Mentorness

MAVEN TOY DATA ANALYSIS

+91 7743927365

GitHub

Table Content

+91 7743927365

GitHub

Introduction

The aim of this SQL project is to analyze the toy sales data to gain insights into inventory management, product performance, sales trends, and store operations. The analysis will be performed using four tables: inventory, products, sales, and stores.

+91 7743927365

GitHub

Dataset Overview

2.1. Inventory Table

- Store_ID: Unique identifier for the store.
- Product_ID: Unique identifier for the product.
- Stock_On_Hand: Number of units available in stock.

2.2. Products Table

- Product_ID: Unique identifier for the product.
- Product_Name: Name of the product.
- Product_Category: Category of the product (e.g., educational, plush, action figures, etc.).
- Product_Cost: Cost price of the product.
- Product_Price: Selling price of the product.

2.3. Sales Table

- Sale_ID: Unique identifier for the sale.
- Date: Date the sale was made.
- Store_ID: Identifier linking to the stores table.
- Product_ID: Identifier linking to the products table.
- Units: Number of units sold.

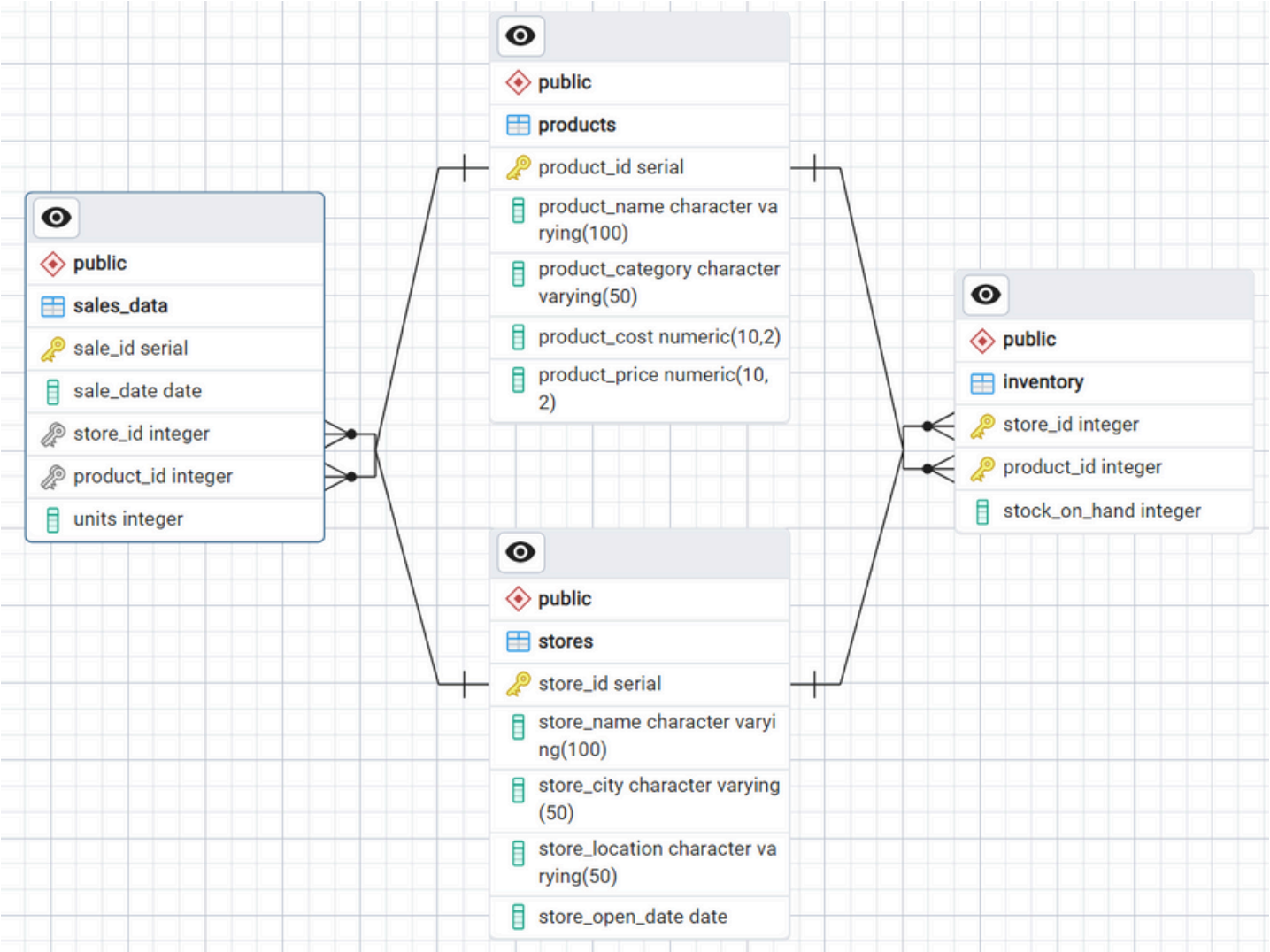
2.4. Stores Table

- Store_ID: Unique identifier for the store.
- Store_Name: Name of the store.
- Store_City: City where the store is located.
- Store_Location: Location details of the store.
- Store_Open_Date: Date when the store was opened.

+91 7743927365

GitHub

Database Schema



+91 7743927365

GitHub

Problem Statement - 01

What is the total sales revenue generated by each store?

```sql
SELECT
    s.Store_ID AS "Store ID",
    s.Store_Name AS "Store Name",
    TO_CHAR(SUM(sd.Units * p.Product_Price),
        '$9,99,999.00') AS "Total Sales Revenue"
FROM
    Sales_Data sd
JOIN
    Stores s ON sd.Store_ID = s.Store_ID
JOIN
    Products p ON sd.Product_ID = p.Product_ID
GROUP BY
    s.Store_ID, s.Store_Name
ORDER BY
    SUM(sd.Units * p.Product_Price) DESC;
```

Store ID integer	Store Name character varying (100)	Total Sales Revenue text
31	Maven Toys Ciudad de Mexic...	$ 5,54,553.43
30	Maven Toys Guadalajara 3	$ 4,49,354.91
9	Maven Toys Ciudad de Mexic...	$ 4,33,556.21
17	Maven Toys Toluca 1	$ 4,11,157.32

Problem Statement - 02

Which products are the top-selling in terms of units sold?

```sql
SELECT
    p.Product_ID AS "Product ID",
    p.Product_Name AS "Product Name",
    concat(SUM(sd.Units), ' Units') AS "Total Units Sold"
FROM
    Sales_Data sd
JOIN
    Products p ON sd.Product_ID = p.Product_ID
GROUP BY
    p.Product_ID, p.Product_Name
ORDER BY
    "Total Units Sold" DESC;
```

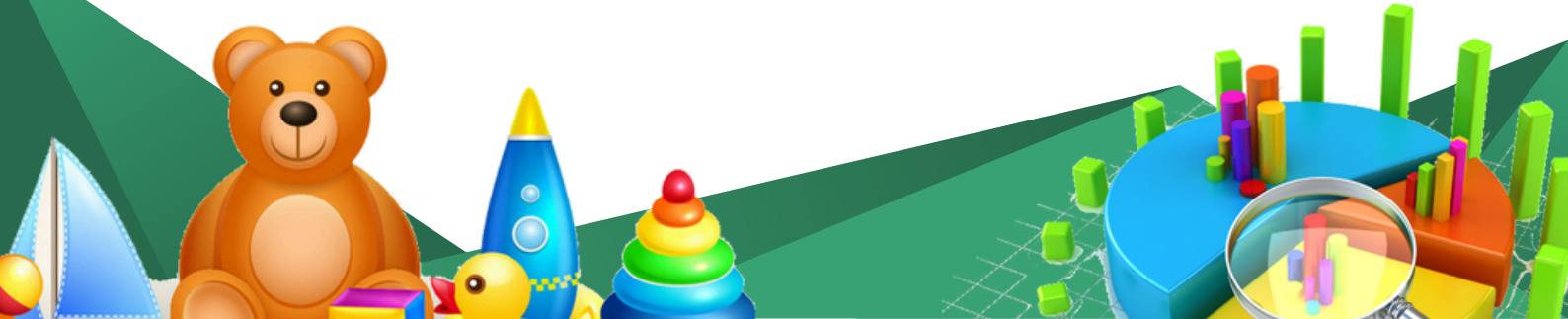
Product ID integer	Product Name character varying (100)	Total Units Sold text
3	Barrel O' Slime	91663 Units
23	Mr. Potatohead	8605 Units
8	Deck Of Cards	84034 Units
12	Foam Disk Launcher	7620 Units
32	Supersoaker Water Gun	6793 Units

Problem Statement - 03

What is the sales performance by product category?

```
SELECT
    p.Product_Category AS "Product Category",
    concat('$ ', SUM(sd.Units * p.Product_Price))
    AS "Total Sales Revenue"
FROM
    Sales_Data sd
JOIN
    Products p ON sd.Product_ID = p.Product_ID
GROUP BY
    p.Product_Category
ORDER BY
    "Total Sales Revenue" DESC;
```

Product Category character varying (50)	Total Sales Revenue text
Toys	$ 5093241.00
Art & Crafts	$ 2705364.26
Electronics	$ 2246771.25
Games	$ 2226836.27
Sports & Outdoors	$ 2172359.57



Problem Statement - 04

What are the current inventory levels for each product at each store?

```sql
SELECT
    s.Store_ID AS "Store ID",
    s.Store_Name AS "Store Name",
    p.Product_ID AS "Product ID",
    p.Product_Name AS "Product Name",
    i.Stock_On_Hand AS "Current Inventory Level"
FROM
    Inventory i
JOIN
    Stores s ON i.Store_ID = s.Store_ID
JOIN
    Products p ON i.Product_ID = p.Product_ID
ORDER BY
    s.Store_ID, p.Product_ID;
```

Store ID integer	Store Name character varying (100)	Product ID integer	Product Name character varying (100)	Current Inventory Level integer
1	Maven Toys Guadalajara 1	1	Action Figure	27
1	Maven Toys Guadalajara 1	2	Animal Figures	0
1	Maven Toys Guadalajara 1	3	Barrel O' Slime	32
1	Maven Toys Guadalajara 1	4	Chutes & Ladders	6
1	Maven Toys Guadalajara 1	5	Classic Dominoes	0

Problem Statement - 05

How do monthly sales trends vary across different stores?

```
SELECT
    s.Store_ID AS "Store ID",
    s.Store_Name AS "Store Name",
    EXTRACT(YEAR FROM sd.Sale_Date) AS "YEAR",
    EXTRACT(MONTH FROM sd.Sale_Date) AS "Month Number",
    TO_CHAR(sd.Sale_Date, 'Month') AS "Month Name",
    SUM(sd.Units * p.Product_Price) AS "Total Sales Revenue"
FROM
    Sales_Data sd
JOIN
    Stores s ON sd.Store_ID = s.Store_ID
JOIN
    Products p ON sd.Product_ID = p.Product_ID
GROUP BY
    s.Store_ID, s.Store_Name,
        EXTRACT(YEAR FROM sd.Sale_Date),
    EXTRACT(MONTH FROM sd.Sale_Date),
        TO_CHAR(sd.Sale_Date, 'Month')
ORDER BY
    s.Store_ID, "YEAR","Month Number";
```

Store ID integer	Store Name character varying (100)	YEAR numeric	Month Number numeric	Month Name text	Total Sales Revenue numeric
1	Maven Toys Guadalajara 1	2017	1	January	8667.37
1	Maven Toys Guadalajara 1	2017	2	February	9063.30
1	Maven Toys Guadalajara 1	2017	3	March	7978.70
1	Maven Toys Guadalajara 1	2017	4	April	13071.32

Problem Statement - 06

Which stores have the highest and lowest sales performance?

```sql
WITH Sales_Performance AS (
    SELECT
        s.Store_ID AS "Store ID",
        s.Store_Name AS "Store Name",
        SUM(sd.Units * p.Product_Price) AS "Total Sales Revenue"
    FROM
        Sales_Data sd
    JOIN
        Stores s ON sd.Store_ID = s.Store_ID
    JOIN
        Products p ON sd.Product_ID = p.Product_ID
    GROUP BY
        s.Store_ID, s.Store_Name)
-- Retrieve highest and lowest performing stores with labels
SELECT
    sp."Store ID",
    sp."Store Name",
    sp."Total Sales Revenue",
    CASE
        WHEN sp."Total Sales Revenue" =
            (SELECT MAX("Total Sales Revenue") FROM Sales_Performance) THEN 'Highest'
        WHEN sp."Total Sales Revenue" =
            (SELECT MIN("Total Sales Revenue") FROM Sales_Performance) THEN 'Lowest'
    END AS "Performance"
FROM
    Sales_Performance sp
WHERE
    sp."Total Sales Revenue" = (SELECT MAX("Total Sales Revenue") FROM Sales_Performance)
    OR sp."Total Sales Revenue" = (SELECT MIN("Total Sales Revenue") FROM Sales Performance);
```

Store ID integer	Store Name character varying (100)	Total Sales Revenue numeric	Performance text
26	Maven Toys Campeche 2	206055.23	Lowest
31	Maven Toys Ciudad de Mexico 2	554553.43	Highest

Problem Statement - 07

What is the profit margin for each product?

```
SELECT
    Product_ID AS "Product ID",
    Product_Name AS "Product Name",
    Product_Cost AS "Cost Price",
    Product_Price AS "Selling Price",
    concat(round(((Product_Price - Product_Cost) /
        Product_Price) * 100,2),' %') AS "Profit Margin (%)"
FROM
    Products;
```

Product ID integer	Product Name character varying (100)	Cost Price numeric (10,2)	Selling Price numeric (10,2)	Profit Margin (%) text
1	Action Figure	9.99	15.99	37.52 %
2	Animal Figures	9.99	12.99	23.09 %
3	Barrel O' Slime	1.99	3.99	50.13 %
4	Chutes & Ladders	1.99	12.99	84.68 %
5	Classic Dominoes	7.99	9.99	20.02 %

Problem Statement - 08

How are sales distributed across different cities?

```sql
SELECT
    s.Store_City AS "City",  -- Store ka city
    concat('$ ', SUM(sd.Units * p.Product_Price)) AS "Total Sales Revenue"
FROM
    Sales_Data sd
JOIN
    Stores s ON sd.Store_ID = s.Store_ID
JOIN
    Products p ON sd.Product_ID = p.Product_ID
GROUP BY
    s.Store_City
ORDER BY
    "Total Sales Revenue" DESC;
```

City character varying (50)	Total Sales Revenue text
Hermosillo	$ 903388.84
Guanajuato	$ 869055.83
Puebla	$ 808710.29
Toluca	$ 633521.68
Xalapa	$ 610119.77

Problem Statement - 09

Which products are out of stock in each store?

```sql
SELECT
    s.Store_Name AS "Store Name",
    p.Product_Name AS "Product Name",
    case when i.stock_on_hand = 0 then 'Out Of Stock'
         else 'Available' end as "Stock Status"
FROM
    Inventory i
JOIN
    Stores s ON i.Store_ID = s.Store_ID
JOIN
    Products p ON i.Product_ID = p.Product_ID
WHERE
    i.Stock_On_Hand = 0;
```

Store Name character varying (100)	Product Name character varying (100)	Stock Status text
Maven Toys Guadalajara 1	Animal Figures	Out Of Stock
Maven Toys Guadalajara 1	Classic Dominoes	Out Of Stock
Maven Toys Guadalajara 1	Toy Robot	Out Of Stock
Maven Toys Monterrey 1	Action Figure	Out Of Stock
Maven Toys Monterrey 1	Dino Egg	Out Of Stock

Problem Statement - 10

How do sales vary by specific dates?

```sql
SELECT
    sd.Sale_Date AS "Sale Date",
    SUM(sd.Units) AS "Total Units Sold",
    SUM(sd.Units * p.Product_Price) AS "Total Sales Revenue"
FROM
    Sales_Data sd
JOIN
    Products p ON sd.Product_ID = p.Product_ID
GROUP BY
    sd.Sale_Date
ORDER BY
    sd.Sale_Date;
```

Sale Date date	Total Units Sold bigint	Total Sales Revenue numeric
2017-01-01	1485	21076.15
2017-01-02	1406	19750.94
2017-01-03	750	11759.50
2017-01-04	960	14814.40
2017-01-05	1370	19791.30

Problem Statement - 11

What is the average cost of products in each category?

```
SELECT
    p.Product_Category AS "Product Category",
    concat('$ ',(round(AVG(p.Product_Cost),2))) AS "Average Cost"
FROM
    Products p
GROUP BY
    p.Product_Category
ORDER BY
    "Average Cost" DESC;
```

Product Category character varying (50)	Average Cost text
Art & Crafts	$ 8.99
Games	$ 7.37
Electronics	$ 14.32
Toys	$ 11.66
Sports & Outdoors	$ 10.28

Problem Statement - 12

What is the sales growth over time for the entire company?

```
SELECT
    TO_CHAR(sd.Sale_Date, 'YYYY-MM') AS "Month",
    SUM(sd.Units * p.Product_Price) AS "Total Sales Revenue",
    LAG(SUM(sd.Units * p.Product_Price)) OVER
        (ORDER BY TO_CHAR(sd.Sale_Date, 'YYYY-MM')) AS "Previous Month Sales",
    concat(round((SUM(sd.Units * p.Product_Price) -
        LAG(SUM(sd.Units * p.Product_Price)) OVER
            (ORDER BY TO_CHAR(sd.Sale_Date, 'YYYY-MM'))) /
                NULLIF(LAG(SUM(sd.Units * p.Product_Price)) OVER
                    (ORDER BY TO_CHAR(sd.Sale_Date, 'YYYY-MM')), 0) * 100,2), ' %')
                        AS "Sales Growth (%)"  -- Sales growth percentage
FROM
    Sales_Data sd
JOIN
    Products p ON sd.Product_ID = p.Product_ID
GROUP BY
    TO_CHAR(sd.Sale_Date, 'YYYY-MM')
ORDER BY
    "Month";
```

Month text	Total Sales Revenue numeric	Previous Month Sales numeric	Sales Growth (%) text
2017-01	542554.91	[null]	%
2017-02	541351.65	542554.91	-0.22 %
2017-03	589485.19	541351.65	8.89 %
2017-04	681072.98	589485.19	15.54 %
2017-05	672369.90	681072.98	-1.28 %

Problem Statement - 13

How does the store open date affect sales performance?

```sql
SELECT
    s.Store_Open_Date AS "Store Open Date",
    s.Store_Name AS "Store Name",
    EXTRACT(YEAR FROM s.Store_Open_Date) AS "Open Year",
    SUM(sd.Units * p.Product_Price) AS "Total Sales Revenue"
FROM
    Sales_Data sd
JOIN
    Stores s ON sd.Store_ID = s.Store_ID
JOIN
    Products p ON sd.Product_ID = p.Product_ID
GROUP BY
    s.Store_Open_Date, s.Store_Name
ORDER BY
    "Open Year" ASC, "Total Sales Revenue" DESC;
```

Store Open Date date	Store Name character varying (100)	Open Year numeric	Total Sales Revenue numeric
1992-09-18	Maven Toys Guadalajara 1	1992	261842.89
1995-04-27	Maven Toys Monterrey 1	1995	277959.14
1999-12-27	Maven Toys Guadalajara 2	1999	262435.02
2000-01-01	Maven Toys Saltillo 1	2000	330408.90
2001-05-31	Maven Toys La Paz 1	2001	210897.83

Problem Statement - 14

What percentage of total sales does each store contribute?

```sql
SELECT
    s.Store_Name AS "Store Name",
    SUM(sd.Units * p.Product_Price) AS "Store Sales",
    ROUND((SUM(sd.Units * p.Product_Price) /
          (SELECT SUM(sd2.Units * p2.Product_Price)
           FROM Sales_Data sd2
           JOIN Products p2 ON sd2.Product_ID = p2.Product_ID))
                  * 100, 2) AS "Sales Percentage (%)"
FROM
    Sales_Data sd
JOIN
    Stores s ON sd.Store_ID = s.Store_ID
JOIN
    Products p ON sd.Product_ID = p.Product_ID
GROUP BY
    s.Store_Name
ORDER BY
    "Sales Percentage (%)" DESC;
```

Store Name character varying (100)	Store Sales numeric	Sales Percentage (%) numeric
Maven Toys Ciudad de Mexic...	554553.43	3.84
Maven Toys Guadalajara 3	449354.91	3.11
Maven Toys Ciudad de Mexic...	433556.21	3.00
Maven Toys Toluca 1	411157.32	2.85
Maven Toys Monterrey 2	372998.82	2.58

Problem Statement - 15

How do sales compare to current stock levels for each product?

```sql
SELECT
    p.Product_Name AS "Product Name",
    SUM(sd.Units) AS "Total Units Sold",
    i.Stock_On_Hand AS "Current Stock Level",
    (i.Stock_On_Hand - SUM(sd.Units)) AS "Stock After Sales"
FROM
    Sales_Data sd
JOIN
    Products p ON sd.Product_ID = p.Product_ID
JOIN
    Inventory i ON p.Product_ID = i.Product_ID
GROUP BY
    p.Product_Name, i.Stock_On_Hand
ORDER BY
    "Stock After Sales" ASC;
```

Product Name character varying (100)	Total Units Sold bigint	Current Stock Level integer	Stock After Sales bigint
Colorbuds	313104	8	-313096
Colorbuds	313104	11	-313093
Colorbuds	313104	17	-313087
Colorbuds	313104	18	-313086
Colorbuds	313104	23	-313081

Conclusion

This analysis provides insights into key business areas, including top-performing stores and products, inventory management, and sales trends. The SQL queries allowed us to understand how different stores and product categories performed and how these factors contribute to the overall success of the business. By continuing to track these metrics, we can drive better business decisions and optimize operations across all store locations.

+91 7743927365

GitHub

Thank You!

Stay Connected:

Name: Bhushan Gawali

Email: bhushanjg14@gmail.com

LinkedIn: linkedin.com/in/Bhushangawali

GitHub: https://github.com/Bhushan148

+91 7743927365

GitHub