

User Story for Data Engineers

Title: Create a Comprehensive Table Joining Customer, Shipping, and Order Tables

As a Senior Data Analyst,
I want a table that joins the Customer, Shipping, and Order tables,
So that I can efficiently generate reports and insights that align with the business’s analytical requirements.

Description:

The goal is to create a new table that integrates data from the Customer, Shipping, and Order tables. This table should allow us to perform analysis and generate reports based on customer purchases, shipping statuses, and order details.

Acceptance Criteria:

- 1. **Table Structure:**
 - The new table should include the following columns:

Column Name	Source Table	Description	Data Type
Customer_ID	Customer	Unique identifier for the customer	Integer
First	Customer	Customer’s first name	String
Last	Customer	Customer’s last name	String
Age	Customer	Customer’s age	Integer
Country	Customer	Customer’s country	String
Order_ID	Order	Unique identifier for the order	Integer
Item	Order	Name of the item purchased	String
Amount	Order		Decimal

Column Name	Source Table	Description	Data Type
		Total amount spent on the order	
Shipping_ID	Shipping	Unique identifier for the shipping	Integer
Status	Shipping	Shipping status (Delivered/Pending)	String

2. Data Integrity:

- Ensure that the Customer_ID field is consistent across all three tables and serves as the key to join these tables.
- Ensure that Order_ID and Shipping_ID are correctly mapped to their respective Customer_ID.
- Address any anomalies identified in the existing data, such as missing order information for certain shipments and discrepancies between orders and shipments.
- Detailed Information is available [here](#).

3. Business Logic:

- The Item should remain in the Orders table, even though the analysis suggested it could be moved to a separate table.
- Amount in the Orders table is considered the unit order amount.

4. Performance:

- The table should be optimized for query performance as it will be used frequently for generating business reports.

5. Data Validation:

- Ensure that there are no null values in primary or foreign key fields.
- Validate that all numeric fields (Amount, Age) fall within expected ranges.
- Ensure data consistency across all three tables before creating the final joined table.

Priority: High

This table will be foundational for our analytical work, enabling the reporting and insights required by the business, including total amounts spent by customers, shipping status analysis, and popular products by customer demographics.