Report

IT300 : Business Intelligence and Database Management Systems

Hardware Retail Store Analysis

Made by:
- Yassine Marrekchi
- Elamine Maaloul
- Raouf Lakhoues

Professors :
- Manel Abdelkader
- Ameni Azzouz

1 - Introduction :

This business intelligence project focuses on analyzing a database containing information about products, categories, pricing, storage, and geographical locations. Our objective is to distill meaningful patterns in product performance, sales, costs, and distribution dynamics. Through the creation of an intuitive Power BI dashboard, we aspire to equip users with visually compelling representations of crucial metrics, fostering informed decision-making for strategic optimization of sales tactics and inventory management.

2 - Implementation :

Data Gathering :

The dataset essential for this project has been obtained by an Oracle data Base integration as a mock dataset for data analysis. It was obtained by extracting information from Kaggle, a prominent platform for datasets and data science resources. " for dataset click here "

Data Transformation :

In our project, the dataset employed has undergone thorough cleaning, and as a result, it did not necessitate any additional operations in this regard. To facilitate efficient data manipulation, we focused on transformation processes, specifically converting the data type from "CSV" to "SQL." This transformation was instrumental in optimizing the dataset for enhanced data management and analysis. During this phase, we strategically rearranged the data into distinct tables, ensuring a structured format conducive to streamlined data manipulation and analysis within our chosen data framework. The utilization of this approach was paramount for maximizing the effectiveness of subsequent analytical processes.

	Table ▲	Action							Rows	Type	Collation	Size	Overhead
☐	category	★	Browse	Structure	Search	Insert	Empty	Drop	1,112	InnoDB	utf8mb4_general_ci	192.0 KiB	-
☐	location	★	Browse	Structure	Search	Insert	Empty	Drop	1,112	InnoDB	utf8mb4_general_ci	352.0 KiB	-
☐	pricing	★	Browse	Structure	Search	Insert	Empty	Drop	1,112	InnoDB	utf8mb4_general_ci	144.0 KiB	-
☐	product	★	Browse	Structure	Search	Insert	Empty	Drop	1,112	InnoDB	utf8mb4_general_ci	272.0 KiB	-
☐	storage	★	Browse	Structure	Search	Insert	Empty	Drop	1,112	InnoDB	utf8mb4_general_ci	192.0 KiB	-
	5 tables	Sum							5,560	InnoDB	utf8mb4_general_ci	1.1 MiB	0 B

Data Storage :

In the data storage phase, we used SQL-Front as our database management system, seamlessly connected to a localhost server. This setup ensures a robust and secure environment for storing and managing our project's datasets. SQL's relational capabilities enable efficient querying and retrieval of information, facilitating smooth data interactions. Attached here is the data structure :

- **Product Table (product):**

  - Columns: PRODUCT_ID*, PRODUCT_NAME, DESCRIPTION - Detail 1, DESCRIPTION - Detail 2, DESCRIPTION - Detail 3, DESCRIPTION - Detail 4.

- **Category Table (category):**

  - Columns: PRODUCT_ID*, CATEGORY_ID*, CATEGORY_NAME.

- **Pricing Table (pricing):**

  - Columns: PRODUCT_ID*, STANDARD_COST, LIST_PRICE.

- **Storage Table (storage):**

  - Columns: PRODUCT_ID*, WAREHOUSE_ID*, QUANTITY, WAREHOUSE_NAME.

- **Location Table (location):**

  - Columns: PRODUCT_ID*, COUNTRY_ID*, REGION_ID*, LOCATION_ID*, ADDRESS, POSTAL_CODE, CITY, STATE, COUNTRY_NAME.

note : Index is mentioned by " * "

Fact :

- The fact in this data is the product line of the hardware retail store represented by the product table

Dimensions :

- Pricing : Includes the needed information related to the pricing of the products.
- Storage : Mentions details about the warehouses where the products are stored.
- Location : Provides the distribution of the products across the warehouses in different counties along with detailed description of the location .
- Category : Represents the category of each product .

Measures : (check the file containing the views code)

Measures were demonstrated through views using SQL , here are the corresponding details :

- Profit (profit) : Shows the profit generated by each product.
- Product by category (productbycategory): Displays product categories along with quantity
- Average Standard Cost by Category (average_standard_cost_by_category): Calculates the average standard cost for each product category.
- Total Sales by Product (total_sales_by_product): Summarizes total sales coming from each product.
- Total Cost by Product (total_cost_by_product): Summarizes total cost by product.
- Product Profit Metrics (product_profit_metrics): Combines total sales, total cost, profit, for the purpose of extracting profit margin for each product.
- Geographic Distribution (geographic_distribution): Shows the distribution of warehouses across different countries and the corresponding quantity in each warehouse.
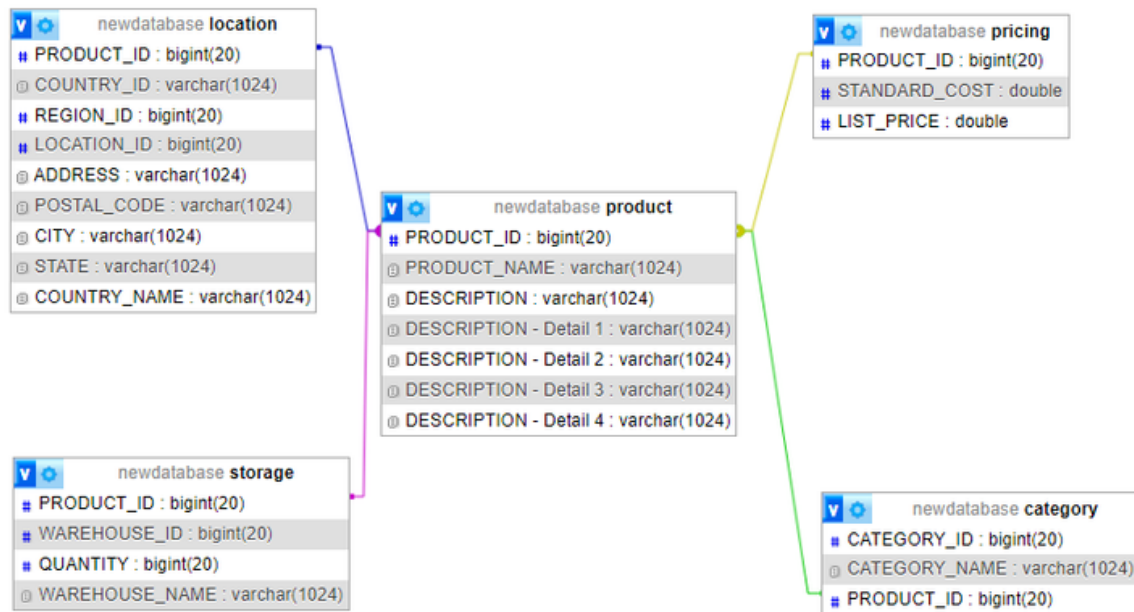
Figure 1: Data Warehouse Schema

As it is shown in the above figure, PRODUCT_ID here is a primary key in each of the dimension tables :
pricing, location, storage, category.
This feature allowed us to reduce the number of foreign keys to implement to our fact table (product) into only one foreign key .
assuring a functioning relational star schema between our tables .

Data Visualization :

Products Menu:

- Displays product names, descriptions, and profit margins.
- Utilizes the product table and the product_profit_metrics view.

Geographical Distribution Map:

- Provides a visual representation of inventory distribution across different countries.
- Uses the geographic_distribution view.

Products Categories Table:

- Lists product categories with quantity and images.
- Enhances user experience with visual representation.
- Relies on the productbycategory view.

Performance Indicators:

- Consists of three cards displaying total sales, total profit, and total cost of goods sold.
- Based on the total_cost_by_product, total_sales_by_product, and profit views.

Cost Allocation by Category:

- Utilizes a stacked column chart for visualizing average costs by product category.
- Draws insights from the average_standard_cost_by_category view.

Top Products by Profit Bar Chart:

- Presents the top products based on maximum profit.
- Utilizes the total_sales_by_product view.

note : to satisfy the condition of having at least two different data type we used in this phase our dataset of type CSV along with the SQL.

INTERPRETIONS :

- The product with the highest profit margin is "Supermicro X10SDV-8C-TLN4F "
- The warehouses are placed in Australia, Canada, China, India, Mexico and The United States of America .
- We have 4 categories of products : CPU, Mother Board, Storage, Video Card , with CPU having the highest stock among the other categories.
- Top products by profit bar chart lead us to "PNY VCQM6000-PB" with the most sales
- Video cards are the most costing products
- We have total sales of $5.43billion, incurring costs of goods sold of $4.33billion and therefore the total profit is $1.1 billion

3 - Conclusion:

In conclusion, our analysis identifies key trends, such as the highest profit margin product and strategic warehouse placements. The dominance of certain product categories, along with insights into top-performing items, showcases valuable information. Additionally, our dashboard effectively reports on overall sales, costs, and profits. To enhance its utility, integrating real-time inventory updates and predictive analytics could provide more actionable insights for informed decision-making.