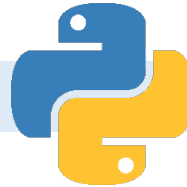


## WEEK 2

### Analysis Questions Phase



#### 1. Introduction

This document outlines the process of conducting a comprehensive data analysis to uncover valuable insights from the provided dataset. The analysis will focus on formulating and answering key questions that are relevant to the organization's decision-making process.

By leveraging SQL and Python libraries like pandas and Matplotlib, we will explore the dataset to identify patterns, trends, and correlations that can inform strategic initiatives and enhance overall business performance.

#### 2. Sales and Revenue

##### 2.1 What is the Impact of Product Category on Sales Performance?

The screenshot shows a SQL query editor window titled 'SQLQuery1.sql - DE...74S3\EI Noby (62))\*'. The query is as follows:

```
SELECT Product_Type,
SUM(S.Revenue_generated) AS Total_Revenue ,
SUM(S.Number_of_products_sold) AS Total_Orders
FROM Sales S
INNER JOIN Products ON S.SKU = Products.SKU
GROUP BY Product_Type
Order By Total_Revenue DESC;
```

Below the query, the 'Results' tab is active, displaying a table with the following data:

	Product_Type	Total_Revenue	Total_Orders
1	skincare	241628.17	20731
2	haircare	174455.42	13611
3	cosmetics	161521.27	11757

## 2.2 How do lead times and shipping times affect sales performance?

SQLQuery1.sql - DE...74S3\EI Noby (62))\*

```
SELECT TOP 10 i.Lead_times, i.Shipping_times,
SUM(s.Revenue_generated) AS total_revenue
FROM Inventory i
JOIN Sales s ON i.sku = s.sku
GROUP BY i.Lead_times, i.Shipping_times
Order By total_revenue DESC;
```

100 %

Results Messages

	Lead_times	Shipping_times	total_revenue
1	5	7	19198.42
2	18	7	19037.37
3	27	8	15405.41
4	17	1	13027.50
5	26	7	11828.05
6	24	8	10766.33
7	6	5	10094.36
8	4	1	9866.47
9	1	5	9692.32
10	17	4	9655.14

## 3. Customer Segmentation

### 3.1 What is Sales Trends Based on Customer Demographics?

SQLQuery1.sql - DE...74S3\EI Noby (62))\*

```
SELECT P.Product_type ,C.Customer_demographics,
SUM(S.Number_of_products_sold) AS Number_Of_Products
FROM Sales S
JOIN Products P ON S.SKU = P.SKU
JOIN Customers C ON C.SKU = P.SKU
GROUP BY P.Product_type, C.Customer_demographics
ORDER BY P.Product_type, Number_Of_Products;
```

100 %

Results Messages

	Product_type	Customer_demographics	Number_Of_Products
1	cosmetics	Male	2304
2	cosmetics	Non-binary	2607
3	cosmetics	Unknown	2834
4	cosmetics	Female	4012
5	haircare	Female	936
6	haircare	Male	2292
7	haircare	Non-binary	2820
8	haircare	Unknown	7563
9	skincare	Male	2911
10	skincare	Unknown	4814
11	skincare	Non-binary	5153
12	skincare	Female	7853

### 3.2 How do customer demographics influence purchasing behavior?

SQLQuery1.sql - DE...74S3\EI Noby (62))\*

```
SELECT C.Customer_demographics,
SUM(S.Revenue_generated) AS Total_Revenue ,
SUM(S.Number_of_products_sold) AS Total_Order
FROM Sales S
JOIN Customers C ON C.SKU = S.SKU
GROUP BY C.Customer_demographics
ORDER BY Total_Revenue;
```

100 %

Results Messages

	Customer_demographics	Total_Revenue	Total_Order
1	Non-binary	116365.81	10580
2	Male	126634.42	7507
3	Female	161514.49	12801
4	Unknown	173090.14	15211

### 3.3 How do customer demographics influence purchasing behavior?

SQLQuery1.sql - DE...74S3\EI Noby (62))\*

```
SELECT C.Customer_demographics,
SUM(S.Revenue_generated) AS Total_Revenue
FROM Sales S
JOIN Customers C ON C.SKU = S.SKU
GROUP BY C.Customer_demographics
ORDER BY Total_Revenue;
```

100 %

Results Messages

	Customer_demographics	Total_Revenue
1	Non-binary	116365.81
2	Male	126634.42
3	Female	161514.49
4	Unknown	173090.14

## 4. Transportation

### 4.1 Which transportation modes are associated with the lowest costs and fastest delivery times?

SQLQuery1.sql - (lo...ply\_chain (sa (54)))

```
SELECT
    t.Transportation_modes,
    AVG(t.Costs) AS Avg_Cost,
    AVG(i.Shipping_times) AS Avg_Shipping_Time
FROM
    Transportation t
JOIN
    Inventory i ON t.SKU = i.SKU
GROUP BY
    t.Transportation_modes
ORDER BY
    Avg_Shipping_Time ASC, Avg_Cost ASC;
```

133 %

Results Messages

	Transportation_modes	Avg_Cost	Avg_Shipping_Time
1	Road	553.386206	4
2	Air	561.711923	5
3	Rail	541.747500	6
4	Sea	417.819411	7

### 4.1 Which Location have the largest amount of order shipping from?

SQLQuery3.sql - DE...74S3\EI Noby (54)\* SQLQuery1.sql - DE...74S3\EI Noby (62))\*

```
SELECT R.Location, sum(S.Number_of_products_sold) as Total_Order
FROM Suppliers R
JOIN Sales S ON S.SKU = R.SKU
GROUP BY R.Location
ORDER BY Total_Order;
```

100 %

Results Messages

	Location	Total_Order
1	Bangalore	5420
2	Chennai	8768
3	Mumbai	9426
4	Delhi	9715
5	Kolkata	12770

## 5. Supplier

### 5.1 Average each Supplier Performance For lead time?

SQLQuery3.sql - DE...74S3\EI Noby (54))\* X SQLQuery1.sql - DE...74S3\EI Noby (62))\*

```
SELECT S.Supplier_name, AVG(S.Lead_time) AS Average_Lead_Time
FROM Suppliers S
GROUP BY S.Supplier_name;
```

100 %

Results Messages

	Supplier_name	Average_Lead_Time
1	Supplier 1	14
2	Supplier 2	18
3	Supplier 3	20
4	Supplier 4	15
5	Supplier 5	18

### 5.1 What are the Location trends in Supplier orders number?

SQLQuery4.sql - DE...74S3\EI Noby (65))\* X SQ

```
SELECT S.Supplier_name, S.Location,
Count(S.SKU) as total_order
FROM Suppliers S
GROUP BY S.Location ,S.Supplier_name
order by S.Supplier_name ,total_order
```

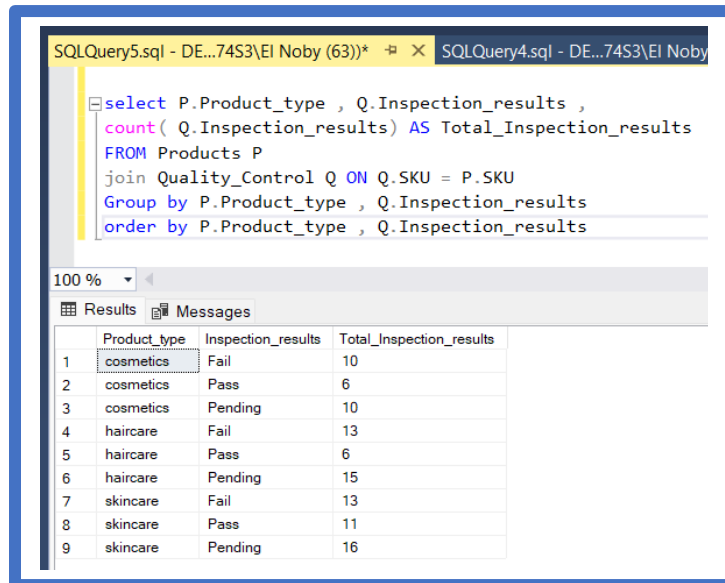
100 %

Results Messages

	Supplier_name	Location	total_order
1	Supplier 1	Chennai	4
2	Supplier 1	Delhi	4
3	Supplier 1	Bangalore	5
4	Supplier 1	Mumbai	6
5	Supplier 1	Kolkata	8
6	Supplier 2	Kolkata	3
7	Supplier 2	Chennai	3
8	Supplier 2	Mumbai	5
9	Supplier 2	Bangalore	5
10	Supplier 2	Delhi	6
11	Supplier 3	Delhi	2
12	Supplier 3	Kolkata	3
13	Supplier 3	Mumbai	3
14	Supplier 3	Bangalore	3
15	Supplier 3	Chennai	4
16	Supplier 4	Bangalore	2
17	Supplier 4	Delhi	2
18	Supplier 4	Chennai	4
19	Supplier 4	Mumbai	4
20	Supplier 4	Kolkata	6
21	Supplier 5	Delhi	1
22	Supplier 5	Bangalore	3
23	Supplier 5	Mumbai	4
24	Supplier 5	Chennai	5
25	Supplier 5	Kolkata	5

## 6. Quality Control

### 6.1 the overall Inspection results regarding each category.



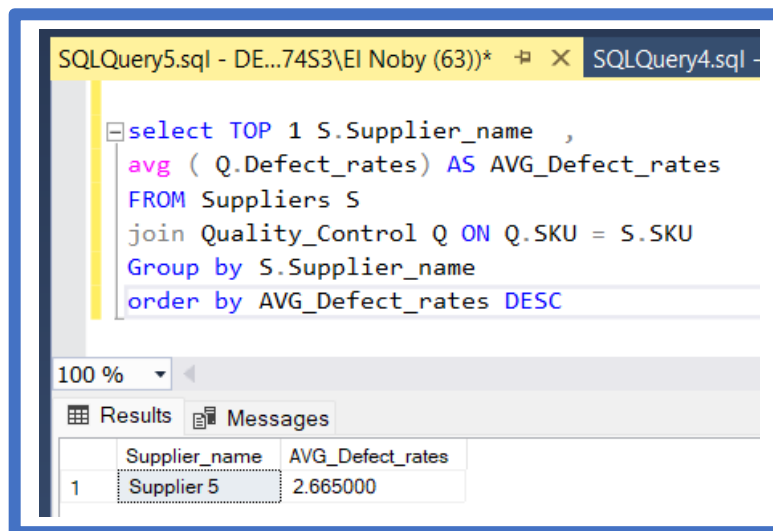
The screenshot shows a SQL query window with the following text:

```
select P.Product_type , Q.Inspection_results ,  
count( Q.Inspection_results) AS Total_Inspection_results  
FROM Products P  
join Quality_Control Q ON Q.SKU = P.SKU  
Group by P.Product_type , Q.Inspection_results  
order by P.Product_type , Q.Inspection_results
```

Below the query, the 'Results' tab is active, displaying a table with 9 rows and 4 columns:

	Product_type	Inspection_results	Total_Inspection_results
1	cosmetics	Fail	10
2	cosmetics	Pass	6
3	cosmetics	Pending	10
4	haircare	Fail	13
5	haircare	Pass	6
6	haircare	Pending	15
7	skincare	Fail	13
8	skincare	Pass	11
9	skincare	Pending	16

### 6.2 the Supplier with the highest average Defect rates



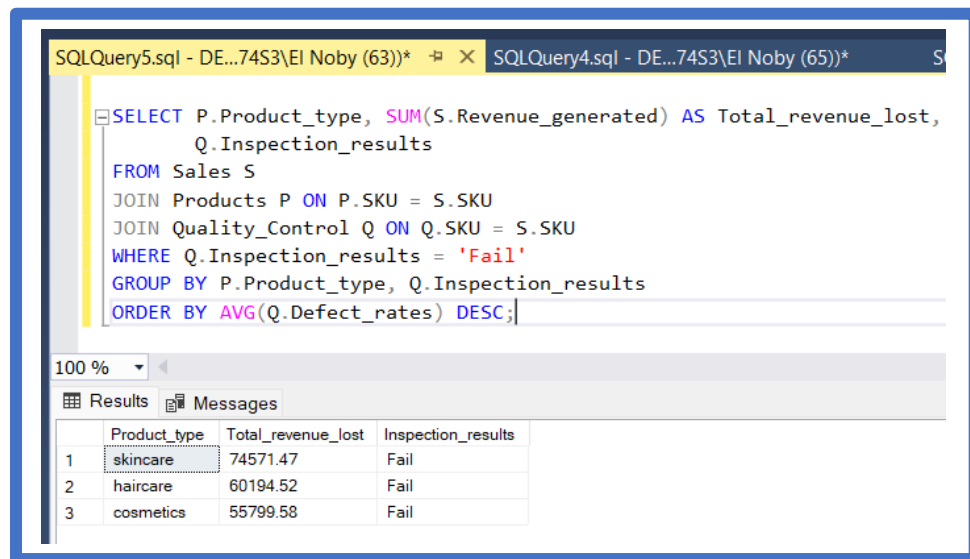
The screenshot shows a SQL query window with the following text:

```
select TOP 1 S.Supplier_name ,  
avg ( Q.Defect_rates) AS AVG_Defect_rates  
FROM Suppliers S  
join Quality_Control Q ON Q.SKU = S.SKU  
Group by S.Supplier_name  
order by AVG_Defect_rates DESC
```

Below the query, the 'Results' tab is active, displaying a table with 2 columns and 1 row:

	Supplier_name	AVG_Defect_rates
1	Supplier 5	2.665000

## 6.3 total revenue lost regarding failed Inspection results



The screenshot shows a SQL query editor with a query that calculates the total revenue lost for failed inspection results, grouped by product type and ordered by average defect rate. The results pane shows three rows of data.

```
SQLQuery5.sql - DE...74S3\EI Noby (63))* X SQLQuery4.sql - DE...74S3\EI Noby (65))* S
SELECT P.Product_type, SUM(S.Revenue_generated) AS Total_revenue_lost,
       Q.Inspection_results
FROM Sales S
JOIN Products P ON P.SKU = S.SKU
JOIN Quality_Control Q ON Q.SKU = S.SKU
WHERE Q.Inspection_results = 'Fail'
GROUP BY P.Product_type, Q.Inspection_results
ORDER BY AVG(Q.Defect_rates) DESC;
```

	Product_type	Total_revenue_lost	Inspection_results
1	skincare	74571.47	Fail
2	haircare	60194.52	Fail
3	cosmetics	55799.58	Fail

## 7. Conclusion

This analysis provides a comprehensive overview of the dataset, leveraging SQL and Python libraries to extract, clean, and analyze the data. The findings reveal valuable insights that can inform strategic decision-making.