

# Walmart Sales Analysis

## Presentation

Identifying the patterns in our products, regions, categories and customer segments for efficiency and profit optimization with SQL

With technology evolving rapidly, establishing robust databases and dedicated departments for interpreting them is essential for staying ahead in efficiency. While execution is crucial, having high-quality inputs ensures even better outputs. As market demand and competition intensify, it's imperative to leverage insights to deliver enhanced value to customers, as value remains paramount.

Our analysis will utilize Excel, SQL, and Tableau. Excel will serve as the initial data repository, SQL will extract meaningful insights, and Tableau will visualize the data effectively. Following the six steps of Data Analysis - Ask, Prepare, Process, Analyze, Share, and Act.

### Step 1: Ask

In this stage, we will articulate the business challenge presented to us, which has been interpreted as "Identifying the optimal products, regions, categories, and customer segments for the Walmart to prioritize or steer clear of, aiming to amplify profitability."

#### Business objectives:

- How can we optimize our profits?
- What are the emerging trends that we can identify?
- How can we take these insights to build recommendations ?

#### Deliverables:

- Concise overview of the business goals.
- Comprehensive documentation detailing data cleansing, manipulation, and analysis.
- Dashboard showcasing visualizations and key findings.
- Recommendations derived from our insights and analysis.

### Step 1: Prepare

In this phase, we will identify and assess the features of our Walmart Dataset:

- It comes with 8400 rows with 8400 being pure data and the other one row being the column headers. It contains data recorded between the 1st of January 2012 (the first order date) to the 30th of December 2015 (the last shipping date). (The last order date is the 30th of December 2015, so we will instead use the order dates range to represent our 4 years of business)
- It contains the data of 795 customers.

- Query The data contains the 23 columns namely; City , Customer Name , Customer Segment , Discount , Number of Records , Order Date , Order ID , Order Priority , Order Quantity , Product Base Margin , Product Category , Product Container , Product Sub Category , Profit , Region , Row ID , Sales , Ship Date , Ship Mode , Shipping Cost , State , Unit Price , Zip Code .

- Moving on to the data processing, we will use Excel , Power for cleaning .

### Step 3: Process

We will process and refine our data utilizing Excel since the file is already in CSV format, making it convenient for us to conduct an initial examination and cleansing of our dataset.

- Reviewing our data
- Identifying missing data using conditional formatting
- Eliminating duplicate rows
- Ensuring proper formatting of columns for streamlined SQL analysis

While examining our dataset, we may observe and take note of the following:

- Our data appears accurate and cohesive. Everything looks well structured for further analysis it just needs a little editing.
- With the command ‘remove duplicates’, there was no instance where the data was duplicated with all the exact parameters for a customer in all columns. So all the rows had some variety to it hence returning no duplicate data for our data set.

Now our dataset is ideal for analysis to discover relationships, trends and patterns that will give us a competitive edge

### Step 4: Analyze

In the analysis phase, we will extract the key components of our data to address our business objectives effectively.

let's load the data into SQL and this query will retrieve the first 5 rows from the Walmart table , allowing us to inspect them and ensure they were imported accurately.

Select \*

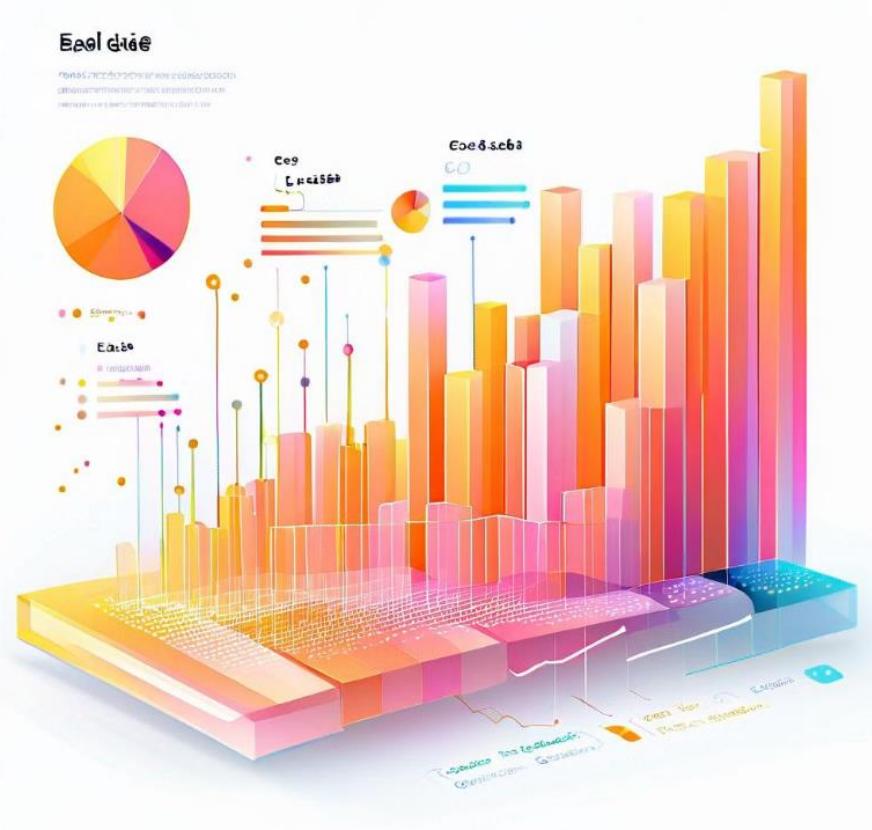
From Walmart

Limit 5;

let's conduct an exploratory data analysis (EDA) on the Walmart dataset. We'll answer a series of tasks, provide the SQL queries for each task, and then present the results. Once we've completed the analysis, we can transition to creating a dashboard summarizing the key insights.

City	Customer_Name	Customer_Segment	Discount	Number of Records	Order Date	Order ID
McKeesport	Jessica Myrick	Small Business	0.1	1	2012-01-01	28774
Bowie	Matt Collister	Home Office	0.08	1	2012-01-01	13729
Napa	Alan Schoenberger	Corporate	0	1	2012-01-02	37537
Montebello	Elizabeth Moffitt	Consumer	0.08	1	2012-01-02	44069
Napa	Alan Schoenberger	Corporate	0.07	1	2012-01-02	37537

The first 5 rows and 23 columns of our dataset



## 1. What are total sales and total profits of each year?

The years were grouped by order date, so we can observe data for the year 2012, 2013, 2014 and 2015.

```
select date_format(order_date,"%Y") as Year,
round(sum(profit),2) as "Total Profit",
round(sum(sales),2) as "Total Sales"
from walmart_sales
group by Year
order by Year asc;
```

Year	Total Profit	Total Sales
2012	434536.3	4209009.11
2013	363587.55	3548319.75
2014	381792.68	3438246.46
2015	341851.44	3720025.79

Total sales and Total profits for each year

## 2. What are the total profits and total sales per quarter?

```
select date_format(order_date,"%Y") as Year,
```

```
case
```

```
when date_format(order_date,"%c") in (1,2,3) then "Q1"
```

```
when date_format(order_date,"%c") in (4,5,6) then "Q2"
```

```
when date_format(order_date,"%c") in (7,8,9) then "Q3"
```

```
else "Q4"
```

```
end as quarter,
```

```
round(sum(profit),2) as "Total Profit",
```

```
round(sum(sales),2) as "Total Sales"
```

```
from walmart_sales group by Year,quarter
```

```
order by Year,quarter asc;
```

Now this table will aid us in knowing what quarters were the most profitable to us from 2012–2015. This can help to pave the way for investment and marketing strategies. Computing the table above through Excel gives us the following :

Year	quarter	Total Profit	Total Sales
2012	Q1	115785.03	1260358.23
2012	Q2	91538.23	886587.58
2012	Q3	102289.11	1033034.65
2012	Q4	124923.93	1029028.65
2013	Q1	75340.03	823069.38
2013	Q2	95763.82	845509.61
2013	Q3	88111.5	857070.81
2013	Q4	104372.2	1022669.95
2014	Q1	82112.95	848800.87
2014	Q2	79134.47	748531.48
2014	Q3	94322.54	837868.84
2014	Q4	126222.72	1003045.27
2015	Q1	106169.18	964845.07
2015	Q2	91162.93	929299.21
2015	Q3	74518.71	862967.53
2015	Q4	70000.62	962913.98

The total sales and total profits for each year per quarter

```
select case
```

```
when date_format(order_date,"%c") in (1,2,3) then "Q1"
```

```
when date_format(order_date,"%c") in (4,5,6) then "Q2"
```

```
when date_format(order_date,"%c") in (7,8,9) then "Q3"
```

```
else "Q4"
```

```
end as Quarter,
```

```
round(sum(profit),0) as "Total Profit",
```

```
round(sum(sales),0) as "Total Sales"
```

```
from walmart_sales group by quarter
```

```
order by quarter;
```

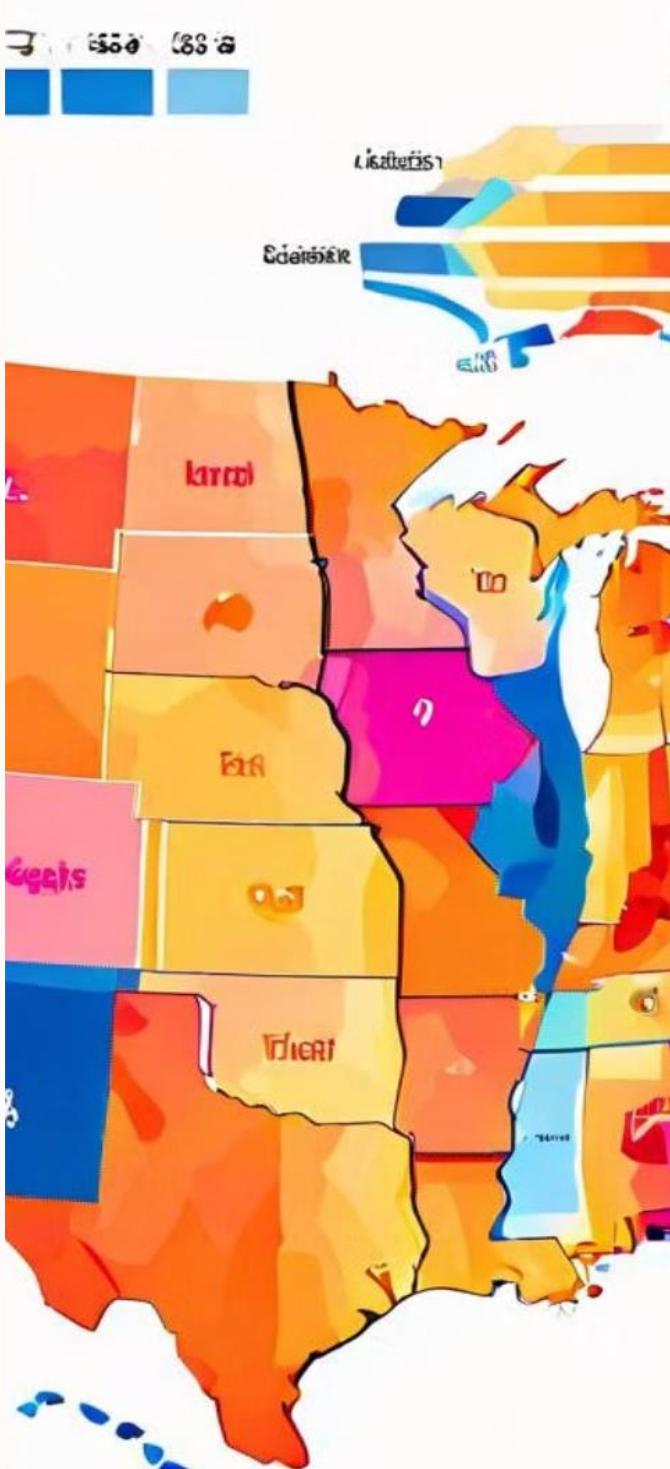
Quarter	Total Profit	Total Sales
Q1	379407	3897074
Q2	357599	3409928
Q3	359242	3590942
Q4	425519	4017658

Most performing quarters from 2012–2015

The data above shows that the period of October, November and December consistently emerge as our peak selling months , generating the highest profits. Analyzing this table provides valuable insights for crafting operational strategies, akin to observing a stock market rally from January to December, followed by a downturn in the initial three months. Let's get into the regions.

### 3. What region generates the highest sales and profits ?

#### Troll Fellon



```
select Region,round(sum(sales),2) as "Total Sales",
round(sum(profit),2) as Total_Profit
from walmart_sales
group by region order by Total_Profit desc;
```

Region	Total Sales	Total_Profit
Central	4699167.35	481891.2
South	3150219.34	422507.13
East	3416466.56	317852.11
West	3649747.86	299517.53

Total Profits and Sales by Region

We can observe above that the Central region is the one with the most sales and brings us in the highest profits. The south region is pretty good looking for our company too because less sales but more profit generate. Those 2 regions are definitely areas of interest if we want to maximize our profits and expand our business. Concerning the East region, we do not gain a lot of revenue but still the profits are there. It is the West region that is quite alarming as we generate way more revenue than the East region but do not make at least the same profits over there. The West region should be on our watchlist as we could start to think on how we could maybe put our resources in the other regions instead. Let's observe each regions profit margins for further analysis with the following code:

```
Select Region,
round((sum(profit)/sum(sales))*100,2)
as Profit_margin
from Walmart_sales
group by Region
order by Profit_margin desc;
```

### Profit Margin Formula

$$\text{Margin Ratio} = \frac{\text{Gross Profit}}{\text{Net Sales}} \times 100$$

Region	Profit_margin
South	13.41
Central	10.25
East	9.3
West	8.21

Profit margins by region

Profit margins are a measure of a company's profitability and are expressed as the percentage of revenue that the company keeps as profit. So we can see that the South and Central are really good in case of sales central region is more selling but less profit\_margin of 10.25% in comparison to South region profit\_margin of 13.41%. The East and west region almost same in terms of profit\_margin. Let's move on and try to pinpoint the data in each region.

## 4. What state and city brings in the highest sales and profits ?

### States

Firstly, Let's discover what states are the top 10 highest and lowest and then we will move on to the cities. For the states, it can be found with the following code:

```
select State,round(sum(sales),2) as "Total Sales",
round(sum(profit),2) as Total_Profit,
round((sum(profit)/sum(sales))*100,2)
as Profit_margin
from walmart_sales
group by State
order by Total_Profit desc limit 10;
```

This produces the following result:-



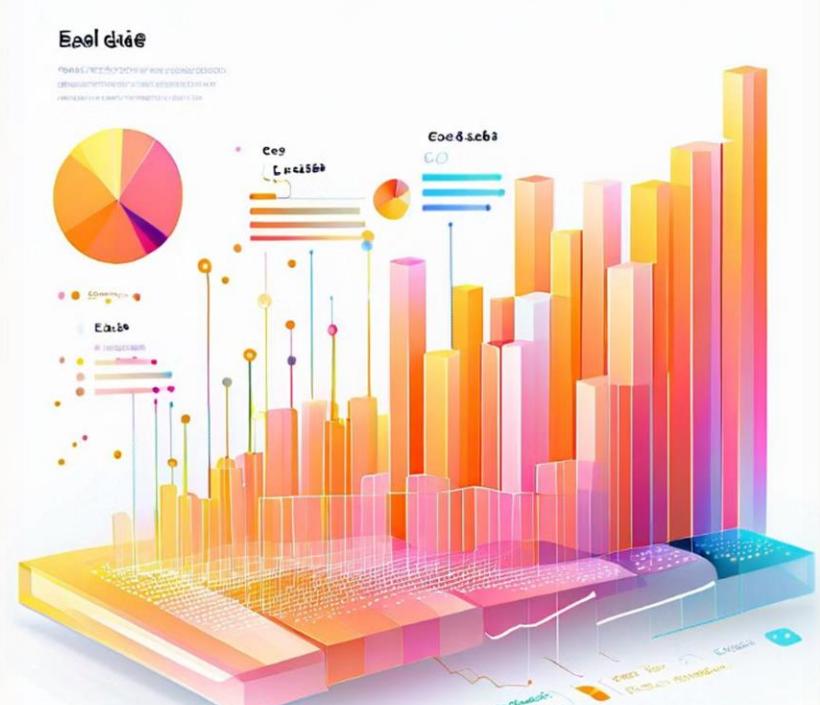
State	Total Sales	Total_Profit	Profit_margin
Illinois	959327.3	108532.04	11.31
Texas	863891.05	93571.53	10.83
Ohio	729426.19	88615.5	12.15
California	1372210.23	87355.51	6.37
New York	738894.21	86396.01	11.69
Florida	777663.94	82571.93	10.62
Virginia	379200.69	61721.94	16.28
Alabama	305010.82	53630	17.58
Oregon	354325.19	45274.03	12.78
Michigan	475170.83	45210.64	9.51

Top 10 State's total sales and profits with their profit margins

The decision was made to incorporate profit margins to gain a fresh perspective. The data reveals the top 10 most lucrative states, showcasing both total sales and profit margins. Profit margins play a pivotal role, enabling us to adopt a long-term mindset as investors, identifying potential high-yield markets. Illinois, Texas, Ohio, and California emerge as our most profitable markets, with significant sales figures. Their substantial sales volumes indicate that achieving higher profit margins would require considerable effort. Nonetheless, these states yield impressive profits, indicating our dominant market presence. Consequently, we should prioritize bolstering our resources and enhancing customer service in these key states.

Let's observe our bottom 10 States:

```
select State,round(sum(sales),2) as "Total Sales",
round(sum(profit),2) as Total_Profit
from walmart_sales
group by State
order by Total_Profit asc limit 10;
```





State	Total Sales	Total_Profit
Montana	76071.98	-9127.2
Delaware	10925.66	585.35
West Virginia	77018.36	1227.62
Nevada	52063.37	3976.83
Vermont	99602.64	6378.93
Kansas	257860.98	6628.44
Rhode Island	50820.48	6793.05
Connecticut	102924.07	6885.14
New Hampshire	95608.53	6979.68
Mississippi	97623.54	7191.02

Bottom 10 State's total sales and profits

The least profitable markets, with Montana, Delaware, and West Virginia topping the list.

In Montana, although sales reach \$76,071.98, unfortunately, our profit stands at a loss of -\$9,127.2. Delaware, with sales of \$10,925.66, fares slightly better, yielding a profit of \$585.35. Similarly, West Virginia's sales reach \$77,018.36, generating a profit of \$1,227.62.

## Cities

The top cities are found with the code below:

```
select City,round(sum(sales),2) as "Total Sales",
round(sum(profit),2) as Total_Profit
,round((sum(profit)/sum(sales))*100,2) as Profit_margin
from walmart_sales
group by City
order by Total_Profit desc limit 10;
```

City	Total Sales	Total_Profit	Profit_margin
Mount Vernon	108225.02	28036.83	25.91
Kokomo	51988.1	12944.31	24.9
Kearney	34552.03	14543.94	42.09
Hicksville	38893.79	13441.07	34.56
Henderson	55510.08	16438.02	29.61
Farmington Hills	29355.29	13725.52	46.76
Easton	65209.36	20764.31	31.84
Dover	55507.69	13388.94	24.12
Columbus	50713.66	13146.56	25.92
Baton Rouge	61726.2	12961.95	21

Top 10 Cities' total sales and profits with their profit margins

The top 3 cities that we should focus on are Mount Vernon, Kokomo and Kearney.

The bottom 10 cities are:

```
select City,round(sum(sales),2) as "Total Sales",
round(sum(profit),2) as Total_Profit
from walmart_sales
group by City
order by Total_Profit asc limit 10;
```

City	Total Sales	Total_Profit
Kettering	21975.83	-13358.02
Des Plaines	25746.84	-13164.03
Frisco	21294.83	-12168.93
Pomona	7698.12	-11646.19
Overland	24549.45	-11201.14
Atlantic City	19125.2	-10227.74
Salem	18922.66	-9709.47
Marblehead	24391.16	-9611.91
Helena	16122.11	-8591.99
Galesburg	43261.01	-8046.78

Bottom 10 Cities' total sales and profits with their profit margins

The bottom 3 are Kettering, Des Plaines and Frisco. We have start redesigning some strategies and how we operate in those cities.

## 5. The relationship between discount and sales and the total discount per category ?

First, let's observe the correlation between discount and average sales to understand how impactful one is to the other.

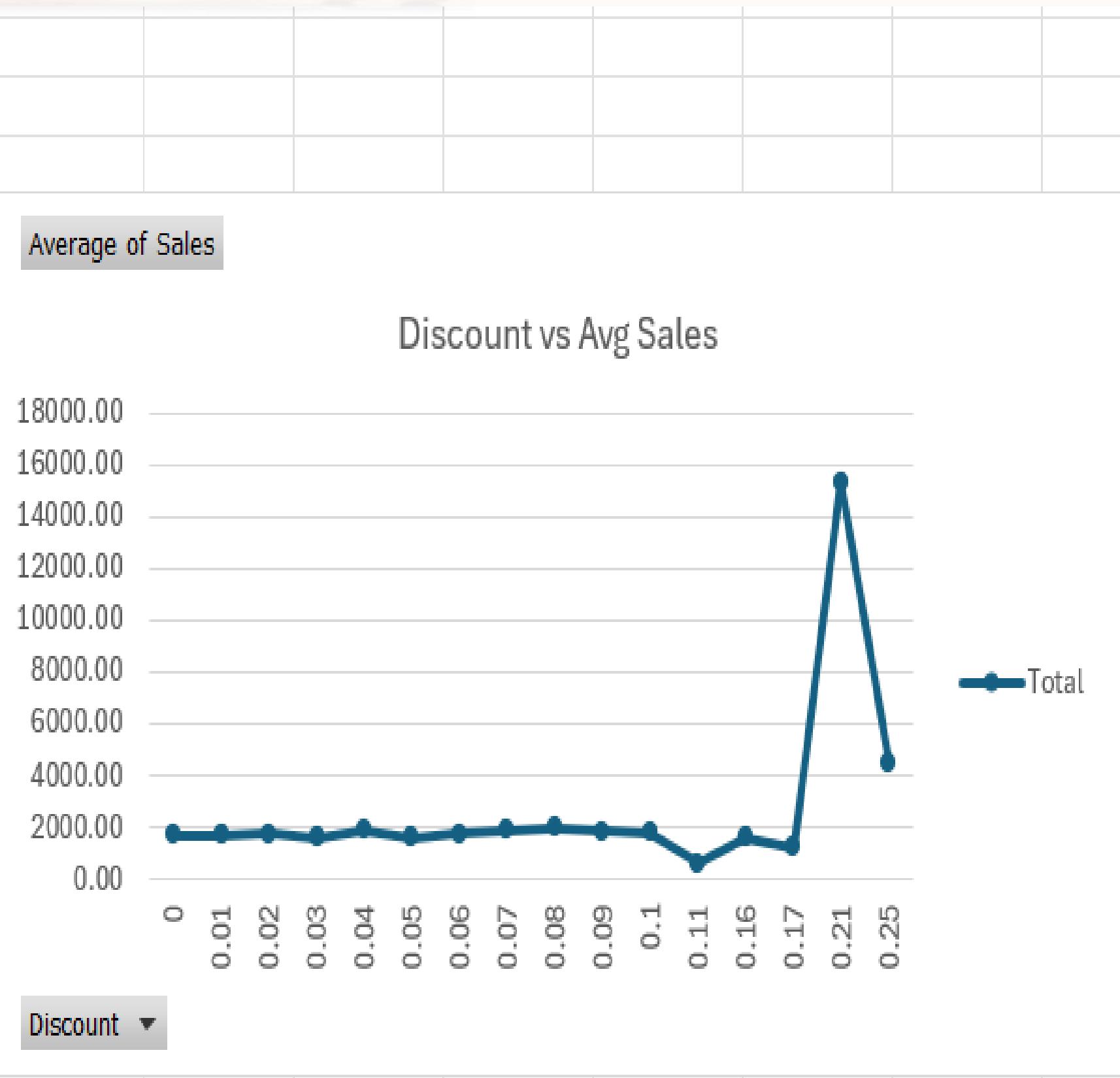
```
select Discount,round(avg(sales),2) as Avg_Sales
from walmart_sales
group by discount
order by discount;
```

Discount	Avg_Sales
0	1710.99
0.01	1710.08
0.02	1749.68
0.03	1611.84
0.04	1870.38
0.05	1619.14
0.06	1744.59
0.07	1868.9
0.08	2001.47
0.09	1844.16
0.1	1799.05
0.11	599.1
0.16	1596.86

Discount vs Avg Sales

Seems that for each discount point, the average sales seem to vary a lot. Let's check the correlation with a graph in Excel

Discount	Average of Sales
0.00	1710.988732
0.01	1710.082457
0.02	1749.676584
0.03	1611.841772
0.04	1870.383091
0.05	1619.144784
0.06	1744.586703
0.07	1868.901507
0.08	2001.469752
0.09	1844.160334
0.10	1799.047879
0.11	599.1
0.16	1596.86
0.17	1217.62
0.21	15341.46
0.25	4462.23
Correlation	0.522902932



Discount vs Avg Sales correlation graph

They almost have no linear relationship. This noted by the correlation coefficient of 0.522 and the shape of the graph.

Let's observe the total discount per product category:

```
select Product_Category,
round(sum(discount),2) as Total_Discount
from Walmart_sales
group by Product_Category
order by Total_Discount desc;
```

Product_Category	Total_Discount
Office Supplies	227.51
Technology	103.01
Furniture	86.65

Most Categories

So Office supplies are the most discounted items Technology and Furniture . We will later dive in into how much profit and sales each generate. Before that, let's zoom in the category section to see exactly what type of products are the most discounted.



```

select Product_Category, Product_Sub_Category,
round(sum(discount),2) as Total_Discount
from Walmart_sales
group by Product_Category, Product_Sub_Category
order by Total_Discount desc;

```

Product_Category	Product_Sub_Category	Total_Discount
Office Supplies	Paper	59.98
Office Supplies	Binders and Binder Accessories	45.29
Technology	Telephones and Communication	42.16
Furniture	Office Furnishings	39.56
Technology	Computer Peripherals	39.05
Office Supplies	Pens and Art Supplies	31.18
Office Supplies	Storage and Organization	26.76
Office Supplies	Appliances	21.82
Furniture	Chairs and Chairmats	19.44
Furniture	Tables	18.21
Technology	Office Machines	17.16
Office Supplies	Labels	14.06
Office Supplies	Envelopes	13.03

Most discounted Prosubcategories (product type)

Paper, Binders and Binder Accessories, Telephones and Office Furnishings are the most discounted items. We should check the sales and profits for the Papers and other items on the list. But first let's move on to the categories per state.

## 6. What category generates the highest sales and profits in each region and state ?

First, let's observe the total sales and total profits of each category with their profit margins:



```
select Product_Category,round(sum(sales),2) as Total_Sales,  

round(sum(profit),2) as Total_Profit,  

round(sum(profit)/sum(sales)*100,2) as Profit_margin  

from walmart_sales  

group by Product_Category  

order by Total_Profit desc;
```

Product_Category	Total_Sales	Total_Profit	Profit_margin
Technology	5984248.5	886313.52	14.81
Office Supplies	3752762.1	518021.42	13.8
Furniture	5178590.51	117433.03	2.27

Product\_Category with their total sales, total profits and profit margins

Among the three categories, it's evident that Technology and Office Supplies stand out for their robust profitability. Plus they seem like a good investment because of their profit margin. To gain deeper insights, let's analyze the regions with the highest total sales and profits within each category. This will provide us with a clearer understanding of each category's regional performance and potential for growth.

Let's observe the highest total sales and total profits per Category in each region:



```
select Region, Product_Category, round(sum(sales),2) as Total_Sales,
round(sum(profit),2) as Total_Profit
from walmart_sales
group by Region, Product_Category
order by Total_Profit desc;
```

Region	Product_Category	Total_Sales	Total_Profit
Central	Technology	1831697.84	286629.68
South	Technology	1421103.72	267851.39
East	Technology	1088313.24	185013.31
Central	Office Supplies	1225756.59	162819.08
West	Technology	1643133.7	146819.14
East	Office Supplies	909889.35	138124.03
South	Office Supplies	819295.45	113316.73
West	Office Supplies	797820.71	103761.58
West	Furniture	1208793.45	48936.81
South	Furniture	909820.17	41339.01
Central	Furniture	1641712.92	32442.44
East	Furniture	1418263.97	-5285.23

Highest total sales and profits per Prod\_Category in each region

Our top-performing Product Categories, particularly in terms of total profits, consistently include Technology across the Central, South, and East regions. Notably, the East region experiences a shortfall with Furniture, failing to achieve profitability and operating at a loss in this category.

Now let's see the highest total sales and total profits per Prod\_Category in each state:

```
select State, Product_Category, round(sum(sales),2) as Total_Sales,
round(sum(profit),2) as Total_Profit
from walmart_sales
group by State, Product_Category
order by Total_Profit desc;
```



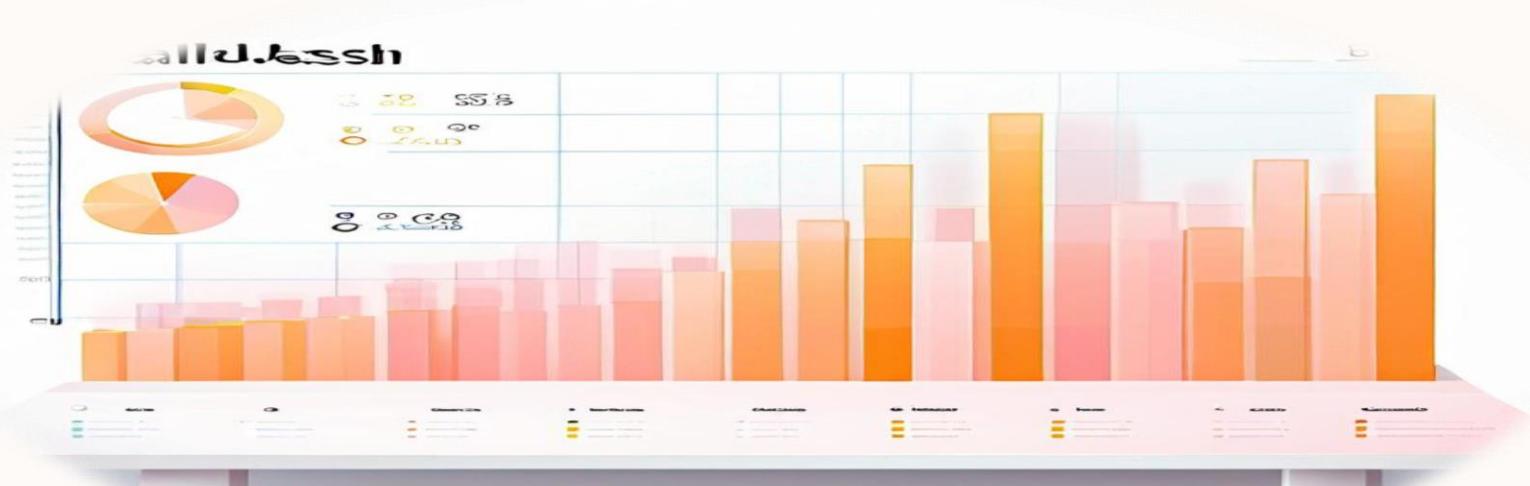
State	Product_Category	Total_Sales	Total_Profit
Ohio	Office Supplies	284592.49	64773.28
New York	Technology	341047.9	61402.95
Illinois	Technology	337190.43	58773.89
Texas	Technology	350243.11	58383.48
California	Office Supplies	366453.99	53482.4
Virginia	Technology	176311.62	42645.05
Minnesota	Technology	213382.27	38795.73
Michigan	Technology	203908.44	37490.39
Oregon	Technology	218764.75	37074.62
Arizona	Technology	130648.36	37014.79
Alabama	Technology	185157.04	36976.15
Florida	Office Supplies	226506.23	36676.61
Illinois	Office Supplies	267117.19	36419.05

Top Highest total sales and profits per Prod\_Category in each state

The table above highlights the top-performing product categories across our states, with Office Supplies leading in Ohio, and Technology reigning supreme in both New York and Illinois. These three categories demonstrate consistent strength across our top markets.

Let's check the least profitable ones by just changing our 'ORDER BY' clause too ascending (ASC)

```
select State, Product_Category,round(sum(sales),2) as Total_Sales,
round(sum(profit),2) as Total_Profit
from walmart_sales group by State,Product_Category order by Total_profit asc;
```





State	Product_Category	Total_Sales	Total_Profit
New Jersey	Furniture	125304.38	-16380.59
Montana	Technology	38573.39	-10279.01
MO	Furniture	80162.49	-6846.34
Michigan	Furniture	164154.62	-6619.04
Utah	Furniture	54195.87	-2889.73
Minnesota	Furniture	179475.87	-2820.61
Mississippi	Furniture	25073.85	-2738.55
Maine	Furniture	139302.25	-2362.46
Pennsylvania	Furniture	179614.9	-2223.99
Arizona	Office Supplies	47297.63	-1683.94
Iowa	Furniture	53472.44	-1396.92
New Hamp...	Furniture	41236.63	-1363.54
Wyoming	Furniture	13456.4	-1031.48

Top Lowest total sales and profits per Prod\_Category in each state



Furniture in New Jersey, Technology in Montana, and Furniture in Michigan and Utah emerge as our most significant losses. Now, let's pivot our focus to Product Categories.

## 7. What Product\_Sub\_Category generates the highest sales and profits in each region ?

Let's observe the total sales and total profits of each Product\_Sub\_Category with their profit margins:

```
select Product_Sub_Category,round(sum(sales),2) as Total_Sales,
round(sum(profit),2) as Total_Profit,
round(sum(profit)/sum(sales)*100,2) as Profit_margin
from walmart_sales group by Product_Sub_Category order by Total_profit desc;
```

Product_Sub_Category	Total_Sales	Total_Profit	Profit_margin
Telephones and Communication	1889314.12	316951.62	16.78
Office Machines	2168697.14	307712.93	14.19
Binders and Binder Accessories	1022957.59	307413.38	30.05
Copiers and Fax	1130361.3	167361.49	14.81
Chairs and Chairmats	1761836.55	149649.73	8.49
Office Furnishings	698093.81	100427.93	14.39
Appliances	736991.54	97158.06	13.18
Computer Peripherals	795875.94	94287.48	11.85
Envelopes	174085.8	48182.6	27.68
Paper	446452.86	45263.2	10.14
Labels	38981.55	13677.17	35.09
Pens and Art Supplies	167107.22	7564.78	4.53
Storage and Organization	1070182.6	6664.15	0.62
Rubber Bands	15006.63	-102.67	-0.68
Scissors Rulers and Trimmers	80996.31	-7799.25	-9.63
Bookcases	822652.04	-33582.13	-4.08

Product\_sub\_category with their total sales, total profits and profit margins

Out of our 17 Product\_sub\_category nationwide, our biggest profits comes from Telephones and communication ,Office Machines, Accessories and Copiers and Fax.

The profits and profit margins on Labels and Binders and Binders Accessories especially are interesting for the long run. Our losses came from Bookcases, Scissors Rulers and Trimmers ,Rubber Bands where we are uncapable of breaking even.

Those 3 should be further reviewed as the sales are there, (except Supplies) but we cannot generate profits from them.

Now let's see the highest total sales and total profits per Product\_sub\_category in each region:



```
select Region,Product_Sub_Category,round(sum(sales),2) as Total_Sales,
round(sum(profit),2) as Total_Profit
from walmart_sales group by Region,Product_Sub_Category order by Total_Profit desc;
```

Region	Product_Sub_Category	Total_Sales	Total_Profit
South	Office Machines	610807.34	155487.64
Central	Office Machines	563394.62	108471.33
Central	Telephones and Communication	613410.11	105941.21
East	Binders and Binder Accessories	294906.73	94617.36
Central	Binders and Binder Accessories	309261.98	90003.92
West	Telephones and Communication	475653.17	81604.44
East	Telephones and Communication	394726.48	64906.37
South	Telephones and Communication	405524.36	64499.6
South	Binders and Binder Accessories	214941.82	61611.92
West	Binders and Binder Accessories	203847.06	61180.18
West	Copiers and Fax	343116.94	59066.14
East	Office Machines	321105.33	55305.43
East	Chairs and Chairmats	469652.2	54180.78
Central	Chairs and Chairmats	651653.98	51534.27
Central	Appliances	317079.05	49445.05
Central	Copiers and Fax	404174.73	44542.3



Top 16 Product\_sub\_category with the highest total sales and total profits in each region

These above display the best Product\_sub\_category per region.

## 8. What segment makes the most of our profits and sales ?

This can be verified with the help of the following query:

```
select *from walmart_sales;
select customer_segment,round(sum(sales),2) as Total_Sales,
round(sum(profit),2) as Total_Profit
from walmart_sales group by customer_segment order by Total_Profit desc;
```

customer_segment	Total_Sales	Total_Profit
Corporate	5716248.36	622373.85
Home Office	3561510.5	384327.97
Small Business	2693463.33	261477.85
Consumer	2944378.92	253588.3

Goods Segment ordered by total profits

The Corporate segment brings in the most profit followed by Home Office, Small Business and Consumer. Let's move on.

## 9. How many customers do we have (unique customer IDs) in total and how much per region and state?

This can be solved with the following;

```
Select
count(distinct(Customer_Name))
as Total_Customer
from Walmart_sales;
```

Total_Customer
795

Total number of customers

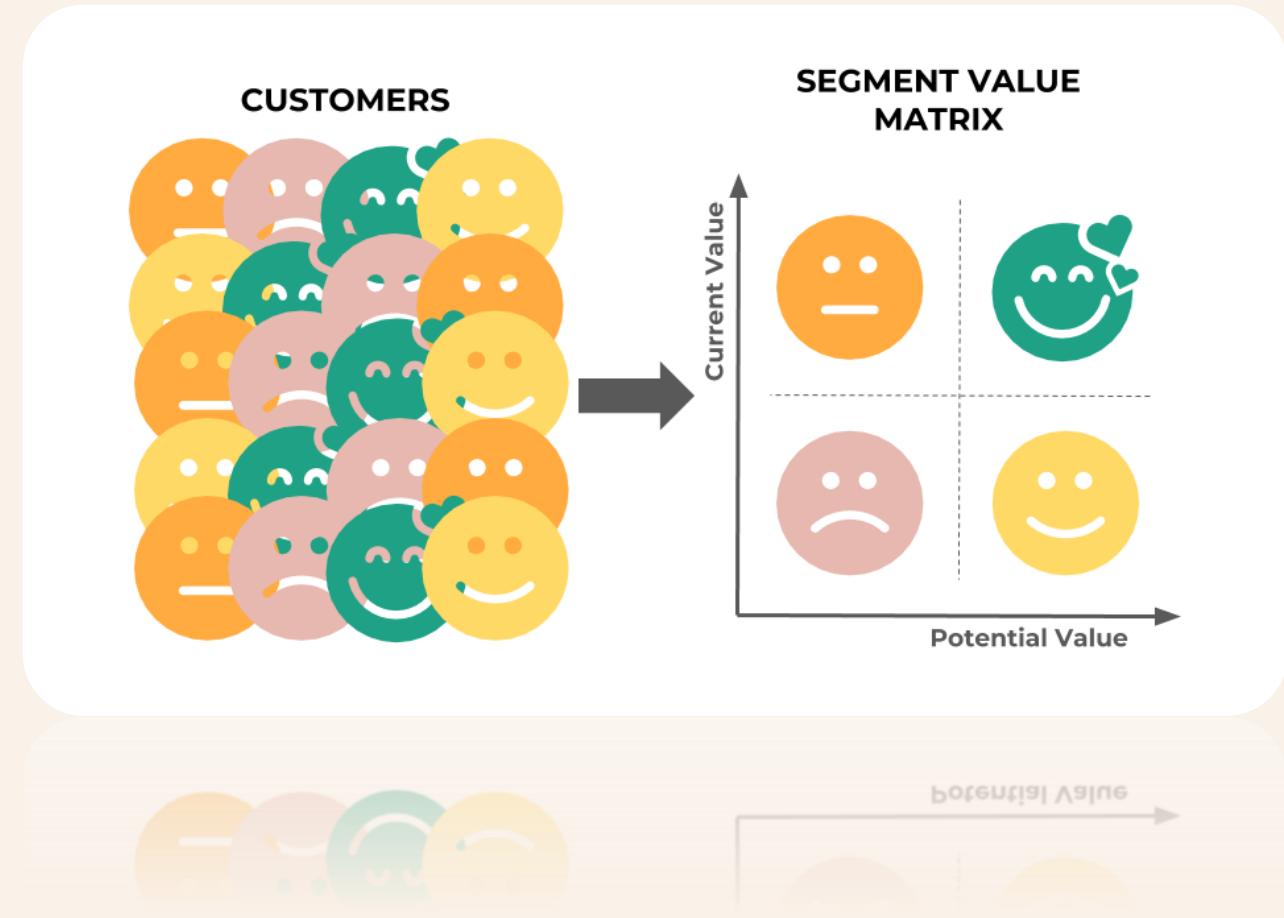
We've had 795 customers between 2012 and 2015. Regionally, we had the following:

```
select Region, count(distinct(Customer_Name)) as Total_Customer
from walmart_sales group by Region
order by Total_Customer desc;
```

Region	Total_Customer
Central	728
West	688
East	659
South	659

Total customers per region

We surely had customers moving around regions which explains why they all do not add up to 795. Since there could be double counting. The Central is the area where we have the biggest market of all. State wise, here are the numbers:



```
select State, count(distinct(Customer_Name)) as Total_Customer
from walmart_sales group by State
order by Total_Customer desc;
```

State	Total_Customer
California	446
Texas	365
Illinois	320
Florida	314
Ohio	281
New York	263
Michigan	218
Washington	187
Minnesota	182
Indiana	180
Pennsylvania	171
Virginia	161
North Carolina	161
New Jersey	146
Georgia	143

We have the most customers in California, Texas and Illinois. The areas where we have the least that passed by there are:

Top 15 states with the most customers

```
select State,count(distinct(Customer_Name)) as Total_Customer
from walmart_sales
group by State
order by Total_Customer asc;
```

State	Total_Customer
Delaware	14
Rhode Island	20
Wyoming	20
South Dakota	26
North Dakota	28
Nevada	36
West Virginia	38
Montana	42
New Hampshire	54
Vermont	57
Nebraska	62
New Mexico	62
Mississippi	66
Kentucky	70
Connecticut	73

Top 15 states with the least customers

## 10. Average shipping time per class and in total.

Finally, the average shipping time regardless of the shipping mode that is chosen is found with the following function:



```
select round(avg(Ship_Date - Order_Date),1) as
Avg_shipping_time
from walmart_sales;
```

Avg_shipping_time
47.0

The shipping time in each shipping mode is:



```
select ship_mode,
round(avg(Ship_Date - Order_Date),1) as Avg_shipping_time
from walmart_sales
group by ship_mode
order by Avg_shipping_time;
```

ship_mode	Avg_shipping_time
Regular Air	44.0
Delivery Truck	52.3
Express Air	59.6

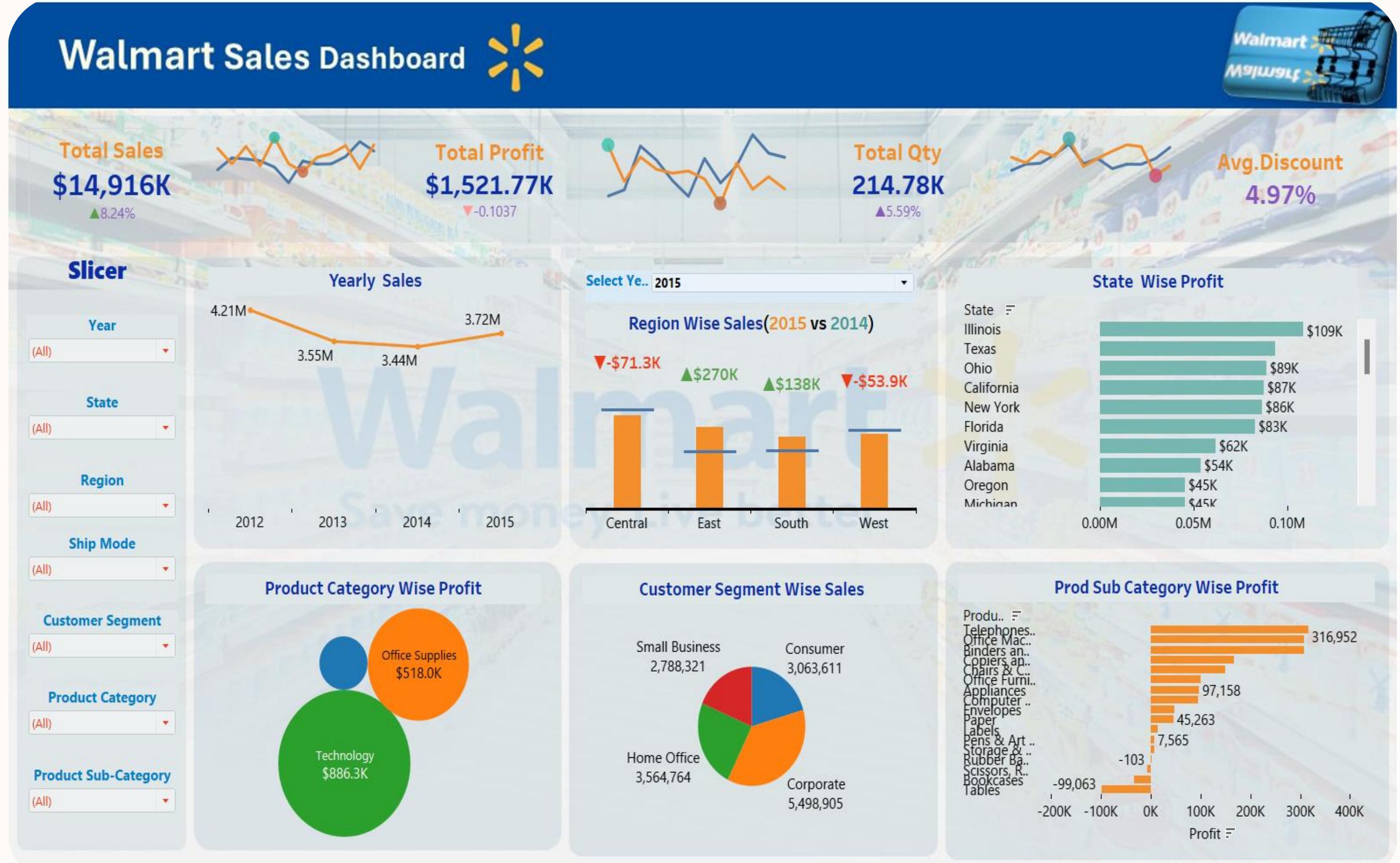
Average shipping time by shipping mode

Here we have all the information we need to transition to our Tableau dashboard. The data visualization is our following step.

## Step 5: Share

- ❖ Onto sharing our observations, below you will find a picture of a screenshot of the interactive dashboard that represents the main KPIs and information on the collected Walmart data which was realized and demonstrated with Tableau.
- ❖ The link to the interactive dashboard will be found under the screenshot below. Here are our findings:
- ❖ This is an in depth analysis of the Walmart dataset. You will find interactive dashboards focused on the yearly sales dashboards displaying data collected from 2012–2015.
- ❖ If there is mainly an interest in the yearly (timeline) data on profits, categories, sub categories, segments and products, then below you will find the direct link to the dashboard:

<https://public.tableau.com/app/profile/goutam.kuir/viz/Walmartretailproject/WalmartSalesDashboard>



- ❖ This is an in depth analysis of the Walmart dataset. You will find interactive dashboards focused on the yearly sales dashboards displaying data collected from 2012–2015.
- ❖ If there is mainly an interest in the yearly (timeline) data on profits, categories, sub categories, segments and products, then below you will find the direct link to the dashboard:

# Conclusions and Future Recommendations

1

## Steady Improvement in Profits and Sales

- ❖ Our profits saw a steady improvement, with sales also experiencing growth despite a brief pause in 2012. Let's strive to maintain this momentum going forward.

2

## Focus on Q4 Profit Maximization

- ❖ Our peak in profitability consistently occurs during Q4. To maximize even more profits, we must make sure to have enough stock and push our marketing and customer service to make the most out of the October – December festive period.

3

## Regional Performance Analysis

- ❖ The most performing regions are the Central then the South, East and West regions in that order. The West region brings in atleast \$100,000 more in sales than the East region but still makes less profits than it. There is work to be done in the West region if we really want to keep that market. However, I believe it is better to take some of the resources in our West region to instead our Central region stores as we are more profitable there and could really establish ourselves as kingpin in that region.

4

## Market and Segment Insights

- ❖ Illinois, Texas and Ohio are most profitable market and most present ones ones especially in terms of sales as states. We have to focus more on them. Our least Profitable markets are Montana, Delaware and West Virginia. Which I believe that we should decrease our presence there or even put a halt at our store locations there as sales in Montana and Delaware are in the \$100,000s but are unable to convert to profits

5

## Clientele and Market Presence

- ❖ Out of the 4 segments, The corporate segment brings in the most profit followed by Home office, Small Business and then Consumer. We must give more importance to the corporate segment even if all the 4 are profitable.

- ❖ Finally, for our clientele, we have 795 customers total, and we have the most customers in California, Texas and Illinois. California and Texas are pretty obvious, we have to be outstanding and be the best of what there is to offer in our respective niche.

**Thank you for your interest and time. Feel free to give your valuable suggestions and connect with me**

<https://www.linkedin.com/in/goutam-kuiiri-949b632a6>