**SALES ANALYSIS PROJECT DOCUMENTATION**

**Workflow Framework**

1. Understand the business case
2. Measurement planning
3. Collection and prep of data
4. Understanding the scope of the data
5. Analysis and visualization
6. Development of data-driven insights
7. Measure, test, and optimize

**Understanding the Business Case**

Questions I asked to understand the goals and objectives of the business:

1. What are the high-level goals of Sasha's Retail Store?
2. Who are the stakeholders for this project?
3. How does my analysis fit into the overall business?
4. What business goals am I trying to impact?
5. What does a successful completion of this project mean to you?
6. What metrics will drive the impact of this project?
7. What are your pain points concerning this project that will be most impactful if answered?
8. Concerning the performance of April-July, are there any known external factors that impacted this, such as the emergence of competition, store shutdowns, low inventory, etc.?

**Measurement Planning**

Has its framework as follows:

1. Know the business
2. Know your audience/stakeholders
3. Define the KPIs
4. Identify data sources

**Know the business:**

1. What are the high-level goals of Sasha’s Retail Store?

Increase overall revenue, improve profitability, identify key products and markets that drive growth, build stronger customer relationships, and increase brand visibility.

1. How does my analysis fit into the overall business?

Enables the business to understand sales trends, identify underperforming areas, and determine where to focus our efforts. This will allow us to make data-driven decisions about marketing, product strategy, and regional focus to align with our revenue and profitability goals.

1. What business goals am I trying to impact?

* Reduce losses in underperforming months.
* Validate or challenge assumptions about high-revenue cities.
* Allocate marketing budgets efficiently.
* Develop actionable strategies to improve monthly and annual sales performance.

**Know your audience:**

1. Who are the stakeholders in this project?

Owner of Sasha’s Retail Store (Leadership Team), Accountant, Assistant Manager, Marketing Team

1. What does a successful completion of this project mean to you?

A successful project would provide:

* Clear and accurate sales performance metrics for 2019.
* Insights into why specific months (April-July) were unprofitable.
* Data-backed recommendations on whether to focus marketing efforts on the suggested cities or explore other high-potential areas.

1. What metrics will drive the impact of this project?

* Monthly revenue comparisons (total and by region/city).
* Revenue trends by city and product category.
* Contribution to total revenue from the suggested cities.
* Marketing return on investment (if historical data is available).

1. What are your pain points concerning this project that will be most impactful if answered?

* Understanding why April-July performed poorly compared to other months.
* Determining whether the assistant manager’s suggestion to focus on specific cities is justified.
* Ensuring that any recommendations for marketing focus are data-driven and address both short-term and long-term business goals.

**Define the KPIs**

|  |  |
| --- | --- |
| **Business Goal** | **KPI Definition(s)** |
| Overall revenue growth | * Total revenue * Monthly sales performance/trend * MoM revenue change * Sales volume |
| Improve profitability | * Total profit * Profit growth |
| Identify key cities that drive growth | * Revenue by city * Sales volume by city * Average Order Value by city |
| Identify key products that drive growth | * Top selling products * Revenue by product * Sales volume by product |

**Identify data sources:**

2019 Sales backup file owned by *Sasha’s Retail Store*

**Collection and prep of data**

Restored the database via a 2019 Sales backup file from *Sasha’s Retail Store* on Microsoft SQL Server.

Import the data into Power BI and utilize Power Query for data transformation

**Quality Assurance (QA):**

|  |  |
| --- | --- |
| **Issue** | **Solution** |
| Inconsistent Data Types | Ensured that all columns had the same data types across all tables (January to December). If "order ID" was a whole number in one table, it was made a whole number in all tables and changed "order ID" in February from a date data type to a whole number to match other months. |
| Null Values | Removed rows where "quantity ordered" was null, as these represented invalid entries. Applied "Remove Empty" to delete rows with null values in the "quantity ordered" column. |
| Incorrect Data Types | Changed "order ID" from decimal to a whole number. Changed "quantity ordered" from decimal to a whole number. Changed "price each" to a fixed decimal number with two decimal places. Changed the data type of "order date" from datetime to date. |
| Incorrect Date Format | Split the "order date" column into month and year columns. Extracted the last two characters from the year column to create a day column. Removed the original "order date" column and created a new "order date" column using the month, day, and a fixed year of 2019. Ensured the month and day columns were whole numbers before creating the "order date" column. |
| Extra Spaces | Trimmed the "purchase address" column to remove extra spaces at the beginning or end of the string. |
| Combined Address | Split the "purchase address" column into multiple columns: street address, city, state, and ZIP code. Used a comma as the delimiter to split the address into three columns. Trimmed the resulting columns before further splitting. Split the "purchase address 3" column by space to separate the state and ZIP code. |
| Column Names | Ensured that all tables had the same column names. |
| Disabled Tables | Disabled the original January to December tables from loading into the report view to avoid confusion and ensure only the combined/cleaned table ("Sales 2019") is used. Grouped hidden queries into a new group called "hidden queries" in the Power Query Editor for better organization. |

**Limitations and Assumptions:**

**Data Limitations:**

No data on product costs is available; hence, calculations for profit metrics cannot be made.

**Business Assumptions:**

The stakeholders understand the limitations and determine that the analysis should proceed with the available data.

**Revised KPI Definitions**

|  |  |
| --- | --- |
| **Business Goal** | **KPI Definition(s)** |
| Overall revenue growth | * Total revenue * Monthly sales performance/trend * MoM revenue change * Sales volume |
| Identify key cities that drive growth | * Revenue by city * Sales volume by city * Average Order Value by city |
| Identify key products that drive growth | * Top selling products * Revenue by product * Sales volume by product |

**Data Analysis and Visualizations:**

**Key Sales Metrics**

The following metrics were generated to assess sales performance:

* Total Revenue
* Total Orders
* Total Products Sold
* Average Order Value
* Month-over-Month Sales Change Percentage

**Data Preparation**

**Measures Table**

* A dedicated measures table was created to store all the calculated measures.

**Date Table**

* A date table was created to facilitate time-based analysis.
* This involved:
  + Creating a custom date table in Power BI using DAX
  + Adding columns for **year, month, day, and month name**
  + Marking the table as the primary date table
  + Establishing a relationship between the date table and the **Sales 2019** table

**Calculations**

**Month-over-Month Sales Change**

* The **CALCULATE** function was used with the **PREVIOUSMONTH** function to compute total sales for the previous month.
* Variables were employed to store:
  + Total revenue for the current month
  + Total revenue for the previous month
* The **DIVIDE** function was used to calculate the **Month-over-Month Sales Change Percentage**.

**Latest Date Extraction**

* The **LASTDATE** function was applied to determine the most recent date in the dataset.

**Dynamic Textboxes for Monthly Insights**

* Dynamic textboxes were created to display the **top three months** and **bottom three months**.
* Steps involved:
  + Creating measures for **total revenue by month**
  + Identifying the **top three** and **bottom three months**
  + Using the **CONCATENATEX** function to join month names with commas

**Visualizations**

**Key Performance Visuals**

* **Card visuals** were used to display key sales metrics.

**Monthly Sales Performance**

* A **line chart** was used to visualize sales performance for each month.

**Sales Across Cities**

* A **clustered bar chart** was created to show total sales per city.

**Top and Bottom Products**

* **Table visuals** were used to display the highest and lowest-performing products.
* **Slicers** were added to allow filtering the number of displayed products.

**Conditional Formatting**

* Applied to **table visuals** to highlight top and bottom-performing products.

**Interactive Filtering**

* **Slicers for city and month** were included for dynamic data filtering.

**Key Considerations**

**Accessibility and Readability**

* **Color Blindness**: Implemented colorblind-friendly palettes in visualizations.
* **Commenting Measures**: Added comments in measures for better understanding and future reference.
* **Measure Organization**: Grouped measures into folders for easier management.

**Uncovering Insights and Data-driven Recommendations**

**Key Insights**

Key Sales Metrics:

* **Total Revenue:** $32.7 million
* **Average Order Value (AOV):** $186.00
* **Number of Orders:** 176,011
* **Number of Products Sold:** 19
* **Latest MoM Sales Change:** 44.19%
* **Average Monthly Revenue:** $2.7 million

Accountants Report on Losses:

**Monthly Sales Performance**

* **Top Performing Months:**
  + December: $4.6 million
  + October: $3.7 million
  + April: $3.4 million
* **Lowest Performing Months:**
  + February: $374.7K
  + January: $1.8 million
  + September: $2.1 million

**Marketing Attention:**

* **Top Performing Cities:**
  + San Francisco, Los Angeles, and New York City contributed 53.4% of total orders and generated $17.4 million.
* **Bottom-Performing Cities:**
  + Austin, Portland, and Seattle contributed 19.5% of total orders with $6.5 million in revenue.

**Top and Bottom Performing Products:**

**Top Selling Products:**

These **﻿5﻿** products made up **﻿66%﻿** of the total sales and a revenue of **﻿$21,509,412.08﻿**

* MacBook Pro Laptop: $7641500
* iPhone: $4547200
* ThinkPad Laptop: $3892961.07
* Google Phone: $3127200
* 27in 4K Gaming Monitor: $2300551

**Bottom Selling Products:**

These **﻿5﻿** products made up **﻿3%﻿** of the total sales and a revenue of **﻿$1,021,039.7﻿**

* AAA Batteries (4 pack): $87798.36
* AA Batteries (4 pack): $100727.04
* Wired Headphones: $233625.15
* USB-C Charging Cables: $270231.40
* Lightning Charging Cable: $328675.75

**Data-Driven Recommendations:**

1. Conduct a thorough investigation into the factors impacting sales during low-performing months. Implement targeted strategies to mitigate these fluctuations, such as promotional campaigns or product enhancements.
2. Optimize marketing and sales strategies for each city, tailoring approaches based on local market conditions. Consider reallocating resources to capitalize on high-performing markets while identifying growth opportunities in underperforming regions.
3. Prioritize inventory management and marketing efforts around top-selling products to maximize revenue. Re-evaluate the profitability and strategic fit of bottom-selling products, considering options such as discontinuation or rebranding.