SuperStore Sales Analysis

Using EXCEL, SQL & MS POWER BI

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 ?

Data Cleaning through Excel

- **Dataset source & overview:**

The dataset from **Kaggle's Superstore dataset** contains **9995 rows, including 9994 data rows and 1 header row.** It includes **793** unique customers and **21 columns.** The data is specific to the **United States.**

- **Data Observation & Preparation:**

1. **Check for Missing Data:** Used Conditional Formatting to highlight blanks (red), confirmed no missing values using filters.
2. **Remove Duplicates:** Applied "Remove Duplicates" across all columns; no duplicates found.
3. **Format Columns:** Formatted **Order Date** and **Ship Date** as dates., **Sales** and **Profit** as currency & **Discount** as a percentage .
4. **Final Check:** Reapplied filters to confirm no blanks or formatting issues.

- **Dataset quality & limitation:** The dataset is reliable, original, and comprehensive, but is 6 years old, which may affect its relevance.

- **Dataset Ready for Analysis**: Clean, formatted, and ready for discovering trends and insights.

Analyze the Dataset

For the analysis phase, I extracted the key components of the data to address the business objectives. I then loaded the data into SQL (Big Query) to review the schema, data types, and rows to ensure everything was imported correctly

1. Total sales and total profits of each year

Untitled query

RUN SAVE DOWN

```sql
SELECT
  EXTRACT(Year FROM OrderDate) AS  year,
  ROUND(SUM(Sales),2) AS total_sales,
  ROUND(SUM(Profit),2) AS total_Profit,
FROM `black-rhino-409407.Superstore.sales`
Group BY
  year
ORDER BY
  year ASC
```

Query results

JOB INFORMATION | RESULTS | CHART | JSON

Row	year	total_sales	total_Profit
1	2014	484247.56	49544.06
2	2015	470532.46	61618.69
3	2016	609205.86	81795.27
4	2017	733215.19	93439.77

2 . Total profits and total sales per quarter

```sql
SELECT
 EXTRACT(Year FROM OrderDate) AS  year,
 EXTRACT(Quarter From OrderDate) AS quarter,
 ROUND(SUM(Sales),2) AS total_sales,
 ROUND(SUM(Profit),2) AS total_Profit
FROM `black-rhino-409407.Superstore.sales`
Group BY
 year,quarter
ORDER BY
 year,quarter ASC
```

Query results

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAIL

Row	year	quarter	total_sales	total_Profit
1	2014	1	74447.86	3811.2
2	2014	2	86538.77	11204.16
3	2014	3	143633.18	12804.73
4	2014	4	179627.75	21723.97
5	2015	1	68851.74	9264.94

3 . Average shipping time per class and in total

Untitled query RUN SAVE DOWNLOAD SHARE

```
1  SELECT
2  | ROUND(AVG(DATE_DIFF(ShipDate, OrderDate, DAY)),1) AS avg_days_to_ship
3  FROM `black-rhino-409407.Superstore.sales`
4
```

Query results

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAIL

Row	avg_days_to_ship
1	4.0

Untitled query RUN SAVE DOWNLOAD SHARE

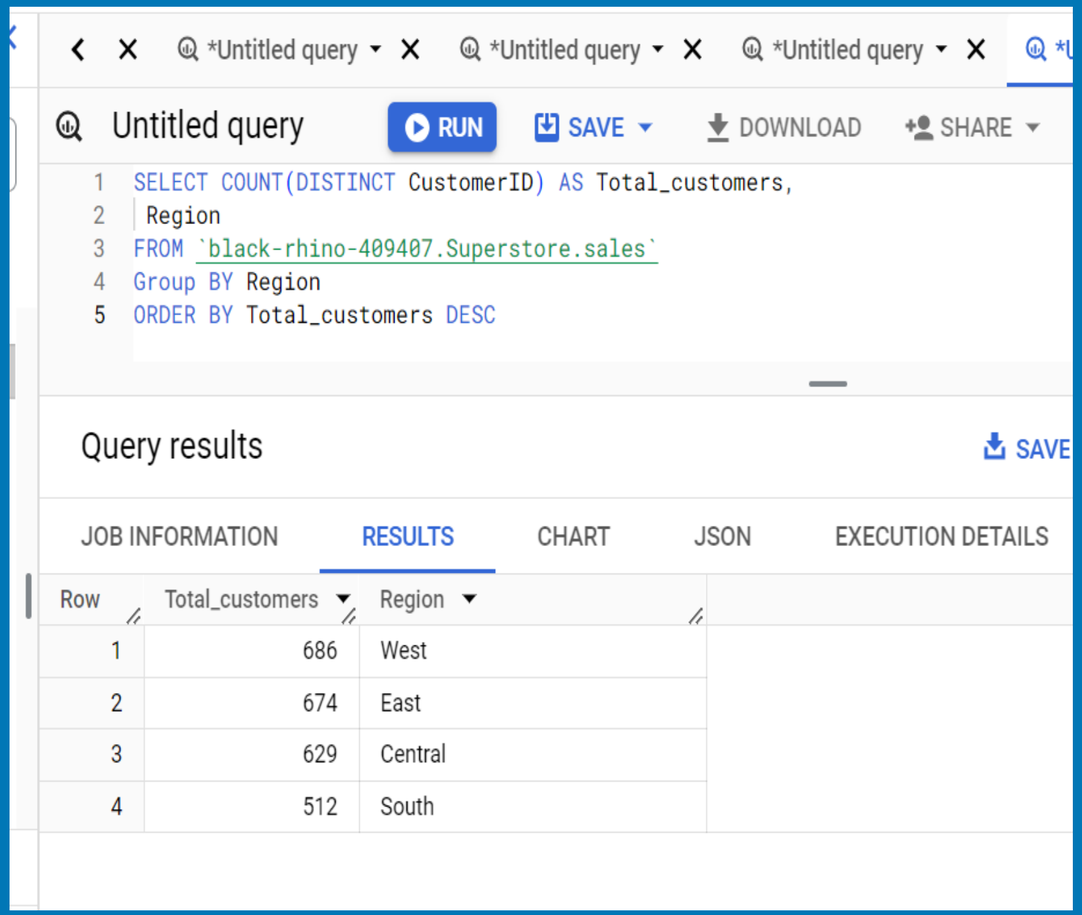
```
1  SELECT
2   ShipMode,
3   ROUND(AVG(DATE_DIFF(ShipDate, OrderDate, DAY)),1) AS avg_days_to_ship
4  FROM `black-rhino-409407.Superstore.sales`
5  GROUP BY ShipMode
6  ORDER BY avg_days_to_ship
```

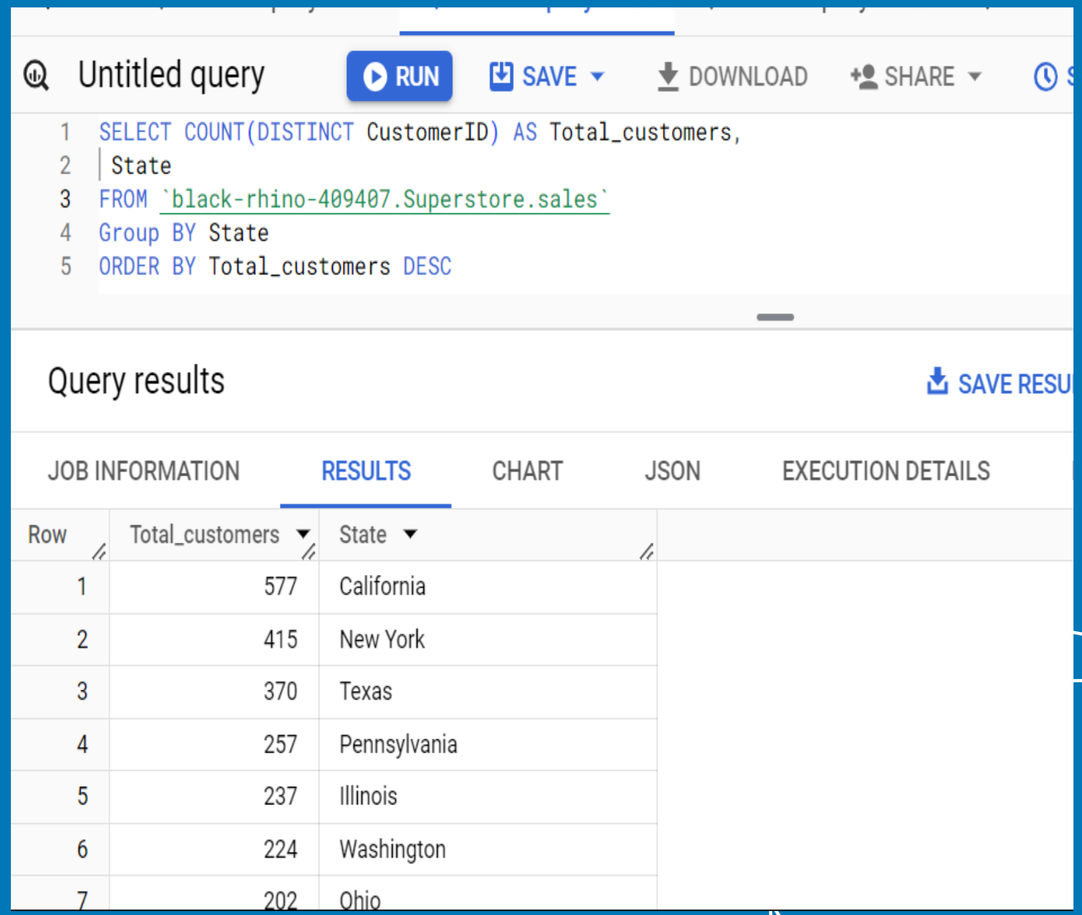
Query results

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETA

Row	ShipMode	avg_days_to_ship
1	Same Day	0.0
2	First Class	2.2
3	Second Class	3.2
4	Standard Class	5.0

4 . Unique customer IDs per region and state





5 . Region that generates the highest sales, profits, and profit margin

Q3.1 RUN SAVE QUERY SHARE SCHEDULE MORE

```sql
SELECT
  Region,
  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 `black-rhino-409407.Superstore.sales`
GROUP BY
  Region
ORDER BY
  Profit_margin DESC
```

Query results SAVE RESULT

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EX

Row	Region	total_sales	total_Profit	Profit_margin
1	West	725457.93	108418.79	14.94
2	East	678781.36	91522.84	13.48
3	South	391721.9	46749.71	11.93
4	Central	501239.88	39706.45	7.92

6 . State that brings in the highest & lowest sales and profits

Q4.2 State ASC RUN SAVE QUERY SHARE SCHEDULE

```sql
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 `black-rhino-409407.Superstore.sales`
GROUP BY State
ORDER BY total_Profit DESC
LIMIT 10
```

Query results SAVE RESULTS

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EX

ow	State	total_sales	total_Profit	Profit_margin
1	California	457687.68	76381.6	16.69
2	New York	310876.2	74038.64	23.82
3	Washington	138641.29	33402.7	24.09
4	Michigan	76269.61	24463.15	32.07
5	Virginia	70636.72	18598.0	26.33
6	Indiana	53555.36	18382.97	34.33

Q4.2 State ASC RUN SAVE QUERY SHARE

```sql
SELECT
  State,
  ROUND(SUM(Sales),2) AS total_sales,
  ROUND(SUM(Profit),2) AS total_Profit
FROM `black-rhino-409407.Superstore.sales`
GROUP BY State
ORDER BY total_Profit ASC
LIMIT 10
```

Query results

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION

Row	State	total_sales	total_Profit
1	Texas	170187.98	-25729.29
2	Ohio	78258.21	-16971.37
3	Pennsylvania	116512.02	-15560.04
4	Illinois	80166.16	-12607.89
5	North Carolina	55603.09	-7490.81
6	Colorado	32108.12	-6527.86

7 . Cities that brings in the highest & lowest sales and profits

Q4.3 CITY DESC

RUN SAVE QUERY SHARE SCHEDULE

```sql
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 `black-rhino-409407.Superstore.sales`
GROUP BY City
ORDER BY total_Profit DESC
LIMIT 10
```

Query results

SAVE RESUL

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS

Row	City	total_sales	total_Profit	Profit_margin
1	New York City	256368.12	62037.08	24.2
2	Los Angeles	175851.33	30440.94	17.31
3	Seattle	119540.74	29156.13	24.39
4	San Francisco	112669.09	17507.39	15.54
5	Detroit	42446.94	13181.79	31.05
6	Lafayette	25036.2	10018.38	40.02

Q4.3 CITY DESC

RUN SAVE QUERY SHARE

```sql
SELECT City,
 ROUND(SUM(Sales),2) AS total_sales,
 ROUND(SUM(Profit),2) AS total_Profit
FROM `black-rhino-409407.Superstore.sales`
GROUP BY City
ORDER BY total_Profit ASC
LIMIT 10
```

Query results

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION

Row	City	total_sales	total_Profit
1	Philadelphia	109077.09	-13837.83
2	Houston	64504.71	-10153.48
3	San Antonio	21843.54	-7299.06
4	Lancaster	9891.48	-7239.08
5	Chicago	48539.59	-6654.55
6	Burlington	21668.08	-3622.88
7	Dallas	20131.9	-2846.55

8 . The relationship between discount and sales and the total discount per category

Q5 Discount

RUN SAVE QUERY

```sql
SELECT
  Discount,
  AVG(Sales) AS AVG_Sales
FROM
  `black-rhino-409407.Superstore.sales`
GROUP BY Discount
ORDER BY Discount
```

Query results

JOB INFORMATION | RESULTS | CHART | JSO

Row	Discount	AVG_Sales
1	0.0	226.7420737807...
2	0.1	578.3968085106...
3	0.15	529.9728846153...
4	0.2	209.0769155045...
5	0.3	454.7434361233...
6	0.32	536.7944444444...
7	0.4	565.1350970873...

```sql
SELECT Category,
  ROUND(SUM(Discount),2) AS total_Discount
FROM `black-rhino-409407.Superstore.sales`
GROUP BY Category
ORDER BY total_Discount desc
```

Query results

JOB INFORMATION | RESULTS | CHART | JSON

ow	Category	total_Discount
1	Office Supplies	947.8
2	Furniture	368.89
3	Technology	244.4

9 . Category that generates the highest sales, profits & profit margin

Untitled query

RUN SAVE DOWNLOAD SHARE SCH

```sql
SELECT
 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 `black-rhino-409407.Superstore.sales`
GROUP BY Category
ORDER BY profit_margin desc
```

Query results

SAVE RESULTS

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EX

Row	Category	total_sales	total_profit	Profit_margin
1	Technology	836154.1	145455.66	17.4
2	Office Supplies	719046.99	122490.88	17.04
3	Furniture	741999.98	18451.25	2.49

10 . Category that generates the highest sales and profits in each region and state

Q6. RUN SAVE QUERY SHARE SCHEDULE MORE

```
1  SELECT Category,Region,
2  Round(SUM(Sales),2) as total_sales,
3  Round(sum(Profit),2) AS total_profit,
4  FROM `black-rhino-409407.Superstore.sales`
5  GROUP BY Category, Region
6  ORDER BY total_profit desc
```

Query results SAVE RESULTS

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GR

Row	Category	Region	total_sales	total_profit
1	Office Supplies	West	220853.2	52609.89
2	Technology	East	264974.04	47462.12
3	Technology	West	251991.86	44303.98
4	Office Supplies	East	205516.15	41014.52
5	Technology	Central	170416.29	33697.6
6	Technology	South	148771.91	19991.96
7	Office Supplies	South	125651.31	19986.52

Q6.1 RUN SAVE QUERY SHARE SCHEDULE MORE

```
1  SELECT Category,State,
2  Round(SUM(Sales),2) AS total_sales,
3  Round(SUM(Profit),2) AS total_profit
4  FROM `black-rhino-409407.Superstore.sales`
5  GROUP BY State,Category
6  ORDER BY total_profit desc
```

Query results SAVE RESULTS

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GR

Row	Category	State	total_sales	total_profit
1	Technology	New York	127483.48	42186.89
2	Office Supplies	California	142351.89	37748.46
3	Technology	California	159271.12	29470.23
4	Office Supplies	New York	90020.03	25994.03
5	Technology	Washington	50536.71	15019.42
6	Office Supplies	Michigan	37723.76	15005.31
7	Office Supplies	Washington	40084.42	11189.48

11. Names of the most and least profitable products

```sql
SELECT ProductName,
 ROUND(SUM(sales),2) AS total_sales,
 ROUND(SUM(profit),2) AS total_profit
FROM `black-rhino-409407.Superstore.sales`
GROUP BY ProductName
ORDER BY total_profit DESC
```

Press Alt+F1 fo

Query results

SAVE RESULTS | EXPLORE

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	ProductName	total_sales	total_profit
1	Canon imageCLASS 2200 Advanced Copier	61599.83	25199.94
2	Fellowes PB500 Electric Punch Plastic Comb Binding Machine with Manual Bind	27453.38	7753.06
3	Hewlett Packard LaserJet 3310 Copier	18839.68	6983.89
4	Canon PC1060 Personal Laser Copier	11619.83	4570.94
5	HP Designjet T520 Inkjet Large Format Printer - 24" Color	18374.9	4094.98
6	Ativa V4110MDD Micro-Cut Shredder	7699.89	3772.95

```sql
SELECT ProductName,
 ROUND(SUM(sales),2) AS total_sales,
 ROUND(SUM(profit),2) AS total_profit
FROM `black-rhino-409407.Superstore.sales`
GROUP BY ProductName
ORDER BY total_profit ASC
```

uery results

SAVE RESULTS

OB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXI

	ProductName	total_sales	total_profit
1	Cubify CubeX 3D Printer Double Head Print	11099.96	-8879.97
2	Lexmark MX611dhe Monochrome Laser Printer	16829.9	-4589.97
3	Cubify CubeX 3D Printer Triple Head Print	7999.98	-3839.99
4	Chromcraft Bull-Nose Wood Oval Conference Tables & B...	9917.64	-2876.11
5	Bush Advantage Collection Racetrack Conference Table	9544.72	-1934.4
6	GBC DocuBind P400 Electric Binding System	17965.07	-1878.17
7	Cisco TelePresence System EX90 Videoconferencing Unit	22638.48	-1811.08

12 . Segment that makes the most of our profits and sales

```sql
SELECT Segment,
 ROUND(SUM(sales),2) AS total_sales,
 ROUND(SUM(profit),2) AS total_profit
FROM `black-rhino-409407.Superstore.sales`
GROUP BY Segment
ORDER BY total_profit DESC
```

Query results

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTIO

ow	Segment	total_sales	total_profit
1	Consumer	1161401.34	134119.33
2	Corporate	706146.44	91979.45
3	Home Office	429653.29	60299.01

13 . Customer rewards program

```sql
SELECT customerid,
 Round(SUM(sales),2) AS total_sales,
 ROUND(SUM(profit),2) AS total_profit
FROM `black-rhino-409407.Superstore.sales`
GROUP BY customerid
ORDER BY total_sales DESC
LIMIT 15
```

Query results

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION

ow	customerid	total_sales	total_profit
1	SM-20320	25043.07	-1980.75
2	TC-20980	19052.22	8981.32
3	RB-19360	15117.35	6976.09
4	TA-21385	14595.62	4703.8
5	AB-10105	14473.57	5444.81
6	KL-16645	14175.23	806.84
7	SC-20095	14142.34	5757.42

Insights & Analysis

1. **Profit Optimization:**
   - Focus on boosting stock and marketing efforts in 4 quarter to maximize festive season profits.
   - Shift resources from the low-profit Central region to the West, focusing on high-performing states like California and New York.
   - Invest more in Technology and Office Supplies, which are the most profitable categories.
2. **Emerging Trends:**
   - The West and East regions are highly profitable, while the Central region lags despite strong sales.
   - Cities like New York, Los Angeles, and Seattle are top performers, while Philadelphia and Houston underperform.
   - Furniture has low margins despite profitability.
3. **Recommendations:**
   - Prioritize profitable regions and reduce stock in underperforming states like Texas and Ohio.
   - Capitalize on the 4 quarter festive season by increasing inventory and marketing.
   - Re-evaluate strategies for low-margin products like Furniture to boost profitability.

To be continued.....

In Part 2, I will upload an interactive dashboard created using **Power BI**, which showcases the main KPIs and insights gathered from the Superstore data. This dashboard provides a visual representation of the key metrics and trends identified during our analysis.

Stay tuned for the upload.