

Informatica ETL Project

Overview

This Informatica project involves Extracting, Transforming, and Loading (ETL) data from a CSV file: superstore_data.csv

Tasks

1. Data Cleaning

- Cleanse the data from the source table and store it in a Cleaned data table.

2. Sales Summary

- Summarize total sales and average sales for each customer. Identify customers with significant contributions to overall sales.

3. Customer Order Analysis

- Analyze customer orders to determine the most frequent buyers and their order patterns.

4. Customer Geography Analysis

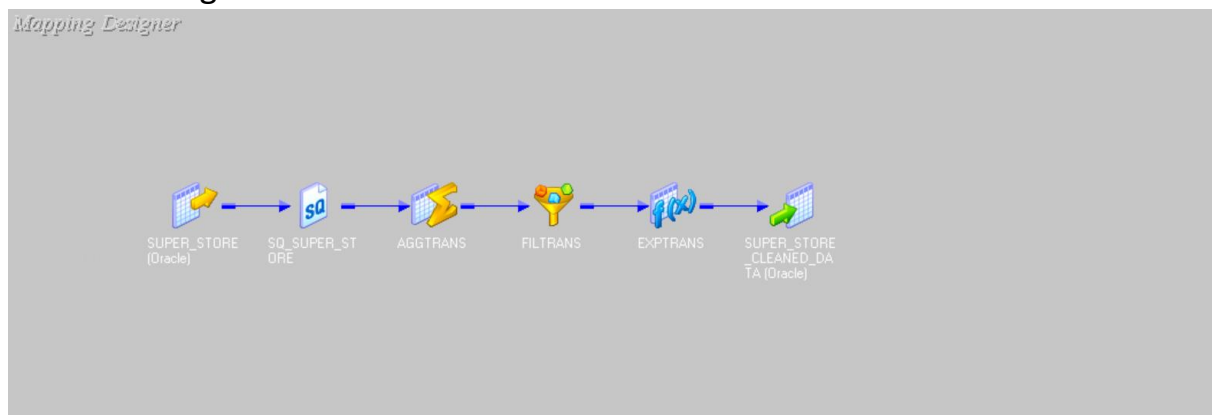
- Analyze customer distribution across different regions to identify potential market segments.

5. Order Processing Time Analysis

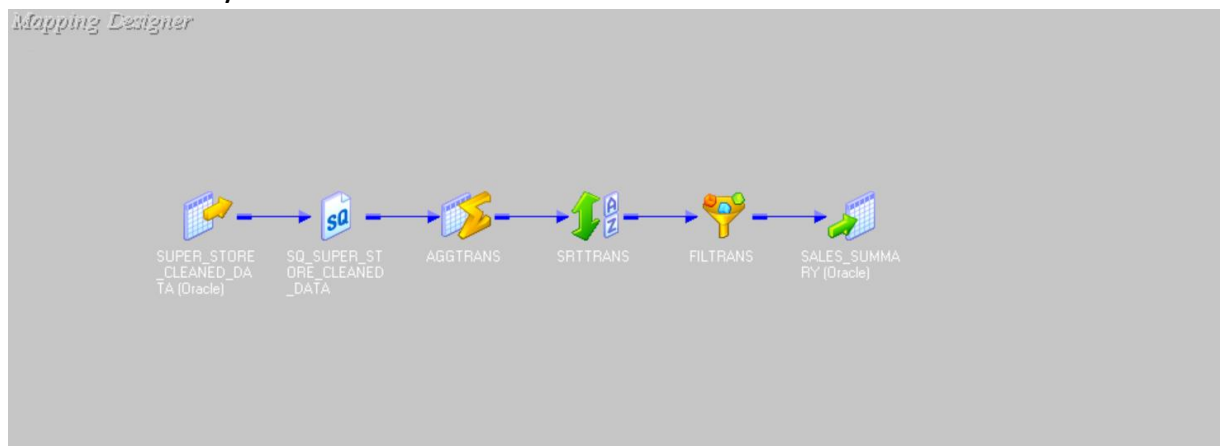
- Evaluate order processing efficiency by analyzing the time taken between order placement and shipment.

Screenshots

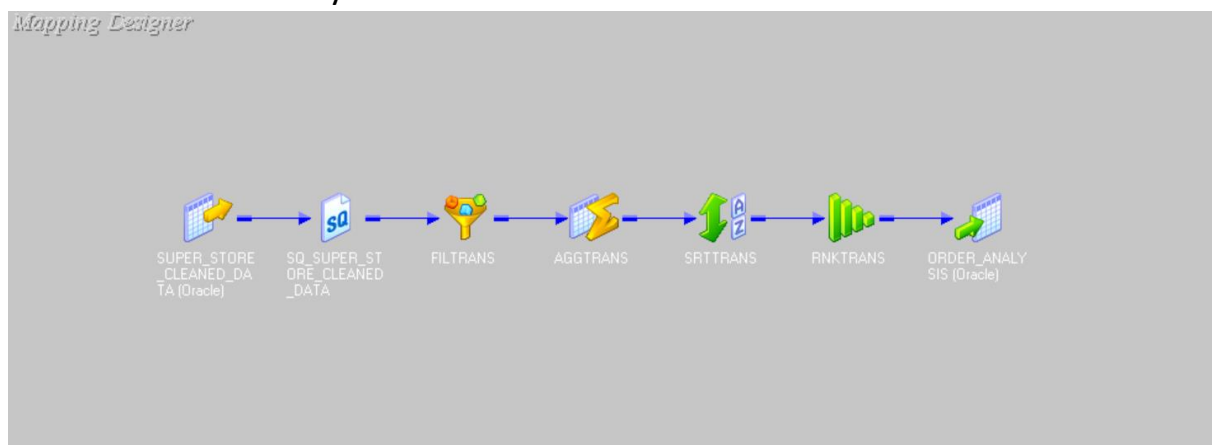
1. Data Cleaning



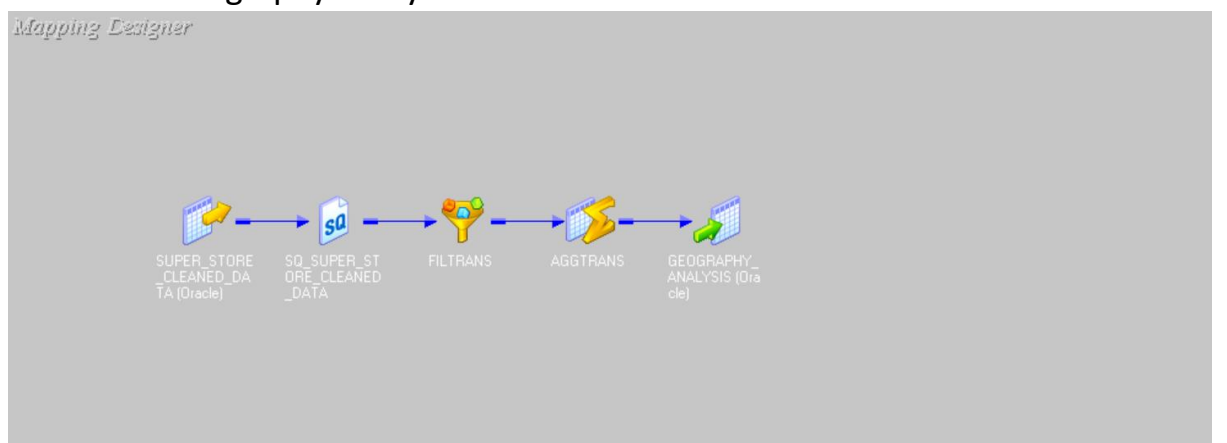
2. Sales Summary



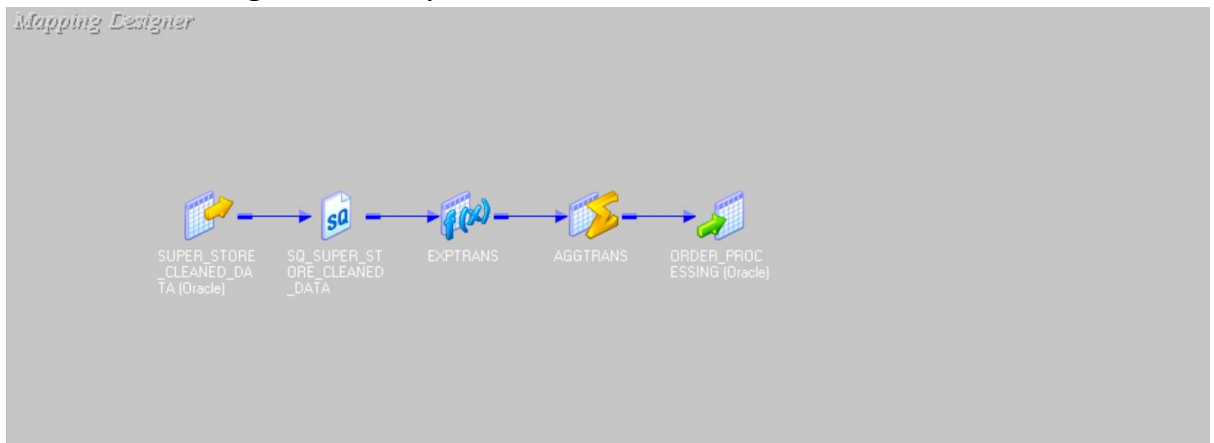
3. Customer Order Analysis



4. Customer Geography Analysis



5. Order Processing Time Analysis



OUTPUTS

1. Super Store Cleaned Data:

CUSTOMER_ID_NAME
21925-Zushuss Donatelli
16585-Ken Black
21520-Tracy Blumstein

2. Sales Summary:

CUSTOMER_ID_NAME || TOTAL_SALES || AVG_SALES
11140-Becky Martin || 10540 || 1506
14635-Grant Thornton || 8167 || 4084
20290-Sean Braxton || 5580 || 2790

3. Order Analysis:

CUSTOMER_ID_NAME || ORDERS_COUNT
15520-Jeremy Pistek || 4
16510-Keith Herrera || 4

4. Geography Analysis:

CUSTOMER_ID_NAME || REGION || STATE || REGION_ORDERS_COUNT
14095-Eudokia Martin || West || California || 182

5. Order Processing:

CATEGORISE_PROCESSING_DAYS || ORDERS_COUNT
Immediate delivery || 32
Long Term delivery || 641

Usage

- Ensure Informatica PowerCenter is installed and configured.
- Import the CSV files superstore_data.csv into Informatica PowerCenter.
- Create mappings, transformations, and workflows to perform the specified tasks.
- Execute the workflows to generate the output files mentioned above.

Note

- Ensure the input CSV files are correctly mapped within Informatica PowerCenter.
- Validate the transformations and mapping to ensure accurate data processing.
- Monitor the workflow execution for any errors or issues.
- Adjust any parameters or configurations as needed to optimize performance and accuracy.