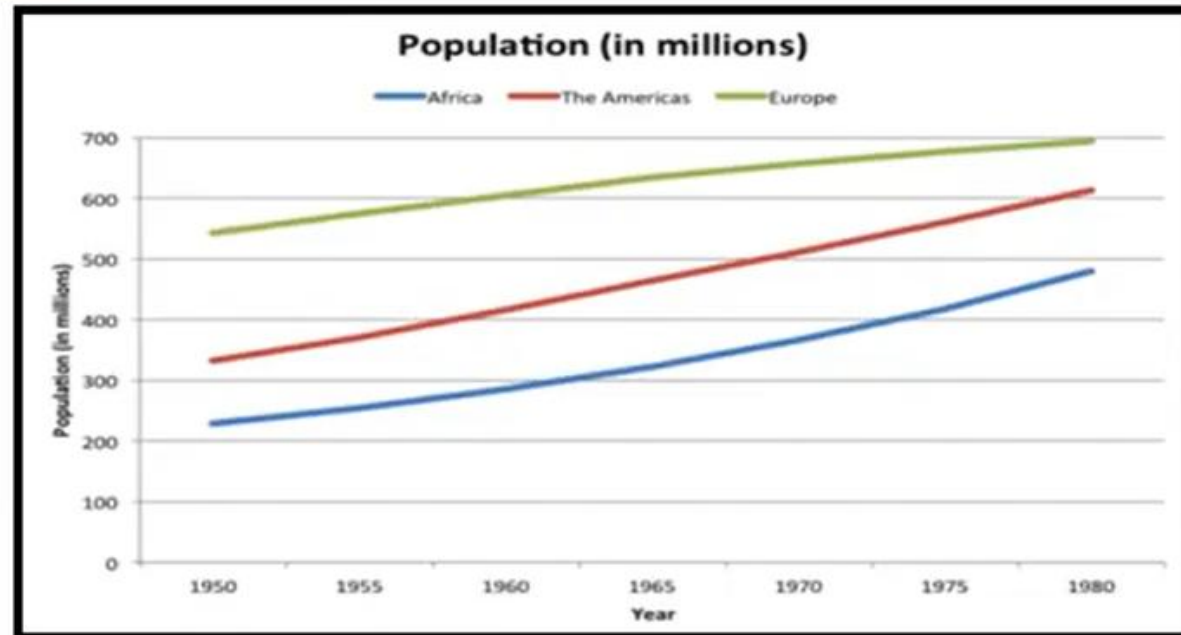


# Multiple Linear Regression

## Multiple Linear Regression

Dependent variable (DV)      Independent variables (IVs)

$$y = b_0 + b_1 * x_1 + b_2 * x_2 + \dots + b_n * x_n$$



# Comparison



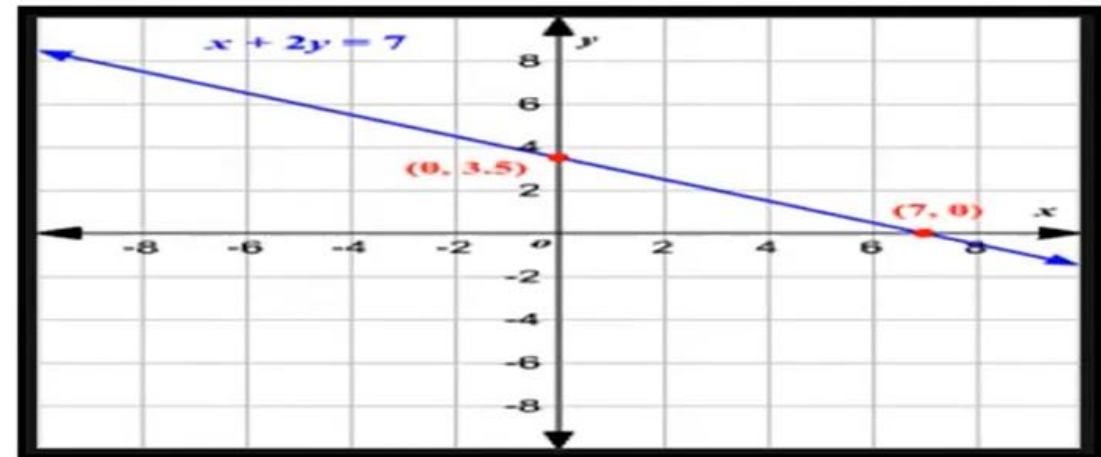
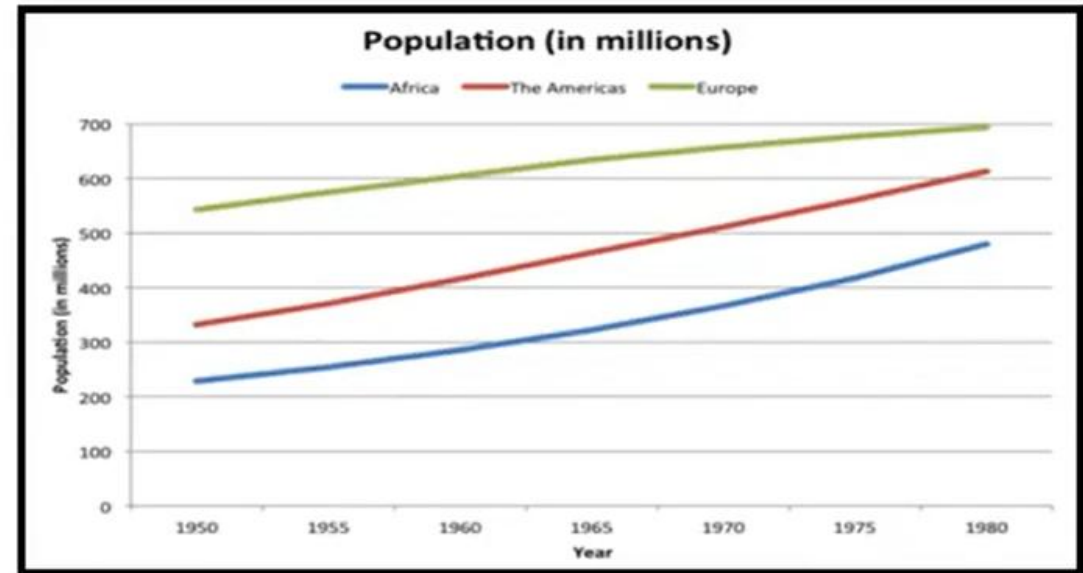
## Multiple Linear Regression

Dependent variable (DV)      Independent variables (IVs)

$$y = b_0 + b_1 * x_1 + b_2 * x_2 + \dots + b_n * x_n$$

## Simple Linear Regression

$$y = b_0 + b_1 * x_1$$



# Problem Statement

## AI in Business Intelligence/Analyst

I/P

	R&D Spend	Administration	Marketing Spend	State	Profit
0	165349.20	136897.80	471784.10	New York	192261.83
1	162597.70	151377.59	443898.53	California	191792.06
2	153441.51	101145.55	407934.54	Florida	191050.39
3	144372.41	118671.85	383199.62	New York	182901.99
4	142107.34	91391.77	366168.42	Florida	166187.94
5	131876.90	99814.71	362861.36	New York	156991.12

O/P

Stage 1  $\rightarrow$  ML  
Stage 2  $\rightarrow$  supervised  
Stage 2  $\rightarrow$  Regression

# Types of Column in the dataset

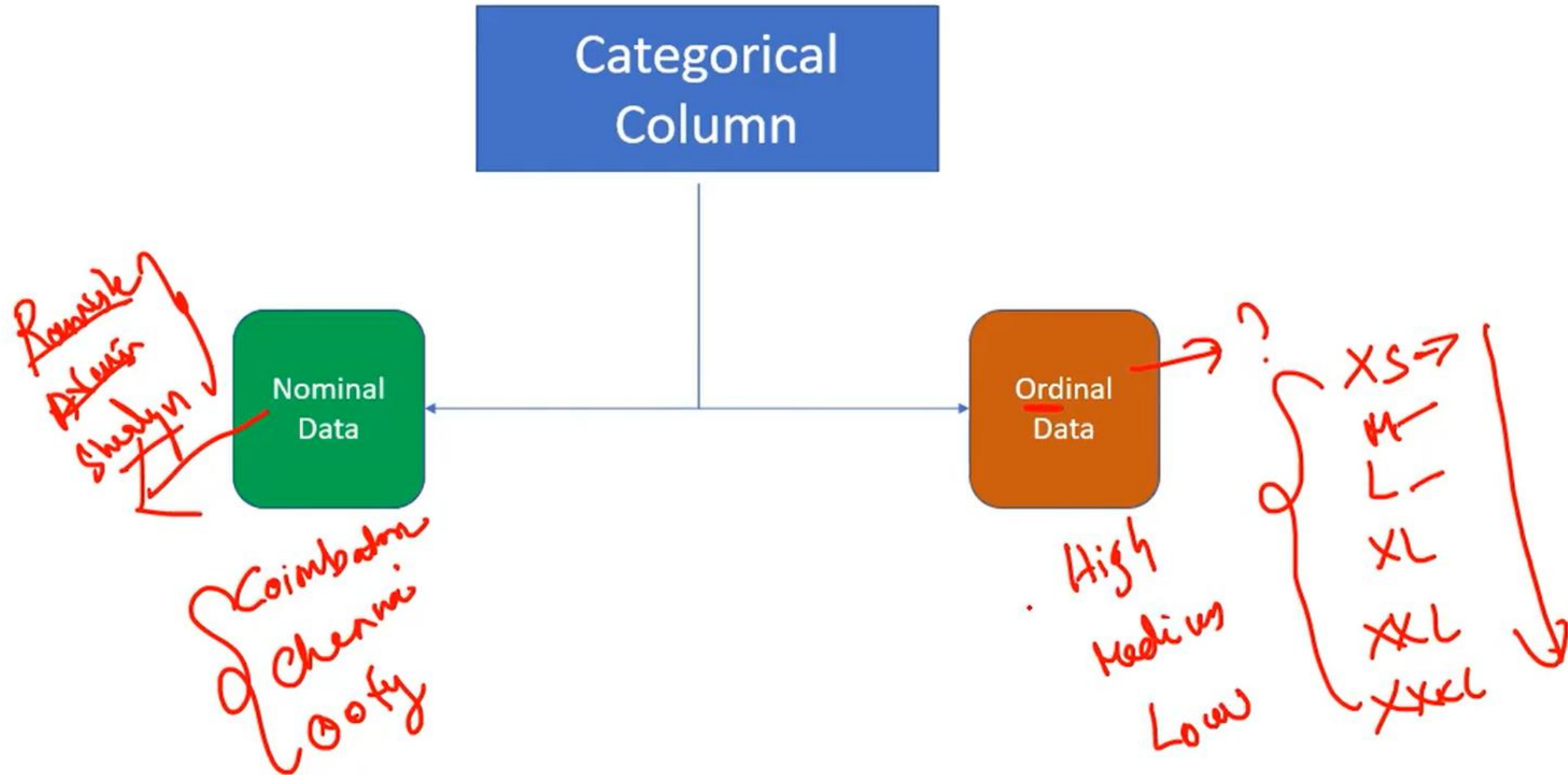
C- categorical value  
Converted in numeric value

	<i>N</i> R&D Spend	<i>N</i> Administration	<i>N</i> Marketing Spend	<i>C</i> State	<i>N</i> Profit
0	165349.20	136897.80	471784.10	New York	192261.83
1	162597.70	151377.59	443898.53	California	91792.06
2	153441.51	101145.55	407934.54	Florida	91050.39
3	144372.41	118671.85	383199.62	New York	182901.99
4	142107.34	91391.77	366168.42	Florida	166187.94
5	131876.90	99814.71	362861.36	New York	156991.12

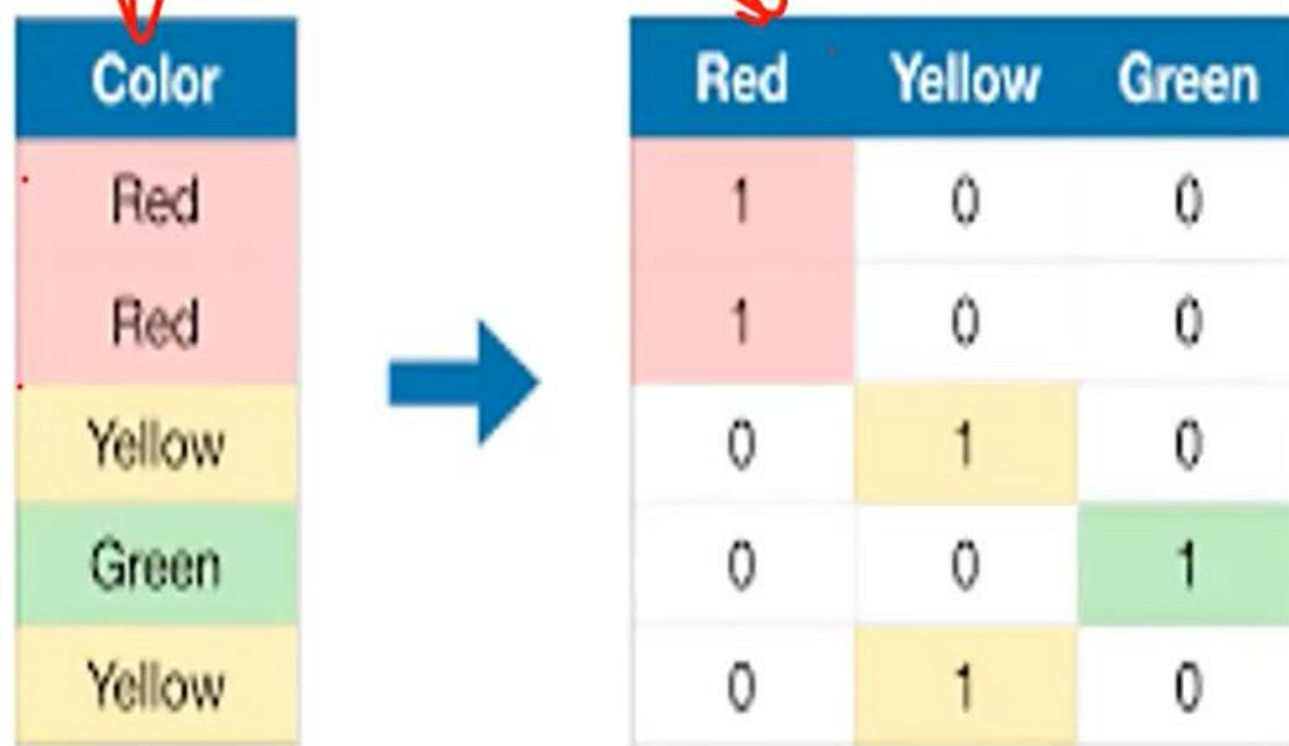
↓  
python → AI → categorical  
data



# How to handle Categorical Column



# Categorical Column- Nominal - One Hot Encoding



The diagram illustrates the process of one-hot encoding for a categorical variable. On the left, a table with a single column 'Color' contains five rows: two 'Red' entries, one 'Yellow' entry, one 'Green' entry, and one 'Yellow' entry. A blue arrow points from this table to a second table on the right. The second table has three columns: 'Red', 'Yellow', and 'Green'. Each row in the second table corresponds to a row in the first table, with a '1' in the column corresponding to the color and '0' in the others. Red arrows from the title point to the 'Color' header and the new column headers 'Red', 'Yellow', and 'Green'.

Color
Red
Red
Yellow
Green
Yellow

Red	Yellow	Green
1	0	0
1	0	0
0	1	0
0	0	1
0	1	0

# Categorical Column- Ordinal –Mapping- Label Encoder

*No Expansion*

SAFETY-LEVEL (TEXT)	SAFETY-LEVEL (NUMERICAL)
None	0
Low	1
Medium	2
High	3
Very-High	4