 8.8

a stem-and-leaf display of these data will be

Leaf Unit = 0.1	
8	6 8
9	1 4
10	2
11	0 7

Slide 14

Example: Leaf Unit = 10

If we have data with values such as

▶ 1806 1717 1974 1791 1682 1910 1838

a stem-and-leaf display of these data will be

Leaf Unit = 10	
16	8
17	1 9
18	0 3
19	1 7

The 82 in 1682 is rounded down to 80 and is represented as an 8. You do this for all of the data.

Slide 15

Histogram

- A graph of the data in a frequency distribution is called a **histogram**
- The **interval endpoints** are shown on the horizontal axis
- the vertical axis is either **frequency**, **relative frequency**, or **percentage**
- Bars of the appropriate heights are used to represent the number of observations within each class

Statistics for Business and Economics, 6e © 2007 Pearson Education, Inc. Chap 2-16

Histogram Example

Interval	Frequency
10 but less than 20	3
20 but less than 30	6
30 but less than 40	5
40 but less than 50	4
50 but less than 60	2

Frequency

Temperature in Degrees

(No gaps between bars)

Statistics for Business and Economics, 6e © 2007 Pearson Education, Inc. Chap 2-17

Pareto Diagram

- Used to portray categorical data
- A bar chart, where categories are shown in descending order of frequency
- A cumulative polygon is often shown in the same graph
- Used to separate the “vital few” from the “trivial many”

Statistics for Business and Economics, 6e © 2007 Pearson Education, Inc. Chap 2-18

Pareto Diagram Example

Example: 400 defective items are examined for cause of defect:

Source of Manufacturing Error	Number of defects
Bad Weld	34
Poor Alignment	223
Missing Part	25
Paint Flaw	78
Electrical Short	19
Cracked case	21
Total	400

Statistics for Business and Economics, 6e © 2007 Pearson Education, Inc. Chap 2-19

Pareto Diagram Example

(continued)

Step 1: Sort by defect cause, in descending order
 Step 2: Determine % in each category

Source of Manufacturing Error	Number of defects	% of Total Defects
Poor Alignment	223	55.75
Paint Flaw	78	19.50
Bad Weld	34	8.50
Missing Part	25	6.25
Cracked case	21	5.25
Electrical Short	19	4.75
Total	400	100%

Statistics for Business and Economics, 6e © 2007 Pearson Education, Inc. Chap 2-20

Pareto Diagram Example

(continued)

Step 3: Show results graphically

Statistics for Business and Economics, 6e © 2007 Pearson Education, Inc. Chap 2-21

Ogive (Cumulative Line Graph)

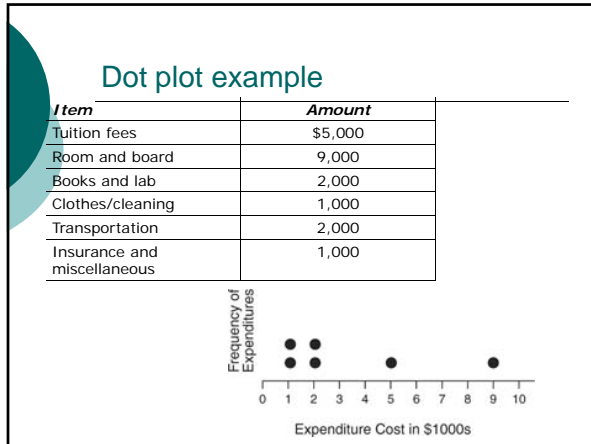
- Data may be expressed using a single line.
- An **ogive** (a cumulative line graph) is best used when you want to display the total at any given time.
- The relative slopes from point to point will indicate greater or lesser increases; for example, a steeper slope means a greater increase than a more gradual slope.
- An ogive, however, is not the ideal graphic for showing comparisons between categories because it simply combines the values in each category and thus indicates an *accumulation*, a growing or lessening total.
- If you simply want to keep track of a total and your individual values are periodically combined, an ogive is an appropriate display.

Ogive example

Statistics for Business and Economics, 6e © 2007 Pearson Education, Inc. Chap 2-22

Dot Plot

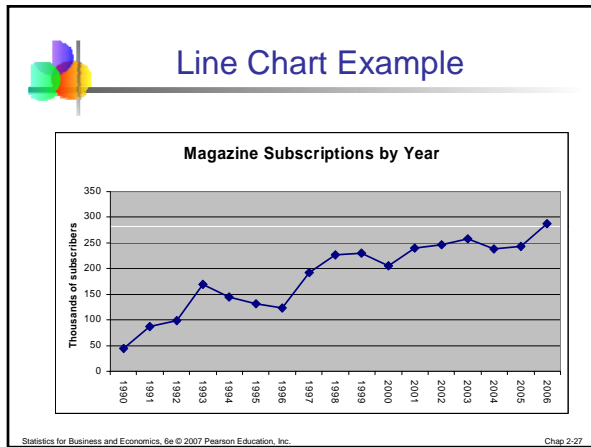
- Dot plots** are similar to bar graphs.
- Typically used for a small set of values, a dot plot uses a *dot* for each unit of measurement;



Graphs for Time-Series Data

- A **line chart** (time-series plot) is used to show the values of a variable over time
- Time is measured on the horizontal axis
- The variable of interest is measured on the vertical axis

Statistics for Business and Economics, 6e © 2007 Pearson Education, Inc. Chap 2-26



Crosstabulations and Scatter Diagrams

- ▶ As we indicated, often a manager is interested in tabular and graphical methods that will help understand the **relationship between two variables**.
- ▶ **Crosstabulation** (or **Contingency Table**) and a **scatter diagram** are two methods for summarizing the data for two (**or more**) variables simultaneously.

Slide 28

Crosstabulation Or Contingency Table

- ▶ Remember that a **crosstabulation** is a tabular summary of data for two variables.
- ▶ **Crosstabulation can be used when:**
 - one variable is qualitative and the other is quantitative,
 - both variables are qualitative, or
 - both variables are quantitative.
- ▶ As we said, the left and top margin labels define the classes for the two variables.

Slide 29

Crosstabulation Or Contingency Table

- **Example: Finger Lakes Homes**
The number of Finger Lakes homes sold for each style and price for the past two years is shown below.

Price Range	Home Style				Total
	Colonial	Log	Split	A-Frame	
< \$99,000	18	6	19	12	55
> \$99,000	12	14	16	3	45
Total	30	20	35	15	100

Slide 30

Crosstabulation Or Contingency Table

Frequency distribution for the price variable

Price Range	Home Style				Total
	Colonial	Log	Split	A-Frame	
≤ \$99,000	18	6	19	12	55
> \$99,000	12	14	16	3	45
Total	30	20	35	15	100

Frequency distribution for the home style variable

Slide 31

Crosstabulation Or Contingency Table

- Insights Gained from Preceding Crosstabulation
 - The greatest number of homes in the sample (19) are a split-level style and priced at less than or equal to \$99,000.
 - Only three homes in the sample are an A-Frame style and priced at more than \$99,000.

Slide 32

Crosstabulation Row Or Column Percentages?

- As we said, converting the entries in the table into row percentages or column percentages can provide additional insight about the relationship between the two variables.

Slide 33

Crosstabulation or Contingency Table (Row %)

Price Range	Home Style				Total
	Colonial	Log	Split	A-Frame	
≤ \$99,000	32.73	10.91	34.55	21.82	100
> \$99,000	26.67	31.11	35.56	6.67	100

Note: row totals are actually 100.01 due to rounding.

$(\text{Colonial and } > \$99\text{K}) / (\text{All } > \$99\text{K}) \times 100 = (12/45) \times 100$

(Cell Count) (100)
Row Total

Slide 34

Crosstabulation or Contingency Table (Column %)

Price Range	Home Style			
	Colonial	Log	Split	A-Frame
≤ \$99,000	60.00	30.00	54.29	80.00
> \$99,000	40.00	70.00	45.71	20.00
Total	100	100	100	100

$(\text{Colonial and } > \$99\text{K}) / (\text{All Colonial}) \times 100 = (12/30) \times 100$

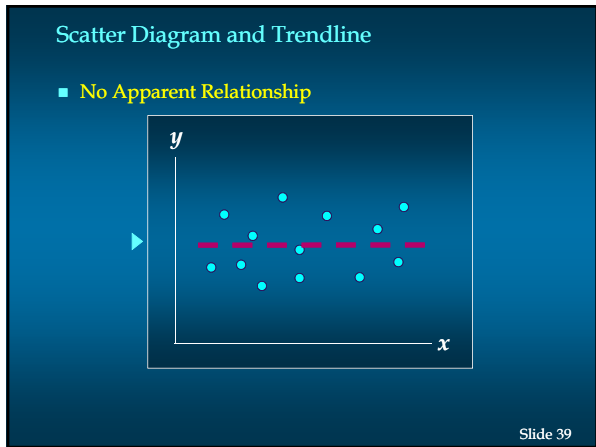
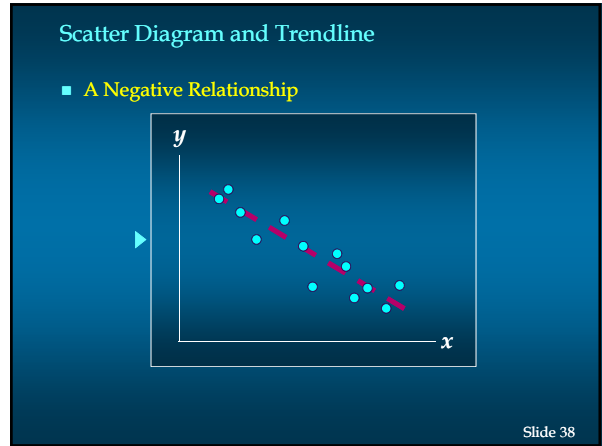
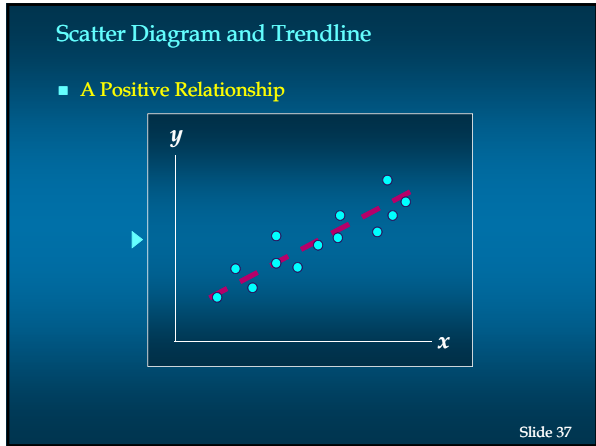
(Cell Count) (100)
Column Total

Slide 35

Scatter Diagram and Trendline

- A scatter diagram is a graphical presentation of the relationship between two quantitative variables.
- One variable is shown on the horizontal axis and the other variable is shown on the vertical axis.
- The general pattern of the plotted points suggests the overall relationship between the variables.
- A trendline is a line that provides an approximation of the relationship.

Slide 36



**“Tell me and I forget,
Teach me and I remember,
Involve me and I Learn”**

-- Ben Franklin --