

# Chapter 2 – Descriptive Statistics

## Organizing and Summarizing Data

### Types of Variables:

**Quantitative (Discrete & Continuous Variable)** Numerically valued data.

**Qualitative** – a non-numerically valued data (arithmetic operations do not make sense).

Examples for the different types of variables:(weight, time, number of TV's in the house, number of m&m's in a bag, marital status, color of hair, ....)

### Section 2.1 Organizing Qualitative Data

### Tabular Summaries:

### **Frequency and Relative Frequency Distributions:**

#### **Example-1 : (Qualitative Variable)**

**The party affiliation for 40 students in math 2620 is given below.**

Democratic Republican Republican Democratic Democratic Republican Republican Republican  
Republican Republican Republican Republican Republican Republican Democratic Democratic  
Democratic Democratic Republican Republican Republican Republican Republican Republican  
Democratic Democratic Republican Democratic Democratic Democratic Democratic Other  
Other Other Other Other Other Other Other Other

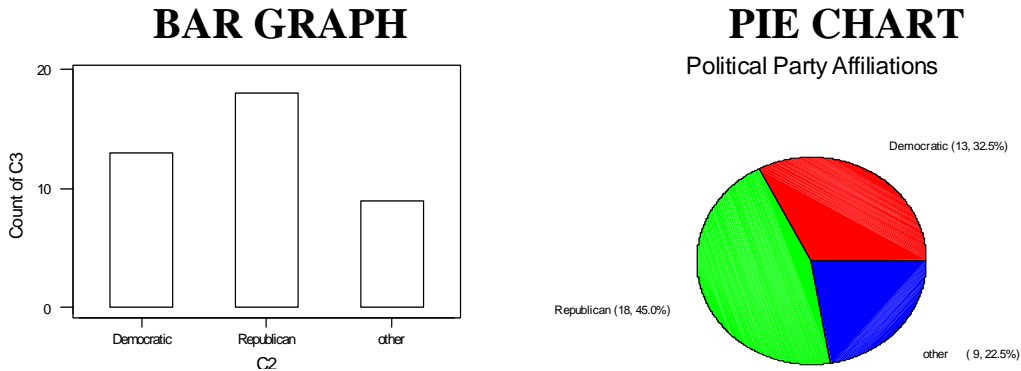
Party	Frequency	Relative Frequency (%)
Democratic	13	32.5
Republican	18	45.0
Other	9	22.5
<b>Total</b>	<b>40</b>	<b>100%</b>

### Graphical Summaries

### **Pie Charts and Bar Graphs**

**Two methods for displaying qualitative data are Pie Charts and Bar Graphs.**

## Example-2 (Reference-Example-1)



**Homework: 17 and 18 on page 72**

## Section 2.2 Organizing Quantitative Data

### Frequency and Relative Frequency Distributions:

#### Example-3 (Quantitative-Discrete Variable)

A nursery school offers programs for 4-year olds ranging from 1-day-a-week program to 5-day-a-week program. To help in planning, the school's director surveyed parents regarding the type of program they prefer. The following data, which represents the number of days, were obtained.

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

Construct the frequency and relative frequency distributions and answer the following.

- (a) What percentage of parents prefer 2-day-a week program?
- (b) What percentage of parents prefers 4-day-a-week or 5-day-a-week program?

### Solution

In the data set there are 4-1s, 10-2s, 6-3s, 3-4s, and 3-5s with 26 total data items. For example, the relative frequency of the data value 1 is  $(4/26)100 = 15.4\%$ .

Number of Days	Frequency	Relative Frequency (%)
1	4	15.4%
2	10	38.5%
3	6	23.1%
4	3	11.5%
5	3	11.5%
Total: 26		100%

- (a) The percentage of parents that prefer 2-day-a-week program is 38.5% .
- (b) The percentage of parents that prefer 4-day-a-week program or 5 day-a- week program is  $11.5 + 11.5 = 23\%$ .

### Guidelines for making a Freq. Distribut. (Continuous Variable):

- a. Choose between 5 and 20 classes (intervals for a histogram). A histogram is sensitive to the number of classes, so you may want to try several possibilities in practice. Rule of thumb: about  $\sqrt{n}$  classes for a Frequency Distribution.
- b. All class widths must be the same.
- c. The lower limit of the smallest class is always less than the smallest data value. The upper limit of the largest class is always greater than the largest value.
- d. Each item goes into one and only class; that is, the classes are non-overlapping.

### Example-4 (Quantitative-Continuous Variable)

We are given the mathematics achievement test scores for a sample of 50 sixth-grade students at Maple Elementary School.



**Example-5 (Quantitative-Discrete Variable) (Kids in the Family)**

**0 3 0 0 3 0 2 2 0 1 2 1 0 0 1 2 4 0 4 2 1 0 1 0 0 2 0 1 3 2**

# of Kids	Frequency	Relative Frequency (%)
0	12	40.0
1	6	20.0
2	7	23.3
3	3	10.0
4	2	6.7
<b>Total</b>	<b>30</b>	<b>100%</b>

**Cumulative Frequency and Cumulative Percent Frequency Distributions:**

Freq. Distribution		Cumulative Freq. Distribution	
# of Kids	Frequency	# of Kids	Cumulative Freq.
0	12	0	12
1	6	1	18
2	7	2	25
3	3	3	28
4	2	4	30

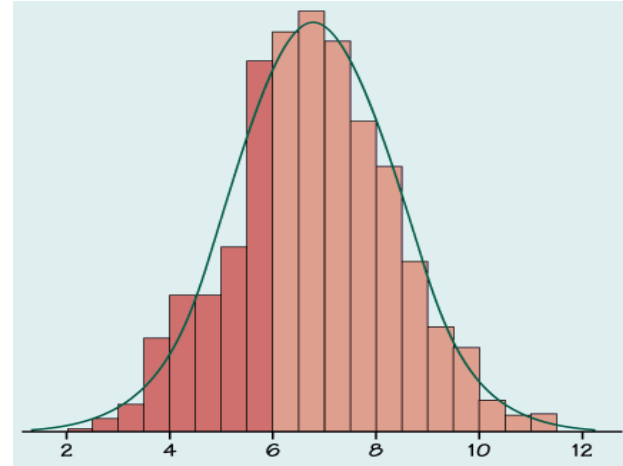
**Cumulative Relative Frequency Distribution**

Relative Freq. Distribut.		Cumulative Relative Freq. Distribut.	
# of Kids	Relat. Freq.	# of Kids	Cumul. Relat.Freq.
0	40.0%	0	40.0%
1	20.0%	1	60.0%
2	23.3%	2	83.3%
3	10.0%	3	93.3%
4	6.7%	4	100.0%

## Graphical Summaries

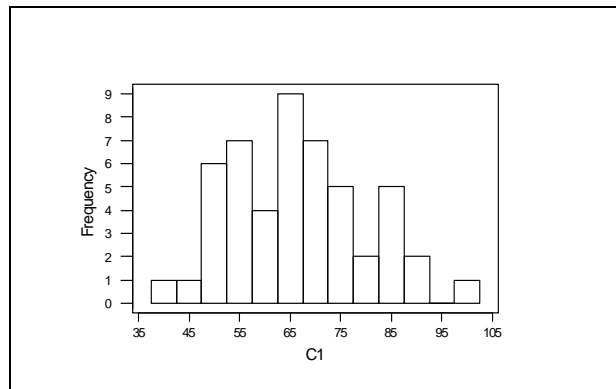
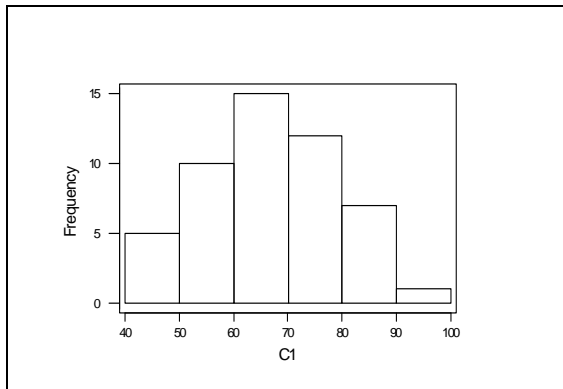
**NOTE: Questions-**What are we looking for when we look at data?

- The shape of the distribution of the data.
- The symmetry or skewness of the data.
- The center of the data.
- The spread of the data.



**Histogram (Use for Quantitative Variable) --** A histogram is a graphical representation of quantitative data that can help answer the questions above.

**Example-5** Draw the histogram for the data in Example-2.



**Homework: 7, 9, 10, 12, 13, 41, and 42 pp. 91-96**