EXPLORATORY DATA ANALYSIS USING SQL

SALES ANALYSIS

Presented by

DENIS M

1.What is the total sales amount and quantity over time?

```sql
SELECT
    d.year,
    CONCAT(ROUND(SUM(sales_amount/1000000),2)," M") AS Total_sales,
    CONCAT(ROUND(SUM(sales_qty/1000000),2)," M") AS Total_Sold_quantity
FROM transactions t
JOIN date d
ON t.order_date=d.date
GROUP BY d.year;
```

year	Total_sales	Total_Sold_quantity
2017	92.88 M	0.18 M
2018	413.69 M	0.78 M
2019	336.02 M	0.65 M
2020	142.22 M	0.27 M

2.Which products generate the highest and lowest revenue?

```sql
WITH CTE AS (
SELECT
    product_code,SUM(sales_amount) AS Total_sales
FROM  transactions
GROUP BY product_code
ORDER BY Total_sales DESC)
SELECT *
FROM CTE
WHERE Total_sales = (SELECT MAX(Total_sales) FROM CTE)
OR Total_sales = (SELECT MIN(Total_sales) FROM CTE);
```

	product_code	Total_sales
▶	Prod318	68967202
	Prod111	65

3. Highest revenue generating products by each year?

```sql
WITH CTE AS (
SELECT
    d.year,
    t.product_code,
    CONCAT(ROUND(SUM(t.sales_amount)/1000000,2)," M") AS Total_sales,
    RANK() OVER(PARTITION BY year ORDER BY year,SUM(t.sales_amount) DESC)
AS rnk
FROM
    transactions t
JOIN date d
ON t.order_date=d.date
GROUP BY
    d.year,t.product_code)
SELECT year,product_code,Total_sales
FROM CTE
WHERE rnk=1;
```

year	product_code	Total_sales
2017	Prod316	8.25 M
2018	Prod318	31.38 M
2019	Prod318	21.31 M
2020	Prod318	8.79 M

4.Which customer type spends the most, and what is their average purchase value?

```sql
SELECT
    c.customer_type, ROUND(AVG(t.sales_amount)) Average_purchase_value
FROM
    transactions t
JOIN
    customers c
ON t.customer_code=c.customer_code
GROUP BY
    c.customer_type
ORDER BY
    Average_purchase_value DESC;
```

customer_type	Average_purchase_value
Brick & Mortar	7740
E-Commerce	4604

5.What is the distribution of profit margins across different products?

```sql
SELECT
    product_code,
    ROUND(AVG(profit_margin)) AS Avg_profit_margin
FROM
    transactions
GROUP BY
    product_code
ORDER BY Avg_profit_margin DESC;
```

product_code	Avg_profit_margin
Prod083	67514
Prod077	50179
Prod153	27962
Prod144	16881
Prod001	16084
Prod067	15163
Prod329	13911
Prod155	13090
Prod171	12458
Prod201	12078
Prod322	11621
Prod012	7412
Prod086	6519
Prod325	5789
Prod087	5010
Prod308	4417
Prod202	4128

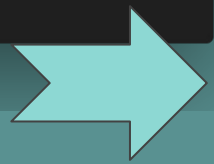
6.How do sales vary across different market zones?

```sql
SELECT m.zone,
    CONCAT(ROUND(SUM(t.sales_amount)/1000000,2)," M") AS Total_sales,
    CONCAT(ROUND(SUM(t.sales_qty)/1000000,2)," M") AS Sold_quantity,
    ROUND(AVG(t.profit_margin)) AS Avg_Profit_margin
FROM
    transactions t
JOIN
    markets m
ON t.market_code=m.markets_code
GROUP BY
    m.zone
ORDER BY SUM(t.sales_amount) DESC;
```

zone	Total_sales	Sold_quantity	Avg_Profit_margin
North	675.53 M	1.27 M	222
Central	263.72 M	0.76 M	119
South	45.56 M	0.40 M	124

7..Which customers have made the most purchases, and what is their total spend?

```sql
SELECT
    c.custmer_name,
    COUNT(c.customer_code) AS Total_Transactions,
    CONCAT(ROUND(SUM(t.sales_amount)/1000000,2)," M") AS Total_purchase_va
lue
FROM
    transactions t
JOIN
    customers c
ON t.customer_code=c.customer_code
GROUP BY
    c.custmer_name
ORDER BY SUM(t.sales_amount) DESC;
```



custmer_name	Total_Transactions	Total_purchase_value
Electricalsara Stores	13819	413.33 M
Electricalslytical	4686	49.64 M
Excel Stores	9236	49.12 M
Premium Stores	19938	44.91 M
Nixon	17327	43.89 M
Info Stores	6017	35.1 M
Control	2182	31.77 M
Surge Stores	4971	28.65 M
Acclaimed Stores	1384	21.2 M
Forward Stores	1612	21.03 M
Epic Stores	5413	18.75 M
Nomad Stores	4514	17.74 M
Electricalsocity	1371	17.49 M
Modular	3615	17.38 M
Atlas Stores	4448	16.67 M
Leader	130	16.53 M
Surface Stores	1704	15.24 M
Integration Stores	1856	13.98 M
Logic Stores	2872	13.2 M
Path	759	13 M
Unity Stores	1177	12.59 M
Electricalsopedia Stores	11721	10.28 M
Flawless Stores	1831	9.16 M

8.What are the peak sales months, and how do they compare year-over-year?

```sql
SELECT
    d.year,
    d.month_name,
    CONCAT(ROUND(SUM(t.sales_amount)/1000000,2)," M") AS Total_sales
FROM
    transactions t
JOIN
    date d
ON t.order_date=d.date
GROUP BY
    d.year,
    d.month_name
ORDER BY d.year,SUM(t.sales_amount) DESC;
```

year	month_name	Total_sales
2017	November	35 M
2017	December	31.8 M
2017	October	26.09 M
2018	January	42.52 M
2018	August	39.45 M
2018	March	38.15 M
2018	July	35.98 M
2018	April	35.89 M
2018	February	35.24 M
2018	June	34.74 M
2018	May	32.19 M
2018	November	31.3 M
2018	December	30.4 M
2018	September	30.1 M
2018	October	27.71 M
2019	July	35.11 M
2019	August	32.2 M
2019	January	31.53 M

9.How does the cost of goods sold (COGS) compare to total sales across products?

```sql
WITH CTE AS (
SELECT
    product_code,
    SUM(sales_amount) Total_sales,
    ROUND(SUM(cost_price),2) Total_cost_price
FROM
    transactions
GROUP BY
    product_code)
SELECT *,ROUND((Total_sales-Total_cost_price),2) AS Gross_Profit,
CASE
WHEN Total_sales>Total_cost_price THEN "Profit"
WHEN Total_sales<Total_cost_price THEN "Loss"
END AS Sales_Result
FROM CTE;
```

product_code	Total_sales	Total_cost_price	Gross_Profit	Sales_Result
Prod279	697698	677304.5	20393.5	Profit
Prod278	544437	531123.65	13313.35	Profit
Prod294	1106577	1065499.52	41077.48	Profit
Prod281	471692	456287.81	15404.19	Profit
Prod292	2411844	2287599.16	124244.84	Profit
Prod290	2824766	2764667.73	60098.27	Profit
Prod280	655320	649041.29	6278.71	Profit
Prod056	2334113	2265100.75	69012.25	Profit
Prod131	7860293	7566494.55	293798.45	Profit
Prod295	850278	832135.38	18142.62	Profit
Prod055	359755	349601.31	10153.69	Profit
Prod286	1306053	1270305.05	35747.95	Profit
Prod271	2598157	2551667.78	46489.22	Profit
Prod263	314180	298743.17	15436.83	Profit
Prod065	16259345	15781705.08	477639.92	Profit
Prod058	3094328	2946306.27	148021.73	Profit
Prod057	7347637	7098367.5	249269.5	Profit
Prod270	491957	486492.27	5464.73	Profit
Prod269	612139	596316.12	15822.88	Profit
Prod053	15135392	14724718.55	410673.45	Profit
Prod054	3969870	3820614.36	149255.64	Profit
Prod300	783669	757237.24	26431.76	Profit
Prod302	200005	198974.88	1030.12	Profit

10.What is the average profit margin per market, and which markets are most profitable?

```sql
SELECT
    m.markets_name,
    ROUND(AVG(t.profit_margin),2) AS Avg_profit_margin
FROM
    transactions t
JOIN
    markets m
ON t.market_code=m.markets_code
GROUP BY
    m.markets_name
HAVING
    Avg_profit_margin > 0
ORDER BY
    Avg_profit_margin DESC;
```

markets_name	Avg_profit_margin
Patna	453.39
Mumbai	430.41
Bhubaneshwar	317.82
Surat	314.34
Lucknow	299.63
Chennai	292.39
Delhi NCR	270.33
Bhopal	171.68
Kochi	142.95
Ahmedabad	142.16
Nagpur	29.62
Hyderabad	23.09

11.How has customer purchasing behavior changed over time?

```sql
SELECT
    c.custmer_name,
    d.year,
    CONCAT(ROUND(SUM(t.sales_amount)/1000000,2)," M") AS Sales
FROM
    transactions t
JOIN
    customers c
ON t.customer_code=c.customer_code
JOIN
    date d
ON t.order_date=d.date
GROUP BY
    c.custmer_name, d.year
ORDER BY c.custmer_name,d.year;
```

custmer_name	year	Sales
Acclaimed Stores	2017	1.7 M
Acclaimed Stores	2018	9.34 M
Acclaimed Stores	2019	7.04 M
Acclaimed Stores	2020	3.12 M
All-Out	2017	0.88 M
All-Out	2018	2.42 M
All-Out	2019	2.06 M
All-Out	2020	0.71 M
Atlas Stores	2017	1.76 M
Atlas Stores	2018	6.35 M
Atlas Stores	2019	6.37 M
Atlas Stores	2020	2.19 M
Control	2017	2.58 M
Control	2018	15 M
Control	2019	10 M
Control	2020	4.18 M
Electricalsara Stores	2017	36.1 M
Electricalsara Stores	2018	173.05 M
Electricalsara Stores	2019	138.54 M
Electricalsara Stores	2020	65.64 M
Electricalsbea Stores	2017	0.01 M
Electricalsbea Stores	2018	0.12 M
Electricalsbea Stores	2019	0.16 M

12.What is the distribution of sales across different products?

```sql
SELECT
    COALESCE(P.product_code, '[Blank]') AS product_code,
    SUM(t.sales_amount) AS Total_sales
FROM
    transactions t
LEFT JOIN
    products p
ON t.product_code=p.product_code
GROUP BY
    t.product_code
ORDER BY SUM(t.sales_amount) DESC;
```

product_code	Total_sales
[Blank]	68967202
[Blank]	60883452
[Blank]	41455364
[Blank]	34381481
[Blank]	31468996
[Blank]	26594157
Prod040	23581969
[Blank]	22188881
[Blank]	17873777
Prod159	17660106
[Blank]	17614884
[Blank]	16472242
Prod065	16259345
[Blank]	16233500
Prod018	15597920
Prod053	15135392

13.What is the distribution of sales across different product_type?

```sql
SELECT
    COALESCE(P.product_type, '[Blank]') AS product_type,
    SUM(t.sales_amount) AS Total_sales
FROM
    transactions t
LEFT JOIN
    products p
ON t.product_code=p.product_code
GROUP BY
    p.product_type
ORDER BY SUM(t.sales_amount) DESC;
```

product_type	Total_sales
[Blank]	468961239
Own Brand	369812648
Distribution	146039576


Thank
You