Measure of

Central Tendency

Average

Lecture 3+4

Average (Central Tendency)

Average is a number which represents a whole data.

The average means

↓

the center of

↓

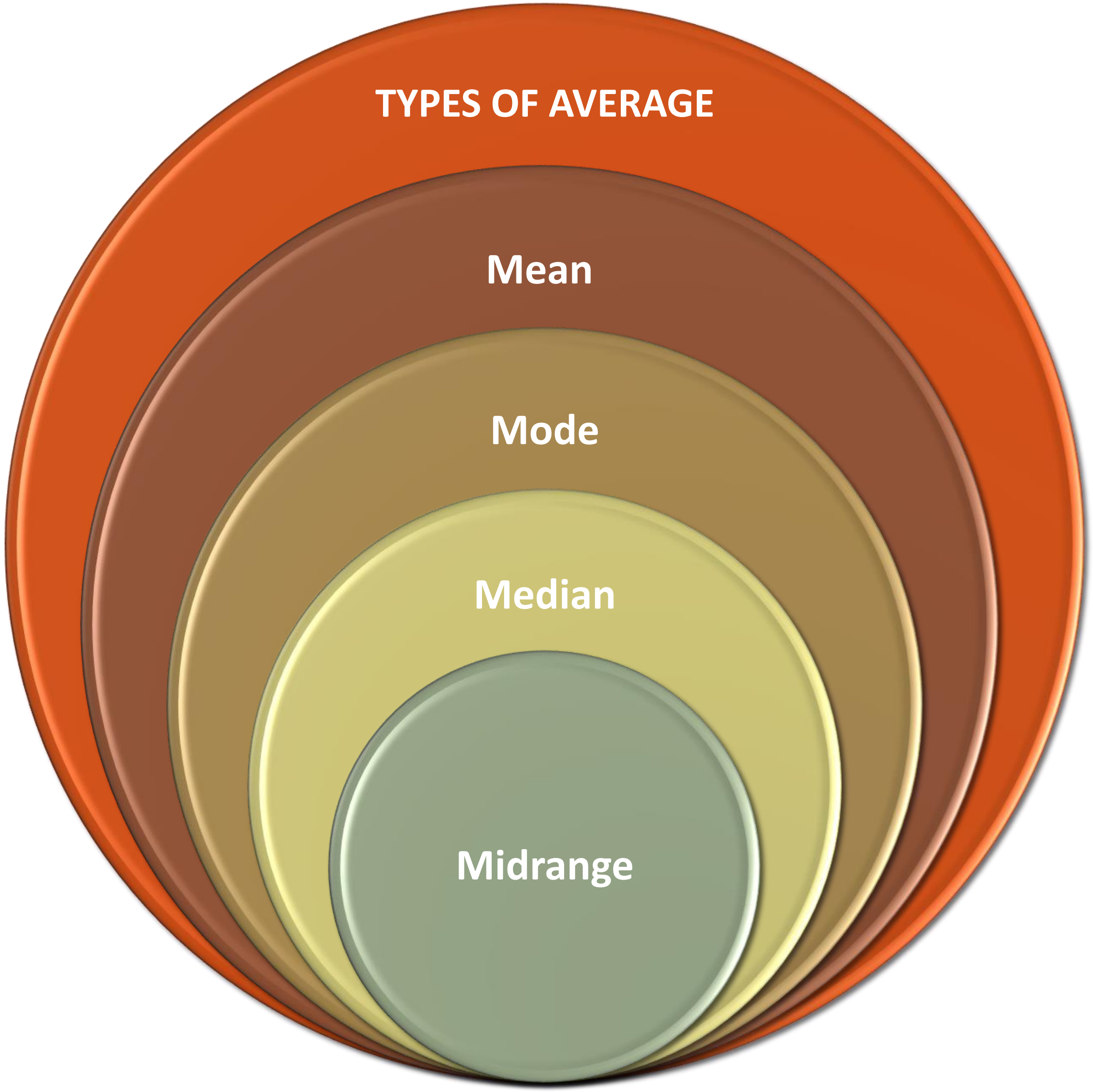
the Values

↓

Observation

↓

Quantities


TYPES OF AVERAGE
Mean
Mode
Median
Midrange

Mean (Arithmetic Mean)

The mean is the sum of the values, divided by the total number of values.

$$Average = \frac{Sum\ of\ values}{Total\ number\ of\ Values}$$

$$\bar{x} = \frac{x_1 + x_2 + x_3 + x_4 + \cdots + x_n}{n}$$

$$\bar{x} = \frac{\sum x_i}{n}$$

Example 1 : Find Mean of 5 , 7 , 8 ,10,5

Solution : As $Average = \frac{Sum\ of\ values}{Total\ number\ of\ Values}$

So $\bar{x} = \frac{5+7+8+10+5}{5}$

$= \frac{35}{5}$

=7 Answer

Example 2 : The number of calls that a local police department responded to for a sample of 9 months is shown. Find the mean. 475, 447, 440, 761, 993, 1052, 783, 671, 621

Solution: $$Average = \frac{Sum\ of\ values}{Total\ number\ of\ Values}$$

So $$\bar{x} = \frac{475+447+440+761+993+1052+783+671+621}{9}$$

$$\bar{x} = \frac{6243}{9}$$

$\bar{x} \approx 693.7$ (Approximately)

Mean of Grouped Data

Mean for grouped Data

$$\bar{x} = \frac{\sum f_i x_i}{n}$$

Where

$\bar{x}$ is mean

f_i is frequency of class

x_i is mid point of class

n is cumulative frequencies

Q 2) find A.M for grouped data of USA 50 states.

CLASS	FREQUENCY	MIDPOINTS	fixi.
100-104	2	102	204
105-109	8	107	856
110-114	19	112	2128
115-119	13	117	1521
120-124	6	122	732
125-129	1	127	127
130-134	1	132.	132.
	$\boxed{\Sigma f_i = 50}$		$\boxed{\Sigma f_i x_i = 5700.}$

$$\bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{5700}{50} = 114.$$

$$\boxed{\bar{x} = 114.}$$

Example: Find Mean of these data represent the net worth (inMillions) of 46 national cooperations.

Class Limits	Frequencies
10-20	3
21-31	8
32-42	15
43-53	7
54-64	10
65-75	3

The Weighted Mean

Sometimes you must find the mean of a data set in which not all values are equally represented.

When there number values of a quantity , we use weighed mean technique to calculate.

The weighted mean of a variable X is calculated by multiplying each value by its corresponding weight and dividing the sum of the products by the sum of the weights.

Example

A student received an A in English Composition I (3 credits), a C in Introduction to Psychology (3 credits), a B in Biology I (4 credits), and a D in Physical Education (2 credits). Assuming A = 4 grade points, B = 3 grade points, C = 2 grade points, D = 1 grade point, and F = 0 grade points, find the student's grade point average.

SOLUTION

Course	Credits (w)	Grade (X)
English Composition I	3	A (4 points)
Introduction to Psychology	3	C (2 points)
Biology I	4	B (3 points)
Physical Education	2	D (1 point)

$$\overline{X} = \frac{\Sigma wX}{\Sigma w} = \frac{3 \cdot 4 + 3 \cdot 2 + 4 \cdot 3 + 2 \cdot 1}{3 + 3 + 4 + 2} = \frac{32}{12} \approx 2.7$$

The grade point average is 2.7.

The Median

The median is the midpoint of the data array.

Before finding this point, the data must be arranged in ascending or increasing order.

When the data set is ordered, it is called a data array.

Example: The number of tornadoes that have occurred in the United States over an 7-year period follows. Find the median. 684, 764, 656, 702, 856, 1133 , 1132

Solution

Step 1 : Arranging the data values in ascending order.

684, 764, 656, 702, 856, 1133 , 1132

Step 2: Data in ascending order

656, 684, 702, 764, 856, 1132, 1133

Step 3: the middle value in the data is 764 .

Hence Median is 764

Example: The number of tornadoes that have occurred in the United States over an 8-year period follows. Find the median. 684, 764, 656, 702, 856, 1133 , 1132,1303

Solution

Step 1 : Arranging the data values in ascending order.

684, 764, 656, 702, 856, 1133 , 1132, 1303

Step 2: Data in ascending order

656, 684, 702, 764, 856, 1132, 1133,1303

Step 3: the middle values in the data are 764 and 856

Hence Median is $\frac{764+856}{2} = \frac{1620}{2} = 810$ Answer

MEDIAN OF GROUPED DATA

Median of Grouped Data

$$= l + \frac{h}{f}\left(\frac{n}{2} - C\right)$$

Where l is lower class boundary

n is sum of frequencies of all classes

f is frequency of median class

h is width of median class

C is cumulative frequency preceding median class

Steps for finding the Median

Step 1: $\frac{n}{2}$ half of total observations

Step 2 : Median Class

Step 3 : terms of median class f and h

Step 4 : Cumulative Frequencies (C) before median class

Example: Find Median of these data represent the net worth (inMillions) of 46 national cooperations.

Class Limits	Frequencies
10-20	3
21-31	8
32-42	15
43-53	7
54-64	10
65-75	3

Solution :

Class Limits	Frequencies	Cummulative Frequency C
10-20	3	3
21-31	8	11
32-42	***15***	***26***
43-53	7	33
54-64	10	43
65-75	3	46
	$\sum f = 46$	

32-42 is median class because C is greater than $\frac{n}{2}$ here.

So Frquency of median class is $f = 15$

Width of of median class is $h = 11 \quad and \quad l = 31.5$

Median for grouped data is given by

$$= l + \frac{h}{f}\left(\frac{n}{2} - C\right)$$

$$= 31.5 + \frac{11}{15}\left(\frac{46}{2} - 11\right)$$

$$= 31.5 + 8.8$$

$= 40.3$ Answer

Mode

The value that occurs most often in a data set is called the mode.

Example : Find the mode of the temperature °C of week of June in Karachi, Sindh

40 , 45 , 44 , 44,43,42 , 44

Solution : Here most frequent number is 44

so 44 is mode

There can be more than one Mode(s) or no mode at all

Example 1: 4,7,7,8,9,3,7   here 7 is mode (Unimodal data)

Example1:   5,6,5,6,5,6   here 5 and 6 are modes (Bimodal data)

Example 2: 3,6,4,3,6,4   here 3,4 are 6 are modes (Multimodal data)

MODE OF GROUPED DATA

Mode of Grouped Data

$$= l + \frac{(f_m - f_1)}{(f_m - f_1) + (f_m - f_2)} . h$$

Where l is lower class boundary of modal class

f_1 is frequency preceding modal class

f_m is frequency of modal class

f_2 is frequency succeding modal class

h is width of median class

Example: Find Mode of these data represent the net worth (inMillions) of 46 national cooperations.

Class Limits	Frequencies
10-20	3
21-31	8
32-42	15
43-53	7
54-64	10
65-75	3

Solution: Table Modal Class is 32-42 with frequency $f_m = 15$

Class Limits	Frequencies
10-20	3
21-31	8
32-42	15
43-53	7
54-64	10
65-75	3

$f_1 = 8$, $f_2 = 7$, $h = 11$ and $l = 31.5$

$$\textbf{As Mode} = l + \frac{(f_m - f_1)}{(f_m - f_1) + (f_m - f_2)} . h$$

$$= 31.5 + \frac{(15-8)}{(15-8)+(15-7)} . 11$$

$$= 31.5 + 5.133$$

$$= 36.633$$

Midrange

The midrange is a rough estimate of the middle

The midrange is defined as the sum of the smallest and largest values in the data set, divided by 2.

$$Midrange = \frac{smallest\ valu\ \ + Larges\ value}{2}$$

Example : The number of bank failures for a recent five-year period is shown. Find the midrange. 3, 30, 148, 157, 71

Solution : As $Midrange = \frac{smallest\ valu + Larges\ value}{2}$

Smallest Value =3 and Largest Value = 157

so $MR = \frac{3+157}{2} = \frac{160}{2}$

$\mathrm{MR} = 80$ Answer

Properties and Uses of Central Tendency

The Mean

1. The mean is found by using all the values of the data.
2. The mean varies less than the median or mode when samples are taken from the same population and all three measures are computed for these samples.
3. The mean is used in computing other statistics, such as the variance.
4. The mean for the data set is unique and not necessarily one of the data values.
5. The mean cannot be computed for the data in a frequency distribution that has an open-ended class.
6. The mean is affected by extremely high or low values, called outliers, and may not be the appropriate average to use in these situations.

The Median

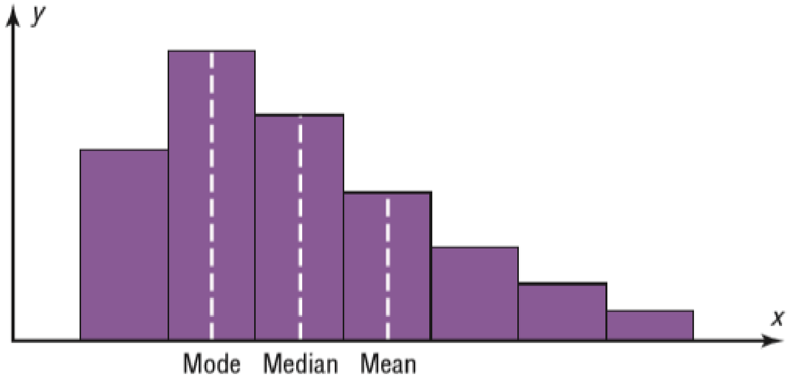
1. The median is used to find the center or middle value of a data set.
2. The median is used when it is necessary to find out whether the data values fall into the upper half or lower half of the distribution.
3. The median is used for an open-ended distribution.
4. The median is affected less than the mean by extremely high or extremely low values.

The Mode

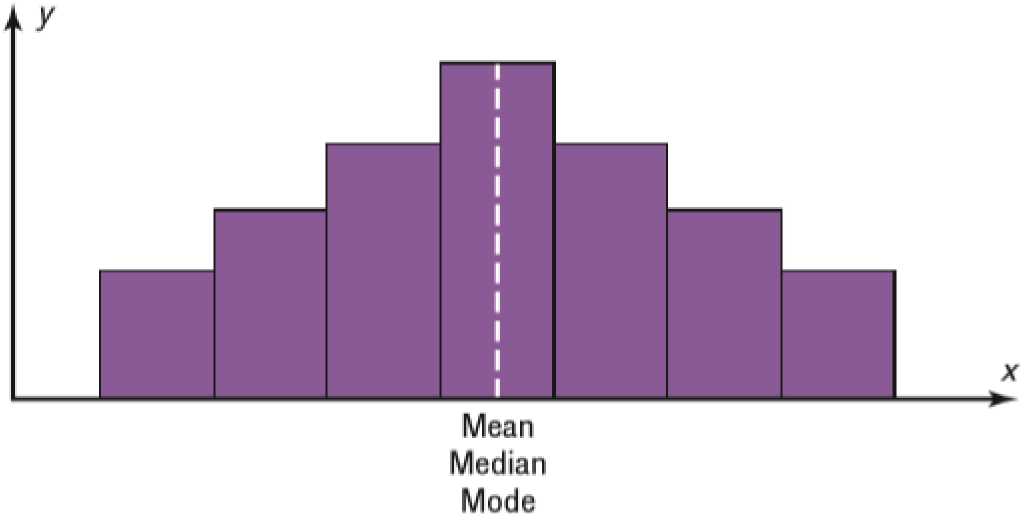
1. The mode is used when the most typical case is desired.
2. The mode is the easiest average to compute.
3. The mode can be used when the data are nominal or categorical, such as religious preference, gender, or political affiliation.
4. The mode is not always unique. A data set can have more than one mode, or the mode may not exist for a data set.

The Midrange

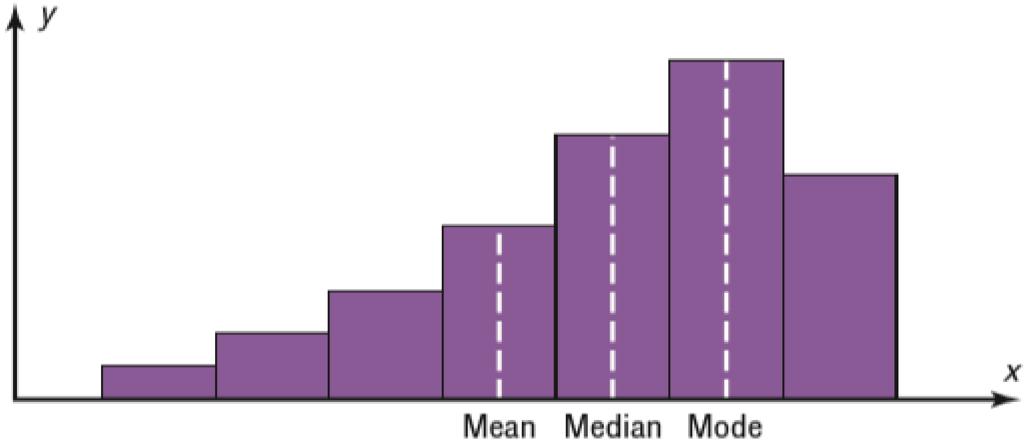
1. The midrange is easy to compute.
2. The midrange gives the midpoint.
3. The midrange is affected by extremely high or low values in a data set.



(a) Positively skewed or right-skewed



(b) Symmetric



(c) Negatively skewed or left-skewed