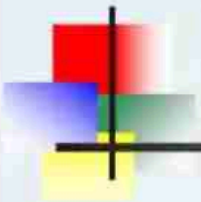
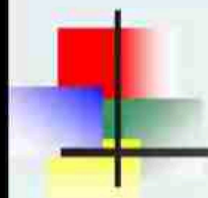


Numerical Descriptive Measure

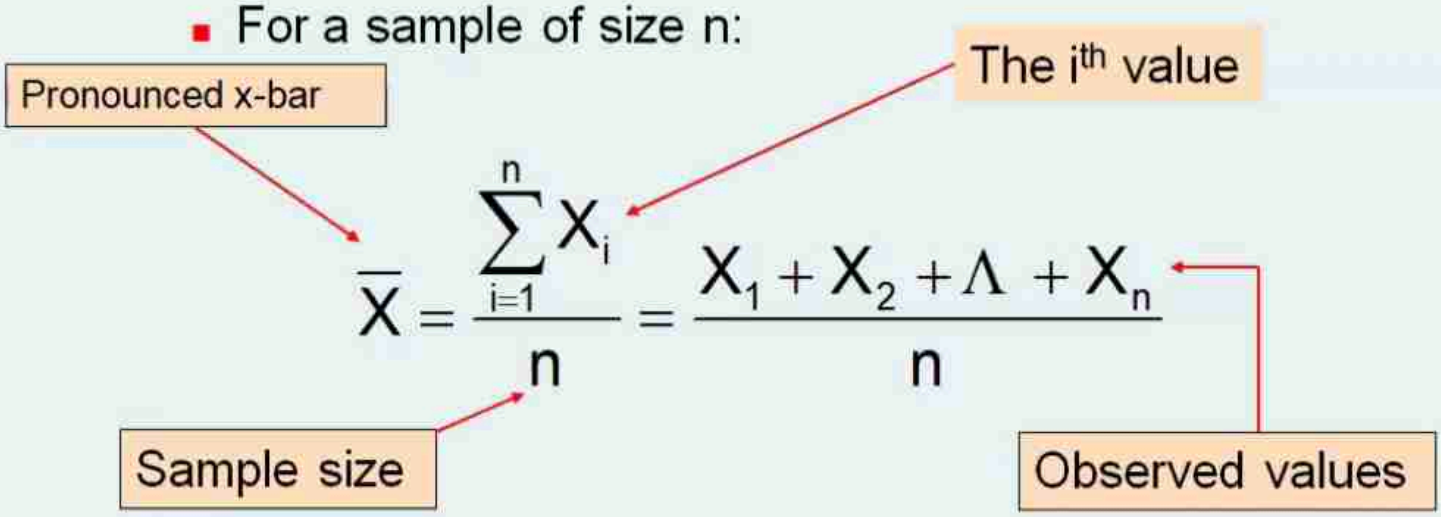


Descriptive Statistics

- The **central tendency** is the extent to which all the data values group around a typical or central value.
- The **variation** is the amount of dispersion or scattering of values
- The **shape** is the pattern of the distribution of values from the lowest value to the highest value.

Measures of Central Tendency: The Mean

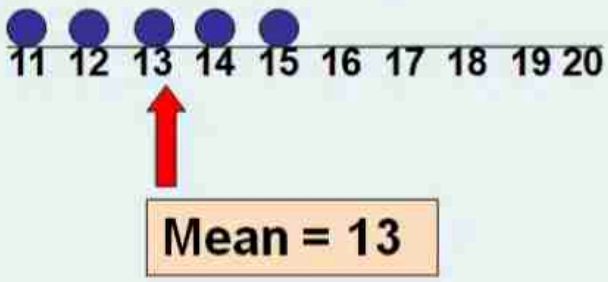
- The arithmetic mean (often just called the "mean") is the most common measure of central tendency
  - For a sample of size n:



Measures of Central Tendency: The Mean

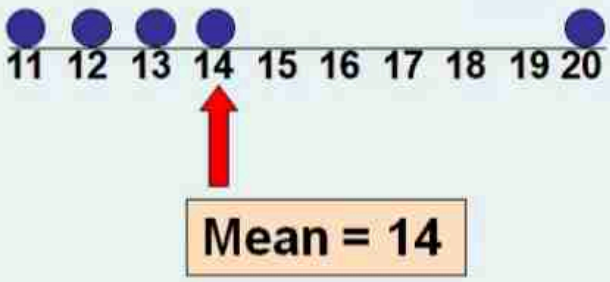
(continued)

- The most common measure of central tendency
- Mean = sum of values divided by the number of values
- Affected by extreme values (outliers)



Mean = 13

$$\frac{11+12+13+14+15}{5} = \frac{65}{5} = 13$$



Mean = 14

$$\frac{11+12+13+14+20}{5} = \frac{70}{5} = 14$$

Mean for Group Data

Formula for Mean is given by $\bar{X} = \frac{\sum f(X)}{n}$

Where

$\bar{X}$ = Mean

$\sum f(X)$ = Sum of cross products of frequency in each class with midpoint X of each class

n = Total number of observations (Total frequency) = $\sum f$

Example

Find the arithmetic mean for the following continuous frequency distribution:

Class	0-1	1-2	2-3	3-4	4-5	5-6
Frequency	1	4	8	7	3	2

	A	B	C	D
1	Class	X (mid pt)	f	fX
2	0-1	0.5	1	0.5
3	1-2	1.5	4	6.0
4	2-3	2.5	8	20.0
5	3-4	3.5	7	24.5
6	4-5	4.5	3	13.5
7	5-6	5.5	2	11.0
8	Totals		25	75.5
9	Mean			3.02

Another Example

Class interval		f
0	49.99	78
50	99.99	123
100	149.99	187
150	199.99	82
200	249.99	51
250	299.99	47
300	349.99	13
350	399.99	9
400	449.99	6
450	499.99	4
		600

By taking mid values as 25, 75,... 475.

$$\bar{X} = \frac{\sum f(X)}{n}$$

Mean: 142.25

f(X) = 85350

n=600

Mean using Coding

Class	f
0-7	2
8-15	6
16-23	3
24-31	5
32-39	2
40-47	2

$$\text{Mean} = x0 + w * \frac{(\text{Summation of } u*f))}{n}$$

w=numerical width of class interval
X0=value of midpoint assigned code 0

Cont'd…

Class	mid	f	Code (u)	u*f
0-7	3.5	2	-2	-4
8-15	11.5	6	-1	-6
16-23	19.5	3	0	
24-31	??	5	1	??
32-39	??	2	2	4
40-47	43.5	2	3	6
		20		5

$$\textbf{Mean} = x_0 + w * \frac{\text{Summation of } u*f}{n}$$

$= 19.5+8* (5) / (20) = 21.5$

w=numerical width of class interval
X0=value of midpoint assigned code 0

Weighted Mean

Weighted mean: A **weighted mean** is a kind of **average**. Instead of each data point contributing **equally** to the final mean, some data points contribute **more "weight" than others**.

To calculate an average that takes into account the importance of each value to the overall cost . Find out **average cost of labor per hour for** each of the product

		Labor hrs per unit of output	
Grade of labor	Hourly wage	Product 1	Product 2
Unskilled	5	1	4
Semiskilled	7	2	3
Skilled	9	5	3

A simple arithmetic mean = (5+7+9) / 3= 7/hr
Using this, labor cost of 1 unit of product 1 to be = 7* (1+2+5) = 56
2 = 7* (4+3+3) = 70
Both are incorrect, the answers must take into account **that different amount of each grade of labor** .

P1= avg cost of labor per hr = (5*1+7*2+9*5) 8 = 8
P2= avg cost of labor per hr = (5*4+7*3+9*3) 10 = 6.80