Analyzing Supply Chain Dynamics in India's Beauty Industry: A Regional Study of Kolkata, Mumbai, Chennai, Bangalore, and Delhi

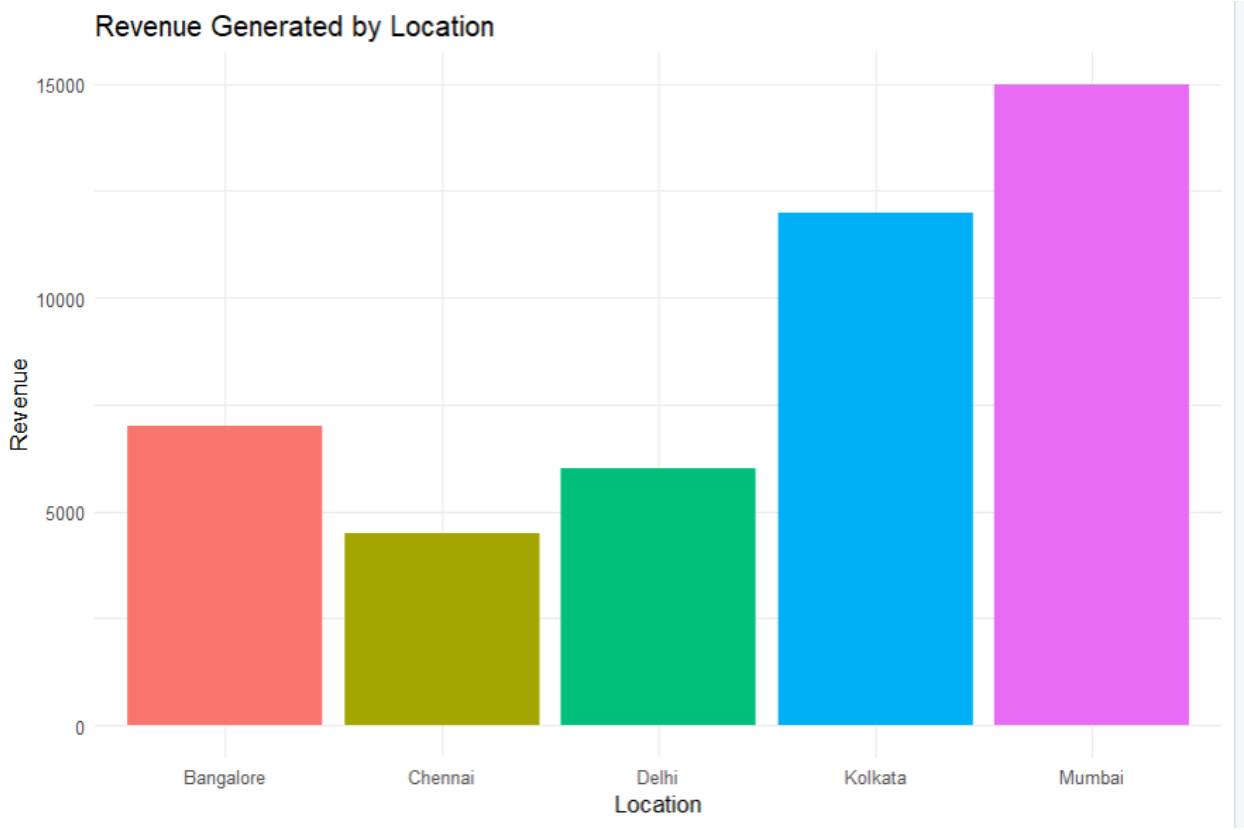
Barbara Afia Asamoah

May 20, 2025

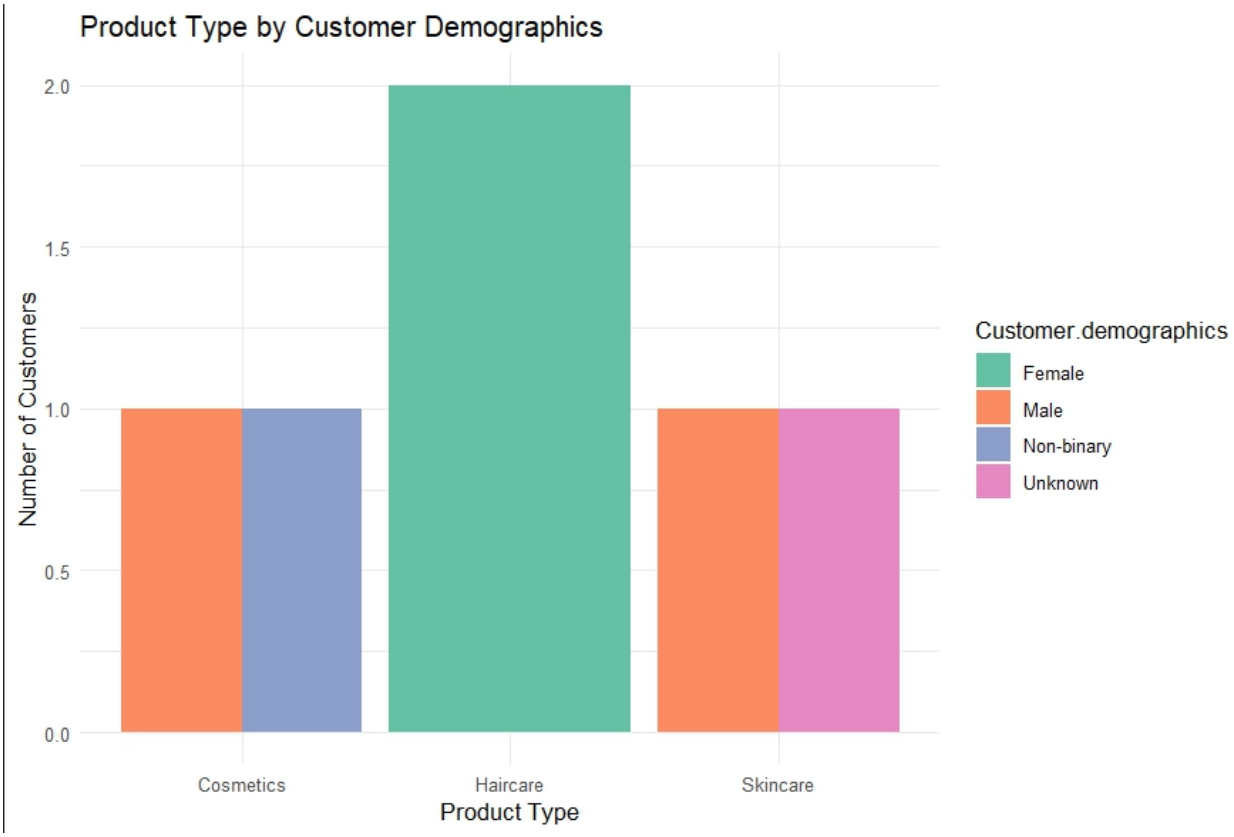
In the beauty and personal care sector, skincare, haircare, and cosmetics are important categories that each have different but related purposes. Cosmetics, which include cosmetics and other items used to improve or change look, have consistently been in demand because of rising beauty consciousness and the influence of social media. Concerns about hair damage, styling, and hygiene are frequently the driving force behind the haircare industry, which includes products like shampoos, conditioners, and treatments that are intended to preserve and enhance the health of hair. With increased knowledge of dermatological health and self-care practices, skincare products, which include moisturizers, cleansers, sunscreens, and other items, are becoming more and more popular.

According to the data, skincare brought in the most money and sold the most products, suggesting a strong customer preference that is probably the result of growing knowledge of the importance of skin health and wellness. Their economic and cultural influence in the personal care business is highlighted by industry reports showing that skincare dominates worldwide beauty sales, closely followed by cosmetics and haircare (Statista, 2024; McKinsey, 2023).

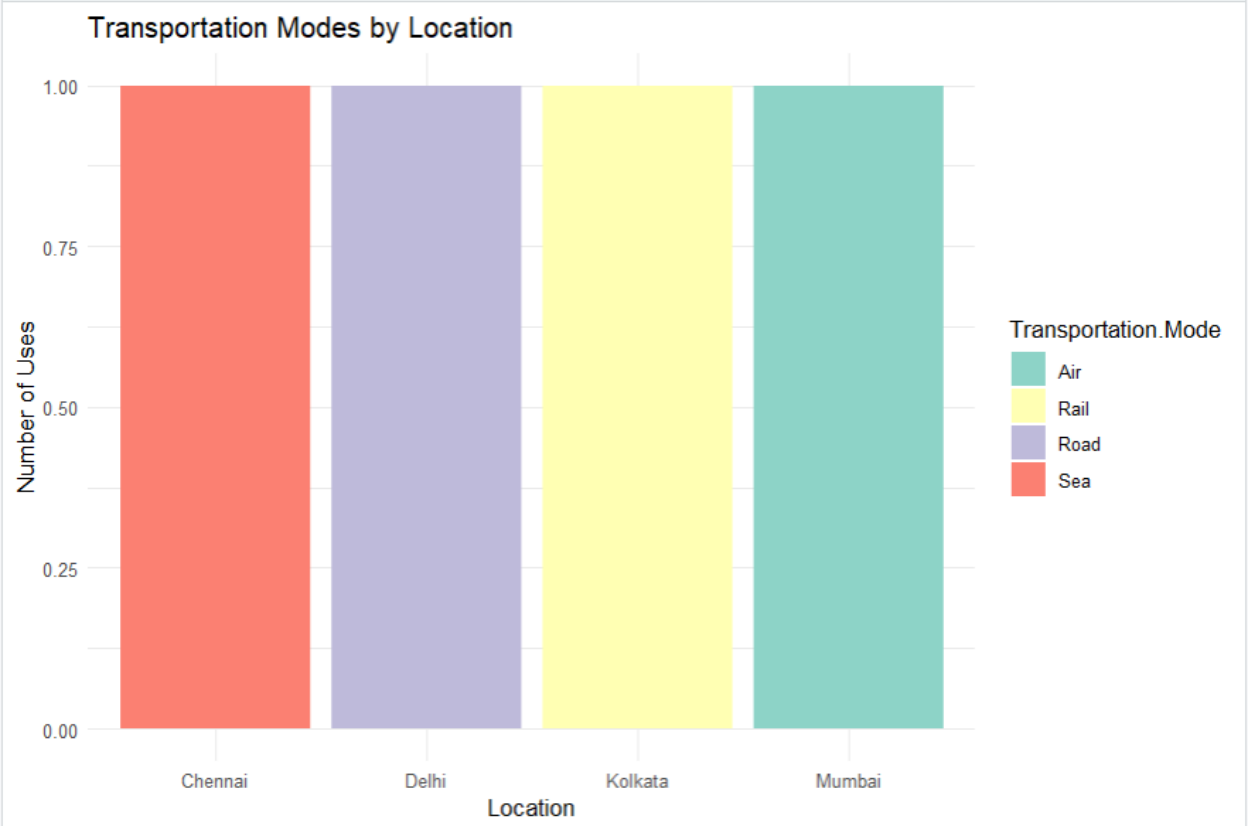

Number of Products by Location and Product Type
Count
1.00
0.75
0.50
0.25
0.00
Bangalore
Chennai
Delhi
Kolkata
Mumbai
Location
Product.type
Cosmetics
Haircare
Skincare


Revenue Generated by Location
Revenue
15000
10000
5000
0
Bangalore
Chennai
Delhi
Kolkata
Mumbai
Location

Customer demographics are crucial in determining product demand and marketing tactics, and the beauty industry survives on comprehending and satisfying the wide range of wants of its clientele. The statistics show that females represent the largest consumer group for haircare products, with overall customer demographics including Female, Male, Non-binary, and Unknown categories. This aligns with long-standing industry trends, as women have traditionally made up a significant portion of the beauty market. However, the increasing presence of male and non-binary consumers highlights a shift toward more inclusive and diverse product offerings, reflecting the evolving landscape of the beauty industry. By comprehending these demographic groups, businesses may better target their advertising, and product offers to appeal to their target audiences, which eventually increases sales and customer loyalty (Chopra & Meindl, 2019).



Another important consideration is location, which has an impact on the demand for the product as well as the logistics needed to service various markets. In addition to being significant consumer centers, cities like Mumbai, Kolkata, Delhi, and Chennai also have an impact on the best forms of transportation. For instance, large shipments to distribution centers are handled by rail and ocean modes, whereas air and truck transportation are more prevalent in metropolitan areas that need rapid replenishment. Both delivery performance and operating expenses are greatly impacted by these choices. Furthermore, the dataset demonstrates geographical differences in income generation, indicating that purchasing power and regional preferences have a direct impact on profitability. Beauty firms can improve the responsiveness and efficiency of their supply chains by matching their product distribution methods to local infrastructure and demand (Christopher, 2016).



Product types, like skincare, cosmetics, and haircare, also have an impact on supply chain choices and consumer behavior. As people's understanding of wellness and self-care has grown, so too has the demand for skincare goods, which brought in the most money in the dataset. These goods frequently need special handling and storage, which affects the decisions made about transportation. For instance, commodities that are sensitive to temperature might require delivery options that are quicker and more regulated. The intricate relationship among client demographics, geographic location, and means of transportation highlights how difficult supply chain planning is in the beauty sector. According to Chopra and Meindl (2019) and Christopher (2016), beauty organizations must incorporate data-driven insights into five fundamental areas to maintain competitive advantage and continue development as they enter new markets and cater to a wider range of demographics.

In summary, the dataset's study shows that the performance of the beauty sector is greatly influenced by several interdependent elements, including client demographics, location, product type, revenue, and modes of transportation. Mumbai stands out above all the insights as the top source of revenue for skincare items, suggesting that the industry there has a lot of potential. Male customers make up the largest demographic sector, which defies conventional wisdom about the beauty industry. Due to the need for prompt delivery and fresh products, air travel is the most popular means of transportation for this sector in Mumbai. These results highlight how crucial it is to match area logistics capabilities and population changes with supply chain objectives.

Based on this research, two important suggestions may be made. To take advantage of the high income and rising demand in this sector, beauty brands should first invest more in men's skincare marketing campaigns in Mumbai. Second, because air transportation is so important, businesses should use better forecasting and strategic warehouse positioning to cut air freight costs while ensuring timely replenishment. Enhancing client happiness and promoting further expansion in one of the most promising sectors for the beauty business are the goals of these initiatives.

The supply chain's means of transportation have a significant impact on how swiftly and affordably goods are delivered to consumers. The dataset illustrates how geography influences logistical decisions by showing how different places, including Mumbai, Kolkata, Delhi, and Chennai, use various modes of transportation, including road, air, rail, and sea. While inland cities may depend more on rail and trucking for regional distribution, those close to large airports or seaports typically favor air and maritime transport for international or accelerated goods. According to the data, supply chain efficiency can be improved, delivery speed can be maximized, and transportation costs can be decreased by matching transportation modes with location benefits (Chopra & Meindl, 2019; Christopher, 2016).

References

- Chopra, S., & Meindl, P. (2019). Supply chain management: Strategy, planning, and operation (7th ed.). Pearson.
- Christopher, M. (2016). Logistics & supply chain management (5th ed.). Pearson UK.
- **McKinsey & Company. (2023).** The beauty market bounces back. https://www.mckinsey.com/industries/consumer-packaged-goods/our-insights/the-beauty-market-bounces-back
- **Statista. (2024).** Revenue of the cosmetics and personal care market worldwide from 2018 to 2028. https://www.statista.com/statistics/585522/global-cosmetics-market-value/

Code Script in R

```
# Read the desired file (e.g., CSV inside the ZIP)
data <- read.csv("C:\\Users\\OneDrive\\Desktop\\supply_chain_data.csv")



## Load the libraries
library(MLmetrics)
library(ranger)
library(vtreat)
library(ggplot2)
library(ggthemes)
library(caret)

# EDA
names(data)
head(data)
summary(data)

# Let see if there are outliers
ggplot(data = data, aes(y=Price)) + geom_boxplot() + theme_gdocs()
```

```
# Load libraries
library(dplyr)



# Show specific columns only
data[, c("Product_Type", "Price", "Availability", "Location", "Revenue_Generated",
"Customer_demographics)]



# View dataset
print(data)

# Load required library
library(dplyr)

product_type_freq <- data %>%
  count(Product.type, sort = TRUE)

print("Frequency of Product type")
print(product_type_freq)
```

```
# 2. Summary statistics for Price
price_summary <- summary(data$Price)
print("Price Summary:")
print(price_summary)

# 3. Availability counts (In Stock vs Out of Stock)
availability_status <- data %>%
  count(Availability, sort = TRUE)
print("Product Availability:")
print(availability_status)

# 4. Number of products per Location
location_distribution <- data %>%
  count(Location, sort = TRUE)
print("Product Count by Location:")
print(location_distribution)

# 5. Average Price by Product Type
avg_price_by_type <- data %>%
  group_by(Product.type) %>%
  summarise(Average_Price = mean(Price))
print("Average Price by Product Type:")
print(avg_price_by_type)
```

```
# Bar plot for product availability
library(ggplot2)
ggplot(data, aes(x = Availability)) +
  geom_bar(fill = "skyblue") +
  ggtitle("Product Availability Distribution") +
  xlab("Availability") + ylab("Count")

# Load libraries
library(dplyr)
library(ggplot2)

# 1. Summary statistics of customer demographics
age_summary <- summary(data$Customer.demographics)
print("Customer demographocs Summary:")


# 2. Gender distribution
gender_dist <- data %>%
  count(Customer.demographics)
print("Gender Distribution:")
print(gender_dist)
```

```
# Load library
library(ggplot2)

# Create a new plot window (for Windows)
windows()

# Create the data
customer_data <- data.frame(
  Customer_Demographics = c("Female", "Male", "Non-binary", "Unknown"),
  Count = c(25, 21, 23, 31)
)

# Create and print the plot
ggplot(customer_data, aes(x = Customer_Demographics, y = Count, fill =
Customer_Demographics)) +
  geom_bar(stat = "identity") +
  theme_minimal() +
  ggtitle("Customer Demographics Distribution") +
  xlab("Customer Demographic") +
  ylab("Number of Customers") +
```

```
  theme(legend.position = "none")


# Load libraries
library(ggplot2)
library(dplyr)


# Sample data with equal lengths
data <- data.frame(
  Product.type = c("Haircare", "Skincare", "Cosmetics", "Haircare", "Skincare",
"Cosmetics"),
  Customer.demographics = c("Female", "Male", "Non-binary", "Female", "Unknown",
"Male")
)


# Count combinations of Product Type and Demographic
product_demo_count <- data %>%
  count(Product.type, Customer.demographics)


# Create grouped bar chart
ggplot(product_demo_count, aes(x = Product.type, y = n, fill = Customer.demographics))
+
  geom_bar(stat = "identity", position = "dodge") +
  theme_minimal() +
```

```
  ggtitle("Product Type by Customer Demographics") +
  xlab("Product Type") +
  ylab("Number of Customers") +
  scale_fill_brewer(palette = "Set2")
```

```
# Load required libraries
library(ggplot2)
library(dplyr)
library(tidyr)

# Example data
data <- data.frame(
  Product.type = c("Haircare", "Skincare", "Cosmetics"),
  Products.sold = c(150, 200, 180),
  Revenue = c(4500, 7000, 6500)
)  # <-- close data.frame here

# Reshape data for plotting
library(tidyr)
library(dplyr)
```

```
long_data <- data %>%
  pivot_longer(cols = c(Products.sold, Revenue),
               names_to = "Metric",
               values_to = "Value")

# Plot grouped bar chart
library(ggplot2)

ggplot(long_data, aes(x = Product.type, y = Value, fill = Metric)) +
  geom_bar(stat = "identity", position = "dodge") +
  theme_minimal() +
  ggtitle("Number of Products Sold vs Revenue Generated") +
  xlab("Product Type") +
  ylab("Value") +
  scale_fill_manual(values = c("Products.sold" = "#1f77b4", "Revenue" = "#ff7f0e"))



# Load ggplot2
library(ggplot2)

# Example data
data <- data.frame(
```

```
  Location = c("Kolkata", "Mumbai", "Chennai", "Bangalore", "Delhi"),
  Revenue = c(12000, 15000, 4500, 7000, 6000)
)

# Plot bar graph
ggplot(data, aes(x = Location, y = Revenue, fill = Location)) +
  geom_bar(stat = "identity") +
  theme_minimal() +
  ggtitle("Revenue Generated by Location") +
  xlab("Location") +
  ylab("Revenue") +
  theme(legend.position = "none")



# Load libraries
library(ggplot2)
library(dplyr)

# Example data
data <- data.frame(
  Location = c("Kolkata", "Mumbai", "Chennai", "Bangalore", "Delhi"),
  Product.type = c("Haircare", "Skincare", "Haircare", "Cosmetics", "Skincare")
)
```

```
# Count number of products by Location and Product.type
count_data <- data %>%
  count(Location, Product.type)

# Plot grouped bar chart
ggplot(count_data, aes(x = Location, y = n, fill = Product.type)) +
  geom_bar(stat = "identity", position = "dodge") +
  theme_minimal() +
  ggtitle("Number of Products by Location and Product Type") +
  xlab("Location") +
  ylab("Count") +
  scale_fill_brewer(palette = "Set2")



# Load required libraries
library(ggplot2)
library(dplyr)

# Example data
data <- data.frame(
```

```
  Location = c("Kolkata", "Mumbai", "Chennai", "Bangalore", "Delhi"),
  Customer.demographics = c("Female", "Male", "Non-binary", "Female", "Unknown")
)

# Count number of customers by Location and Demographic
count_data <- data %>%
  count(Location, Customer.demographics)

# Plot grouped bar chart
ggplot(count_data, aes(x = Location, y = n, fill = Customer.demographics)) +
  geom_bar(stat = "identity", position = "dodge") +
  theme_minimal() +
  ggtitle("Customer Demographics by Location") +
  xlab("Location") +
  ylab("Number of Customers") +
  scale_fill_brewer(palette = "Pastel1")

# Load required libraries
library(ggplot2)
library(dplyr)

# Example data
data <- data.frame(
```

```
  Location = c("Kolkata", "Mumbai", "Chennai", "Bangalore", "Delhi"),
  Route = c("Route A", "Route B", "Route A", "Route C", "Route A")
)

# Count number of routes by location
route_counts <- data %>%
  count(Location, Route)

# Create bar chart
ggplot(route_counts, aes(x = Location, y = n, fill = Route)) +
  geom_bar(stat = "identity", position = "dodge") +
  theme_minimal() +
  ggtitle("Routes by Location") +
  xlab("Location") +
  ylab("Number of Routes") +
  scale_fill_brewer(palette = "Pastel2")



# Load required libraries
library(ggplot2)
library(dplyr)

# Example data
```

```
data <- data.frame(
  Location = c("Kolkata", "Mumbai", "Chennai","Delhi"),
  Transportation.Mode = c("Rail", "Air", "Sea", "Road")
)

# Count how many times each transportation mode is used per location
mode_counts <- data %>%
  count(Location, Transportation.Mode)

# Create the bar graph
ggplot(mode_counts, aes(x = Location, y = n, fill = Transportation.Mode)) +
  geom_bar(stat = "identity", position = "dodge") +
  theme_minimal() +
  ggtitle("Transportation Modes by Location") +
  xlab("Location") +
  ylab("Number of Uses") +
  scale_fill_brewer(palette = "Set3")
```

http://127.0.0.1:39791/graphics/183095aa-73b7-411c-b378-17716d146872.png