

SESSION 2 OF

STATISTICS FOR BUSINESS

DR. SURESH KUMAR GADI

UNIVERSIDAD DE LAS AMÉRICAS PUEBLA



WHAT ARE WE LEARNING TODAY?

Grouped data

- Relative frequency
- Mean
- Median, Quartiles, Deciles, and Percentile
- Mode
- Standard deviation and Variance

Multiple variables

HOW TO PREPARE GROUPED DATA

EXCEL EXAMPLE

TERMINOLOGY

Class boundaries	Lower boundary (L_i)	Upper boundary (U_i)	Midpoint (x_i)	Limits	Frequency (f_i)	Cumulative frequency (c_i)
0 – 10	0	10	5	≥ 0 and < 10	3	3
10 – 20	10	20	15	≥ 10 and < 20	21	24
20 – 30	20	30	25	≥ 20 and < 30	26	50
30 – 40	30	40	35	≥ 30 and < 40	9	59
Total (n)					59	

Class interval = class width = class height (h) = 10

RELATIVE FREQUENCY (RF)

The relative frequency (f_R) is given by

$$f_R = \frac{f_i}{\sum f_i} = \frac{f_i}{n}$$

where,

f_i is frequency of the i^{th} class and

n is total number of data elements

RELATIVE FREQUENCY

L	U	f	f_R
0	10	62	$62/714=0.087$
10	20	102	$102/714=0.143$
20	30	220	$220/714=0.308$
30	40	167	$167/714=0.234$
40	50	89	$89/714=0.125$
50	60	48	$48/714=0.067$
60	70	19	$19/714=0.027$
70	80	6	$6/714=0.008$
80	90	1	$1/714=0.001$
90	100	0	$0/714=0.000$
Sum		714	1

MEAN

Mean ($\mu = \bar{x}$) of grouped data is

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

where,

m_i is midpoint of the class-interval,

f_i is the frequency.

EXAMPLE

01

Generate
Random data

02

Group them

03

Obtain mean
from the
ungrouped
data

04

Obtain mean
from the
grouped data

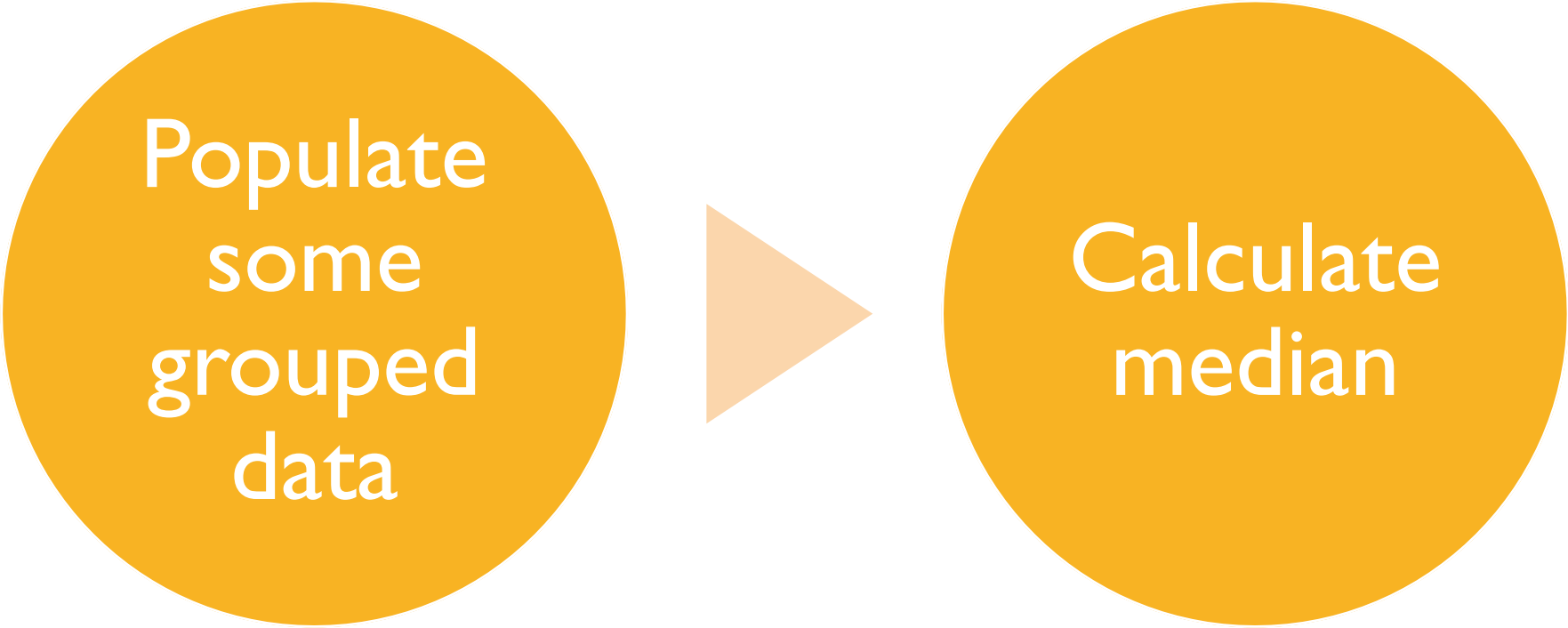
MEDIAN

Median (\tilde{x}) of the grouped data is

$$\tilde{x} = L_m + \left[\frac{\frac{n}{2} - c_{m-1}}{f_m} \right] h,$$

where,

L_m is the lower boundary of the median class



```
graph LR; A((Populate some grouped data)) --> B((Calculate median))
```

Populate
some
grouped
data

Calculate
median

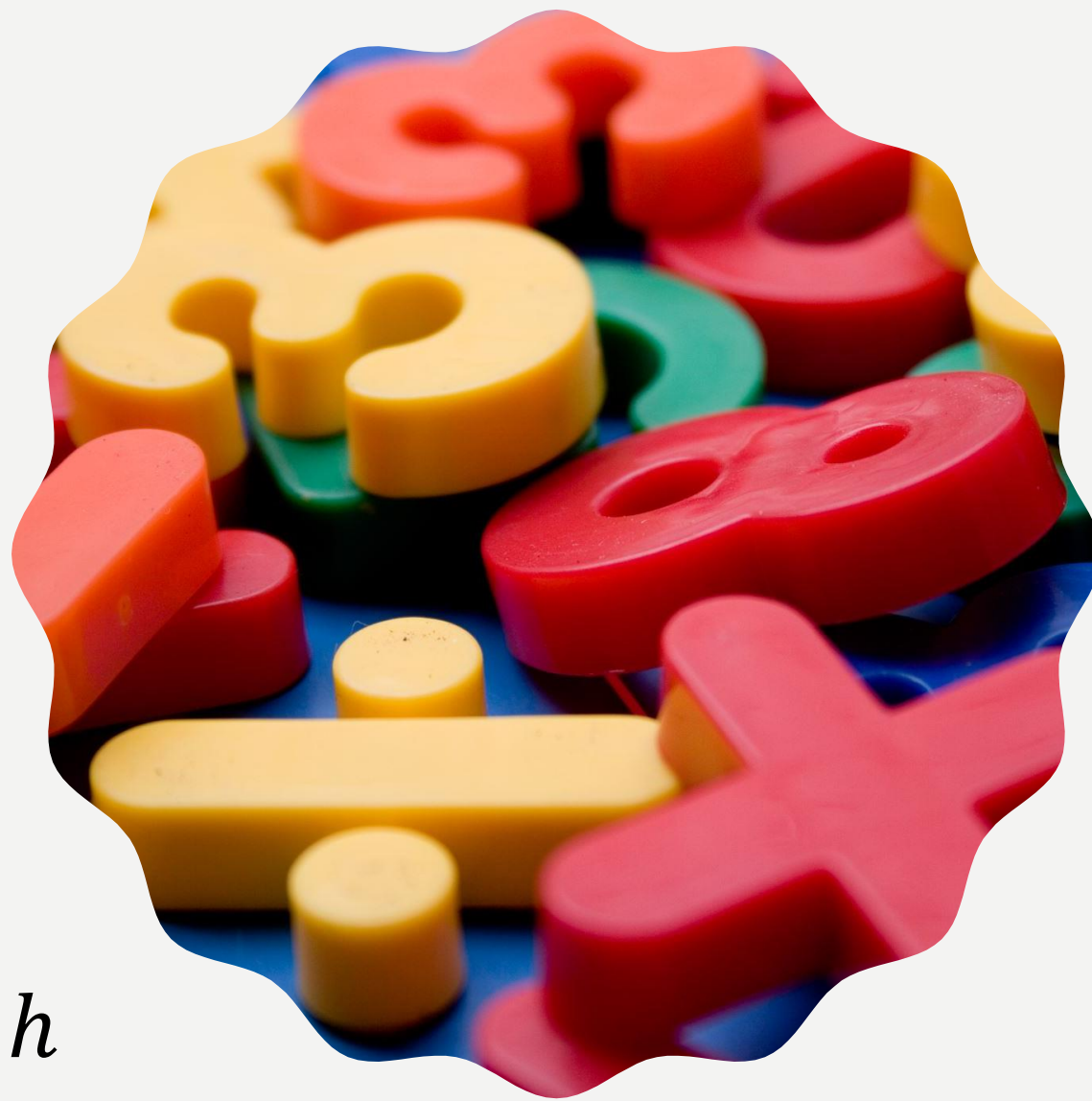
EXAMPLE

QUARTILES, DECILES, AND PERCENTILE

Calculations are like Median.

Replace the L_m and $\frac{n}{2}$ with the correct proportion.

$$L_{\left(\frac{i}{*}\right)} + \left[\frac{\left(\frac{i}{*}\right)n - c_{m-1}}{f_m} \right] h$$



EXAMPLE

1. Generate random grouped data
2. Find 3rd quartile
3. Find 84th percentile
4. Find 7th decile



MODE

$$\text{Mode} = L_M + \left[\frac{f_M - f_{M-1}}{(f_M - f_{M+1}) + (f_M - f_{M-1})} \right] h,$$

where

L_M is the lower boundary of the class with maximum frequency

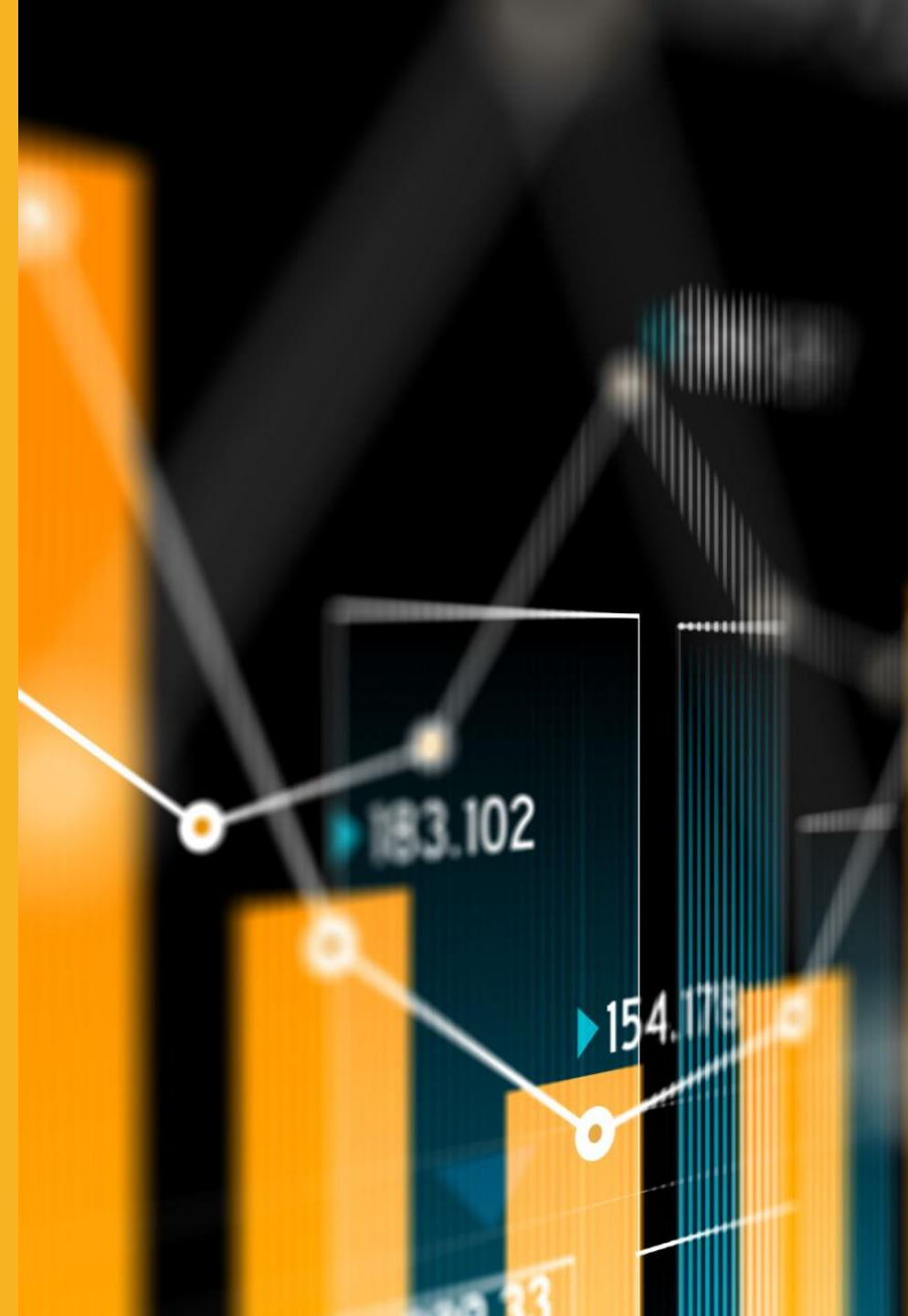
f_M is the maximum frequency

f_{M-1} is the frequency of the class before the maximum frequency

f_{M+1} is the frequency of the class after the maximum frequency

EXAMPLE

FIND MODE FOR A GROUPED DATA



VARIANCE (UNGROUPED DATA)

- Population variance (σ^2) formula is

$$\sigma^2 = \frac{\sum[(\bar{x} - x_i)^2]}{n}$$

- Sample variance (σ^2) formula is

$$\sigma^2 = \frac{\sum[(\bar{x} - x_i)^2]}{n - 1}$$

VARIANCE (GROUPED DATA)

- Population variance (σ^2) formula is

$$\sigma^2 = \frac{\sum[f_i(\bar{x} - x_i)^2]}{n}$$

- Sample variance (σ^2) formula is

$$\sigma^2 = \frac{\sum[f_i(\bar{x} - x_i)^2]}{n - 1}$$

STANDARD DEVIATION

- Population standard deviation(σ) formula is

$$\sigma = \sqrt{\sigma^2}$$

- Sample standard deviation (σ) formula is

$$\sigma = \sqrt{\sigma^2}$$

EXAMPLE

**CALCULATE STANDARD DEVIATION AND
VARIANCE**



MULTIPLE VARIABLES

COVARIANCE (UNGROUPED DATA)

- Population covariance formula is

$$\text{cov}(X, Y) = \frac{\sum[(\bar{x} - x_i)(\bar{y} - y_i)]}{n}$$

- Sample covariance formula is

$$\text{cov}(X, Y) = \frac{\sum[(\bar{x} - x_i)(\bar{y} - y_i)]}{n - 1}$$

COVARIANCE (GROUPED DATA)

- Population covariance formula is

$$\text{cov}(X, Y) = \frac{\sum [f_i(\bar{x} - x_i)(\bar{y} - y_i)]}{n}$$

- Sample covariance formula is

$$\text{cov}(X, Y) = \frac{\sum [f_i(\bar{x} - x_i)(\bar{y} - y_i)]}{n - 1}$$

EXAMPLE

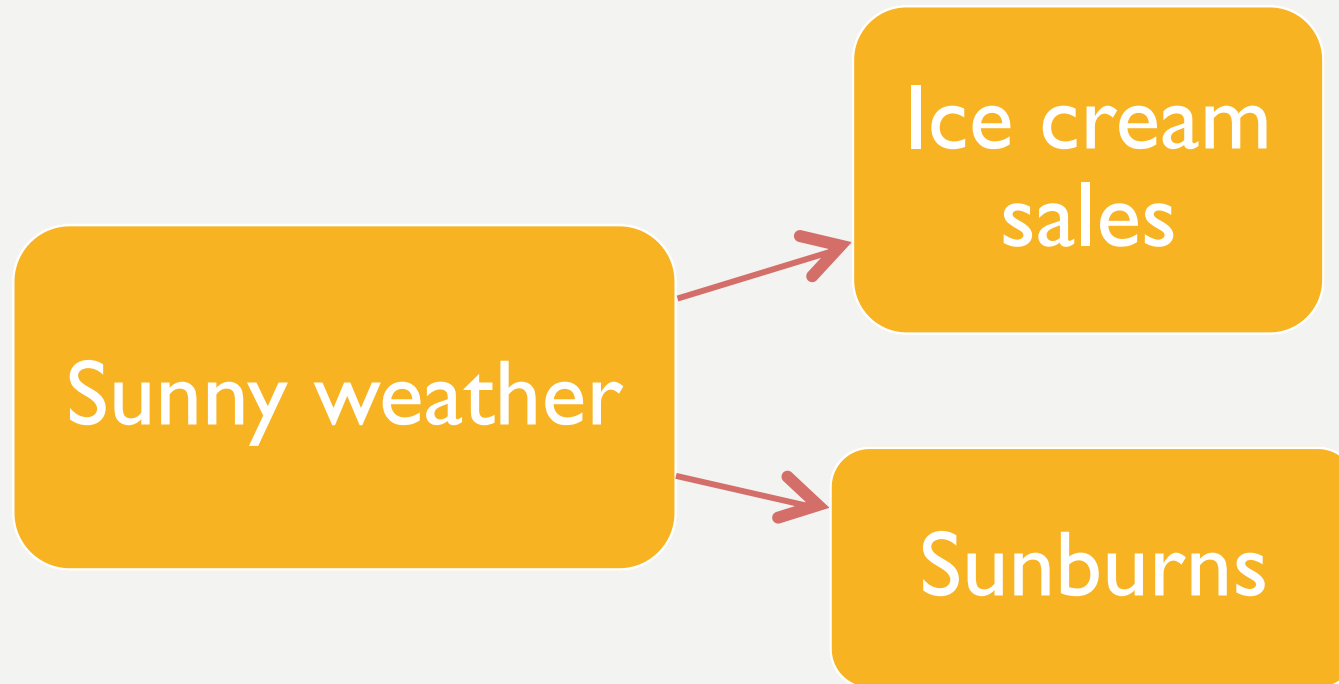


GENERATE GROUPED
DATA FOR TWO VARIABLES



CALCULATE COVARIANCE

CORRELATION



PEARSON CORRELATION COEFFICIENT

Pearson correlation coefficient (ρ) is

$$\rho = \frac{\text{cov}(X, Y)}{\sigma_x \sigma_y}$$

where,

σ_x is standard deviation of variable X

σ_y is standard deviation of variable Y

PROPERTIES OF FACTOR OF CORRELATION

Value between
1 and -1

Significance of
positive and
negative signs

THANK YOU

