Smart Inventory Management in Retail and Warehousing: A Data-Driven Approach

Team Data Wranglers:

Reddy Narendranath Reddy
Machanapally Sanjay Kumar
Pocha Jaya Simha Reddy
Veluru Jayanth
Yalamanchili Nihitha

Introduction:

In the dynamic world of retail and warehousing, efficient inventory management is crucial for operational effectiveness and customer satisfaction. The complexities associated with inventory management, such as overstocking, stockouts, and resource misallocation, pose significant challenges. This summary outlines a data-driven approach to tackle these issues, leveraging advanced tools and techniques across the data lifecycle.

Generation and Collection of Data:

Data-driven techniques are essential for efficient inventory management in retail and storage. This includes combining data from several sources, including barcoding, RFID, and point-of-sale systems. The deployment of IoT devices makes real-time monitoring possible. Predictive analytics and cooperative efforts with suppliers enable accurate demand forecasts, averting stockouts and overstock. Using cloud-based solutions guarantees efficient operations with enhanced scalability and accessibility. Sensitive information is protected by strict adherence to legislation, security measures, and data quality assurance. A robust and flexible inventory management system is developed by ongoing process monitoring, personnel training programs, and the gradual integration of new technology. This eventually improves overall performance and customer satisfaction.

Data Lifecycle Steps and Tools Used:

1. Data Cleaning and Wrangling with OpenRefine:
    - The first step involves using OpenRefine, a powerful tool for data cleaning and preparation.
    - This process is essential for ensuring data quality, which includes handling missing values, correcting inconsistencies, and standardizing data formats.

2. Storing Pre-processed Dataset in Google Cloud Platform (GCP) Bucket:
    - After cleaning, the data is securely stored in a GCP Bucket.

- This cloud storage solution offers scalability and robustness, essential for handling large datasets typically encountered in retail and warehousing scenarios.

3. Loading Dataset into Hadoop Distributed File System (HDFS):
   - For further processing and analysis, the dataset is loaded into HDFS.
   - HDFS offers a reliable and distributed storage method, crucial for managing big data in a distributed computing environment.

4. Analysis Using SQL Queries in BigQuery:
   - The core analysis is performed through SQL queries in BigQuery, Google's enterprise data warehouse.
   - This allows for efficient processing of large datasets and complex analytical queries.

5. Visualizing the Dataset Using Google Data Studio:
   - Visualizing data is a critical step for interpreting and presenting the analysis results.
   - Google Data Studio provides a versatile platform for creating interactive reports and dashboards, enhancing the understanding and communication of data insights.

6. Interpreting Analysis Results:
   - The final step involves interpreting the results of the analysis.
   - This step is crucial for drawing actionable insights and making informed decisions.

Use Cases of Dataset:

The processed and analyzed data serves multiple use cases in inventory management:

1. Monthly Sales Summary:
   - Provides an overview of sales performance, highlighting trends and patterns monthly.
2. Supplier Performance Analysis:
   - Assesses the reliability and efficiency of suppliers, influencing procurement strategies.
3. Item Type Sales and Stock Analysis:
   - Offers insights into the sales volume and stock levels of different item categories, aiding in inventory optimization.
4. Identifying Top Selling Items:
   - Helps in recognizing the most popular items, guiding marketing and stock replenishment decisions.
5. Seasonal Sales Analysis:
   - Analyzes sales variations across different seasons, assisting in seasonal stock planning.
6. Yearly Supplier Contribution:
   - Evaluates the annual performance and contribution of suppliers to the business.

Conclusion and Next Steps:

The implementation of this data-driven approach in inventory management represents a significant advancement in addressing the challenges faced by retail and warehousing sectors. The next steps involve continuous monitoring and refinement of the analysis. The data science or analyst teams should focus on integrating real-time data analytics for more dynamic decision-making and exploring advanced analytical techniques like machine learning for predictive analytics. This ongoing process will ensure sustained improvements in inventory management, leading to increased profitability and customer satisfaction.

Supporting Documents:

One of the review's most important findings is how smart warehouse operations management is becoming increasingly important for companies looking to cut costs and increase productivity. The authors emphasize the development and implementation of intelligent warehouse solutions using a variety of technologies, including robots, artificial intelligence, and the Internet of Things (IoT).

ZHEN, L., & LI, H. (2022). A literature review of smart warehouse operations management. *Front. Eng. Manag. 2022, 9(1): 31–55*. https://doi.org/10.1007/s42524-021-0178-9

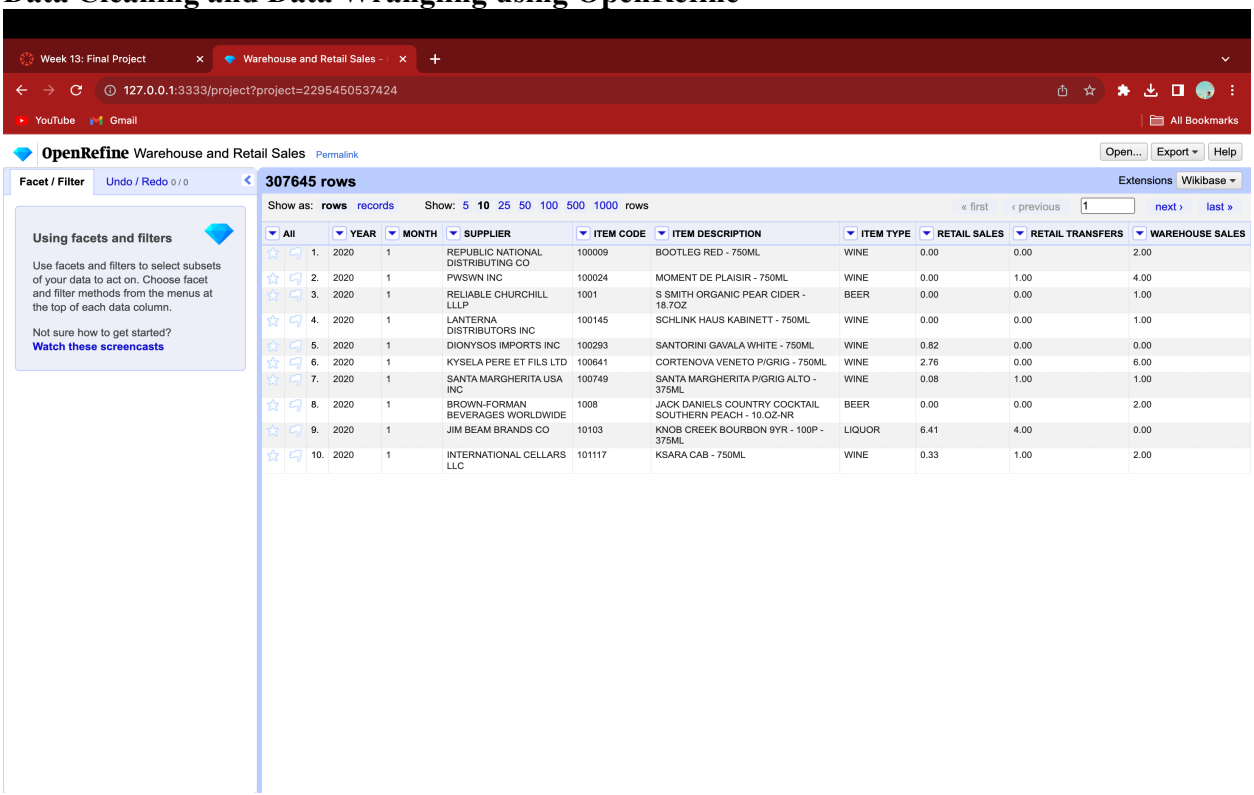
This describes a warehouse inventory management system that combines Iot tracking with open-source software. It is scalable and efficient, making it suited to a wide range of sectors. The system is flexible to fit warehouses of different sizes and employs open-source software to create inventory reports and track item movement. It provides flexibility in inventory management and may be used in a variety of industries, including manufacturing, distribution, and retail.

B., S. S. T., & S., N. (2018). Warehouse inventory management system using IoT and open-source framework. *Alexandria Engineering Journal Volume 57, Issue 4, December 2018, Pages 3817-3823*. https://doi.org/10.1016/j.aej.2018.02.003

In "Exploring Cyber-Physical Systems in Industry 4.0: A Comprehensive Examination of the Core Technologies in Smart Warehousing," the vital technologies that enable the integration of Cyber-Physical Systems (CPS) in smart warehouses are examined. In line with Industry 4.0 principles, the study explores the revolutionary effect of CPS on warehouse intelligence and operational efficacy.

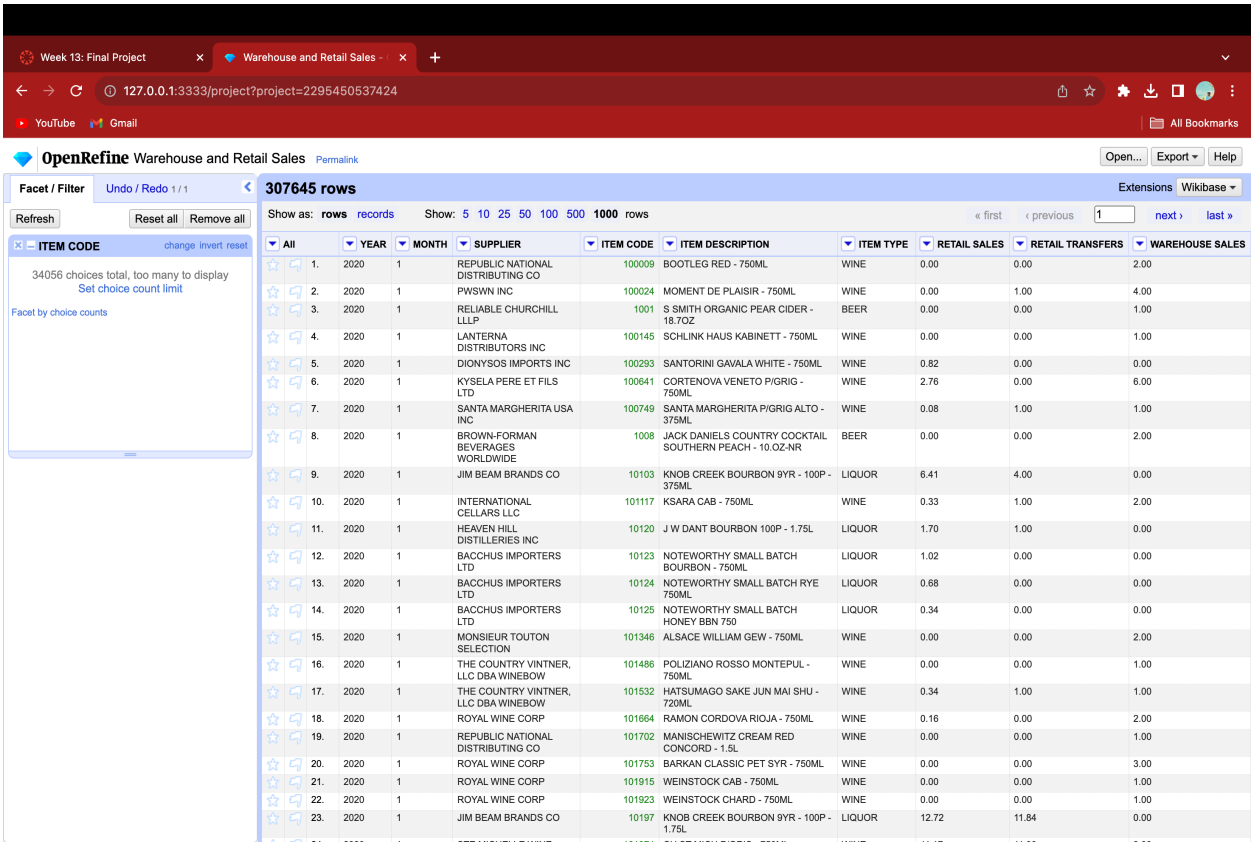
Liu, X., Cao, J., Yang, Y., & Jiang, S. (2018). CPS-Based Smart Warehouse for Industry 4.0: A Survey of the Underlying Technologies. Computers 2018, 7(1), 13. https://doi.org/10.3390/computers7010013

Data Cleaning and Data Wrangling using OpenRefine



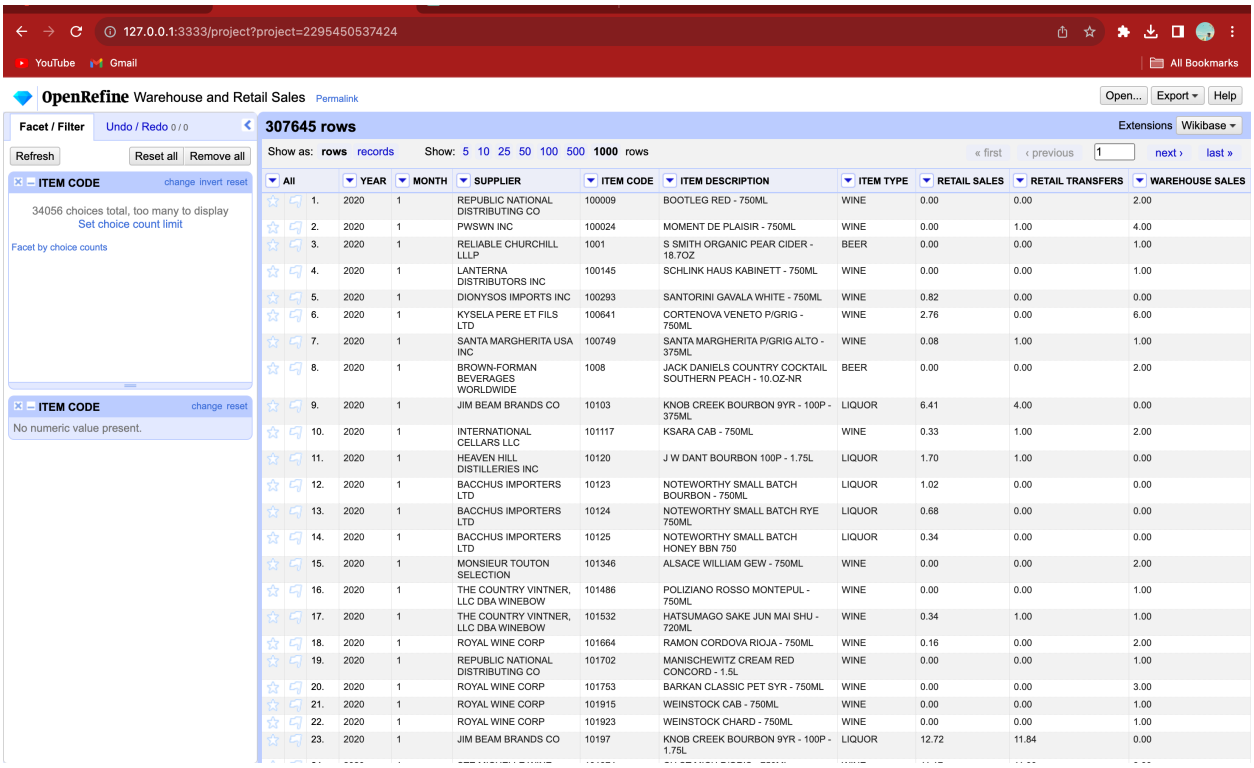
	All	YEAR	MONTH	SUPPLIER	ITEM CODE	ITEM DESCRIPTION	ITEM TYPE	RETAIL SALES	RETAIL TRANSFERS	WAREHOUSE SALES
1.		2020	1	REPUBLIC NATIONAL DISTRIBUTING CO	100009	BOOTLEG RED - 750ML	WINE	0.00	0.00	2.00
2.		2020	1	PWSWN INC	100024	MOMENT DE PLAISIR - 750ML	WINE	0.00	1.00	4.00
3.		2020	1	RELIABLE CHURCHILL LLLP	1001	S SMITH ORGANIC PEAR CIDER - 18.7OZ	BEER	0.00	0.00	1.00
4.		2020	1	LANTERNA DISTRIBUTORS INC	100145	SCHLINK HAUS KABINETT - 750ML	WINE	0.00	0.00	1.00
5.		2020	1	DIONYSOS IMPORTS INC	100293	SANTORINI GAVALA WHITE - 750ML	WINE	0.82	0.00	0.00
6.		2020	1	KYSELA PERE ET FILS LTD	100641	CORTENOVA VENETO P/GRIG - 750ML	WINE	2.76	0.00	6.00
7.		2020	1	SANTA MARGHERITA USA INC	100749	SANTA MARGHERITA P/GRIG ALTO - 375ML	WINE	0.08	1.00	1.00
8.		2020	1	BROWN-FORMAN BEVERAGES WORLDWIDE	1008	JACK DANIELS COUNTRY COCKTAIL SOUTHERN PEACH - 10.OZ-NR	BEER	0.00	0.00	2.00
9.		2020	1	JIM BEAM BRANDS CO	10103	KNOB CREEK BOURBON 9YR - 100P - 375ML	LIQUOR	6.41	4.00	0.00
10.		2020	1	INTERNATIONAL CELLARS LLC	101117	KSARA CAB - 750ML	WINE	0.33	1.00	2.00

Text Facet for Item Code

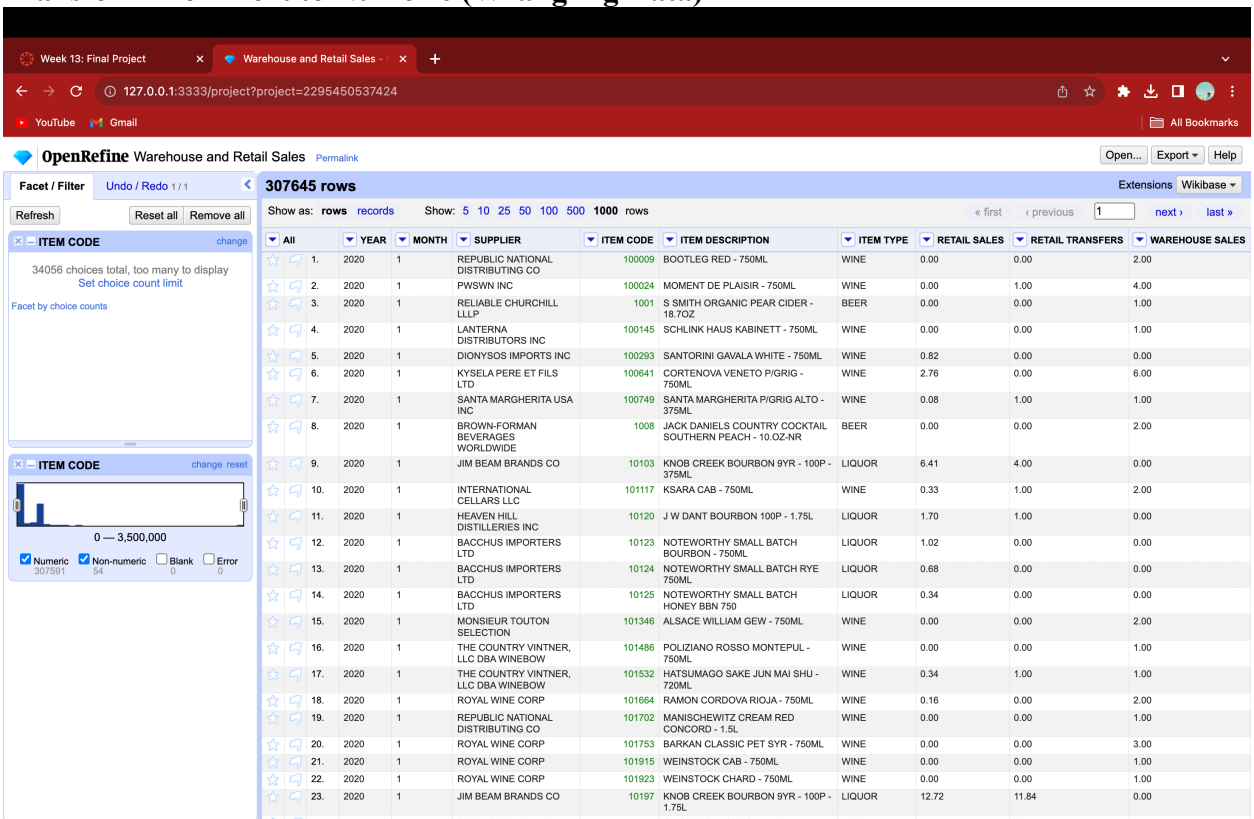


	All	YEAR	MONTH	SUPPLIER	ITEM CODE	ITEM DESCRIPTION	ITEM TYPE	RETAIL SALES	RETAIL TRANSFERS	WAREHOUSE SALES
1.		2020	1	REPUBLIC NATIONAL DISTRIBUTING CO	100009	BOOTLEG RED - 750ML	WINE	0.00	0.00	2.00
2.		2020	1	PWSWN INC	100024	MOMENT DE PLAISIR - 750ML	WINE	0.00	1.00	4.00
3.		2020	1	RELIABLE CHURCHILL LLLP	1001	S SMITH ORGANIC PEAR CIDER - 18.7OZ	BEER	0.00	0.00	1.00
4.		2020	1	LANTERNA DISTRIBUTORS INC	100145	SCHLINK HAUS KABINETT - 750ML	WINE	0.00	0.00	1.00
5.		2020	1	DIONYSOS IMPORTS INC	100293	SANTORINI GAVALA WHITE - 750ML	WINE	0.82	0.00	0.00
6.		2020	1	KYSELA PERE ET FILS LTD	100641	CORTENOVA VENETO P/GRIG - 750ML	WINE	2.76	0.00	6.00
7.		2020	1	SANTA MARGHERITA USA INC	100749	SANTA MARGHERITA P/GRIG ALTO - 375ML	WINE	0.08	1.00	1.00
8.		2020	1	BROWN-FORMAN BEVERAGES WORLDWIDE	1008	JACK DANIELS COUNTRY COCKTAIL SOUTHERN PEACH - 10.OZ-NR	BEER	0.00	0.00	2.00
9.		2020	1	JIM BEAM BRANDS CO	10103	KNOB CREEK BOURBON 9YR - 100P - 375ML	LIQUOR	6.41	4.00	0.00
10.		2020	1	INTERNATIONAL CELLARS LLC	101117	KSARA CAB - 750ML	WINE	0.33	1.00	2.00
11.		2020	1	HEAVEN HILL DISTILLERIES INC	10120	J W DANT BOURBON 100P - 1.75L	LIQUOR	1.70	1.00	0.00
12.		2020	1	BACCHUS IMPORTERS LTD	10123	NOTEWORTHY SMALL BATCH BOURBON - 750ML	LIQUOR	1.02	0.00	0.00
13.		2020	1	BACCHUS IMPORTERS LTD	10124	NOTEWORTHY SMALL BATCH RYE 750ML	LIQUOR	0.68	0.00	0.00
14.		2020	1	BACCHUS IMPORTERS LTD	10125	NOTEWORTHY SMALL BATCH HONEY BBN 750	LIQUOR	0.34	0.00	0.00
15.		2020	1	MONSIEUR TOUTON SELECTION	101346	ALSACE WILLIAM GEW - 750ML	WINE	0.00	0.00	2.00
16.		2020	1	THE COUNTRY VINTNER, LLC DBA WINEBOW	101486	POLIZIANO ROSSO MONTEPUL - 750ML	WINE	0.00	0.00	1.00
17.		2020	1	THE COUNTRY VINTNER, LLC DBA WINEBOW	101532	HATSUMAGO SAKE JUN MAI SHU - 720ML	WINE	0.34	1.00	1.00
18.		2020	1	ROYAL WINE CORP	101664	RAMON CORDOVA RIOJA - 750ML	WINE	0.16	0.00	2.00
19.		2020	1	REPUBLIC NATIONAL DISTRIBUTING CO	101702	MANISCHEWITZ CREAM RED CONCORD - 1.5L	WINE	0.00	0.00	1.00
20.		2020	1	ROYAL WINE CORP	101753	BARKAN CLASSIC PET SYR - 750ML	WINE	0.00	0.00	3.00
21.		2020	1	ROYAL WINE CORP	101915	WEINSTOCK CAB - 750ML	WINE	0.00	0.00	1.00
22.		2020	1	ROYAL WINE CORP	101923	WEINSTOCK CHARD - 750ML	WINE	0.00	0.00	1.00
23.		2020	1	JIM BEAM BRANDS CO	10197	KNOB CREEK BOURBON 9YR - 100P - 1.75L	LIQUOR	12.72	11.84	0.00

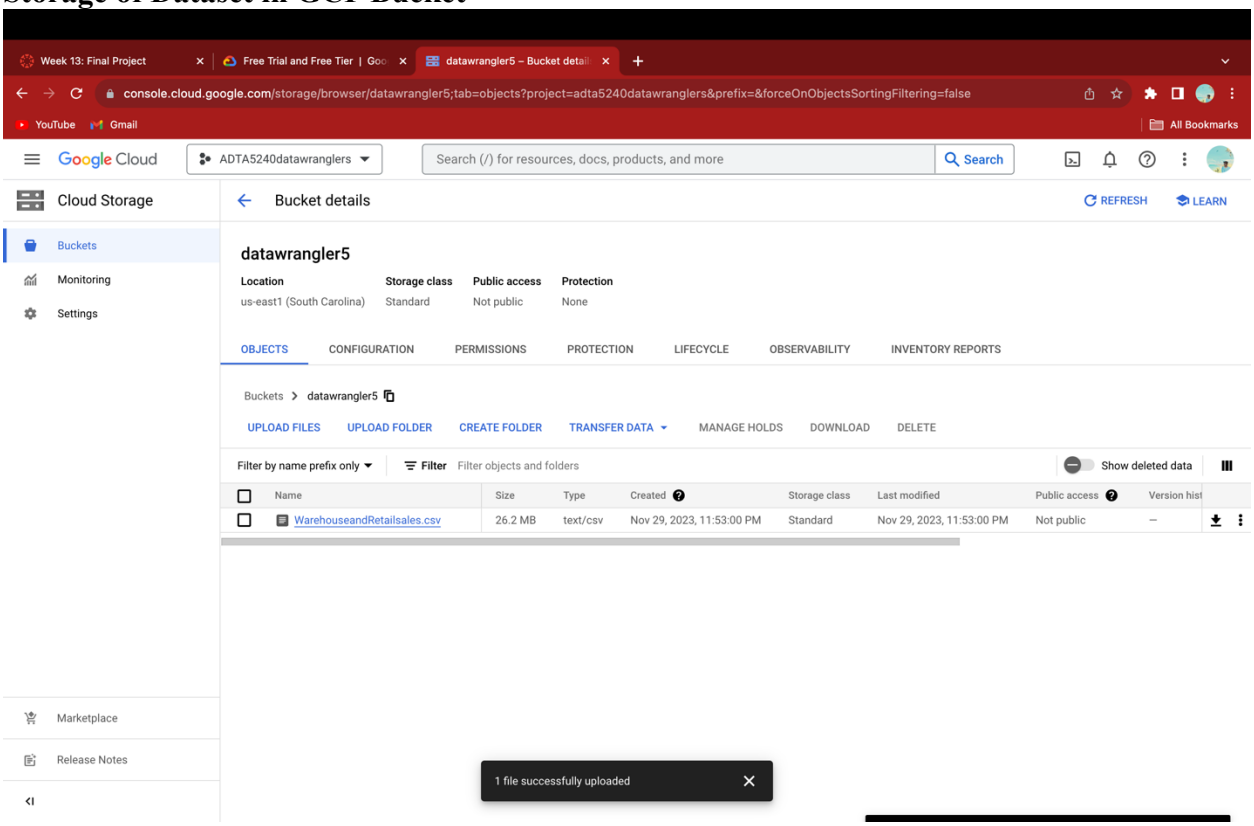
Numeric Facet for Item Code



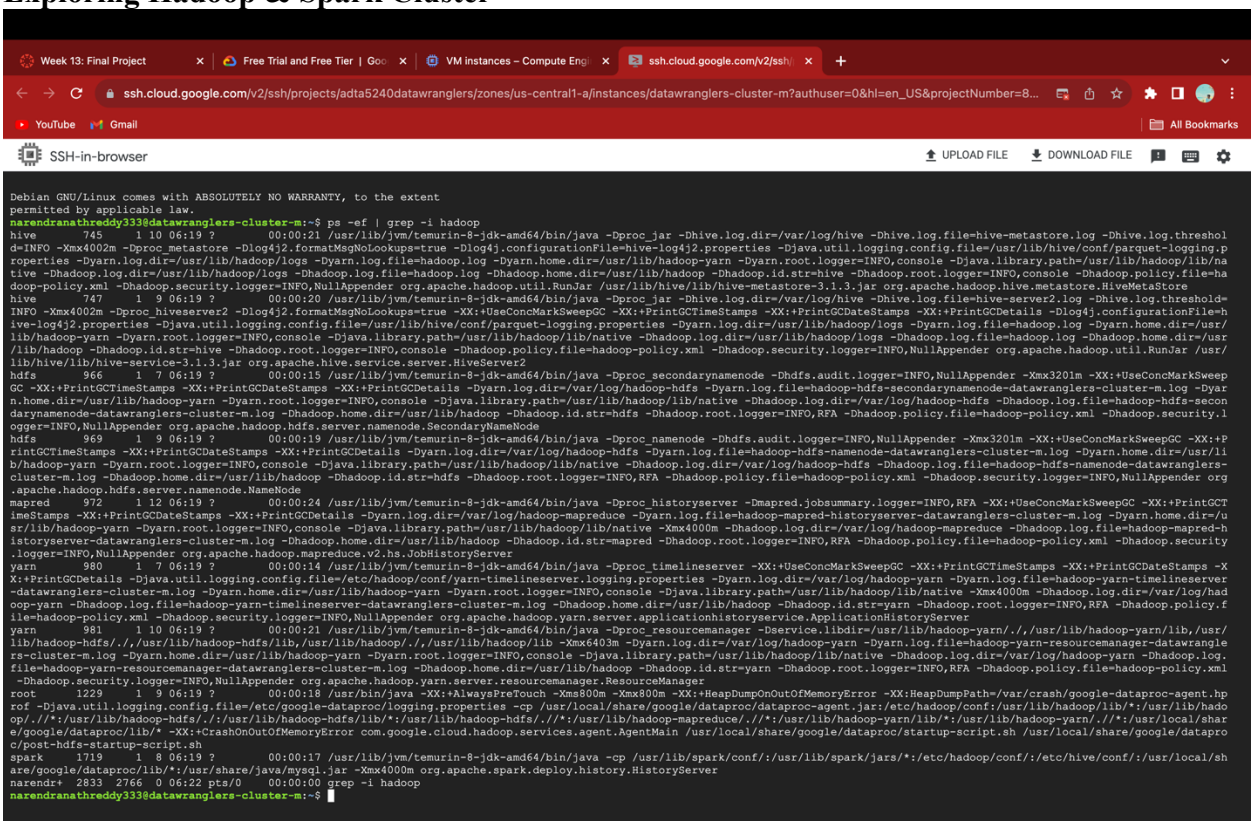
Transform from Text to Numeric (Wrangling Data)



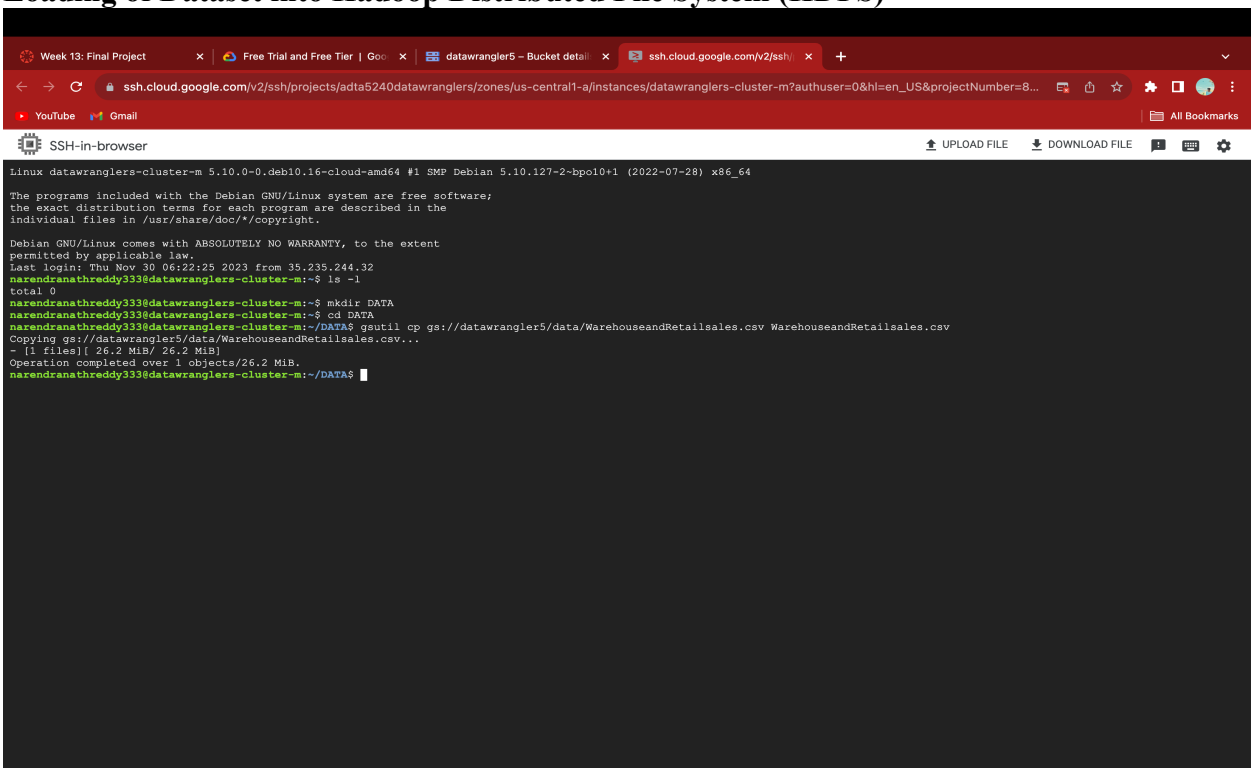
Storage of Dataset in GCP Bucket



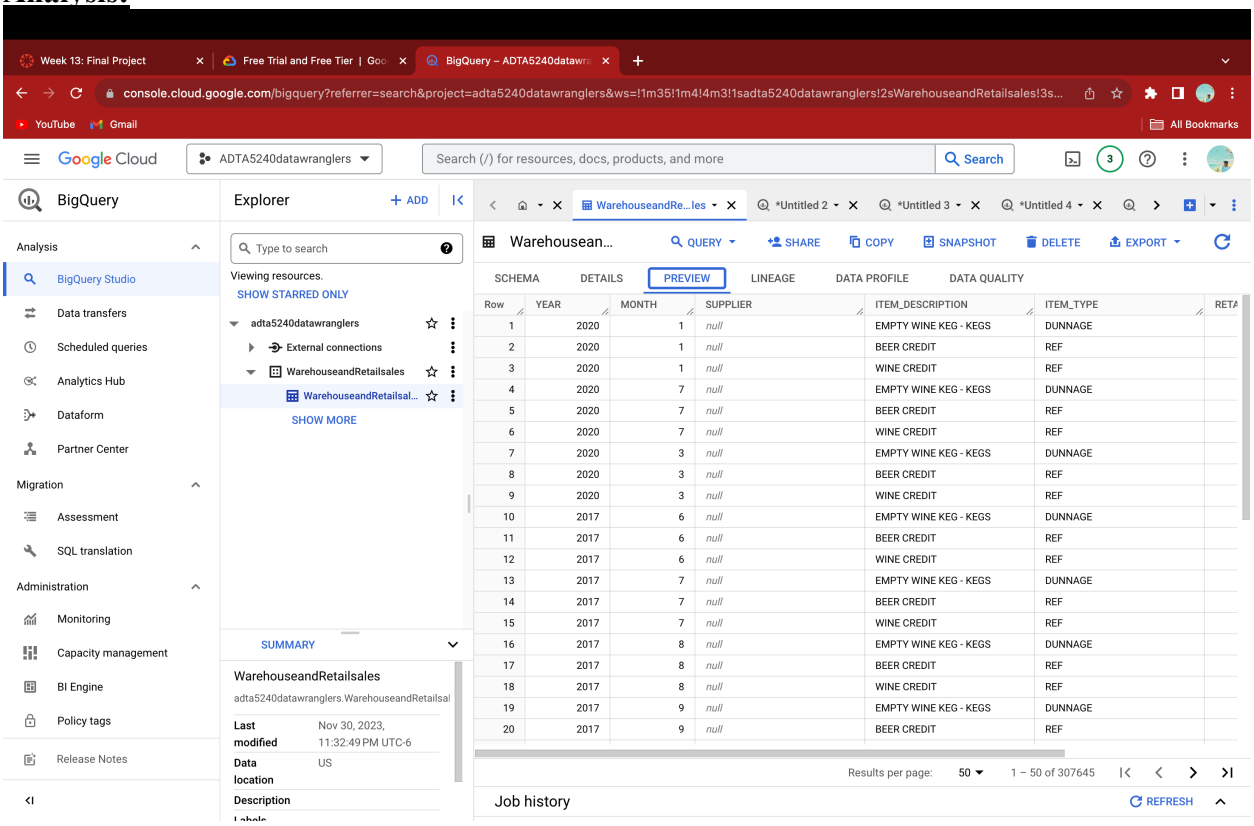
Exploring Hadoop & Spark Cluster



Loading of Dataset into Hadoop Distributed File System (HDFS)

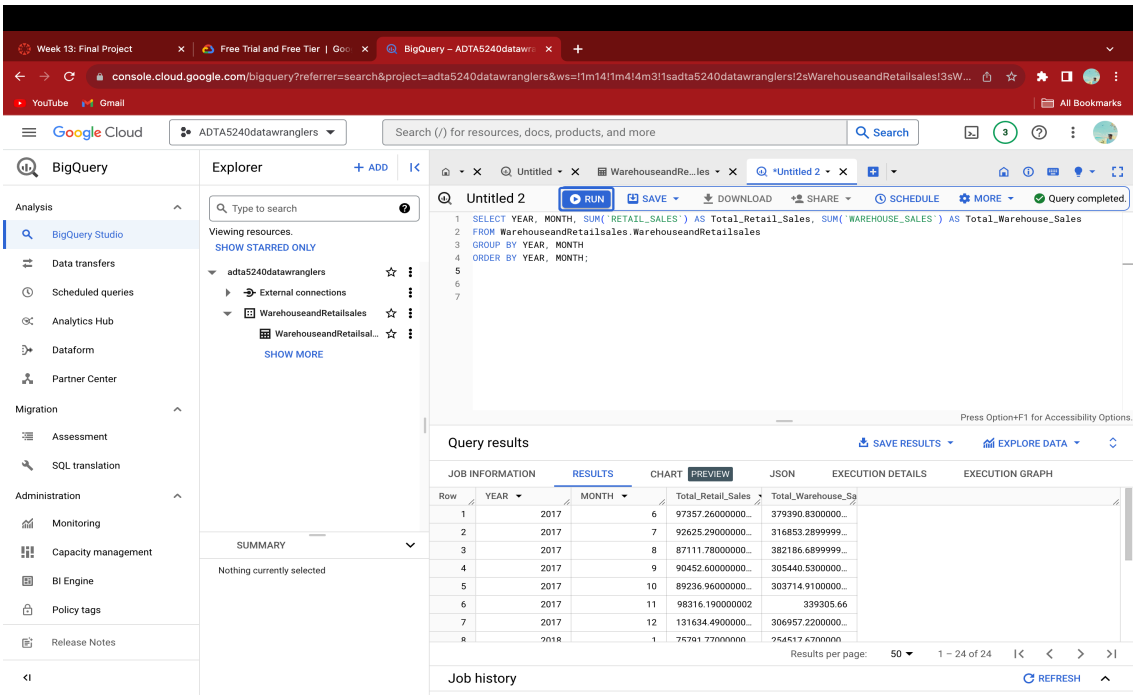


Analysis:

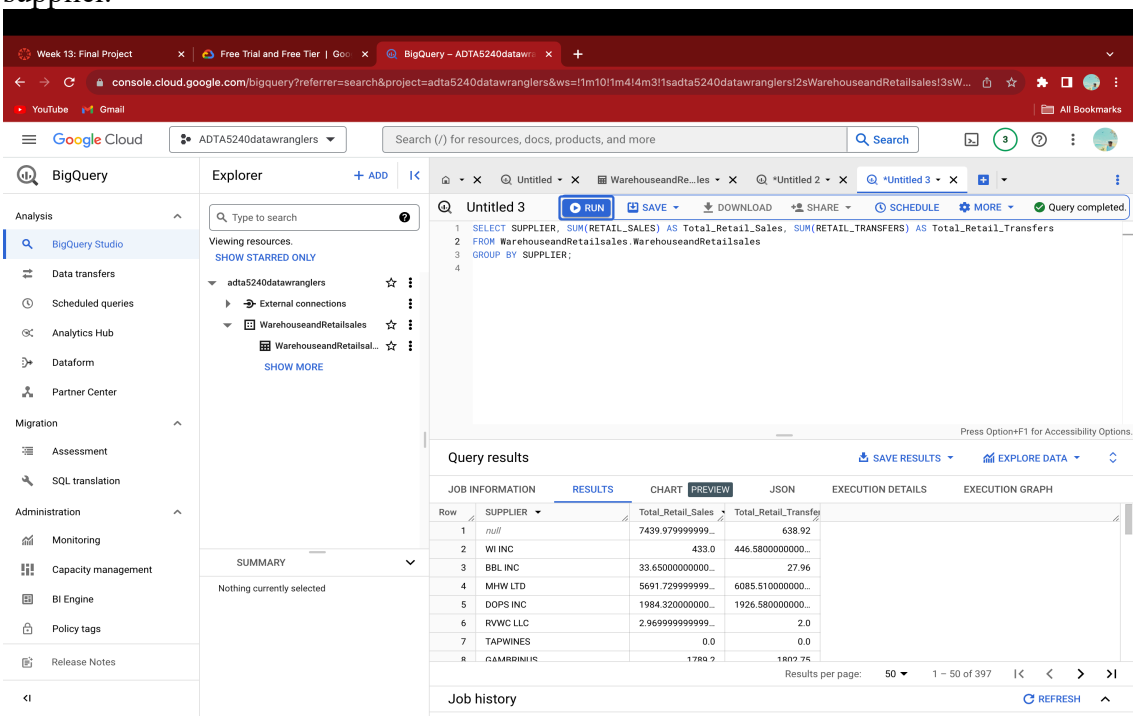


Use cases:

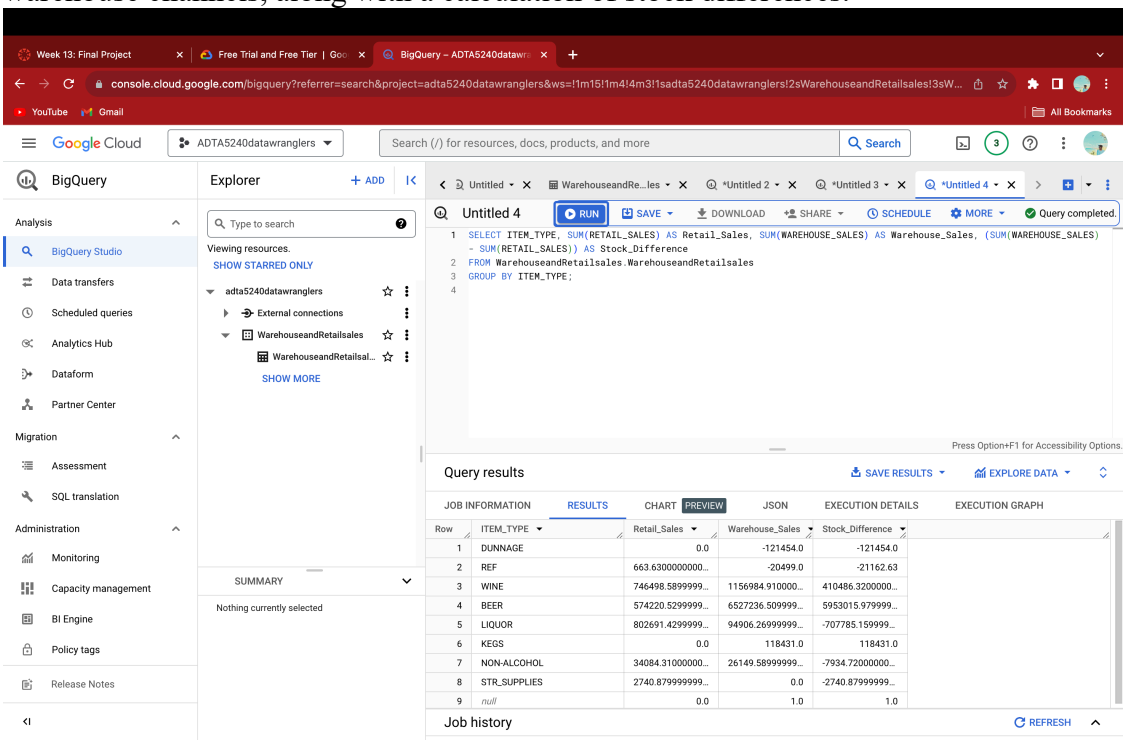
The purpose of this query is to provide a clear overview of sales trends over time, differentiating between retail and warehouse sales.



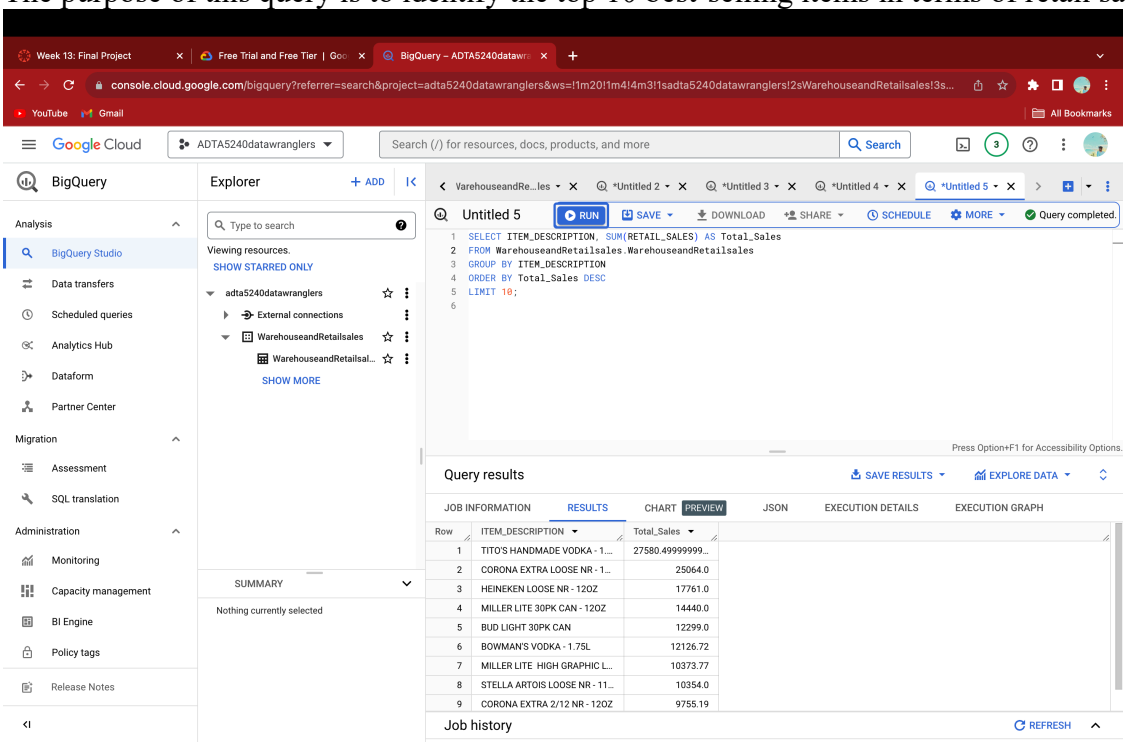
The purpose of this query to provide a summary of total retail sales and transfers for each supplier.



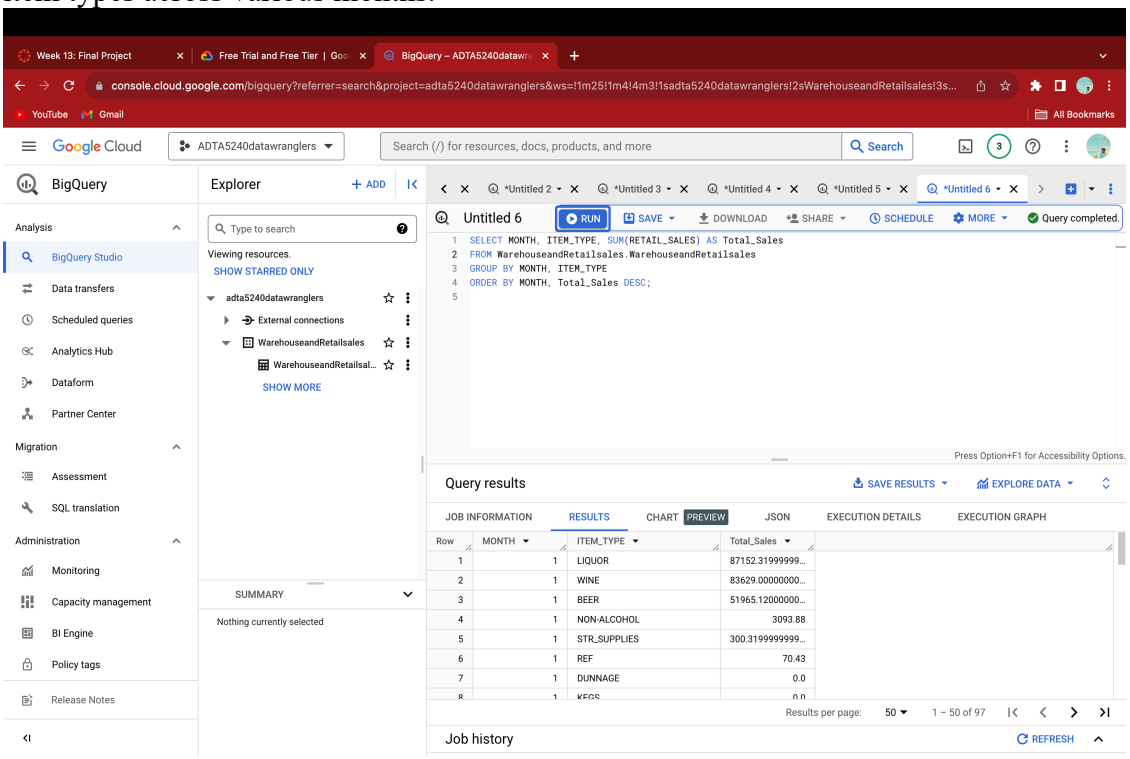
The purpose of this query is to provide an item-type-wise breakdown of sales across retail and warehouse channels, along with a calculation of stock differences.



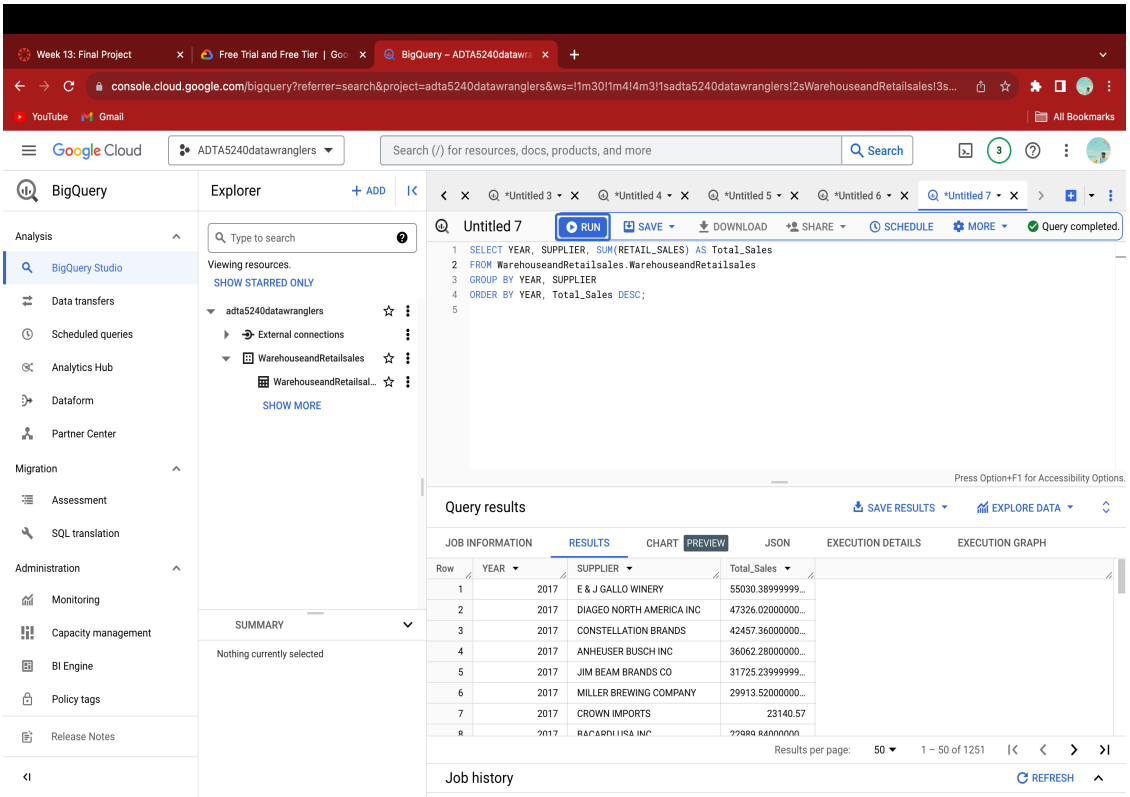
The purpose of this query is to identify the top 10 best-selling items in terms of retail sales.



The purpose of this query is to provide a detailed overview of sales performance for different item types across various months.



The purpose of this query is to provide a yearly comparison of supplier performance based on retail sales.



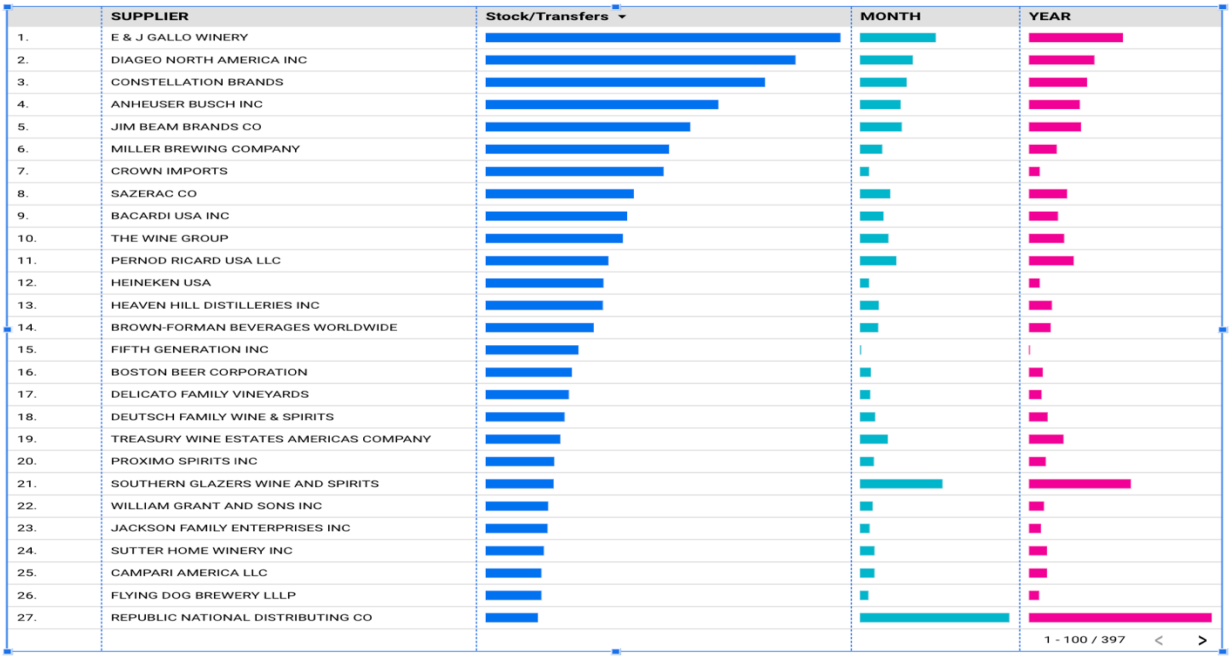
SQL Queries:

```
SELECT YEAR, MONTH, SUM(`RETAIL_SALES`) AS Total_Retail_Sales,
SUM(`WAREHOUSE_SALES`) AS Total_Warehouse_Sales
FROM WarehouseandRetailsales.WarehouseandRetailsales
GROUP BY YEAR, MONTH
ORDER BY YEAR, MONTH;


SELECT SUPPLIER, SUM(RETAIL_SALES) AS Total_Retail_Sales,
SUM(RETAIL_TRANSFERS) AS Total_Retail_Transfers
FROM WarehouseandRetailsales.WarehouseandRetailsales
GROUP BY SUPPLIER;


SELECT ITEM_TYPE, SUM(RETAIL_SALES) AS Retail_Sales,
SUM(WAREHOUSE_SALES) AS Warehouse_Sales, (SUM(WAREHOUSE_SALES) -
SUM(RETAIL_SALES)) AS Stock_Difference
FROM WarehouseandRetailsales.WarehouseandRetailsales
GROUP BY ITEM_TYPE;

SELECT ITEM_DESCRIPTION, SUM(RETAIL_SALES) AS Total_Sales
FROM WarehouseandRetailsales.WarehouseandRetailsales
GROUP BY ITEM_DESCRIPTION
ORDER BY Total_Sales DESC
LIMIT 10;


SELECT MONTH, ITEM_TYPE, SUM(RETAIL_SALES) AS Total_Sales
FROM WarehouseandRetailsales.WarehouseandRetailsales
GROUP BY MONTH, ITEM_TYPE
ORDER BY MONTH, Total_Sales DESC;


SELECT YEAR, SUPPLIER, SUM(RETAIL_SALES) AS Total_Sales
FROM WarehouseandRetailsales.WarehouseandRetailsales
GROUP BY YEAR, SUPPLIER
ORDER BY YEAR, Total_Sales DESC;
```

Visualization (Use Cases)

Monthly Sales Summary

The objective of the Monthly Sales Summary is to provide a comprehensive overview of sales for each month across different years.

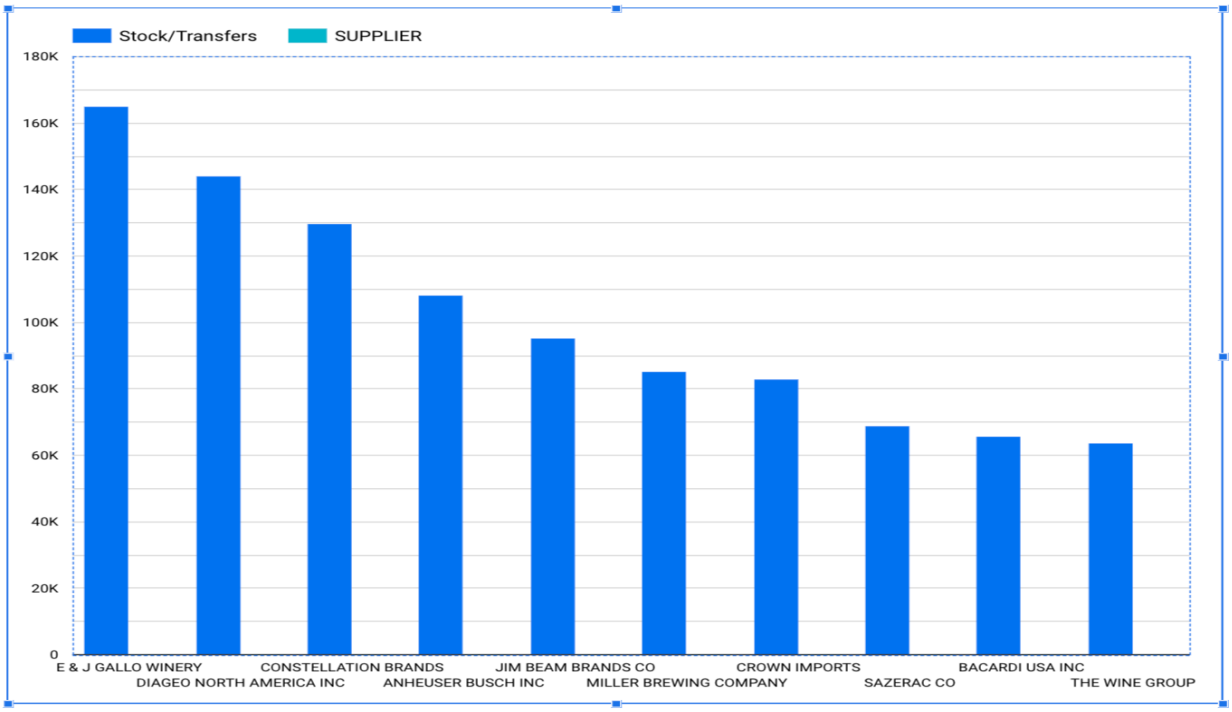


	SUPPLIER	Stock/Transfers	MONTH	YEAR
1.	E & J GALLO WINERY			
2.	DIAGEO NORTH AMERICA INC			
3.	CONSTELLATION BRANDS			
4.	ANHEUSER BUSCH INC			
5.	JIM BEAM BRANDS CO			
6.	MILLER BREWING COMPANY			
7.	CROWN IMPORTS			
8.	SAZERAC CO			
9.	BACARDI USA INC			
10.	THE WINE GROUP			
11.	PERNOD RICARD USA LLC			
12.	HEINEKEN USA			
13.	HEAVEN HILL DISTILLERIES INC			
14.	BROWN-FORMAN BEVERAGES WORLDWIDE			
15.	FIFTH GENERATION INC			
16.	BOSTON BEER CORPORATION			
17.	DELICATO FAMILY VINEYARDS			
18.	DEUTSCH FAMILY WINE & SPIRITS			
19.	TREASURY WINE ESTATES AMERICAS COMPANY			
20.	PROXIMO SPIRITS INC			
21.	SOUTHERN GLAZERS WINE AND SPIRITS			
22.	WILLIAM GRANT AND SONS INC			
23.	JACKSON FAMILY ENTERPRISES INC			
24.	SUTTER HOME WINERY INC			
25.	CAMPARI AMERICA LLC			
26.	FLYING DOG BREWERY LLLP			
27.	REPUBLIC NATIONAL DISTRIBUTING CO			

1 - 100 / 397 < >

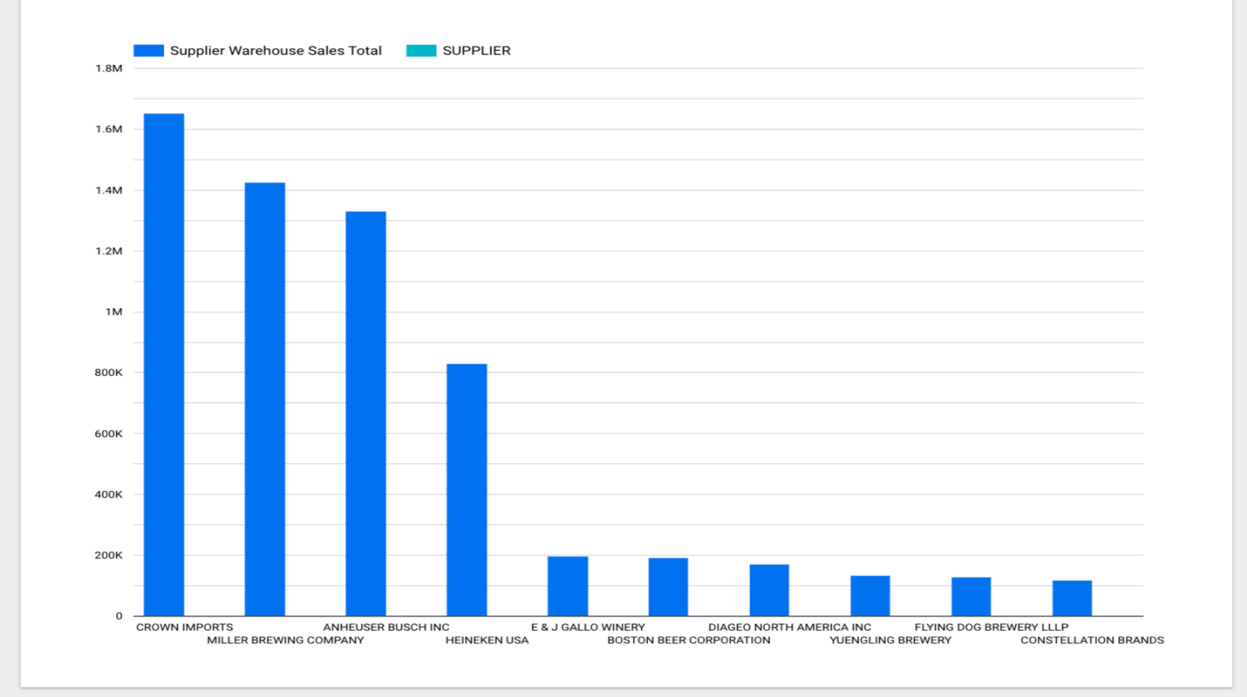
Supplier Performance Analysis

The objective of the Supplier Performance Analysis is to assess the performance of each supplier by calculating their average retail sales.



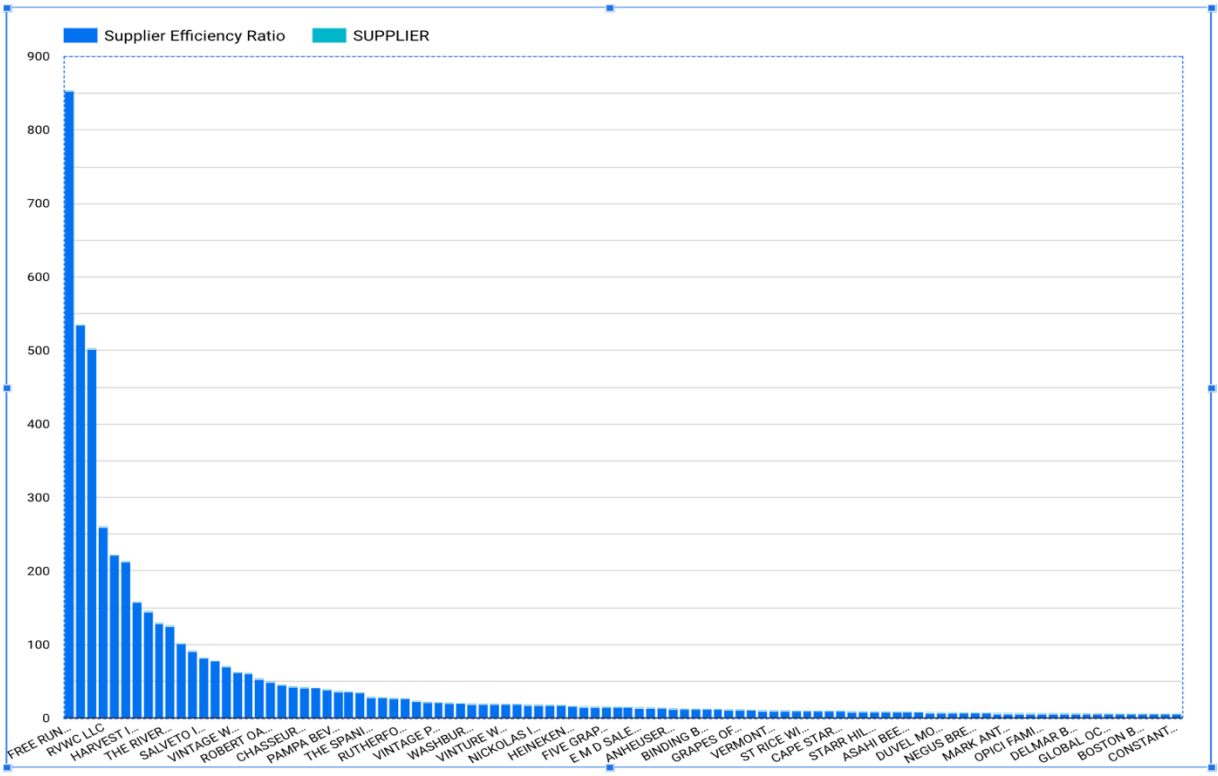
Supplier Warehouse Sales Total:

The objective of the Supplier Warehouse Sales Total analysis is to ascertain the total sales generated in the warehouse that are attributed to each supplier.



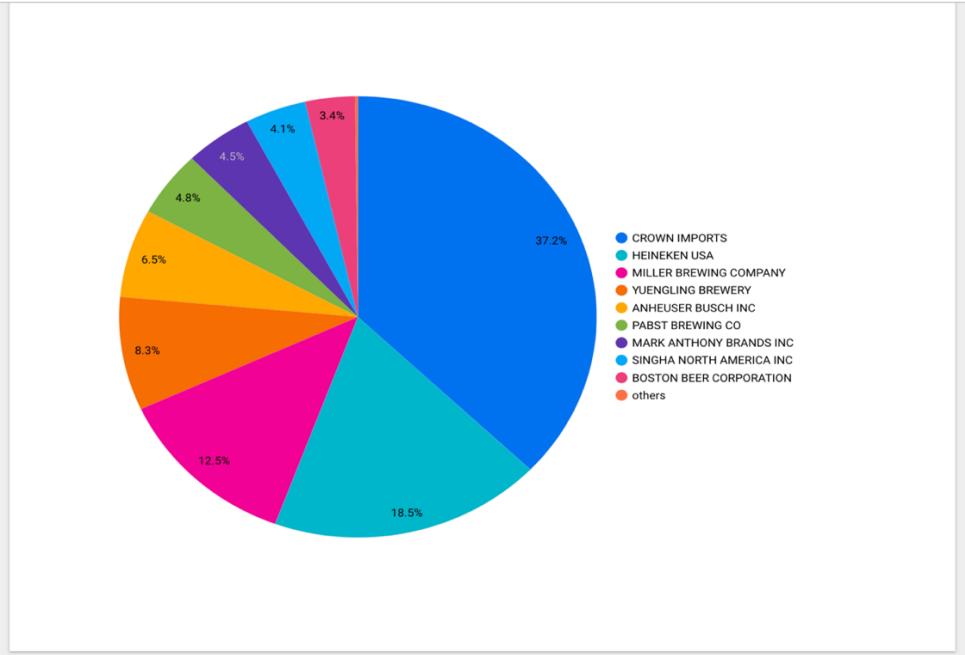
Supplier Efficiency Ratio:

The objective of the Monthly Sales Summary is to provide a comprehensive overview of sales for each month across different years.



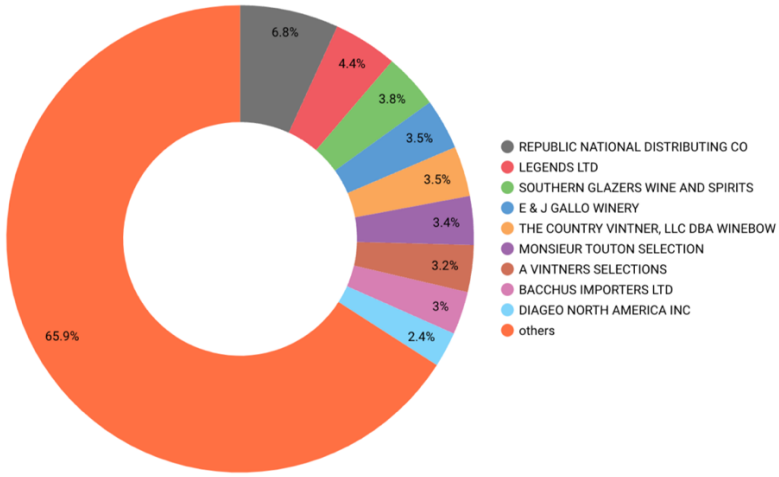
Average Warehouse Sales per Supplier:

The objective of calculating the "Average Warehouse Sales per Supplier" is to assess and compare the performance of suppliers by computing the average sales each supplier contributes to the warehouse.



Average Monthly Transfers

The objective of the "Average Monthly Transfers" analysis is to determine the average number of items transferred to retail outlets each month.



Interpretation of analysis:

Understanding Data Patterns: By studying the data visualization outputs, such as those from Google Data Studio, the team identifies patterns and trends in sales, stock levels, and supplier performance.

Identifying Key Insights: This includes recognizing top-selling items, understanding seasonal sales variations, and assessing monthly sales summaries. These insights are critical for strategic planning and operational adjustments.

Evaluating Supplier Performance: The data helps in assessing the reliability and efficiency of suppliers. This evaluation influences procurement strategies, ensuring better inventory control and cost management.

Analyzing Sales and Stock for Different Item Types: The analysis provides a detailed view of how different items perform in terms of sales and stock levels, aiding in optimal stock management and marketing strategies.

Actionable Decision-Making: Based on these insights, the report likely suggests specific actions to improve inventory management. This might include recommendations for stock replenishment, identifying potential areas for cost reduction, and strategies for improving sales and customer satisfaction.

Future Planning: The insights derived from the analysis also inform long-term strategic decisions, like exploring new market opportunities, adjusting inventory levels based on seasonal trends, and enhancing supplier relationships.