

# Final Project

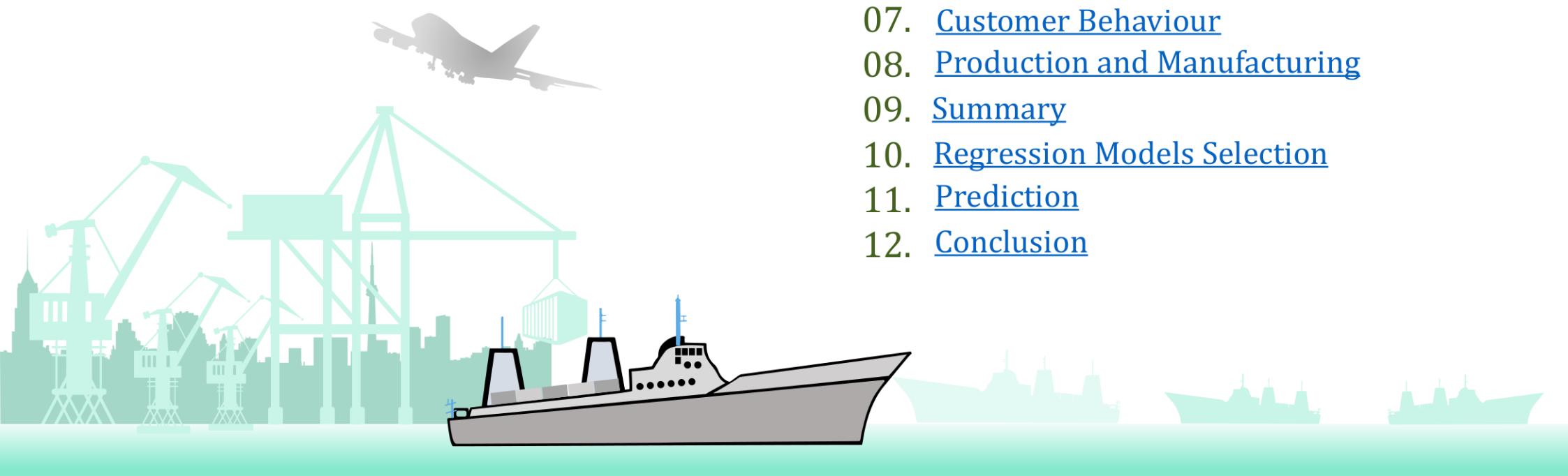
## Supply Chain

### Analysis and Modelling



**Bernat Sitinjak**  
June, 11th 2023

# Contents

- 
01. [Introduction](#)
  02. [Dataset Overview](#)
  03. [Data Quick-Glance](#)
  04. [Problem Framework](#)
  05. [Logistics Management](#)
  06. [Demand and Products delivery](#)
  07. [Customer Behaviour](#)
  08. [Production and Manufacturing](#)
  09. [Summary](#)
  10. [Regression Models Selection](#)
  11. [Prediction](#)
  12. [Conclusion](#)

### Background

- ▶ Managing the flow of products, information and finance in an integrated manner involving parties from upstream to downstream consisting of suppliers, factories, distribution networks and logistics services
- ▶ Create efficiency and effectiveness in the business process activities carried out by a company
- ▶ Produce and distribute products in the right quantities, in the right locations, and at the right time to reduce costs and meet customer needs

### Goals

- ▶ Balancing between demand and supply to be more effective and efficient.
- ▶ Procurement management, supplier management, customer relationship management, problem identification and problem response, risk management.
- ▶ Win the competition, at least the company can survive in the midst of intense market competition.
- ▶ Providing products at low prices, quality, on time, and also more variety.



## Dataset Overview

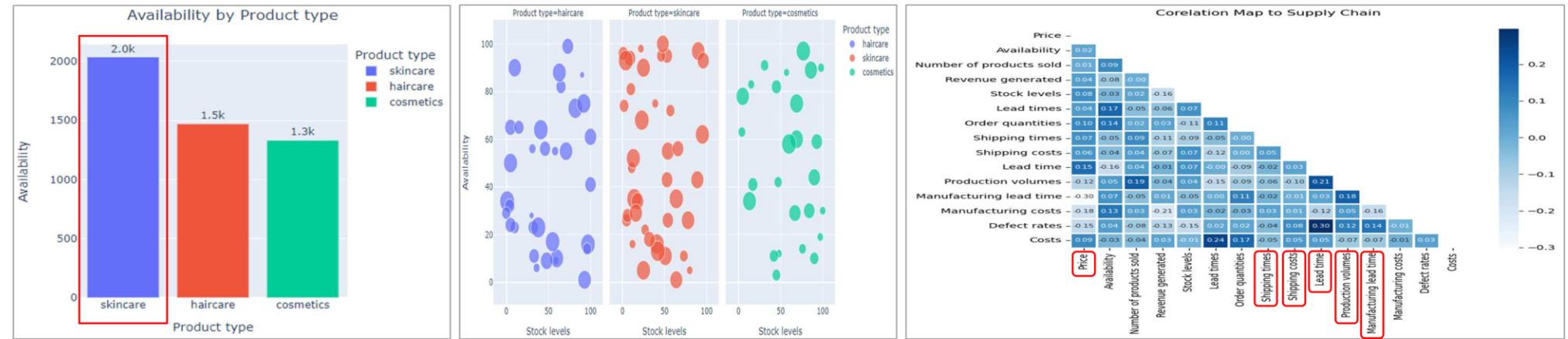
### Source

### [Supply Chain Analysis](#)

### Data Dictionary

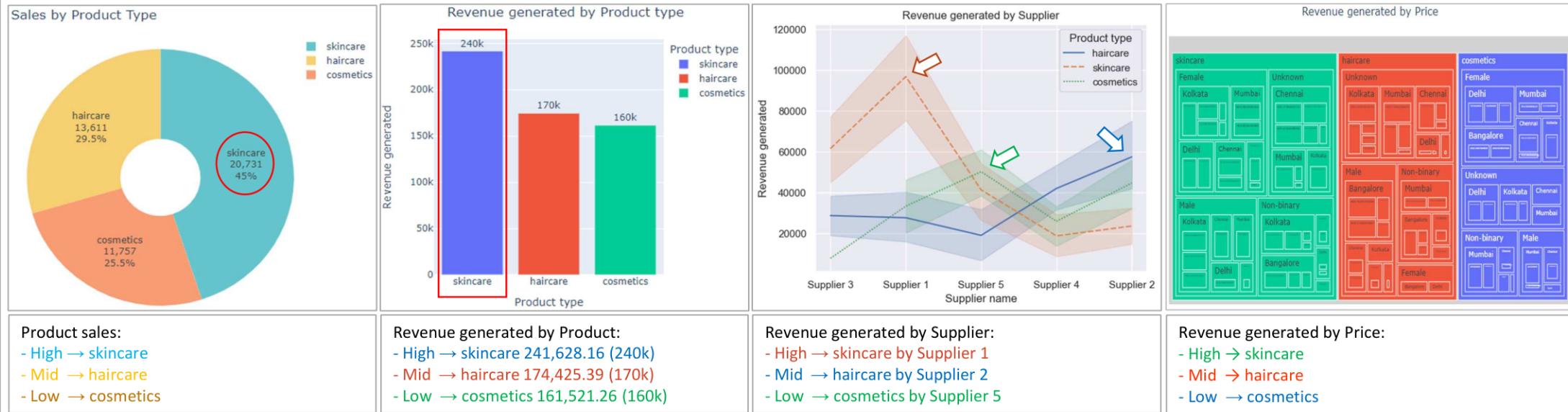
No.	Variables	Description	No.	Variables	Description
1	Product type	Types or names of goods available	13	Shipping costs	Item delivery cost
2	SKU	Stock Keeping Unit or unique code that becomes a marker for each item	14	Supplier name	Supplier of goods or services
3	Price	Price of goods	15	Location	Location of buyer
4	Availability	The number of items ready to be sold	16	Lead time	Total order time
5	Number of products sold	The number of items that have been sold	17	Production volumes	The number of products a makes
6	Revenue generated	Income earned from each process according to activities	18	Manufacturing lead time	The amount of time (start of the process to completion)
7	Customer demographics	Categories of consumer populations useful for marketing and product design	19	Manufacturing costs	Cost of Production
8	Stock levels	The limit amount of inventory that must be provided	20	Inspection results	The status of the inspection results of each item produced
9	Lead times	The amount of time for initiation and process completion	21	Defect rates	Established product quality standards
10	Order quantities	Number of customer orders	22	Transportation modes	The means supporting the mobility of freight
11	Shipping times	Number of orders shipped	23	Routes	Routes of transportation mode
12	Shipping carriers	Goods delivery service	24	Costs	The amount of money spent to produce a product

## Data Quick-Glance



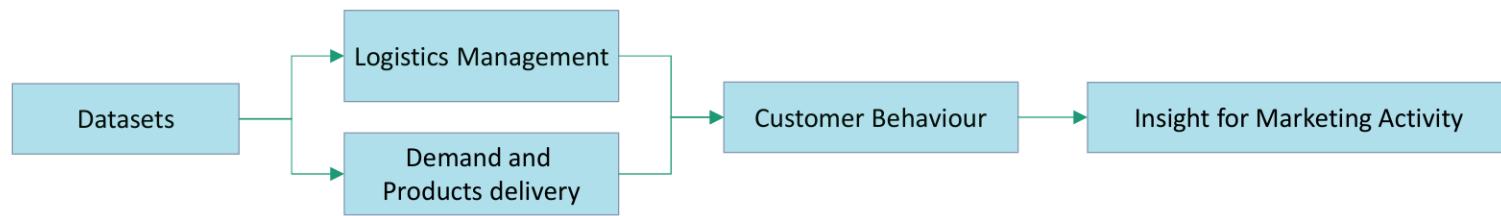
**Stock levels** is the number of items a company has in its entire **logistics or distribution network**. **Stock quantity** located in **warehouses, logistics centers, and physical stores**. Maintaining **Stock levels** are **efficiency and management strategy** to meet **sustainable customer needs**.

The correlation matrix is the relationship between variables that affect resistance (strength). Revenue generated in this dataset is more strongly influenced by variables: **Price, Shipping times, Shipping costs, Lead time, Production volumes, Manufacturing lead time** ( $>= 0.5$  atau  $\geq -0.05$ ).

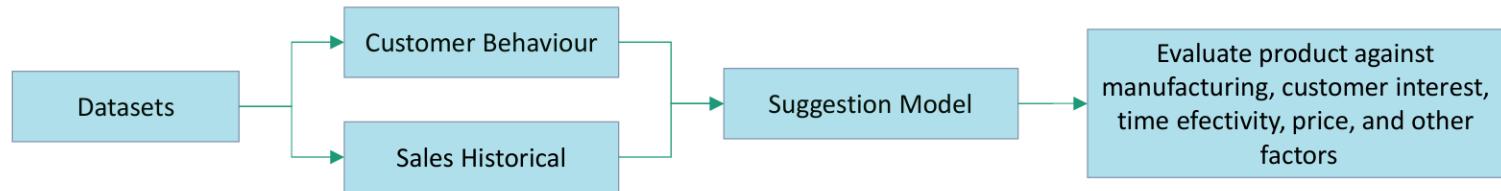


## Problem Framework

### ► Identifying Customer Behaviour



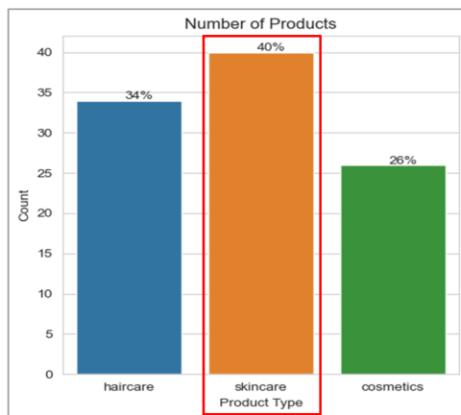
### ► Data Dictionary



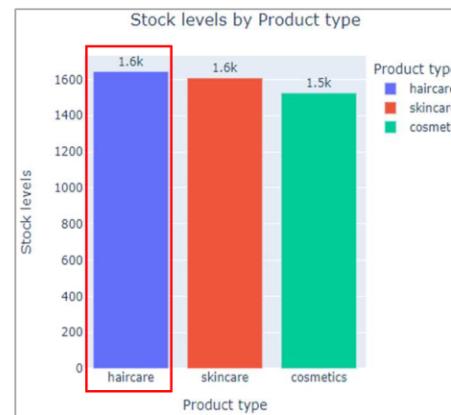
## Logistics Management

► To segment based on stock availability.

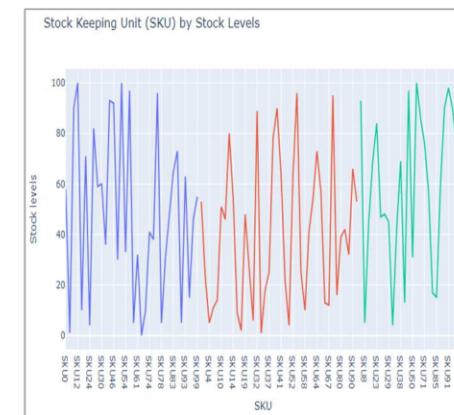
- Identify and meet customer needs from the number of products sold
- Correlation between Stock Levels , Prpduct type, and SKU



The higher **number of products** based on product type is **skincare** 40%



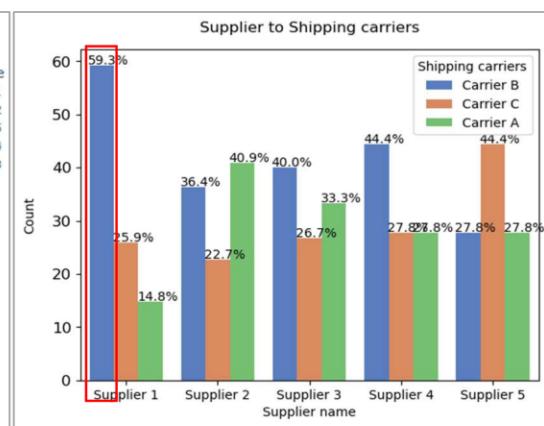
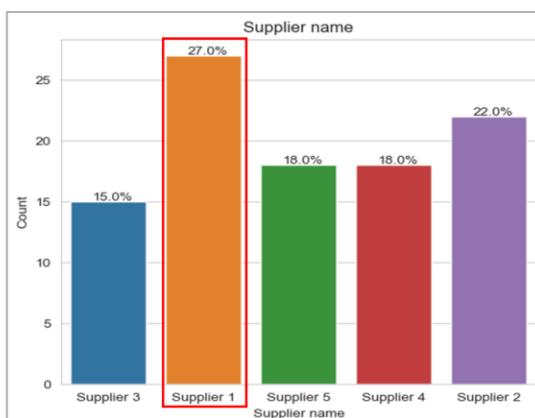
The highest **stock levels** is **haircare** as 1,644 pcs (1.6k)



## Demand and Products delivery

### ► To segment based on stock demand and delivery

- Correlation between Supplier, order quantities, and delivery time
- Selection of Shipping, transportation modes, routes, and shipping costs



Supplier 1 is the top quantities order category reaching 27.0%

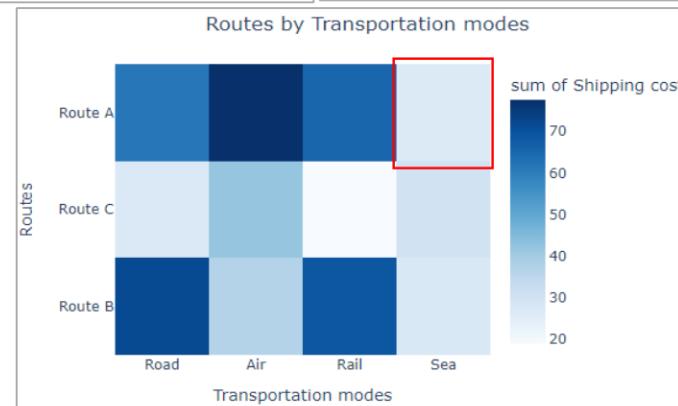
The highest order quantities reaching 1,458 pcs (1.5k) were made by Supplier 1

The shipping carriers that received the highest deliveries were Carrier B, around 59.3% from Supplier1

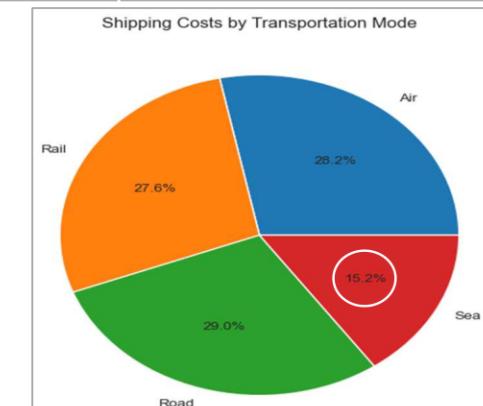
The shipping carriers highest delivery is Carrier C around 31.0% to Kolkata city.



Most shipping times are 9 times to Mumbai locations using shipping carriers carrier B



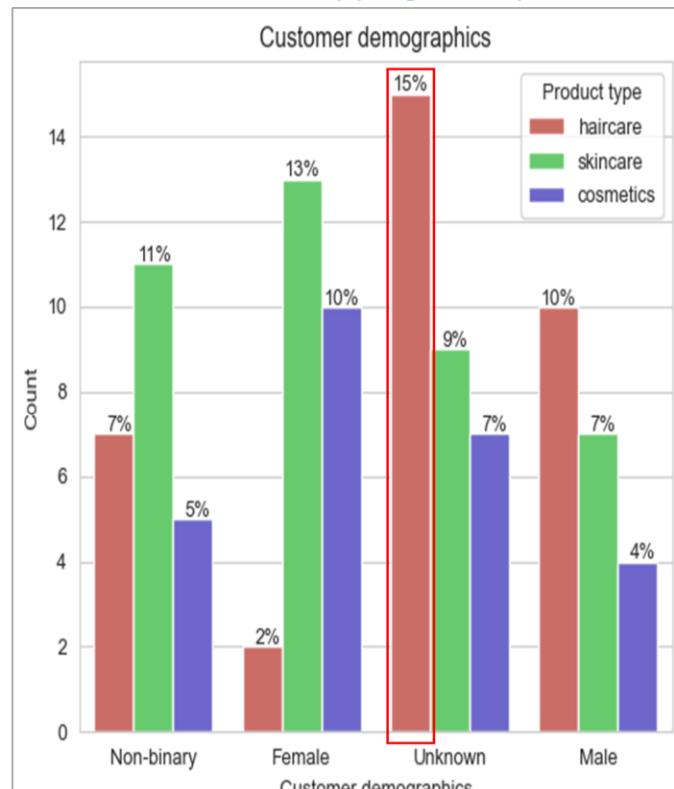
The lowest sum of shipping costs is 20 using the Sea transportation mode via Route A



Shipping costs as low as 15.2% using Sea transportation mode.

## Customer Behaviour

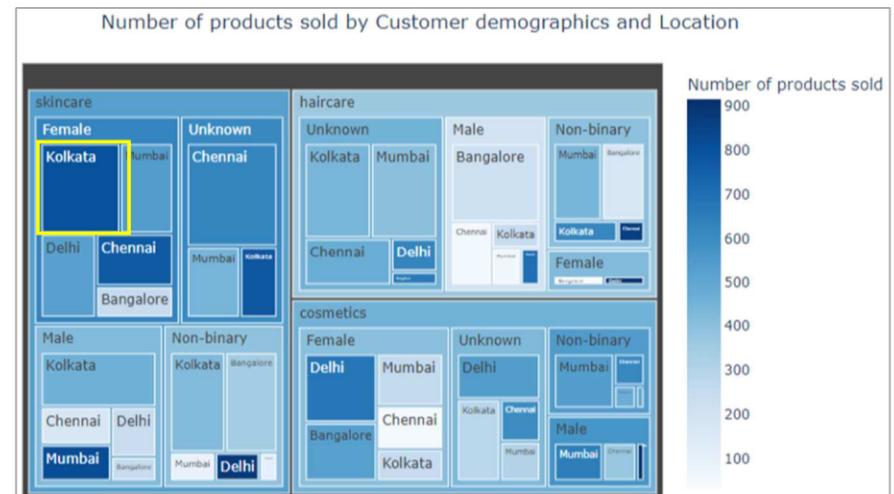
- To segment based on stock demand and delivery
  - Correlation between Supplier, order quantities, and delivery time
  - Selection of Shipping, transportation modes, routes, and shipping costs



The **customer demographic** for **haircare** products is the highest (15%). They are customers from **Unknown** sources (the data is not completed)

The highest number was found in :

- Number of products sold **805,6300**
- **Skincare** products
- **Female** and
- the city of **Kolkata**



Revenue generated obtained from customer demand are from products:

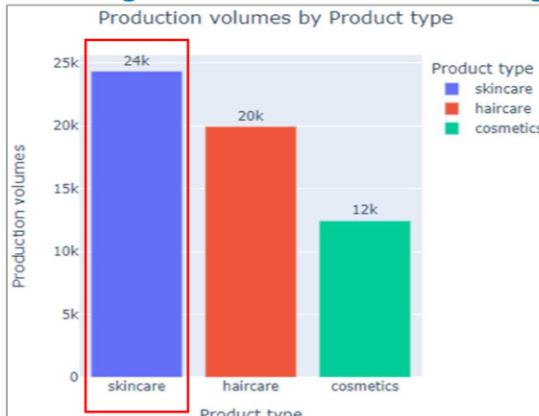
- **skincare**
- max: **9,571.55**
- mean: **8,128,028**
- source: **Unknown**



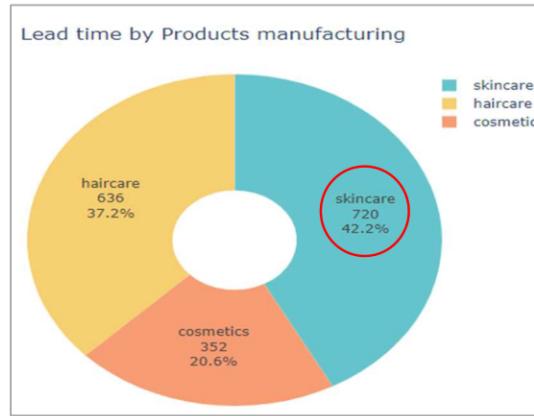
## Production and Manufacturing

► Producing sufficient orders and ensuring sufficient quantities are in stock.

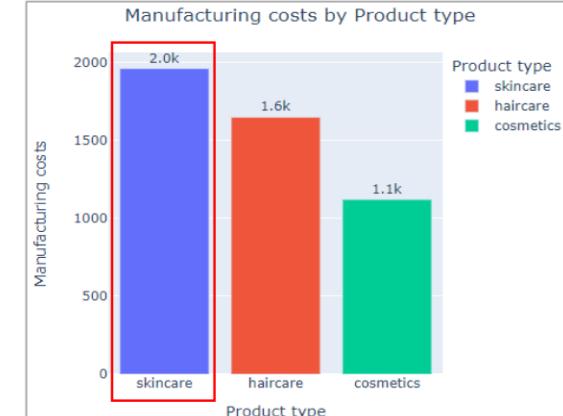
- Correlation between Production volume, and Lead time
- Management costs and revenue generated



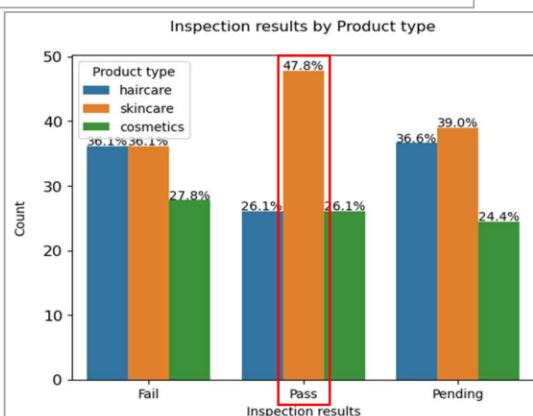
The highest total **production volumes** produced were **skincare** types, **24,366 pcs (24k)**



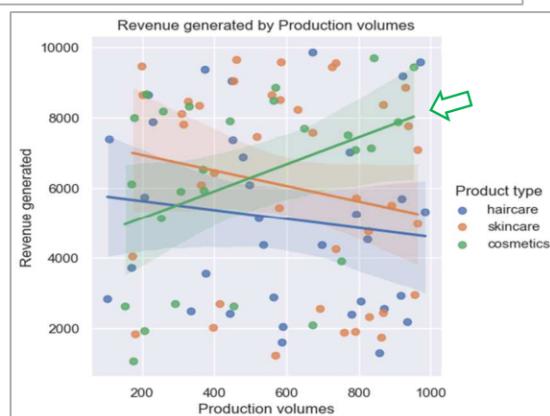
The highest for **product manufacturing** is **skincare**. The lead time for producing **42.2%** of the total production takes **720**



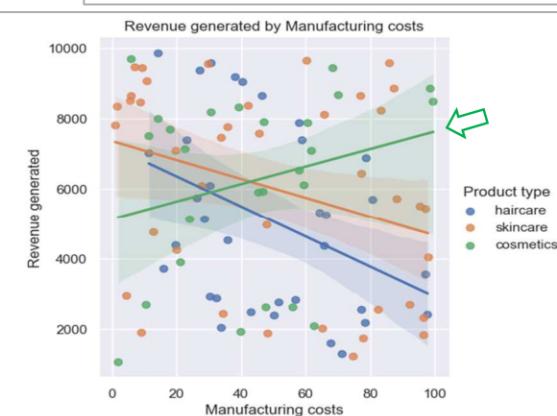
The largest total **production cost** for the type of **skincare** product is **1,959.73 (2.0k)**



The highest **Inspection results** for every manufacturing product passed by **skincare** **47.8%**



The company derives more **revenue** from **cosmetics** products, and the **higher the price of cosmetics** products, the more **revenue generated**.



## Summary

### ► Logistics Management

- Larger skincare supplies up to 40%
- Based on Stock Levels, larger haircare products reached 1,644 pcs (1.6k)
- All supplies have uniq code (SKU)
- Number of products sold was higher for skincare products, reaching 20,731 pcs (21k)
- The highest supply Availability for skincare reached 2,037 pcs (2.0k)

### ► Demand and Products delivery

- The highest order is Supplier 1 reaching 27.0% reaching 1,458 pcs (1.5k)
- Shipping carriers by Supplier 1 use Carrier B with a rate of 59.3%
- Kolkata is the highest customer using Carrier C reaching 31.0%
- The Mumbai location used the highest shipping carriers carrier B at 9 times
- Lowest shipping costs of 20 using sea transportation through Route A of 15%

### ► Customer Behaviour

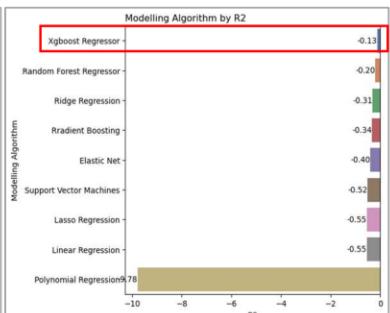
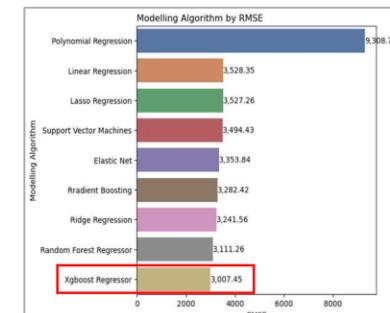
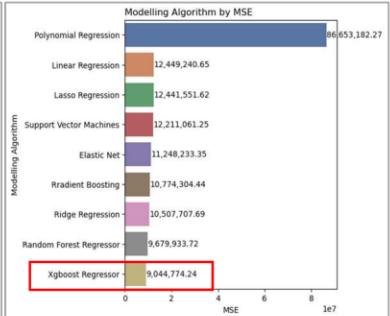
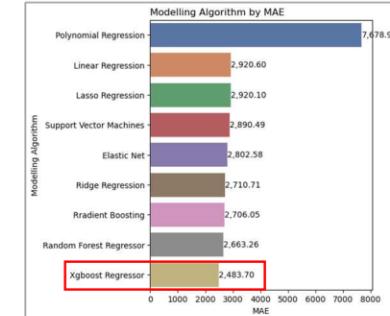
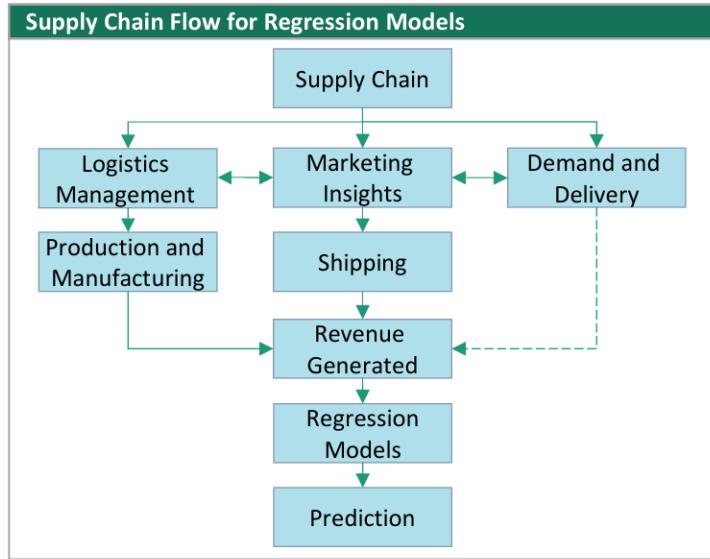
- The highest product customer in haircare is from Unknown at 15%
- Highest number of products sold Skincare = 805.6300, by Women from Kolkata
- Revenue generated obtained from customer demand are from products: skincare, max: 9,571.55, means: 8,128,028, source: Unknown

### ► Production and Manufacturing

- The highest skincare production volume reached 24,366 pcs (24k)
- The lead time for manufacturing skincare products of 720 requires 42.2% of all total production
- The highest total production cost of 1,959.73 (2.0k) for producing skincare

## Regression Models Selection

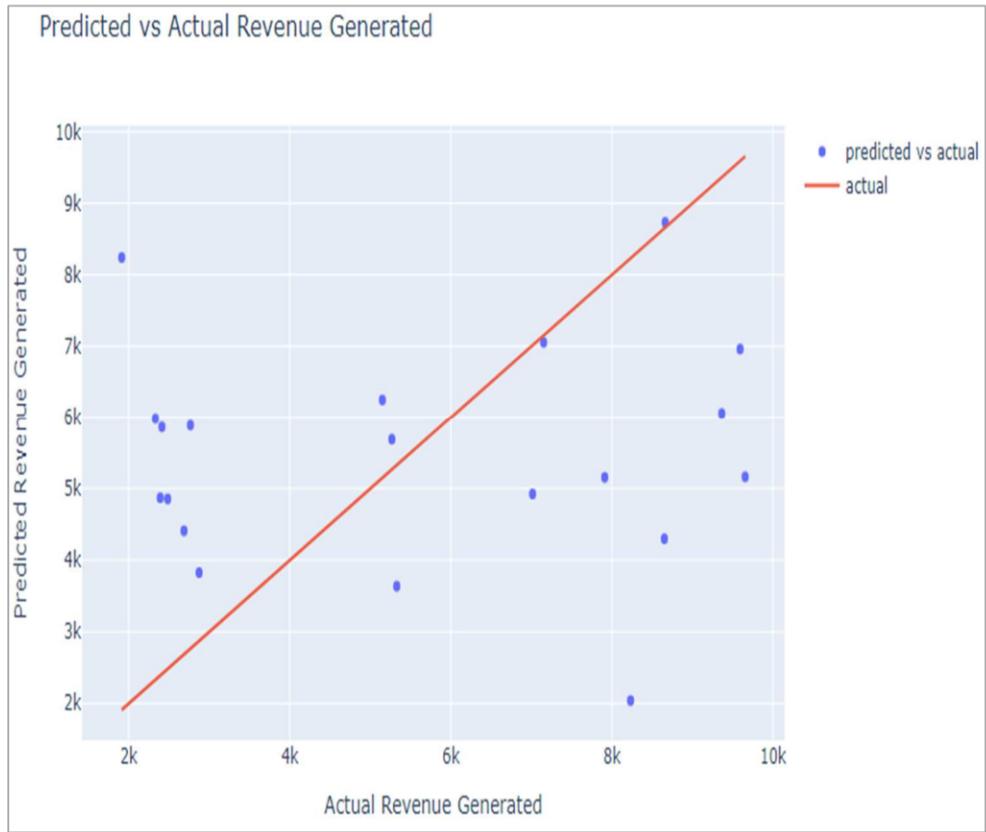
- ▶ Data Sets:
  - Supply Chain
- ▶ Features:
  - Product type
  - Price
  - Availability
  - Stock levels
  - Customer demographics
  - Order quantities
  - Supplier name
  - Shipping carriers
  - Location
  - Production volumes
  - Manufacturing costs
  - Lead time
  - Transportation modes
  - Routes
- ▶ Target
  - Revenue generated



Based on the evaluation metrics (MAE, MSE, RMSE, and R2), it appears that the model with **the best performance** is the **Xgboost Regressor**.

This model has the **lowest MAE, MSE, RMSE, and R2 values**.

## Prediction



From the graph above, manually calculated:

**Actual: 45,292.82, Predicted: 64,607.803**

The difference in revenue generated: **19,314.983**

- Haircare and Skincare sales are **high** in Kolkata.
- The **manufacturing cost, revenue & number of orders for Skincare** are **highest**.
- **Women (female)** and **Non-binary** gender people purchased more Skincare products. **Undisclosed gender (unknown)** people purchased more Haircare products. **Men** purchased all three products more or **less equally**.
- **SKU51** has **highest revenue**.
- **Shipments** are recommended to use **sea freight** via **Route A** because shipping **costs are cheap**
- There is a negative correlation between **Order quantities** and **Stock levels** which means that as the **stock levels increase**, the **order quantities decrease**, and **vice versa**. In the context of a product sales dataset, this may suggest that **as the stock levels of a particular product increase, there is less need to place large orders for that product since there is already a sufficient amount in stock**. On the other hand, if **the stock levels are low**, the company may need to place larger orders to ensure that there is enough inventory to meet demand.

## Conclusion

Based on the analysis of the given dataset, it is evident that there is a wealth of information that can be extracted to gain insights into the supply chain operations. Through the use of various visualizations, **we were able to identify the most popular products, revenue generation, transportation modes, shipping costs, lead times, and manufacturing costs, among other metrics.**

This analysis can help inform strategic decision-making and identify areas for improvement in the supply chain operations. For example, it may be possible to optimize **inventory management, reduce waste and inefficiencies, improve product qualityand strengthen supplier relationships.**

Overall, the insights gained from analyzing this dataset **can help drive positive change in the supply chain, leading to improved efficiency, sustainability, and profitability.**

**Thank You**

