

**Final Project Report**  
**Return Policy Evaluation**  
**Company – Sephora**

Mai Ngo

June 11, 2024

# Tables of Contents

<u>Company Background</u> .....	1
<u>Business Objectives &amp; Imperatives</u> .....	2
<u>Problem</u> .....	3
<u>Process Map</u> .....	4
<u>Dimension Model Diagram</u> .....	6
<u>Data Quality</u> .....	7
Error Checking .....	7
Standards .....	7
<u>Architecture</u> .....	8
End to End Architecture .....	8
Source Systems .....	8
ETL Components .....	9
BI Tools .....	10
BI Users .....	11
<u>Dashboard &amp; Report</u> .....	13
<u>Kimball's Roadmap</u> .....	15
<u>Submission Format</u> .....	16

## Company Background

Sephora is a leading French multinational omni-retailer specializing in beauty products, including cosmetics, skincare and hair care. Founded in Limoges, France, in 1969, Sephora has grown to dominate the beauty retail sector with a global presence. Today, Sephora operates over 2,700 stores across 35 countries, offering a diverse range of over 300 brands along with its own private label. The company's innovative and experiential retail approach sets it apart from competitors. The stores are designed to be interactive beauty hubs where customers can test products, receive personalized recommendations, and attend beauty workshops. This hands-on shopping experience is complemented by Sephora's robust online platform, which provides a seamless omni-channel experience for customers worldwide.

In the United States, Sephora's headquarters is located in San Francisco, California. The US market is a significant part of Sephora's operations, with numerous stores strategically placed across the country in prime retail locations.

In addition, Sephora is a subsidiary of LVMH Moët Hennessy Louis Vuitton, a renowned luxury goods conglomerate headquartered in Paris. LVMH's portfolio includes prestigious brands across various sectors, enhancing Sephora's standing in the luxury beauty market. Under LVMH's ownership, Sephora continues to expand its global footprint and innovate within the beauty industry, maintaining its position as a market leader.

# **Business Objectives & Imperatives**

## ***Objectives***

The beauty industry is a highly lucrative market driven by a loyal customer base of makeup and skincare enthusiasts who are continually seeking new products. Sephora effectively meets this demand by offering a diverse retail space stocked with a wide variety of products from popular and well-known beauty brands. The company is also recognized for its commitment to diversity and inclusion across all aspects of its operations.

Despite these strengths, Sephora faces several in-store customer issues. To further enhance its appeal, Sephora employs a highly incentivized membership scoring system that offers attractive rewards and gifts, encouraging customers to make additional purchases. This strategy, coupled with the inherently enticing nature of beauty products, often leads to impulse buys, resulting in customers spending more than they initially planned. Consequently, Sephora enjoys higher average receipts per customer.

Additionally, using the data to optimize inventory levels based on return patterns and sales forecasts can reduce costs and improve turnover ratios. Lastly, utilizing customer data to create personalized marketing campaigns and promotions can improve customer engagement, retention, and lifetime value, leading to higher revenue.

## ***Imperatives***

This problem derived from the existing product return policy of Sephora. The company's dominant products are high-end make-up and personal care. The main objective of this DW is to tackle the product return operation of a Sephora, how it effect their inventory level and provide meaningful statistic to make improvement on their return policy, customer relationship, and supplier product contract. Since Sephora acts as an intermediary between beauty companies and customers, this system aims to evaluate the status and values of returning products based on the nature/feature of the products, categories, purchasing window, value, price, etc. Since these products are perishable, there will be cases where customers may have already used the product thus, determine whether it can be resale again. Ultimately, the bigger picture is to account the relationships between Sephora to both customers and suppliers. For example, customer have high product returning patterns, or the product is actually bad.

## Problem

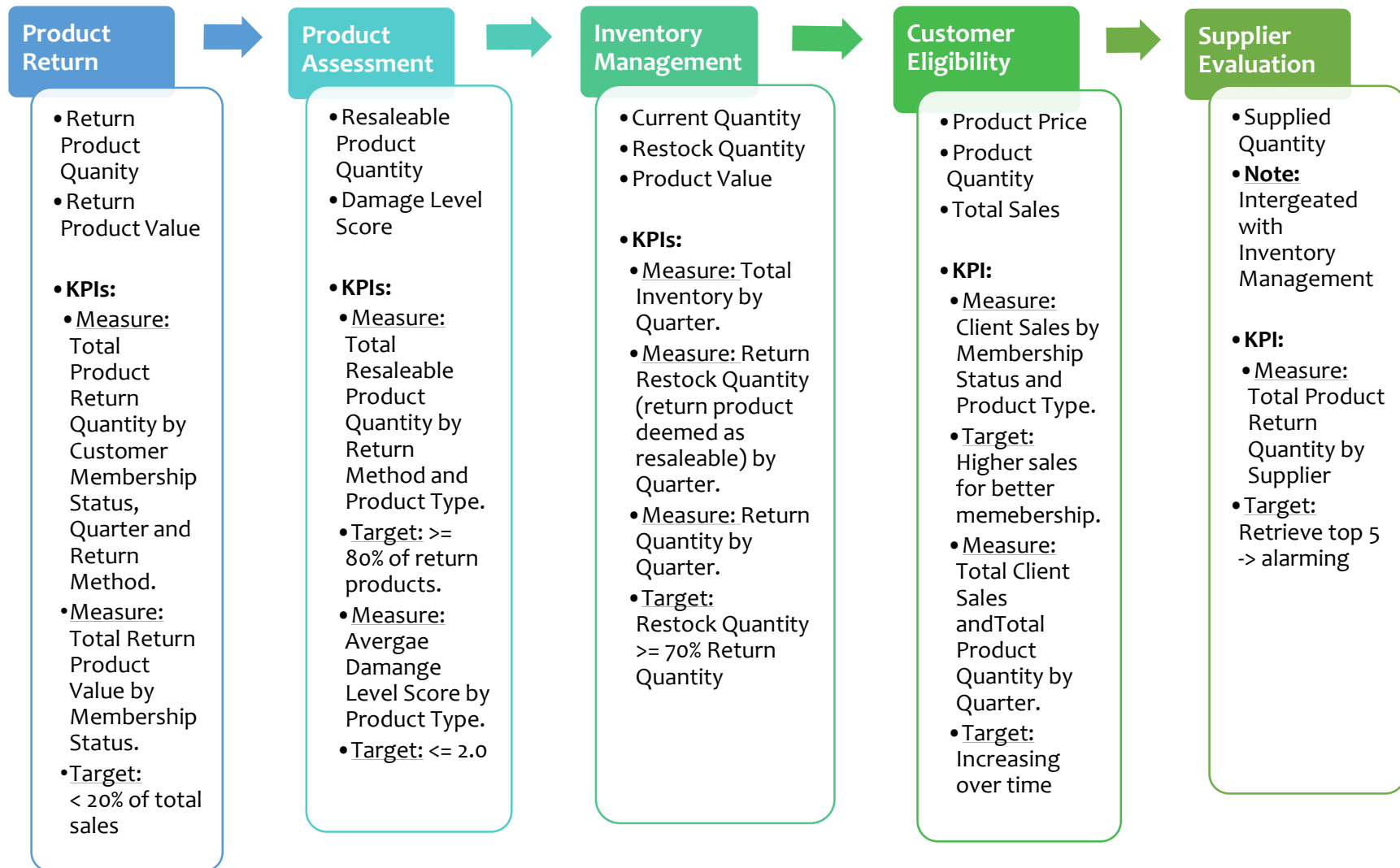
Due to volatile purchasing volume and aim to boost sale, Sephora then offer product returns which ultimately created controversy that the company was not being transparent enough towards their consumers in the return policy. Some customers received warnings of potential bans when returning products in store, the policy does not provide sufficient clarification on the reason why or which criteria decides their ban status, leaving customers feel confused and frustrated. To address this issue, the company will need to create a data warehouse focused on tracking returned products. This database will primarily provide insights into the product categories being returned and whether these returns originated from consumers or beauty suppliers side. Using this data, Sephora can make better decision to refine and transform their return policy, ultimately maintaining their client base.

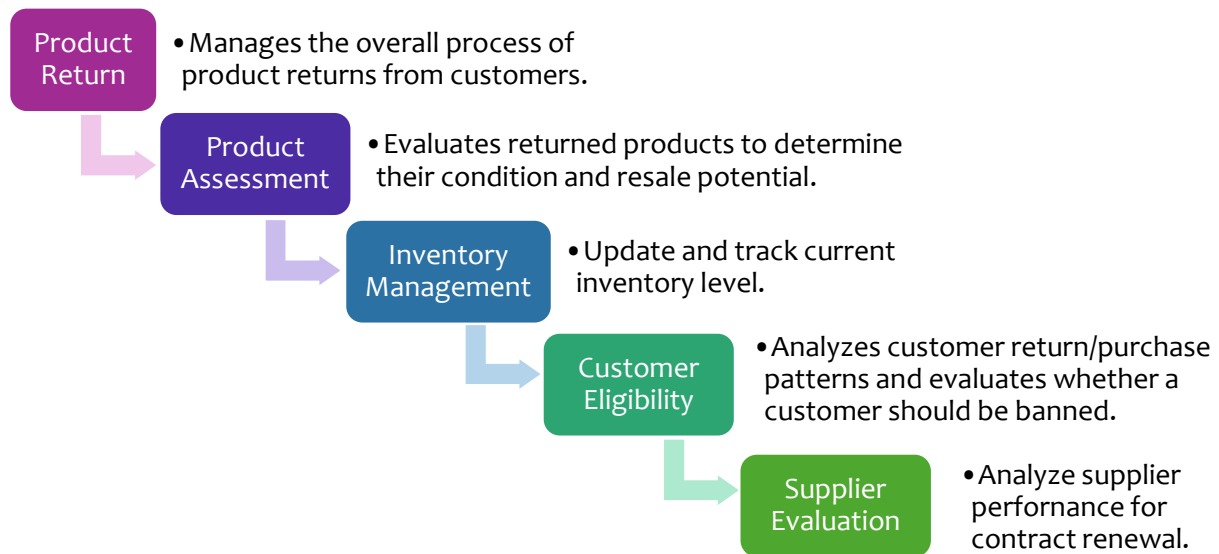
Another problem to be considered is whether a product was returned online or in-store. By including returning information as an additional dimension table, we can further tackle this issue since sometimes store staff tends to boost their sales performance by promising the product is returnable within a two-week period even if intensively used. This could jeopardize Sephora's inventory turnover ratio and lead to return policy abuse. Having store data would allow decision-makers to determine whether this problem arises from the consumer or company side. Furthermore, since we are looking at the problem and creating data at a micro level, including survey data on customer experience with the product and services can provide valuable analytical insights. This can help in understanding customer satisfaction, improving product offerings, and enhancing service quality.

By addressing these areas, the data model can serve as a powerful strategic business decision tool, helping Sephora improve operational efficiency, enhance customer satisfaction, and stay competitive in the market. The two main focuses are:

- 1) Customer Satisfaction and Loyalty** - Lack of comprehensive understanding of customer satisfaction and the factors influencing repeat purchases, returns, and loyalty.
- 2) Supplier Performance Evaluation** - Challenges in evaluating supplier performance and managing supplier relationships effectively.

## Process Map





### **Identify Common Dimensions:**

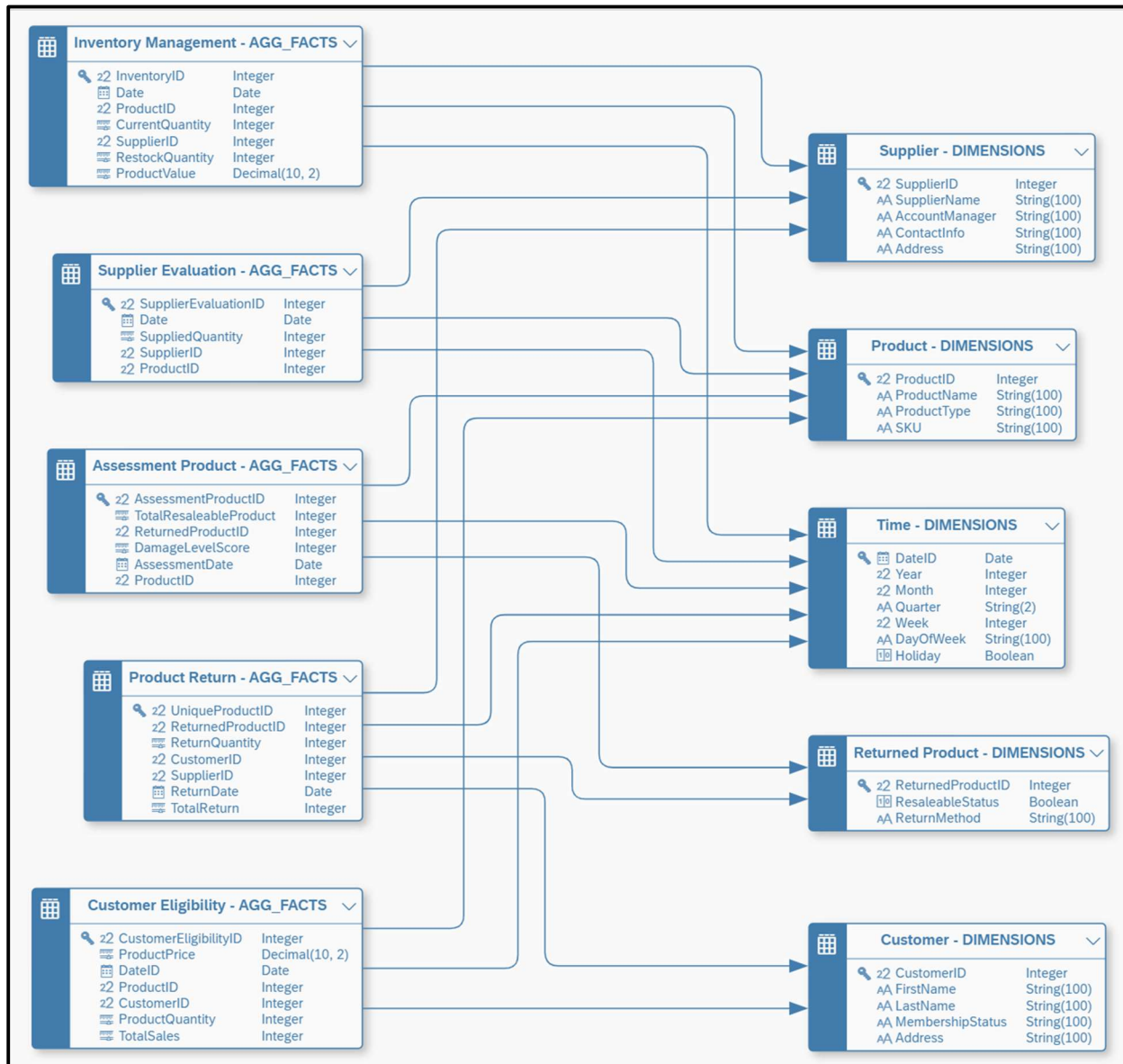
	Time	Returned Product	Product	Supplier	Customer
<b>Product Return</b>	✓	✓		✓	✓
<b>Assessment Product</b>	✓	✓	✓		
<b>Inventory Management</b>	✓		✓	✓	
<b>Customer Eligibility</b>	✓		✓		✓
<b>Supplier Evaluation</b>	✓		✓	✓	

### **Naming Convention Note:**

Per our discussion, since SAP allows users to set up default aggregation methods for measures (for example: Sales, Quantity) in fact tables. I have utilized this feature for convenience, reducing my KPIs calculation steps.

With that being said, 'TableName – FACTS' for fact tables are all named as 'TableName – AGG\_FACTS' table to correctly reflect this set up.

# Dimension Model Diagram



## Data View:

Data Preview: Product - DIMENSIONS (300)			
ProductID	ProductName	ProductType	SKU
1	Foundation	Makeup	1IHYOLXI
2	Serum	Skincare	B50GBOD5
3	Highlighter	Makeup	ILJCVBZH
4	Conditioner	Haircare	2HIXEKDV
5	Texture Spray	Haircare	W56H2A4J
6	Texture Spray	Haircare	AY0JNAN6



## Data Quality

### *Error Checking – using ETL Tools*

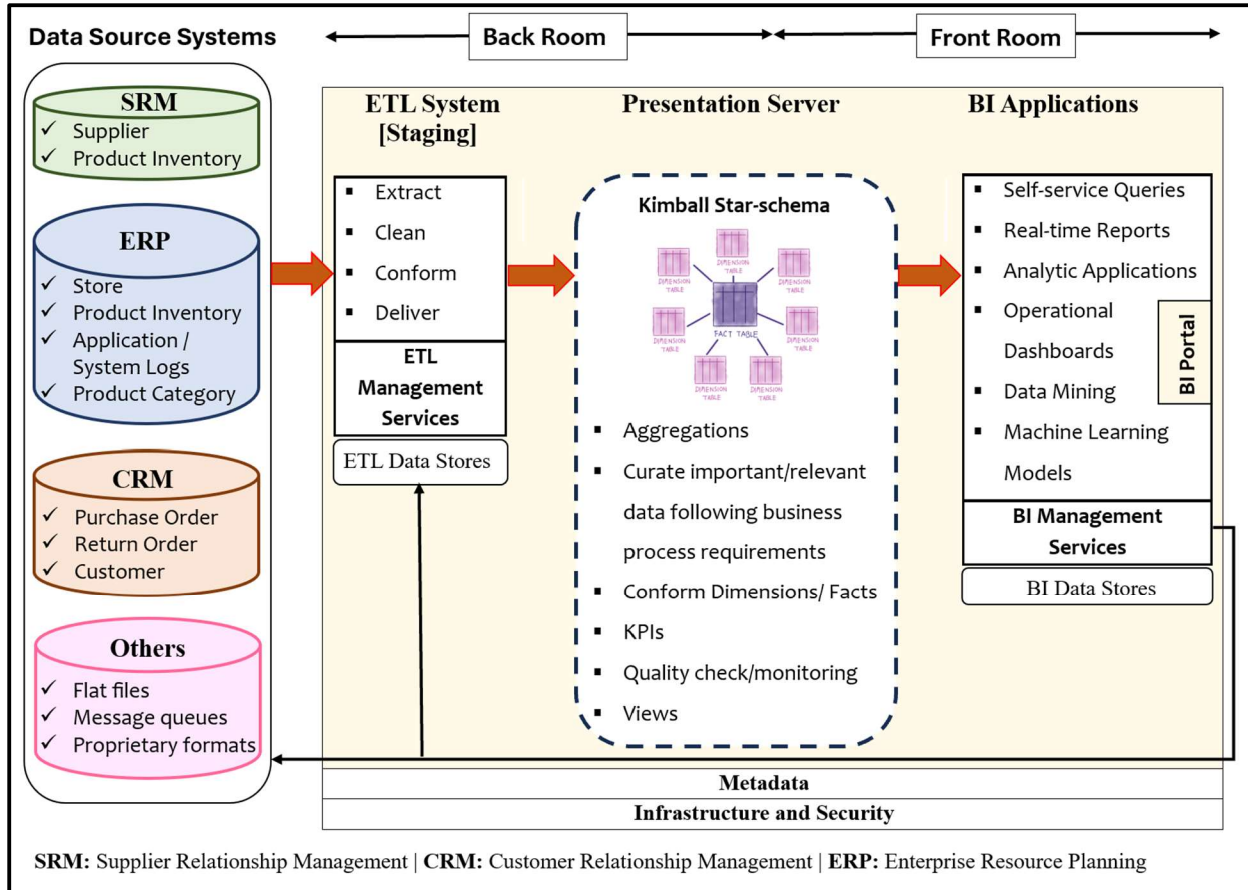
Errors	Description
<b>Date Format</b>	Dates are conformed to the expected format (e.g., yyyy/mm/dd). ☞ Use regular expressions to validate and transform date formats.
<b>Data Type</b>	Data matches the expected data type of each attribute.
<b>Missing Values</b>	Identify and handle missing values by replacing them with default values indicates in DWC. ☞ <u>Example:</u> Number is '99999', Decimal is '9999.99'
<b>Data Length</b>	Data do not exceed predetermined length.
<b>Range and Constraint Validation</b>	Ensure numerical values fall within acceptable digit ranges and foreign keys integrity are maintained. ☞ <u>Example:</u> Product Price are within a realistic range, correct currency and unique table IDs.
<b>Referential Integrity</b>	Foreign keys in fact tables reference to primary keys in dimension tables are consistent.
<b>Data Source</b>	From correct source and Dimension tables data are correctly up to date. ☞ <u>Example:</u> Product Type correctly reflect Sephora current product categories.

### *Standards*

Errors	Description
<b>Data and Table Naming Conventions</b>	Clear and consistent. ☞ <u>Example:</u> Prefixes like 'TableName – DIMENSIONS' for dimensions, 'TableName – FACTS' for fact tables and 'TableName – AGG_FACTS' for aggregate fact tables.
<b>Documentation</b>	<b>MUST</b> have an Excel file keeps all table connections and information. ☞ <u>Example:</u> Mapping, foreign – primary keys relationship, data type and standard with example.
<b>Logging and Monitoring</b>	More for ETL tool to keep check of new errors detection for system updates, backup and recovery.
<b>2NF</b>	Keep tables relation in this format to reduce redundancy and dependency.

# Architecture

## *End to End Architecture*

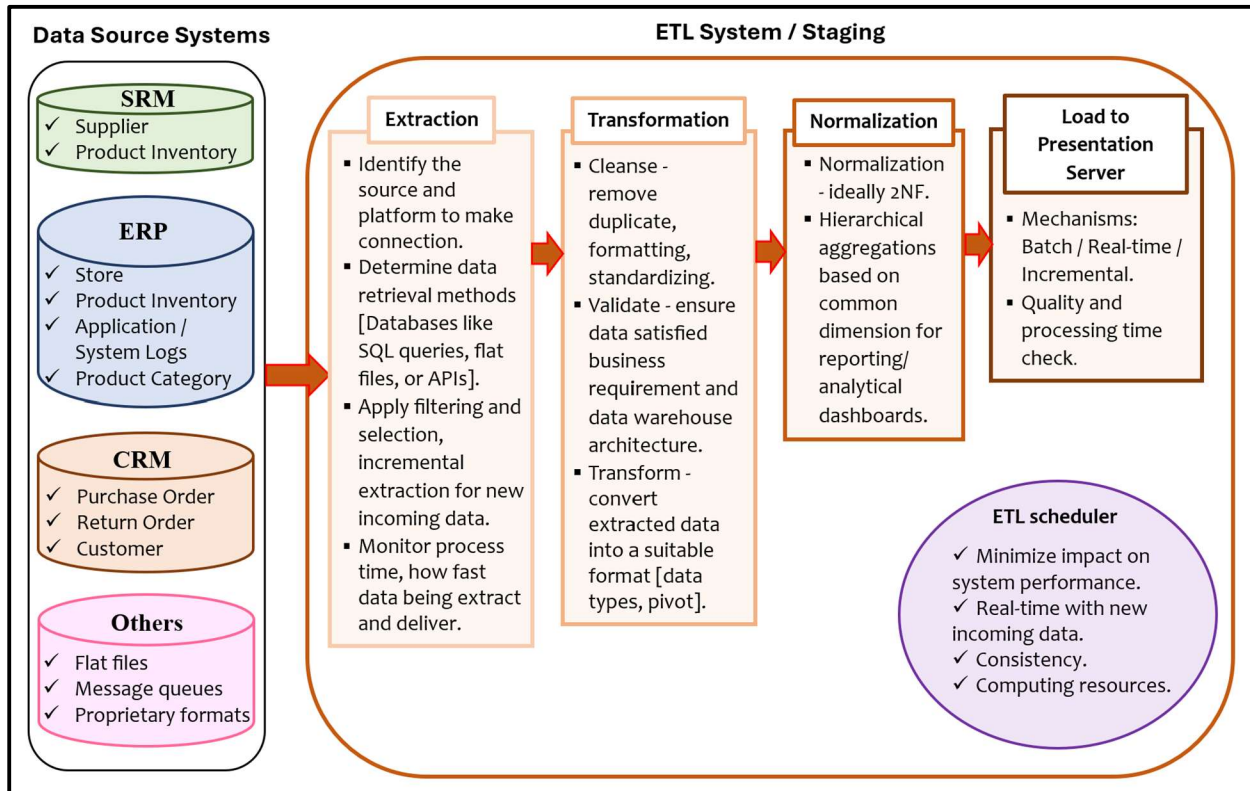


## *Source Systems*

Source (Internal)	Business Owner	IS Owner	Platform	Location	Description
Supplier	Mai N.	Luke P.	SQL	HQ Illinois	Supplier profiles
Customer	Arielis R.	Maddie K.	SQL	HQ Illinois	Customer profiles
Product	Jonathan P.	Gabriel G.	SAP	Product warehouse Texas	Product and Returned Product profiles
Inventory	Amir B.	Michael G.	Linux	Product warehouse Texas	Inventory management system
Sales	Luke P.	Raghav C.	SAP	HQ Illinois	Sale data from both online and in-store

## ETL Components

For ETL – Extract, Load, Transform, data will need to be cleansed to follow format and data structure requirements. The key is to have a selection of data that is meaningful, relevant and important for the business process.



### Key points:

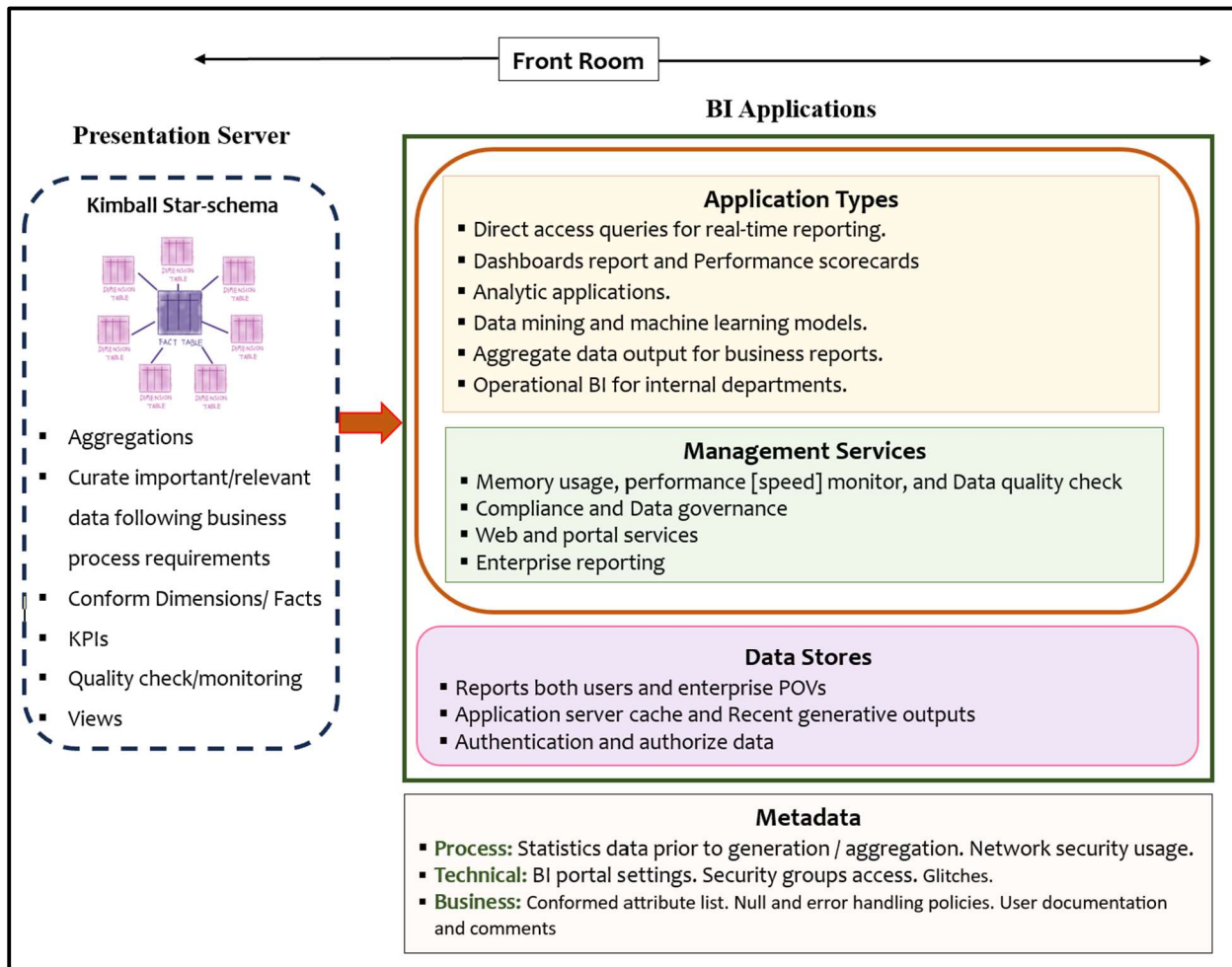
☞ Extract relevant data from the data sources. Relevancy depends on the requirements of data warehouse, which was determined prior to creating Data Architecture diagram. For instance, data would make up a customer/supplier profile.

☞ Clean, load and transform data will be performed using automated ETL tools for efficiency, constantly running to process new data coming in. For instance, date format, reason for product returns.

☞ Apply normalization, ideally 2NF to organize data efficiency, reduce redundancy and maintain integrity.

☞ Define dimensions, considered common dimensions and facts for the data to be loaded in presentation server. Considering KPIs for each business process and aggregation tables, filters if needed. This will help in measuring and analyzing business processes effectively.

## BI Tools



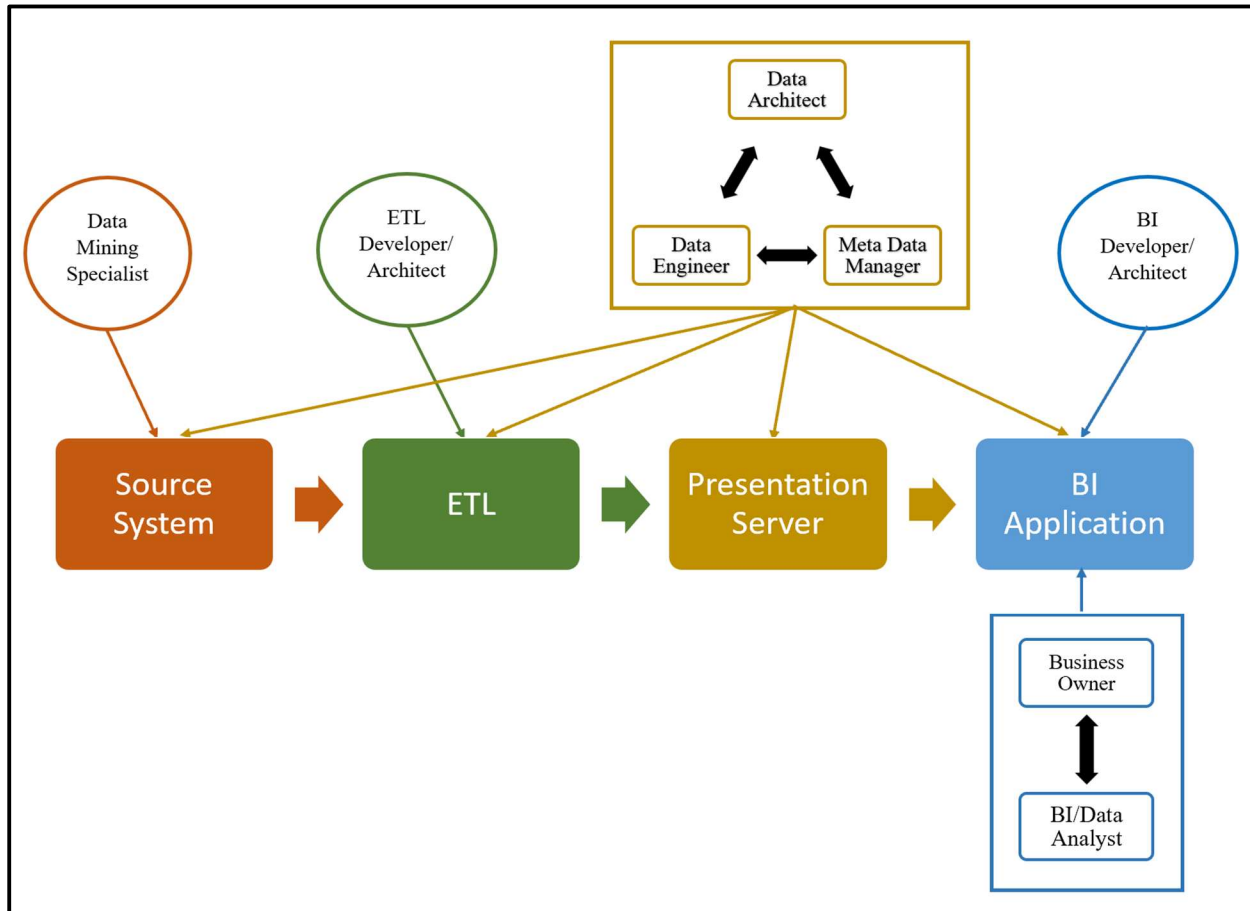
### Key points:

At this stage, the data should be ready for reports and analytic purposes, furthermore with machine learning models for business process development. The business intelligence layer, which mainly for self-service users encompasses BI portal with various application types, BI Management Services, and BI Data Stores all constructed on top of the metadata and security infrastructure. This is a very simple BI architecture diagram; real-world application should expect to be much more complex than this. BI usage for this problem would be:

- ☞ Top returning product, categories and brands.
- ☞ Customer churn, membership distribution and retention.
- ☞ Percentage of product return out of sale volume.

☞ Purchasing history of banned customers → Re-examine returning product threshold to determine banned status.

## **BI Users**



Given big picture of the data warehouse system above, there are some key users who interact across the entire system, while some users have a more designated interactions with specific data stages and uses. While each role has their own focus, the overall goal is to ensure performance efficiency, data integrity and accuracy while align with business process requirements:

☞ **Data Mining Specialist:** Focus on extracting data from source systems using expertise depends on source platform.

☞ **ETL Developer/Architect:** Focus on implementation of ETL tools after necessary data got extracted.

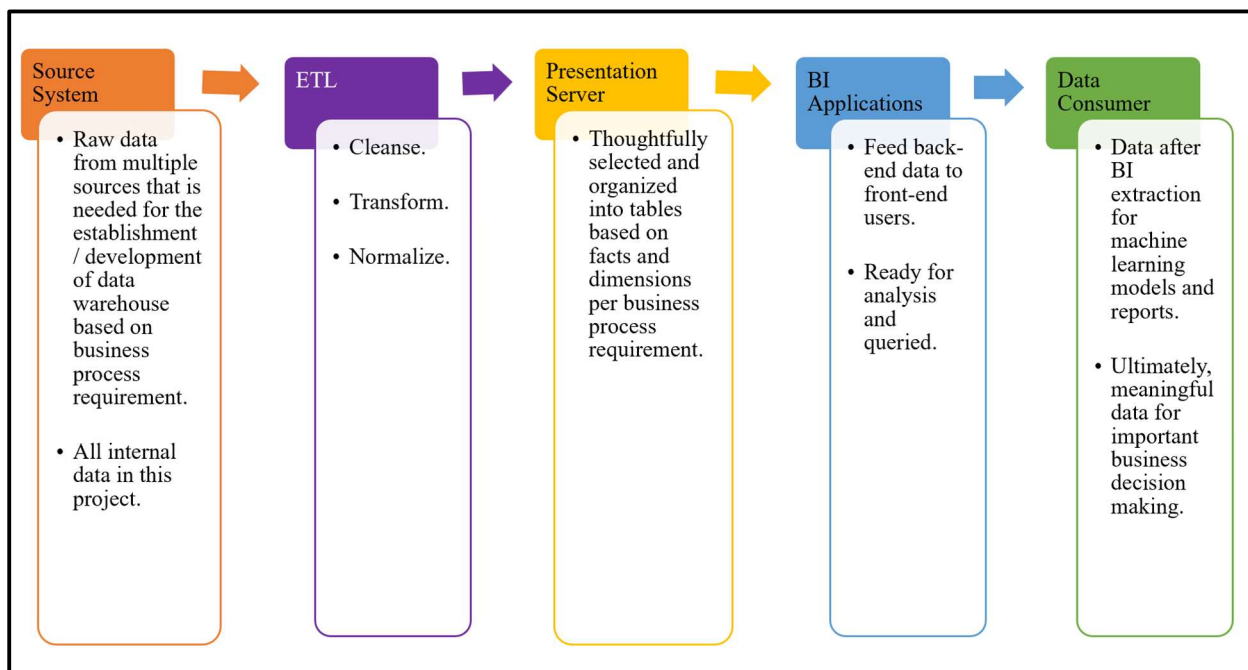
☞ **Data Architect/ Engineer:** Crucial role. Ensure data get into the presentation server and produce out to BI application, create a functional data warehouse system. Most likely will be involved in every step.

☞ **Meta Data Manager:** Focus on data integrity and accuracy throughout the entire end-to-end process.

☞ **BI Developer/Architect:** Focus on implementation and creation of BI tools. Ensure efficiency and user-friendly interface.

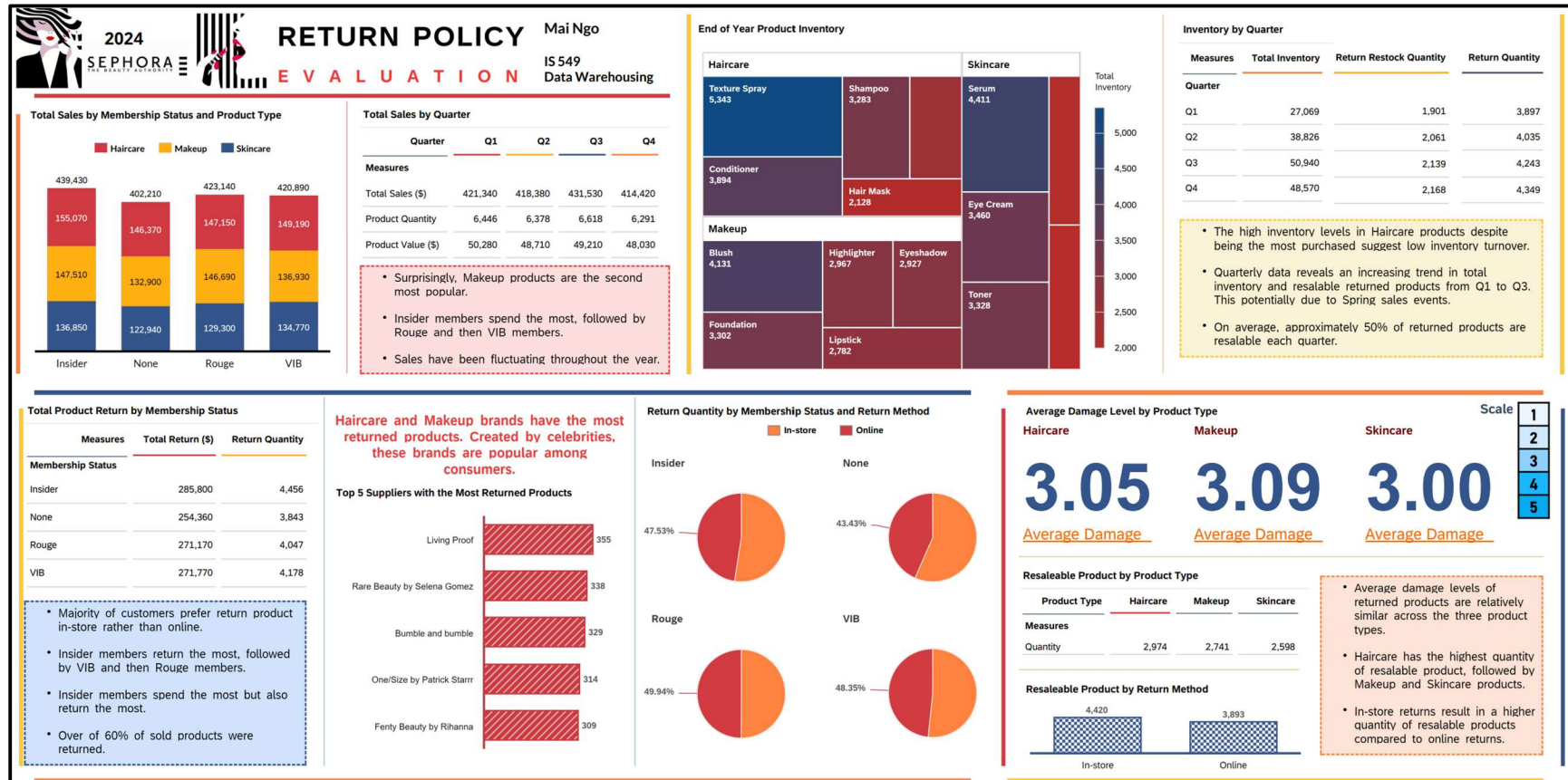
☞ **BI/Data Analyst and Business Owner:** Generate reports and analysis using BI applications to make important business decisions.

### **Data Flows from Ingestion to BI and Consumption:**





# Dashboard and Report



### Total Product Return by Membership Status

Measures	Total Return (\$)	Return Quantity
Membership Status		
Insider	285,800	4,456
None	254,360	3,843
Rouge	271,170	4,047
VIB	271,770	4,178

- Majority of customers prefer return product in-store rather than online.
- Insider members return the most, followed by VIB and then Rouge members.
- Insider members spend the most but also return the most.
- Over of 60% of sold products were returned.

### Haircare and Makeup brands have the most returned products. Created by celebrities, these brands are popular among consumers.

#### Top 5 Suppliers with the Most Returned Products

Supplier	Returned Products
Living Proof	355
Rare Beauty by Selena Gomez	338
Bumble and bumble	329
One/Size by Patrick Starr	314
Fenty Beauty by Rihanna	309

### Return Quantity by Membership Status and Return Method

Membership Status	In-store	Online
Insider	47.53%	
None	43.43%	
Rouge	49.94%	
VIB	48.35%	

### Average Damage Level by Product Type

Product Type	Average Damage
Haircare	3.05
Makeup	3.09
Skincare	3.00

#### Resalable Product by Product Type

Product Type	Haircare	Makeup	Skincare
Measures			
Quantity	2,974	2,741	2,598

#### Resalable Product by Return Method

Return Method	Resalable Product
In-store	4,420
Online	3,893

- Average damage levels of returned products are relatively similar across the three product types.
- Haircare has the highest quantity of resalable product, followed by Makeup and Skincare products.
- In-store returns result in a higher quantity of resalable products compared to online returns.

## Department and Users:

- 👉 **Departments:** Sales | Supply Chain | Operation Management | **Users:** Department managers, data analysts, and management.
- 👉 Certainly, other department can get involved as well depends how they want to tackle return policy and extra data information.

### Viewing format:

☞ View the dashboard through Desktop layout.

**\* Source** – Product Return Aggregate fact table, Customer and Return Product dimension tables.

**KPIs:** Return Product Value (\$) and Return Product (quantity).

☞ **Total Product Return by Membership Status:** A table displaying the total return quantity and return percentage for each membership status.

☞ **Return Quantity by Membership Status and Return Method:** Pie charts showing the distribution of return quantities by return method (in-store vs. online) for each membership status.

☞ **Top 5 Suppliers with the Most Returned Products: - Embed Supplier Evaluation fact table here:** A bar chart listing the top suppliers with the highest number of returned products.

**\* Source** – Customer Eligibility Aggregate fact table, Product and Time dimension tables.

**KPI: Client Sales**

☞ **Total Sales by Membership Status and Product Type:** Bar charts displaying sales figures for different product types segmented by membership status.

☞ **Total Sales by Quarter:** A summary table showing total sales, product quantity, and product value broken down by quarter.

**\* Source** – Inventory Management Aggregate fact table, Time dimension table.

**KPIs:** Restock Quantity, Total Inventory, and Return Quantity

☞ **End of Year Product Inventory:** A tree map visualization showing the inventory levels of various product categories.

☞ **Inventory by Quarter:** A table summarizing total inventory, return restock quantity, and return quantity for each quarter.

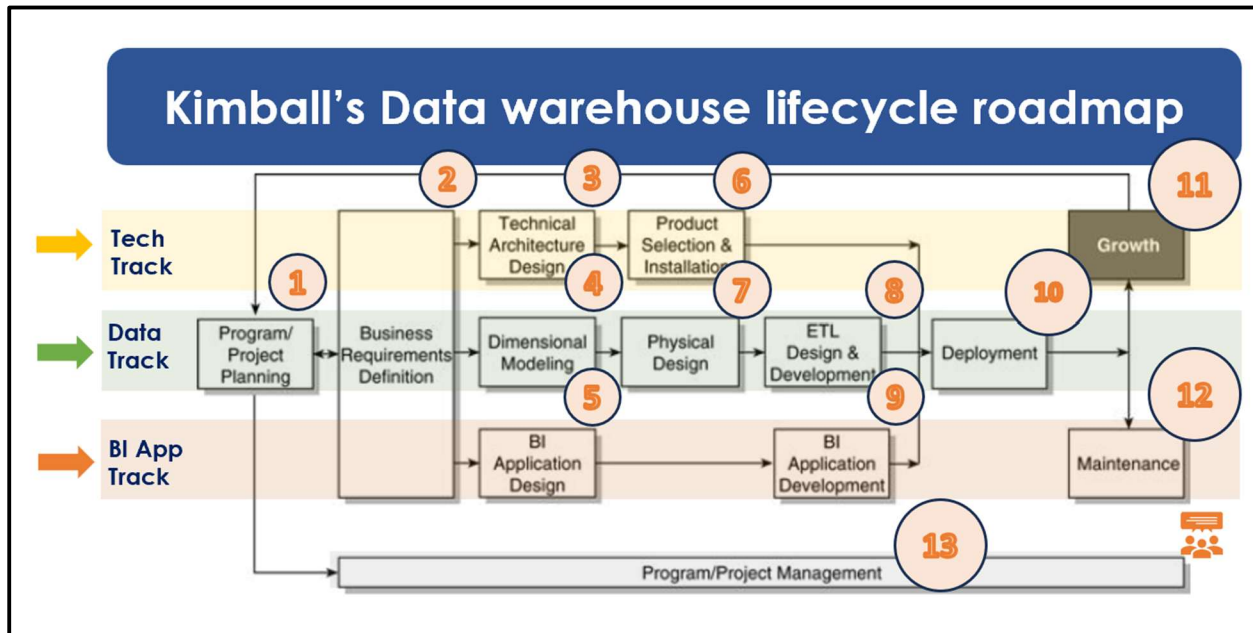
**\* Source** – Assessment Product Aggregate fact table, Product and Return Product dimension tables. **KPIs:** Restock Quantity, Total Inventory, and Return Quantity

☞ **Average Damage Level by Product Type:** Average damage levels for different product types presented in a numerical format with a scale for context.

☞ **Resalable Product by Product Type and Return Method:** Summary tables displaying the quantity of resalable products categorized by product type and return method.



## Kimball's Roadmap



### Understanding key points:

**1. Program/Project Planning:** Define business background, objectives, imperatives and problem of the company Sephora.

**2. Business Requirements Definition:** Gather business process requirements from stakeholders through interviews and meetings, which are made up in this project.

**3. Technical Architecture Design:** Create an End-to-End Architecture diagram of the data warehouse, including additional diagrams to emphasize internal processes like ETL.

**4. Dimensional Modeling:** Design data warehouse schema in DWC using fact and dimension tables which were defined in an Excel file with relations.

**5. BI Application Design:** Design BI dashboards and reports set up to display pre-determined KPIs during business process and activities set up.

**6. Product Selection & Installation:** Consider the uses of ETL tools and additional resources which were made up in this project.

**7. Physical Design:** Create logical design from dimensional model, which are creating and adding data in DWC, consider data schema and indexing – primary keys.

**8. ETL Design & Development:** Design and develop ETL implementation based on potential errors and standards.

**9. BI Application Development:** Implement BI Application Design, specifically creating dashboard and reports, ensuring correctly showcase pre-determined KPIs and tabular.

**10. Deployment:** Testing and training the data warehouse and BI applications to users – using format, including creating manual which are not included in this project.

**11. Growth:** Account new data sources and expansion of the data warehouse.

**12. Maintenance:** Monitor, logging system performance and troubleshooting, account also for ETL tool performance.

**13. Program/Project Management:** Manage efficiency, resources, risks and develop contingency plans.