Describing Data

Chapter 2

Learning Objectives

- Make a frequency table for a set of data.

- Organize data into bar chart

- Present a set of data in pie chart.

- Create a frequency distribution for a data set

- Understand a relative frequency distribution.

- Present data from a frequency distribution in a histogram or frequency polygon.

- Construct and interpret a cumulative frequency distribution.

Frequency Tables

- Frequency table: A grouping of qualitative data into mutually exclusive classes showing the number of observations in each class.

Example: In the Professional Saudi League season 2013/2014 there were 671 yellow cards.

Player position	Number of yellow cards
Goalkeeper	31
Defender	276
Midfielder	260
Striker	104

Frequency Tables

- Relative Frequency: captures the relationship between a class and the total number of observation.

Example: In the Professional Saudi League season 2013/2014 there were 671 yellow cards.

Player position	Number of yellow cards	Relative Frequency
Goalkeeper	31	0.05
Defender	276	0.41
Midfielder	260	0.39
Striker	104	0.15

Frequency Tables

A, A, A, O, O, AB, O, O, AB, A, A, B, B, B, O, O, O , B, B, O, O, O, AB, AB, AB,

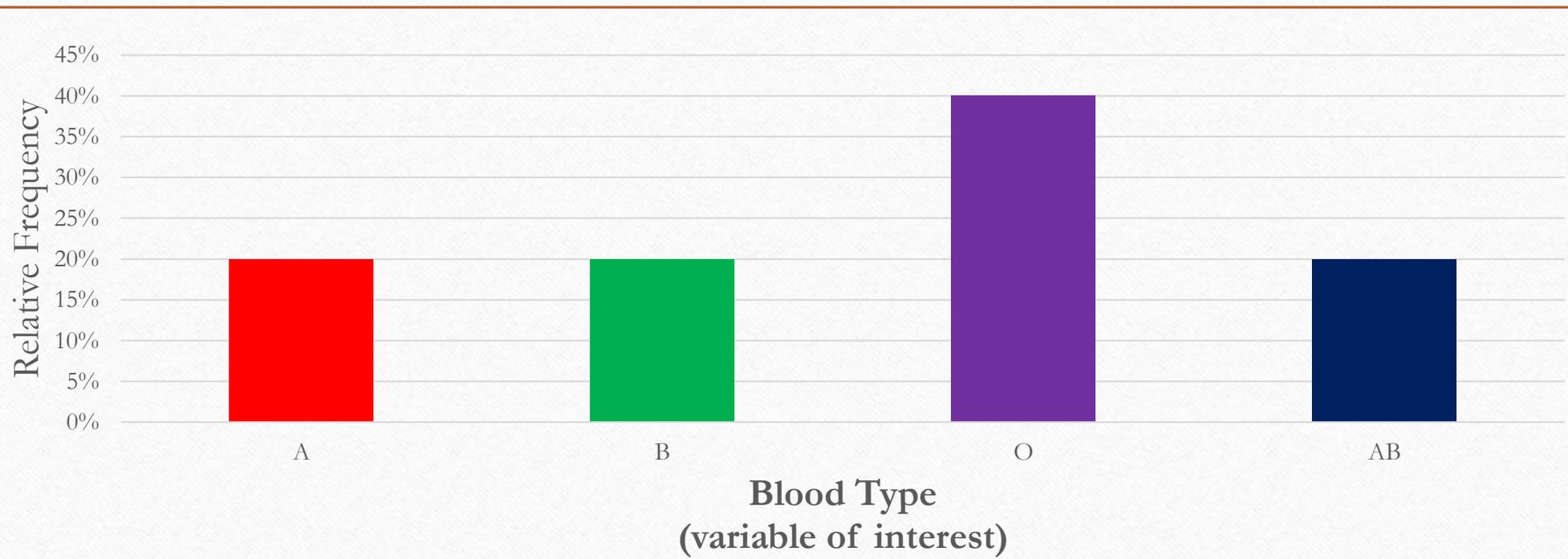| Blood Type	Frequency	Relative Frequency
A	5	0.2
B	5	0.2
O	10	0.4
AB	5	0.2

Bar Charts

- Bar Chart: A graph in which the classes are reported on horizontal axis and the class frequencies on the vertical axis. The class frequencies are proportional to the heights of the bars.
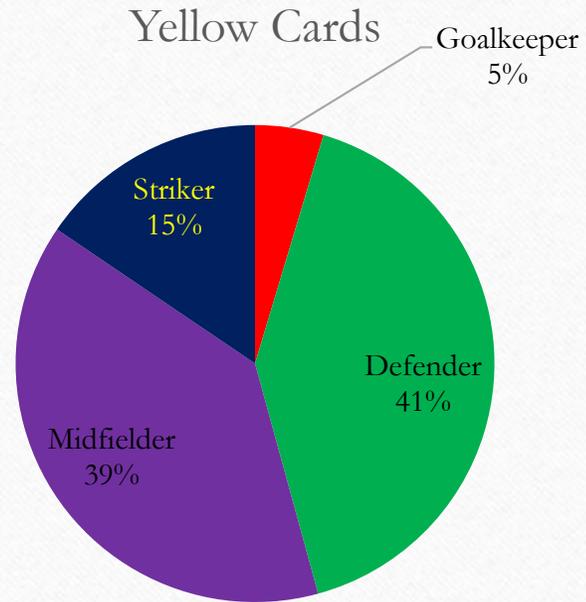
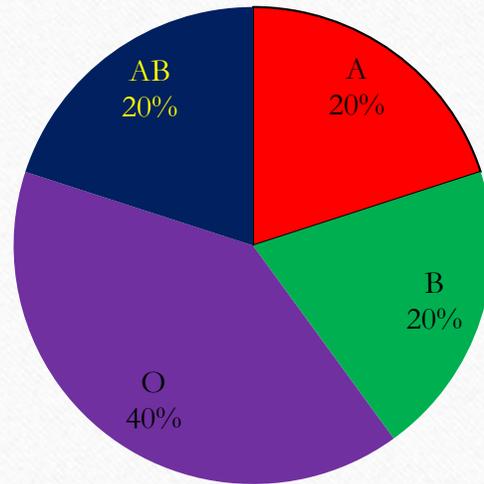# Bar Charts

Bar Charts

Pie Charts

- Pie chart: A chart that shows the proportion or percent that each class represents of the total number of frequencies.

Pie Charts

Yellow Cards

Goalkeeper
5%

Striker
15%

Defender
41%

Midfielder
39%

Pie Charts

Blood Types

Frequency Distribution

- Frequency Distribution: A grouping of data into mutually exclusive classes showing the number of observations in each class.

Example: In an event we asked the audience about their ages and we construct the following table:

Class	Frequency
5 up to 10	10
10 up to 15	2
15 up to 20	4
20 up to 25	3
25 up to 30	1

Constructing a Frequency Distribution

- Example:

Ms. Kathryn Ball of AutoUSA wants to develop tables, charts, and graphs to show the typical selling price on various dealer lots. The table down reports only the price of the 80 vehicles sold last month at Whitner Autoplex.

TABLE 2—4 Prices of Vehicles Sold Last Month at Whitner Autoplex

						Lowest
$23,197	$23,372	$20,454	$23,591	$26,651	$27,453	$17,266
18,021	28,683	30,872	19,587	23,169	35,851	19,251
20,047	24,285	24,324	24,609	28,670	15,546	15,935
19,873	25,251	25,277	28,034	24,533	27,443	19,889
20,004	17,357	20,155	19,688	23,657	26,613	20,895
20,203	23,765	25,783	26,661	32,277	20,642	21,981
24,052	25,799	15,794	18,263	35,925	17,399	17,968
20,356	21,442	21,722	19,331	22,817	19,766	20,633
20,962	22,845	26,285	27,896	29,076	32,492	18,890
21,740	22,374	24,571	25,449	28,337	20,642	23,613
24,220	30,655	22,442	17,891	20,818	26,237	20,445
21,556	21,639	24,296				Highest

Constructing a Frequency Distribution

- Step 1: Decide on the number of classes.

A useful recipe to determine the number of classes (k) is the "2 to the k rule." Such that $2^k > n$.

There were 80 vehicles sold. So n = 80. If we try k = 6, which means we would use 6 classes, then $2^6 = 64$, somewhat less than 80. Hence, 6 is not enough classes. If we let k = 7, then $2^7 = 128$, which is greater than 80. So the recommended number of classes is 7.

Constructing a Frequency Distribution

- Step 2: Determine the class interval or width.

The formula is: $i \geq \frac{H-L}{k}$, where:

i is the class interval H is the highest observed value

L is the lowest observed value k is the number of classes.

$$\frac{\$35,925 - \$15,546}{7} = \$2,911$$

Round up to some convenient number, such as a multiple of 10 or 100. Use a class width of $3,000

Constructing a Frequency Distribution

- Step 3: Set the individual class limits.

$15,000 up to 18,000
18,000 up to 21,000
21,000 up to 24,000
24,000 up to 27,000
27,000 up to 30,000
30,000 up to 33,000
33,000 up to 36,000

Constructing a Frequency Distribution

- **Step 4:** Tally the vehicle selling prices into the classes.

Class	Tallies
$15,000 up to $18,000	JHT III
$18,000 up to $21,000	JHT JHT JHT JHT III
$21,000 up to $24,000	JHT JHT JHT II
$24,000 up to $27,000	JHT JHT JHT III
$27,000 up to $30,000	JHT III
$30,000 up to $33,000	IIII
$33,000 up to $36,000	II

Constructing a Frequency Distribution

- Step 5: Count the number of items in each class.

Selling Prices ($ thousands)	Frequency
15 up to 18	8
18 up to 21	23
21 up to 24	17
24 up to 27	18
27 up to 30	8
30 up to 33	4
33 up to 36	2
Total	80

Relative Frequency Distribution

- To convert a frequency distribution to a relative frequency distribution, each of the class frequencies is divided by the total number of observations.

TABLE 2–8 Relative Frequency Distribution of the Prices of Vehicles Sold Last Month at Whitner Autoplex

Selling Price ($ thousands)	Frequency	Relative Frequency	Found by
15 up to 18	8	0.1000	8/80
18 up to 21	23	0.2875	23/80
21 up to 24	17	0.2125	17/80
24 up to 27	18	0.2250	18/80
27 up to 30	8	0.1000	8/80
30 up to 33	4	0.0500	4/80
33 up to 36	2	0.0250	2/80
Total	80	1.0000	

Frequency Distribution

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

- Class frequency: The number of observations in each class.

- Class interval: The class interval is obtained by subtracting the lower limit of a class from the lower limit of the next class.

Cumulative Frequency Distribution

- Cumulative frequency distribution: Is the sum of the class and all classes below it in a frequency distribution.

TABLE 2–9 Cumulative Frequency Distribution for Vehicle Selling Price

Selling Price ($ thousands)	Frequency	Cumulative Frequency	Found by
15 up to 18	8	8	
18 up to 21	23	31	← 8 + 23
21 up to 24	17	48	8 + 23 + 17
24 up to 27	18	66	8 + 23 + 17 + 18
27 up to 30	8	74	⋮
30 up to 33	4	78	
33 up to 36	2	80	
Total	80		

Histogram

- Histogram for a frequency distribution based on quantitative data is very similar to the bar chart showing the distribution of qualitative data. 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.

CHART 2–4 Histogram of the Selling Prices of 80 Vehicles at Whitner Autoplex

Frequency Polygon

- A **frequency polygon** also shows the shape of a distribution and is similar to a histogram.

- It consists of line segments connecting the points formed by the intersections of the class midpoints and the class frequencies.

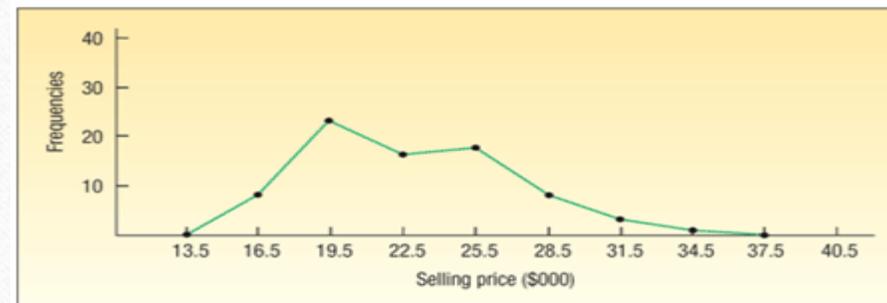| Selling Price ($ thousands)	Midpoint	Frequency
15 up to 18	16.5	8
18 up to 21	19.5	23
21 up to 24	22.5	17
24 up to 27	25.5	18
27 up to 30	28.5	8
30 up to 33	31.5	4
33 up to 36	34.5	2
Total		80

CHART 2–5 Frequency Polygon of the Selling Prices of 80 Vehicles at Whitner Autoplex
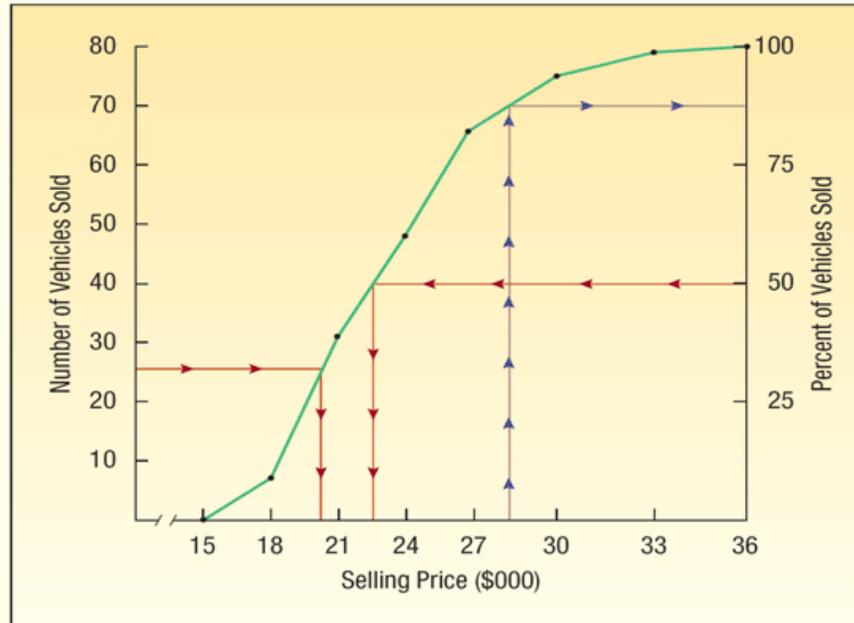
Cumulative Frequency Polygon

CHART 2–7 Cumulative Frequency Distribution for Vehicle Selling Price