GOVERNMENT COLLEGE OF ENGINEERING ERODE

அரசினர் பொறியியல் கல்லூரி, ஈரோடு
Government College of Engineering, Erode
(Approved by AICTE, New Delhi and Affiliated to Anna University, Chennai)

B.E Electronics and Communication Engineering

PRODUCT SALES ANALYSIS

Name of the Students:

Surendhar S

University Register no:

731121106049

Under the mentor of

Dr.M.Poongothai

Department of Information Technology(IT)

Department of Electronics and Communication Engineering

Government College of Engineering

Erode ,PO ,near Vasavi College,TamilNadu-638316,

Affiliated to Anna University ,Chennai.

Technology name: PRODUCT SALES ANALYSIS

Project Definition:

Product sales analysis is the process of examining and evaluating data related to the sales of a particular product or a group of products within a business. The primary goal of product sales analysis is to gain insights into how well a product is performing in the market and to make informed decisions to improve sales and profitability.

Project Overview:

The Product Sales Analysis Data Analytics Project with Cognos is a strategic initiative aimed at leveraging the power of IBM Cognos, a robust business intelligence and analytics platform, to gain deep insights into a company's product sales performance. This project involves the systematic collection, processing, analysis, and visualization of sales data to drive data-driven decision-making, enhance business profitability, and optimize product-related strategies.

Project Objectives:

Sales Performance Assessment:

The primary objective of this project is to assess the performance of products in terms of revenue, units sold, and profitability. This involves a comprehensive analysis of historical sales data.

Customer Segmentation:

Utilize Cognos to segment customers based on demographics, purchasing behavior, and geography. Understanding customer segments helps tailor marketing strategies and product offerings.

Trend Analysis:

Identify sales trends, seasonality, and cyclical patterns using historical data. This enables the business to anticipate and plan for future sales fluctuations.

Competitive Benchmarking:

Compare the company's product sales with those of competitors. Identify strengths, weaknesses, opportunities, and threats in the market.

Inventory Management:

Optimize inventory levels by analyzing sales trends and forecasting future demand. Prevent overstocking or understocking of products.

Pricing Strategy Optimization:

Use data analytics to determine the most effective pricing strategies that maximize revenue while maintaining competitiveness.

Marketing Campaign Effectiveness:

Analyze the impact of marketing campaigns on product sales. Identify which campaigns generated the highest ROI and customer engagement.

Project Phases:

Data Gathering and Integration:

Collect relevant sales data from various sources, including CRM systems, POS terminals, and online sales platforms. Integrate and clean the data to ensure accuracy.

Data Modeling:

Create data models within IBM Cognos to support analysis. This involves defining data structures and relationships for efficient querying.

Descriptive Analysis:

Generate descriptive reports and dashboards in Cognos to provide an overview of sales performance, highlighting key metrics and KPIs.

Advanced Analytics:

Apply advanced analytics techniques such as regression analysis, time series forecasting, and customer segmentation within Cognos to derive actionable insights.

Visualization:

Develop interactive and visually appealing dashboards and reports using Cognos' reporting and visualization capabilities. These visuals should make complex data easy to understand for stakeholders.

Predictive Modeling:

Utilize predictive modeling to forecast future sales trends, allowing for proactive decision-making.

Monitoring and Continuous Improvement:

Implement ongoing monitoring of sales data and regularly update the analysis to adapt to changing market conditions and business goals.

MACHINE LEARNING

DEFINITION OF MACHINE LEARNING :

Machine learning is a subfield of artificial intelligence (AI) that focuses on the development of algorithms and models that enable computers to learn and make predictions or decisions without being explicitly programmed. Machine learning systems use data and statistical techniques to improve their performance on a specific task over time. The primary goal of machine learning is to develop algorithms that can generalize from data, allowing them to make accurate predictions or decisions on new, unseen data.

LINEAR REGRESSION : [Supervised]

Linear regression is one of the fundamental techniques in machine learning and statistics used for modeling the relationship between a dependent variable (target) and one or more independent variables (features or predictors). It is a type of supervised learning algorithm that aims to find the linear relationship between the features and the target variable. Linear regression is widely used for tasks such as predicting house prices, estimating sales figures, and understanding the relationship between variables in various fields.

ABOUT DATASET

CONTEXT :

This dataset contain various details of products sold at a store. These type of datasets are studied to find out the patterns in the selling structure and profit earned from them.

CONTENT :

- Order_ID : A specific ID given to each product
- Order_Priority : Priority of the product
- Order_Quantity: No of product items sold.
- Ship_Mode: Divided in two categories - Express Air and Regular Air
- Profit: Profit earned from the sale
- Customer_Name: Name of the customer purchasing the products
- Region: Region to which the customer belongs
- Customer_Segment: Divided as per the size of business
- Product_Category: Divided according to the usage of the product
- Product_Sub-Category: Divided according to the usage of the product
- Product_Name: Name of the product
- Product_Container: Type of container in which the product is shipped.

LINEAR REGRESSION – PYTHON PROGRAM

```
# Import necessary libraries

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt
```

```
from sklearn.model_selection import train_test_split

from sklearn.linear_model import LinearRegression

from sklearn.preprocessing import OneHotEncoder

from sklearn.compose import ColumnTransformer


# Load and preprocess the sales data

data = pd.read_csv("D:\IBM PROJECT\productsalesanalysis.csv")


# Select the features and target variable

X = data[['Product_Name', 'Region', 'Customer_Segment', 'Product_Category', 'Product_Sub-Category', 'Sales', 'Product_Container', 'Ship_Mode']]

Y = data['Profit']


# Specify which features are categorical

categorical_features = ['Product_Name', 'Region', 'Customer_Segment', 'Product_Category', 'Product_Sub-Category', 'Product_Container', 'Ship_Mode']


# One-hot encode categorical features

ct = ColumnTransformer(transformers=[('encoder', OneHotEncoder(), categorical_features)], remainder='passthrough')

X = ct.fit_transform(X)


# Avoiding the dummy variable trap by dropping one column for each one-hot encoded feature

X = X[:, 1:]


# Split the data into training and testing sets

X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_state=42)


# Train the linear regression model

model = LinearRegression()

model.fit(X_train, Y_train)


# Ask the user for input: Region

region_input = input("Enter a Region: ")
```

```
        # Filter the dataset for the specified region
        filtered_data = data[data['Region'] == region_input]

        # If there are no records for the specified region, inform the user
        if filtered_data.shape[0] == 0:
            print("No records found for the specified region.")
        else:
            # Use the model to predict profit for the filtered data
            X_filtered = filtered_data[['Product_Name', 'Region', 'Customer_Segment', 'Product_Category', 'Product_Sub-Category', 'Sales', 'Product_Container', 'Ship_Mode']]
            X_filtered = ct.transform(X_filtered)
            X_filtered = X_filtered[:, 1:]
            y_pred_filtered = model.predict(X_filtered)

            # Sort the filtered data by predicted profit in descending order
            sorted_data = filtered_data.copy()
            sorted_data['Predicted_Profit'] = y_pred_filtered
            sorted_data = sorted_data.sort_values(by='Predicted_Profit', ascending=False)

            # Get the top 5 and bottom 5 records
            top_5 = sorted_data.head(5)
            bottom_5 = sorted_data.tail(5)

            # Print the top 5 predicted profit values and corresponding Order_IDs, Product Names
            print("\nTop 5 Predicted Profit Values:")
            for index, row in top_5.iterrows():
                print(f"Order_ID: {row['Order_ID']}, Product Name: {row['Product_Name']}, Predicted Profit: {row['Predicted_Profit']:.2f}")

            # Print the bottom 5 predicted profit values and corresponding Order_IDs, Product Names
            print("\nBottom 5 Predicted Profit Values:")
            for index, row in bottom_5.iterrows():
                print(f"Order_ID: {row['Order_ID']}, Product Name: {row['Product_Name']}, Predicted Profit: {row['Predicted_Profit']:.2f}")
```

```
# Create a scatter plot for visualization
plt.figure(figsize=(8, 6))
plt.scatter(filtered_data['Profit'], y_pred_filtered, color='blue', marker='o')
plt.title('Actual vs. Predicted Profit for the Specified Region')
plt.xlabel('Actual Profit')
plt.ylabel('Predicted Profit')

# Add a diagonal line for reference (perfect predictions)
plt.plot([filtered_data['Profit'].min(), filtered_data['Profit'].max()], [filtered_data['Profit'].min(), filtered_data['Profit'].max()], 'k--', lw=2)

# Show the plot
plt.show()
```

OUTPUT 1 :

Enter a Region: Atlantic

Top 5 Predicted Profit Values:

Order_ID: 25315, Product Name: Holmes Harmony HEPA Air Purifier for 17 x 20 Room, Predicted Profit: 3450.39

Order_ID: 4416, Product Name: GBC DocuBind P100 Manual Binding Machine, Predicted Profit: 2529.74

Order_ID: 28836, Product Name: 688, Predicted Profit: 2340.31

Order_ID: 7015, Product Name: T65, Predicted Profit: 2253.34

Order_ID: 59234, Product Name: Kensington 7 Outlet MasterPiece Power Center, Predicted Profit: 2059.58

Bottom 5 Predicted Profit Values:

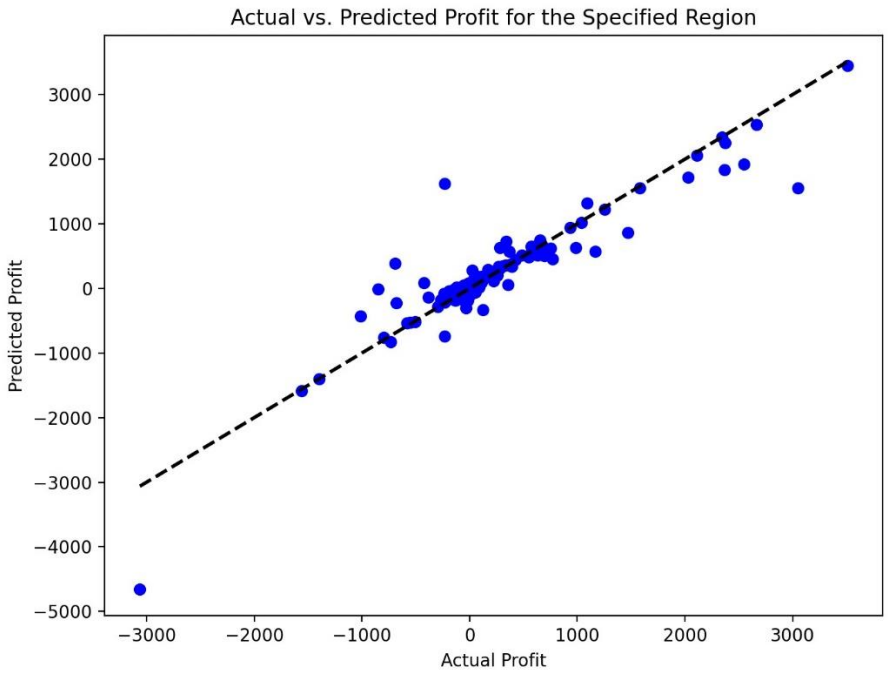
Order_ID: 8007, Product Name: StarTAC 8000, Predicted Profit: -764.00

Order_ID: 36677, Product Name: Canon Image Class D660 Copier, Predicted Profit: -825.96

Order_ID: 13604, Product Name: Eldon ClusterMat Chair Mat with Cordless Antistatic Protection, Predicted Profit: -1399.30

Order_ID: 53894, Product Name: Lesro Sheffield Collection Coffee Table, End Table, Center Table, Corner Table, Predicted Profit: -1581.44

Order_ID: 32199, Product Name: Canon imageCLASS 2200 Advanced Copier, Predicted Profit: -4656.02



OUTPUT 2:

Enter a Region: Northwest Territories

Top 5 Predicted Profit Values:

Order_ID: 21383, Product Name: Polycom ViaVideo™ Desktop Video Communications Unit, Predicted Profit: 7432.30

Order_ID: 12419, Product Name: Polycom ViewStation™ Adapter H323 Videoconferencing Unit, Predicted Profit: 5309.53

Order_ID: 39364, Product Name: Fellowes PB500 Electric Punch Plastic Comb Binding Machine with Manual Bind, Predicted Profit: 4914.49

Order_ID: 9927, Product Name: Lifetime Advantage™ Folding Chairs, 4/Carton, Predicted Profit: 3379.05

Order_ID: 30658, Product Name: Sharp AL-1530CS Digital Copier, Predicted Profit: 3345.52

Bottom 5 Predicted Profit Values:

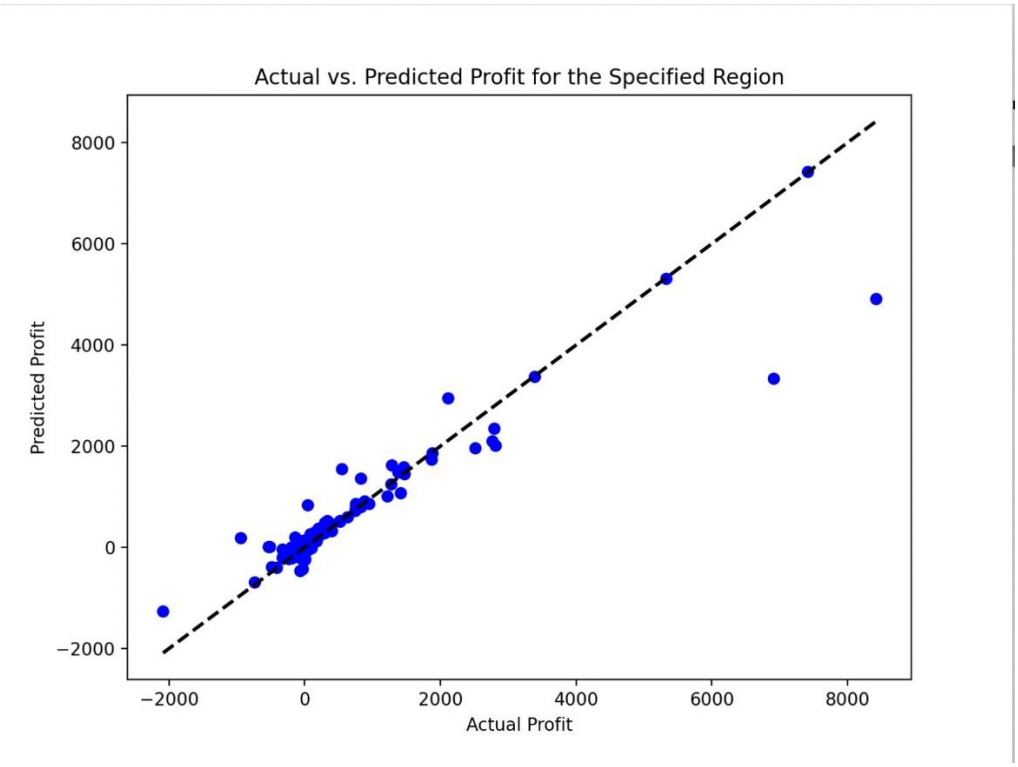
Order_ID: 36134, Product Name: R280, Predicted Profit: -400.19

Order_ID: 19138, Product Name: Recycled Eldon Regeneration Jumbo File, Predicted Profit: -430.95

Order_ID: 18144, Product Name: Eldon Portable Mobile Manager, Predicted Profit: -463.42

Order_ID: 36646, Product Name: Tennsco Industrial Shelving, Predicted Profit: -686.57

Order_ID: 41696, Product Name: Hoover Portapower™ Portable Vacuum, Predicted Profit: -1263.36



OUTPUT 3 :

Enter a Region: Nunavut

Top 5 Predicted Profit Values:

Order_ID: 7110, Product Name: SAFCO Arco Folding Chair, Predicted Profit: 1724.97

Order_ID: 483, Product Name: R380, Predicted Profit: 1253.75

Order_ID: 7430, Product Name: #10 White Business Envelopes,4 1/8 x 9 1/2, Predicted Profit: 389.30

Order_ID: 1344, Product Name: LX 788, Predicted Profit: 380.23

Order_ID: 7906, Product Name: 3M Office Air Cleaner, Predicted Profit: 297.28

Bottom 5 Predicted Profit Values:

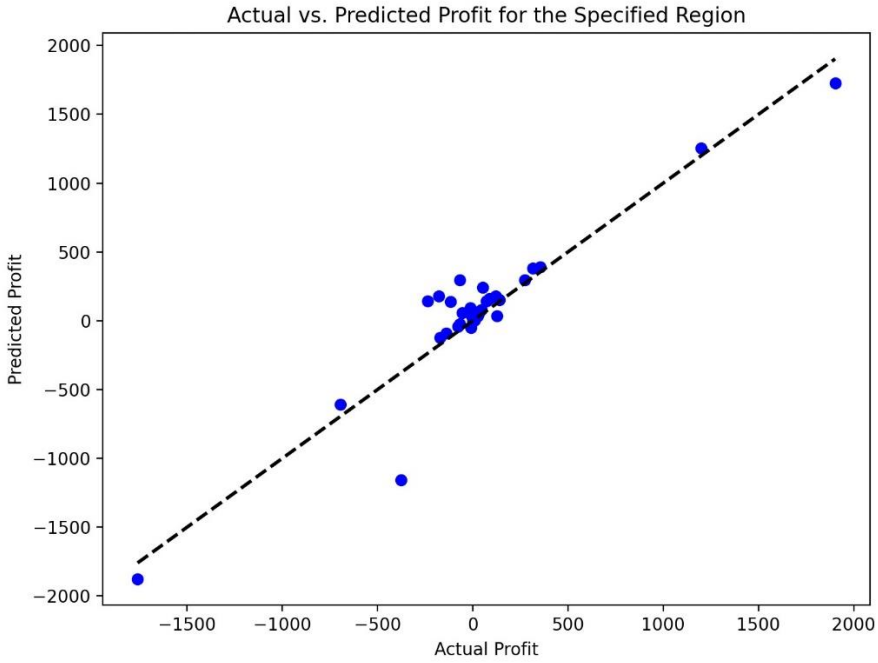
Order_ID: 6916, Product Name: Crate-A-Files™, Predicted Profit: -90.12

Order_ID: 1539, Product Name: GBC Pre-Punched Binding Paper, Plastic, White, 8-1/2" x 11", Predicted Profit: -124.12

Order_ID: 643, Product Name: SAFCO Commercial Wire Shelving, Black, Predicted Profit: -606.73

Order_ID: 4612, Product Name: Hoover Portapower™ Portable Vacuum, Predicted Profit: -1157.06

Order_ID: 6116, Product Name: High Speed Automatic Electric Letter Opener, Predicted Profit: -1876.85



OUTPUT 4 :

Enter a Region: Ontario

Top 5 Predicted Profit Values:

Order_ID: 41059, Product Name: Polycom VoiceStation 100, Predicted Profit: 5457.91

Order_ID: 33250, Product Name: Kensington 7 Outlet MasterPiece Power Center, Predicted Profit: 2494.59

Order_ID: 31238, Product Name: Hoover WindTunnel™ Plus Canister Vacuum, Predicted Profit: 2259.20

Order_ID: 3109, Product Name: Talkabout T8097, Predicted Profit: 1505.17

Order_ID: 56101, Product Name: GBC Therma-A-Bind 250T Electric Binding System, Predicted Profit: 1344.24

Bottom 5 Predicted Profit Values:

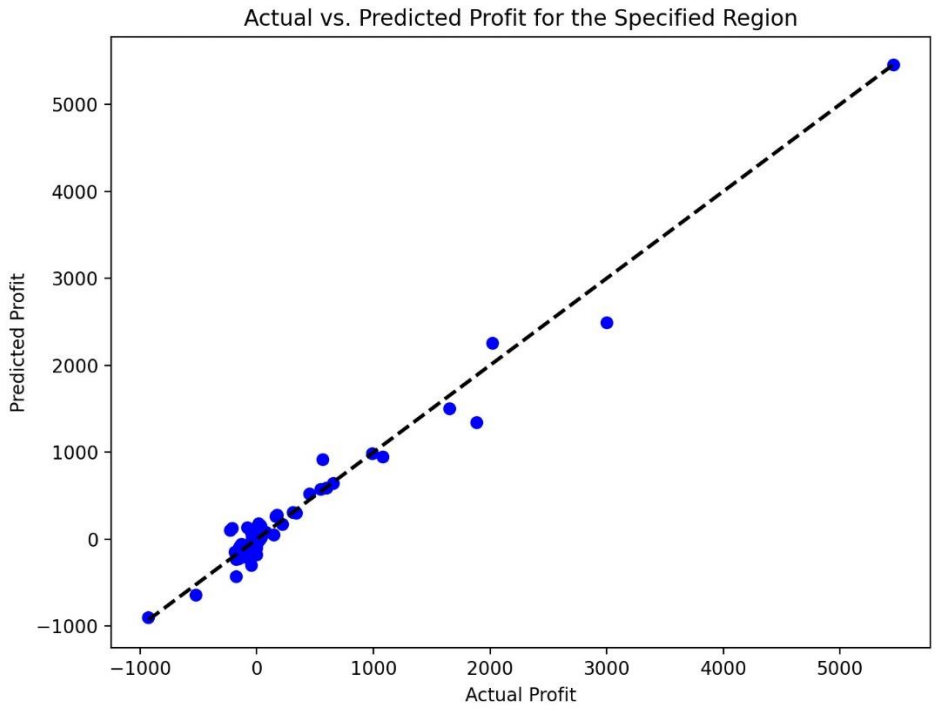
Order_ID: 51780, Product Name: Fellowes Twister Kit, Gray/Clear, 3/pkg, Predicted Profit: -226.34

Order_ID: 33255, Product Name: Xerox 4200 Series MultiUse Premium Copy Paper (20Lb. and 84 Bright), Predicted Profit: -301.82

Order_ID: 28290, Product Name: Fellowes Super Stor/Drawer® Files, Predicted Profit: -428.80

Order_ID: 29318, Product Name: SAFCO Commercial Wire Shelving, Black, Predicted Profit: -640.77

Order_ID: 50533, Product Name: Space Solutions Commercial Steel Shelving, Predicted Profit: -900.21



OUTPUT 5 :

Enter a Region: Prarie

Top 5 Predicted Profit Values:

Order_ID: 48614, Product Name: T28 WORLD, Predicted Profit: 1673.59

Order_ID: 24386, Product Name: Gyration Ultra Cordless Optical Suite, Predicted Profit: 1659.34

Order_ID: 54753, Product Name: T39m, Predicted Profit: 1374.12

Order_ID: 45248, Product Name: T28 WORLD, Predicted Profit: 922.79

Order_ID: 50404, Product Name: 24 Capacity Maxi Data Binder Racks, Pearl, Predicted Profit: 913.25

Bottom 5 Predicted Profit Values:

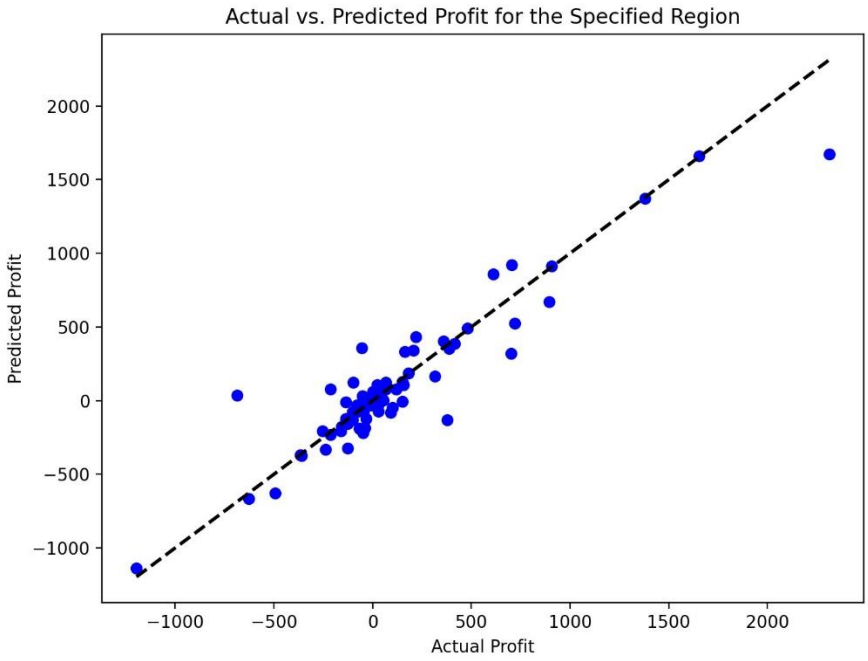
Order_ID: 50404, Product Name: Lesro Round Back Collection Coffee Table, End Table, Predicted Profit: -366.73

Order_ID: 26016, Product Name: Micro Innovations Media Access Pro Keyboard, Predicted Profit: -372.62

Order_ID: 5538, Product Name: 8290, Predicted Profit: -626.80

Order_ID: 55138, Product Name: Tennsco Industrial Shelving, Predicted Profit: -663.71

Order_ID: 54501, Product Name: Tennsco Commercial Shelving, Predicted Profit: -1138.62



OUTPUT 6 :

Enter a Region: West

Top 5 Predicted Profit Values:

Order_ID: 7427, Product Name: Hoover WindTunnel™ Plus Canister Vacuum, Predicted Profit: 4276.72

Order_ID: 52035, Product Name: Canon PC1060 Personal Laser Copier, Predicted Profit: 3912.76

Order_ID: 1221, Product Name: GBC DocuBind TL300 Electric Binding System, Predicted Profit: 3674.94

Order_ID: 22657, Product Name: Hayes Optima 56K V.90 Internal Voice Modem, Predicted Profit: 3352.21

Order_ID: 5988, Product Name: Sharp AL-1530CS Digital Copier, Predicted Profit: 3280.85

Bottom 5 Predicted Profit Values:

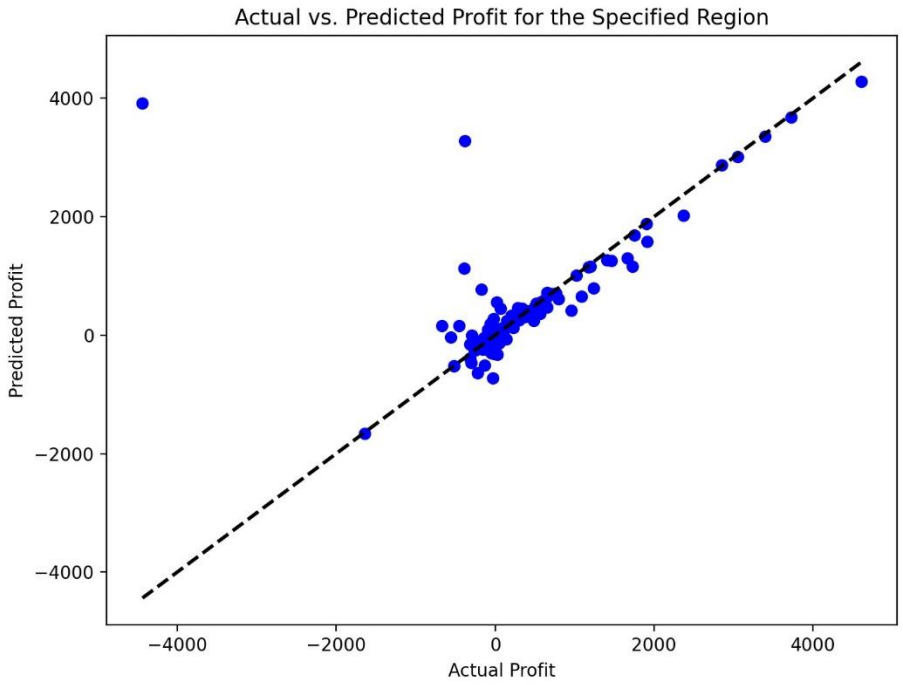
Order_ID: 11553, Product Name: Gould Plastics 9-Pocket Panel Bin, 18-3/8w x 5-1/4d x 20-1/2h, Black, Predicted Profit: -516.39

Order_ID: 39301, Product Name: Adesso Programmable 142-Key Keyboard, Predicted Profit: -519.57

Order_ID: 38758, Product Name: Euro Pro Shark Stick Mini Vacuum, Predicted Profit: -640.95

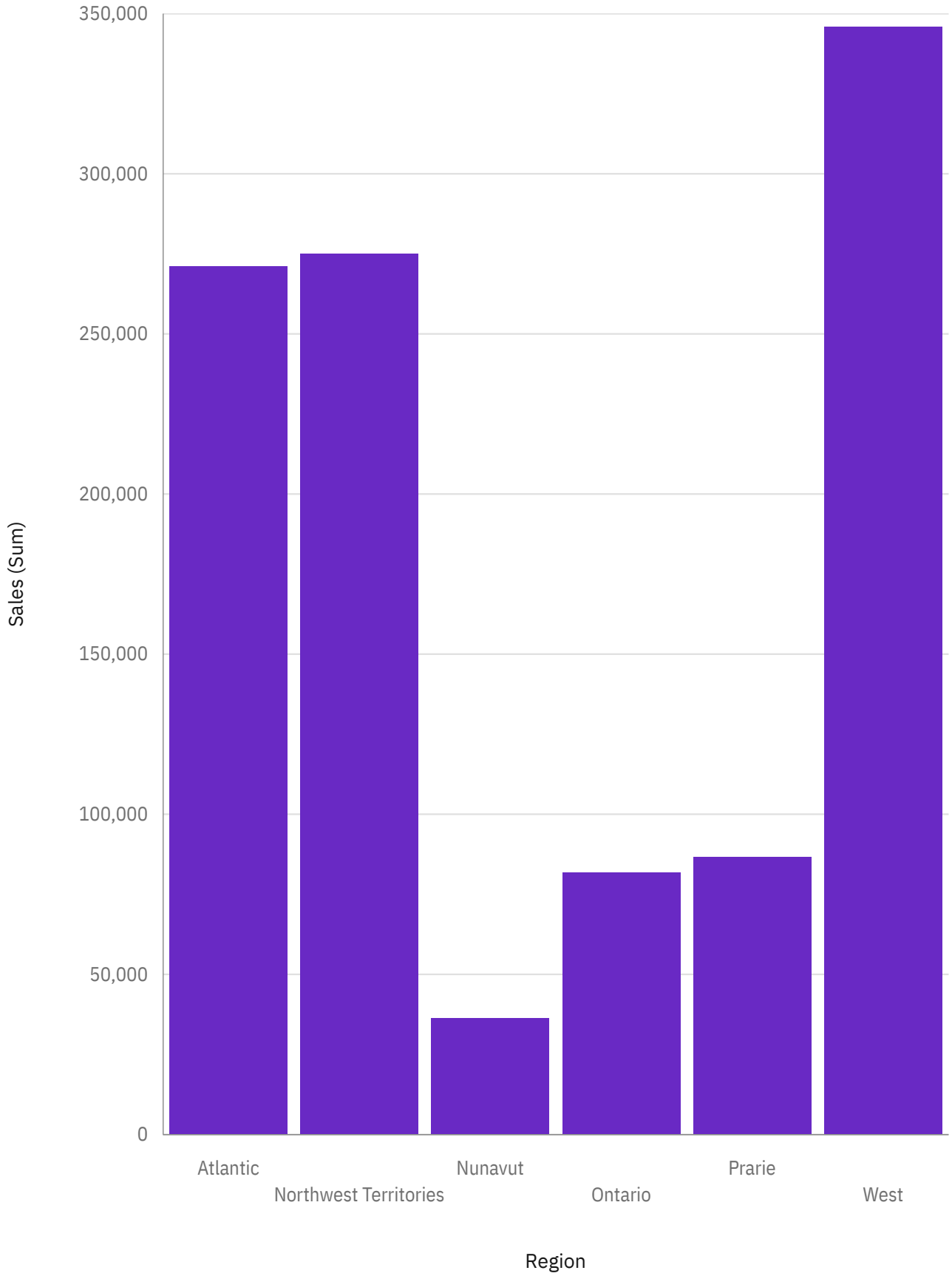
Order_ID: 22980, Product Name: Eldon Simplefile® Box Office®, Predicted Profit: -727.96

Order_ID: 44320, Product Name: Laminate Occasional Tables, Predicted Profit: -1665.66


Actual vs. Predicted Profit for the Specified Region
Predicted Profit
Actual Profit
4000
2000
0
−2000
−4000
−4000
−2000
0
2000
4000

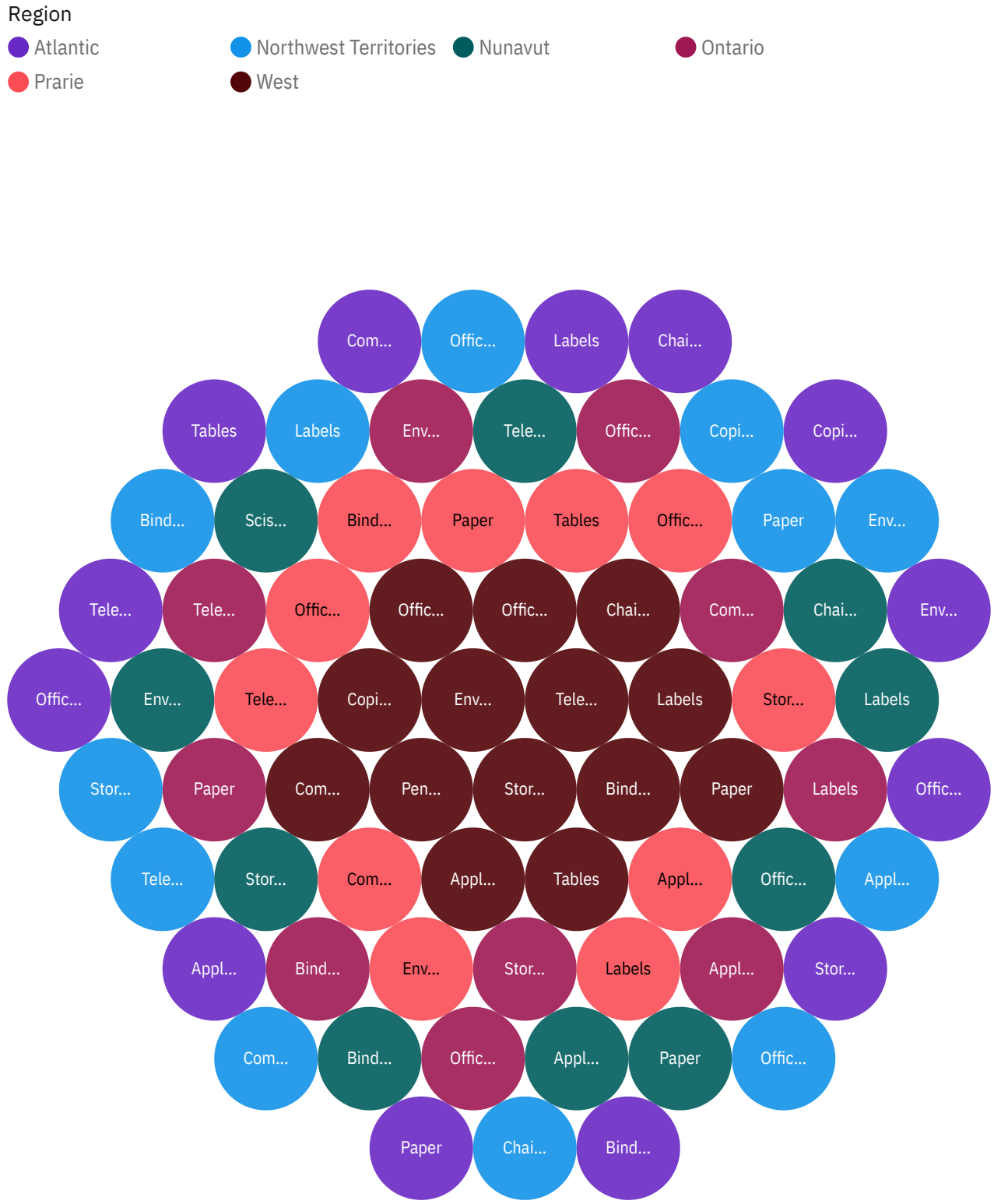
Tab 1

Sales by Region



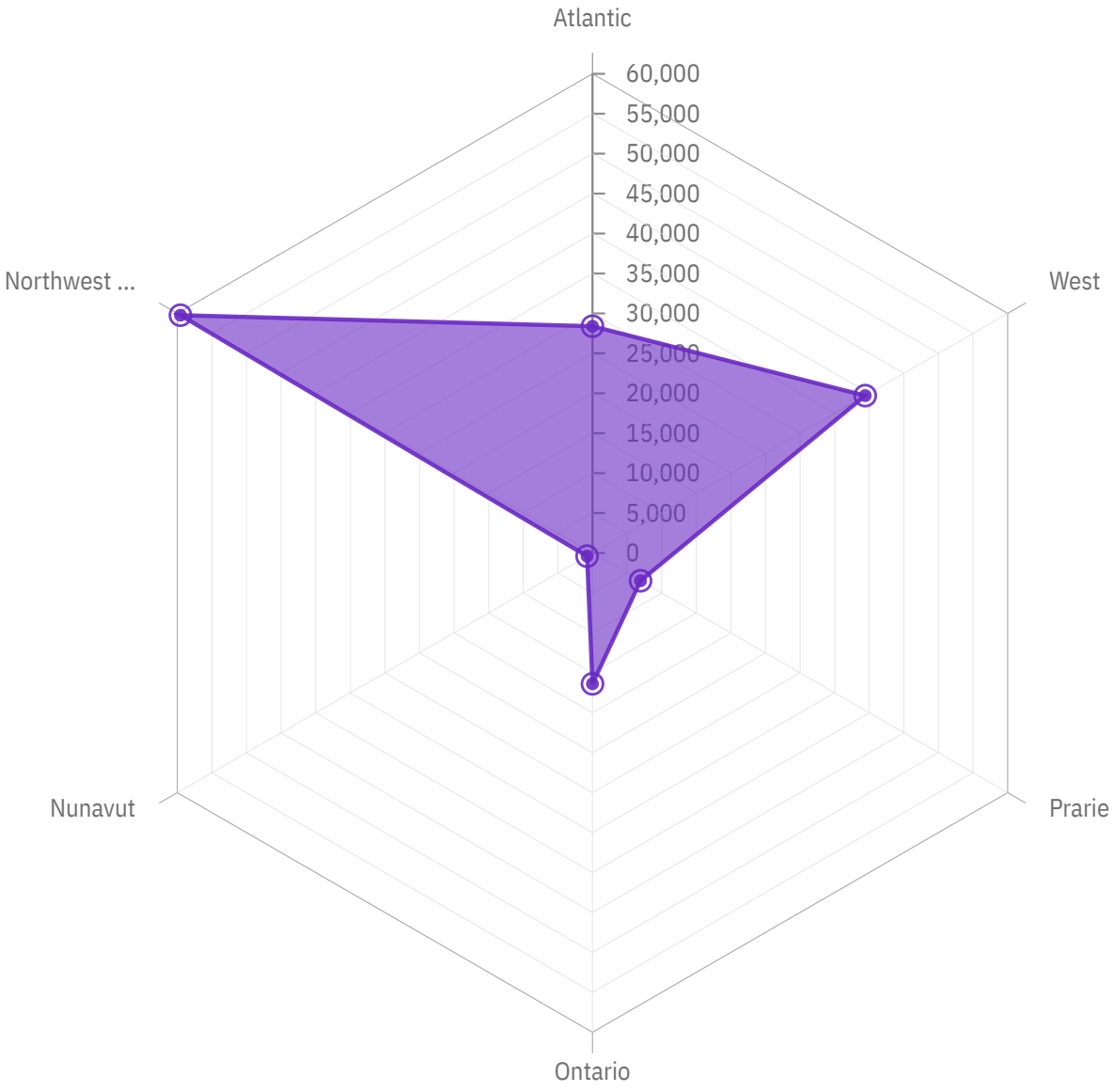
Tab 3

Product_Sub-Category colored by Region



Tab 2

Profit by Region

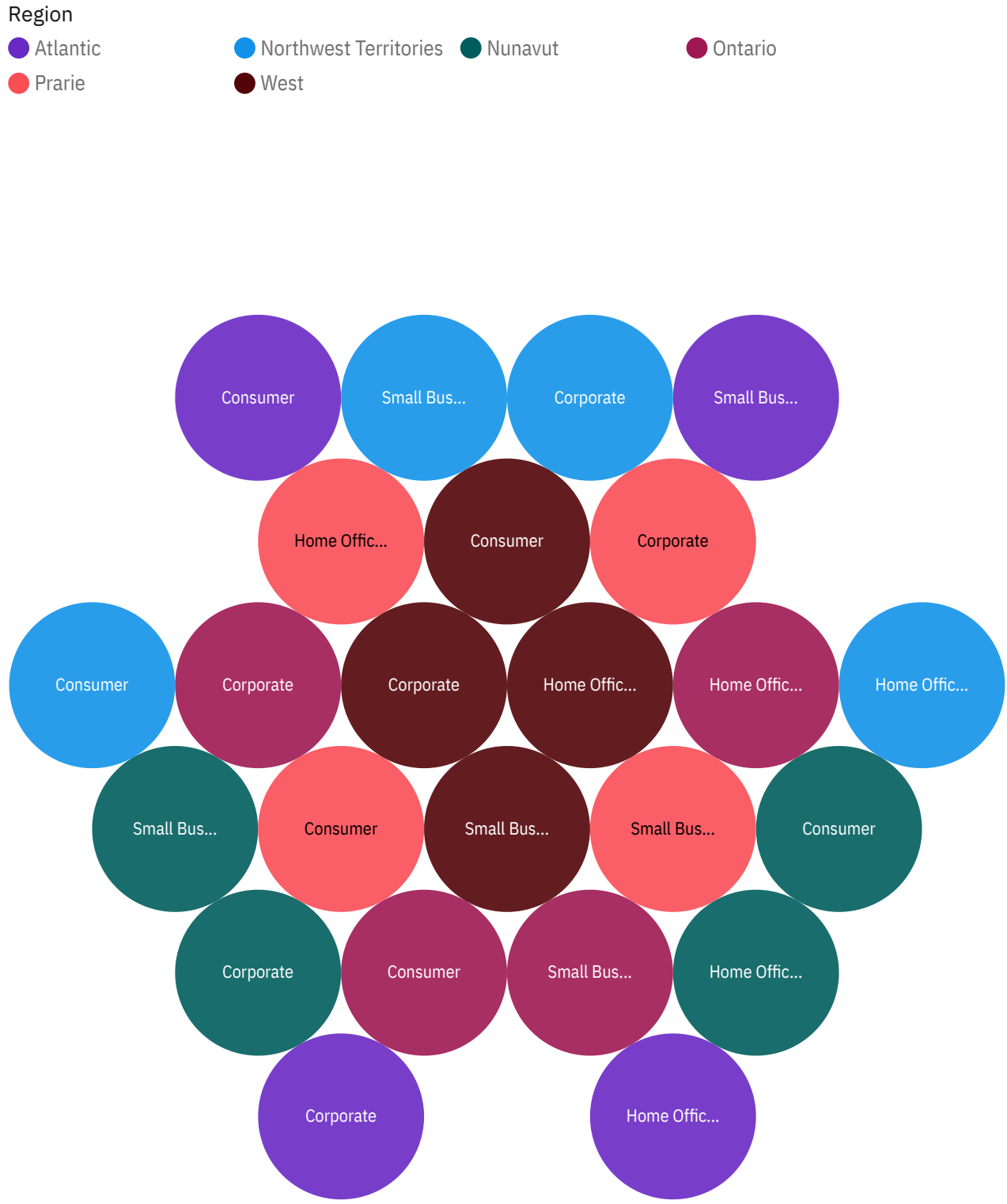


Tab 4

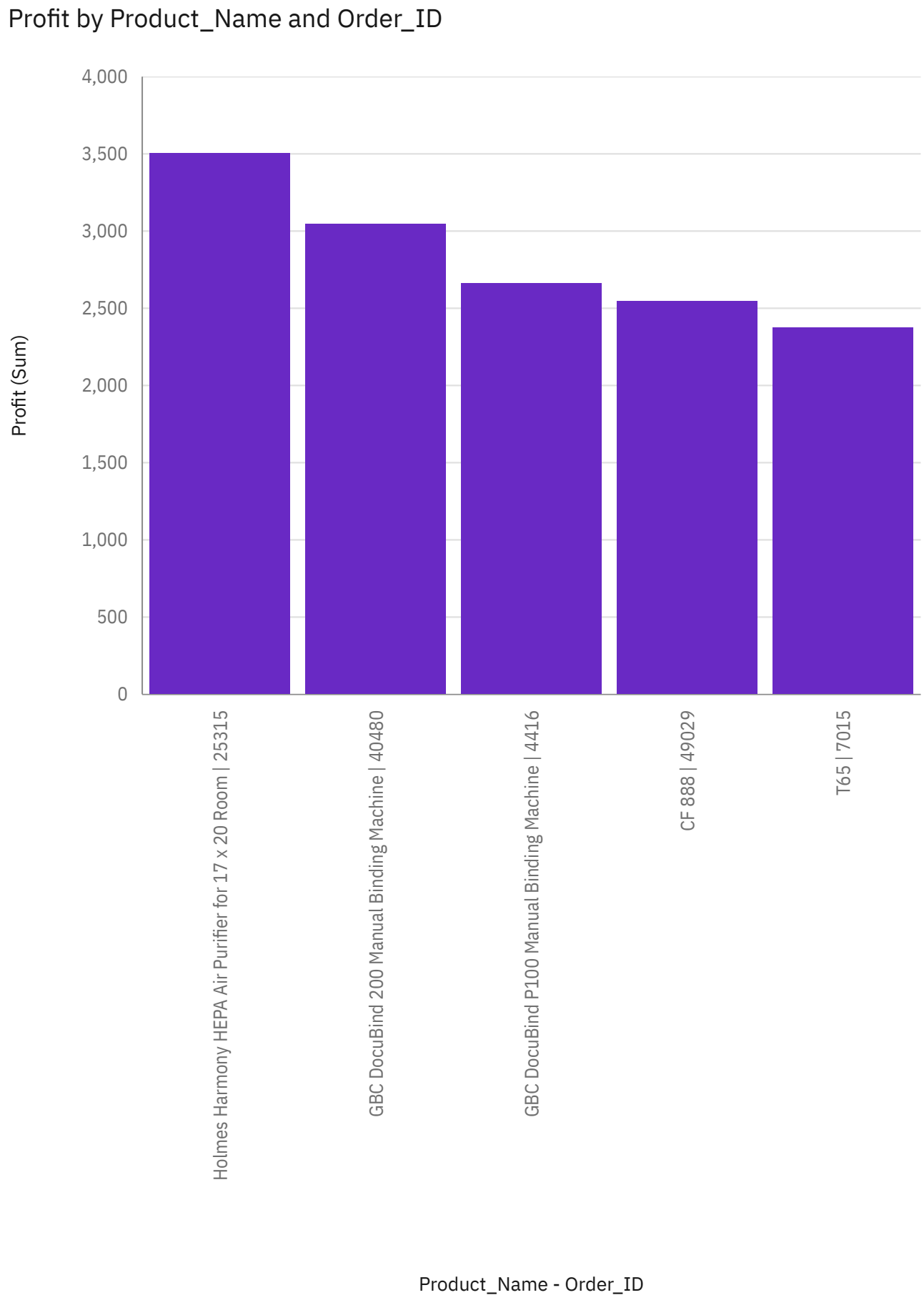
Product_Category colored by Region

Tab 5

Customer_Segment colored by Region

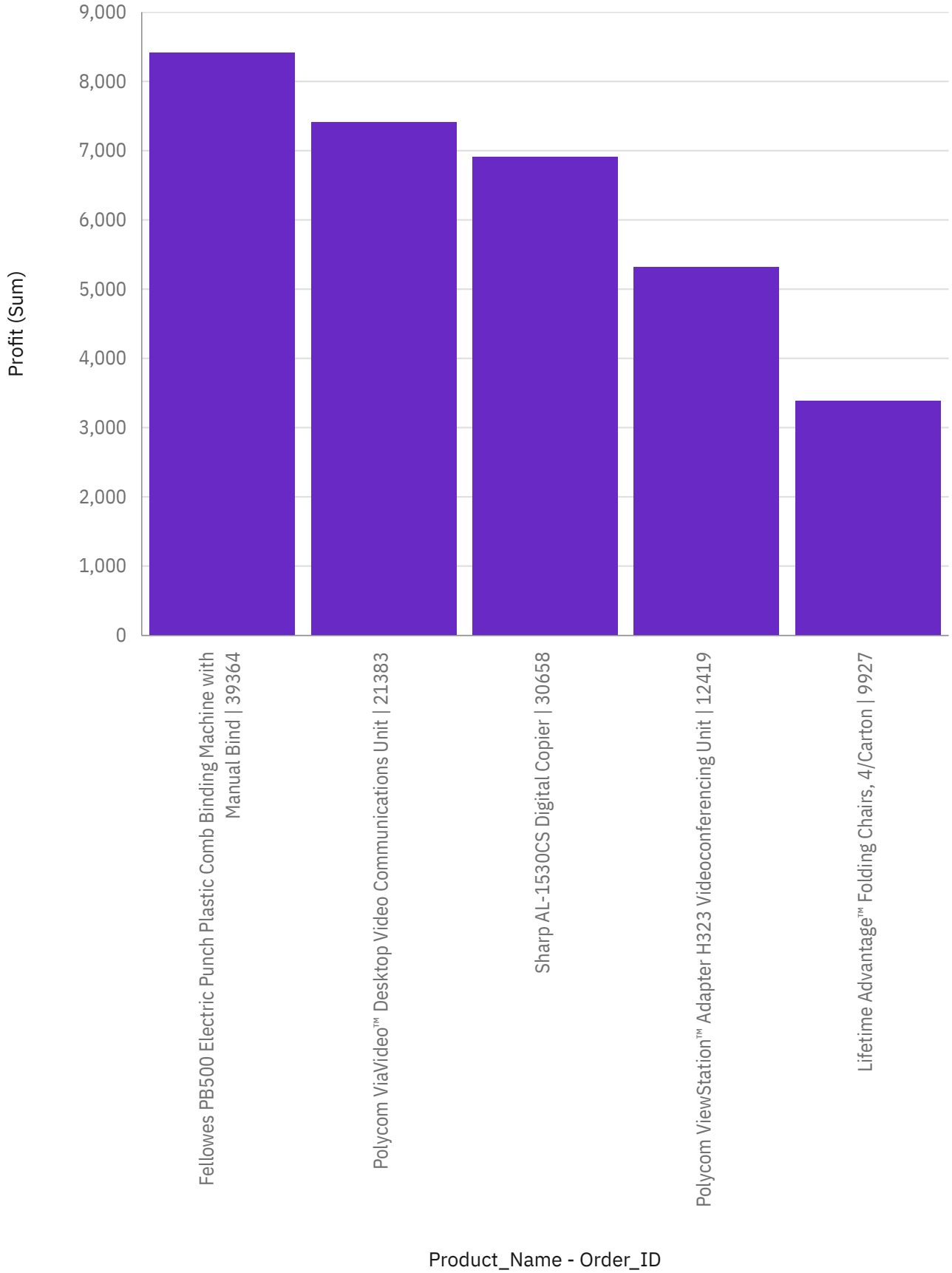


Atlantic-P

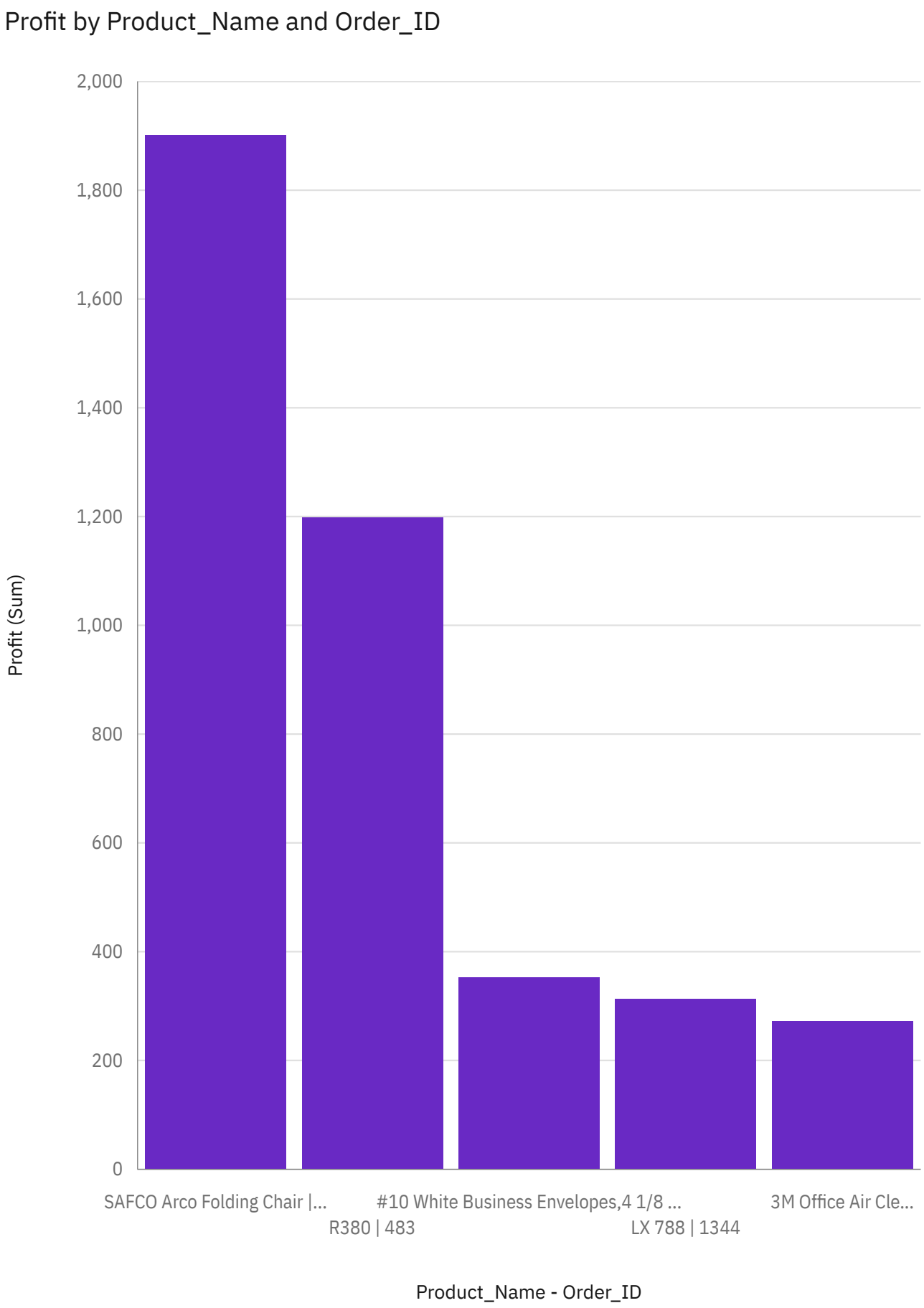

Profit by Product_Name and Order_ID
Profit (Sum)
4,000
3,500
3,000
2,500
2,000
1,500
1,000
500
0
Holmes Harmony HEPA Air Purifier for 17 x 20 Room | 25315
GBC DocuBind 200 Manual Binding Machine | 40480
GBC DocuBind P100 Manual Binding Machine | 4416
CF 888 | 49029
T65 | 7015
Product_Name - Order_ID

Northwest territories-P

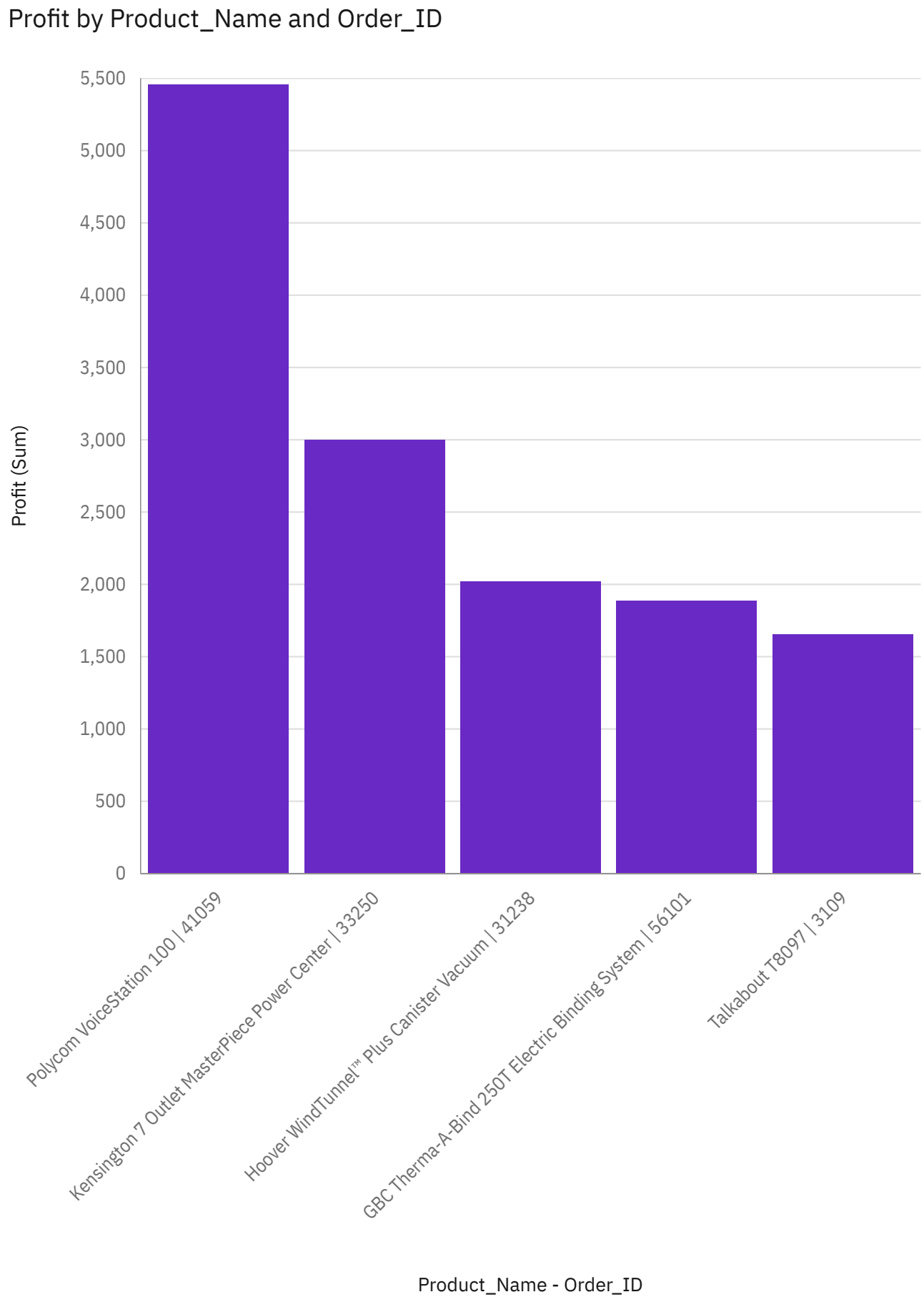
Profit by Product_Name and Order_ID



Nunavut-P

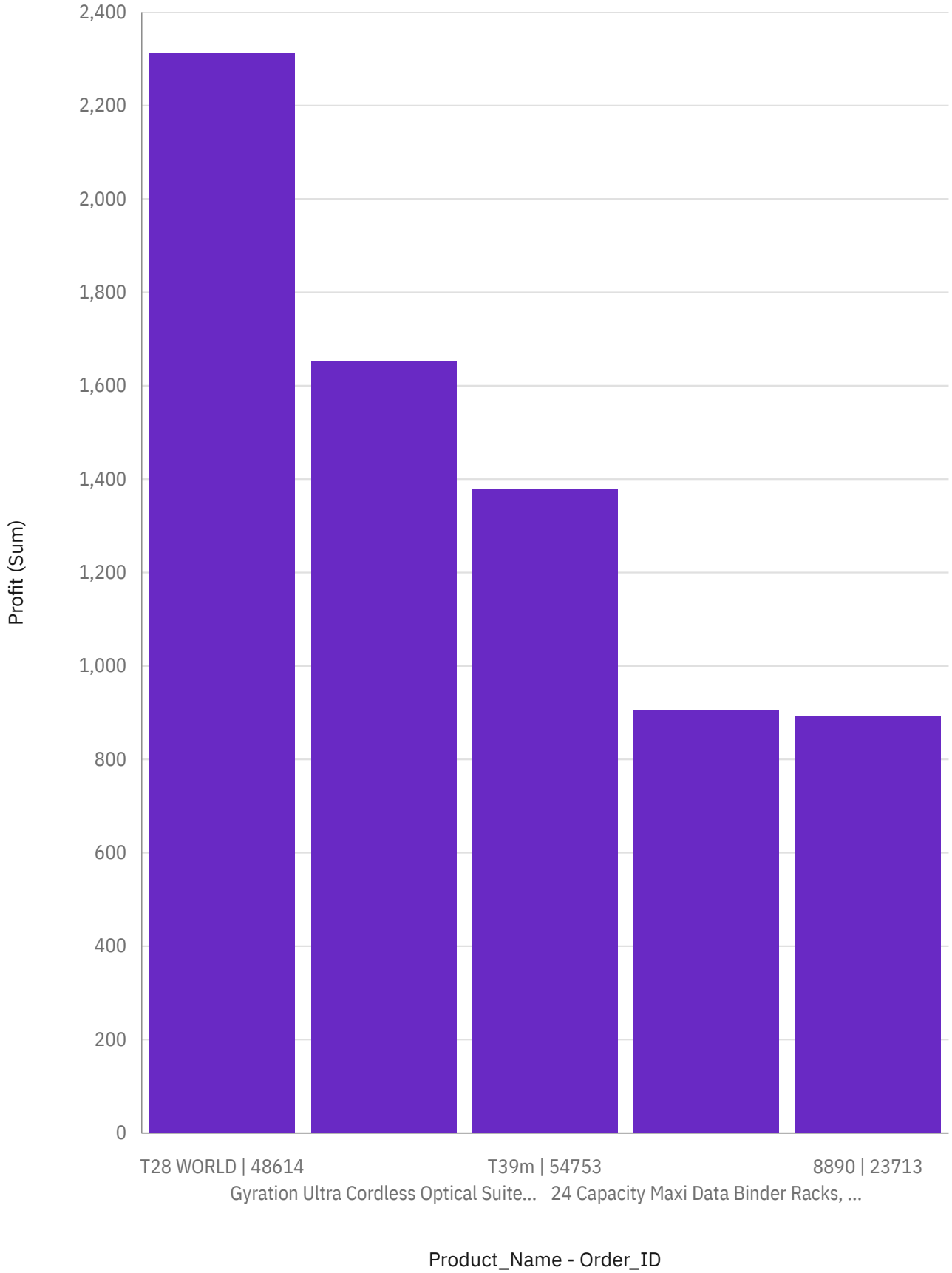

Profit by Product_Name and Order_ID
2,000
1,800
1,600
1,400
1,200
1,000
800
600
400
200
0
Profit (Sum)
SAFCO Arco Folding Chair |...
R380 | 483
#10 White Business Envelopes,4 1/8 ...
LX 788 | 1344
3M Office Air Cle...
Product_Name - Order_ID

Ontario-P


Profit by Product_Name and Order_ID
Profit (Sum)
5,500
5,000
4,500
4,000
3,500
3,000
2,500
2,000
1,500
1,000
500
0
Polycom VoiceStation 100 | 41059
Kensington 7 Outlet MasterPiece Power Center | 33250
Hoover WindTunnel™ Plus Canister Vacuum | 31238
GBC Therma-A-Bind 250T Electric Binding System | 56101
Talkabout T8097 | 3109
Product_Name - Order_ID

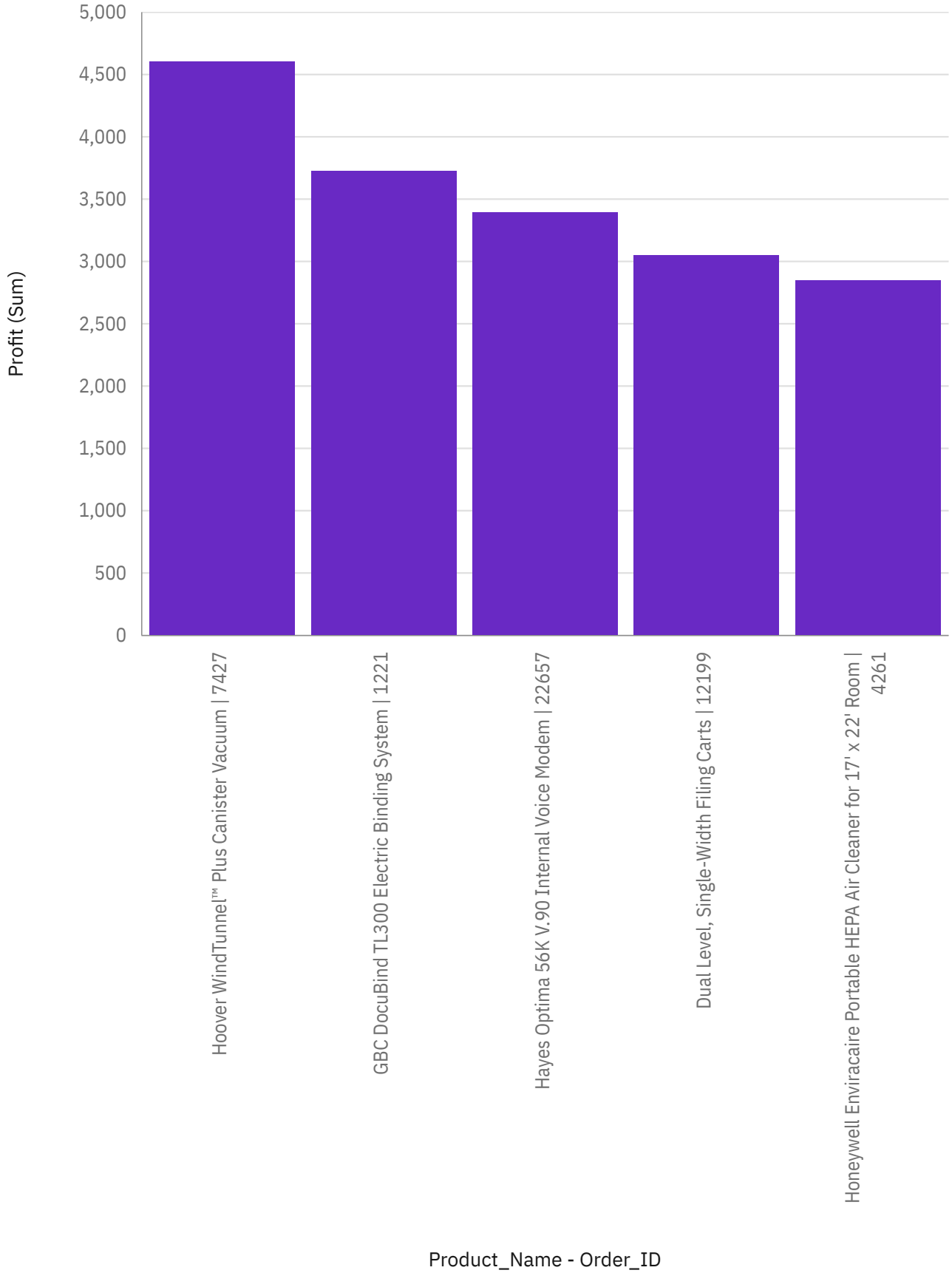
Prarie-P

Profit by Product_Name and Order_ID



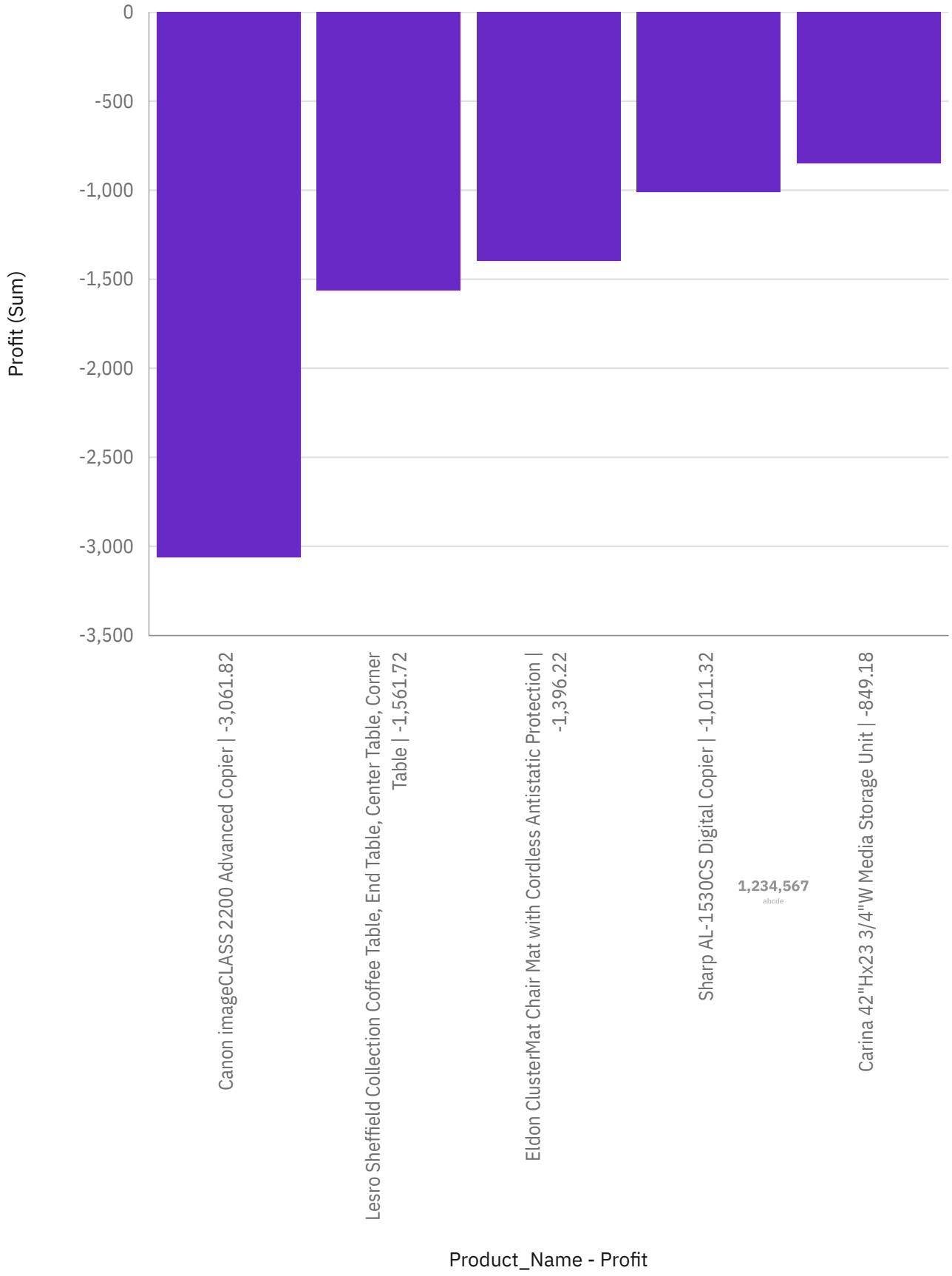
West-P

Profit by Product_Name and Order_ID



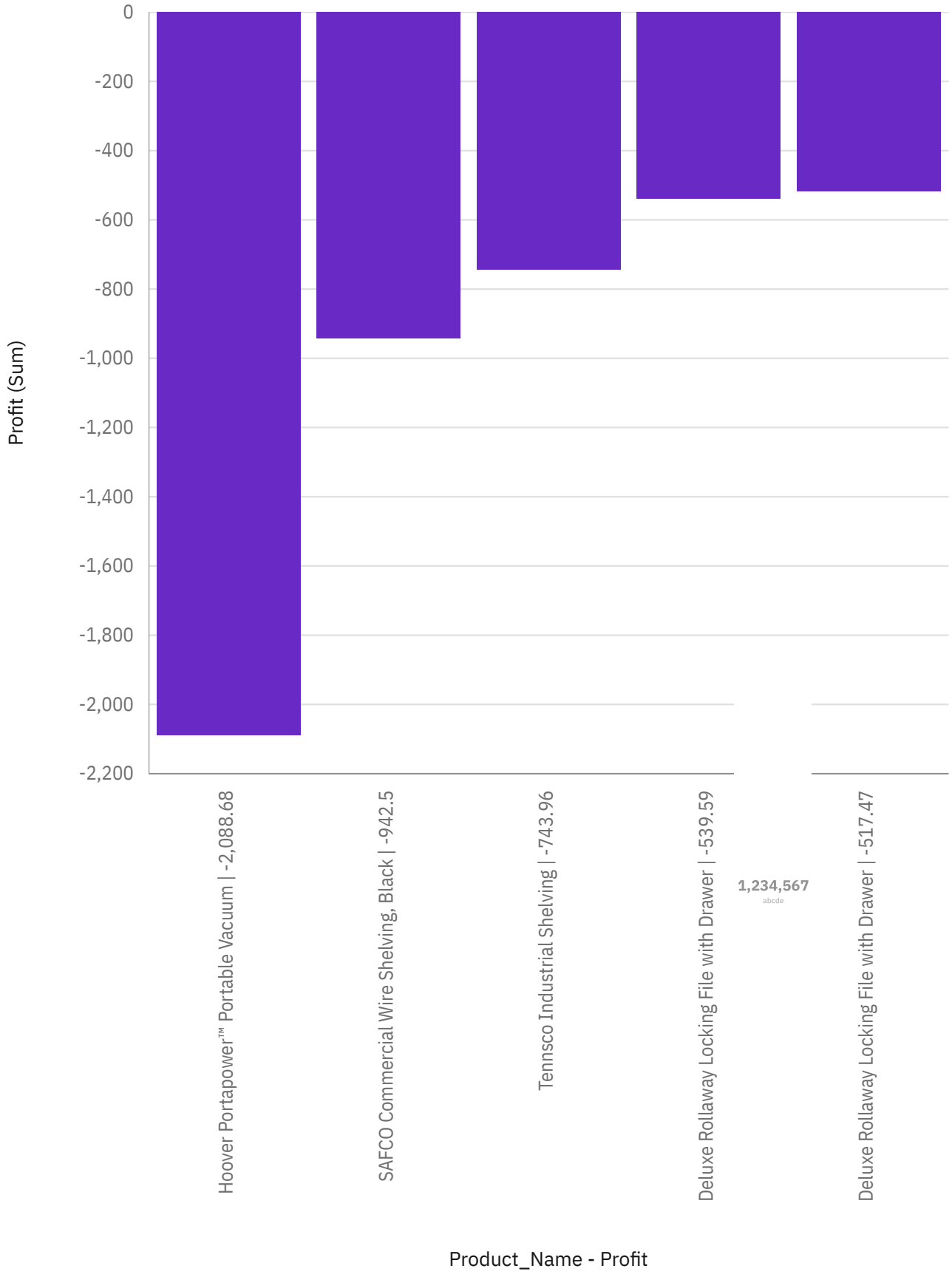
Atlantic-N

Profit by Product_Name and Profit

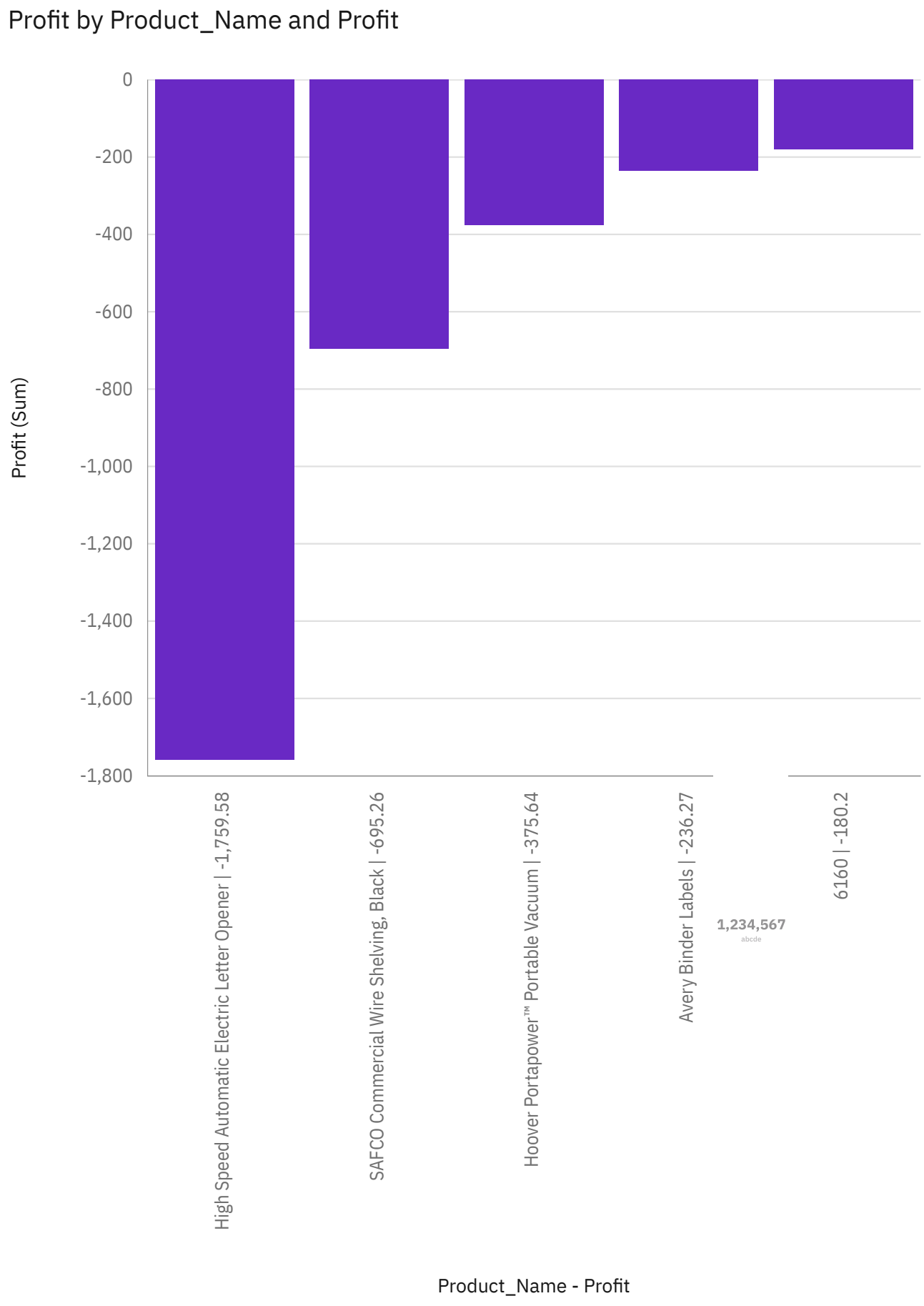


Northwest Territories-N

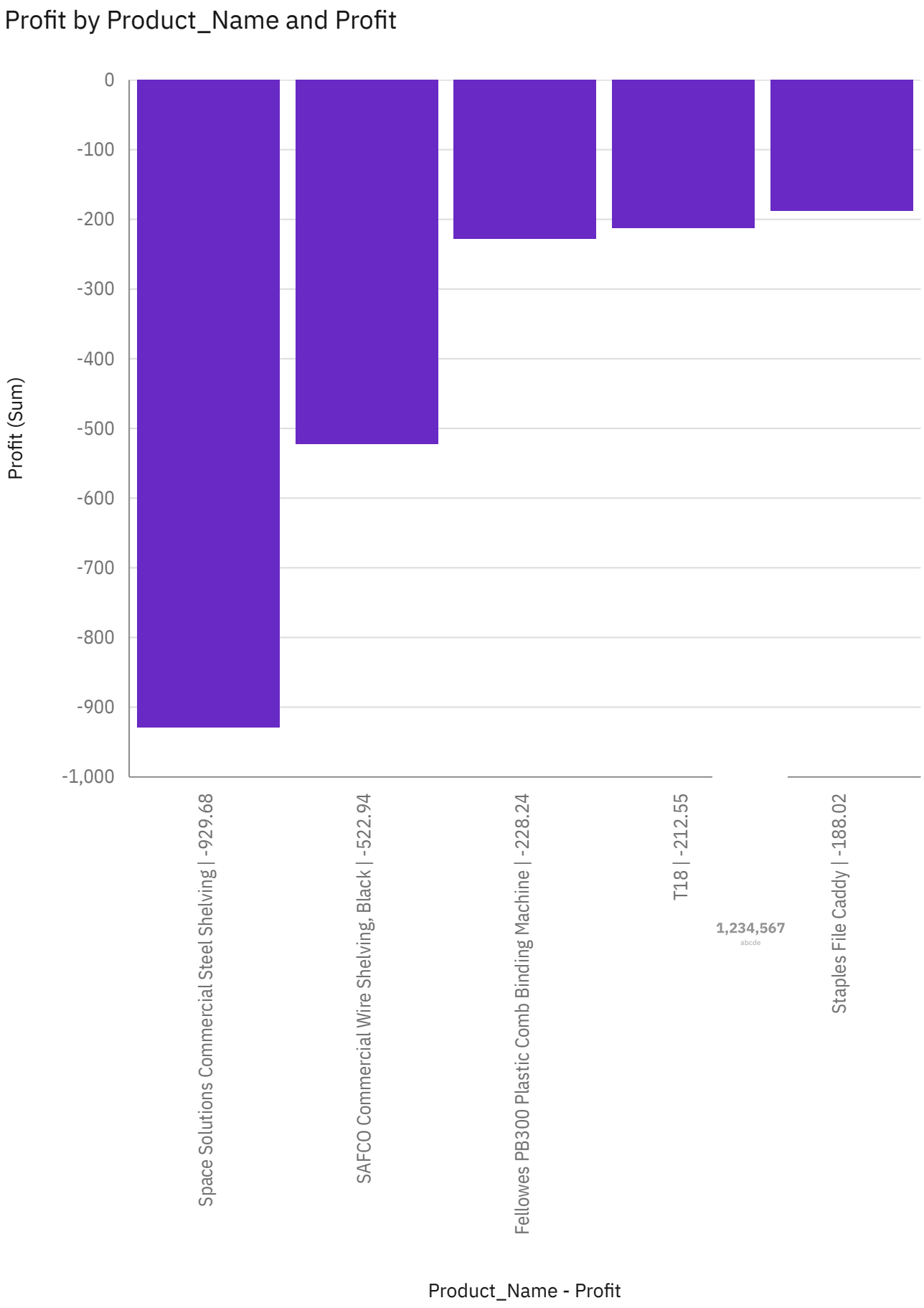
Profit by Product_Name and Profit



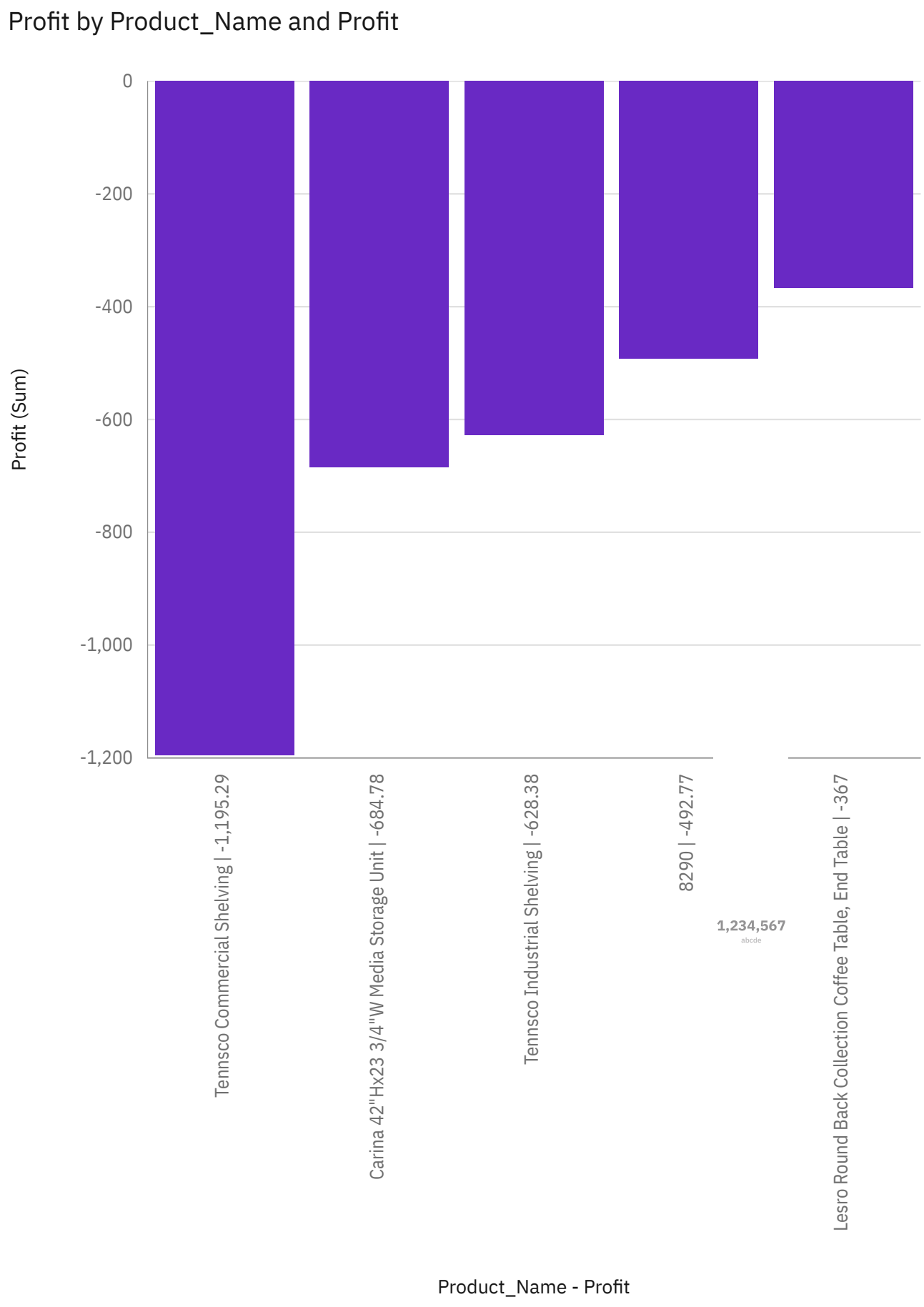
Nunavut-N


Profit by Product_Name and Profit
Profit (Sum)
0
-200
-400
-600
-800
-1,000
-1,200
-1,400
-1,600
-1,800
High Speed Automatic Electric Letter Opener | -1,759.58
SAFCO Commercial Wire Shelving, Black | -695.26
Hoover Portapower™ Portable Vacuum | -375.64
Avery Binder Labels | -236.27
1,234,567
abcde
6160 | -180.2
Product_Name - Profit

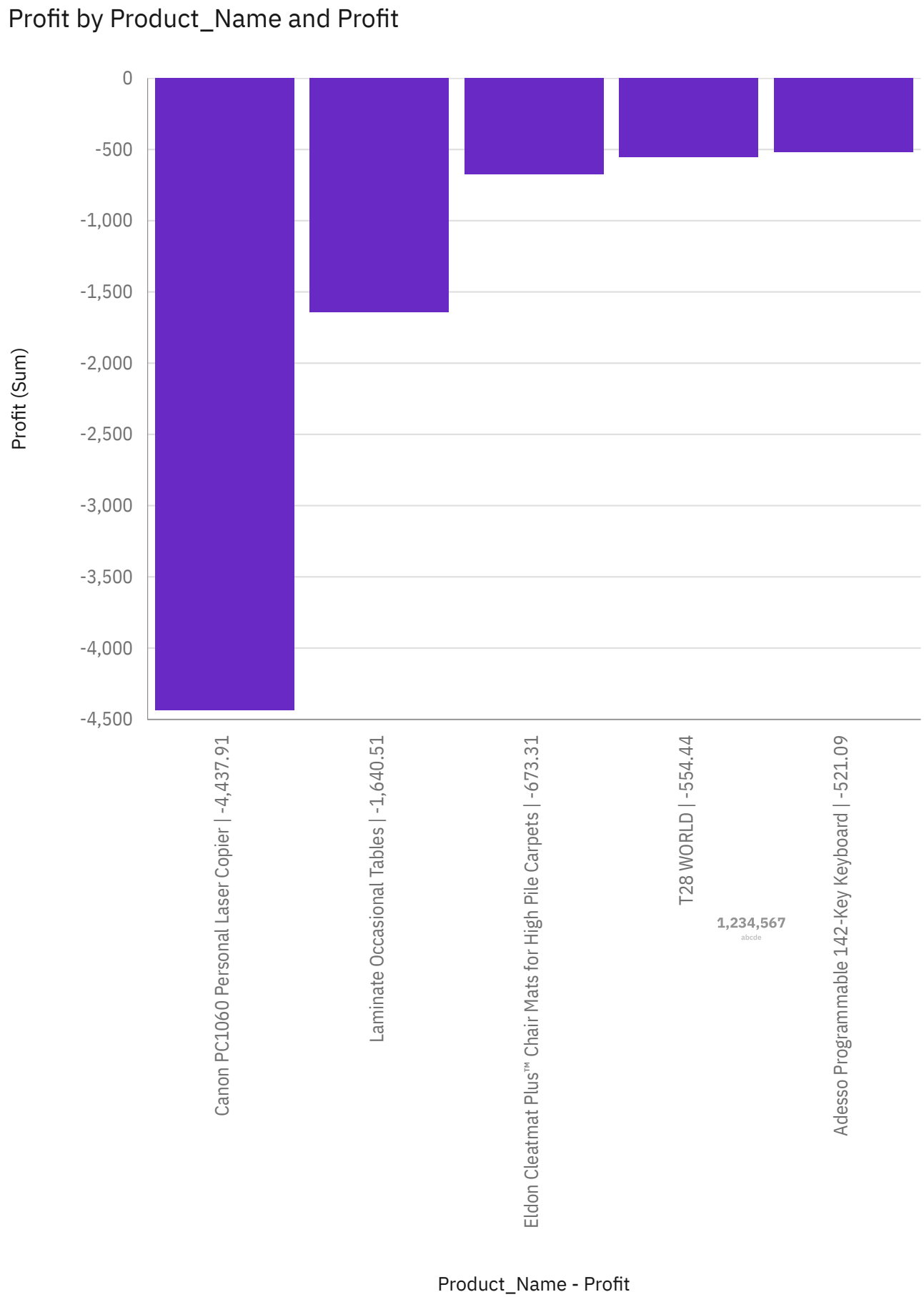
Ontario-N


Profit by Product_Name and Profit
0
-100
-200
-300
-400
-500
-600
-700
-800
-900
-1,000
Profit (Sum)
Space Solutions Commercial Steel Shelving | -929.68
SAFCO Commercial Wire Shelving, Black | -522.94
Fellowes PB300 Plastic Comb Binding Machine | -228.24
T18 | -212.55
1,234,567
abcde
Staples File Caddy | -188.02
Product_Name - Profit

Prarie-N


Profit by Product_Name and Profit
0
-200
-400
-600
-800
-1,000
-1,200
Profit (Sum)
Tennsco Commercial Shelving | -1,195.29
Carina 42"Hx23 3/4"W Media Storage Unit | -684.78
Tennsco Industrial Shelving | -628.38
8290 | -492.77
1,234,567
abcde
Lesro Round Back Collection Coffee Table, End Table | -367
Product_Name - Profit

West-N


Profit by Product_Name and Profit
0
-500
-1,000
-1,500
-2,000
-2,500
-3,000
-3,500
-4,000
-4,500
Profit (Sum)
Canon PC1060 Personal Laser Copier | -4,437.91
Laminate Occasional Tables | -1,640.51
Eldon Cleatmat Plus™ Chair Mats for High Pile Carpets | -673.31
T28 WORLD | -554.44
1,234,567
abcde
Adesso Programmable 142-Key Keyboard | -521.09
Product_Name - Profit

CONCUSION:

Thus using the **IBM COGNOS ANALYTICS** , the visualization of product sales by the given dataset was done region wise. By using these visualization we can clearly understand that which **top 5** products produces more profit and which **bottom 5** products produces less profit in each region.