Optimizing operations for Lenskart

Introduction:

Lenskart, a leading eyewear retailer in India, aims to enhance its operational efficiency by centralizing product information, pricing details, customer reviews, and improving customer and employee databases. The objective of this project is to increase sales, identify underperforming stores, and visualize sales trends over the years. This report outlines the steps taken in data extraction, transformation, and visualization using various tools and methodologies.

ETL(Extract, Transform, Load) :

The ETL process involved several critical steps to ensure data integrity and readiness for analysis:

1. ***Extract :***
   - **Scraping Data:** The Lenskart website was scraped using Python and BeautifulSoup to extract detailed information on eyeglasses, sunglasses, and contact lenses, along with pricing details, discounts, and promotions and also to get the loaction of stores and their details in India.

   - **Client Data:** Additional data, including transaction files and customer details, were obtained from the client.

2. ***Transform :***
   - **Data Cleaning:** The extracted and client-provided data were cleaned to remove null values and duplicates, ensuring consistency and accuracy.
   - **Data Structuring:** The data was structured to facilitate easy loading into the database. This involved ensuring proper data types, removing inconsistencies, and formatting data as per the requirements.

3. ***Load :***
   - **Azure Blob Storage:** The cleaned data was uploaded to Azure Blob Storage using BlobServiceClient in azure.storage.blob in python.

   - **Azure Data Factory**: Using Azure Data Factory, the data was ingested into a SQL database. Pipelining ensured data integrity and consistency during the loading process.

Data Models :

The project involves several data models to store and process the necessary information:

1. ***Orders:***

    - **Columns:** order_id, transaction_id, customer_id, product_id, store_id, quantity, order_date, payment_method
    - Contains transaction details for each order.

2. ***Products****:*

    - **Columns:** ProductID, ProductImageURL, ProductURL, ProductColor, ProductSize, ProductWidth, MarketPrice, LenskartPrice, ProductBrandName, ProductRating, TotalRatings.
    - Stores details about each product available on Lenskart.

3. ***StoreDetails:***

    - **Columns:** number, Name, Address, Location, Timings, PhoneNumber, Rating, NumberofRatings.
    - Information about each Lenskart store in India.

4. ***CustomerDetails****:*

    - **Columns:** customer_id, first name, last name, email, DOB, address, city, region.
    - Details about Lenskart customers.

Data presentation :

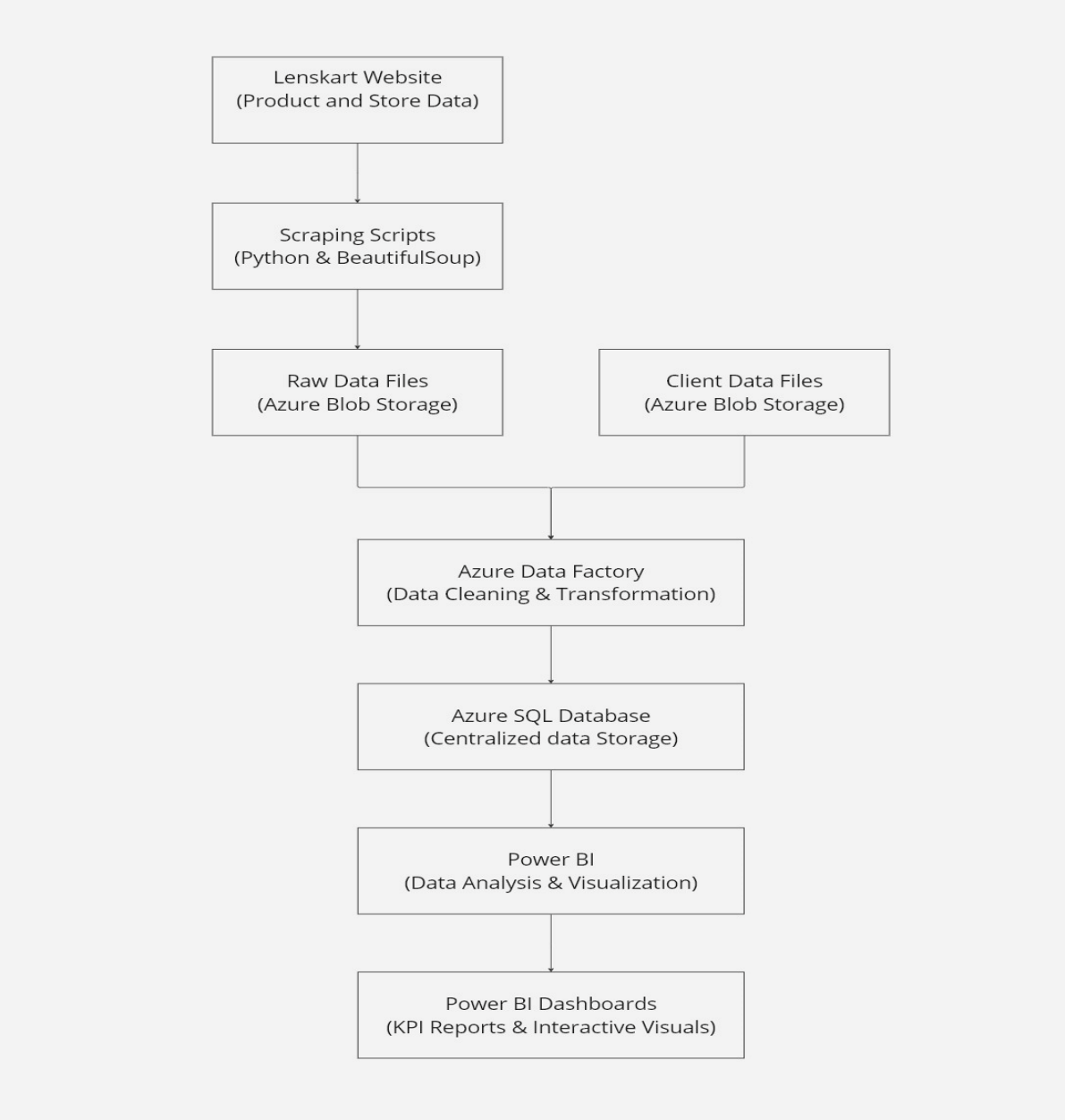
1. ***Azure SQL database :***
    - Written some sql queries for the KPI's to get the ouput data.

2. ***Connecting to Azure SQL Database****:*
    - Connected Power BI to the Azure SQL Database to access the data.

3. ***Creating Visualizations****:*
    - Used DAX queries to generate visualizations for the specified KPIs.
    - Created interactive dashboards to display sales trends, top-performing products, store comparisons, and customer insights.

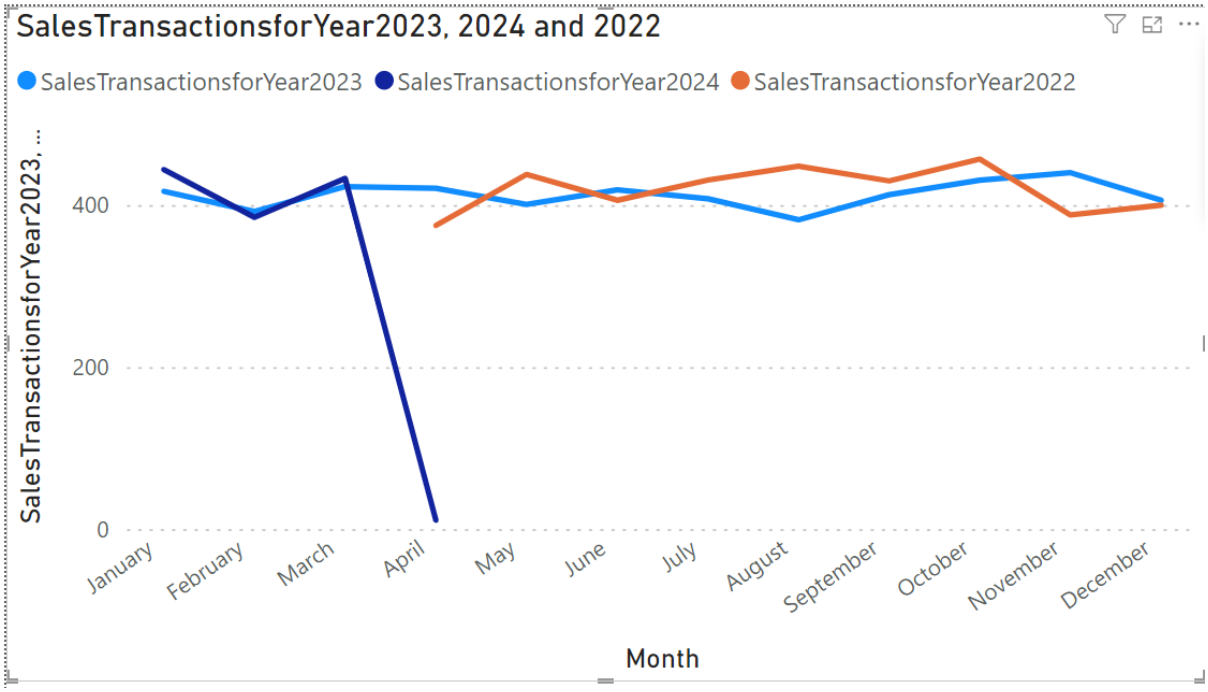
4. ***KPI Visualizations :***
    - Compared the count of sales transactions over a period between the current year and the previous year.
    - Analyzed average transaction value by payment type.
    - Compared current year revenue to the previous year.
    - Identified the top 10 products by revenue in the last 30 days.
    - Calculated the average unit purchase of products in the last 30 days.
    - Listed the top 15 highly rated products.

Architecture :

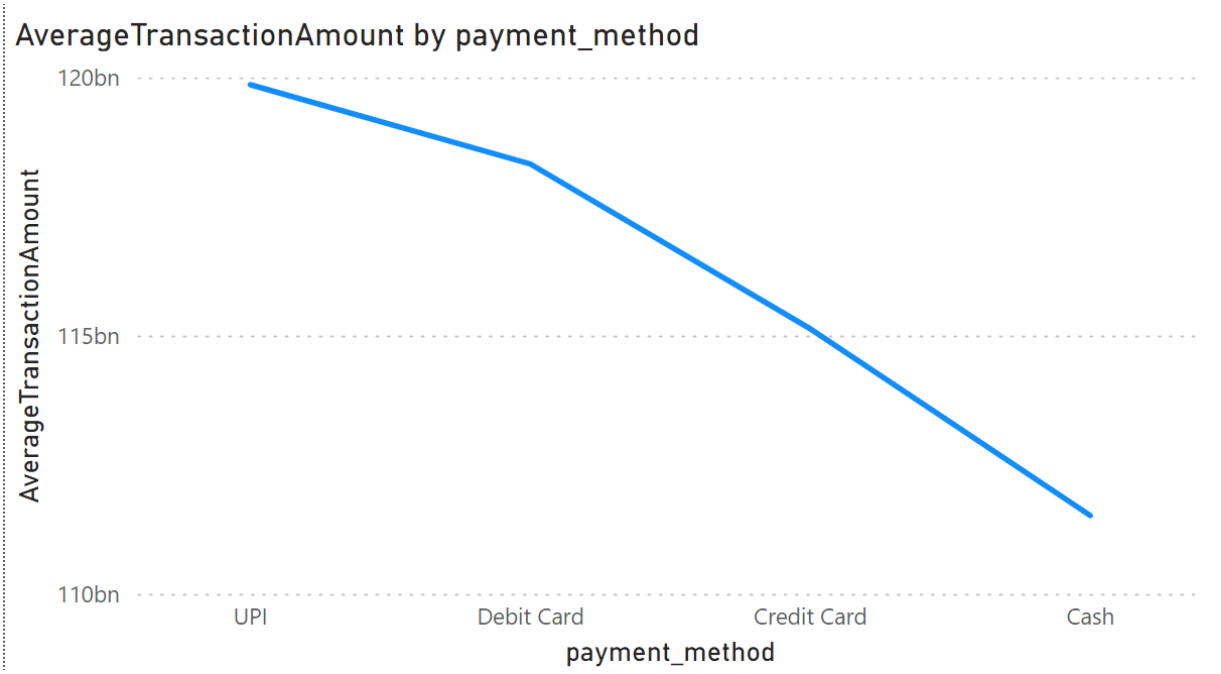


Some Power BI visualizations :

1. Compared the count of sales transactions over a period between the current year and the previous year.



2. Analyze average transaction value by payment type.



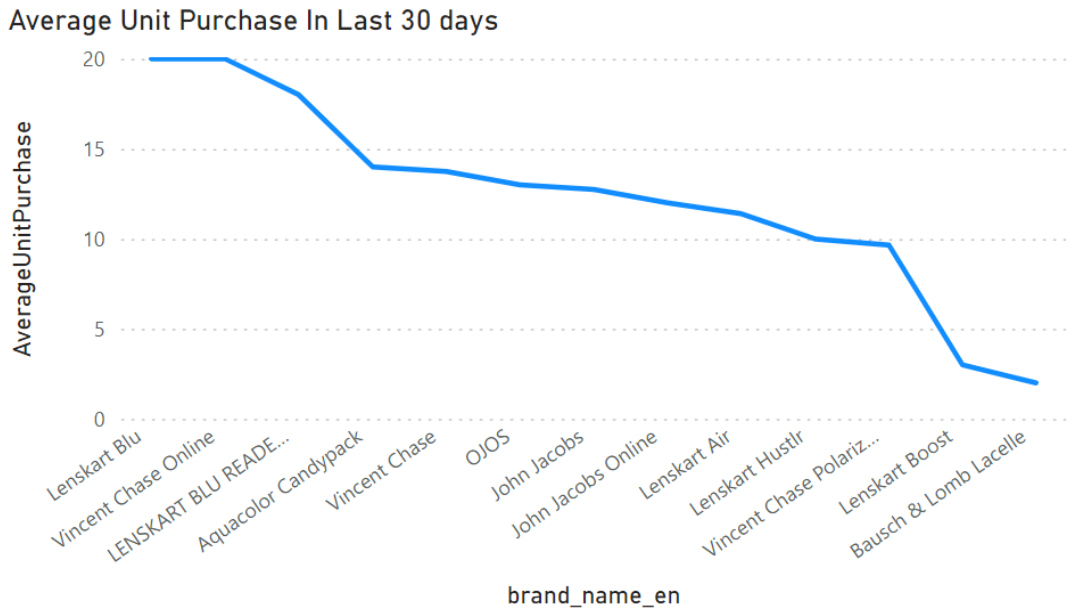
3. Compared current year revenue to the previous year.

228.63bn	352.59bn	34.93bn
Revenue_for_Year_2022	Revenue_for_Year_2023	Revenue_for_the_Year2024

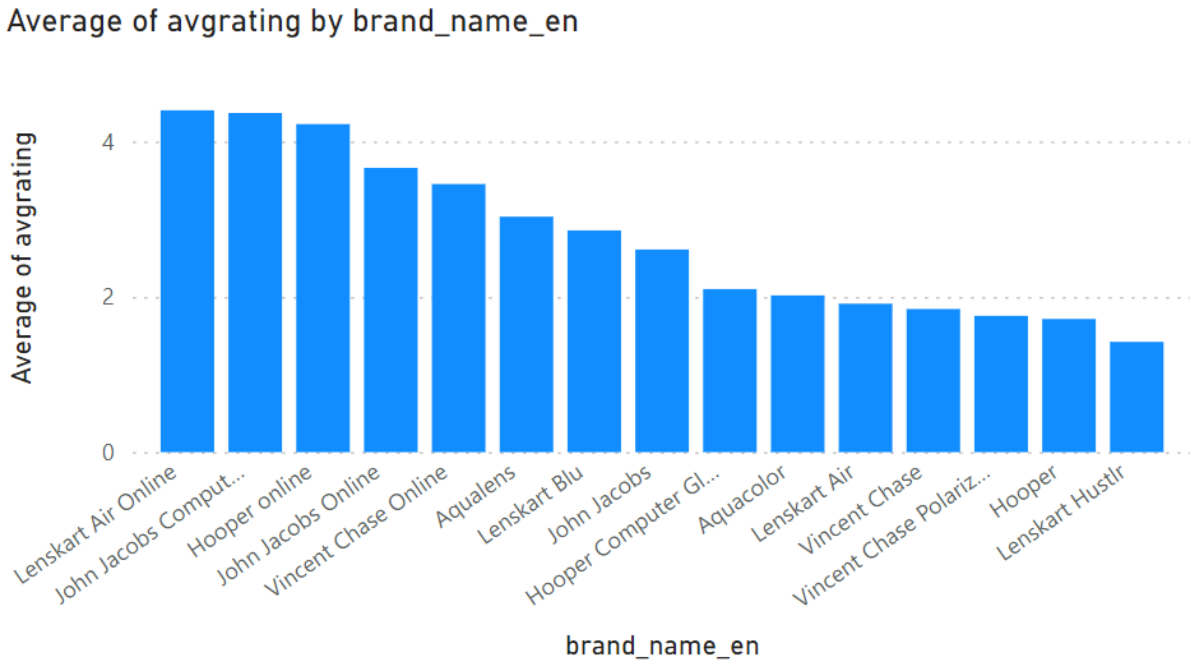
4. Identified the top 10 products by revenue in the last 30 days.

RevenueLast30Days	model_name	brand_name_en
1,33,62,000.00	LB E14058-W	Lenskart Hustlr
88,40,000.00	LA E15417-W	Lenskart Hustlr
60,02,700.00	LA E15417-N	Lenskart Hustlr
58,32,000.00	LKJ E10062	Hooper online
56,91,000.00	VC 5158/P	Vincent Chase Polarized
48,07,600.00	HP D15011	Lenskart Hustlr
45,50,000.00	JJ S12504	John Jacobs
43,00,000.00	VC S15999	Lenskart Hustlr
12,60,000.00	JJ E10118	John Jacobs Computer Glasses
10,40,000.00	JJ E12787	John Jacobs Computer Glasses
9,15,000.00	JJ E12787	John Jacobs
3,22,000.00	JJ E10118	John Jacobs
1,22,500.00	JJ E10118	John Jacobs Online
42,500.00	LB E14058-W	Lenskart Blu
28,000.00	JJ E12787	John Jacobs Online
64,95,57,900.00		

5. Calculated the average unit purchase of products in the last 30 days.



6. Listed the top 15 highly rated products by their brand name.



Conclusion :

The project successfully streamlined Lenskart's data management process, providing valuable insights into sales performance and customer behavior. By leveraging Python for data extraction, Azure for data storage and transformation, and Power BI for visualization, Lenskart can now make informed decisions to boost sales and improve overall operational efficiency.