

Informatica Challenge Kick off Jan-25

1. Oracle SQL Setup

- **Table Creation Queries: 1.1 Superstore Table**

```
CREATE TABLE Superstore (  
    Category VARCHAR(255),  
    City VARCHAR(255),  
    Country_Region VARCHAR(255),  
    Customer_ID VARCHAR(255),  
    Customer_Name VARCHAR(255),  
    Order_Date VARCHAR(255),  
    Order_ID VARCHAR(255),  
    Order_Profitable VARCHAR(255),  
    Postal_Code NUMBER,  
    Product_ID VARCHAR(255),  
    Product_Name VARCHAR(255),  
    Region VARCHAR(255),  
    Row_ID NUMBER,  
    Segment VARCHAR(255),  
    Ship_Date VARCHAR(255),  
    Ship_Mode VARCHAR(255),  
    Ship_Status VARCHAR(255),  
    State VARCHAR(255),  
    Sub_Category VARCHAR(255),  
    Days_to_Ship_Actual NUMBER,  
    Days_to_Ship_Scheduled NUMBER,  
    Discount VARCHAR(255),  
    Profit NUMBER,  
    Quantity NUMBER,  
    Sales NUMBER,  
    Sales_Forecast NUMBER  
);
```

1.2 Superstore_Cleaned_Data Table

```
CREATE TABLE Superstore_Cleaned_Data (  
    Category VARCHAR(255),  
    City VARCHAR(255),  
    Customer_ID VARCHAR(255),  
    Customer_Name VARCHAR(255),
```

```

    Order_Date VARCHAR(255),
    Order_ID VARCHAR(255),
    Order_Profitable VARCHAR(255),
    Product_ID VARCHAR(255),
    Product_Name VARCHAR(255),
    Region VARCHAR(255),
    Row_ID NUMBER,
    Segment VARCHAR(255),
    Ship_Date VARCHAR(255),
    Ship_Mode VARCHAR(255),
    Ship_Status VARCHAR(255),
    State VARCHAR(255),
    Sub_Category VARCHAR(255),
    Days_to_Ship_Actual NUMBER,
    Days_to_Ship_Scheduled NUMBER,
    Discount VARCHAR(255),
    Profit NUMBER,
    Quantity NUMBER,
    Sales NUMBER,
    Sales_Forecast NUMBER
);

```

1.3 Top_15 Table

```

CREATE TABLE Top_15 (
    Category VARCHAR(255),
    City VARCHAR(255),
    Customer_ID VARCHAR(255),
    Customer_Name VARCHAR(255),
    Order_Date VARCHAR(255),
    Order_ID VARCHAR(255),
    Order_Profitable VARCHAR(255),
    Region VARCHAR(255),
    Discount VARCHAR(255),
    Profit NUMBER,
    Quantity NUMBER,
    Sales NUMBER,
    Total_Profit NUMBER
);

```

1.4 Discount_Analysis Table

```

CREATE TABLE Discount_Analysis (
    Category VARCHAR(255),
    City VARCHAR(255),

```

```

Customer_ID VARCHAR(255),
Customer_Name VARCHAR(255),
Order_Date VARCHAR(255),
Order_ID VARCHAR(255),
Order_Profitable VARCHAR(255),
Product_ID VARCHAR(255),
Product_Name VARCHAR(255),      Region VARCHAR(255),
Row_ID NUMBER,
Segment VARCHAR(255),
Ship_Date VARCHAR(255),
Ship_Mode VARCHAR(255),
Ship_Status VARCHAR(255),
State VARCHAR(255),
Sub_Category VARCHAR(255),
Days_to_Ship_Actual NUMBER,
Days_to_Ship_Scheduled NUMBER,
Discount VARCHAR(255),
Profit NUMBER,
Quantity NUMBER,
Sales NUMBER,
Sales_Forecast NUMBER,
Discount_Updated NUMBER,
Avg_Profit NUMBER,
Avg_Sales NUMBER
);

```

2. Data Preparation

- Dataset Location:
~/Desktop/Project/kickoffs-information-tableau-superstore/Superstore.csv
- Load Data: Use Oracle SQL Developer to import Superstore.csv into the Superstore table.

3. Informatica Tasks

Task 1: Cleaned Data

- Mapping Name: m_cleaned_data
- Workflow Name: w_cleaned_data
- Session Name: s_cleaned_data
- Target Table: Superstore_Cleaned_Data
- CSV Name: superstore_cleaned_data.csv

Operations:

1. Import the Superstore table as the source.
2. Remove all rows with null values using Expression Transformation (e.g.,
ISNULL(Column_Name) = 0).
3. Delete columns Country_Region and Postal_Code.
4. Load cleaned data into Superstore_Cleaned_Data.

Sample output

Table count: 977

CATEGORY	CITY	CUSTOMER_ID	CUSTOMER_NAME	ORDER_DATE	ORDER_ID	ORDER_PROFITABLE	PRODUCT_ID
PRODUCT_NAME	REGION	ROW_ID	SEGMENT	SHIP_DATE	SHIP_MODE	SHIP_STATUS	STATE
SUB_CATEGORY	DAYS_TO_SHIP_ACTUAL	DAYS_TO_SHIP_SCHEDULED	DISCOUNT	PROFIT	QUANTITY	SALES	SALES_FORECAST
Furniture	New York City	C0037	Customer 47	2/17/2022	CA-2022-228926	Profitable	FUR-2681
Global Deluxe High-Back Manager's Chair	East	1	Corporate	2/25/2022	Standard Class	Shipped on Time	California
Furniture	4	2	0.15%	93.92	5	434.26	413.98

Task 2: Top 15 Customers by Category

- Mapping Name: m_top_15
- Workflow Name: w_top_15
- Session Name: s_top_15
- Target Table: Top_15
- CSV Name: top_15.csv

Operations:

1. Import Superstore_Cleaned_Data as the source.

2. Split data into three categories: Technology, Office Supplies, Furniture.
3. Filter rows where Order_Profitable = "Profitable".
4. Calculate Total_Profit (sum of Profit) grouped by Customer_Name.
5. Select Top 5 Customers per category based on Total_Profit.
6. Load results into the Top_15 table.

Sample output

Table count: 15

CATEGORY	CITY	CUSTOMER_ID	CUSTOMER_NAME	ORDER_DATE	ORDER_ID	ORDER_PROFITABLE	REGION	DISCOUNT	PROFIT	QUANTITY	SALLES	TOTAL_PROFIT
Office Supplies	Philadelphia	C0075	Customer 11	2/15/2022	CA-2022-546500	Profitable	East	0.01%	127.64	5	855.2	371

Task 3: Discount Analysis

- Mapping Name: m_discount_analysis
 - Workflow Name: w_discount_analysis
 - Session Name: s_discount_analysis
 - Target Table: Discount_Analysis
 - CSV Name: discount_analysis.csv
- Operations:**
1. Import Superstore_Cleaned_Data as the source.
 2. Create Discount_Updated by removing "%" from the Discount column (e.g., 0.00% → 0.00).
 3. Split data into segments: Consumer and Corporate.
 4. Calculate Avg_Profit and Avg_Sales grouped by Region.
 5. Load results into the Discount_Analysis table.

Sample output:

Table count: 8

CATEGORY	CITY	CUSTOMER_ID	CUSTOMER_NAME	ORDER_DATE	ORDER_ID	ORDER_PROFITABLE	PRODUCT_ID	PRODUCT_NAME
----------	------	-------------	---------------	------------	----------	------------------	------------	--------------

REGION	ROW_ID	SEGMENT	SHIP_DATE	SHIP_MODE	SHIP_STATUS	STATE	SUB_CATEGORY	DAYS_TO_SHIP_ACTUAL
DAYS_TO_SHIP_SCHEDULED	DISCOUNT	PROFIT	QUANTITY	SALES	SALES_FORECAST	DISCOUNT_UPDATED	AVG_PROFIT	AVG_SALES
Furniture	Indianapolis	C0087	Customer 92	1/28/2022	CA-2022-619416	Profitable	FUR-6289	Hon Deluxe Fabric Micro Chair
Central	969	Corporate	2/14/2022	Second Class	Shipped Late	Texas	Furniture	3
4	0.05%	143.73	3	172.91	189.03	0	45	554

4. Export CSV Files

1. Export Steps in Oracle SQL Developer:

- Run SELECT * FROM [Table_Name];
- Right-click query results → Export → Format: CSV.
- Save to output folder.

Required CSV Files:

- superstore_cleaned_data.csv
- top_15.csv
- discount_analysis.csv