Statistics 1
Chapter 2

Describing Data

Ankara University, Faculty of Political Science, Department of Economics, Onur Özsoy
4/4/2018

Chapter Two

Describing Data: Frequency Distributions and Graphic Presentation

GOALS

When you have completed this chapter, you will be able to:

ONE

Organize data into a frequency distribution.

TWO

Portray a frequency distribution in a histogram, frequency polygon, and cumulative frequency polygon.
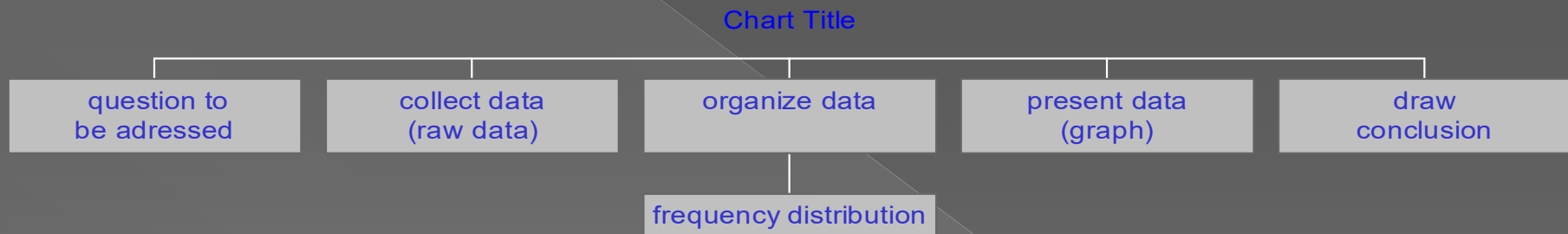
THREE

Develop a stem-and-leaf display.

FOUR

Present data using such graphic techniques as line charts, bar charts, and pie charts.

Frequency Distribution

⦿ Frequency distribution: A grouping of data into categories showing the number of observations in each mutually exclusive category.

Construction of a Frequency Distribution

Chart Title

| question to be adressed | collect data (raw data) | organize data | present data (graph) | draw conclusion |

frequency distribution

Frequency Distribution

- Class mark (midpoint): A point that divides a class into two equal parts. This is the average between the upper and lower class limits.

- Class interval: For a frequency distribution having classes of the same size, the class interval is obtained by subtracting the lower limit of a class from the lower limit of the next class.

EXAMPLE 1

- Dr. Tillman is the dean of the school of business and wishes to determine the amount of studying business school students do. He selects a random sample of 30 students and determines the number of hours each student studies per week: 15.0, 23.7, 19.7, 15.4, 18.3, 23.0, 14.2, 20.8, 13.5, 20.7, 17.4, 18.6, 12.9, 20.3, 13.7, 21.4, 18.3, 29.8, 17.1, 18.9, 10.3, 26.1, 15.7, 14.0, 17.8, 33.8, 23.2, 12.9, 27.1, 16.6.

- Organize the data into a frequency distribution.

EXAMPLE 1 *continued*

Consider the classes 8-12 and 13-17. The class marks are 10 and 15. The class interval is 5 (13-8).

Hours studying	Frequency, f
8-12	1
13-17	12
18-22	10
23-27	5
28-32	1
33-37	1

Suggestions on Constructing a Frequency Distribution

- The class intervals used in the frequency distribution should be equal.

- Determine a suggested class interval by using the formula: *i = (highest value-lowest value)/number of classes.*

Suggestions on Constructing a Frequency Distribution

- Use the computed suggested class interval to construct the frequency distribution.
  Note: this is a *suggested* class interval; if the computed class interval is 97, it may be better to use 100.
- Count the number of values in each class.

Relative Frequency Distribution

⊙ The relative frequency of a class is obtained by dividing the class frequency by the total frequency.

Hours	Frequency, f	Relative Frequency
8 - 1 2	1	1 / 3 0 = .0 3 3 3
1 3 - 1 7	1 2	1 2 / 3 0 = .4 0 0
1 8 - 2 2	1 0	1 0 / 3 0 = .3 3 3
2 3 - 2 7	5	5 / 3 0 = .1 6 6 7
2 8 - 3 2	1	1 / 3 0 = .0 3 3 3
3 3 - 3 7	1	1 / 3 0 = .0 3 3 3
T O T A L	3 0	3 0 / 3 0 = 1

T

Stem-and-Leaf Displays

- Stem-and-Leaf Display: A statistical technique for displaying a set of data. Each numerical value is divided into two parts: the leading digits become the stem and the trailing digits the leaf.

- *Note:* An advantage of the stem-and-leaf display over a frequency distribution is we do not lose the identity of each observation.

EXAMPLE 2

- Colin achieved the following scores on his twelve accounting quizzes this semester: 86, 79, 92, 84, 69, 88, 91, 83, 96, 78, 82, 85. Construct a stem-and-leaf chart for the data.

| stem	leaf
6	9
7	8 9
8	2 3 4 5 6 8
9	1 2 6

Graphic Presentation of a Frequency Distribution

- The three commonly used graphic forms are histograms, frequency polygons, and a cumulative frequency distribution (ogive).

- Histogram: A graph in which the classes are marked on the horizontal axis and the class frequencies on the vertical axis. The class frequencies are represented by the heights of the bars and the bars are drawn adjacent to each other.

Graphic Presentation of a Frequency Distribution

⦿ A frequency polygon consists of line segments connecting the points formed by the class midpoint and the class frequency.

⦿ A cumulative frequency distribution (ogive) is used to determine how many or what proportion of the data values are below or above a certain value.
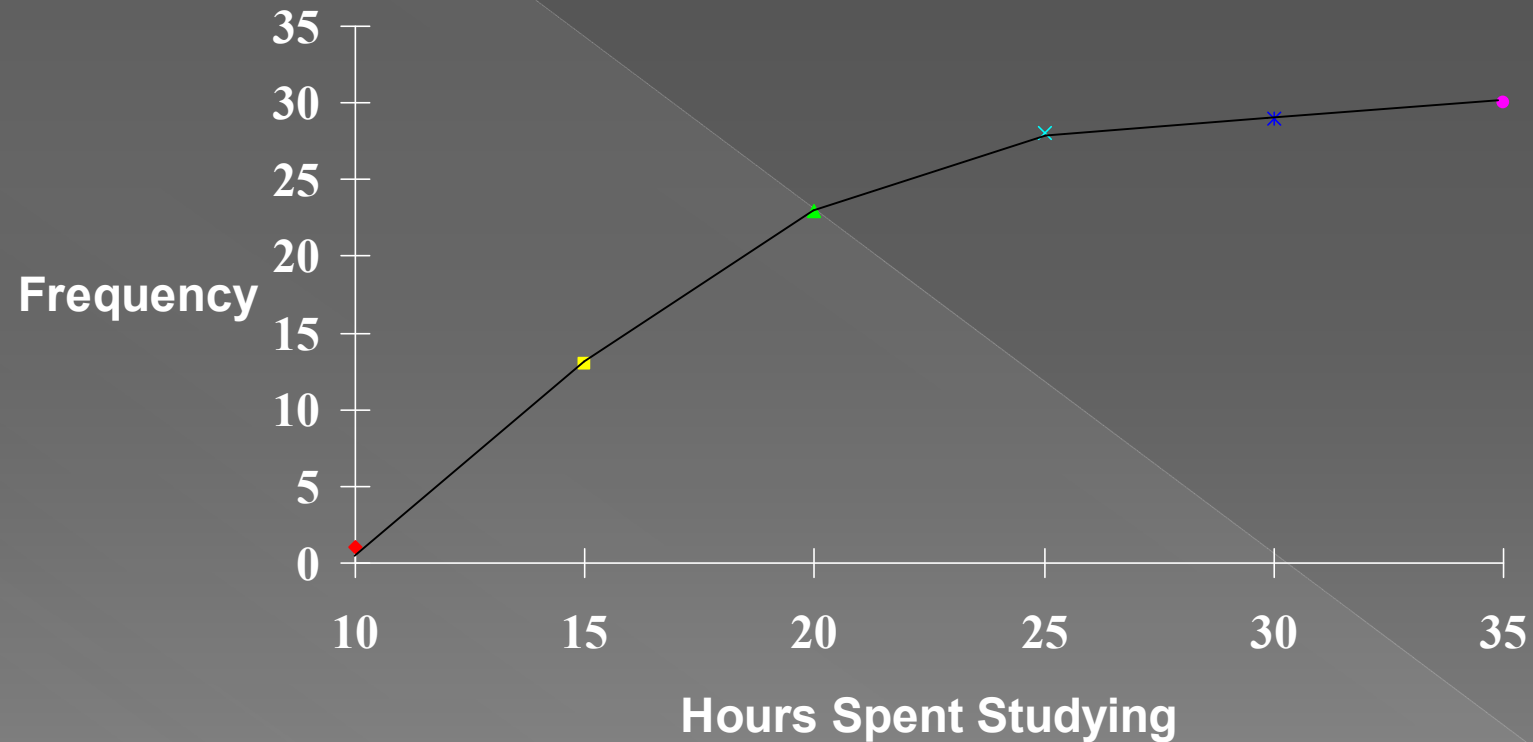
Histogram for Hours Spent Studying
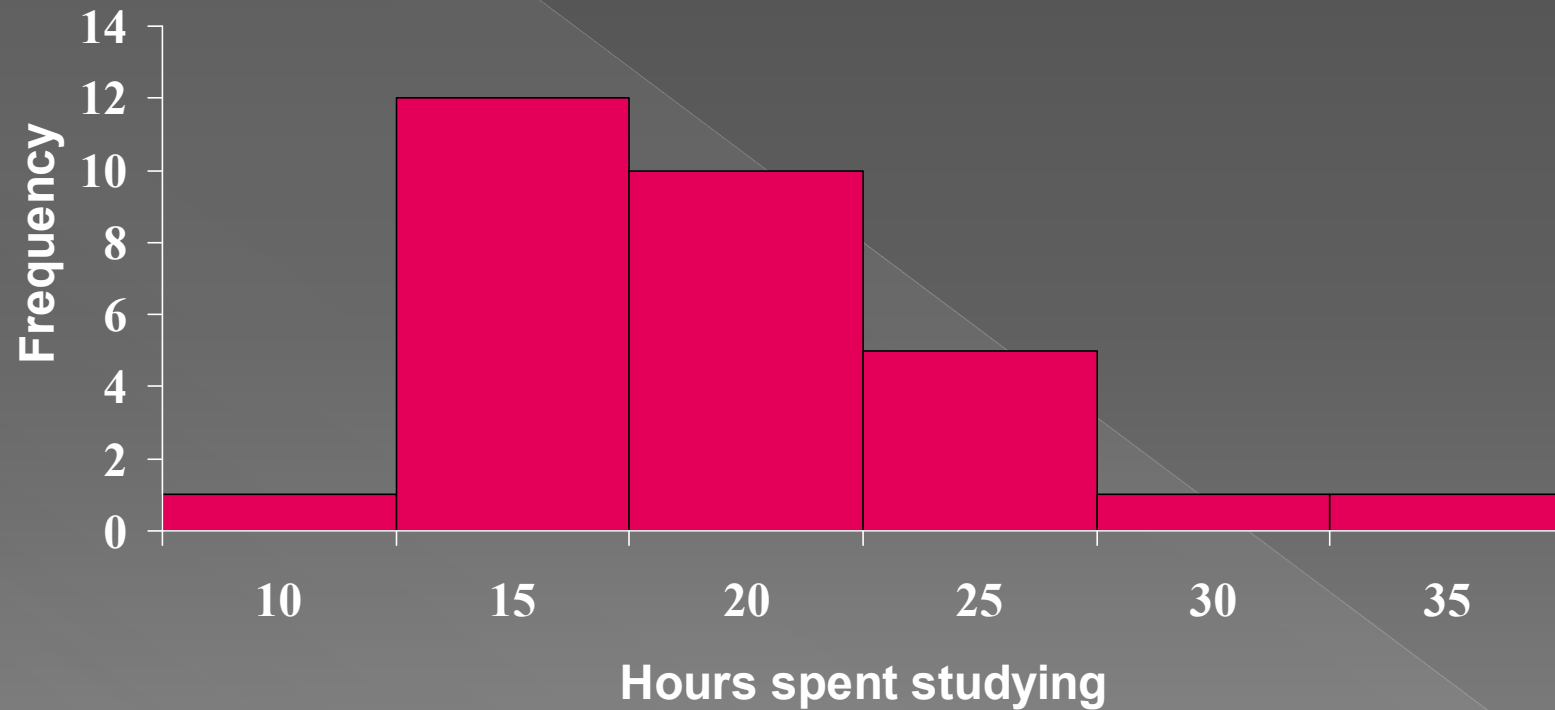
Frequency Polygon for Hours Spent Studying
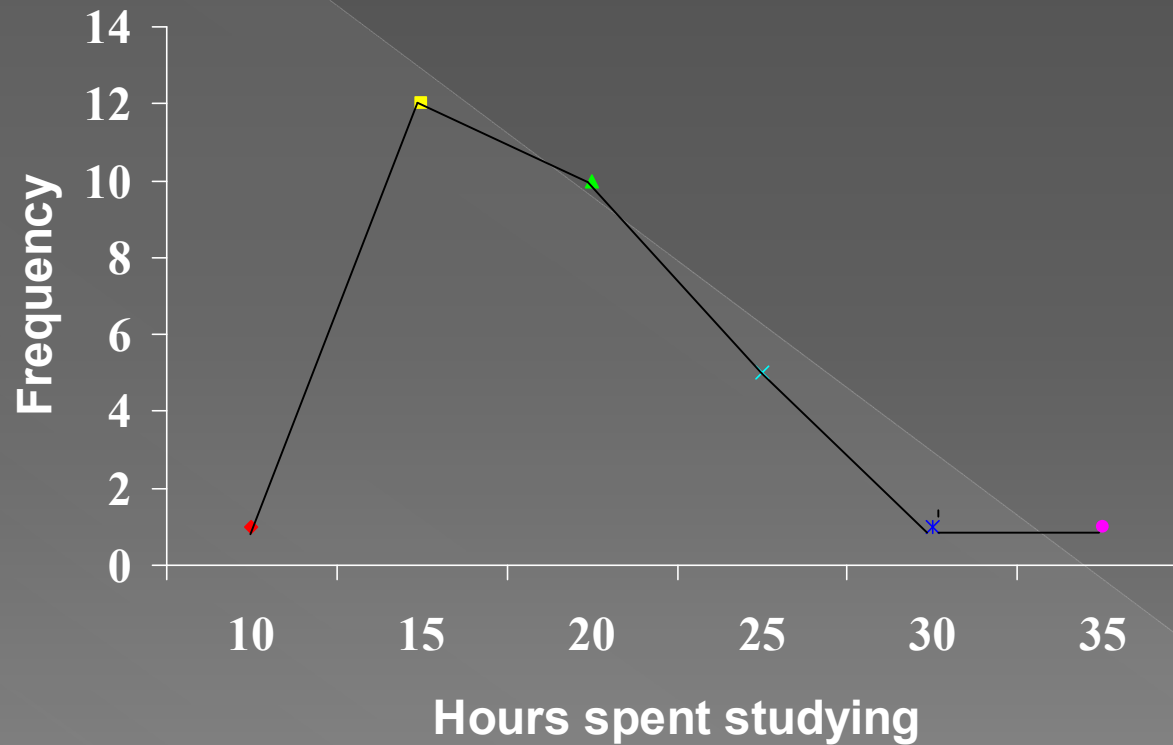
Less Than Cumulative Frequency Distribution For Hours Studying

Frequency

Hours Spent Studying

Bar Chart

- A bar chart can be used to depict any of the levels of measurement (nominal, ordinal, interval, or ratio).

- EXAMPLE 3: Construct a bar chart for the number of unemployed people per 100,000 population for selected cities of 1995.

EXAMPLE 3 *continued*

City	Number of unemployed per 100,000 population
Atlanta, GA	7300
Boston, MA	5400
Chicago, IL	6700
Los Angeles, CA	8900
New York, NY	8200
Washington, D.C.	8900

Bar Chart for the Unemployment Data

Pie Chart

- A pie chart is especially useful in displaying a relative frequency distribution. A circle is divided proportionally to the relative frequency and portions of the circle are allocated for the different groups.

- EXAMPLE 4: A sample of 200 runners were asked to indicate their favorite type of running shoe.

EXAMPLE 4 *continued*

- Draw a pie chart based on the following information.

Type of shoe	# of runners
Nike	92
Adidas	49
Reebok	37
Asics	13
Other	9

Pie Chart for Running Shoes