



# US Candy : Sales Insights

Presented by: Samantha Beltran.

# US Candy Sales Analysis

---

## About

This SQL project explores sales data from a U.S. national candy distributor to gain insights into top-performing product divisions, sales trends, and customer behavior. By analyzing this data, the goal is to identify opportunities to improve and optimize sales strategies. The dataset includes sales and geospatial information related to factory-to-customer shipments, encompassing customer and factory locations, sales orders and goals, and detailed product information.

Data Source: [Maven Analytics](#).

## Objective

The main objective of this project is to analyze sales performance, profitability, and customer trends using SQL. This involves answering key business questions that support data-driven decision-making for sales optimization and operational improvements.

## Data

The dataset, titled **"US Candy Distributor"**, was obtained from Maven Analytics and consists of five tables with a total of **10,194 records** and **39 fields**. The data was imported into SQL Server for analysis. As the dataset was already clean and well-structured, no additional preprocessing or cleaning was necessary.

### Table Candy\_Sales

Column	Description	Data Type
Row ID	Unique row identifier	SMALLINT
Order ID	Unique order identifier	NVARCHAR(50)
Order Date	Date of order	DATE
Ship Date	Date of shipment	DATE
Ship Mode	Shipping method of order	NVARCHAR(50)
Customer ID	Unique customer identifier	INT
Country/Region	Country or region of customer	NVARCHAR(50)
City	City of customer	NVARCHAR(50)
State/Province	State/province of customer	NVARCHAR(50)
Postal Code	Postal code / zip code of customer	NVARCHAR(50)
Division	Product division	NVARCHAR(50)
Region	Region of customer	NVARCHAR(50)
Product ID	Unique product identifier	NVARCHAR(50)
Product Name	Product long name	NVARCHAR(50)

Sales	Total sales value of order	FLOAT
Units	Total units of order	TINYINT
Gross Profit	Gross profit of order ( Sales - Cost )	FLOAT
Cost	Cost to manufacture	FLOAT

**Table Candy\_Factories**

Column	Description	Data Type
Factory	Factory Name	NVARCHAR(50)
Latitude	Latitude Coordinates	FLOAT
Longitude	Longitude Coordinates	FLOAT

**Table Candy\_Products**

Column	Description	Data Type
Division	Product Division	NVARCHAR(50)
Product Name	Product Descriptive Name	NVARCHAR(50)
Factory	Factory Name	NVARCHAR(50)
Product ID	Product Unique Identifier	NVARCHAR(50)
Unit Price	Selling Price of Product	FLOAT
Unit Cost	Cost to Produce Product	FLOAT

**Table Candy\_Targets**

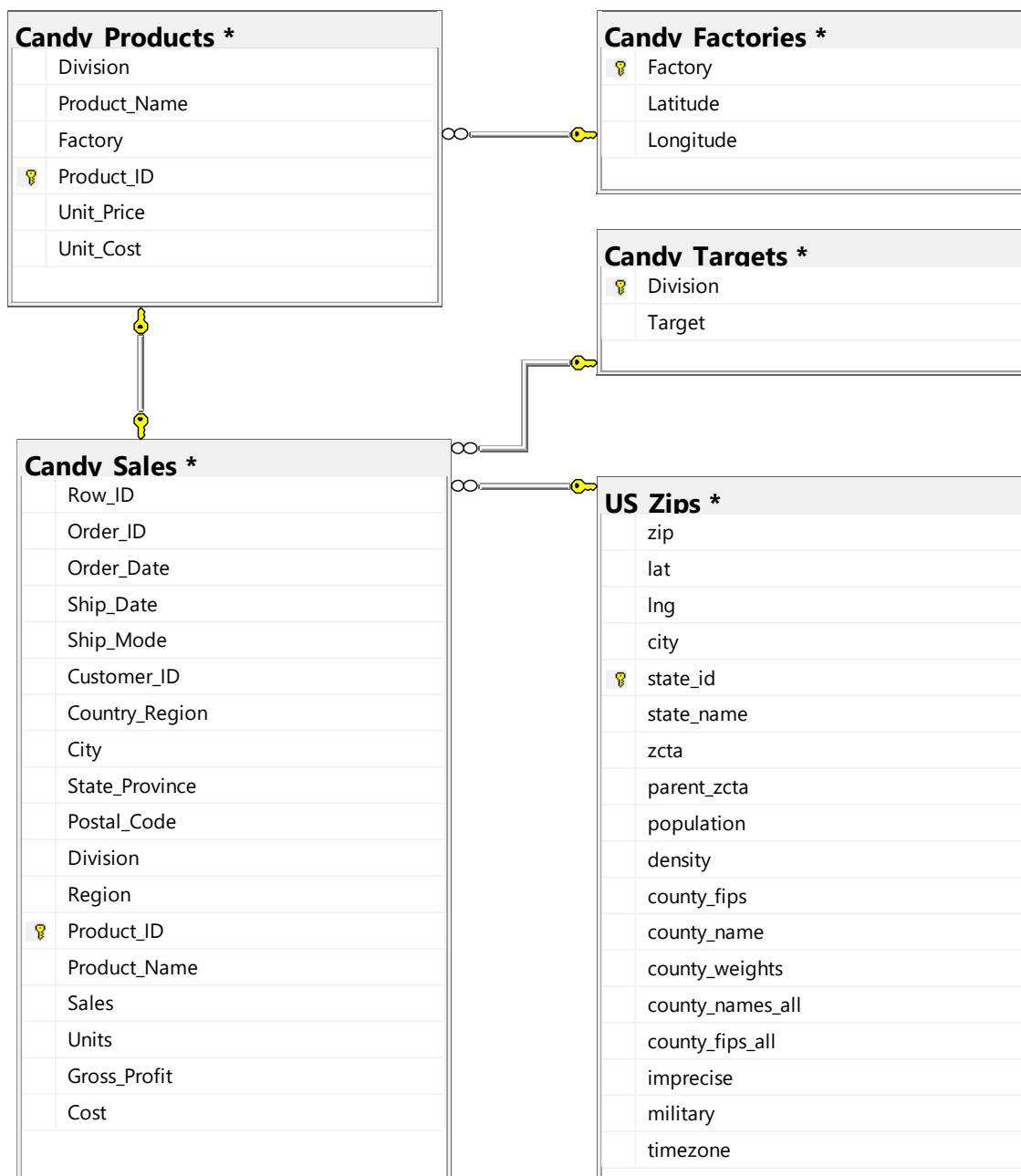
Column	Description	Data Type
Division	Product Division	NVARCHAR(50)
Target	Sales Target 2024	SMALLINT

**Table Candy\_US\_Zips**

Column	Description	Data Type
zip	The 5-digit zip code assigned by the U.S. Postal Service.	INT
lat	The latitude of the zip code (learn more).	FLOAT
lng	The longitude of the zip code (learn more).	FLOAT
city	The official USPS city name.	NVARCHAR(50)
state_id	The official USPS state abbreviation.	NVARCHAR(50)
state_name	The state's name.	NVARCHAR(50)
zcta	TRUE if the zip code is a Zip Code Tabulation area (learn more).	NVARCHAR(50)
parent_zcta	The ZCTA that contains this zip code. Only exists if zcta is FALSE. Useful for making inferences about a zip codes that is a point from the ZCTA that contains it.	NVARCHAR(1)
population	An estimate of the zip code's population. Only exists if zcta is TRUE.	INT
density	The estimated population per square kilometer. Only exists if zcta is TRUE.	FLOAT
county_fips	The zip's primary county in the FIPS format.	INT

county_name	The name of the county_fips.	NVARCHAR (50)
county_weights	A JSON dictionary listing all county_fips and their weights (by area) associated with the zip code.	NVARCHAR (100)
imprecise	TRUE if the lat/Ing has been geolocated using the city (rare).	NVARCHAR (50)
military	TRUE if the zip code is used by the US Military (lat/Ing not available).	NVARCHAR (50)
timezone	The city's time zone in the tz database format. (e.g. America/Los_Angeles)	NVARCHAR (50)

## ER Diagram



# Contents

---

<b>US Candy Project</b>	<b>1</b>
About	1
Objective	1
Data	1
<b>Overview</b>	<b>5</b>
a. Annual Sales by Division (2021-2024)	5
b. Annual Profit Margin by Division (2021-2024)	5
c. Total Orders by Division (2021-2024)	6
<b>Business Questions:</b>	<b>7</b>
<b>Part I: Sales Performance And Profitability</b>	<b>7</b>
1. What are the total sales, orders, cost, gross profit, and profit margin by Country for 2024?	7
2. What are the quarterly sales trends for 2024?	8
3. What is the month-over-month sales growth for 2024?	8
4. Which product divisions are generating the highest and lowest profits?	10
5. What are the monthly sales trends per product division? (E.g. Chocolate)	10
6. Which products generate the highest profit margin?	11
7. What is the most popular shipping method?	15
8. Which customers have placed the most orders and contributed the highest revenue?	15
<b>Part II: Product And Factory Insights</b>	<b>16</b>
1. What is the average profit margin for each product division and name?	16
2. What is the most profitable product, and which factory produces it?	17
3. Which products contribute to 80% of the total profit?	18
4. Which factories produce the highest profitable products?	18
5. Which divisions are underperforming in sales relative to their 2024 targets?	19
<b>Part III: Geographical And Demographic Analysis</b>	<b>20</b>
1. What is the distribution of sales by country and division?	20
2. Which provinces/states in Canada and USA generate the highest profitability for candy sales in 2024, and how does their profit margin percentage compare across these regions?	21
3. What is the distribution of customers by state/province, and which states generate the most revenue?	22
4. How does the population density of a region (US Zips) correlate with the number of orders?	24
<b>Insights &amp; Recommendations</b>	<b>26</b>

# Overview

## a. Annual Sales by Division (2021-2024)

This query retrieves the total sales for each division across the years 2021 to 2024, providing a breakdown of annual sales performance for easy comparison.

```
WITH SourceTable AS (  
    SELECT  
        Division,  
        YEAR(Order_Date) AS Sales_Year,  
        CAST(SUM(Sales) AS NUMERIC(18, 2)) AS Total_Sales  
    FROM Candy_Sales  
    WHERE YEAR(Order_Date) BETWEEN 2021 AND 2024  
    GROUP BY Division, YEAR(Order_Date)  
)  
SELECT  
    Division,  
    ISNULL([2021], 0) AS Sales_2021,  
    ISNULL([2022], 0) AS Sales_2022,  
    ISNULL([2023], 0) AS Sales_2023,  
    ISNULL([2024], 0) AS Sales_2024  
FROM SourceTable  
PIVOT (  
    SUM(Total_Sales)  
    FOR Sales_Year IN ([2021], [2022], [2023], [2024])  
) AS PivotTable;
```

Annual Sales Performance by Division: 2021-2024

	Division	Sales_2021	Sales_2022	Sales_2023	Sales_2024
1	Chocolate	26356.36	27763.20	34193.38	43379.96
2	Other	2549.25	1207.75	2461.75	3444.50
3	Sugar	47.94	31.70	204.19	143.65

## b. Annual Profit Margin by Division (2021-2024)

This query calculates the profit margin percentage for each division from 2021 to 2024. It then pivots the results to display the annual profit margin percentages side by side for easy comparison across years.

```
WITH SourceTable AS (  
    SELECT  
        Division,
```

```

        YEAR(Order_Date) AS PM_Year,
        CAST((SUM(Gross_Profit) / NULLIF(SUM(Sales), 0)) * 100 AS NUMERIC(18, 2)) AS
Profit_Margin_Percentage
FROM Candy_Sales
WHERE YEAR(Order_Date) BETWEEN 2021 AND 2024
GROUP BY Division, YEAR(Order_Date)
)
SELECT
    Division,
    CAST(ISNULL([2021], 0) AS NUMERIC(18, 2)) AS PM_2021,
    CAST(ISNULL([2022], 0) AS NUMERIC(18, 2)) AS PM_2022,
    CAST(ISNULL([2023], 0) AS NUMERIC(18, 2)) AS PM_2023,
    CAST(ISNULL([2024], 0) AS NUMERIC(18, 2)) AS PM_2024
FROM SourceTable
PIVOT (
    AVG(Profit_Margin_Percentage)
    FOR PM_Year IN ([2021], [2022], [2023], [2024])
) AS PivotTable;

```

Annual Profit Margin by Division: 2021-2024

	Division	PM_2021	PM_2022	PM_2023	PM_2024
1	Chocolate	67.38	67.40	67.53	67.46
2	Other	45.33	39.71	44.84	46.29
3	Sugar	56.82	56.31	71.60	65.05

## c. Total Orders by Division (2021-2024)

This query retrieves the number of unique orders placed in each division for the years 2021 to 2024. The result provides an annual breakdown of order volumes by division, allowing for year-over-year comparisons of order activity.

```

WITH SourceTable AS (
    SELECT
        Division,
        YEAR(Order_Date) AS Sales_Year,
        COUNT(DISTINCT Order_ID) AS Order_Count
    FROM Candy_Sales
    WHERE YEAR(Order_Date) BETWEEN 2021 AND 2024
    GROUP BY Division, YEAR(Order_Date)
)
SELECT
    Division,
    ISNULL([2021], 0) AS Orders_2021,
    ISNULL([2022], 0) AS Orders_2022,
    ISNULL([2023], 0) AS Orders_2023,
    ISNULL([2024], 0) AS Orders_2024
FROM SourceTable
PIVOT (

```



```

SUM(Order_Count)
FOR Sales_Year IN ([2021], [2022], [2023], [2024])
) AS PivotTable;

```

#### Total Orders by Division from 2021-2024

	Division	Orders_2021	Orders_2022	Orders_2023	Orders_2024
1	Chocolate	1635	1708	2138	2724
2	Other	77	50	72	105
3	Sugar	6	6	13	15

The Chocolate Division has shown steady year-over-year growth, with a notable increase in 2024, solidifying its position as the primary revenue driver and highlighting strong market demand. In contrast, the Other Division has exhibited inconsistent performance, marked by a decline in 2022 followed by moderate recovery, suggesting potential volatility or challenges that may warrant further analysis. Meanwhile, the Sugar Division consistently reports low sales, with only a slight peak in 2023, which could reflect low market interest or a reduced strategic focus. Overall, the Chocolate Division continues to lead revenue generation, while the Other and Sugar divisions present areas for potential improvement.

After reviewing the sales trends from 2021 to 2024, the focus will now shift specifically to the performance in 2024.

## Business Questions:

### Part I: Sales Performance And Profitability

#### 1. What are the total sales, orders, cost, gross profit, and profit margin by Country for 2024?

```

SELECT
    Country_Region,
    CAST(SUM(Sales) AS NUMERIC(18, 2)) AS Total_Sales,
    COUNT(DISTINCT Order_ID) AS Total_Orders,
    CAST(SUM(Cost) AS NUMERIC(18, 2)) AS Total_Costs,
    CAST(SUM(Gross_Profit) AS NUMERIC(18, 2)) AS Total_Gross_Profit,
    CAST(SUM(Gross_Profit) / NULLIF(SUM(Sales), 0) * 100 AS NUMERIC(18, 2)) AS
Profit_Margin_Percentage
FROM Candy_Sales
WHERE YEAR(Order_Date) = 2024
GROUP BY Country_Region

```



```
ORDER BY Profit_Margin_Percentage DESC;
```

#### Sales by Country - 2024

	Country_Region	Total_Sales	Total_Orders	Total_Costs	Total_Gross_Profit	Profit_Margin_Percentage
1	United States	45959.35	2790	15665.76	30293.59	65.91
2	Canada	1008.76	54	351.56	657.20	65.15

## 2. What are the quarterly sales trends for 2024?

This query calculates the total sales for each quarter of 2024 from the Candy\_Sales table. It extracts the quarter from Order\_Date, filters records to include only those from 2024, sums the Sales for each quarter, formats the totals to two decimal places, and sorts the results by quarter.

```
SELECT
    DATEPART(QUARTER, Order_Date) AS Quarter,
    CAST(SUM(Sales) AS NUMERIC(18, 2)) AS Total_Sales
FROM Candy_Sales
WHERE YEAR(Order_Date) = 2024
GROUP BY DATEPART(QUARTER, Order_Date)
ORDER BY Quarter;
```

#### Quarterly Sales Summary for 2024

	Quarter	Total_Sales
1	1	6943.32
2	2	9416.15
3	3	12643.28
4	4	17965.36

The results show a steady increase in sales throughout 2024, peaking in Q4 at 17,965.36, likely due to seasonal demand (holiday-driven sales). The consistent growth from Q1 to Q3 indicates a positive sales trajectory, with opportunities to analyze Q4 drivers for replication in other quarters.

## 3. What is the month-over-month sales growth for 2024?

This query displays total sales for each month in 2024, along with the previous month's sales and the percentage growth. It helps identify trends in sales performance and measure month-over-month growth

```

WITH Monthly_Sales AS (
    SELECT
        MONTH(Order_Date) AS Month,
        CAST(SUM(Sales) AS NUMERIC(18, 2)) AS Total_Sales
    FROM Candy_Sales
    WHERE YEAR(Order_Date) = 2024
    GROUP BY MONTH(Order_Date)
)
SELECT
    Month,
    Total_Sales,
    LAG(Total_Sales) OVER (ORDER BY Month) AS Previous_Month_Sales,
    CAST(((Total_Sales - LAG(Total_Sales) OVER (ORDER BY Month))
    / NULLIF(LAG(Total_Sales) OVER (ORDER BY Month), 0)) * 100 AS NUMERIC(18, 2)) AS
Growth_Percentage
FROM Monthly_Sales;

```

#### Monthly Sales Growth for 2024

	Month	Total_Sales	Previous_Month_Sales	Growth_Percentage
1	1	2169.42	NULL	NULL
2	2	1495.34	2169.42	-31.07
3	3	3278.56	1495.34	119.25
4	4	2663.00	3278.56	-18.78
5	5	3216.05	2663.00	20.77
6	6	3537.10	3216.05	9.98
7	7	2991.39	3537.10	-15.43
8	8	3208.37	2991.39	7.25
9	9	6443.52	3208.37	100.83
10	10	4541.81	6443.52	-29.51
11	11	6747.64	4541.81	48.57
12	12	6675.91	6747.64	-1.06

The sales data shows significant fluctuations throughout the year, with sharp declines after periods of high growth, such as in February, April, October, and December. Notable growth spikes occurred in March (119.25%) and September (100.83%), likely due to seasonal demand, while mid-year months (May to August) exhibited more stable, moderate growth. These patterns suggest that while there are seasonal opportunities for growth, maintaining consistent sales performance throughout the year may require addressing the factors contributing to the sharp declines following peak months.

#### 4. Which product divisions are generating the highest and lowest profits?

```
SELECT
    Division,
    CAST(SUM(Sales) AS NUMERIC(18, 2)) AS Total_Sales,
    SUM(Units) AS Total_Units_Sold,
    CAST(SUM(Gross_Profit) AS NUMERIC(18, 2)) AS Total_Gross_Profit,
    CAST(SUM(Gross_Profit) / NULLIF(SUM(Sales), 0) * 100 AS NUMERIC(18, 2)) AS
Profit_Margin_Percentage
FROM Candy_Sales
WHERE YEAR(Order_Date) = 2024
GROUP BY Division
ORDER BY Profit_Margin_Percentage DESC;
```

Sales and Profit Analysis by Division for 2024

	Division	Total_Sales	Total_Units_Sold	Total_Gross_Profit	Profit_Margin_Percentage
1	Chocolate	43379.96	12274	29262.84	67.46
2	Sugar	143.65	51	93.45	65.05
3	Other	3444.50	412	1594.50	46.29

The results show that the Chocolate division generates the highest total sales and profit margin at 67.46%, followed by Sugar with a 65.05% margin. Despite its high profit margin, Sugar shows very low sales and units sold, indicating limited market interest or focus. The Other division, while contributing significantly to total sales and units sold, has a lower profit margin of 46.29%. This indicates that while Chocolate leads in both revenue and profitability, the Sugar and Other divisions still provide valuable contributions, though there is room for improvement in profitability.

#### 5. What are the monthly sales trends per product division? (Eg. Chocolate)

```
WITH Monthly_Sales AS (
    SELECT
        Division,
        MONTH(Order_Date) AS Month,
        CAST(SUM(Sales) AS NUMERIC(18, 2)) AS Total_Sales
    FROM Candy_Sales
    WHERE YEAR(Order_Date) = 2024
    GROUP BY Division, MONTH(Order_Date)
)
```

```

SELECT
    Division,
    Month,
    Total_Sales,
    LAG(Total_Sales) OVER (ORDER BY Month) AS Previous_Month_Sales,
    CAST(((Total_Sales - LAG(Total_Sales) OVER (ORDER BY Month))
        / NULLIF(LAG(Total_Sales) OVER (ORDER BY Month), 0)) * 100 AS NUMERIC(18, 2)) AS
Growth_Percentage
FROM Monthly_Sales
WHERE Division LIKE 'Chocolate' --Change it as needed
ORDER BY Month;

```

Monthly Sales Growth Per Division- 2024

	Division	Month	Total_Sales	Previous_Month_Sales	Growth_Percentage
1	Chocolate	1	2148.68	NULL	NULL
2	Chocolate	2	1205.34	2148.68	-43.90
3	Chocolate	3	3074.56	1205.34	155.08
4	Chocolate	4	2499.50	3074.56	-18.70
5	Chocolate	5	3122.05	2499.50	24.91
6	Chocolate	6	3205.60	3122.05	2.68
7	Chocolate	7	2935.39	3205.60	-8.43
8	Chocolate	8	3113.37	2935.39	6.06
9	Chocolate	9	5664.79	3113.37	81.95
10	Chocolate	10	3925.06	5664.79	-30.71
11	Chocolate	11	6189.21	3925.06	57.68
12	Chocolate	12	6296.41	6189.21	1.73

## 6. Which products generate the highest profit margin?

This query calculates the profit margin percentage for each product in the Candy\_Sales table for the year 2024. The results are ordered by the highest profit margin to identify the most profitable products.

```

SELECT
    Product_Name, -- Let's check the profit margin of all products
    CAST((SUM(Gross_Profit) / NULLIF(SUM(Sales), 0)) * 100 AS NUMERIC(18, 2)) AS
Profit_Margin_Percentage
FROM Candy_Sales
WHERE YEAR(Order_Date) = 2024
GROUP BY Product_Name
ORDER BY Profit_Margin_Percentage DESC;

```

### Profit Margin by Product Name in 2024

Results Messages		
	Product_Name	Profit_Margin_Percentage
1	Everlasting Gobstopper	80.00
2	Hair Toffee	77.78
3	Wonka Bar - Nutty Crunch Surprise	71.35
4	Wonka Bar - Scrumdiddlyumptious	69.44
5	Wonka Bar - Fudge Mallows	66.67
6	Wonka Bar - Triple Dazzle Caramel	65.33
7	Wonka Bar - Milk Chocolate	64.92
8	Laffy Taffy	62.31
9	Fizzy Lifting Drinks	60.00
10	Wonka Gum	52.00
11	Lickable Wallpaper	50.00
12	SweeTARTS	46.67
13	Nerds	46.67
14	Kazookles	7.69

Next Step: Once the most profitable products are identified, it's valuable to determine which factories produce them. The following query lists the most profitable product along with its respective factory.

```
SELECT
    TOP 1 WITH TIES
    s.Product_Name, -- Let's check the most profitable product and its factory.
    p.Factory,
    CAST((SUM(s.Gross_Profit) / NULLIF(SUM(s.Sales), 0)) * 100 AS NUMERIC(18, 2)) AS
Profit_Margin_Percentage
FROM Candy_Sales AS s
JOIN Candy_Products AS p ON s.Product_ID = p.Product_ID
WHERE YEAR(s.Order_Date) = 2024
GROUP BY
    s.Product_Name,
    p.Factory
ORDER BY Profit_Margin_Percentage DESC;
```

### Most Profitable Product in 2024 and Its Factory

Results Messages			
	Product_Name	Factory	Profit_Margin_Percentage
1	Everlasting Gobstopper	Secret Factory	80.00

Further Step: To gain a deeper understanding, we can perform a Pareto analysis to identify the products that contribute to 80% of the total gross profit. The following query will help in determining these key products.

```
WITH Product_Profit AS (
    SELECT
        Product_Name,
        CAST(SUM(Gross_Profit) AS NUMERIC(18, 2)) AS Total_Gross_Profit
    FROM Candy_Sales
```

```

WHERE YEAR(Order_Date) = 2024
GROUP BY Product_Name
),

Cumulative_Profit AS (
    SELECT
        Product_Name,
        Total_Gross_Profit,
        SUM(Total_Gross_Profit) OVER (ORDER BY Total_Gross_Profit DESC) AS Cumulative_Profit,
        SUM(Total_Gross_Profit) OVER () AS Total_Profit
    FROM Product_Profit
)
SELECT
    Product_Name,
    Total_Gross_Profit,
    CAST((Cumulative_Profit / Total_Profit) * 100 AS NUMERIC(18, 2)) AS
Cumulative_Profit_Percentage
FROM Cumulative_Profit
WHERE
    CAST((Cumulative_Profit / Total_Profit) * 100 AS NUMERIC(18, 2)) <= 80
ORDER BY Cumulative_Profit_Percentage;

```

#### 2024 Pareto Analysis: Gross Profit Contribution by Product

	Product_Name	Total_Gross_Profit	Cumulative_Profit_Percentage
1	Wonka Bar -Scrumdiddlyumptious	6462.50	20.88
2	Wonka Bar - Triple Dazzle Caramel	6227.90	41.00
3	Wonka Bar - Milk Chocolate	5682.23	59.36
4	Wonka Bar - Nutty Crunch Surprise	5550.21	77.29

Final Step: To further analyze profitability, we could cross-check the cost-to-revenue ratio by product. In theory, products with lower cost-to-revenue ratios are likely the ones that generate higher profitability for the company. This analysis will provide additional insights into the financial efficiency of each product.

```

SELECT
    Product_Name, -- Let's check cost-to-price ratio
    CAST(Unit_Price AS NUMERIC(18, 2)) AS Unit_Price,
    CAST(Unit_Cost AS NUMERIC(18, 2)) AS Unit_Cost,
    CAST((Unit_Cost * 100.0) / NULLIF(Unit_Price, 0) AS NUMERIC(18, 2)) AS Cost_To_Price_Ratio
FROM Candy_Products
ORDER BY Cost_To_Price_Ratio ASC;

```

Cost-to-Ratio by Product 2024

	Product_Name	Unit_Price	Unit_Cost	Cost_To_Price_Ratio
1	Everlasting Gobstopper	10.00	2.00	20.00
2	Hair Toffee	4.50	1.00	22.22
3	Wonka Bar - Nutty Crunch Surprise	3.49	1.00	28.65
4	Wonka Bar - Scrumdiddlyumptious	3.60	1.10	30.56
5	Wonka Bar - Fudge Mallows	3.60	1.20	33.33
6	Wonka Bar - Triple Dazzle Caramel	3.75	1.30	34.67
7	Wonka Bar - Milk Chocolate	3.25	1.14	35.08
8	Laffy Taffy	1.99	0.75	37.69
9	Fizzy Lifting Drinks	3.75	1.50	40.00
10	Wonka Gum	1.25	0.60	48.00
11	Lickable Wallpaper	20.00	10.00	50.00
12	SweeTARTS	1.50	0.80	53.33
13	Nerds	1.50	0.80	53.33
14	Fun Dip	1.50	0.90	60.00
15	Kazookles	3.25	3.00	92.31

Focusing on the profit margin percentage, Everlasting Gobstopper and Hair Toffee stand out as the candies with the highest profit margins. This aligns with the cost-to-price ratio analysis, where these products also have the lowest cost-to-price ratios. In terms of gross profit, the products contributing 80% of the total profit are the Wonka Bar candies, highlighting their significant role in overall profitability.

```
WITH Order_Counts AS (
    SELECT
        Ship_Mode,
        COUNT(DISTINCT Order_ID) AS Total_Orders,
        CAST(SUM(Gross_Profit) AS NUMERIC(18, 2)) AS Total_Gross_Profit
    FROM Candy_Sales
    WHERE YEAR(Order_Date) = 2024
    GROUP BY Ship_Mode
)
SELECT
    Ship_Mode,
    CAST((Total_Orders * 100.0) / SUM(Total_Orders) OVER () AS NUMERIC(5, 2)) AS
    Order_Percentage,
    Total_Gross_Profit
FROM Order_Counts
ORDER BY Total_Orders DESC;
```



## 7. What is the most popular shipping method?

```
WITH Order_Counts AS (  
    SELECT  
        Ship_Mode,  
        COUNT(DISTINCT Order_ID) AS Total_Orders,  
        CAST(SUM(Gross_Profit) AS NUMERIC(18, 2)) AS Total_Gross_Profit  
    FROM Candy_Sales  
    WHERE YEAR(Order_Date) = 2024  
    GROUP BY Ship_Mode  
)  
SELECT  
    Ship_Mode,  
    CAST((Total_Orders * 100.0) / SUM(Total_Orders) OVER () AS NUMERIC(5, 2)) AS  
    Order_Percentage,  
    Total_Gross_Profit  
FROM Order_Counts  
ORDER BY Total_Orders DESC;
```

Shipping Methods Impact on Sales and Profit - 2024

	Ship_Mode	Order_Percentage	Total_Gross_Profit
1	Standard Class	58.51	17735.15
2	Second Class	19.13	6338.06
3	First Class	17.09	5256.12
4	Same Day	5.27	1621.46

## 8. Which customers have placed the most orders and contributed the highest revenue?

```
WITH Orders_by_Customer AS (  
    SELECT  
        Customer_ID,  
        COUNT(DISTINCT Order_ID) AS Total_Orders,  
        CAST(SUM(Sales) AS NUMERIC(18, 2)) AS Total_Sales  
    FROM Candy_Sales  
    WHERE YEAR(Order_Date) = 2024  
    GROUP BY Customer_ID  
)  
SELECT  
    Customer_ID,  
    Total_Orders,  
    Total_Sales,  
    DENSE_RANK() OVER (ORDER BY Total_Orders DESC) AS Order_Rank  
FROM Orders_by_Customer  
ORDER BY Order_Rank;
```

### Top Customers and Order Volume - 2024

Results		Messages		
	Customer_ID	Total_Orders	Total_Sales	Order_Rank
1	131807	10	294.00	1
2	163790	6	208.95	2
3	166695	6	72.30	2
4	164756	5	304.14	3
5	153948	5	49.24	3
6	157966	5	126.22	3
7	161956	5	93.14	3
8	135937	5	160.34	3
9	100111	5	182.41	3
10	102925	5	79.46	3
11	104220	5	70.90	3
12	118087	5	75.04	3
13	119284	5	107.40	3
14	121293	5	60.46	3
15	122035	4	79.40	4
16	119438	4	43.32	4
17	118772	4	26.28	4

## Part II: Product And Factory Insights

### 1. What is the average profit margin for each product division and name?

```
SELECT
    Division,
    Product_Name,
    CAST(AVG(NULLIF(Gross_Profit / NULLIF(Sales, 0), 0)) * 100 AS NUMERIC(18, 2)) AS
    Avg_Profit_Margin
FROM Candy_Sales
WHERE YEAR(Order_Date) = 2024
GROUP BY Division, Product_Name
ORDER BY Division, Avg_Profit_Margin DESC;
```

### Product Profitability Performance for 2024

Results Messages			
	Division	Product_Name	Avg_Profit_Margin
1	Chocolate	Wonka Bar - Nutty Crunch Surprise	71.35
2	Chocolate	Wonka Bar - Scrumdiddlyumptious	69.44
3	Chocolate	Wonka Bar - Fudge Mallows	66.67
4	Chocolate	Wonka Bar - Triple Dazzle Caramel	65.33
5	Chocolate	Wonka Bar - Milk Chocolate	64.92
6	Other	Wonka Gum	52.00
7	Other	Lickable Wallpaper	50.00
8	Other	Kazookles	7.69
9	Sugar	Everlasting Gobstopper	80.00
10	Sugar	Hair Toffee	77.78
11	Sugar	Laffy Taffy	62.31
12	Sugar	Fizzy Lifting Drinks	60.00
13	Sugar	SweeTARTS	46.67
14	Sugar	Nerds	46.67

The Sugar division offers the highest average profit margins, with Everlasting Gobstopper at 80% and Hair Toffee at 77.78%. Meanwhile, the Chocolate division maintains strong and consistent profitability, with Nutty Crunch Surprise at 71.35%. However, certain "Other" category products, such as Kazookles, have a significantly low 7.69% margin. This suggests that while some products in this category, like Wonka Gum (52%), are lucrative, others may need a pricing review or volume-based sales strategy to remain viable.

## 2. What is the most profitable product, and which factory produces it?

```
SELECT
  TOP 1 WITH TIES
    s.Product_Name,
    p.Factory,
    CAST((SUM(s.Gross_Profit) / NULLIF(SUM(s.Sales), 0)) * 100 AS NUMERIC(18, 2)) AS
Profit_Margin_Percentage
FROM Candy_Sales AS s JOIN Candy_Products AS p ON s.Product_ID = p.Product_ID
WHERE YEAR(s.Order_Date) = 2024
GROUP BY s.Product_Name, p.Factory
ORDER BY Profit_Margin_Percentage DESC;
```

### Most Profitable Product and its Factory (2024)

Results Messages			
	Product_Name	Factory	Profit_Margin_Percentage
1	Everlasting Gobstopper	Secret Factory	80.00

The most profitable product in 2024 is Everlasting Gobstopper, with a profit margin of 80%, supplied by the Secret Factory.

### 3. Which products contribute to 80% of the total profit?

```
WITH Product_Profit AS (  
    SELECT  
        Product_Name,  
        CAST(SUM(Gross_Profit) AS NUMERIC(18, 2)) AS Total_Gross_Profit  
    FROM Candy_Sales  
    WHERE YEAR(Order_Date) = 2024  
    GROUP BY Product_Name  
) ,  
Cumulative_Profit AS (  
    SELECT  
        Product_Name,  
        Total_Gross_Profit,  
        SUM(Total_Gross_Profit) OVER (ORDER BY Total_Gross_Profit DESC) AS Cumulative_Profit,  
        SUM(Total_Gross_Profit) OVER () AS Total_Profit  
    FROM Product_Profit  
)  
SELECT  
    Product_Name,  
    Total_Gross_Profit,  
    CAST((Cumulative_Profit / Total_Profit) * 100 AS NUMERIC(18, 2)) AS  
Cumulative_Profit_Percentage  
FROM Cumulative_Profit  
WHERE CAST((Cumulative_Profit / Total_Profit) * 100 AS NUMERIC(18, 2)) <= 80  
ORDER BY Cumulative_Profit_Percentage;
```

Products Contributing to 80% of Total Profit (2024)

	Product_Name	Total_Gross_Profit	Cumulative_Profit_Percentage
1	Wonka Bar - Scrumdiddlyumptious	6462.50	20.88
2	Wonka Bar - Triple Dazzle Caramel	6227.90	41.00
3	Wonka Bar - Milk Chocolate	5682.23	59.36
4	Wonka Bar - Nutty Crunch Surprise	5550.21	77.29

### 4. Which factories produce the highest profitable products?

```
SELECT  
    p.Factory,  
    s.Product_Name,  
    CAST((SUM(s.Gross_Profit) / NULLIF(SUM(s.Sales), 0)) * 100 AS NUMERIC(18, 2)) AS  
Profit_Margin_Percentage  
FROM Candy_Sales AS s JOIN Candy_Products AS p ON s.Product_ID = p.Product_ID  
WHERE YEAR(s.Order_Date) = 2024  
GROUP BY p.Factory, s.Product_Name  
ORDER BY Factory, Profit_Margin_Percentage DESC;
```

```

SELECT
    p.Factory, -- Let's check profitability by factory
    CAST((SUM(s.Gross_Profit) / NULLIF(SUM(s.Sales), 0)) * 100 AS NUMERIC(18, 2)) AS
Profit_Margin_Percentage
FROM Candy_Sales AS s JOIN Candy_Products AS p ON s.Product_ID = p.Product_ID
WHERE YEAR(s.Order_Date) = 2024
GROUP BY p.Factory
ORDER BY Profit_Margin_Percentage DESC;

```

#### Factory-Level Revenue Performance (2024)

	Factory	Product_Name	Profit_Margin_Percentage
1	Lot's O' Nuts	Wonka Bar - Nutty Crunch Surprise	71.35
2	Lot's O' Nuts	Wonka Bar -Scrumdiddlyumptious	69.44
3	Lot's O' Nuts	Wonka Bar - Fudge Mallows	66.67
4	Secret Factory	Everlasting Gobstopper	80.00
5	Secret Factory	Wonka Gum	52.00
6	Secret Factory	Lickable Wallpaper	50.00
7	Sugar Shack	Laffy Taffy	62.31
8	Sugar Shack	Fizzy Lifting Drinks	60.00
9	Sugar Shack	SweeTARTS	46.67
10	Sugar Shack	Nerds	46.67
11	The Other Factory	Hair Toffee	77.78
12	The Other Factory	Kazookles	7.69
13	Wicked Choccy's	Wonka Bar - Triple Dazzle Caramel	65.33
14	Wicked Choccy's	Wonka Bar - Milk Chocolate	64.92

	Factory	Profit_Margin_Percentage
1	Lot's O' Nuts	69.15
2	Wicked Choccy's	65.14
3	Sugar Shack	54.72
4	Secret Factory	50.42
5	The Other Factory	14.12

## 5. Which divisions are underperforming in sales relative to their 2024 targets?

```

WITH Division_Sales AS (
    SELECT
        Division,
        CAST(SUM(Sales) AS NUMERIC(18, 2)) AS Total_Sales
    FROM Candy_Sales
    WHERE YEAR(Order_Date) = 2024
    GROUP BY Division

```

```

)
SELECT
    ds.Division, ds.Total_Sales, ct.Target,
    (ds.Total_Sales - ct.Target) AS Difference,
    CASE
        WHEN ds.Total_Sales >= ct.Target THEN 'Outperformed'
        ELSE 'Underperformed'
    END AS Performance
FROM Division_Sales AS ds JOIN Candy_Targets AS ct ON ds.Division = ct.Division;

```

Division Performance Against Targets (2024)

	Division	Total_Sales	Target	Difference	Performance
1	Chocolate	43379.96	27000	16379.96	Outperformed
2	Sugar	143.65	15000	-14856.35	Underperformed
3	Other	3444.50	3000	444.50	Outperformed

## Part III: Geographical And Demographic Analysis

### 1. What is the distribution of sales by country and division?

```

SELECT
    COALESCE(Country_Region, 'All Contries') AS Country,
    COALESCE(Division, 'All Divisions') AS Division,
    CAST(SUM(Sales) AS NUMERIC(18, 2)) AS Total_Sales
FROM Candy_Sales
WHERE YEAR(Order_Date) = 2024
GROUP BY CUBE (Country_Region, Division);

```

Total Sales by Country and Division (2024) with Subtotals

	Country	Division	Total_Sales
1	Canada	Chocolate	888.76
2	United States	Chocolate	42491.20
3	All Contries	Chocolate	43379.96
4	Canada	Other	120.00
5	United States	Other	3324.50
6	All Contries	Other	3444.50
7	United States	Sugar	143.65
8	All Contries	Sugar	143.65
9	All Contries	All Divisions	46968.11
10	Canada	All Divisions	1008.76
11	United States	All Divisions	45959.35

The United States leads in overall sales, contributing most of the total revenue (approximately 98%), with the Chocolate division being the dominant performer across all countries and regions. This highlights the strong market presence of chocolate products in the U.S. and potentially suggests a high consumer demand for these products. While Canada's sales are minimal in comparison, with Canada accounting for only 2% of the total, the Other category's low sales in Canada could be indicative of a potential untapped market or limited product penetration. The underperformance of Sugar products across all regions, especially in the U.S., calls for an evaluation of the product's appeal and its marketing strategies. Targeting specific regions in Canada with potential for growth could be a good strategy for increasing sales in the Other and Sugar categories.

## 2. Which provinces/states in Canada and USA generate the highest profitability for candy sales in 2024, and how does their profit margin percentage compare across these regions?

```
SELECT
    State_Province, --Let's check the most profitable province in Canada
    CAST(SUM(Sales) AS NUMERIC(18, 2)) AS Total_Sales,
    CAST((SUM(Gross_Profit) / NULLIF(SUM(Sales), 0)) * 100 AS NUMERIC(18, 2)) AS
Profit_Margin_Percentage
FROM Candy_Sales
WHERE YEAR(Order_Date) = 2024 AND Country_Region = 'Canada'
GROUP BY Country_Region, State_Province
ORDER BY Profit_Margin_Percentage DESC;

SELECT
    State_Province, --Let's check the most profitable state in USA
    CAST(SUM(Sales) AS NUMERIC(18, 2)) AS Total_Sales,
    CAST((SUM(Gross_Profit) / NULLIF(SUM(Sales), 0)) * 100 AS NUMERIC(18, 2)) AS
Profit_Margin_Percentage
FROM Candy_Sales
WHERE YEAR(Order_Date) = 2024 AND Country_Region = 'United States'
GROUP BY Country_Region, State_Province
ORDER BY Profit_Margin_Percentage DESC;
```

Most Profitable Provinces in Canada (2024)

Results		Messages	
	State_Province	Total_Sales	Profit_Margin_Percentage
1	Newfoundland and Labrador	57.60	69.44
2	New Brunswick	43.18	68.74
3	British Columbia	100.74	68.24
4	Prince Edward Island	65.05	68.15
5	Ontario	240.75	66.70
6	Manitoba	74.90	65.34
7	Nova Scotia	27.50	65.09
8	Saskatchewan	22.75	64.92
9	Quebec	202.29	63.90
10	Alberta	174.00	59.20



### Most Profitable State in USA (2024)

100 % ▾			
Results Messages			
	State_Province	Total_Sales	Profit_Margin_Percentage
1	Vermont	55.14	69.42
2	Mississippi	211.34	68.91
3	South Dakota	115.40	68.80
4	West Virginia	63.97	68.61
5	Missouri	393.19	68.05
6	Connecticut	359.88	67.84
7	Maryland	551.13	67.83
8	Michigan	899.57	67.59
9	Nebraska	111.91	67.56
10	North Dakota	109.66	67.44
11	Iowa	179.98	67.44
12	Minnesota	487.81	67.40
13	Oklahoma	224.97	67.38
14	Indiana	786.82	67.26
15	Louisiana	171.75	67.25

In Canada, the most profitable provinces are Newfoundland and Labrador and New Brunswick, which both achieve high profit margins (over 68%). These regions, despite their smaller total sales, reflect efficient cost-to-revenue ratios, suggesting that targeted regional strategies could yield high returns in these areas. However, Ontario and Quebec, with larger sales volumes, show a slight decline in margin percentages, indicating that higher sales may come at the expense of profitability, possibly due to higher operational costs or discount practices. In the United States, similar patterns are observed, with states like California and Texas showing high sales, but smaller states with higher margins may warrant further investment in niche markets for greater profitability. Exploring ways to improve profitability in high-revenue states could involve reducing overheads or optimizing pricing strategies.

## 3. What is the distribution of customers by state/province, and which states generate the most revenue?

```
SELECT
    COALESCE(State_Province, 'All States') AS State,
    COALESCE(City, 'All Cities') AS City,
    COUNT(DISTINCT(Customer_ID)) AS Total_Customers,
    CAST(SUM(Sales) AS NUMERIC(18, 2)) AS Total_Sales
FROM Candy_Sales
WHERE YEAR(Order_Date) = 2024 AND Country_Region LIKE 'Canada'
GROUP BY ROLLUP (State_Province, City)
ORDER BY State_Province;
```

```
SELECT
    COALESCE(State_Province, 'All States') AS State,
    COALESCE(City, 'All Cities') AS City,
    COUNT(DISTINCT(Customer_ID)) AS Total_Customers,
    CAST(SUM(Sales) AS NUMERIC(18, 2)) AS Total_Sales
```

```
FROM Candy_Sales
WHERE YEAR(Order_Date) = 2024 AND Country_Region LIKE 'United States'
GROUP BY ROLLUP (State_Province, City)
ORDER BY State_Province;
```

Customer and Sales Summary by Province and City in Canada (2024)

Results Messages				
	State	City	Total_Customers	Total_Sales
1	All States	All Cities	36	1008.76
2	Alberta	All Cities	2	174.00
3	Alberta	Calgary	2	126.90
4	Alberta	Edmonton	1	47.10
5	British Columbia	All Cities	5	100.74
6	British Columbia	Vancouver	5	100.74
7	Manitoba	All Cities	2	74.90
8	Manitoba	Winnipeg	2	74.90
9	New Brunswick	All Cities	2	43.18
10	New Brunswick	Moncton	2	43.18
11	Newfoundland and Labrador	All Cities	2	57.60
12	Newfoundland and Labrador	St John's	2	57.60
13	Nova Scotia	All Cities	2	27.50
14	Nova Scotia	Halifax	2	27.50

Customer and Sales Summary by State and City in USA (2024)

Results Messages				
	State	City	Total_Customers	Total_Sales
1	All States	All Cities	1687	45959.35
2	Alabama	All Cities	11	236.63
3	Alabama	Auburn	1	7.20
4	Alabama	Decatur	2	40.05
5	Alabama	Florence	1	6.50
6	Alabama	Hoover	2	54.00
7	Alabama	Huntsville	2	56.48
8	Alabama	Montgomery	1	40.00
9	Alabama	Tuscaloosa	2	32.40
10	Arizona	All Cities	32	997.04
11	Arizona	Chandler	4	46.90
12	Arizona	Gilbert	1	160.34

The Ontario province emerges as the top performer in Canada in terms of total sales, accounting for approximately 24% of total sales revenue. Similarly, in the United States, California, and Texas dominate, with both states contributing significantly to total sales. This indicates that larger populations and more populous states are making most sales. Notably, regions like Quebec and Alberta show strong revenue per customer, despite smaller customer bases, which could suggest that these areas might have higher-value customers or more targeted, high-demand markets. Strategic marketing or promotional campaigns in areas with smaller but profitable customer bases (such as Quebec) could help to boost growth in those provinces and states.

#### 4. How does the population density of a region (US Zips) correlate with the number of orders?

```
WITH City_Sales AS (  
    SELECT  
        UPPER(City) AS City,  
        COUNT(DISTINCT Customer_ID) AS Total_Customers,  
        COUNT(DISTINCT Order_ID) AS Total_Orders,  
        CAST(SUM(Sales) AS NUMERIC(18, 2)) AS Total_Sales  
    FROM Candy_Sales  
    WHERE YEAR(Order_Date) = 2024 AND Country_Region = 'United States'  
    GROUP BY City  
)  
City_Population AS (  
    SELECT  
        UPPER(City) AS City,  
        SUM(Population) AS Total_Population  
    FROM US_Zips  
    GROUP BY City  
)  
SELECT  
    CS.City,  
    CS.Total_Customers,  
    CS.Total_Orders,  
    CS.Total_Sales,  
    COALESCE(CP.Total_Population, 0) AS Total_Population  
FROM City_Sales CS  
LEFT JOIN City_Population CP  
    ON CS.City = CP.City  
ORDER BY Total_Population DESC;
```

Sales, Customers, and Orders by U.S. City with Population Data (2024)

Results		Messages			
	City	Total_Customers	Total_Orders	Total_Sales	Total_Population
1	HOUSTON	50	92	1350.14	3236469
2	CHICAGO	67	103	1641.76	2705664
3	LOS ANGELES	123	179	2948.68	2406666
4	MIAMI	12	13	144.27	1877027
5	SAN ANTONIO	6	8	134.02	1849089
6	LAS VEGAS	2	4	67.85	1629464
7	PHILADELPHIA	91	151	2384.14	1624777
8	DALLAS	37	55	861.61	1536971
9	PHOENIX	8	19	247.68	1466272
10	SAN DIEGO	19	30	473.29	1338191
11	COLUMBUS	40	68	1351.17	1252113
12	AUSTIN	8	20	233.08	1198430
13	JACKSONVILLE	23	40	492.05	1155514
14	DENVER	6	12	280.30	1147657

Cities with higher population densities like Los Angeles, Houston, and Chicago show significantly higher total sales, which supports the hypothesis that larger populations drive more consumer activity. However, interestingly, cities like San Francisco generate high sales despite fewer customers, suggesting that factors like affluent populations, niche market demand, or higher product value per order may contribute to this pattern. This calls for further investigation into consumer behaviors, local economic conditions, and market segmentation to identify whether certain cities are more inclined to purchase premium or specialized products. Exploring the relationship between population density and revenue generation, possibly by calculating a correlation coefficient, would offer more concrete insights into how demographics influence purchasing behaviors. Additionally, examining external factors like local economic conditions or consumer preferences in each city could provide a clearer understanding of sales dynamics.



## Insights & Recommendations

---

The 2024 sales performance data reveals a strong market position for the Chocolate Division, consistent quarterly sales growth, and strategic opportunities in regional and product-specific areas. While high-margin products such as Everlasting Gobstopper and Nutty Crunch Surprise drive profitability, underperforming products and divisions like Sugar and certain "Other" items present improvement opportunities.

## I. Sales Performance And Profitability Insights

- **Strongest Division:** The Chocolate Division is the primary revenue driver, showing steady growth and high profit margins (67.46%).
- **Underperformance:** The Sugar Division underperformed significantly, generating only \$143.65 in 2024, a -99% shortfall from its \$15,000 sales target.
- **Fluctuating Sales:** Sales show notable monthly volatility with sharp declines post-peak months (e.g., February, April, October, December).
- **Top Products:** Products contributing 80% of total profit are all chocolate-based: Scrumdiddlyumptious, Triple Dazzle Caramel, Milk Chocolate, and Nutty Crunch Surprise.
- **Top Supplier:** Lot's O' Nuts provides high-margin chocolate products (69.15%), while The Other Factory underperforms with low-margin items (e.g., Kazookles at 7.69%).

## II. Product And Factory Insights

- **Most Profitable Product:** Everlasting Gobstopper (80% profit margin), produced by Secret Factory.
- **Product Concentration Risk:** Heavy reliance on chocolate-based products for profitability presents a risk if market trends shift.
- **Supplier Performance:** Strong suppliers include Lot's O' Nuts, Wicked Choccy's, and Sugar Shack. Underperformers, like The Other Factory, need review.

Recommendations:

1. **Focus on High-Margin Items:** Expand production and inventory for high-profit products like Everlasting Gobstopper and Nutty Crunch Surprise.
2. **Reevaluate Low-Margin Products:** Discontinue or renegotiate Kazookles and other underperforming SKUs.
3. **Strengthen Key Supplier Relationships:** Explore long-term agreements with top suppliers to secure pricing and supply consistency.

## II. Geographical And Demographic Insights

- **Top Countries:** USA contributes 98% of total revenue, driven by strong chocolate product performance.
- **Canadian Opportunities:** Low sales suggest untapped potential. Newfoundland and Labrador, New Brunswick show high profit margins (>68%).
- **Top Regions:** Ontario (Canada), California and Texas (USA) lead in sales. High population cities like Los Angeles and Houston correlate with high sales.

- **Hidden Gems:** Cities like San Francisco show high sales with fewer customers, indicating strong per-order value.

Recommendations:

1. **Target High-Margin Regions:** Boost marketing and distribution in Newfoundland, New Brunswick, and profitable U.S. states.
2. **Invest in Niche Markets:** Leverage high-value markets like San Francisco and Quebec with premium or specialty offerings.
3. **Analyze Regional Behaviors:** Perform detailed correlation analysis between population density and sales to identify expansion opportunities.

### III. Strategic Recommendations

- **Diversify Product Portfolio:** Reduce overreliance on Chocolate by innovating high-margin items in Sugar and Other divisions.
- **Optimize Inventory:** Use sales forecasts and past seasonal data to stock appropriately, avoiding overstock during low-demand periods.
- **Marketing and Promotion:** Bundle low-performing Sugar items with top-selling chocolates to drive upsell opportunities.
- **Refine Pricing Strategy:** Review pricing and cost structures for low-margin products and high-cost regions.
- **Replicate Q4 Success:** Identify factors that led to Q4 spike and apply similar strategies during other quarters.

The data-driven insights point to a highly profitable core business driven by chocolate products and certain suppliers. However, unlocking further growth will require diversification, strategic regional focus, and a sharper approach to low-margin products. Implementation of these recommendations will better position the company to sustain and expand its profitability in future years.