

EXPLORATORY DATA ANALYSIS OF MALL TRAFFIC & SALES DATA

PRESENTED BY
DENIS M



-- Sales Performance & Trends

-- total sales and revenue fluctuation on a yearly and monthly basis

```
WITH monthly_sales AS (  
    SELECT YEAR(`invoice date`) AS `Year`,  
           MONTH(`invoice date`) AS `Monthnumber`,  
           MONTHNAME(`invoice date`) AS `Month`,  
           SUM(total_sales) AS `Sales`  
FROM sales  
GROUP BY YEAR(`invoice date`), MONTH(`invoice date`), MONTHNAME(`invoice date`)  
ORDER BY `Year`, `Monthnumber`  
  
SELECT `Month`,  
       CONCAT(ROUND(SUM(CASE WHEN `Year`=2021 THEN Sales ELSE 0 END)/1000000,2), " M") AS "2021 Sales",  
       CONCAT(ROUND(SUM(CASE WHEN `Year`=2022 THEN Sales ELSE 0 END)/1000000,2), " M") AS "2022 Sales",  
       CONCAT(ROUND(SUM(CASE WHEN `Year`=2023 THEN Sales ELSE 0 END)/1000000,2), " M") AS "2023 Sales"  
FROM monthly_sales  
GROUP BY `Month`;
```

Month	2021 Sales	2022 Sales	2023 Sales
January	9.31 M	9.92 M	6.53 M
February	8.81 M	8.59 M	6.57 M
March	10.06 M	9.86 M	0.90 M
April	9.73 M	9.62 M	1.01 M
May	9.77 M	9.90 M	0.94 M
June	9.49 M	9.73 M	1.13 M
July	10.14 M	9.66 M	0.92 M
August	9.49 M	9.41 M	0.91 M
September	8.91 M	9.60 M	0.73 M
October	10.16 M	10.07 M	0.82 M
November	9.06 M	9.15 M	0.49 M
December	9.63 M	9.92 M	0.56 M

INSIGHTS

- ❖ **October 2021 recorded the highest sales at 10.13 million, reflecting strong market performance, and November 2023 marked the lowest point in sales, dropping to 0.49 million, indicating a significant downturn.**
- ❖ **The Sales remained strong throughout 2021 and early 2022, but began to gradually decline heading into 2023, suggesting a loss in momentum.**
- ❖ **A notable drop in sales during the first two quarters of 2023 raises concerns and highlights the need for a thorough investigation into internal and external factors.**



```
-- year-over-year sales growth rate
```

```
WITH yearly_sales AS (  
    SELECT YEAR(`invoice date`) AS `Year`, SUM(total_sales) AS Sales  
    FROM sales  
    GROUP BY YEAR(`invoice date`)  
,  
yoy_calc AS (  
    SELECT  
        cur.`Year`,  
        cur.Sales AS current_year_sales,  
        pre.Sales AS previous_year_sales,  
        ROUND(  
            ((cur.Sales - pre.Sales) / pre.Sales) * 100, 2) AS yoy_growth_rate  
    FROM yearly_sales AS cur  
    LEFT JOIN yearly_sales AS pre  
        ON cur.`Year` = pre.`Year` + 1  
)  
SELECT *  
FROM yoy_calc  
WHERE previous_year_sales IS NOT NULL  
ORDER BY `Year`;
```

	Year	current_year_sales	previous_year_sales	yoy_growth_rate
▶	2022	115439155	114562735	0.77
	2023	21508868	115439155	-81.37

INSIGHTS

- ❖ 2020 witnessed a modest growth with a YoY increase of 0.77%, indicating stable growth.
- ❖ 2023 experienced a dramatic decline, with a YoY growth rate plummeting to -81.37%, signaling a significant drop in overall sales performance compared to the previous year.



-- sales performance across different shopping malls, the top three revenue-generating malls

```
WITH Sales_by_shopping_mall AS (  
    SELECT sm.shopping_mall,  
           SUM(s.total_sales) AS Sales  
    FROM shopping_mall AS sm  
    JOIN sales AS s  
    ON sm.shopping_mall=s.shopping_mall  
    GROUP BY sm.shopping_mall  
    ORDER BY Sales DESC  
)  
SELECT shopping_mall AS "Top 3 Malls"  
FROM Sales_by_shopping_mall  
LIMIT 3;
```

	Top 3 Malls
▶	Del Amo Fashion Center
	South Coast Plaza
	Westfield Century City

INSIGHTS

- ❖ **Top 3 revenue-generating Malls are**
Del Amo Fashion Center
South Coast Plaza
Westfield Century City .



```
-- distribution of sales across different product categories, and three categories  
perform best in each shopping mall
```

```
SELECT category, CONCAT(ROUND(SUM(total_sales)/1000000, 2), " M") AS Sales  
FROM sales  
GROUP BY category  
ORDER BY SUM(total_sales) DESC;
```

```
WITH Category_wise_sales AS (  
    SELECT sm.shopping_mall,  
           s.category,  
           SUM(s.total_sales) AS Sales  
    FROM shopping_mall AS sm  
    JOIN sales AS s  
    ON sm.shopping_mall=s.shopping_mall  
    GROUP BY sm.shopping_mall ,s.category  
    ORDER BY shopping_mall ,Sales DESC  
)  
top_performing_categories AS (SELECT *,  
ROW_NUMBER() OVER(PARTITION BY shopping_mall) AS Performance_rank  
FROM Category_wise_sales)  
SELECT shopping_mall , category, Sales  
FROM top_performing_categories  
WHERE Performance_rank<=3;
```

shopping_mall	category	Sales
Beverly Center	Clothing	5792330
Beverly Center	Shoes	3328001
Beverly Center	Technology	2516850
Del Amo Fashion Center	Clothing	22946863
Del Amo Fashion Center	Shoes	13468089
Del Amo Fashion Center	Technology	11828250
Fashion Valley	Clothing	5697186
Fashion Valley	Shoes	3535670
Fashion Valley	Technology	2987250
Glendale Galleria	Clothing	5604461
Glendale Galleria	Shoes	3194774
Glendale Galleria	Technology	3066000
Irvine Spectrum Center	Clothing	5706177
Irvine Spectrum Center	Shoes	3243983
Irvine Spectrum Center	Technology	3051300
South Coast Plaza	Clothing	22608983
South Coast Plaza	Shoes	13383448
South Coast Plaza	Technology	11944800
Stanford Shopping Cen...	Clothing	11567790
Stanford Shopping Cen...	Shoes	7149963
Stanford Shopping Cen...	Technology	5327700
The Grove	Clothing	5590365
The Grove	Technology	3094350
The Grove	Shoes	3089754
Westfield Century City	Clothing	17226258
Westfield Century City	Shoes	9519492

INSIGHTS

- ❖ Clothing emerged as the top-performing category, significantly ahead of all other categories.
- ❖ Shoes and Technology also contributed substantially, showing strong customer interest and consistent purchasing trends.
- ❖ Books and Souvenirs recorded the lowest sales, indicating either limited demand.



-- How does total sales revenue differ across regions, and regions contribute the most to overall sales

```
WITH Regional_sales AS (  
    SELECT sm.location,  
           SUM(s.total_sales) AS Sales  
    FROM shopping_mall AS sm  
    JOIN sales AS s  
    ON sm.shopping_mall=s.shopping_mall  
    GROUP BY sm.location  
    ORDER BY Sales DESC  
)  
Overall_sales AS (  
    SELECT SUM(total_sales) AS Total_regional_Sales  
    FROM sales  
)  
SELECT location,Sales, CONCAT(ROUND((Sales*100/Total_regional_Sales)), " %")  
AS percentage  
FROM Regional_sales  
CROSS JOIN Overall_sales;
```

	location	Sales	percentage
▶	Los Angeles	62014064	25 %
	Torrance	50873460	20 %
	Costa Mesa	50555257	20 %
	Palo Alto	25380412	10 %
	Santa Clara	24619303	10 %
	San Diego	12901309	5 %
	Irvine	12645364	5 %
	Glendale	12521589	5 %

INSIGHTS

- ❖ Los Angeles stands out as the top-performing region, contributing a significant 25% of the total sale highlighting its dominant customer base
- ❖ Torrance and Costa Mesa each contribute 20% to overall sales.
- ❖ Palo Alto and Santa Clara also hold a 10% share collectively.
- ❖ Palo Alto and Santa Clara also hold a 10% share collectively.



```
-- total sales quantity by category, and percentage of total sales does each category contribute
```

```
WITH sold_quantity_by_category AS (  
    SELECT category, SUM(quantity) Total_sold_quantity  
    FROM sales  
    GROUP BY category  
    ORDER BY Total_sold_quantity DESC  
),  
Over_all_sales AS(  
    SELECT SUM(quantity) AS Over_all_sold_quantity  
    FROM sales  
)  
SELECT category, Total_sold_quantity, CONCAT(ROUND((Total_sold_quantity*100/  
Over_all_sold_quantity)), " %") AS percentage  
FROM sold_quantity_by_category  
CROSS JOIN Over_all_sales;
```

	category	Total_sold_quantity	percentage
▶	Clothing	103558	35 %
	Cosmetics	45465	15 %
	Food & Beverage	44277	15 %
	Toys	30321	10 %
	Shoes	30217	10 %
	Technology	15021	5 %
	Books	14982	5 %
	Souvenir	14871	5 %

INSIGHTS

- ❖ Clothing dominates product category sales, generating a remarkable ₦113.99 million, making it the clear leader in customer demand.
- ❖ Shoes and Technology categories also saw strong sales performance, confirming their popularity.
- ❖ Books and Souvenirs were the least purchased.



```
-----  
-- Customer Preferences & Behavior  
-----
```

```
-- Purchasing behavior of customers based on age group across product categories
```

```
WITH sales_by_age_group AS (  
    SELECT c.age_group, s.category, SUM(total_sales) AS Sales  
    FROM customers AS c  
    JOIN sales AS s  
    ON c.customer_id=s.customer_id  
    GROUP BY c.age_group, s.category  
    ORDER BY c.age_group, Sales DESC  
)  
SELECT category,  
SUM(CASE WHEN age_group="Youth" THEN Sales ELSE 0 END) AS "Youth",  
SUM(CASE WHEN age_group="Middle Age" THEN Sales ELSE 0 END) AS "Middle Age",  
SUM(CASE WHEN age_group="Adults" THEN Sales ELSE 0 END) AS "Adults",  
SUM(CASE WHEN age_group="Seniors" THEN Sales ELSE 0 END) AS "Seniors"  
FROM sales_by_age_group  
GROUP BY category;
```

category	Youth	Middle Age	Adults	Seniors
Clothing	26490111	43837924	21809290	21856711
Shoes	15787787	25604974	12629013	12533009
Technology	12721800	23200800	10619700	11320050
Cosmetics	1524393	2662055	1310205	1301332
Toys	957880	1510252	733573	778339
Food & Beverage	193895	333524	162960	160159
Books	201867	315688	158537	158240
Souvenir	149773	246904	119961	120052

INSIGHTS

- ❖ Clothing, Shoes, and Technology emerge as the most preferred product categories across all age groups, reflecting high consumer demand.
- ❖ Middle-aged customers contribute the highest spending overall, showcasing strong purchasing power.
- ❖ Youth (18–30) follow closely behind, displaying consistent spending across categories, especially in Clothing and Technology, indicating high interest and trend sensitivity.



```
-- payment method most commonly used by customers, and how does it vary  
by shopping mall and purchase amount
```

```
-- Payment methods based on number of transactions  
SELECT c.payment_method, COUNT(*) AS Total_payments  
FROM customers AS c  
JOIN sales AS s  
ON c.customer_id=s.customer_id  
GROUP BY c.payment_method  
ORDER BY Total_payments DESC;
```

```
-- payment method used by customers in each shopping mall  
SELECT sm.shopping_mall, c.payment_method, COUNT(*) AS Total_payments  
FROM customers AS c  
JOIN sales AS s  
ON c.customer_id=s.customer_id  
JOIN shopping_mall AS sm  
ON s.shopping_mall=sm.shopping_mall  
GROUP BY sm.shopping_mall, c.payment_method  
ORDER BY sm.shopping_mall, Total_payments DESC;
```

```
-- payment method used by customers by sales category  
SELECT s.sales_category, c.payment_method, COUNT(*) AS Total_payments  
FROM customers AS c  
JOIN sales AS s  
ON c.customer_id=s.customer_id  
JOIN shopping_mall AS sm  
ON s.shopping_mall=sm.shopping_mall  
GROUP BY s.sales_category, c.payment_method  
ORDER BY s.sales_category, Total_payments DESC;
```

payment_method	Total_payments
Cash	44447
Credit Card	34931
Debit Card	20079

shopping_mall	payment_method	Total_payments
Beverly Center	Cash	2183
Beverly Center	Credit Card	1750
Beverly Center	Debit Card	1014
Del Amo Fashion Center	Cash	8894
Del Amo Fashion Center	Credit Card	7019
Del Amo Fashion Center	Debit Card	4030
Fashion Valley	Cash	2324
Fashion Valley	Credit Card	1760
Fashion Valley	Debit Card	991
Glendale Galleria	Cash	2231
Glendale Galleria	Credit Card	1721
Glendale Galleria	Debit Card	962
Irvine Spectrum Center	Cash	2228

sales_category	payment_method	Total_payments
High	Cash	14691
High	Credit Card	11482
High	Debit Card	6624
Low	Cash	20802
Low	Credit Card	16426
Low	Debit Card	9441
Medium	Cash	8954
Medium	Credit Card	7023
Medium	Debit Card	4014

INSIGHTS

- ❖ Cash remains the dominant payment method, leading across all malls and accounting for the majority of transactions.
- ❖ Credit card payments follow as the second most preferred method, particularly gaining traction in medium-value purchases.
- ❖ High-value transactions are predominantly paid in cash, suggesting either customer preference for direct payment.



- How do male and female customers differ in purchasing behavior, including category preferences and average transaction size?

```
WITH demographic_sales AS (  
  SELECT c.gender,s.category, COUNT(*) AS Purchases  
  FROM customers AS c  
  JOIN sales AS s  
  ON c.customer_id=s.customer_id  
  GROUP BY c.gender,s.category  
  ORDER BY c.gender, Purchases DESC)  
  SELECT category,  
  SUM(CASE WHEN gender="Female" THEN Purchases ELSE 0 END) AS "Female",  
  SUM(CASE WHEN gender="Male" THEN Purchases ELSE 0 END) AS "Male"  
  FROM demographic_sales  
  GROUP BY category;  
  
WITH demographic_sales AS (  
  SELECT  
    c.gender,  
    s.category,  
    COUNT(*) AS Purchases  
  FROM customers AS c  
  JOIN sales AS s ON c.customer_id = s.customer_id  
  GROUP BY c.gender, s.category  
)  
SELECT  
  category,  
  ROUND(  
    SUM(CASE WHEN gender = 'Female' THEN Purchases ELSE 0 END) * 100.0 /  
    NULLIF(SUM(Purchases), 0)) AS female_percentage,  
  ROUND(  
    SUM(CASE WHEN gender = 'Male' THEN Purchases ELSE 0 END) * 100.0 /  
    NULLIF(SUM(Purchases), 0)) AS male_percentage  
  FROM demographic_sales  
  GROUP BY category;
```

	category	Female	Male
▶	Clothing	20652	13835
	Cosmetics	9070	6027
	Food & Beverage	8804	5972
	Toys	6085	4002
	Shoes	5967	4067
	Souvenir	3017	1982
	Technology	2981	2015
	Books	2906	2075

category	female_percentage	male_percentage
Clothing	60	40
Food & Beverage	60	40
Shoes	59	41
Souvenir	60	40
Cosmetics	60	40
Technology	60	40
Toys	60	40
Books	58	42

INSIGHTS

❖ **Female customers account for approximately 60% of total sales across all categories, clearly indicating that women are the primary drivers of in-store purchases.**



-- Which malls are the top performers in terms of transaction volume and Sales?

```
WITH Category_wise_sales AS (  
    SELECT sm.shopping_mall,  
           COUNT(*) AS Transactions,  
           SUM(total_sales) AS Sales  
    FROM shopping_mall AS sm  
    JOIN sales AS s  
    ON sm.shopping_mall=s.shopping_mall  
    GROUP BY sm.shopping_mall  
    ORDER BY Transactions DESC  
)  
SELECT *  
FROM Category_wise_sales;
```

shopping_mall	Transactions	Sales
Del Amo Fashion Center	19943	50873460
South Coast Plaza	19823	50555257
Westfield Century City	15011	37303511
Stanford Shopping Center	10161	25380412
Westfield Valley Fair	9781	24619303
Fashion Valley	5075	12901309
Irvine Spectrum Center	4991	12645364
Beverly Center	4947	12304172
Glendale Galleria	4914	12521589
The Grove	4811	12406381

INSIGHTS

- ❖ The top 3 malls were Del Amo Fashion Center, South Coast Plaza, and Westfield Century City, lead both in total sales revenue and transaction volume.
- ❖ These malls demonstrate consistently high customer footfall and purchasing behavior, making them strategic locations for promotions and new product launches.

**THANK
YOU**