

# AMAZON SALES DATA ANALYSIS



Presented By

Vikash Kumar Mahapatra

# OBJECTIVE



The goal of the project is to analyze the sale data and provide the sales and profit trends on the yearly as well as monthly basis for better use of sources for maximize the sales and profit.

# PROBLEM STATEMENT

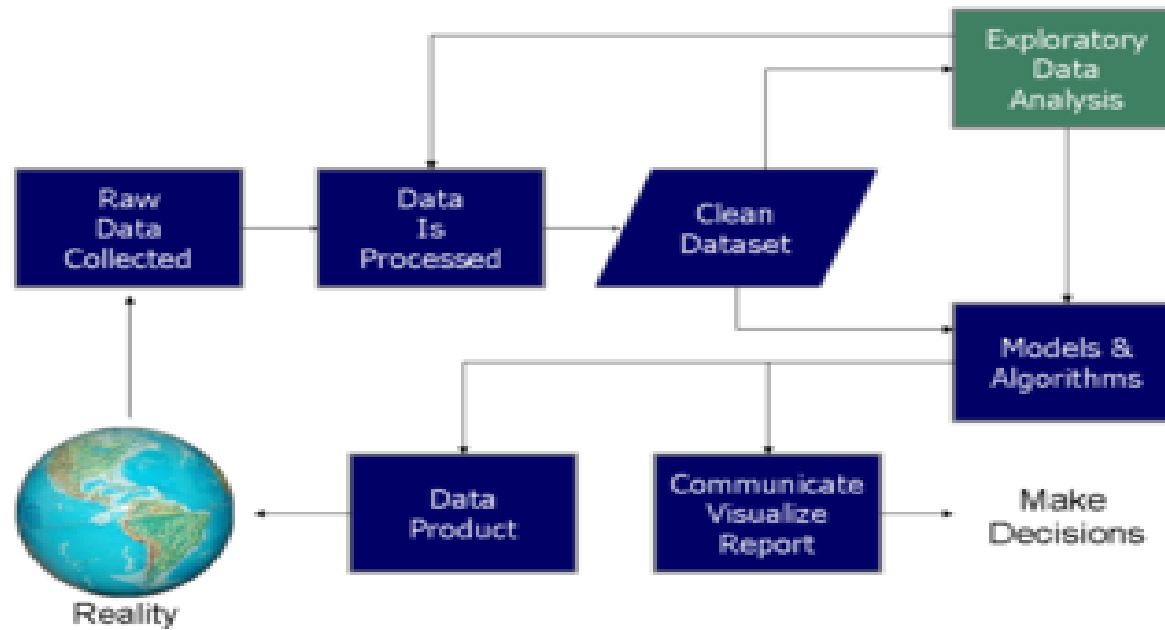


The problem statement aims to analyze amazon sales from 2017-2019 to know more in depth information about the items that are in high demand, high profit so that ways are find out to improve methods of distribution to reduce cost ,stay ahead in the competition, have progressive growth as well as further increase in the sales. As sale management today is the most important function in a commercial and business enterprise. To achieve the goal, we used a data set that is given and analyze most important parameter that are responsible for it. Also get some meaningful insights from the given information regarding Sales Data.

# ARCHITECTURE



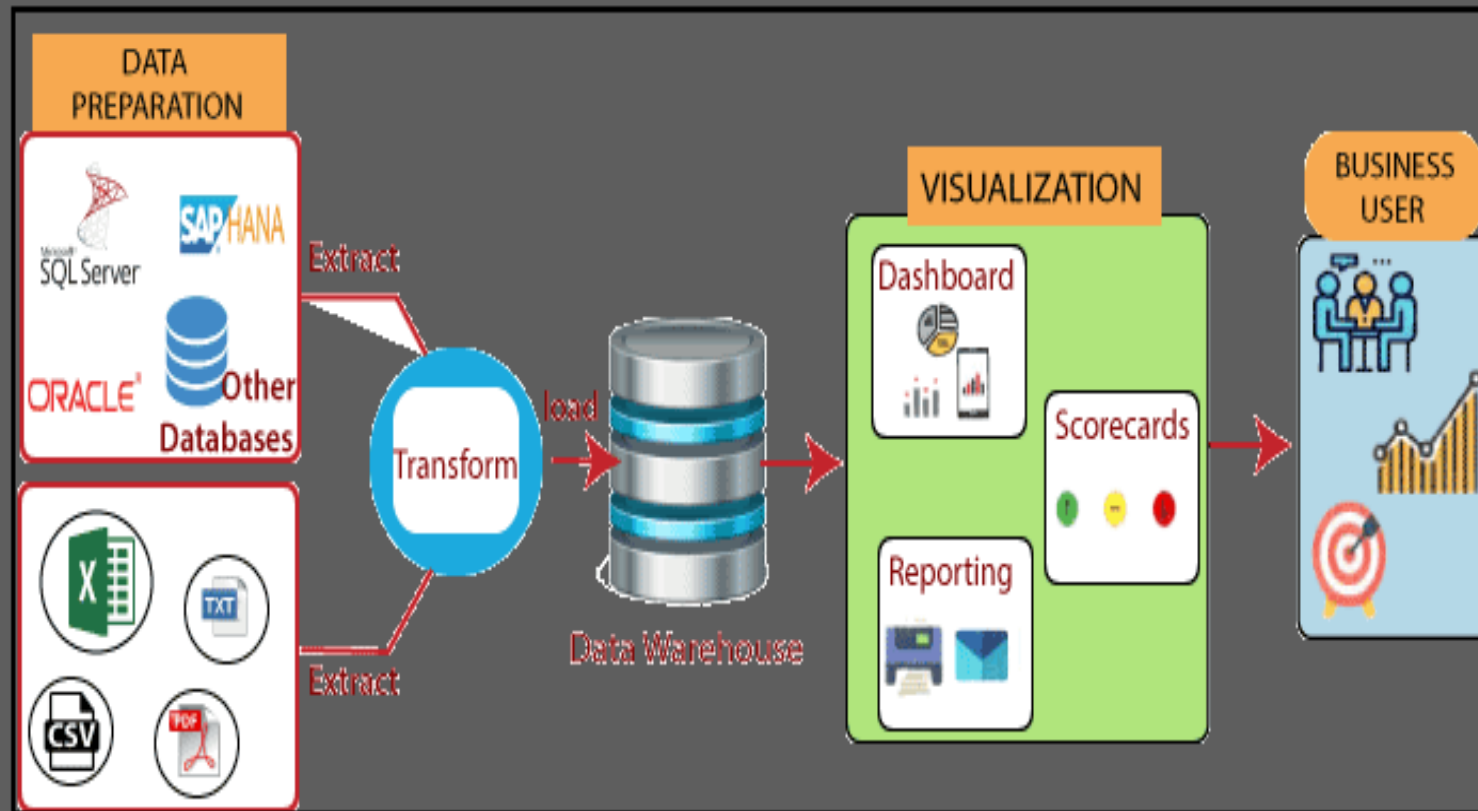
## Data Science Process



# POWER-BI ARCHITECTURE



## POWER BI ARCHITECTURE



# DATASET INFORMATION



This is a Sales related dataset that contains Information like Sales Amount, Cost Amount, Sales Price, List Price, Sales Margin Amount(Profit), Sale Quantity, Invoice Date, Order Date, Delivery Date ,Item Name, Customer Key etc. Sample Below

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	P	Q	R	S	T
1	CustKey	DateKey	Discount	Invoice Da	Invoice Numbr	Item Class	Item Nr	Item	Line Numbr	List Price	Order Number	Promised Deliv	Sales Amount	Sales Ai	Sales Margin Ai	Sales Price	Sales	Sales Rep	U/M
2	1000481	30-04-17	-237.91	2017-04-30	100012			Urban Large Eggs	2000	0	200015	30-04-17	237.91	0	237.91	237.91	1	184	EA
3	10002220	14-07-17	368.79	2017-07-14	100233	P01	20910	Moms Sliced Turkey	1000	824.96	200245	14-07-17	456.17	824.96	456.17	456.17	1	127	EA
4	10002220	17-10-17	109.73	2017-10-17	116165	P01	38076	Cutting Edge Foot-Long Hot Dogs	1000	548.66	213157	16-10-17	438.93	548.66	438.93	438.93	1	127	EA
5	10002489	03-06-17	-211.75	2017-06-03	100096			Kiwi Lox	1000	0	200107	03-06-17	211.75	0	211.75	211.75	1	160	EA
6	10004516	27-05-17	96627.9	2017-05-27	103341	P01	60776	High Top Sweet Onion	1000	408.52	203785	28-05-17	89248.66	185877	89248.66	196.1509	455	124	SE
7	10004516	30-05-17	-1950	2017-05-30	103610			Best Choice Fudge Brownies	2000	0	203785	30-05-17	1950	0	1950	1950	1	124	EA
8	10007866	03-09-17	371.014	2017-09-03	100403	P01	20910	Moms Sliced Turkey	2000	795.314	200436	03-09-17	424.3	795.314	424.3	424.3	1	149	EA
9	10009356	18-06-17	608.08	2017-06-18	105481	P01	62550	Tell Tale Garlic	29000	575	205213	18-06-17	541.92	1150	541.92	270.96	2	103	EA
10	10009356	18-06-17	424.8	2017-06-18	105481	P01	60794	High Top Walnuts	18000	51.88	205213	18-06-17	353.4	778.2	353.4	23.56	15	103	EA
11	10009356	18-06-17	13492.8	2017-06-18	105481	P01	36001	Big Time Frozen Cheese Pizza	9000	412.03	205213	18-06-17	11229	24721.8	11229	187.15	60	103	EA
12	10009356	18-06-17	10481.1	2017-06-18	105481	P01	38076	Cutting Edge Foot-Long Hot Dogs	13000	548.66	205213	18-06-17	8722	19203.1	8722	249.2	35	103	EA
13	10009356	18-06-17	404.147	2017-06-18	105481	P01	61484	Super Creamy Peanut Butter	37000	50.5051	205213	18-06-17	353.43	757.577	353.43	23.562	15	103	EA
14	10009606	16-09-17	1287.35	2017-09-16	100445	P01	17801	Better Fancy Canned Sardines	3000	1379.79	200478	16-09-17	1472.24	2759.59	1472.24	736.12	2	118	EA
15	10009606	16-09-17	4764.33	2017-09-16	100445	P01	48500	Red Spade Low Fat Cole Slaw	1000	1134.77	200478	16-09-17	5448.6	10212.9	5448.6	605.4	9	118	EA
16	10009633	10-06-17	-526.64	2017-06-10	104708			Jumbo Large Eggs	1000	0	204575	06-06-17	526.64	0	526.64	526.64	1	127	EA
17	10009633	20-08-17	0	2017-08-20	100379	P01	60449	High Top Corn on the Cob	1000	346.53	200397	18-08-17	346.53	346.53	346.53	346.53	1	127	EA
18	10009645	17-11-17	0	2017-11-17	100684			Moms Low Fat Bologna	2000	258.01	200733	14-11-17	3096.12	3096.12	3096.12	258.01	12	155	EA
19	10009652	14-05-17	91	2017-05-14	102256			Landslide Columbian Coffee	1000	5.61	202989	14-05-17	1872.5	1963.5	1872.5	5.35	350	145	EA
20	10009652	14-05-17	-700	2017-05-14	102256			High Top Dried Mushrooms	2000	0	202989	14-05-17	700	0	700	2	350	145	EA
21	10009907	06-11-17	786.6	2017-11-06	117815	P01	38050	Gorilla Chocolate Milk	2000	301.23	200706	06-11-17	418.32	1204.92	418.32	104.58	4	167	EA
22	10010884	19-06-17	499.44	2017-06-19	105615	P01	38007	Gorilla Jack Cheese	1000	1103	205323	19-06-17	603.56	1103	603.56	603.56	1	105	EA
23	10010884	08-07-17	409.77	2017-07-08	107052	P01	28401	Ebony Prepared Salad	10000	966.44	205041	13-06-17	556.67	966.44	556.67	556.67	1	105	EA
24	10010884	08-07-17	186.43	2017-07-08	107052	P01	26361	Bravo Canned Yams	7000	439.7	205041	13-06-17	253.27	439.7	253.27	253.27	1	105	EA
25	10010884	08-07-17	349.78	2017-07-08	107052	P01	20910	Moms Sliced Turkey	4000	824.96	205041	13-06-17	475.18	824.96	475.18	475.18	1	105	EA
26	10010884	08-07-17	325.53	2017-07-08	107052	P01	45880	Red Spade Low Fat Bologna	17000	767.75	205041	13-06-17	442.22	767.75	442.22	442.22	1	105	EA
27	10010884	08-07-17	313.27	2017-07-08	107052	P01	29394	Pearl Chardonnay	12000	123.14	205041	13-06-17	425.57	738.84	425.57	70.92833	6	105	EA
28	10010884	08-07-17	174.7	2017-07-08	107052	P01	36001	Big Time Frozen Cheese Pizza	13000	412.03	205041	13-06-17	237.33	412.03	237.33	237.33	1	105	EA

# INSIGHTS



Overall Sales

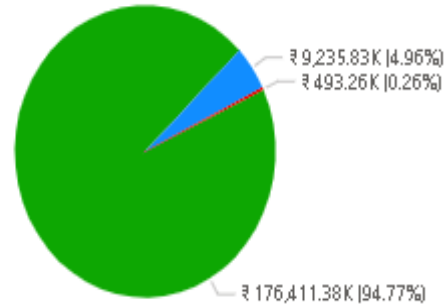
186.14M

Overall Profit

77.71M

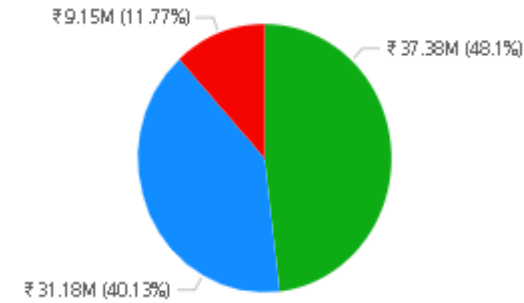
Sum of Sales Amount by U/M

U/M ● EA ● SE ● PR



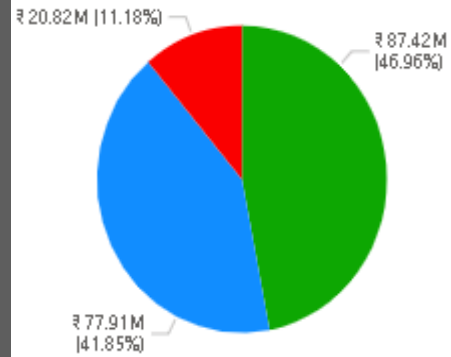
Sum of Sales Margin Amount by Billed\_Year

Billed\_Year ● 2017 ● 2019 ● 2018

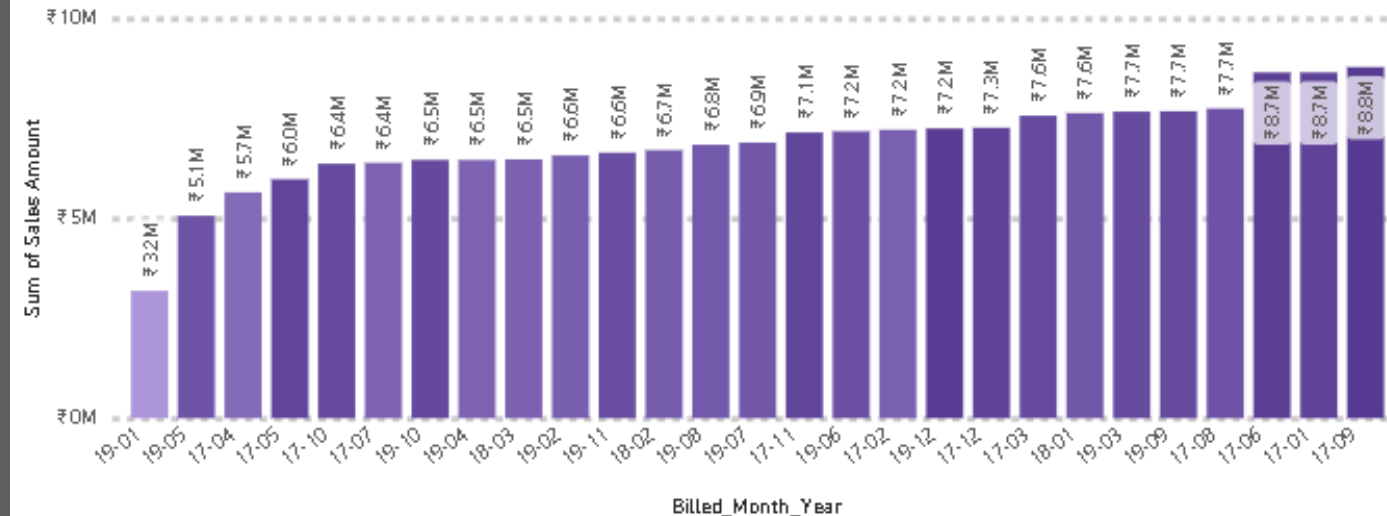


Sum of Sales Amount by Billed\_Year

Billed\_Year ● 2017 ● 2019 ● 2018



Sum of Sales Amount by Billed\_Month\_Year



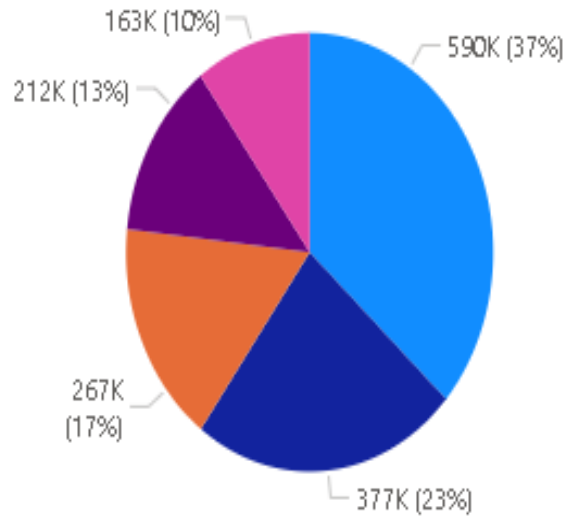
- ❑ From above figure which contains two number card one of them is total no of sales which is 186.14 Million and second one is total profit 77.71 Million between the year 2017 to 2019.
- ❑ In the above figure pie chart which is below the number card display the share of total sales amount between the year 2017 to 2019 i.e. in year 2017 is highest with INR 87.42 M(46.96%) second come year 2019 with INR 77.19 M(41.85%) least come the year 2018 with INR 20.82 M (11.18%).
- ❑ In the above figure two pie charts which are right side of number card . First pie chart share quantity sold on the basis of unit of measurement max was EA(Each) INR 176,411.4 K(94.88%), SE(Some item in Kilos or Gallons) INR 9,235.83 K(4.96%), PR(Pair) INR 493.26 K(0.26%). Second pie chart share profit between 2017 to 2019 i.e. highest profit in year 2017 INR 37.38 M(48.1%) second come the year 2019 INR 31.18 M(40.13%) last come the year 2018 INR 9.15 M(11.77%).
- ❑ In the above figure stack bar chart which display item sold on the month year basis. In the year 2019 month Jan sold least i.e. 32 M and March, Sept have sold highest i.e. 77 M . In the year 2018 March sold least i.e. 65 M and Jan sold highest i.e. 76 M. In the year 2017 April month sold least i.e. 57M and Sept sold highest 88 M.



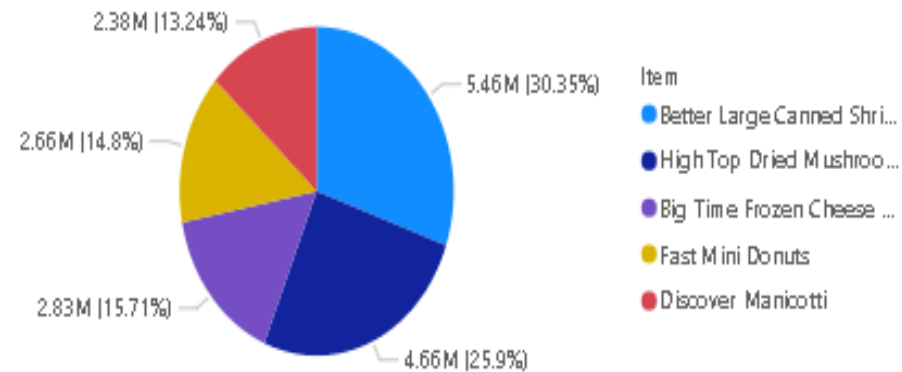
Top\_Five\_Purchases\_Items by Item

Item

- Better Large Cann...
- High Top Dried Mu...
- Better Canned Tun...
- Walrus Chardonnay
- Red Spade Piment...



Most\_Profitable\_Items by Item



□ From the top pie chart which display share of top five most purchased items over 3 years. Major share i.e. 37% Better Large Canned Shrimp then comes High Top Dried Mushrooms i.e. 23% and Lastly Red Spade Pimento Loaf i.e. 10%

□ From the bottom pie chart which display share of top five most profitable items over 3 years. Major share i.e. 30.35% Better Large Canned Shrimp then comes High Top Dried Mushrooms i.e. 25.9% and Lastly Discover Manicotti i.e. 13.24%

BBB Best Columbian Coffee	2017	1
Blue Label Fancy Canned Clams	2018	1
Fast Fudge Brownies	2019	1
Carlson Chocolate Milk	2019	1
Blue Medal Large Brown Eggs	2017	1
Carlson Low Fat Sour Cream	2018	1
Super Apple Jam	2017	1
Kiwi Lox	2018	1
Blue Label Vegetable Soup	2017	1
Landslide Apple Preserves	2017	1
Just Right Chicken Noodle Soup	2019	1
Carlson Chocolate Milk	2017	1
Choice Bubble Gum	2017	1
Just Right Chicken Noodle Soup	2017	1
BBB Best Hot Chocolate	2017	1
High Top Shitake Mushrooms	2017	1
Fantastic Blueberry Muffins	2017	1
BBB Best Corn Oil	2017	1
Club Mild Cheddar Cheese	2017	1
Landslide Salt	2019	-1018.41
	2017	-1184.57
Blue Label Rice Soup	2019	-1278.00
Carlson Blueberry Yogurt	2017	-1361.19
Just Right Chicken Soup	2018	-1381.00
Landslide Low Fat Apple Butter	2017	-1507.97
Good Chablis Wine	2017	-1610.10
Landslide Low Fat Apple Butter	2019	-2023.17
Good Chablis Wine	2019	-2150.52
Carlson Blueberry Yogurt	2019	-2248.84
Best Choice Fondue Mix	2018	-2488.89
Fast Lemon Cookies	2018	-3085.25
	2017	-7587.60
Just Right Chicken Soup	2019	-8235.16
	2017	-9917.30
Best Choice Fondue Mix	2019	-10020.34
	2017	-20585.56
Fast Lemon Cookies	2019	-22435.78

❑ The top image display least purchased items over 3 years.

❑ The bottom image display least profitable items over 3 years. In that Fast Lemon Cookies are the most least profitable then follows the Best Choice Fondue Mix then Just Right Chicken Soup.

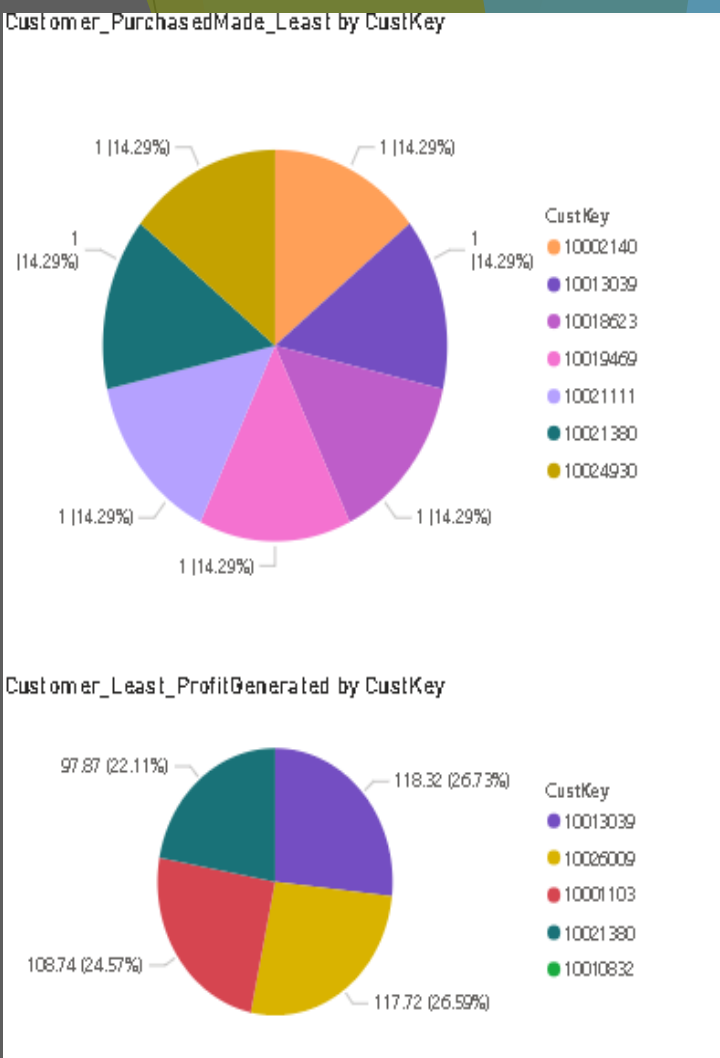
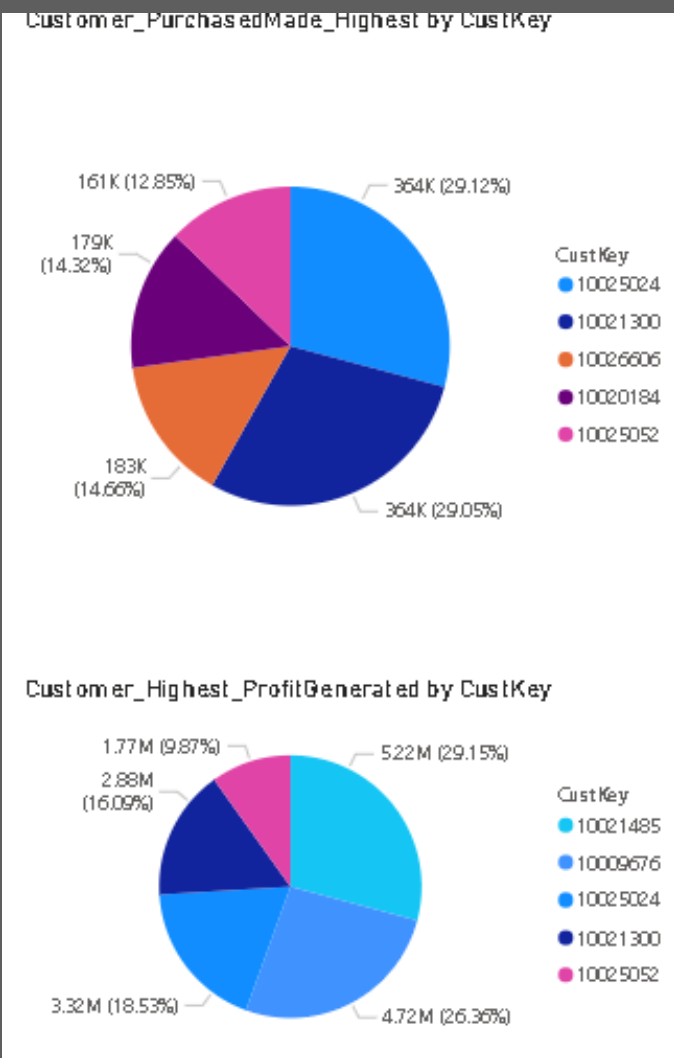
Item	Billed_Month_Year	
Better Large Canned Shrimp	19-03	38546
	19-09	38458
	19-08	32617
	17-08	32613
	17-02	32613
	19-02	32529
High Top Dried Mushrooms	18-01	31340
	17-06	30140
	19-04	30087
Better Large Canned Shrimp	18-02	30056
	17-09	29964
	17-03	29964
High Top Dried Mushrooms	18-01	29868
Better Large Canned Shrimp	18-03	25212
	19-05	24208
	17-01	24036
	17-07	24031
Landslide Hot Chocolate	19-06	22684

Sales Margin Amount		
Item	Billed_Month_Year	
Fast Mini Donuts	19-07	410520.45
High Top Dried Mushrooms	18-01	369730.26
Fast Mini Donuts	17-12	328386.96
Better Large Canned Shrimp	19-03	324178.73
	19-09	323438.89
	18-01	322267.57
High Top Dried Mushrooms	19-04	312223.88
Fast Mini Donuts	17-01	310527.12
Better Large Canned Shrimp	17-08	309201.71
	17-02	309201.71
	18-02	308579.99
	17-06	301970.74

- ❑ The top image display most purchased items on the basis of year month. Better large Canned shrimp max purchase in the year 2019 (Mar, Sept&Aug), 2017 (Aug&Feb), 2018 (Jan) then comes the High Top Dried Mushrooms then lastly Landslide Hot Chocolate 2019 June.
- ❑ The bottom image display most profitable items on the basis of year month. Fast Mini Donuts in July 2019, Dec 2017 then High Top Dried Mushrooms in Jan 2018 then lastly Better Large Canned Shrimp in Sept 2019, Jan 2018.

Cutting Edge Pimento Loaf	19-12	1
Fabulous Berry Juice	17-06	1
	19-09	1
	19-03	1
Landslide Vegetable Oil	18-03	1
	17-09	1
	17-03	1
Best Corn Puffs	19-05	1
	17-05	1
Landslide Sesame Oil	19-01	1
	17-12	1
Excellent Strawberry Drink	17-11	1
Discover Spaghetti	19-01	1

Fast Lemon Cookies	19-03	-2321.54
	19-09	-2321.54
	17-06	-2450.44
	17-05	-2450.44
	19-06	-2476.20
Best Choice Fondue Mix	18-02	-2488.89
	17-03	-2508.21
	17-09	-2508.21
	19-03	-2533.97
	19-09	-2533.97
	17-12	-2879.53
Just Right Chicken Soup	19-12	-3036.22
	19-10	-3036.22
	17-09	-3378.71
	19-11	-3932.93
	17-12	-4335.59
Fast Lemon Cookies	19-10	-4724.99
	19-12	-4724.99
	17-10	-7944.40



- ❑ The left most image display least purchased items on the basis of year month. Some items are Landslide Vegetable oil, Best Corn Puff ,Fabulous berry juice and Discover Spaghetti.
- ❑ The second left image display least profitable items on the basis of year month. Fast Lemon Cookies in year 2019(Oct&Dec) ,2017 Oct then Just Right Chicken Soup in year 2019(Oct),2017 Sept lastly Best Choice Fondue Mix 2018 Feb , 2017 (Mar&Sept).
- ❑ The second right top image with pie chart display customer brought max items in 3 years . First comes 10025024(29.12%), second comes 10021300(29.05%) and last comes 10025052(12.85%) .
- ❑ The second right bottom image with pie chart displays customer provide max profit in 3 years . First comes 10021485(29.15%),second comes 10009676(26.36%) and last comes 10025052(9.87%) .
- ❑ The right most top image with pie chart display customer who brought least items in 3 years .10002104,10013039,10018623,10019469,10021111,10021380 and 10024390 each of them have 14.29% .
- ❑ The right most bottom image with pie chart display customer who provide least profit in 3 years . First comes 10013039(26.73%),second comes 10026009(26.59%) and last 10010832(22.11%)

# Q & A

- ❖ What's the source of data? The Dataset was taken from project document <https://drive.google.com/drive/folders/1FkmFVL8w1JmQWP1z52TD8Plh0JhiTyl?usp=sharing>  
What was the type of data? It was combination of numerical and categorical values.
- ❖ What were the libraries that you used in Python? I used Pandas , NumPy and Matplotlib libraries of python.
- ❖ Do we have to download these libraries? To check whether you have it or not try pip list (you need to have pip install in system execute this command)if don't find then download using pip install library name.
- ❖ What techniques were you using for data? 1.Finding null values and datatypes. 2.Visualizing relationship between independent variables with each other and output variables. 3.Removing unwanted attributes, removing outliers, impute missing if required otherwise drop from dataset. 4. Transform your data the get desired result.

*THANK YOU*