



**Centurion**  
**UNIVERSITY**  
*Shaping Lives...*  
*Empowering Communities...*

## LECTURE-5, 6 & 7



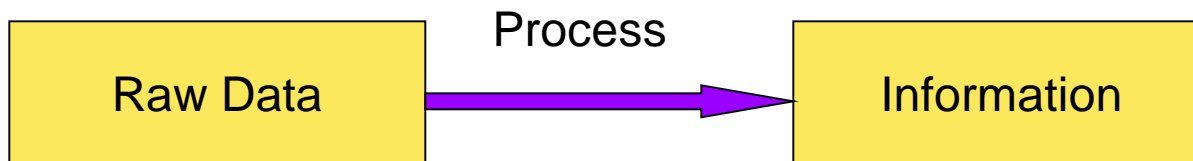
**Centurion**  
**UNIVERSITY**  
*Shaping Lives...  
Empowering Communities...*

## Organizing Numerical Data

- The Ordered Array
- Stem-and-leaf Display
- Example for the above two.



- Raw data represents numbers and facts in the original format in which data have been collected.
- Large and massive raw data tend to bewilder so much that the overall patterns are obscured.
- Thus raw data is to be processed to give useful information.





## Example of Raw Data:

The weekly sales ('000 units of a product in a region over the past year are-

52 61 59 55 63 70 59 77 81 83 69 91 73  
83 90 81 77 77 74 65 33 77 64 49 49 52  
50 45 42 46 39 29 38 41 43 23 26 27 22  
29 31 29 31 30 30 29 40 44 45 46 47 53.

# Group Data

- Dividing data into group or classes or intervals

Groups may be

- Mutually exclusive
- Exhaustive
- Equal width

# Tabulation of Numerical Data

## The frequency distribution

--- It is a summary table in which the data are arranged into numerically ordered class groups or intervals.

- Selecting the number of classes
- Obtaining the class interval
- Determining the width of the class
- Establishing the class boundaries
- Relative frequency distribution
- Cumulative distribution
- Advantage and Disadvantage.



# Frequency Distribution

- Table with two columns listing:
  - Each and every group or class or interval of values
  - Associated **frequency** of each group
    - Number of observations assigned to each group
    - Sum of frequencies is number of observations
      - N for population
      - n for sample
- Class **midpoint** is the middle value of a group or class or interval
- **Relative frequency** is the percentage of total observations in each class
  - Sum of relative frequencies = 1



# Example: Frequency Distribution

<b>x</b>	<b>f(x)</b>	<b>f(x)/n</b>
<b>Spending Class (\$)</b>	<b>Frequency (number of customers)</b>	<b>Relative Frequency</b>
0 to less than 100	30	0.163
100 to less than 200	38	0.207
200 to less than 300	50	0.272
300 to less than 400	31	0.168
400 to less than 500	22	0.120
500 to less than 600	13	0.070
	<hr/> 184	<hr/> 1.000

- Example of relative frequency:  $30/184 = 0.163$
- Sum of relative frequencies = 1

# Cumulative Frequency Distribution

<b>x</b>	<b>F(x)</b>	<b>F(x)/n</b>
<b>Spending Class (\$)</b>	<b>Cumulative Frequency</b>	<b>Cumulative Relative Frequency</b>
0 to less than 100	30	0.163
100 to less than 200	68	0.370
200 to less than 300	118	0.641
300 to less than 400	149	0.810
400 to less than 500	171	0.929
500 to less than 600	184	1.000

The **cumulative frequency** of each group is the sum of the frequencies of that and all preceding groups.



# Graphical Representation of Data

- The Histogram
- The Polygon
- Stem and Leaf Display
- The Bar chart
- The Scatter Diagram
- The Pie Chart
- The Pareto Diagram
- The Side-by-side Bar Chart



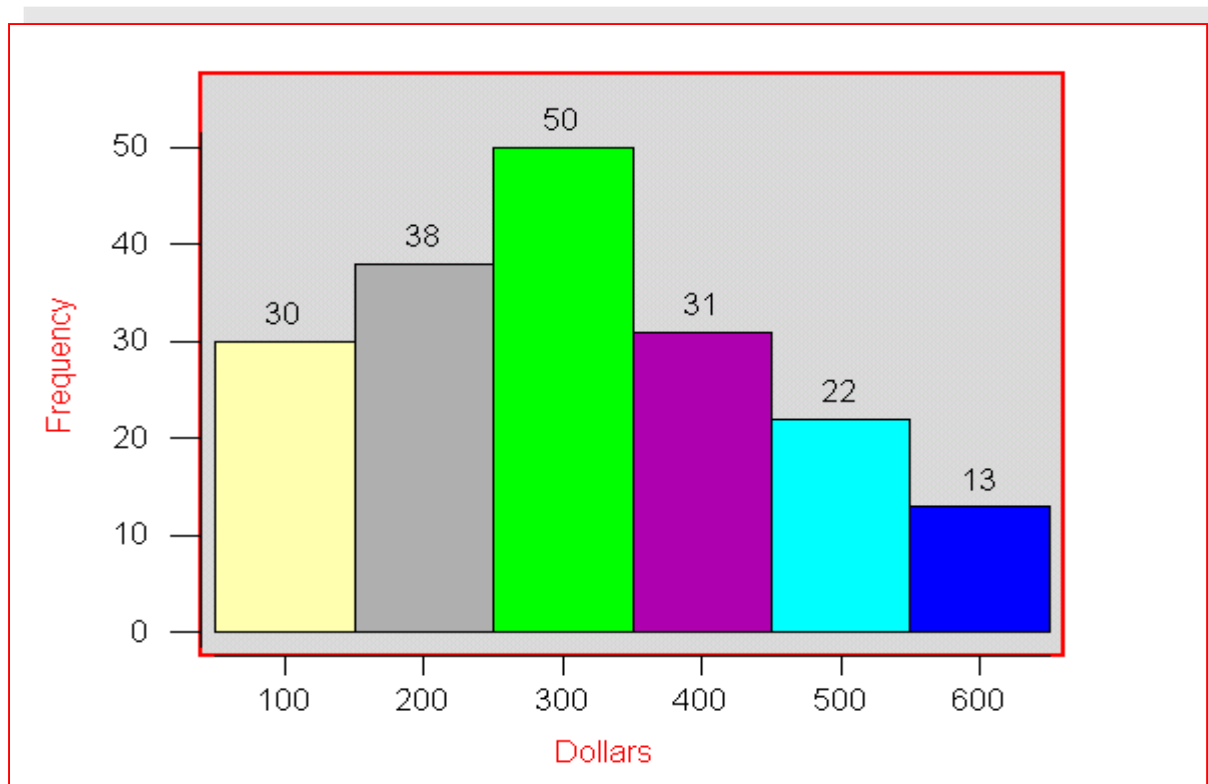
# Histogram

- A **histogram** is a chart made of bars of different heights.
  - ✓ Widths and locations of bars correspond to widths and locations of data groupings
  - ✓ Heights of bars correspond to frequencies or relative frequencies of data groupings



# Histogram Example

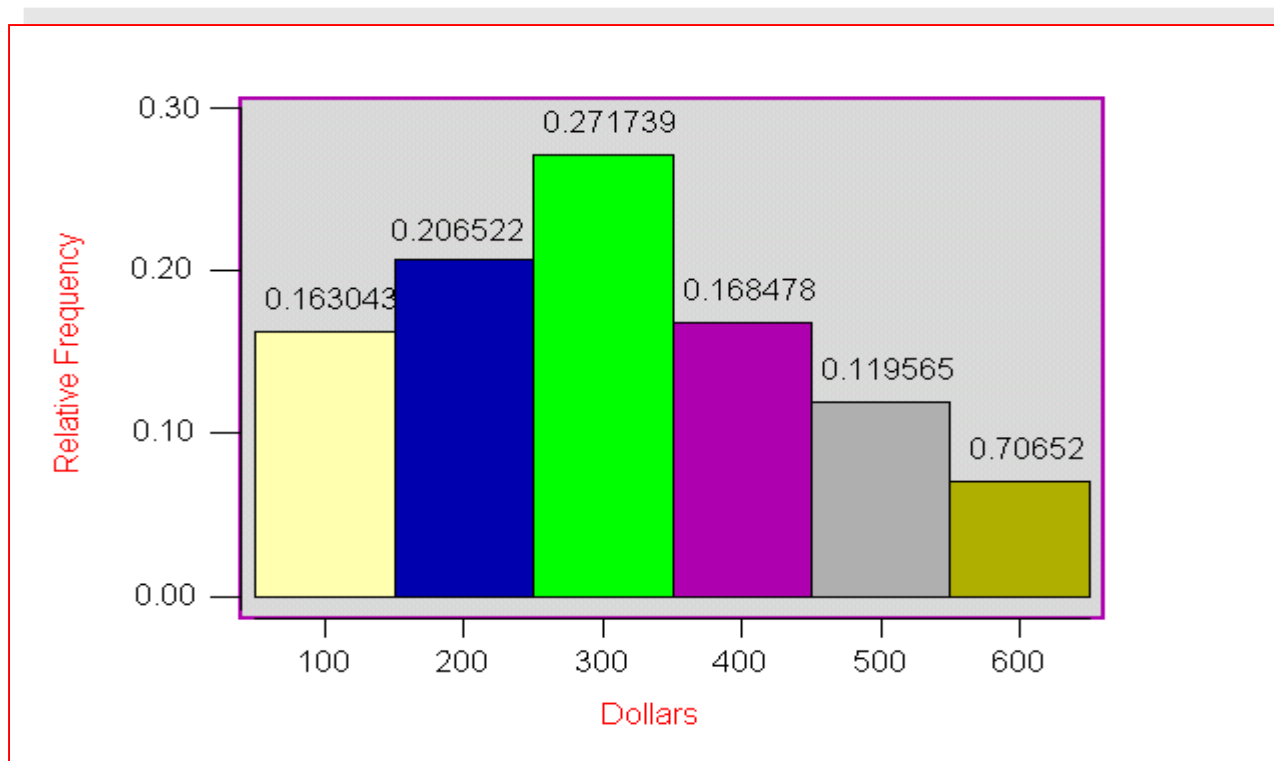
## Frequency Histogram





# Histogram Example

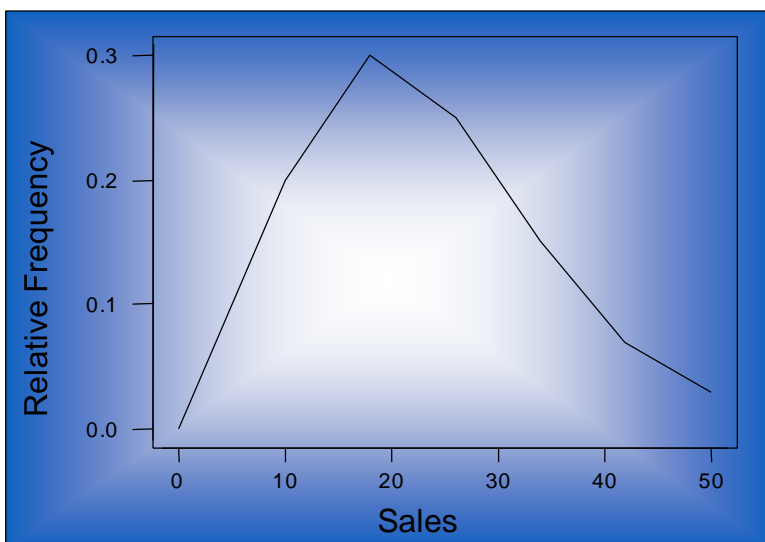
## Relative Frequency Histogram



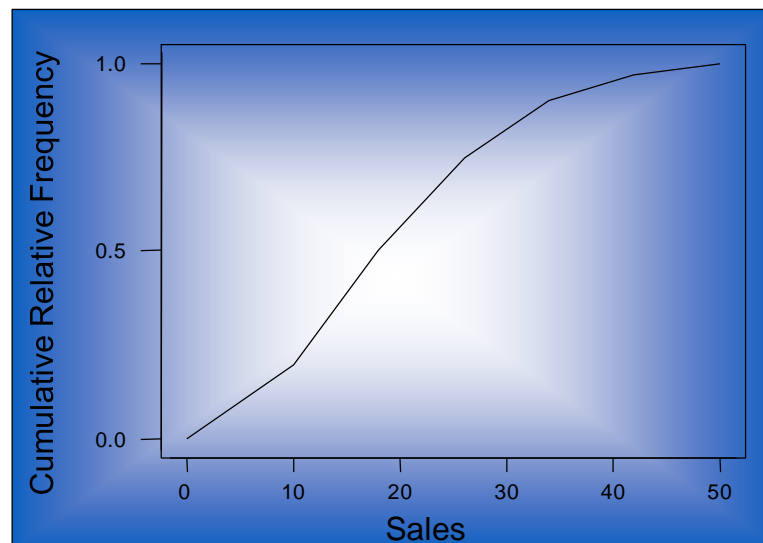


# Frequency Polygon and Ogive

## Relative Frequency Polygon



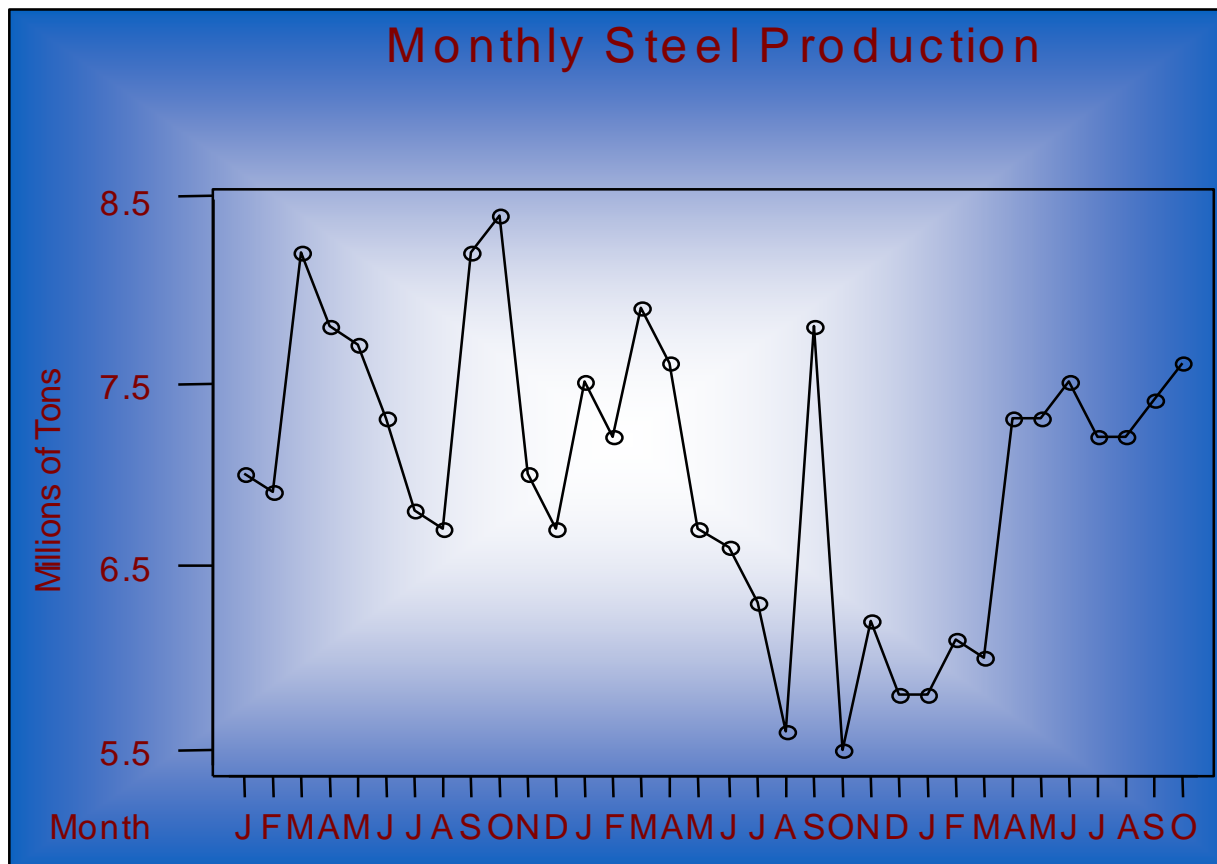
## Ogive



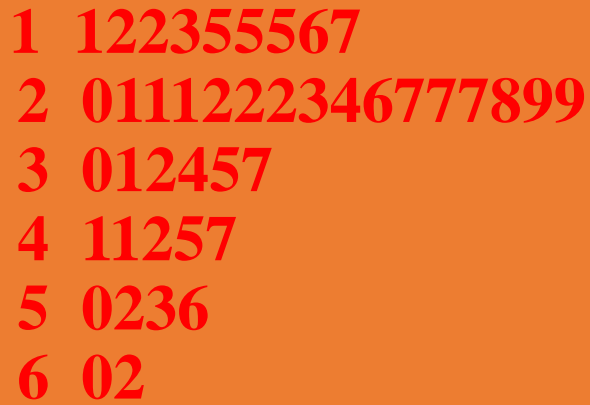
(Cumulative frequency or  
relative frequency graph)



# Time Plot



# Example: Stem-and-Leaf Display



A stem-and-leaf display showing task performance times. The stems are 1 through 6, and the leaves are the digits of the times. The display is presented in red text on an orange background.

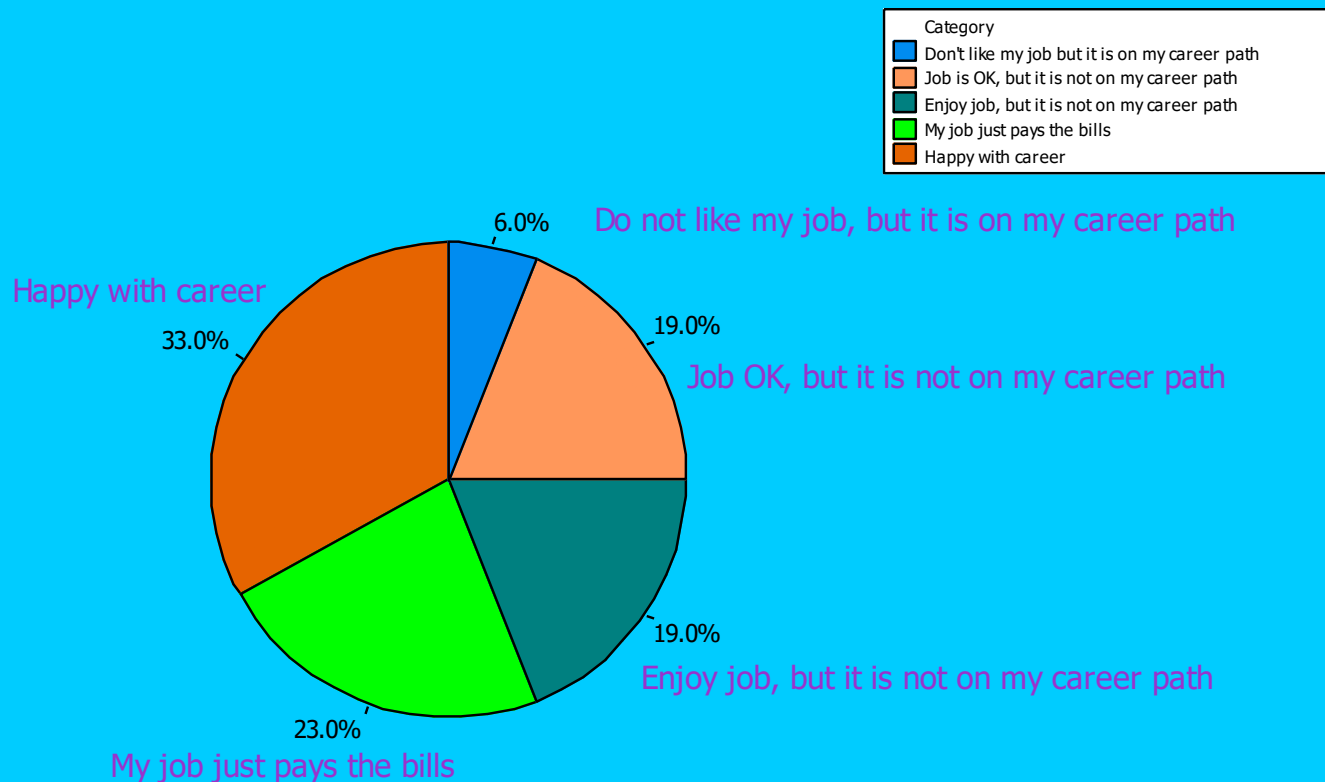
Stem	Leaves
1	1 2 2 3 5 5 5 6 7
2	0 1 1 1 2 2 2 3 4 6 7 7 7 8 9 9
3	0 1 2 4 5 7
4	1 1 2 5 7
5	0 2 3 6
6	0 2

Figure 1-17: Task Performance Times



# Pie Chart

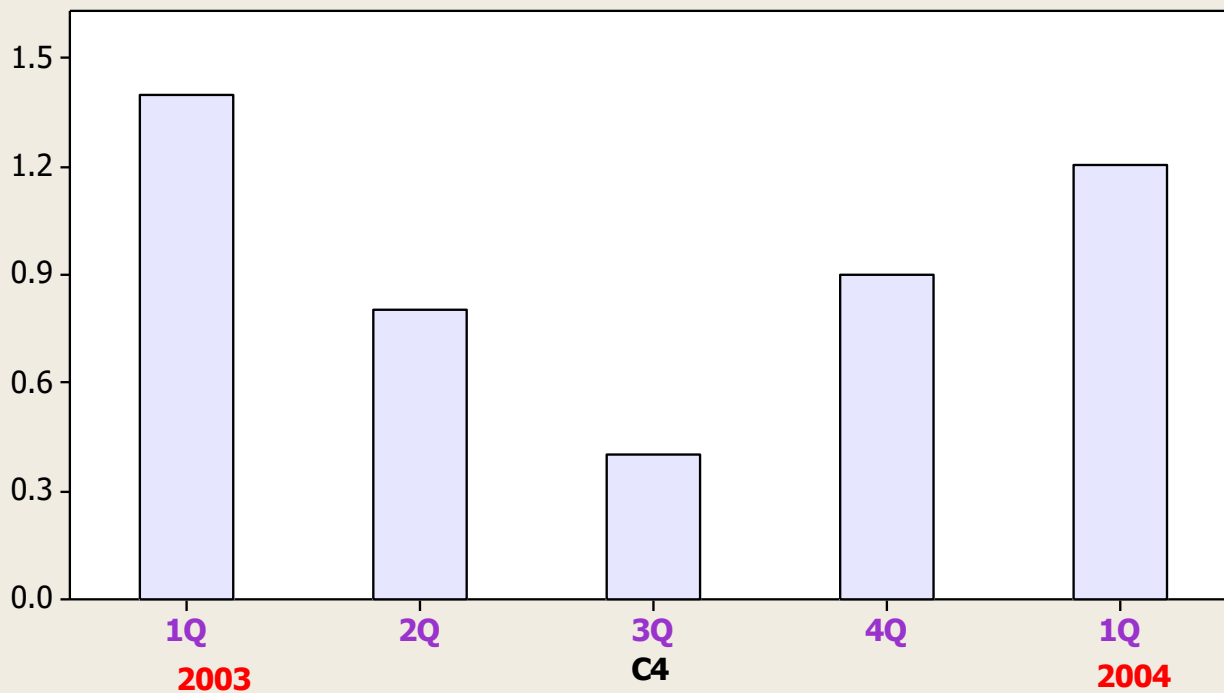
**Figure 1-10: Twentysomethings split on job satisfaction**





# Bar Chart

**Figure 1-11: SHIFTING GEARS**  
Quarterly net income for General Motors (in billions)



# Presentation of Data

Tabulating and Graphing  
Bivariate Categorical Data:

Example: Contingency Table  
or cross-classification  
table.

		Population (in lakhs)				
		1961	1971	1981	1991	2001
Rural		200	350	450	580	700
Urban		150	200	300	350	400
Total		350	550	750	920	1100

# Further Discussions

- False Base Line
- Net balance graph
- Band graph
- Range Graph
- Subdivided bar diagram
- Broken bars
- Deviation Bar Diagram
- Duo directional bar diagram
- Sliding bar diagram
- Pyramid diagram
- Rectangular Diagram
- Percentage Rectangular Diagram
- Pictograms
- Cartograms