

Tunis Business School

IT 300

**Business Intelligence Analysis for Superstore Sales Data (2021-2024)**

PREPARED BY : EYA ARFAOUI

MAJOR/MINOR : FINANCE/IT

1. INTRODUCTION:

The purpose of this project is to analyze four years (2021-2024) of sales data from Vertex superstore data warehouse to provide actionable insights for strategic decision-making. The analysis aims to support the manager in evaluating the business's performance and determining areas for improvement or expansion. Key objectives include identifying top-performing regions, customers, and products, comparing sales metrics over time, and assessing profitability to make informed investment and operational decisions. This study will serve as a foundation for potential actions such as:

* Opening new stores in promising regions.
* Closing or transferring underperforming stores.
* Expanding the product catalog or discontinuing non-profitable items.
* Reorganizing marketing strategies to target high-performing demographics.
* Planning territorial extensions into new regions.

The deliverable includes a unified Business Intelligence (BI) dashboard to visualize KPIs and support decision-making.

1. PHASE 1: Data Gathering
2. Source Systems : data was collected from 2 main sources

* A CSV file containing fields: Order ID, Order Date, Ship Date, Ship Mode, Customer Name, Segment, State, Country, Market, Region, Product ID, Category, Sub-Category, Product Name, Sales, Quantity, Discount, Profit, Shipping Cost, Order Priority, Year.
* A JSON file containing fields: Row ID, Order ID, Order Date, Ship Date, Ship Mode, Customer ID, Customer Name, Segment, City, State, Country, Postal Code, Market, Region, Product ID, Category, Sub-Category, Product Name, Sales, Quantity, Discount, Profit, Shipping Cost, Order Priority, Year.

1. Initial Data Review :

* A preliminary review of the data was conducted to identify key attributes relevant to sales, profitability, customer demographics, and geographic regions.
* Data inconsistencies, such as missing values or duplicate records, were noted for cleaning in the next phase.

1. PHASE 2:Data Preparation (ETL Process)
2. Extraction: Extracted data from disparate sources, including:

* Historical sales records (2021-2024).
* Customer profiles and purchase history.
* Regional and seasonal sales trends.

1. Transformation: Structured and cleaned the data:

* Handled missing or erroneous entries such as : ( Row ID, Customer ID, City, Postal Code)
* Standardized formats for dates, currencies, and product categories.

1. Loading:

* Loaded the transformed data into a centralized data warehouse (DWH) for further analysis.

1. PHASE 3 :Data Storage and Modeling
2. **Data Warehouse Schema:** Implemented a star schema with the following components:

* **Fact Table:** Sales data including: sales amount, profit, quantity sold, discount, and shipping costs.
* **Dimension Tables:** Customers, Products and Time.

A diagram of a product

Description automatically generated

1. **OLAP Process:**

* Adopted ROLAP (Relational Online Analytical Processing) for efficient querying and reporting.

1. PHASE 4 :Data Analysis
2. **Metrics and KPIs:**

* **Sales Revenue:** Annualy, quarterly, and monthly.
* **Top 10 Countries:** By sales volume.
* **Top 10 Customers:** Based on purchase frequency and revenue contribution.
* **Top 5 Most Profitable Products:** Highest profit.
* **Yearly Comparisons:** Highlight the most profitable year from 2021-2024
* **Regional Performance:** Analyze profitability by region, focusing on the most profitable region in each country for 2024 to help in future investment decisions.
* **Top 5 Most Profitable Product Categories:** Identify categories with the highest profit to guide strategic decisions, such as:( Increasing the supply of high-margin products, prioritizing marketing efforts for profitable categories, allocating shelf space or storage based on profitability trends).
* **Most Profitable Segments :** To help in future marketing strategies.
* **Discount Analysis: Identify the effect of discount on sales.**

1. **Visualization Tools:**

* Created a BI dashboard using Power BI to present insights through visualizations, including:

1. Bar charts for profits per product.
2. Line graphs for profitability trend and discount effect.
3. Pie charts for customer segmentation and product categories profitability.
4. Heatmaps for countries’ performance.
5. CONCLUSION :

* The profitability **trend** is **increasing** reaching more than **1 million** dollars in 2024 which indicates a **good financial position coming from operations.**
* Most of the store’s revenue is coming from the USA, Australia and France so the company must **allocate more efforts in maintaining the supply to these countries either by enhancing its shipping strategies or considering opening new stores there. Also, the firm should consider also expanding into more regions by working more on its delivery image.**
* The **discount effect** on the store’s profits is **negative**: the higher the discount rate the lower the profits so the company must **minimize its discount offers.**
* The most profitable **customer segment** is the **Consumer representing 51.03% of the company’s total profits**, so the store must **allocate more marketing efforts to this segment while also try to attract more the others.**
* The most profitable **product category** is **Technology** contributing by **1.33** **million** dollars to the overall profit followed by Office Supplies then Furniture contributing respectively by **1.04 million** dollars and **0.57** **million** dollars. **As a conclusion the company should maintain high levels of supply for these categories.**

1. **CHALLENGES AND FUTURE ENHANCEMENTS :**
2. **Challenges:**

* Handling incomplete or inconsistent data from legacy systems.
* Optimizing the ETL process for large datasets.
* Ensuring data accuracy and consistency across multiple sources.

1. **Future Enhancements:**

* Incorporating predictive analytics for sales forecasting.
* Expanding the data warehouse to include more granular data, such as marketing performance.
* Developing automated alerts for key performance indicator (KPI) thresholds.