PREPERATION OF FREQUENCY DISTRIBUTION TABLE & DIAGRAMS

Describing, Exploring and Comparing Data

Presented by

Dr. Golam Mostafa

Assistant Professor of Geography

Kabi Sukanta Mahavidyalaya

Angus, Bhadreswar, Hooghly

Outline

- 2-1 Introduction
- 2-2 Frequency Distributions
- 2-3 Visualizing Data

Frequency Distributions

- When data are collected in original form, they are called raw data.
- When the raw data is organized into a frequency distribution, the frequency will be the number of values in a specific class of the distribution.

Frequency Distributions

- A frequency distribution is the organizing of raw data in table form, using classes and frequencies.
- The following slide shows an example of a frequency distribution.

Frequency Distributions

- Categorical frequency distributions - can be used for data that can be placed in specific categories, such as nominal- or ordinal-level data.
- Examples - political affiliation, religious affiliation, blood type etc.

Frequency Distribution

Example

Class	Frequency	Percent
A	5	20
B	7	28
O	9	36
AB	4	16

Grouped Frequency Distributions

- Grouped frequency distributions - can be used when the range of values in the data set is very large. The data must be grouped into classes that are more than one unit in width.
- Examples - the life of boat batteries in hours.

Lifetimes of Boat Batteries (hours)

Example

Class limits	Class Boundaries	Frequency	Cumulative frequency
24 - 30	23.5 - 37.5	4	4
38 - 51	37.5 - 51.5	14	18
52 - 65	51.5 - 65.5	7	25

Terms Associated with a Grouped Frequency Distribution

- Class limits represent the smallest and largest data values that can be included in a class.
- In the lifetimes of boat batteries example, the values 24 and 30 of the first class are the class limits.
- The lower class limit is 24 and the upper class limit is 30.
- The class boundaries are used to separate the classes so that there are no gaps in the frequency distribution

Terms Associated with a Grouped Frequency Distribution

- The class width for a class in a frequency distribution is found by subtracting the lower (or upper) class limit of one class minus the lower (or upper) class limit of the previous class.

Guidelines for Constructing a Frequency Distribution

- There should be between 5 and 20 classes.
- The class width should be an odd number.
- The classes must be mutually exclusive.
- The classes must be continuous.
- The classes must be exhaustive.
- The class must be equal in width.

Procedure for the Construction of a Grouped Frequency Distribution

- Find the highest and lowest value.
- Find the range.
- Select the number of classes desired.
- Find the width by dividing the range by the number of classes and rounding up.
- Select a starting point (usually the lowest value); add the width to get the lower limits.
- Find the upper class limits.
- Find the boundaries.
- Tally the data, find the frequencies and find the cumulative frequency.

Grouped Frequency Distribution

Example

In a survey of 20 patients who smoked, the following data were obtained. Each value represents the number of cigarettes the patient smoked per day. Construct a frequency distribution using six classes

10	8	6	14
22	13	17	19
11	9	18	14
13	12	15	15
5	11	16	11

Grouped Frequency Distribution Example

- Step 1: Find the highest and lowest values: $H = 22$ and $L = 5$.
- Step 2: Find the range:

  $R = H - L = 22 - 5 = 17.$
- Step 3: Select the number of classes desired. In this case it is equal to 6.
- Step 4: Find the class width by dividing the range by the number of classes. Width = 17/6 = 2.83. This value is rounded up to 3.
- Step 5: Select a starting point for the lowest class limit. For convenience, this value is chosen to be 5, the smallest data value. The lower class limits will be 5, 8, 11, 14, 17 and 20.
- Step 6: The upper class limits will be 7, 10, 13, 16, 19 and 22. For example, the upper limit for the first class is computed as 8 - 1, etc.
- Step 7: Find the class boundaries by subtracting 0.5 from each lower class limit and adding 0.5 to the upper class limit.
- Step 8: Tally the data, write the numerical values for the tallies in the frequency column and find the cumulative frequencies.

Note: The dash "-" represents "to".

Class Limits	Class Boundaries	Frequency	Cumulative Frequency
05 to 07	4.5 - 7.5	2	2
08 to 10	7.5 - 10.5	3	5
11 to 13	10.5 - 13.5	6	11
14 to 16	13.5 - 16.5	5	16
17 to 19	16.5 - 19.5	3	19
20 to 22	19.5 - 22.5	1	20

Visualization of Data

The three most commonly used graphs in Geography are:

- The histogram.
- The frequency polygon.
- The frequency Curve.

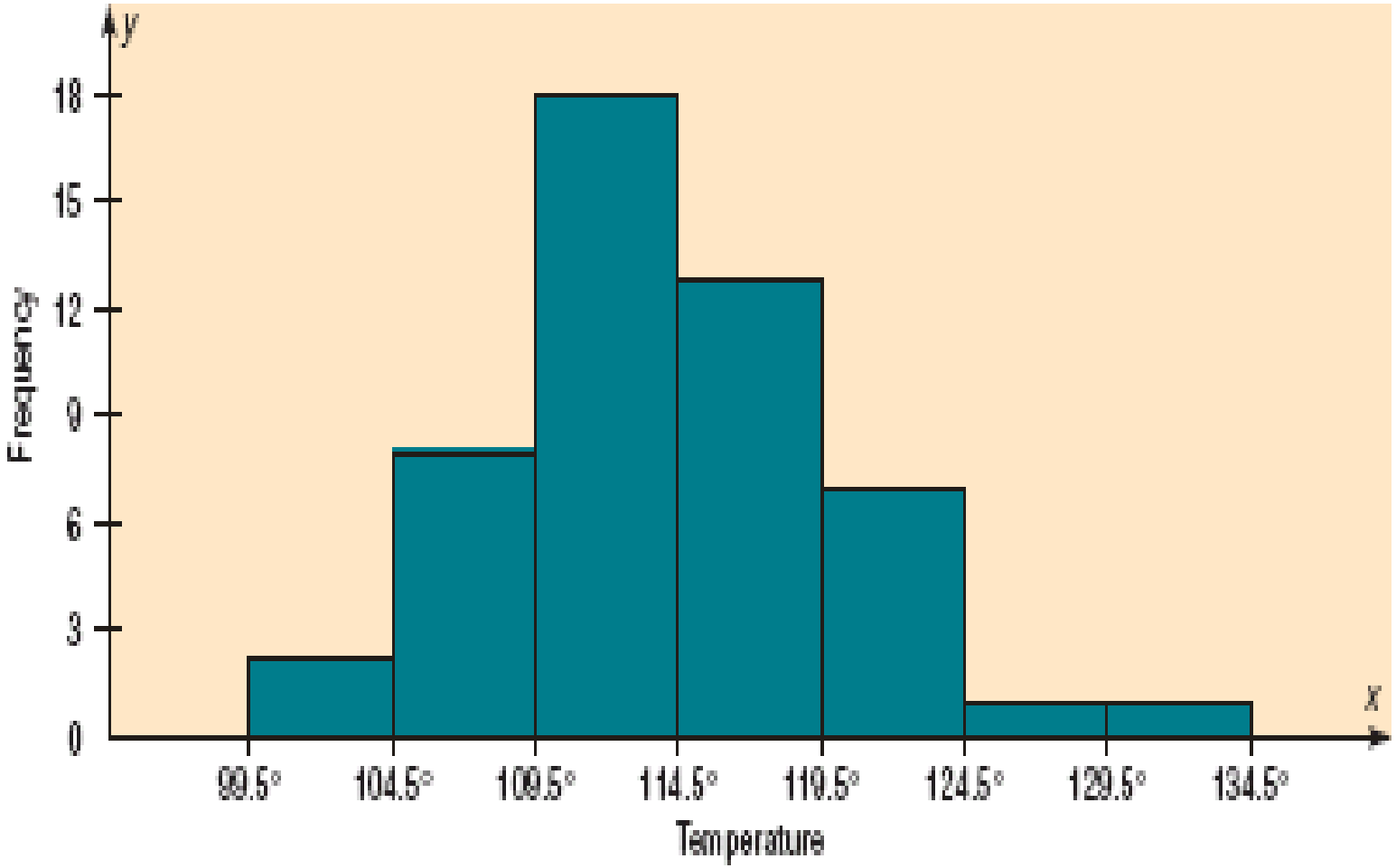
What is Histogram?

A graphic representation of the frequency distribution of a continuous variable. Rectangles are drawn in such a way that their bases lie on a linear scale representing different intervals, and their heights are proportional to the frequencies of the values within each of the intervals.

Uses for a Histogram

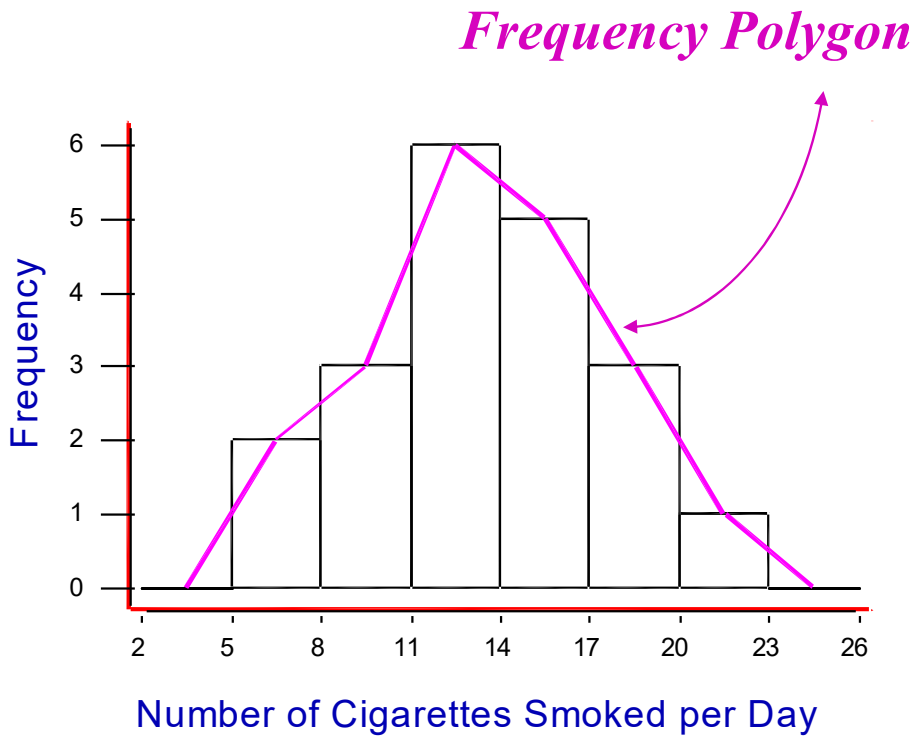
A Histogram can be used:

- to display large amounts of data values in a relatively simple chart form.
- to tell relative frequency of occurrence.
- to easily see the distribution of the data.
- to see if there is variation in the data.
- to make future predictions based on the data.

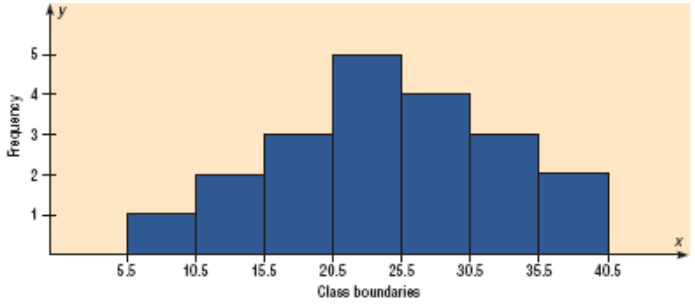

y
18
15
12
9
6
3
0
Frequency
x
99.5°
104.5°
109.5°
114.5°
119.5°
124.5°
129.5°
134.5°
Temperature

The **Frequency Polygon** is a graph that displays the data by using lines that connect points plotted for the frequencies at the midpoints of the classes. The frequencies are represented by the heights of the points.

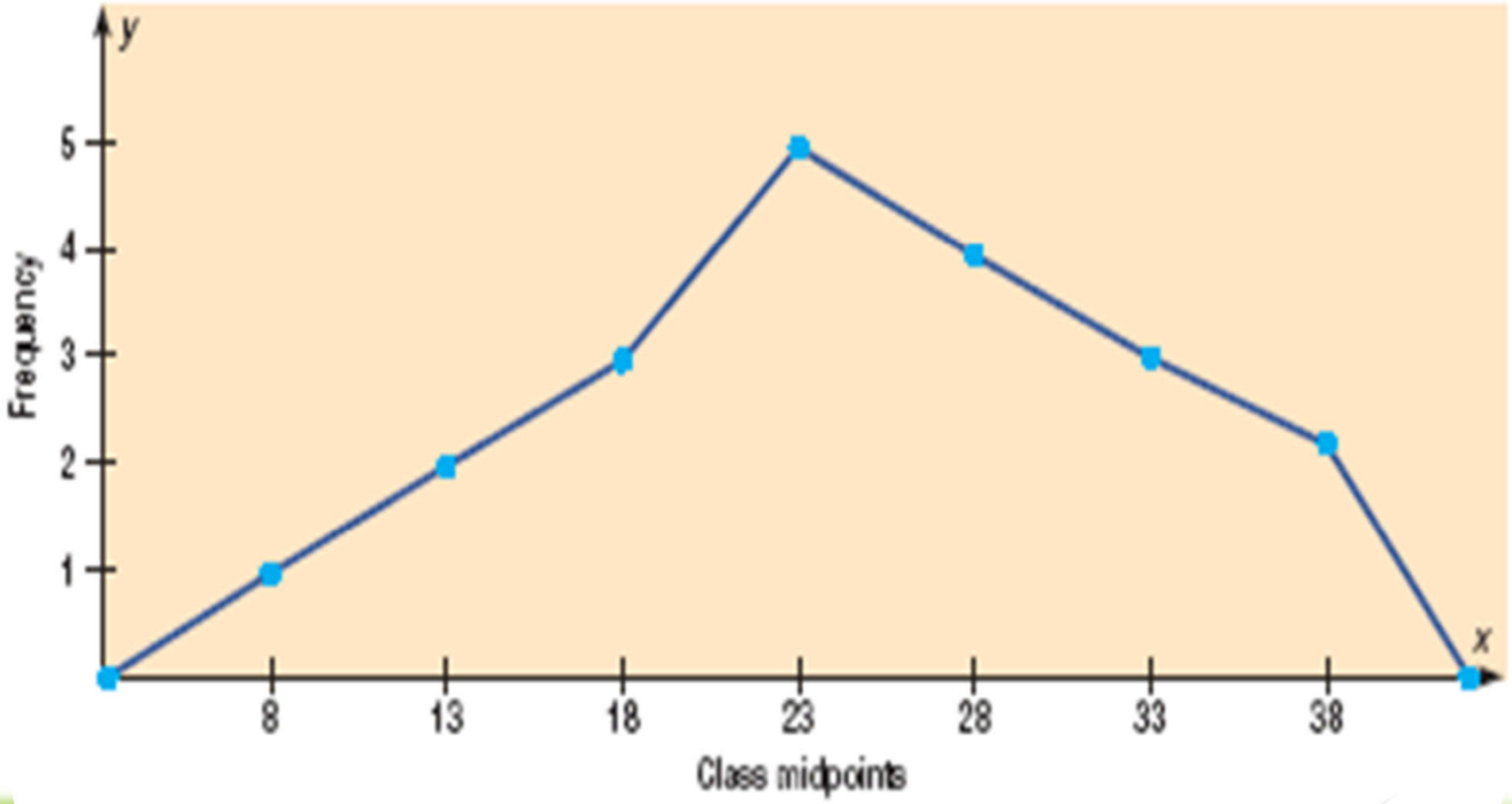
Example of Frequency Polygon



Frequency Polygon



Frequency Polygon



The **Frequency Curve** is a graph that displays the data by using Smooth lines that connect points plotted for the frequencies at the midpoints of the classes. The frequencies are represented by the heights of the points.