

Amazon Sales Data Analysis

Detailed Project Report

Sahil & Siddhi

PROJECT DETAIL

Project Title	Amazon Sales Data Analysis
Domain name	E-Commerce
Project Difficulties Level	Advanced
Tools	Jupyter notebook, , Ms Excel, Ms PowerPoint,python.

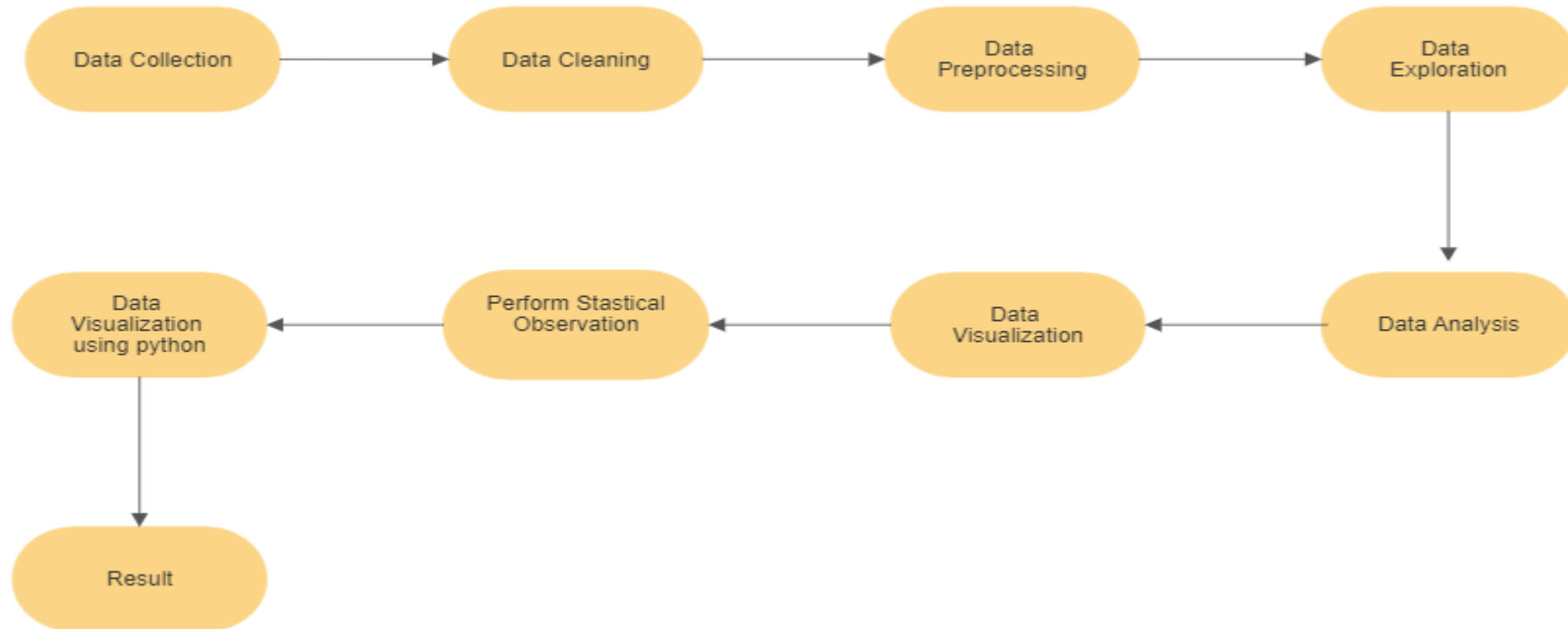
OBJECTIVE

Extract-Transform-Load some Amazon dataset and find
Sales-trend -> month wise , year wise , yearly_month wise

PROBLEM STATEMENT

- Sales management has gained importance to meet increasing competition and the need
- for improved methods of distribution to reduce cost and to increase profits. Sales
- management today is the most important function in a commercial and business
- enterprise.
- Do ETL : Extract-Transform-Load some Amazon dataset and find for me
- Sales-trend -> month wise , year wise , yearly_month wise
- Find key metrics and factors and show the meaningful relationships between attributes.

ARCHITECTURE



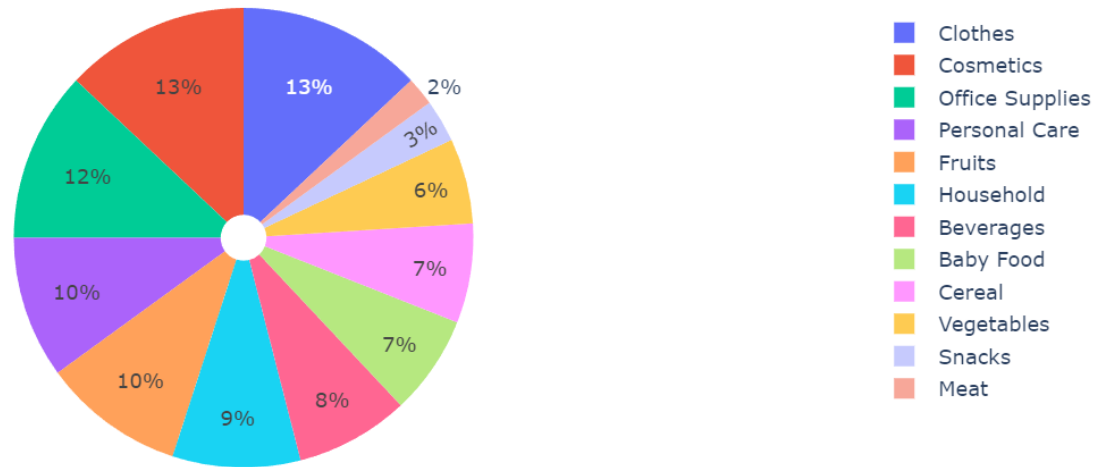
DATASET INFORMATION

The Dataset is taken from iNeuron's provided dataset-

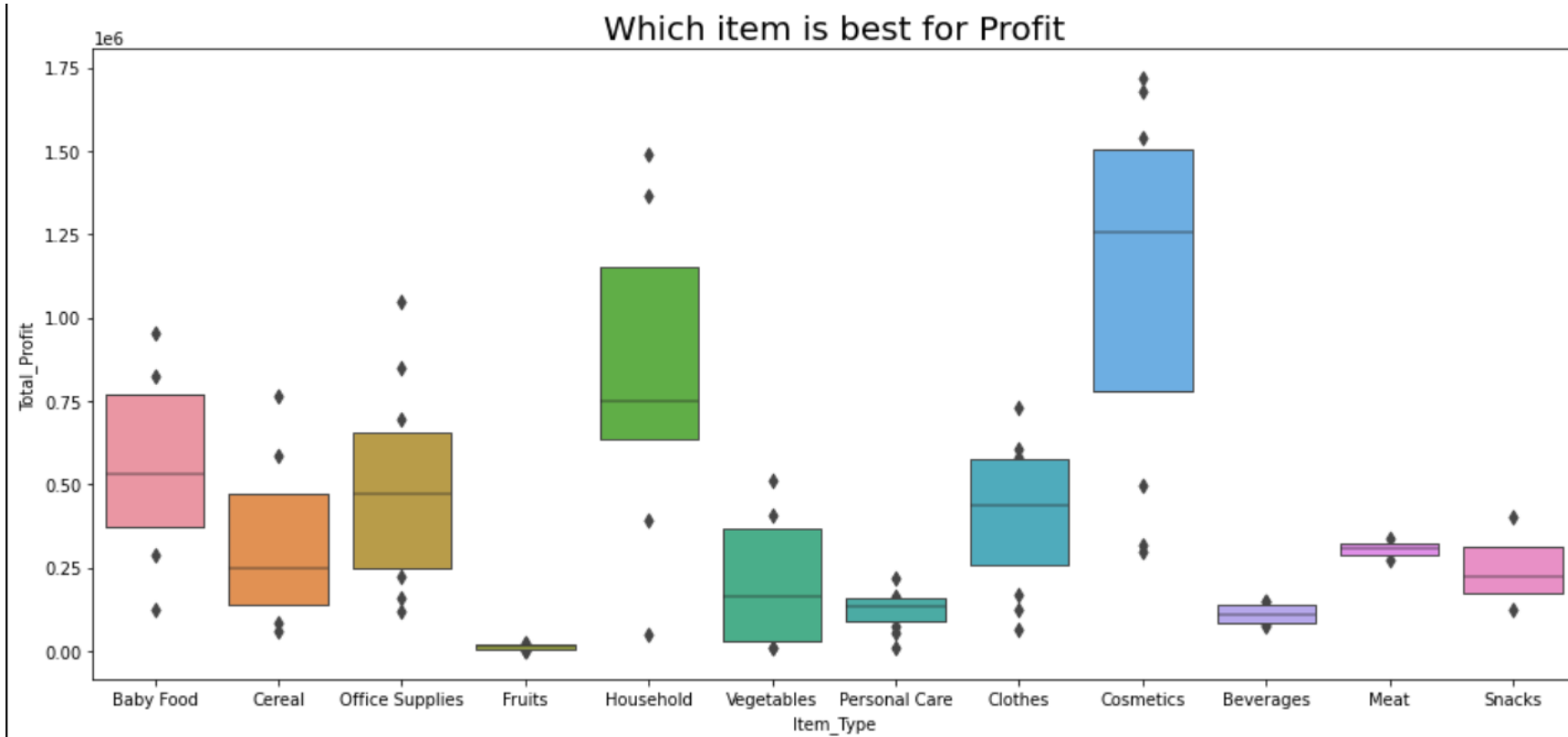
<https://drive.google.com/drive/folders/1FkmFVL8wIjMqWP1z52TD8PIhOJhitTyl?usp=sharing>

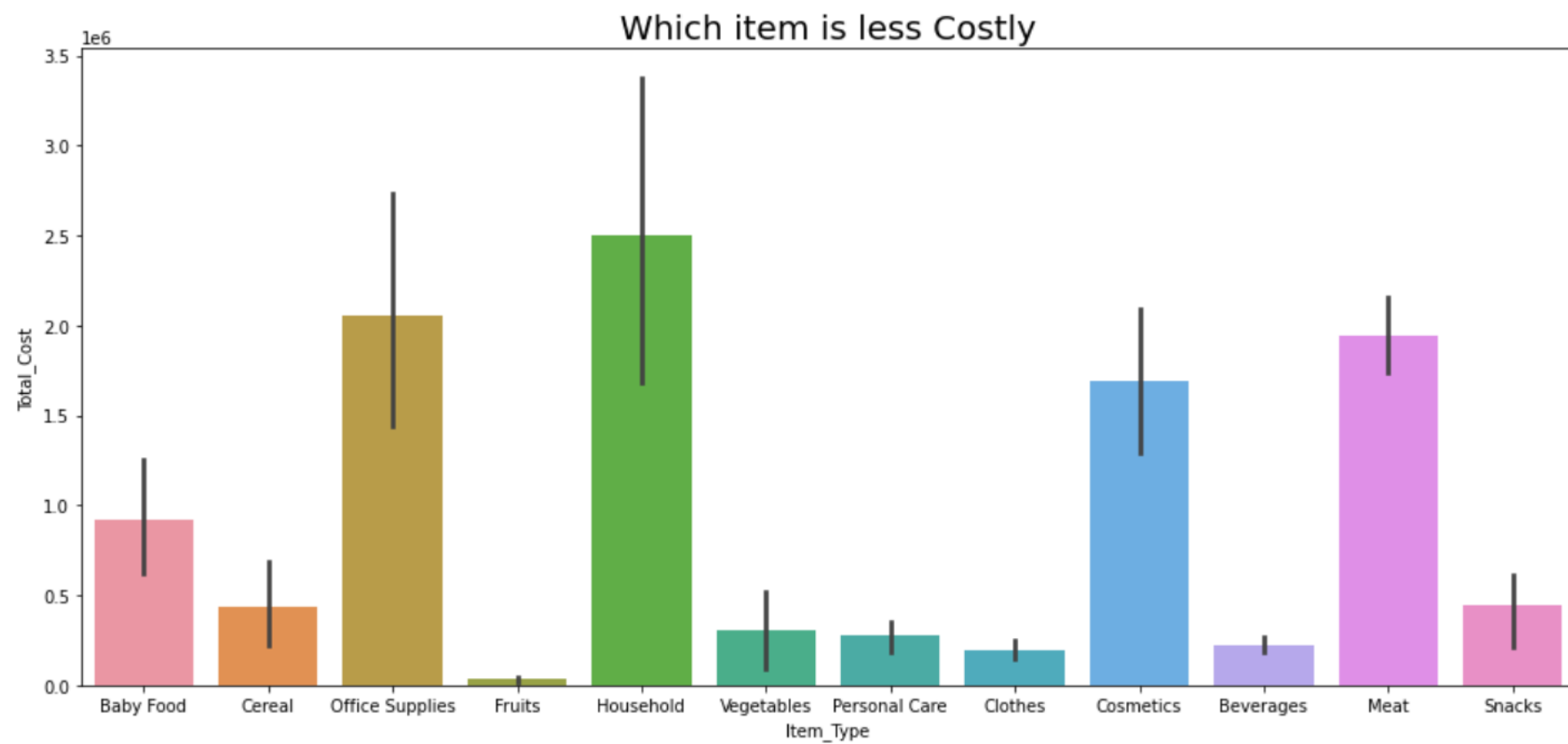
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1		Region	Country	Item_Type	Sales_Channel	Order_Priority	Order_Date	Order_ID	Ship_Date	Units_Sold	Unit_Price	Unit_Cost	Total_Revenue	Total_Cost	Total_Profit				
2	0	Australia &	Tuvalu	Baby Food	Offline	H	#####	6.69E+08	#####	9925	255.28	159.42	2533654	1582244	951410.5				
3	1	Central America	Grenada	Cereal	Online	C	#####	9.64E+08	#####	2804	205.7	117.11	576782.8	328376.4	248406.4				
4	2	Europe	Russia	Office Supplies	Offline	L	#####	3.41E+08	#####	1779	651.21	524.96	1158503	933903.8	224598.8				
5	3	Sub-Saharan Africa	Sao Tome & Principe	Fruits	Online	C	#####	5.14E+08	#####	8102	9.33	6.92	75591.66	56065.84	19525.82				
6	4	Sub-Saharan Africa	Rwanda	Office Supplies	Offline	L	#####	1.15E+08	#####	5062	651.21	524.96	3296425	2657348	639077.5				
7	5	Australia & Oceania	Solomon Islands	Baby Food	Online	C	#####	5.48E+08	#####	2974	255.28	159.42	759202.7	474115.1	285087.6				
8	6	Sub-Saharan Africa	Angola	Household Appliances	Offline	M	#####	1.35E+08	#####	4187	668.27	502.54	2798046	2104135	693911.5				
9	7	Sub-Saharan Africa	Burkina Faso	Vegetables	Online	H	#####	8.72E+08	#####	8082	154.06	90.93	1245113	734896.3	510216.7				
10	8	Sub-Saharan Africa	Republic of Congo	Personal Care	Offline	M	#####	7.7E+08	#####	6070	81.73	56.67	496101.1	343986.9	152114.2				
11	9	Sub-Saharan Africa	Senegal	Cereal	Online	H	#####	6.17E+08	#####	6593	205.7	117.11	1356180	772106.2	584073.9				
12	10	Asia	Kyrgyzstan	Vegetables	Online	H	#####	8.15E+08	#####	124	154.06	90.93	19103.44	11275.32	7828.12				
13	11	Sub-Saharan Africa	Cape Verde	Clothes	Offline	H	#####	9.4E+08	#####	4168	109.28	35.84	455479	149381.1	306097.9				
14	12	Asia	Bangladesh	Clothes	Online	L	#####	1.87E+08	#####	8263	109.28	35.84	902980.6	296145.9	606834.7				
15	13	Central America	Honduras	Household Appliances	Offline	H	#####	5.23E+08	#####	8974	668.27	502.54	5997055	4509794	1487261				
16	14	Asia	Mongolia	Personal Care	Offline	C	#####	8.32E+08	#####	4901	81.73	56.67	400558.7	277739.7	122819.1				
17	15	Europe	Bulgaria	Clothes	Online	M	#####	9.72E+08	#####	1673	109.28	35.84	182825.4	59960.32	122865.1				
18	16	Asia	Sri Lanka	Cosmetics	Offline	M	#####	4.19E+08	#####	6952	437.2	263.33	3039414	1830670	1208744				
19	17	Sub-Saharan Africa	Cameroon	Beverages	Offline	C	#####	5.2E+08	#####	5430	47.45	31.79	257653.5	172619.7	85033.8				
20	18	Asia	Turkmenistan	Household Appliances	Offline	L	#####	4.42E+08	#####	3830	668.27	502.54	2559474	1924728	634745.9				

Distribution of Item Type

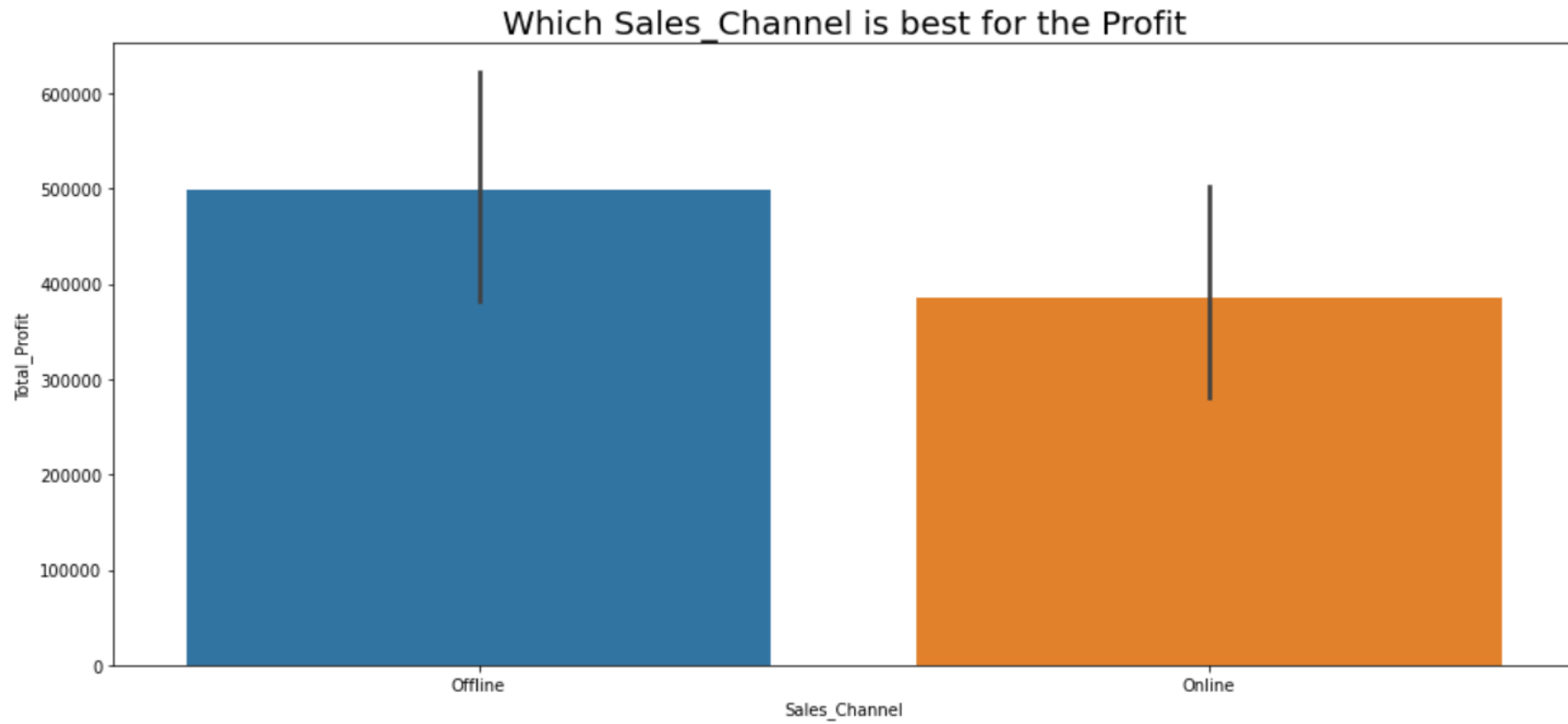


Which item is best for Profit

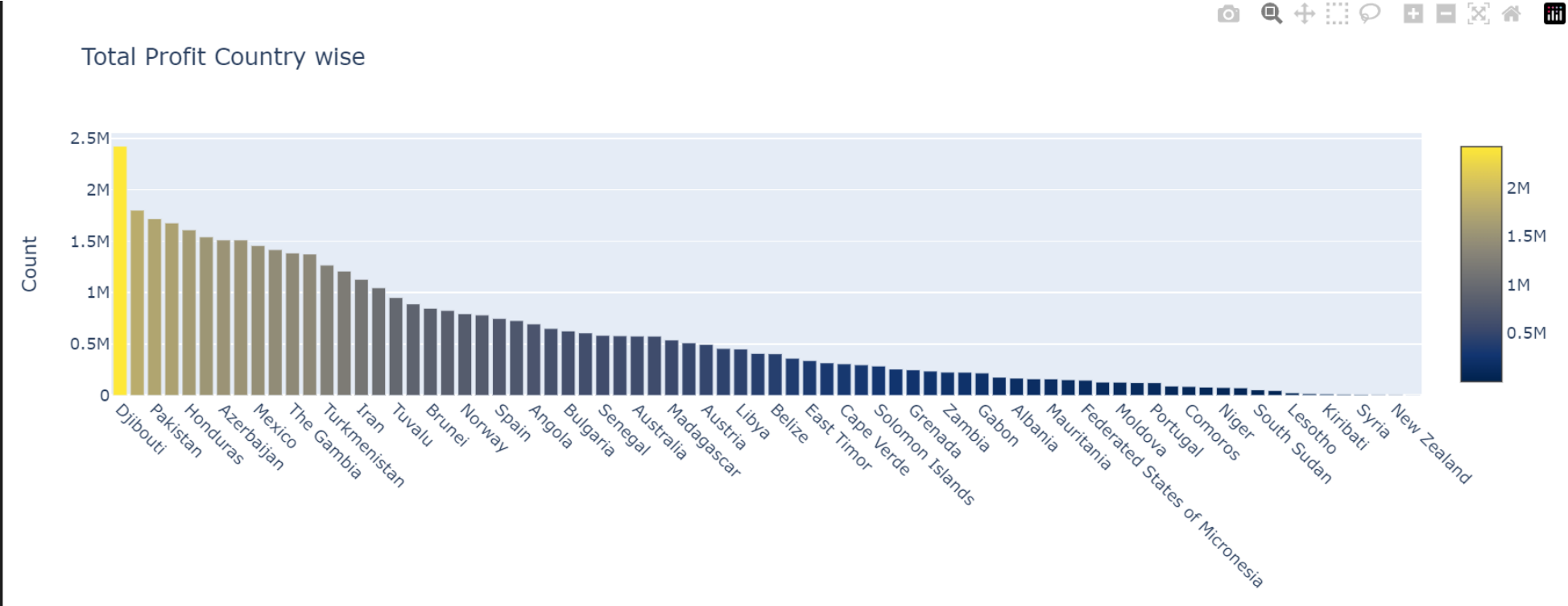




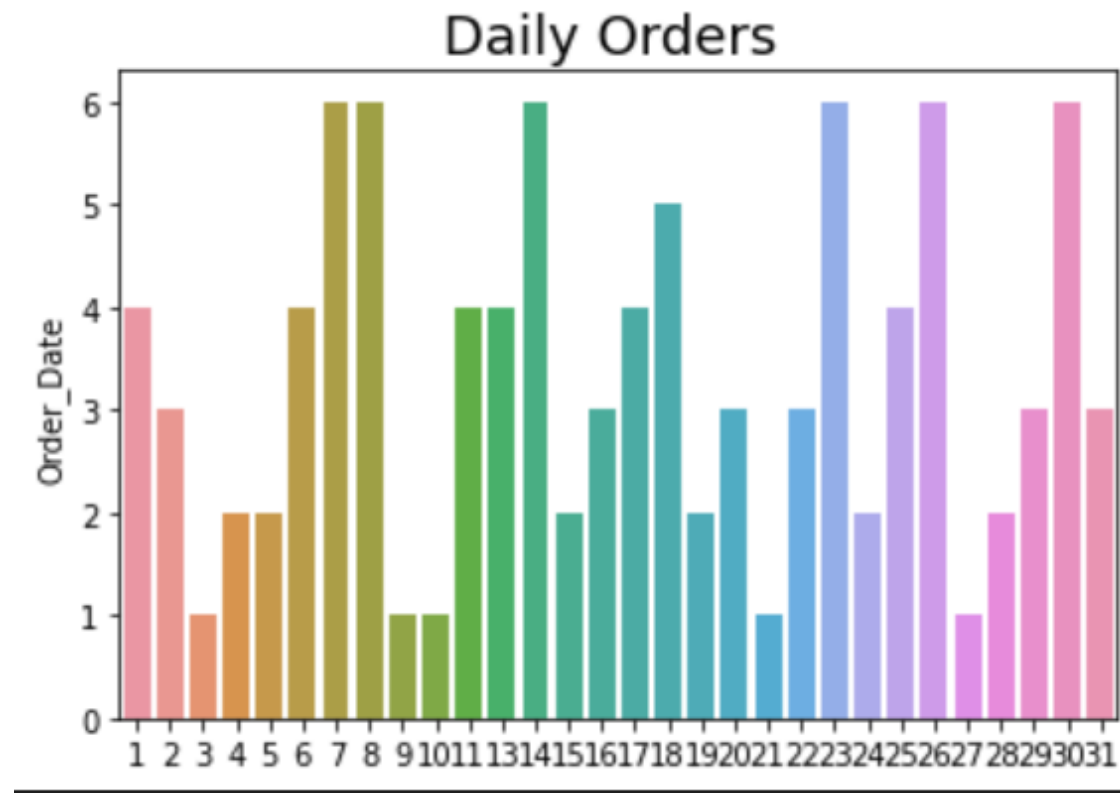
which Sales Channel is best for the Profit



Total Profit country wise



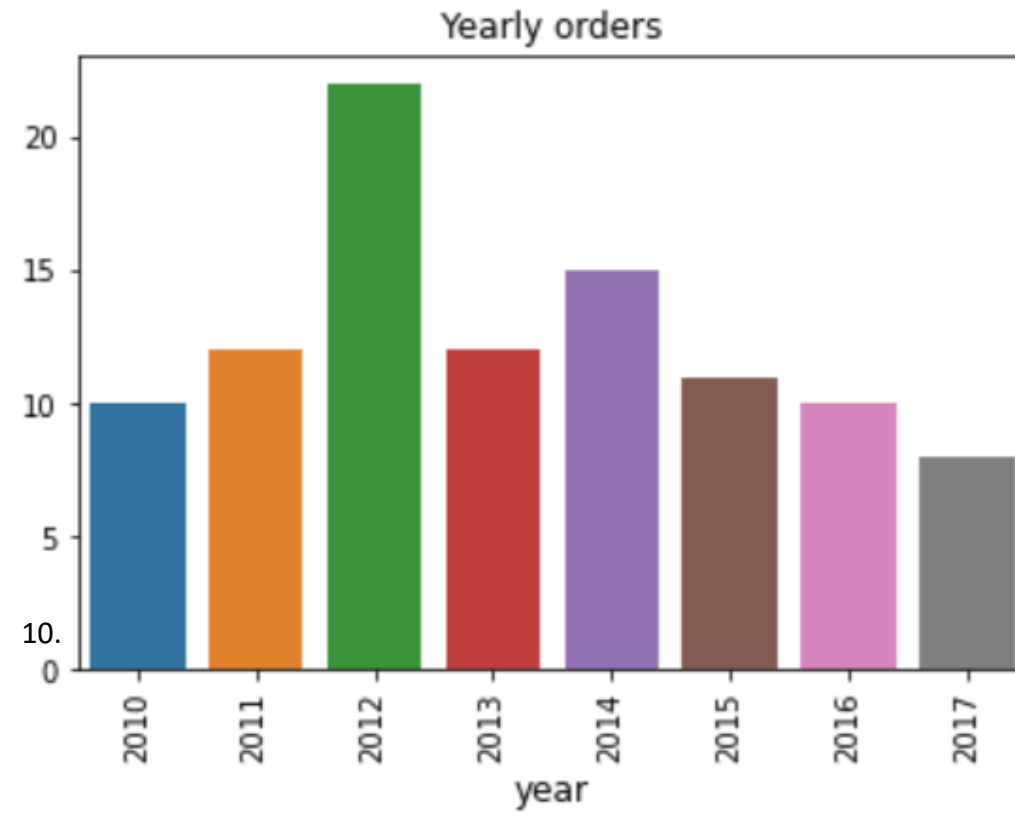
Daily orders



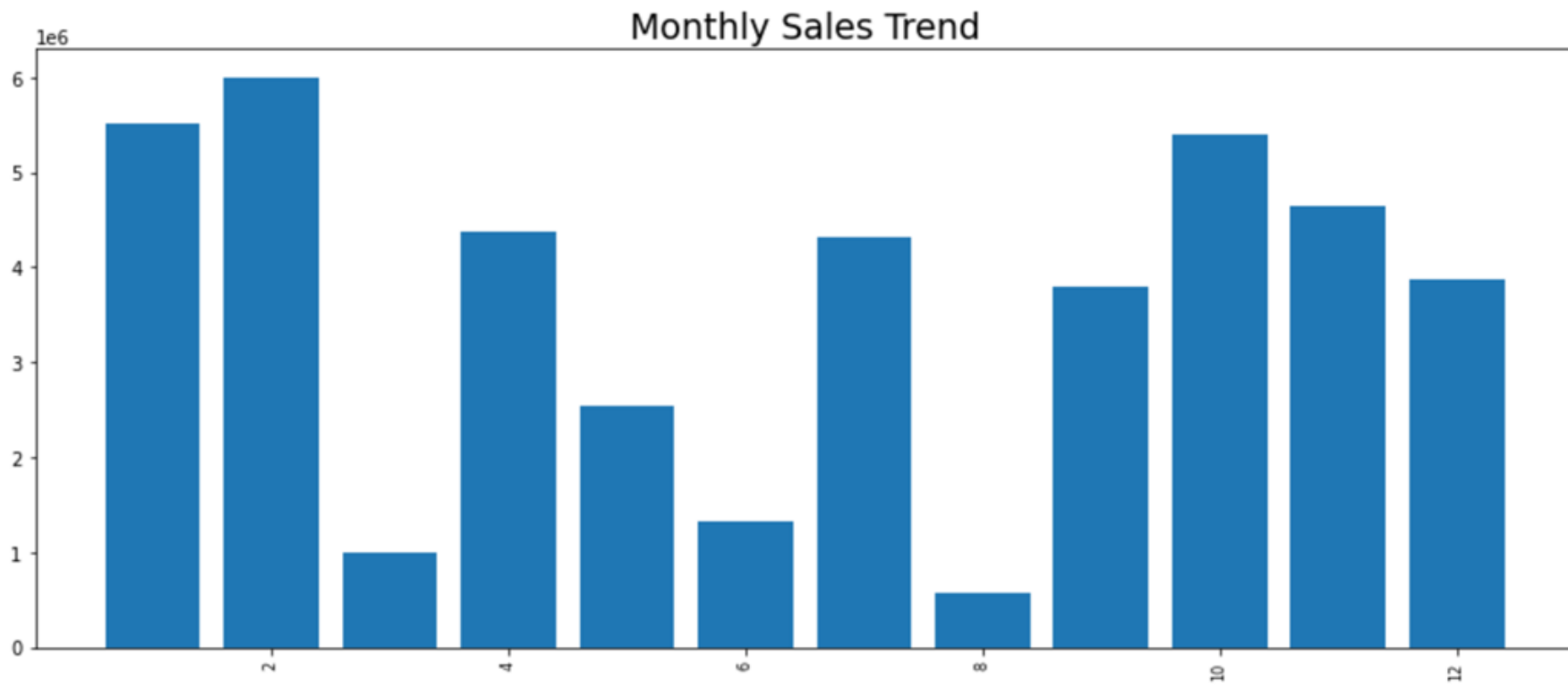
Monthly Orders



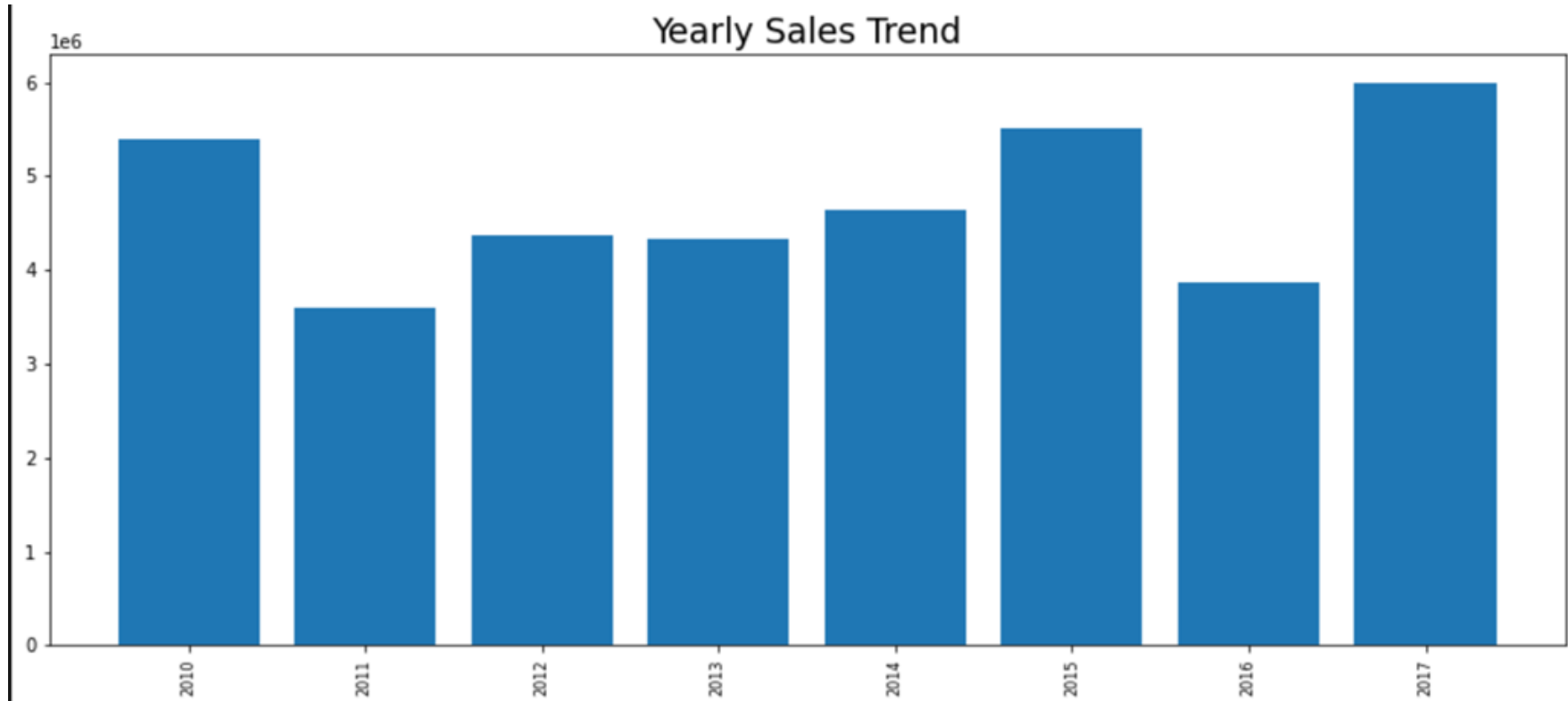
Yearly orders



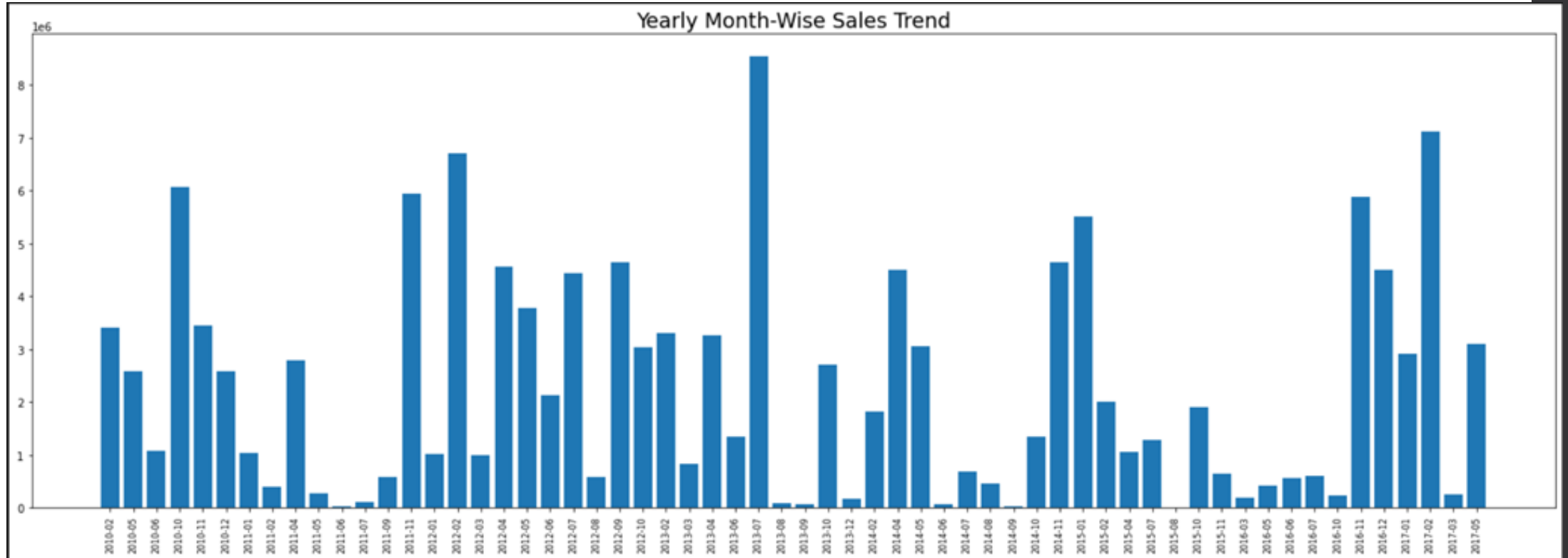
Monthly Sales Trends :-



Yearly Sales Trend



Yearly Month-Wise Sales Trend



Total Revenue

Total_Revenue	
Item_Type	
Cosmetics	36601509.60
Office Supplies	30585380.07
Household	29889712.29
Baby Food	10350327.60
Clothes	7787292.80
Cereal	5322898.90
Meat	4503675.75
Personal Care	3980904.84
Vegetables	3089057.06
Beverages	2690794.60

QUESTIONS AND ANSWERS

Q1) What's the source of data?

Ans) The Dataset is taken from iNeuron's provided dataset-

<https://drive.google.com/drive/folders/1FkmFVL8wlJmQWP1z52TD8PlhOJhitTyI?usp=sharing>

Q2) What was the type of data?

Ans) The data was a combination of numerical and Categorical values.

Q 3) What was the complete flow you followed in this Project?

Ans) Refer slide 5th for better Understanding

Q4) What techniques were you using for data?

Ans) -Removing unwanted attributes.

-Visualizing relation of independent variables with each other and output variables.

-Checking and changing distribution of continuous values.

-Removing outliers

-Cleaning data and imputing if null values are present.

-Transforming data to yield the desired result.

THANK YOU