DATAMITES INTERNSHIP

SUPPLY CHAIN ANALYSIS

PROJECT TEAM ID: **PTID-CDA-SEP-25-761**

PROJECT ID: **CDACL003 -SUPPLY CHAIN ANALYSIS**

PROJECT BY: **ARUN PANDIYAN R.A**

DATABASE MANAGEMENT SYSTEM: **SQL SERVER**

VISUALIZATION TOOL: **POWER BI**

TABLE OF CONTENTS

INTRODUCTION:

This project "Supply Chain Analysis" is about the analysis indulges in the product sales. With the dataset we have explored the end-to-end process and what are all the trade happened and how the sales performed.

we have pre-processed the dataset for basic exploration, we have created few **DAX queries** and created sophisticated new fields to explore valuable insights by visually explaining the story with interactive dashboard using **POWERBI**.

DATASET EXPLORATION:

First, we have accessed the dataset with the provided login credentials (host id, username and password) from **SQL Server** using **My SQL Workbench** and explored the database **'project_supply_chain_data"** and the table **"supply_chain_table"** using the following SQL queries.

(Note: The SQL queries are mention in the colour – SQL QUERIES)

- **ACCESSED SQL SERVER DATABASE:**

  SHOW databases;

  USE project_supply_chain_data;

  SHOW Tables;

Database
▶ information_schema
project_supply_chain_data

Tables_in_project_supply_chain_data
▶ supply_chain_table

It shows the available database as **'project_supply_chain_data'** and the table present within the database as **'supply_chain_table'.**

- **EXPLORING DATASET FIELDS:**

  DESCRIBE supply_chain_table;

Field	Type	Null	Key	Default	Extra
Product_type	varchar(9)	NO		NULL	
SKU	varchar(5)	NO		NULL	
Price	varchar(16)	NO		NULL	
Availability	int(11)	NO		NULL	
Number_of_products_sold	int(11)	NO		NULL	
Revenue_generated	varchar(16)	NO		NULL	
Customer_demographics	varchar(10)	NO		NULL	
Stock_levels	int(11)	NO		NULL	
Lead_times	int(11)	NO		NULL	
Order_quantities	int(11)	NO		NULL	
Shipping_times	int(11)	NO		NULL	
Shipping_carriers	varchar(9)	NO		NULL	
Shipping_costs	varchar(16)	NO		NULL	
Supplier_name	varchar(40)	NO		NULL	
Location	varchar(9)	NO		NULL	
Lead_time	int(11)	NO		NULL	
Production_volumes	int(11)	NO		NULL	
Manufacturing_lead_time	int(11)	NO		NULL	
Manufacturing_costs	varchar(16)	NO		NULL	
Inspection_results	varchar(7)	NO		NULL	
Defect_rates	varchar(18)	NO		NULL	
Transportation_modes	varchar(4)	NO		NULL	
Routes	varchar(7)	NO		NULL	
Costs	varchar(16)	NO		NULL	

It shows the available **fields, datatypes of the fields and also it shows whether there is null values present in the dataset or not**.

The Fields present in the provided dataset are **Product_type, SKU, Price, Availability, Number_of_products_sold, Revenue_generated, Customer_demographics, Stock_levels, Lead_times, Order_quantities, Shipping_times, Shipping_carriers, Shipping_costs, Supplier_name, Location, Lead_time, Production_volumes, Manufacturing_lead_time, Manufacturing_costs, Inspection_results, Defect_rates, Transportation_modes, Routes and Costs.**

If there's any discrepancies, we need to rectify it. If there is any Null value present, we need to replace it.

If the dataset doesn't have any null values, then we can use it to explore more and more in POWER BI.

Before going into POWER BI, we need to explore it using SQL QUERIES so that we can do preprocessing, data preparation and get some knowledge on the dataset. We will move on to next step now.

- **TOTAL COUNT OF ROWS:**

  SELECT count(*) FROM supply_chain_table;

	count(*)
▶	100

  It shows that the total available rows in the dataset is **100 Rows.**

- **DISTINCT SKU and COUNT OF SKU:**

  SELECT distinct SKU From supply_chain_table;

  SELECT Count(*) SKU FROM supply_chain_table;

SKU
SKU3
SKU4
SKU5
SKU6
SKU7
SKU8
SKU9
SKU10
SKU11
SKU12
SKU13
SKU14
SKU15
SKU16
SKU17

  … Table Continues

	SKU
▶	100

  The **count of SKU is 100**, it shows that the SKU has Distinct Values. So there's no any duplication or missing in it.

- **SUPPLIER DETAILS:**

  SELECT distinct Supplier_name From supply_chain_table;

Supplier_name
▶ Supplier 3
Supplier 1
Supplier 5
Supplier 4
Supplier 2

  It displays the details of the unique suppliers.

- **ROUTES:**

  SELECT DISTINCT Routes FROM supply_chain_table;

Routes
▶ Route B
Route C
Route A

  It displays the Route details.

- **CUSTOMER DEMOGRAPHIC DETAILS:**

  SELECT distinct Customer_demographics From supply_chain_table;

Customer_demographics
▶ Non-binary
Female
Unknown
Male

  It displays the details of customer demographic.

- **PRODUCT TYPE DETAILS:**

  SELECT distinct Product_type From supply_chain_table;

Product_type
▶ haircare
skincare
cosmetics

  It displays the details of product types.

- **TRANSPORTATION MODES:**

  SELECT distinct Transportation_modes From supply_chain_table;

	Transportation_modes
▶	Road
	Air
	Rail
	Sea

  It displays the transportation mode in which shipping are done.

- **MAXIMUM PRODUCTS SOLD:**

  SELECT Number_of_products_sold AS Max_products_sold
  FROM supply_chain_table
  WHERE Number_of_products_sold = (SELECT MAX(Number_of_products_sold)
  FROM supply_chain_table);

	Max_products_sold
▶	996

  It converts the Field name 'Number_of_products_sold as 'Max_product_sold' and displays the maximum value in the field.

- **MINIMUM PRODUCTS SOLD:**

  SELECT Number_of_products_sold AS Min_products_sold
  FROM supply_chain_table
  WHERE Number_of_products_sold = (SELECT MIN(Number_of_products_sold)
  FROM supply_chain_table);

	Min_products_sold
▶	8

  It converts the Field name 'Number_of_products_sold as 'Min_product_sold' and displays the minimum value in the field.

- **DEFECT RATES LESS THAN 1:**

  SELECT Defect_rates, Transportation_modes FROM supply_chain_table
  WHERE Transportation_modes = 'Air' AND Defect_rates <1;
  SELECT Defect_rates, Transportation_modes FROM supply_chain_table

WHERE Transportation_modes = 'Sea' AND Defect_rates <1;
SELECT Defect_rates, Transportation_modes FROM supply_chain_table
WHERE Transportation_modes = 'Road' AND Defect_rates <1;
SELECT Defect_rates, Transportation_modes FROM supply_chain_table
WHERE Transportation_modes = 'Rail' AND Defect_rates <1;

Defect_rates	Transportation_modes
0.0211698213729943	Air
0.100682851565093	Air
0.102020754918176	Air
0.0186075676310149	Air
0.815757079295672	Air
0.613326899164507	Air
0.372304767985097	Air
0.447194015463823	Air
0.159486314717514	Air
0.165871627480608	Air

Defect_rates	Transportation_modes
0.541154098060581	Sea
0.0453022623982596	Sea
0.131955444311814	Sea
0.722204401882931	Sea

Defect_rates	Transportation_modes
0.226410360849925	Road
0.39817718685065	Road
0.965394705352393	Road
0.646604559372054	Road
0.773006134067247	Road

Defect_rates	Transportation_modes
0.380573586713213	Rail
0.396612724109935	Rail
0.626001858209394	Rail
0.333431825224739	Rail
0.346027290705503	Rail

It displays the details of the Defect_rates which is less than 1 in various
transport_modes like Air, Sea, Road and Rail so that we can able to identify in which
mode of transport the defect is very less.

- **TOTAL UNITS OF HAIRCARE, SKINCARE AND COSMETICS:**

SELECT Count(*) AS total_haircare_units FROM supply_chain_table
WHERE Product_type = 'haircare';
SELECT Count(*) AS total_skincare_units FROM supply_chain_table
WHERE Product_type = 'skincare';
SELECT Count(*) AS total_cosmetics_unit FROM supply_chain_table
WHERE Product_type = 'cosmetics';

	total_haircare_units
▶	34

	total_skincare_units
▶	40

	total_cosmetics_unit
▶	26

It displays the total units of haircare, skincare and cosmetics in SKU (Stock Keeping Unit).

So that we have explored the major details of the dataset and pre-processed the fields present in the dataset and gained information about total no of rows in the dataset, Product names, Supplier names, Routes, Transportation Modes, Maximum and Minimum products sold, Defects <1 in various transportation modes, total SKU units on various Products.

Based on these insights we have gained basic details of the dataset using this knowledge we have to create chart distributions of various fields and do comparisons and find out deep insights by storytelling using Power BI.

IMPORTING DATA IN POWERBI:

After exploring and gaining some basic knowledge on the provided data from database. We are now importing the database into Power BI.

There are many ways to import data into Power BI **by importing the spreadsheets, by connecting to data from multiple sources**, **by entering the data manually, by connecting organizational data with Data Lake, by connecting to SQL Server, by connect with Dataverse through SQL endpoint, by manage and connect with recent resources.**

*Here we are connecting to **SQL Server** with given login credentials **and then the database will be accessed for further explorations**.*

After connecting with SQL server, the database "project_supply_chain_data" has been displayed in 'Data Pane'.

In the 'Table View' section the "supply_chain_table" is displayed in the table we are going to change few details as per our convenient so that we can able to identify the column easily for exploring insights.

RENAMING THE DATA FIELD NAMES:

We have Renamed few fields, they are

SKU - **Products_Unique_ID**
Price - **Product_Price_USD**
Availability - **Products_Availability_for_purchase**
Customer_demographics – **Customer_gender**
Stock_levels - **Stock_in_inventory**
Lead_times - **Customer_Lead_times**
Order_quantities - **Quantities_ordered_by_customer**
Shipping_times - **Shipping_times_seller_to_customer**
Shipping_carriers - **Shipping_carriers_seller_to_customer**
Shipping_costs - **Shipping_costs_seller_to_customer**
Location – **Supplier_location_facility**
Insperction_results – **Quality_inspection_results**
Defect_rates - **Precentage_of_defects**
Transportation_modes – **Transportation**
Costs – **Transportation_cost**

Likewise we have renamed the fields present in the table as per our understanding which will be feasible for creating dashboard charts with more understandable way and also we have created new columns with DAX Measures.

DAX MEASURES:

DAX is called as DATA ANALYSIS EXPRESSIONS and it is used for data analysis and calculations after the data has been loaded into Power BI or Excel. It is used to create measures, calculated columns, and perform complex aggregations.

Here we have created new columns with DAX measures. The columns which we created are **Revenue_on_products, Profit, Total_cost, Defects_in_%, Average_Revenue_on_Products, Average_Cost, Lead_Time_Bin.**

(Note: The DAX Measures are mention in the colour – DAX MEASURES)

EXPLORING NEW COLUMNS WITH DAX MEASURES:

- **REVENUE ON PRODUCTS**:

    The new column **Revenue_on_products** has been created with DAX measure as

    Revenue_on_products = 'project_supply_chain_data supply_chain_table'[Number_of_products_sold] * 'project_supply_chain_data supply_chain_table'[Product_Price_USD]

    The Revenue on products has been calculated by number of products sold with product price.

- **AVERAGE REVENUE ON PRODUCTS:**

    The new column **Average_Revenue_on_Products** has been created with DAX measure as

    Avg_Revenue_on_products = AVERAGE('project_supply_chain_data supply_chain_table'[Revenue_on_products])

    The Average Revenue on Products has been calculated with this measure.

- **TOTAL COST:**

    The new column **Total Cost** has been created with DAX measure as

    Totoal_cost = ('project_supply_chain_data supply_chain_table'[Number_of_products_sold])*('project_supply_chain_data supply_chain_table'[Manufacturing_costs]+'project_supply_chain_data supply_chain_table'[Shipping_costs_seller_to_customer])

    The **Total Cost** is calculated by adding **Manufacturing Cost** and **Shipping Cost** and multiplying with **No of products sold**.

- **AVERAGE COST ON PRODUCTS:**

   The new column **Average Cost** has been created with DAX measure as

   Average_Cost = AVERAGE('project_supply_chain_data supply_chain_table'[Totoal_cost])

   The **Average_Cost** is calculated by averaging the Total_cost.

- **LEAD TIME BIN:**

   The new column **Lead_Time_Bin** has been created with DAX measure as

   Lead_Time_Bin =

   SWITCH(

      TRUE(),

      'project_supply_chain_data supply_chain_table'[Customer_Lead_times] <= 10, "0-10 Days",

      'project_supply_chain_data supply_chain_table'[Customer_Lead_times] <= 20, "11-20 Days",

      'project_supply_chain_data supply_chain_table'[Customer_Lead_times] <= 30, "21-30 Days",

      "30+ Days"

)

   The Bins are separated as **"0-10 Days", "11-20 Days", "21-30 Days" and "30+ Days"**

- **AVERAGE INVENTORY:**

   The new column **Average_Inventory** has been created with DAX measure as

   Average_Inventory =
   AVERAGE('project_supply_chain_data supply_chain_table'[Stocks_in_inventory])

   It displays the average of inventory stocks.

- # **DASHBOARD**
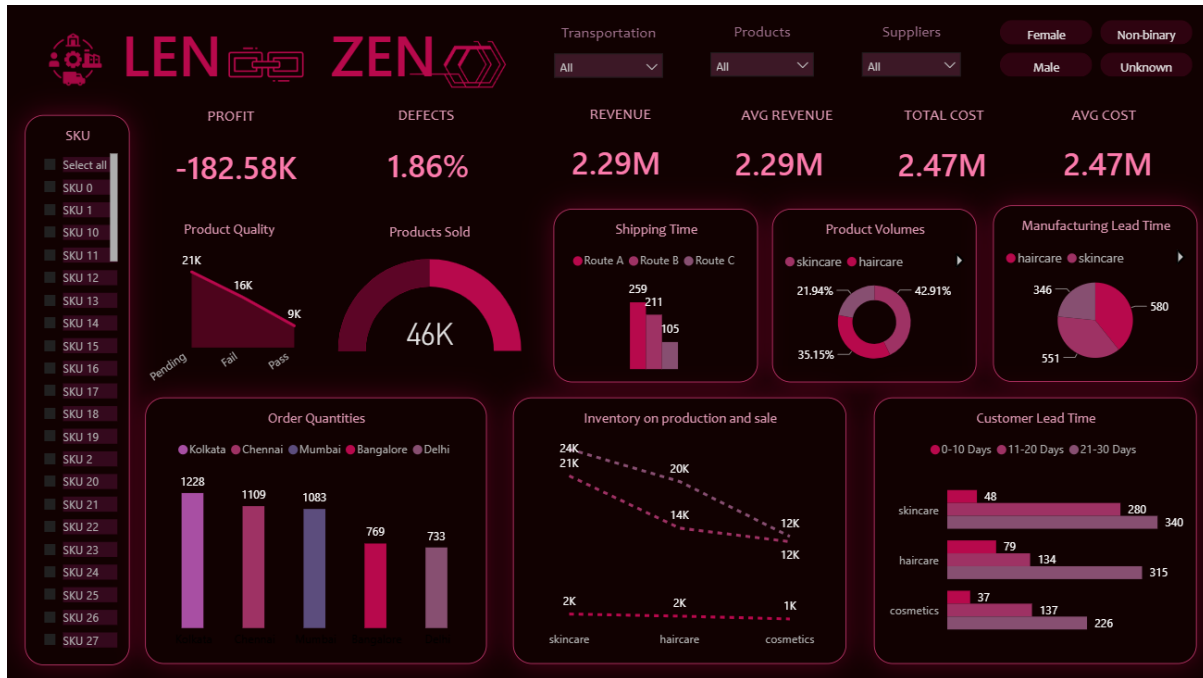
Dashboard Consists of:

- • Have imagined a Word mark logo for the company as **LEN ZEN**.
- • In the top right there is **Transportation, Products and Suppliers** as dropdown list box.
- • Then followed by **PROFIT, DEFECTS, REVENUE, AVG REVENUE, TOTAL COST, AVG COST**,
- • Then in the middle **product Quality, Products Sold, Shopping Time, Product Volumes, Manufacturing Lead Time**.
- • Then in the Last row its **Order Quantities** and **Customer Lead Time**.
- • **SKU** unit IDs are placed in the Left Corner with check box options.

DASHBOARD DETAILS:
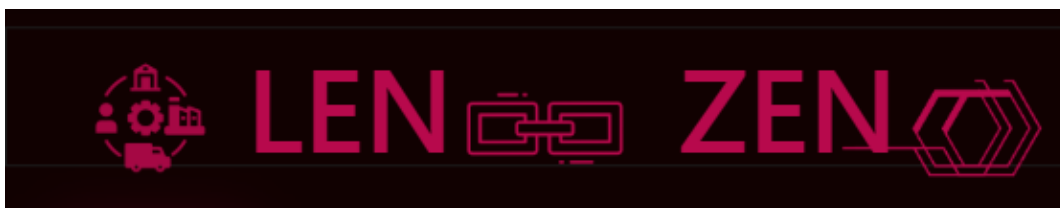
For this supply chain analysis dashboard, I have chosen an imaginary name for the company and designed logo called LEN ZEN.

PROFIT	DEFECTS	REVENUE	AVG REVENUE	TOTAL COST	AVG COST
-108.61K	1.86%	1.23M	1.37M	1.34M	1.48M

It displays **PROFIT, PERCENTAGE OF DEFECTS, REVENUE, AVERAGE REVENUE, TOTAL COST** and **AVG COST**.
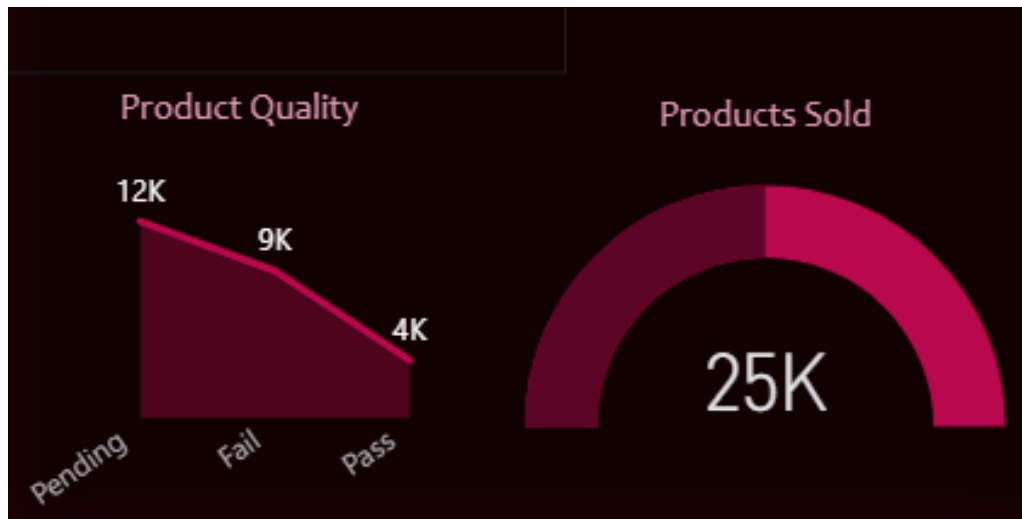
It shows the quality of the products whether the quality inspection is pending or fail or pass. We can able to check the quality differentiation in this Stacked area chart.

Under the Products Sold, we can able to see the details of total number of products sold, when we select haircare product it will show the total no of haircare products sold which follows the same for skincare and cosmetic while selecting those products to get displayed.
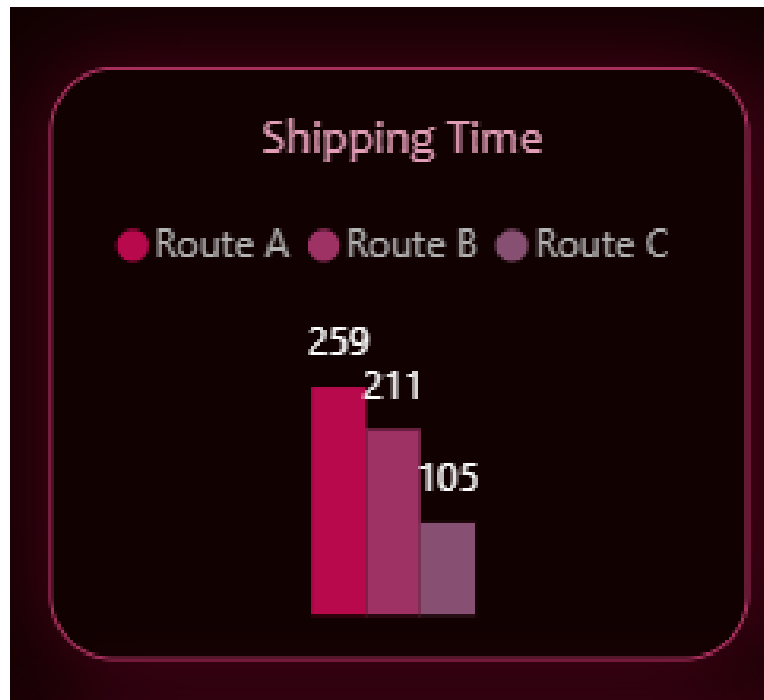
In this Clustered Colum Chart, we can able to know about the Total amount of Shipping Time taken based on the Routes A, B and C. So that we can able to identify the fastest shipping times on the Routes.
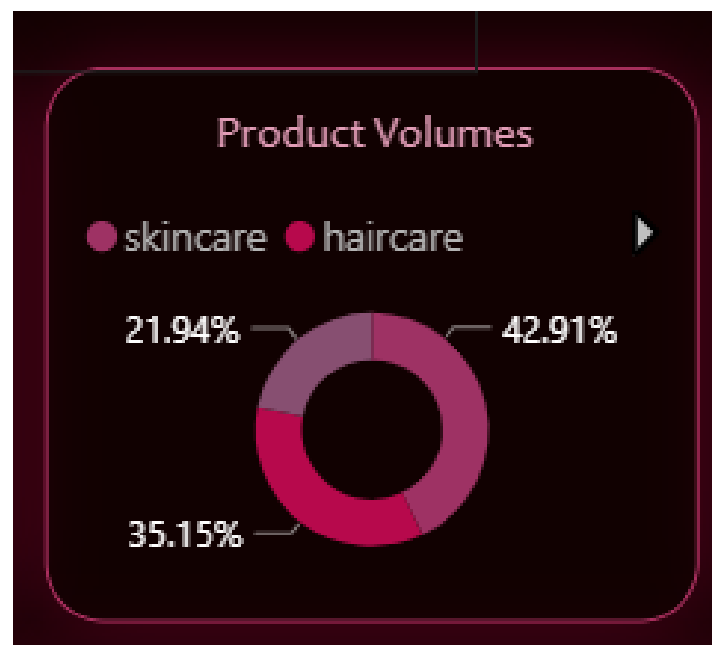
In this Product Volumes Donut Chart, we can able to see the variations in product volumes based on manufacturing, based on orders, based on shipping times, based on the products by selecting it. It shows the volumes in Percentages.

It displays the Total manufacturing Lead time based on products, using this Pie chart we can easily able to know which products took maximum manufacturing time and minimum manufacturing time.
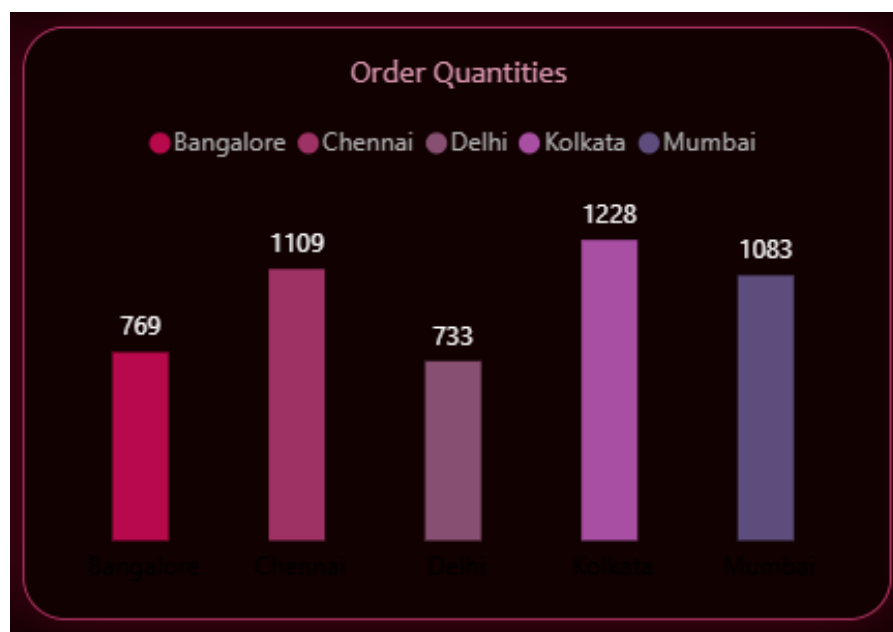
This Order Quantities Ribbon chart displays the Quantity of orders done by customers based on Supplier Locations. Here we can see the various supplier locations as Bangalore, Chennai, Delhi, Kolkata, Mumbai. In this way we can able to check which Supplier Locations Ordered more.
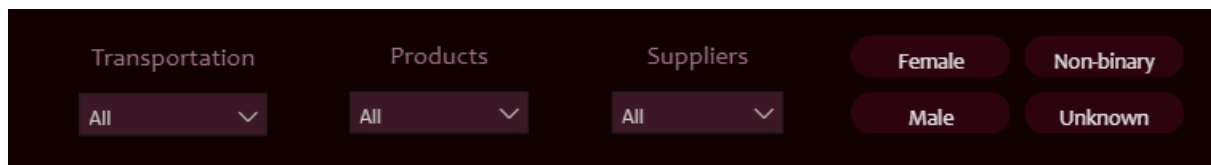
Inventory on production and sale

The Line chart shows the trends on inventory based on production volume and the sale of products.

Customer Lead Time

The Clustered Bar chart shows the customer lead time taken from time of ordering the products to time of delivery. The days have been agreegated into bins.

It shows the slicers which has drop down filters of Transportation mode, Product categories, supplier names and genders. We can choose our filters and check the details according to our wish.
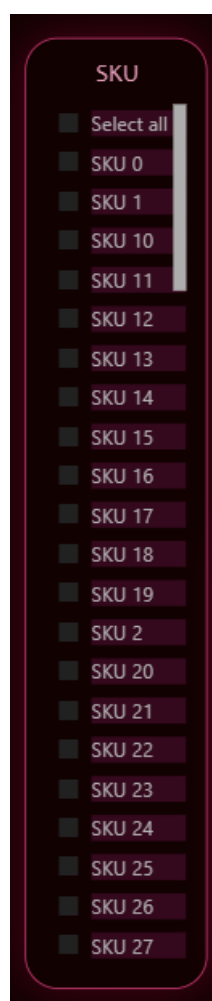
The Filtered slicer shows the List of SKU (STOCK KEEPING UNIT) we can select the units which we needed more insights and details.

CONCLUSION:

With the preprocessing of data, using DAX measures have created Interactive dashboard for supply chain analysis, With the dashboard I'm concluding some insights as follows

- **Total Cost** is **2.4 million**
- **Total Revenue** is **2.29 million**
- Overall **Highest number of orders** is from **Kolkata**
- Overall **Lowest number of orders** is from **Delhi**
- No of **Product passed in quality check is 9045** products out of **46099 overall products**
- **Skincare** products dominate with highest number of inventories.
- **Skincare** products sold in **highest numbers** while comparing to other products with **20731 skincare products with 42.91 %**
- **Route C** is the **fastest shipping Route** (Less Shipping Time) with generated **Revenue** as **379.32 K**
- **Chennai** location received the highest number of order quantity (1109 **Orders**) in **Route C** in **Less Shipping Time** as less than and equal to 10 days of **Customer Lead Time**.
- **Highest Production Volume** is for **Skincare Products with 24366 Products** in which 1370 products passed the Quality checks.
- **Cosmetics Products** consumes Less Manufacturing Time
- **Overall Defects** with total production are **1.86 %**
- **Total no of products failed in Quality** Check is **15610** Products
- **Supplier 1** has highest no of product sold **(11080 products)** in which **4321 products passed quality checks**.
- By the way of **Road Transportation** mode **32979 Products had been sold** by which the **profit is 51.90 K** in which **Skincare has the more production volume transported in Road**. **Chennai** has **gained highest number of Order Quantities by Road ways**.
- In the **Customer Demographic, male** is the more profit generator than female, non-binary and unknown.
- In the Location wise **Chennai is the most profit generating Location** based on Quantity of Order placed**.**
- From **Chennai Location 519 Order quantities** were placed by **Unknown Persons**.

So that we were concluding our Insights. In this way LEN and ZEN's Supply Chain Analysis has been done.