GROUP 3 PROJECT PROPOSAL

BRANDON, PARTH, STEVE, YANNICK, ATEEQ

Table of Contents

PURPOSE AND SCOPE

We were tasked with a project to analyze the revenue of the Global Superstore during a given period (2011-2014). The business needs to ensure that it is profitable as it is important for investors and managers to ensure that the business can continue to finance its operations and grow. The analysis should determine whether the company's current business approach is effective and efficient, if not recommend ways in which the company can improve its performance in the future.

Global Superstore is an imaginary international e-commerce giant that sells and ships various products worldwide. The purpose of this project is to find out which market is the most profitable based on the top five profitable products and which market is the least profitable. Armed with this knowledge, Global Superstore stakeholders can decide where and how to focus their activities in future.

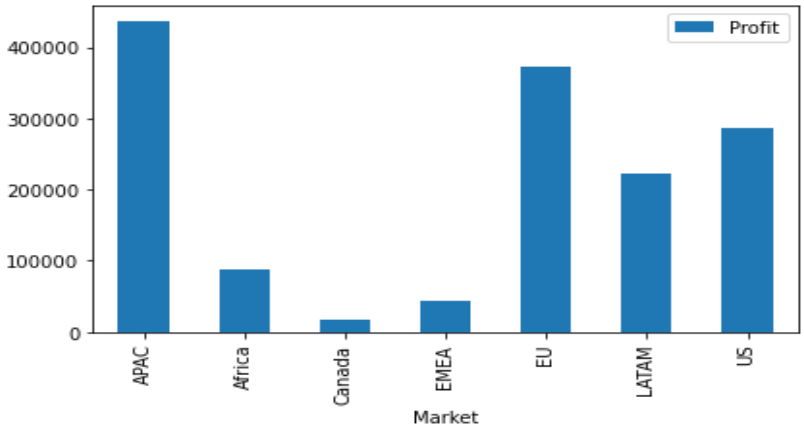
This analysis will be done by collecting data from the CSV Global-Superstore. file, which is a dataset of previous sales, and filter it, using python, to organize the top five and bottom five by market.

The markets are divided as follows:

- **APAC - Asia-Pacific**
- **US - United States**
- **EU - European Union**
- **CANADA - Canada**
- **EMEA - Europe, Middle East and Africa**
- **AFRICA - Africa**
- **LATAM - Latin America**

After analyzing the dataset, we found that the Asian market is where the store makes the most profit and the list in descending order is as follows:

Market	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
APAC	283802091	3.585744e+06	41226	1637.530	436000.0	387165.11
EU	152945000	2.938089e+06	37773	1031.050	372830.0	309422.35
US	362717239	2.297201e+06	37873	1561.090	286397.0	238173.79
LATAM	52988365	2.164605e+06	38526	1395.158	221643.0	234138.61
Africa	212025742	7.837732e+05	10564	718.800	88872.0	88139.47
EMEA	233028207	8.061613e+05	11517	986.100	43898.0	88375.73
Canada	17851051	6.692817e+04	833	0.000	17817.0	7405.63



After determining the most profitable markets, we decided to dive into the details to find out which products were driving the store's profits in those regions, then we discovered the top 5 most profitable products in each region (as seen in Appendix).

Despite this performance, we decided to investigate and find a way for the business to be more efficient by discovering the underperformers that were draining the company's cash flow, so we discovered the 5 least profitable products on each market (as seen in Appendix).

In addition, we made an interesting discovery in the Canadian market which is the least profitable market where the store did not lose money because there are no negative profit products despite the insignificant profit which is produced there.

EXTERNAL ENVIRONMENT

This section is devoted to the results of our research in the process of understanding whether external economic factors may have affected the business activities of global supermarkets and dictated their sales results.

The growth of the internet and digital technologies has increased global business exponentially. Multinational corporations, small businesses, and sole proprietorships are involved in global commerce. This is driving the demand of photocopiers, VOIP phones, and smartphones; the top selling products we see in our analysis. Below are some excerpts of global trend articles taken from around the era of the data used for analysis (2011-2014).

- "The enormous potential impact of digitization is only beginning to emerge. Consider that international Skype-call minutes grew to 40 percent of the present level of traditional international calls in just a decade. Or that cross-border e-commerce has grown to represent more than 10 percent of trade in goods in less than a decade".

- "The copier market is driven by emerging photocopying and fast technologies, government organizations, an increase in the number of educational institutes and offices, product innovation in the field of photocopying, availability of photocopier models and cheaper products, and increased demand for multifunctional devices".

- "The value of mobile phone calls made using the internet is expected to rise to $30 billion over the next three years, according to Deloitte".

DATA PROCESS AND CLEANING

After investigating the dataset, we decided to remove the postal code column because it is missing 80.51% of necessary data. Overall, the rest of the data was clean.

CONCLUSION

After determining the five most profitable products in each market, we discovered that some products can be profitable in one market but not in another market such as the full-size Apple smartphone, which is the most profitable product in Africa, but was the least profitable product in the Asian market. Global Superstore must therefore be careful not to assume that a product performance can be replicated in all markets.

Moreover, we discovered after analysis that there is a negative linear correlation between the discount strategy and product sales as a higher discount did not necessarily translate to increased sales (as seen in Appendix – Scatter 1.1, 1.2): because the more discounts the store offered, the fewer sales it made. Therefore, we recommend the store to change its discount marketing strategy to improve its earnings performance.
However, we have found that there is a positive correlation between number of sales and profit to some degree with most sales below 5000 and profit below 2000. There are outliers with negative profits even when there is a sale, which is understandable because there are many variables that dictate profit, such as product cost, shipping, or discount.

This investigation was eye-opening because even though the store is overall profitable but was losing money selling certain products in many markets except Canada, there is room for improvement by implementing the strategies we recommended, and which can be listed as follows:

1. Stop selling products with negative profits
2. Reallocated capital to increase the number of units sold for profitable products
3. Change the discount marketing strategy to avoid offering more discounts that don't necessarily translate to more sales

REFERENCES

- Manyika, James. “Global Flows in a Digital Age.” *McKinsey.com, 1* Apr. 2014, https://www.mckinsey.com/capabilities/strategy-and-corporate-finance/our-insights/global-flows-in-a-digital-age

- “Copier industry market scope, growth trends and forecast.” *medium.com, 8 Sep. 2020,* https://medium.com/@1800copier/copier-industry-trendscopier-industry-market-scope-growth-trends-and-forecast-33a1dac37161

- Ryan, Vivienne. “Mobile VOIP value to rise to $30b: Deloitte.” *theist.com.au*, 27 Jan. 2010, https://thewest.com.au/business/finance/mobile-voip-value-to-rise-to-30b-deloitte-ng-ya-223531

APPENDIX

This section contains tables and graphs of the most important results of the analysis of the dataset to provide visualization and better understanding of our results.

Table 1.1 - Top 5 most profitable products in the US

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Canon imageCLASS 2200 Advanced Copier	184702	61599.824	20	0.6	25199.9280	1276.32
Fellowes PB500 Electric Punch Plastic Comb Binding Machine with Manual Bind	374819	27453.384	31	2.4	7753.0390	3197.94
Hewlett Packard LaserJet 3310 Copier	291812	18839.686	38	1.6	6983.8836	2764.80
Canon PC1060 Personal Laser Copier	150363	11619.834	19	0.6	4570.9347	678.96
HP Designjet T520 Inkjet Large Format Printer - 24" Color	113133	18374.895	12	0.5	4094.9766	961.07

Table 1.2 - Least 5 profitable products in the US

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Cubify CubeX 3D Printer Double Head Print	109032	11099.963	9	1.60	-8879.9704	681.59
Lexmark MX611dhe Monochrome Laser Printer	129603	16829.901	18	1.60	-4589.9730	1314.41
Cubify CubeX 3D Printer Triple Head Print	31980	7999.980	4	0.50	-3839.9904	674.82
Chromcraft Bull-Nose Wood Oval Conference Tables & Bases	184541	9917.640	27	1.40	-2876.1156	1133.62
Bush Advantage Collection Racetrack Conference Table	253622	9544.725	33	2.45	-1934.3976	736.08

Table 2.1 - Top 5 most profitable products in the Africa

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Apple Smart Phone, Full Size	416986	16946.406	28	1.4	4103.406	1664.97
Canon Wireless Fax, Laser	138360	6058.560	16	0.0	2968.320	746.82
Sauder Classic Bookcase, Traditional	291181	6539.850	15	0.0	2681.100	691.31
Cisco Smart Phone, Cordless	225557	8404.092	15	2.1	2540.592	801.54
Ikea Classic Bookcase, Traditional	193359	4117.500	10	0.0	1564.500	262.71

Table 2.2 - Least 5 profitable products in the Africa

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Breville Microwave, Red	90306	1025.814	9	0.7	-1445.226	73.33
Rogers Lockers, Blue	190339	1058.100	12	1.4	-1303.860	69.06
Bevis Training Table, Fully Assembled	42222	605.610	6	0.7	-1150.830	105.10
Harbour Creations Rocking Chair, Set of Two	138032	970.710	17	1.4	-1111.110	125.37
Nokia Smart Phone, Full Size	330433	4588.920	10	2.1	-1083.780	945.89

Table 3.1 - Top 5 most profitable products in the Canada

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Hewlett Wireless Fax, High-Speed	101411	3783.90	10	0.0	1854.00	479.09
Motorola Smart Phone, Full Size	46366	3863.88	6	0.0	1159.02	546.56
SanDisk Router, USB	50299	1539.36	6	0.0	646.38	93.09
Breville Microwave, White	45096	1799.10	6	0.0	593.64	274.75
Rogers Lockers, Industrial	130077	1484.70	7	0.0	564.06	148.47

Table 3.2 - Least 5 profitable products in the Canada

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Avery 3-Hole Punch, Recycled	48214	28.23	1	0.0	0.0	3.39
Belkin Memory Card, USB	48412	460.08	4	0.0	0.0	35.03
Cuisinart Microwave, Silver	49953	558.42	2	0.0	0.0	60.80
Enermax Memo Slips, 8.5 x 11	46238	19.80	1	0.0	0.0	5.81
Kleencut Shears, Easy Grip	50058	48.39	1	0.0	0.0	5.48

Table 4.1 - Top 5 most profitable products in the Asian Pacific Market

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Motorola Smart Phone, Full Size	340608	34261.5270	59	1.61	7616.9670	3441.64
Office Star Executive Leather Armchair, Black	248381	15289.6353	37	1.02	6123.2553	1335.59
Samsung Smart Phone, VoIP	148533	16796.7360	31	0.60	5356.8060	1046.17
Nokia Smart Phone, with Caller ID	302086	30808.2402	54	1.24	5277.0402	3988.66
Novimex Executive Leather Armchair, Adjustable	316144	18832.2576	46	1.19	5047.9176	2997.20

Table 4.2 - Least 5 profitable products in the Asian Pacific Market

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Apple Smart Phone, Full Size	313608	27251.1561	52	1.69	-2739.1239	2297.23
Hon Conference Table, Rectangular	109480	4769.8695	14	2.17	-2564.1705	267.26
Bevis Wood Table, with Bottom Storage	129253	7043.6520	19	1.60	-2284.3980	585.74
Breville Microwave, Silver	98589	7434.4668	34	0.94	-1924.0332	435.06
Chromcraft Round Table, Fully Assembled	49160	3326.8470	13	1.10	-1851.1830	207.17

Table 5.1 - Top 5 most profitable products in the EMEA Market

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Cisco Smart Phone, Full Size	279266	12376.410	19	0.0	2722.320	1305.40
Barricks Training Table, Adjustable Height	100245	7050.780	22	0.0	2255.880	281.75
Motorola Smart Phone, Full Size	477587	9715.455	18	2.3	2019.435	1334.24
Bush Classic Bookcase, Pine	94122	3726.270	9	0.0	1825.740	235.84
Canon Wireless Fax, Digital	180774	3783.000	10	0.0	1664.400	429.95

Table 5.2 – Least 5 profitable products in the EMEA Market

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Motorola Smart Phone, Cordless	190037	7584.804	19	0.6	-3773.586	312.69
Barricks Conference Table, Rectangular	96533	3075.810	9	0.7	-2460.810	436.01
Office Star Executive Leather Armchair, Adjustable	44795	1503.744	8	0.6	-1842.096	126.17
Hoover Stove, White	91707	1926.474	7	0.6	-1841.556	133.11
Bevis Round Table, Adjustable Height	45920	1248.912	6	0.6	-1779.768	18.28

Table 6.1 - Top 5 most profitable products in the EU Market

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Hoover Stove, Red	93578	21147.0840	38	0.3	10345.5840	2640.76
Nokia Smart Phone, Full Size	267771	33511.3275	61	2.0	6890.7675	3191.68
Hamilton Beach Stove, Silver	115604	18247.8240	38	1.2	5452.4640	1132.12
SAFCO Executive Leather Armchair, Black	84919	12301.2000	30	0.6	5003.1000	1029.36
Safco Classic Bookcase, Metal	90074	10371.3570	25	0.3	4681.8570	1742.30

Table 6.2 – Least 5 profitable products in the EU Market

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Hoover Stove, White	78354	11728.827	31	1.20	-4958.163	1377.60
Office Star Executive Leather Armchair, Black	74425	6497.277	22	1.40	-3066.783	818.07
Fellowes File Cart, Single Width	215157	5809.482	60	3.80	-2430.918	557.44
Barricks Conference Table, Adjustable Height	34460	3728.007	9	1.05	-2082.393	463.58
Bevis Conference Table, Fully Assembled	11684	555.138	4	0.85	-1924.542	62.75

Table 7.1 – Top 5 profitable products in The Latin America Market

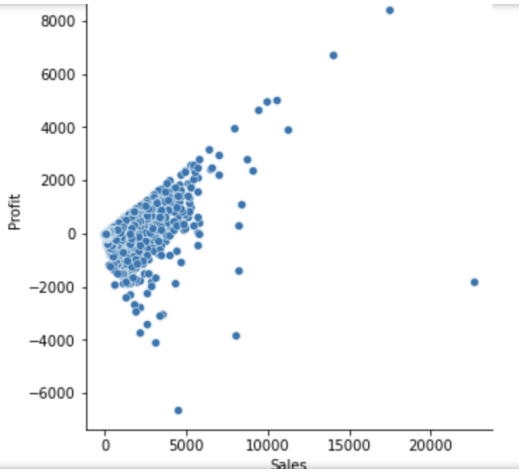
Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Cisco Smart Phone, Full Size	44018	22088.616	52	0.6	5914.696	2127.49
Nokia Smart Phone, Full Size	65149	19873.280	51	1.0	4146.100	2367.21
Harbour Creations Executive Leather Armchair, Red	31392	8184.792	28	0.4	3387.832	527.45
Dania Classic Bookcase, Traditional	24190	7612.752	31	1.2	3283.292	777.28
Nokia Smart Phone, with Caller ID	22118	10392.936	28	0.8	3236.696	776.41

Table 7.2 – Least 5 profitable products in The Latin America Market

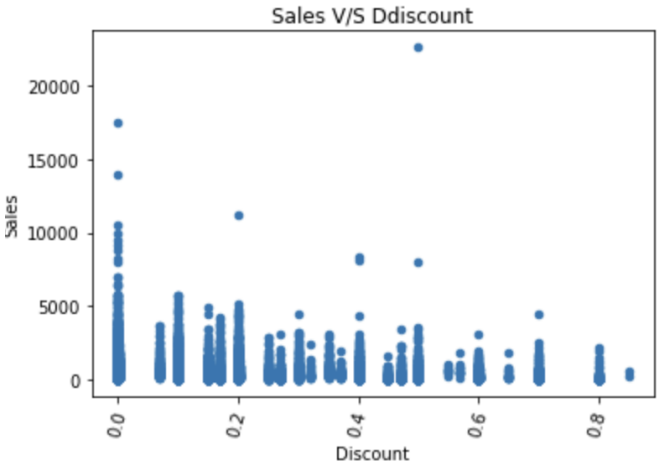
Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Samsung Smart Phone, Cordless	68573	17636.400	54	2.2	-3757.320	2933.95
Lesro Computer Table, Fully Assembled	14351	1199.352	14	2.1	-2798.488	115.22
Chromcraft Conference Table, with Bottom Storage	12842	1047.204	6	1.4	-1780.356	181.30
Ikea Library with Doors, Pine	41053	4620.040	28	3.2	-1780.200	598.47
Bush Library with Doors, Pine	24516	2632.824	18	1.2	-1667.736	253.75

UNCOVERING THE CORRELATION

SCATTER 1.1 - SALES AND PROFIT



SCATTER 1.2 - DISCOUNT AND SALES



GLOBAL SUPERSTORE PROJECT BY ATEEQ, PARTH, BRANDON, YANNICK AND STEVE

In [2]:
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [3]:
```
df=pd.read_csv("Global-Superstore.csv",encoding='latin1')
```

Business Question : What are the top five most profitable products in each market and how can the Global Superstore improve its profitability in each market based on past performance?

In [4]:
```
df.info ()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 51290 entries, 0 to 51289
Data columns (total 24 columns):
 #   Column          Non-Null Count  Dtype  
---  ------          --------------  -----  
 0   Row ID          51290 non-null  int64  
 1   Order ID        51290 non-null  object 
 2   Order Date      51290 non-null  object 
 3   Ship Date       51290 non-null  object 
 4   Ship Mode       51290 non-null  object 
 5   Customer ID     51290 non-null  object 
 6   Customer Name   51290 non-null  object 
 7   Segment         51290 non-null  object 
 8   City            51290 non-null  object 
 9   State           51290 non-null  object 
 10  Country         51290 non-null  object 
 11  Postal Code     9994 non-null   float64
 12  Market          51290 non-null  object 
 13  Region          51290 non-null  object 
 14  Product ID      51290 non-null  object 
 15  Category        51290 non-null  object 
 16  Sub-Category    51290 non-null  object 
 17  Product Name    51290 non-null  object 
 18  Sales           51290 non-null  float64
 19  Quantity        51290 non-null  int64  
 20  Discount        51290 non-null  float64
 21  Profit          51290 non-null  float64
 22  Shipping Cost   51290 non-null  float64
 23  Order Priority  51290 non-null  object 
dtypes: float64(5), int64(2), object(17)
memory usage: 9.4+ MB
```

AFTER INVESTIGATING THE DATA SET, WE DECIDED TO REMOVE THE POSTAL CODE COLUMN BECAUSE THERE ARE MANY

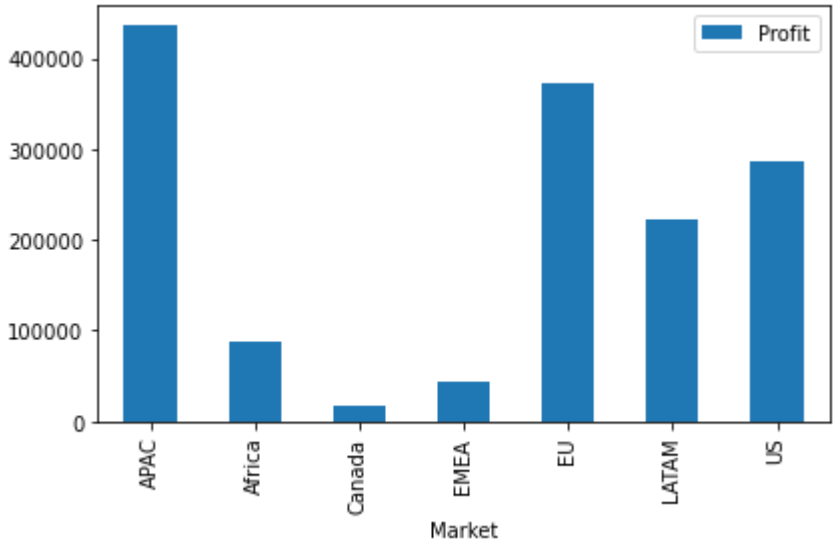
MISSING DATA

In [5]:
```
df=df.drop(['Postal Code'], axis=1)
```

CHART OF THE MARKETS BY PROFIT

In [6]:
```
df.pivot_table(values = 'Profit', index = 'Market', aggfunc='sum').plot(kind='bar')
plt.tight_layout()
plt.show()

# The graph shows that Asia-Pacific is the most profitable market and the Canadian mar
```



The graph shows that Asia-Pacific is the most profitable market and the Canadian market is the one where the store makes the least profit

UNITED STATES MARKET

The Top 5 most profitable products in the US

In [7]:
```
df_US= df[df['Market'].str.contains('US')]
df_top5grp1_US=df_US.groupby('Product Name').sum()
df_top5grp_US=df_top5grp1_US.nlargest(5,'Profit')
df_top5grp_US.head()
```

```
/var/folders/k7/31gr9yq52q53zqt_23tjp2q00000gn/T/ipykernel_7861/70210780.py:2: Future
Warning: The default value of numeric_only in DataFrameGroupBy.sum is deprecated. In
a future version, numeric_only will default to False. Either specify numeric_only or
select only columns which should be valid for the function.
  df_top5grp1_US=df_US.groupby('Product Name').sum()
```

Out[7]:

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Canon imageCLASS 2200 Advanced Copier	184702	61599.824	20	0.6	25199.9280	1276.32
Fellowes PB500 Electric Punch Plastic Comb Binding Machine with Manual Bind	374819	27453.384	31	2.4	7753.0390	3197.94
Hewlett Packard LaserJet 3310 Copier	291812	18839.686	38	1.6	6983.8836	2764.80
Canon PC1060 Personal Laser Copier	150363	11619.834	19	0.6	4570.9347	678.96
HP Designjet T520 Inkjet Large Format Printer - 24" Color	113133	18374.895	12	0.5	4094.9766	961.07

This table informs us that the Canon Imgae 2200 Advanced copier is the most profitable product in the United States, so the company should focus selling more of this product.

The 5 least profitable products in the US

In [8]:
```
df_US= df[df['Market'].str.contains('US')]
df_top5grp1_US=df_US.groupby('Product Name').sum()
df_top5grp_US=df_top5grp1_US.nsmallest(5,'Profit')
df_top5grp_US.head()
```

```
/var/folders/k7/31gr9yq52q53zqt_23tjp2q00000gn/T/ipykernel_7861/4250059105.py:2: Futu
reWarning: The default value of numeric_only in DataFrameGroupBy.sum is deprecated. I
n a future version, numeric_only will default to False. Either specify numeric_only o
r select only columns which should be valid for the function.
  df_top5grp1_US=df_US.groupby('Product Name').sum()
```

Out[8]:

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Cubify CubeX 3D Printer Double Head Print	109032	11099.963	9	1.60	-8879.9704	681.59
Lexmark MX611dhe Monochrome Laser Printer	129603	16829.901	18	1.60	-4589.9730	1314.41
Cubify CubeX 3D Printer Triple Head Print	31980	7999.980	4	0.50	-3839.9904	674.82
Chromcraft Bull-Nose Wood Oval Conference Tables & Bases	184541	9917.640	27	1.40	-2876.1156	1133.62
Bush Advantage Collection Racetrack Conference Table	253622	9544.725	33	2.45	-1934.3976	736.08

Clearly the store is losing a lot of money in the US with the Cubify CubeX #D printer and should stop selling it and reallocate capital to selling more of the Canon Imgae 2200 Advanced copier.

AFRICAN MARKET

The Top 5 most profitable products in Africa

In [9]:
```
df_Africa= df[df['Market'].str.contains('Africa')]
df_top5grp1_Africa=df_Africa.groupby('Product Name').sum()
df_top5grp_Africa=df_top5grp1_Africa.nlargest(5,'Profit')
df_top5grp_Africa.head()
```

```
/var/folders/k7/31gr9yq52q53zqt_23tjp2q00000gn/T/ipykernel_7861/2366592338.py:2: FutureWarning: The default value of numeric_only in DataFrameGroupBy.sum is deprecated. In a future version, numeric_only will default to False. Either specify numeric_only or select only columns which should be valid for the function.
  df_top5grp1_Africa=df_Africa.groupby('Product Name').sum()
```

Out[9]:

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Apple Smart Phone, Full Size	416986	16946.406	28	1.4	4103.406	1664.97
Canon Wireless Fax, Laser	138360	6058.560	16	0.0	2968.320	746.82
Sauder Classic Bookcase, Traditional	291181	6539.850	15	0.0	2681.100	691.31
Cisco Smart Phone, Cordless	225557	8404.092	15	2.1	2540.592	801.54
Ikea Classic Bookcase, Traditional	193359	4117.500	10	0.0	1564.500	262.71

The table shows that the sale of Apple Smart Phones (Full Size) created more profits in the African market. The store should sell more of this product, however they should not neglect the other four profitable products.

The 5 least profitable products in Africa

In [10]:
```
df_AF= df[df['Market'].str.contains('Africa')]
df_top5grp1_AF=df_AF.groupby('Product Name').sum()
df_top5grp_AF=df_top5grp1_AF.nsmallest(5,'Profit')
df_top5grp_AF.head()
```

```
/var/folders/k7/31gr9yq52q53zqt_23tjp2q00000gn/T/ipykernel_7861/948495512.py:2: FutureWarning: The default value of numeric_only in DataFrameGroupBy.sum is deprecated. In a future version, numeric_only will default to False. Either specify numeric_only or select only columns which should be valid for the function.
  df_top5grp1_AF=df_AF.groupby('Product Name').sum()
```

Out[10]:

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Breville Microwave, Red	90306	1025.814	9	0.7	-1445.226	73.33
Rogers Lockers, Blue	190339	1058.100	12	1.4	-1303.860	69.06
Bevis Training Table, Fully Assembled	42222	605.610	6	0.7	-1150.830	105.10
Harbour Creations Rocking Chair, Set of Two	138032	970.710	17	1.4	-1111.110	125.37
Nokia Smart Phone, Full Size	330433	4588.920	10	2.1	-1083.780	945.89

The table above shows the top five unprofitable products in the African market, with the Red Breville microwave being the product the store lost the most money on. The store should stop selling these products and reallocate capital to selling more of the 5 products that brings in the most profit.

CANADA MARKET

The Top 5 most profitable products in Canada

In [11]:
```
df_CA= df[df['Market'].str.contains('Canada')]
df_top5grp1_CA=df_CA.groupby('Product Name').sum()
df_top5grp_CA=df_top5grp1_CA.nlargest(5,'Profit')
df_top5grp_CA.head()
```

```
/var/folders/k7/31gr9yq52q53zqt_23tjp2q00000gn/T/ipykernel_7861/692202079.py:2: FutureWarning: The default value of numeric_only in DataFrameGroupBy.sum is deprecated. In a future version, numeric_only will default to False. Either specify numeric_only or select only columns which should be valid for the function.
  df_top5grp1_CA=df_CA.groupby('Product Name').sum()
```

Out[11]:

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Hewlett Wireless Fax, High-Speed	101411	3783.90	10	0.0	1854.00	479.09
Motorola Smart Phone, Full Size	46366	3863.88	6	0.0	1159.02	546.56
SanDisk Router, USB	50299	1539.36	6	0.0	646.38	93.09
Breville Microwave, White	45096	1799.10	6	0.0	593.64	274.75
Rogers Lockers, Industrial	130077	1484.70	7	0.0	564.06	148.47

The table shows that in the Canadian market the sale of the Hewelett high-speed wireless fax machine generated the most

profit, however the other four products should also be considered.

The 5 least profitable products in Canada

In [12]:
```
df_CA= df[df['Market'].str.contains('Canada')]
df_top5grp1_CA=df_CA.groupby('Product Name').sum()
df_top5grp_CA=df_top5grp1_CA.nsmallest(5,'Profit')
df_top5grp_CA.head()
```

```
/var/folders/k7/31gr9yq52q53zqt_23tjp2q00000gn/T/ipykernel_7861/2603876016.py:2: Futu
reWarning: The default value of numeric_only in DataFrameGroupBy.sum is deprecated. I
n a future version, numeric_only will default to False. Either specify numeric_only o
r select only columns which should be valid for the function.
  df_top5grp1_CA=df_CA.groupby('Product Name').sum()
```

Out[12]:

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Avery 3-Hole Punch, Recycled	48214	28.23	1	0.0	0.0	3.39
Belkin Memory Card, USB	48412	460.08	4	0.0	0.0	35.03
Cuisinart Microwave, Silver	49953	558.42	2	0.0	0.0	60.80
Enermax Memo Slips, 8.5 x 11	46238	19.80	1	0.0	0.0	5.81
Kleencut Shears, Easy Grip	50058	48.39	1	0.0	0.0	5.48

This graph not only shows the five least profitable products in Canada, but also shows that the store did not lose money selling in this market. The store should probably find a new distributor offering these products at a more attractive price or stop selling them

ASIAN PACIFIC MARKET

The Top 5 most profitable products in The Asian Pacific Market

In [13]:
```
df_APAC= df[df['Market'].str.contains('APAC')]
df_top5grp1_APAC=df_APAC.groupby('Product Name').sum()
df_top5grp_APAC=df_top5grp1_APAC.nlargest(5,'Profit')
df_top5grp_APAC.head()
```

```
/var/folders/k7/31gr9yq52q53zqt_23tjp2q00000gn/T/ipykernel_7861/3869578424.py:2: Futu
reWarning: The default value of numeric_only in DataFrameGroupBy.sum is deprecated. I
n a future version, numeric_only will default to False. Either specify numeric_only o
r select only columns which should be valid for the function.
  df_top5grp1_APAC=df_APAC.groupby('Product Name').sum()
```

Out[13]:

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Motorola Smart Phone, Full Size	340608	34261.5270	59	1.61	7616.9670	3441.64
Office Star Executive Leather Armchair, Black	248381	15289.6353	37	1.02	6123.2553	1335.59
Samsung Smart Phone, VoIP	148533	16796.7360	31	0.60	5356.8060	1046.17
Nokia Smart Phone, with Caller ID	302086	30808.2402	54	1.24	5277.0402	3988.66
Novimex Executive Leather Armchair, Adjustable	316144	18832.2576	46	1.19	5047.9176	2997.20

The table shows that the sale of the Full Size Motorola Smart Phone generated more profits in the Asian Pacific market. The store should sell more of this product without neglecting the other four profitable products.

The 5 least profitable products in The Asian Pacific Market

In [14]:
```
df_AP= df[df['Market'].str.contains('APAC')]
df_top5grp1_AP=df_AP.groupby('Product Name').sum()
df_top5grp_AP=df_top5grp1_AP.nsmallest(5,'Profit')
df_top5grp_AP.head()
```

```
/var/folders/k7/31gr9yq52q53zqt_23tjp2q00000gn/T/ipykernel_7861/1840463907.py:2: Futu
reWarning: The default value of numeric_only in DataFrameGroupBy.sum is deprecated. I
n a future version, numeric_only will default to False. Either specify numeric_only o
r select only columns which should be valid for the function.
  df_top5grp1_AP=df_AP.groupby('Product Name').sum()
```

Out[14]:

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Apple Smart Phone, Full Size	313608	27251.1561	52	1.69	-2739.1239	2297.23
Hon Conference Table, Rectangular	109480	4769.8695	14	2.17	-2564.1705	267.26
Bevis Wood Table, with Bottom Storage	129253	7043.6520	19	1.60	-2284.3980	585.74
Breville Microwave, Silver	98589	7434.4668	34	0.94	-1924.0332	435.06
Chromcraft Round Table, Fully Assembled	49160	3326.8470	13	1.10	-1851.1830	207.17

The table above shows the five most unprofitable products in the Asian Pacific Market , with the Full Size Apple Smart Phone

being the product the store lost the most money on. The store should stop selling these products in this market and reallocate capital to selling more products that bring in the most profit

EMEA MARKET

The Top 5 most profitable products in The EMEA Market

In [15]:
```
df_EMEA= df[df['Market'].str.contains('EMEA')]
df_top5grp1_EMEA=df_EMEA.groupby('Product Name').sum()
df_top5grp_EMEA=df_top5grp1_EMEA.nlargest(5,'Profit')
df_top5grp_EMEA.head()
```

```
/var/folders/k7/31gr9yq52q53zqt_23tjp2q00000gn/T/ipykernel_7861/2431907235.py:2: FutureWarning: The default value of numeric_only in DataFrameGroupBy.sum is deprecated. In a future version, numeric_only will default to False. Either specify numeric_only or select only columns which should be valid for the function.
  df_top5grp1_EMEA=df_EMEA.groupby('Product Name').sum()
```

Out[15]:

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Cisco Smart Phone, Full Size	279266	12376.410	19	0.0	2722.320	1305.40
Barricks Training Table, Adjustable Height	100245	7050.780	22	0.0	2255.880	281.75
Motorola Smart Phone, Full Size	477587	9715.455	18	2.3	2019.435	1334.24
Bush Classic Bookcase, Pine	94122	3726.270	9	0.0	1825.740	235.84
Canon Wireless Fax, Digital	180774	3783.000	10	0.0	1664.400	429.95

The table shows that the sale of the Full Size Cisco Smart Phone generated more profits in the EMEA market. The store should sell more of this product without neglecting the other four profitable products.

The bottom 5 least profitable products in The EMEA Market

In [16]:
```
df_EM= df[df['Market'].str.contains('EMEA')]
df_top5grp1_EM=df_EM.groupby('Product Name').sum()
df_top5grp_EM=df_top5grp1_EM.nsmallest(5,'Profit')
df_top5grp_EM.head()
```

```
/var/folders/k7/31gr9yq52q53zqt_23tjp2q00000gn/T/ipykernel_7861/4267688179.py:2: Futu
reWarning: The default value of numeric_only in DataFrameGroupBy.sum is deprecated. I
n a future version, numeric_only will default to False. Either specify numeric_only o
r select only columns which should be valid for the function.
  df_top5grp1_EM=df_EM.groupby('Product Name').sum()
```

Out[16]:

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Motorola Smart Phone, Cordless	190037	7584.804	19	0.6	-3773.586	312.69
Barricks Conference Table, Rectangular	96533	3075.810	9	0.7	-2460.810	436.01
Office Star Executive Leather Armchair, Adjustable	44795	1503.744	8	0.6	-1842.096	126.17
Hoover Stove, White	91707	1926.474	7	0.6	-1841.556	133.11
Bevis Round Table, Adjustable Height	45920	1248.912	6	0.6	-1779.768	18.28

The table above shows the five most unprofitable products in the EMEA market, with the cordless Motorola Smart Phone being the product the store lost the most money on. The store should stop selling these products in this market and reallocate capital to selling more products that bring in the most profit.

EUROPEAN MARKET

The Top 5 most profitable products in The EU Market

In [17]:
```
df_EU= df[df['Market'].str.contains('EU')]
df_top5grp1_EU=df_EU.groupby('Product Name').sum()
df_top5grp_EU=df_top5grp1_EU.nlargest(5,'Profit')
df_top5grp_EU.head()
```

```
/var/folders/k7/31gr9yq52q53zqt_23tjp2q00000gn/T/ipykernel_7861/3817986505.py:2: Futu
reWarning: The default value of numeric_only in DataFrameGroupBy.sum is deprecated. I
n a future version, numeric_only will default to False. Either specify numeric_only o
r select only columns which should be valid for the function.
  df_top5grp1_EU=df_EU.groupby('Product Name').sum()
```

Out[17]:

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Hoover Stove, Red	93578	21147.0840	38	0.3	10345.5840	2640.76
Nokia Smart Phone, Full Size	267771	33511.3275	61	2.0	6890.7675	3191.68
Hamilton Beach Stove, Silver	115604	18247.8240	38	1.2	5452.4640	1132.12
SAFCO Executive Leather Armchair, Black	84919	12301.2000	30	0.6	5003.1000	1029.36
Safco Classic Bookcase, Metal	90074	10371.3570	25	0.3	4681.8570	1742.30

The table shows that the sale of the red Hoover Stove generated more profits in the EU market. The store should sell more of this product without neglecting the other four profitable products.

The bottom 5 least profitable products in The EU

In [18]:
```
df_EU= df[df['Market'].str.contains('EU')]
df_top5grp1_EU=df_EU.groupby('Product Name').sum()
df_top5grp_EU=df_top5grp1_EU.nsmallest(5,'Profit')
df_top5grp_EU.head()
```

```
/var/folders/k7/31gr9yq52q53zqt_23tjp2q00000gn/T/ipykernel_7861/1027496343.py:2: FutureWarning: The default value of numeric_only in DataFrameGroupBy.sum is deprecated. In a future version, numeric_only will default to False. Either specify numeric_only or select only columns which should be valid for the function.
  df_top5grp1_EU=df_EU.groupby('Product Name').sum()
```

Out[18]:

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Hoover Stove, White	78354	11728.827	31	1.20	-4958.163	1377.60
Office Star Executive Leather Armchair, Black	74425	6497.277	22	1.40	-3066.783	818.07
Fellowes File Cart, Single Width	215157	5809.482	60	3.80	-2430.918	557.44
Barricks Conference Table, Adjustable Height	34460	3728.007	9	1.05	-2082.393	463.58
Bevis Conference Table, Fully Assembled	11684	555.138	4	0.85	-1924.542	62.75

The table above shows the five most unprofitable products in the EU market, with the white Hoover Stove being the product the store lost the most money on. The store should stop selling these products in this market and reallocate capital to selling more products that bring in the most profit.

LATIN AMERICA MARKET

The Top 5 most profitable products in The Latin America Market

In [19]:
```
df_LATAM= df[df['Market'].str.contains('LATAM')]
df_top5grp1_LATAM=df_LATAM.groupby('Product Name').sum()
df_top5grp_LATAM=df_top5grp1_LATAM.nlargest(5,'Profit')
df_top5grp_LATAM.head()
```

```
/var/folders/k7/31gr9yq52q53zqt_23tjp2q00000gn/T/ipykernel_7861/3885008566.py:2: FutureWarning: The default value of numeric_only in DataFrameGroupBy.sum is deprecated. In a future version, numeric_only will default to False. Either specify numeric_only or select only columns which should be valid for the function.
  df_top5grp1_LATAM=df_LATAM.groupby('Product Name').sum()
```

Out[19]:

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Cisco Smart Phone, Full Size	44018	22088.616	52	0.6	5914.696	2127.49
Nokia Smart Phone, Full Size	65149	19873.280	51	1.0	4146.100	2367.21
Harbour Creations Executive Leather Armchair, Red	31392	8184.792	28	0.4	3387.832	527.45
Dania Classic Bookcase, Traditional	24190	7612.752	31	1.2	3283.292	777.28
Nokia Smart Phone, with Caller ID	22118	10392.936	28	0.8	3236.696	776.41

The table shows that the sale of the Full Size Cisco Smart Phone generated more profits in Latin America. The store should sell more of this product without neglecting the other four profitable products.

The 5 least profitable products in The Latin America Market

In [20]:
```
df_LA= df[df['Market'].str.contains('LATAM')]
df_top5grp1_LA=df_LA.groupby('Product Name').sum()
df_top5grp_LA=df_top5grp1_LA.nsmallest(5,'Profit')
df_top5grp_LA.head()
```

```
/var/folders/k7/31gr9yq52q53zqt_23tjp2q00000gn/T/ipykernel_7861/852715349.py:2: FutureWarning: The default value of numeric_only in DataFrameGroupBy.sum is deprecated. In a future version, numeric_only will default to False. Either specify numeric_only or select only columns which should be valid for the function.
  df_top5grp1_LA=df_LA.groupby('Product Name').sum()
```

Out[20]:

Product Name	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
Samsung Smart Phone, Cordless	68573	17636.400	54	2.2	-3757.320	2933.95
Lesro Computer Table, Fully Assembled	14351	1199.352	14	2.1	-2798.488	115.22
Chromcraft Conference Table, with Bottom Storage	12842	1047.204	6	1.4	-1780.356	181.30
Ikea Library with Doors, Pine	41053	4620.040	28	3.2	-1780.200	598.47
Bush Library with Doors, Pine	24516	2632.824	18	1.2	-1667.736	253.75

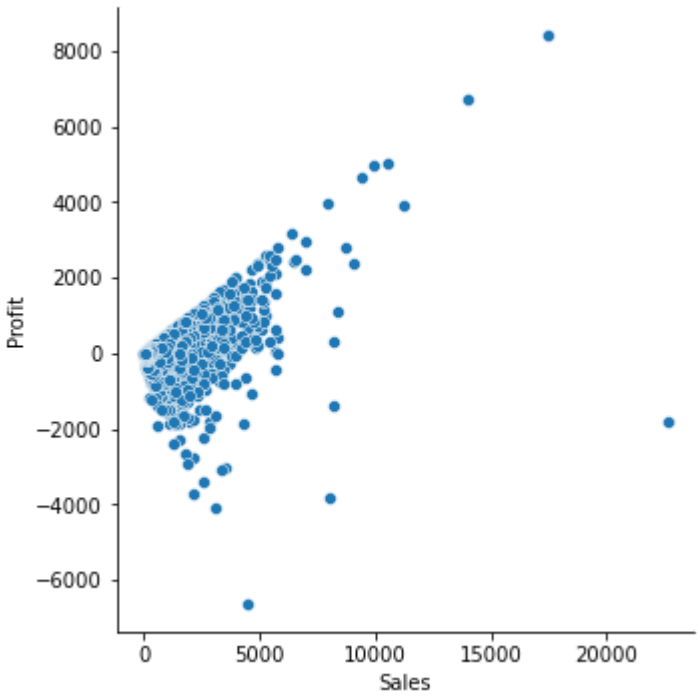
The table above shows the five most unprofitable products in the EU market, with the cordless Samsung Smart Phone being the product the store lost the most money on. The store should stop selling these products in this market and reallocate capital to selling more products that bring in the most profit.

UNCOVERING THE CORRELATION BETWEEN SALES AND PROFIT

In [21]:
```
sns.pairplot(df, x_vars=['Sales'], y_vars='Profit', size=5)
```

```
/Users/stevembikayi/opt/anaconda3/lib/python3.9/site-packages/seaborn/axisgrid.py:2076: UserWarning: The `size` parameter has been renamed to `height`; please update your code.
  warnings.warn(msg, UserWarning)
```

Out[21]:
```
<seaborn.axisgrid.PairGrid at 0x7ff6b2ead310>
```

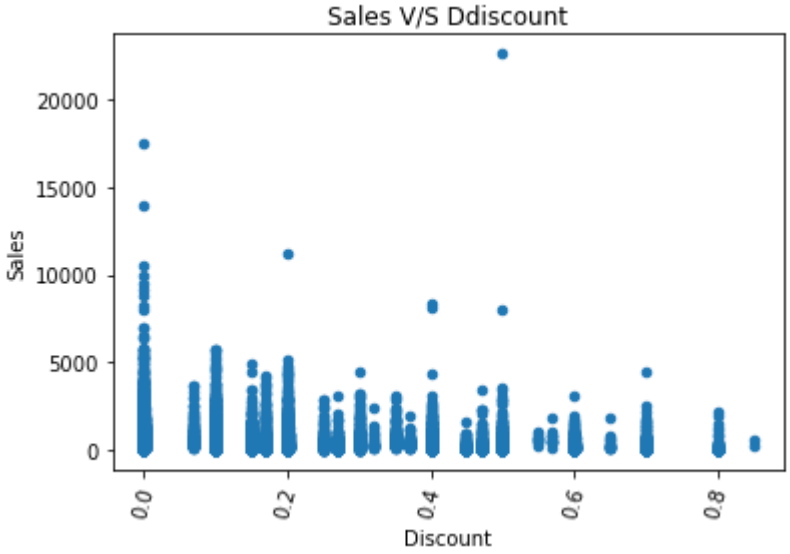


Looking at the sales and profit scatter plot, we can see a strong positive linear correlation between the two variables with most of the sales under the 5000 mark and a profit of under 2000. Also we can see that there are outliers with negative profits even when there is a sale, which is understandable as there are many variables that dictate profit such as product cost, shipping cost or discount.

CORRELATION BETWEEN DISCOUNT AND SALES

In [22]:
```
df.plot(kind='scatter',  y='Sales', x='Discount', rot=75, title='Sales V/S Ddiscount')
```

Out[22]:
```
<AxesSubplot:title={'center':'Sales V/S Ddiscount'}, xlabel='Discount', ylabel='Sale
s'>
```



Looking at the scatterplot of discounts and sales, we can see a negative linear correlation between the two variables. It is obvious that a higher discount does not necessarily translate into more sales, because the more discounts, the less sales the store makes.

In [ ]: