كلية العلـــوم

قســـم الانظـمـة الـطبية الذكـية
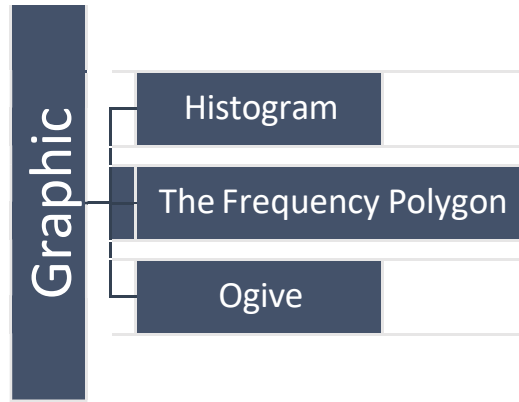
Lecture: (3)

Subject: Statistics and probability
Class: Second
Lecturer: Asst. Lecturer Nabaa Ali

CHAPTER THREE

Graphic

The three most commonly used graphs in research are:

Graphic

- Histogram
- The Frequency Polygon
- Ogive

Purpose of graphs in statistics is to convey the **data** to the viewers **in pictorial form**

- **Easier** for most people to understand the **meaning of data** in form of graphs

- They can also be used to discover a **trend or pattern** in a situation over a period of time

- Useful in **getting** the audience's **attention** in a publication or a speaking presentation

➢ Histogram

The histogram is a graph that displays the data by using contiguous vertical bars (unless the frequency of a class is 0) of various heights to represent the frequencies of the classes.

❑The <u>class boundaries</u> are represented on the horizontal axis

Example 2-4: Construct a histogram to represent the data for the record high temperatures for each of the 50 states (see Example 2–2 for the data).

Class Limits	Class Boundaries	Frequency
100 - 104	99.5 - 104.5	2
105 - 109	104.5 - 109.5	8
110 - 114	109.5 - 114.5	18
115 - 119	114.5 - 119.5	13
120 - 124	119.5 - 124.5	7
125 - 129	124.5 - 129.5	1
130 - 134	129.5 - 134.5	1

Histograms use class boundaries and frequencies of the classes

Record High Temperatures
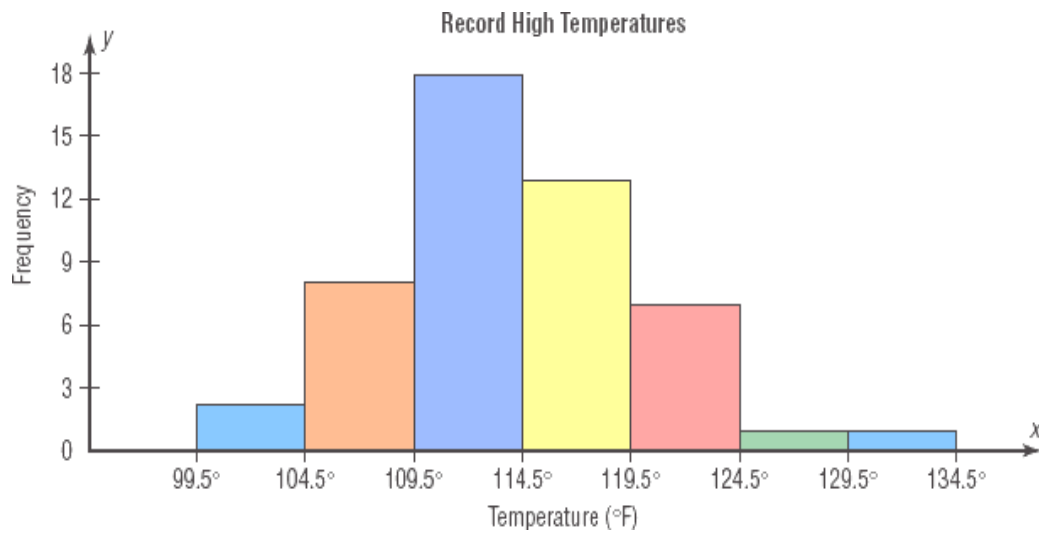
➤ Frequency Polygons

- The *frequency polygon* is a graph that displays the data by using lines that connect points plotted for the frequencies at the class midpoints. The frequencies are represented by the heights of the points.

- The ~~class midpoints~~ are represented on the horizontal axis.

Construct a frequency polygon to represent the data for the record high temperatures for each of the 50 states.

Step 1: find the midpoints of each class (Recall that midpoints are found by adding the upper and lower boundaries and dividing by 2).

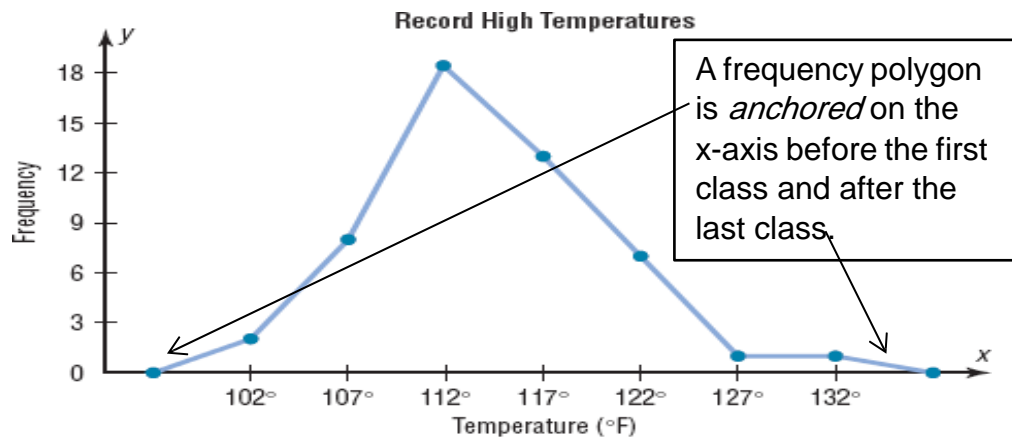$$\frac{99.5+104.5}{2}=102 \qquad \frac{104.5+109.5}{2}=107$$

Step 2 Draw the x and y axes. Label the x axis with the midpoint of each class, and then use a suitable scale on the y axis for the frequencies.

Step 3 Using the midpoints for the x values and the frequencies as the y values, plot the points.

Step 4 Connect adjacent points with line segments. Draw a line back to the x axis at the beginning and end of the graph, at the same distance that the previous and next midpoints would be located, as shown in Figure 2–3.

Frequency polygons use class midpoints and frequencies of the classes.

Class Limits	Class Midpoints	Frequency
100 - 104	102	2
105 - 109	107	8
110 - 114	112	18
115 - 119	117	13
120 - 124	122	7
125 - 129	127	1
130 - 134	132	1

Record High Temperatures

A frequency polygon is *anchored* on the x-axis before the first class and after the last class.

4

> Ogive
>> ■ The *ogive* is a graph that represents the cumulative frequencies for the classes in a frequency distribution.
>> ■ The ~~upper class boundaries~~ are represented on the horizontal axis.

Construct an ogive to represent the data for the record high temperatures for each of the 50 states.

Ogives use upper class boundaries and cumulative frequencies of the classes.

Class Limits	Class Boundaries	Frequency	Cumulative Frequency
100 - 104	99.5 - 104.5	2	2
105 - 109	104.5 - 109.5	8	10
110 - 114	109.5 - 114.5	18	28
115 - 119	114.5 - 119.5	13	41
120 - 124	119.5 - 124.5	7	48
125 - 129	124.5 - 129.5	1	49
130 - 134	129.5 - 134.5	1	50

Class Boundaries	Cumulative Frequency
Less than 104.5	2
Less than 109.5	10
Less than 114.5	28
Less than 119.5	41
Less than 124.5	48
Less than 129.5	49
Less than 134.5	50

Record High Temperatures

5
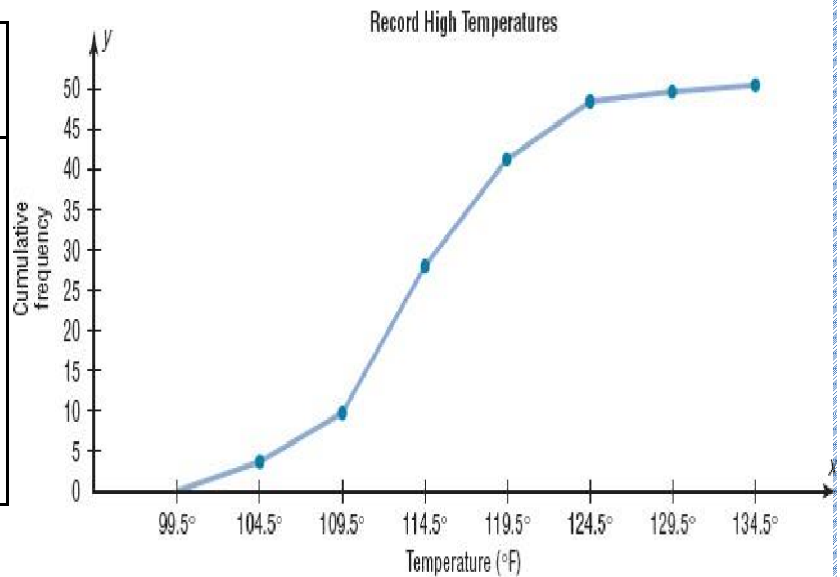
➤ Pareto charts

When the variable displayed on the horizontal axis is qualitative or categorical, a *Pareto chart* can also be used to represent the data.

A **Pareto chart** is used to represent a frequency distribution for a categorical variable, and the frequencies are displayed by the heights of vertical bars, which are arranged in order from highest to lowest.

The data shown here consist of the number of homeless people for a sample of selected cities. Construct and analyze a Pareto chart for the data.

City	Number
Atlanta	6832
Baltimore	2904
Chicago	6680
St. Louis	1485
Washington	5518

Solution

Step 1 Arrange the data from the largest to smallest according to frequency.

City	Number
Atlanta	6832
Chicago	6680
Washington	5518
Baltimore	2904
St. Louis	1485

Step 2 Draw and label the x and y axes.

Step 3 Draw the bars corresponding to the frequencies. The graph shows that the number of homeless people is about the same for Atlanta and Chicago and a lot less for Baltimore and St. Louis.

Number of Homeless People for Large Cities

6
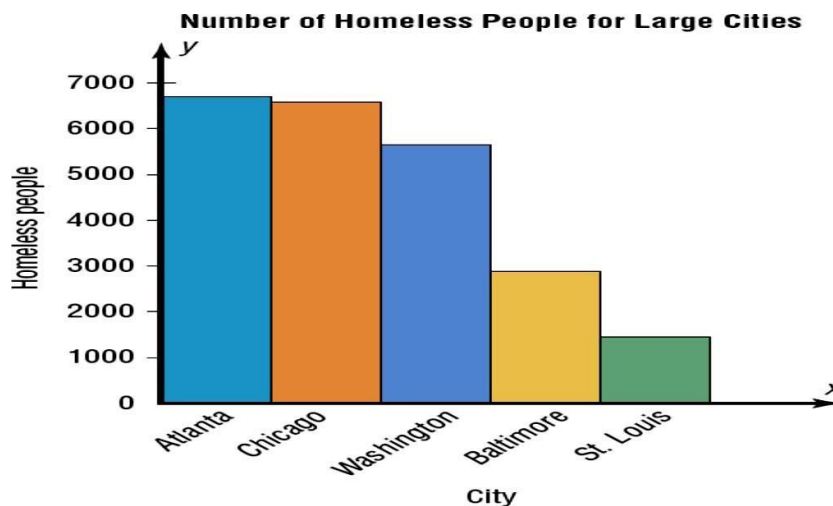
➤ **The Pie Graph**

Pie graphs are used extensively in statistics. The purpose of the pie graph is to show the relationship of the parts to the whole by visually comparing the sizes of the sections. Percentages or proportions can be used. The variable is nominal or categorical.

A **pie graph** is a circle that is divided into sections or wedges according to the percentage of frequencies in each category of the distribution.

This frequency distribution shows the number of pounds of each snack food eaten during the Super Bowl. Construct a pie graph for the data.

Snack	**Pounds (frequency)**
Potato chips	11.2 million
Tortilla chips	8.2 million
Pretzels	4.3 million
Popcorn	3.8 million
Snack nuts	2.5 million
Total $n = 30.0$ million	

Solution

Step 1 Since there are 360_ in a circle, the frequency for each class must be converted into a proportional part of the circle. This conversion is done by using the formula $\text{Degrees} = \frac{f}{n} \bullet 360°$

where f _ frequency for each class and n _ sum of the frequencies. Hence, the following conversions are obtained. The degrees should sum to 360_.*

Potato chips	$\frac{11.2}{30} \cdot 360° = 134°$
Tortilla chips	$\frac{8.2}{30} \cdot 360° = 98°$
Pretzels	$\frac{4.3}{30} \cdot 360° = 52°$
Popcorn	$\frac{3.8}{30} \cdot 360° = 46°$
Snack nuts	$\frac{2.5}{30} \cdot 360° = 30°$
Total	$360°$

Step 2 Each frequency must also be converted to a percentage. Recall from Example 2–1 that this conversion is done by using the formula

$$\% = \frac{f}{n} \cdot 100$$

Hence, the following percentages are obtained. The percentages should sum to 100%.[†]

Potato chips	$\dfrac{11.2}{30} \cdot 100 =$	37.3%
Tortilla chips	$\dfrac{8.2}{30} \cdot 100 =$	27.3%
Pretzels	$\dfrac{4.3}{30} \cdot 100 =$	14.3%
Popcorn	$\dfrac{3.8}{30} \cdot 100 =$	12.7%
Snack nuts	$\dfrac{2.5}{30} \cdot 100 =$	8.3%
Total		99.9%

Step 3 Next, using a protractor and a compass, draw the graph using the appropriate degree measures found in step 1, and label each section with the name and percentages, as shown in Figure 2–14.

Super Bowl Snacks

Popcorn 12.7%
Snack nuts 8.3%
Pretzels 14.3%
Potato chips 37.3%
Tortilla chips 27.3%