

In [2]: !pip install plotly

Requirement already satisfied: plotly in c:\users\neha\appdata\local\programs\python\python313\lib\site-packages (6.0.1)
Requirement already satisfied: narwhals>=1.15.1 in c:\users\neha\appdata\local\programs\python\python313\lib\site-packages (from plotly) (1.38.0)
Requirement already satisfied: packaging in c:\users\neha\appdata\local\programs\python\python313\lib\site-packages (from plotly) (24.2)

In [3]: !pip install statsmodels

Collecting statsmodels

Downloading statsmodels-0.14.4-cp313-cp313-win_amd64.whl.metadata (9.5 kB)
Requirement already satisfied: numpy<3,>=1.22.3 in c:\users\neha\appdata\local\programs\python\python313\lib\site-packages (from statsmodels) (2.2.4)

Collecting scipy!=1.9.2,>=1.8 (from statsmodels)

Downloading scipy-1.15.2-cp313-cp313-win_amd64.whl.metadata (60 kB)
Requirement already satisfied: pandas!=2.1.0,>=1.4 in c:\users\neha\appdata\local\programs\python\python313\lib\site-packages (from statsmodels) (2.2.3)

Collecting patsy>=0.5.6 (from statsmodels)

Downloading patsy-1.0.1-py2.py3-none-any.whl.metadata (3.3 kB)
Requirement already satisfied: packaging>=21.3 in c:\users\neha\appdata\local\programs\python\python313\lib\site-packages (from statsmodels) (24.2)

Requirement already satisfied: python-dateutil>=2.8.2 in c:\users\neha\appdata\local\programs\python\python313\lib\site-packages (from pandas!=2.1.0,>=1.4->statsmodels) (2.9.0.post0)

Requirement already satisfied: pytz>=2020.1 in c:\users\neha\appdata\local\programs\python\python313\lib\site-packages (from pandas!=2.1.0,>=1.4->statsmodels) (2025.2)

Requirement already satisfied: tzdata>=2022.7 in c:\users\neha\appdata\local\programs\python\python313\lib\site-packages (from pandas!=2.1.0,>=1.4->statsmodels) (2025.2)

Requirement already satisfied: six>=1.5 in c:\users\neha\appdata\local\programs\python\python313\lib\site-packages (from python-dateutil>=2.8.2->pandas!=2.1.0,>=1.4->statsmodels) (1.17.0)

Downloading statsmodels-0.14.4-cp313-cp313-win_amd64.whl (9.8 MB)

```
----- 0.0/9.8 MB ? eta -:--:--
- ----- 0.3/9.8 MB ? eta -:--:--
-- ----- 0.5/9.8 MB 1.8 MB/s eta 0:00:06
---- ----- 1.0/9.8 MB 2.0 MB/s eta 0:00:05
----- ----- 1.6/9.8 MB 2.2 MB/s eta 0:00:04
----- ----- 2.4/9.8 MB 2.4 MB/s eta 0:00:04
----- ----- 2.9/9.8 MB 2.5 MB/s eta 0:00:03
----- ----- 3.4/9.8 MB 2.6 MB/s eta 0:00:03
----- ----- 3.9/9.8 MB 2.6 MB/s eta 0:00:03
----- ----- 5.0/9.8 MB 2.8 MB/s eta 0:00:02
----- ----- 5.8/9.8 MB 2.9 MB/s eta 0:00:02
----- ----- 6.6/9.8 MB 3.0 MB/s eta 0:00:02
----- ----- 7.3/9.8 MB 3.0 MB/s eta 0:00:01
----- ----- 8.4/9.8 MB 3.2 MB/s eta 0:00:01
----- -- 9.2/9.8 MB 3.2 MB/s eta 0:00:01
----- ----- 9.7/9.8 MB 3.3 MB/s eta 0:00:01
----- ----- 9.8/9.8 MB 3.1 MB/s eta 0:00:00
```

Downloading patsy-1.0.1-py2.py3-none-any.whl (232 kB)

Downloading scipy-1.15.2-cp313-cp313-win_amd64.whl (41.0 MB)

```
----- 0.0/41.0 MB ? eta -:--:--
----- 0.8/41.0 MB 3.4 MB/s eta 0:00:12
-- ----- 2.1/41.0 MB 4.8 MB/s eta 0:00:09
-- ----- 2.9/41.0 MB 4.5 MB/s eta 0:00:09
----- ----- 3.7/41.0 MB 4.4 MB/s eta 0:00:09
----- ----- 4.5/41.0 MB 4.3 MB/s eta 0:00:09
----- ----- 5.5/41.0 MB 4.3 MB/s eta 0:00:09
----- ----- 6.3/41.0 MB 4.2 MB/s eta 0:00:09
----- ----- 7.1/41.0 MB 4.3 MB/s eta 0:00:08
----- ----- 8.1/41.0 MB 4.3 MB/s eta 0:00:08
----- ----- 8.9/41.0 MB 4.2 MB/s eta 0:00:08
----- ----- 10.0/41.0 MB 4.3 MB/s eta 0:00:08
----- ----- 11.0/41.0 MB 4.4 MB/s eta 0:00:07
----- ----- 12.3/41.0 MB 4.5 MB/s eta 0:00:07
----- ----- 13.1/41.0 MB 4.4 MB/s eta 0:00:07
----- ----- 14.2/41.0 MB 4.5 MB/s eta 0:00:06
```


[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

Successfully installed patsy-1.0.1 scipy-1.15.2 statsmodels-0.14.4

Import Libraries

```
In [4]: import pandas as pd
import plotly.express as px
import plotly.io as pio
import plotly.graph_objects as go
pio.templates.default = "plotly white"
```

Read Data


```
In [5]: data = pd.read_csv("C:/Users/Neha/Desktop/portfolio projects05/Supply chain anal
```

```
In [6]: print(data.head())
```

| | Product type | SKU | Price | Availability | Number of products sold | \ |
|---|--------------|------|-----------|--------------|-------------------------|---|
| 0 | haircare | SKU0 | 69.808006 | 55 | 802 | |
| 1 | skincare | SKU1 | 14.843523 | 95 | 736 | |
| 2 | haircare | SKU2 | 11.319683 | 34 | 8 | |
| 3 | skincare | SKU3 | 61.163343 | 68 | 83 | |
| 4 | skincare | SKU4 | 4.805496 | 26 | 871 | |

| | Revenue generated | Customer demographics | Stock levels | Lead times | \ |
|---|-------------------|-----------------------|--------------|------------|---|
| 0 | 8661.996792 | Non-binary | 58 | 7 | |
| 1 | 7460.900065 | Female | 53 | 30 | |
| 2 | 9577.749626 | Unknown | 1 | 10 | |
| 3 | 7766.836426 | Non-binary | 23 | 13 | |
| 4 | 2686.505152 | Non-binary | 5 | 3 | |

| | Order quantities | ... | Location | Lead time | Production volumes | \ |
|---|------------------|-----|----------|-----------|--------------------|---|
| 0 | 96 | ... | Mumbai | 29 | 215 | |
| 1 | 37 | ... | Mumbai | 23 | 517 | |
| 2 | 88 | ... | Mumbai | 12 | 971 | |
| 3 | 59 | ... | Kolkata | 24 | 937 | |
| 4 | 56 | ... | Delhi | 5 | 414 | |

| | Manufacturing lead time | Manufacturing costs | Inspection results | \ |
|---|-------------------------|---------------------|--------------------|---|
| 0 | 29 | 46.279879 | Pending | |
| 1 | 30 | 33.616769 | Pending | |
| 2 | 27 | 30.688019 | Pending | |
| 3 | 18 | 35.624741 | Fail | |
| 4 | 3 | 92.065161 | Fail | |

| | Defect rates | Transportation modes | Routes | Costs |
|---|--------------|----------------------|---------|------------|
| 0 | 0.226410 | Road | Route B | 187.752075 |
| 1 | 4.854068 | Road | Route B | 503.065579 |
| 2 | 4.580593 | Air | Route C | 141.920282 |
| 3 | 4.746649 | Rail | Route A | 254.776159 |
| 4 | 3.145580 | Air | Route A | 923.440632 |

[5 rows x 24 columns]

Descriptive Statistics

```
In [7]: print(data.describe())
```

| | Price | Availability | Number of products sold | Revenue generated \ |
|-------|------------|--------------|-------------------------|---------------------|
| count | 100.000000 | 100.000000 | 100.000000 | 100.000000 |
| mean | 49.462461 | 48.400000 | 460.990000 | 5776.048187 |
| std | 31.168193 | 30.743317 | 303.780074 | 2732.841744 |
| min | 1.699976 | 1.000000 | 8.000000 | 1061.618523 |
| 25% | 19.597823 | 22.750000 | 184.250000 | 2812.847151 |
| 50% | 51.239831 | 43.500000 | 392.500000 | 6006.352023 |
| 75% | 77.198228 | 75.000000 | 704.250000 | 8253.976921 |
| max | 99.171329 | 100.000000 | 996.000000 | 9866.465458 |

| | Stock levels | Lead times | Order quantities | Shipping times \ |
|-------|--------------|------------|------------------|------------------|
| count | 100.000000 | 100.000000 | 100.000000 | 100.000000 |
| mean | 47.770000 | 15.960000 | 49.220000 | 5.750000 |
| std | 31.369372 | 8.785801 | 26.784429 | 2.724283 |
| min | 0.000000 | 1.000000 | 1.000000 | 1.000000 |
| 25% | 16.750000 | 8.000000 | 26.000000 | 3.750000 |
| 50% | 47.500000 | 17.000000 | 52.000000 | 6.000000 |
| 75% | 73.000000 | 24.000000 | 71.250000 | 8.000000 |
| max | 100.000000 | 30.000000 | 96.000000 | 10.000000 |

| | Shipping costs | Lead time | Production volumes \ |
|-------|----------------|------------|----------------------|
| count | 100.000000 | 100.000000 | 100.000000 |
| mean | 5.548149 | 17.080000 | 567.840000 |
| std | 2.651376 | 8.846251 | 263.046861 |
| min | 1.013487 | 1.000000 | 104.000000 |
| 25% | 3.540248 | 10.000000 | 352.000000 |
| 50% | 5.320534 | 18.000000 | 568.500000 |
| 75% | 7.601695 | 25.000000 | 797.000000 |
| max | 9.929816 | 30.000000 | 985.000000 |

| | Manufacturing lead time | Manufacturing costs | Defect rates | Costs |
|-------|-------------------------|---------------------|--------------|------------|
| count | 100.00000 | 100.000000 | 100.000000 | 100.000000 |
| mean | 14.77000 | 47.266693 | 2.277158 | 529.245782 |
| std | 8.91243 | 28.982841 | 1.461366 | 258.301696 |
| min | 1.00000 | 1.085069 | 0.018608 | 103.916248 |
| 25% | 7.00000 | 22.983299 | 1.009650 | 318.778455 |
| 50% | 14.00000 | 45.905622 | 2.141863 | 520.430444 |
| 75% | 23.00000 | 68.621026 | 3.563995 | 763.078231 |
| max | 30.00000 | 99.466109 | 4.939255 | 997.413450 |

Product type and Price

Analyzing the Supply Chain by looking at the relationship between the price of the products and the revenue generated by them:

```
In [8]: data.columns
```

```
Out[8]: Index(['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', 'Costs'],
             dtype='object')
```

```
In [10]: fig = px.scatter(data, x='Price',
                        y='Revenue generated',
                        color='Product type',
                        hover_data=['Number of products sold'],
                        trendline="ols")
fig.show()
```

Sales by Product Type

The company derives more revenue from skincare products, and the higher the price of skincare products, the more revenue they generate. Now let's have a look at the sales by product type:

```
In [14]: sales_data = data.groupby('Product type')['Number of products sold'].sum().reset_index()

pie_chart = px.pie(sales_data, values='Number of products sold', names='Product type',
                  title='Sales by Product Type',
                  hover_data=['Number of products sold'],
                  hole=0.5,
                  color_discrete_sequence=px.colors.qualitative.Pastel)

pie_chart.update_traces(textposition='inside', textinfo='percent+label')
pie_chart.show()
```

Analysis:- So 45% of the business comes from skincare products, 29.5% from haircare, and 25.5% from cosmetics.

Total Revenue by Shipping Carrier

```
In [16]: total_revenue = data.groupby('Shipping carriers')['Revenue generated'].sum().reset_index()
fig = go.Figure()
fig.add_trace(go.Bar(x=total_revenue['Shipping carriers'],
                     y=total_revenue['Revenue generated'])))
fig.update_layout(title='Total Revenue by Shipping Carrier',
                  xaxis_title='Shipping Carrier',
                  yaxis_title='Revenue Generated')
fig.show()
```

Product type

The company is using three carriers for transportation, and Carrier B helps the company in generating more revenue. Now let's have a look at the Average lead time and Average Manufacturing Costs for all products of the company:

```
In [18]: avg_lead_time = data.groupby('Product type')['Lead time'].mean().reset_index()
avg_manufacturing_costs = data.groupby('Product type')['Manufacturing costs'].me
result = pd.merge(avg_lead_time, avg_manufacturing_costs, on='Product type')
result.rename(columns={'Lead time': 'Average Lead Time', 'Manufacturing costs':
print(result)
```

| | Product type | Average Lead Time | Average Manufacturing Costs |
|---|--------------|-------------------|-----------------------------|
| 0 | cosmetics | 13.538462 | 43.052740 |
| 1 | haircare | 18.705882 | 48.457993 |
| 2 | skincare | 18.000000 | 48.993157 |

Analyzing SKUs

SKU stands for Stock Keeping Units. They're like special codes that help companies keep track of all the different

things they have for sale. Imagine you have a large toy store with lots of toys. Each toy is different and has its name and price, but when you want to know how many you have left, you need a way to identify them. So you give each toy a unique code, like a secret number only the store knows. This secret number is called SKU

Revenue generated by SKU

```
In [19]: revenue_chart = px.line(data, x='SKU',  
                                y='Revenue generated',  
                                title='Revenue Generated by SKU')  
revenue_chart.show()
```

Stock Levels by SKU

Stock levels refer to the number of products a store or business has in its inventory. Now let's have a look at the stock levels of each SKU:

```
In [20]: stock_chart = px.line(data, x='SKU',  
                                y='Stock levels',  
                                title='Stock Levels by SKU')  
stock_chart.show()
```

Order Quantity by SKU

```
In [21]: order_quantity_chart = px.bar(data, x='SKU',  
                                         y='Order quantities',  
                                         title='Order Quantity by SKU')  
order_quantity_chart.show()
```

Shipping Costs by Carrier

```
In [22]: shipping_cost_chart = px.bar(data, x='Shipping carriers',  
                                     y='Shipping costs',  
                                     title='Shipping Costs by Carrier')  
shipping_cost_chart.show()
```


Analysis : In one of the above visualizations, we discovered that Carrier B helps the company in more revenue. It is also the most costly Carrier among the three.

Cost Distribution by Transportation Mode

```
In [23]: transportation_chart = px.pie(data,
                                         values='Costs',
                                         names='Transportation modes',
                                         title='Cost Distribution by Transportation Mode',
                                         hole=0.5,
                                         color_discrete_sequence=px.colors.qualitative.Pastel1,
                                         title_text_color='black')
transportation_chart.show()
```

Analysis: So the company spends more on Road and Rail modes of transportation for the transportation of Goods.

Analyzing Defect Rate

```
In [24]: defect_rates_by_product = data.groupby('Product type')['Defect rates'].mean().re

fig = px.bar(defect_rates_by_product, x='Product type', y='Defect rates',
              title='Average Defect Rates by Product Type')
fig.show()
```

Analysis: So the defect rate of haircare products is higher.

Defect Rates by Transportation Mode

```
In [25]: pivot_table = pd.pivot_table(data, values='Defect rates',
                                       index=['Transportation modes'],
                                       aggfunc='mean')

transportation_chart = px.pie(values=pivot_table["Defect rates"],
                              names=pivot_table.index,
                              title='Defect Rates by Transportation Mode',
                              hole=0.5,
                              color_discrete_sequence=px.colors.qualitative.Pastel1)
transportation_chart.show()
```

Analysis: Road transportation results in a higher defect rate, and Air transportation has the lowest defect rate.

Summary :-

Supply Chain Analysis means analyzing various components of a Supply Chain to understand how to improve the effectiveness of the Supply Chain to create more value for customers.

In []: