business-case-study

April 23, 2024

[64]:
```
import pandas as pd
import numpy as np
import plotly.express as px
import plotly.graph_objects as go
import plotly.io as pio
import plotly.colors as colors

pio.templates.default = "plotly_white"
```

[4]:
```
data = pd.read_csv("Business Case Study.csv", encoding= 'Latin-1')
```

[5]:
```
data
```

[5]:

	Row ID	Order ID	Order Date	Ship Date	Ship Mode	\
0	1	CA-2016-152156	11/8/2016	11/11/2016	Second Class	
1	2	CA-2016-152156	11/8/2016	11/11/2016	Second Class	
2	3	CA-2016-138688	6/12/2016	6/16/2016	Second Class	
3	4	US-2015-108966	10/11/2015	10/18/2015	Standard Class	
4	5	US-2015-108966	10/11/2015	10/18/2015	Standard Class	
…	…	…	…	…	…	
9989	9990	CA-2014-110422	1/21/2014	1/23/2014	Second Class	
9990	9991	CA-2017-121258	2/26/2017	3/3/2017	Standard Class	
9991	9992	CA-2017-121258	2/26/2017	3/3/2017	Standard Class	
9992	9993	CA-2017-121258	2/26/2017	3/3/2017	Standard Class	
9993	9994	CA-2017-119914	5/4/2017	5/9/2017	Second Class	

	Customer ID	Customer Name	Segment	Country	City	\
0	CG-12520	Claire Gute	Consumer	United States	Henderson	
1	CG-12520	Claire Gute	Consumer	United States	Henderson	
2	DV-13045	Darrin Van Huff	Corporate	United States	Los Angeles	
3	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	
4	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	
…	…	…	…	…	…	
9989	TB-21400	Tom Boeckenhauer	Consumer	United States	Miami	
9990	DB-13060	Dave Brooks	Consumer	United States	Costa Mesa	
9991	DB-13060	Dave Brooks	Consumer	United States	Costa Mesa	
9992	DB-13060	Dave Brooks	Consumer	United States	Costa Mesa	

```
9993    CC-12220      Chris Cortes   Consumer  United States      Westminster

      … Postal Code  Region       Product ID         Category Sub-Category  \
0     …       42420   South  FUR-BO-10001798        Furniture    Bookcases
1     …       42420   South  FUR-CH-10000454        Furniture       Chairs
2     …       90036    West  OFF-LA-10000240  Office Supplies       Labels
3     …       33311   South  FUR-TA-10000577        Furniture       Tables
4     …       33311   South  OFF-ST-10000760  Office Supplies      Storage
…   …         …     …              …               …           …
9989  …       33180   South  FUR-FU-10001889        Furniture  Furnishings
9990  …       92627    West  FUR-FU-10000747        Furniture  Furnishings
9991  …       92627    West  TEC-PH-10003645       Technology       Phones
9992  …       92627    West  OFF-PA-10004041  Office Supplies        Paper
9993  …       92683    West  OFF-AP-10002684  Office Supplies   Appliances

                                              Product Name     Sales  Quantity  \
0                    Bush Somerset Collection Bookcase  261.9600         2
1     Hon Deluxe Fabric Upholstered Stacking Chairs,…  731.9400         3
2     Self-Adhesive Address Labels for Typewriters b…   14.6200         2
3         Bretford CR4500 Series Slim Rectangular Table  957.5775         5
4                         Eldon Fold 'N Roll Cart System   22.3680         2
…                                                      …       …       …
9989                             Ultra Door Pull Handle   25.2480         3
9990  Tenex B1-RE Series Chair Mats for Low Pile Car…   91.9600         2
9991                               Aastra 57i VoIP phone  258.5760         2
9992  It's Hot Message Books with Stickers, 2 3/4" x 5"   29.6000         4
9993  Acco 7-Outlet Masterpiece Power Center, Wihtou…  243.1600         2

      Discount    Profit
0         0.00   41.9136
1         0.00  219.5820
2         0.00    6.8714
3         0.45 -383.0310
4         0.20    2.5164
…          …       …
9989      0.20    4.1028
9990      0.00   15.6332
9991      0.20   19.3932
9992      0.00   13.3200
9993      0.00   72.9480

[9994 rows x 21 columns]
```

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

[6]:
```
   Row ID        Order ID  Order Date   Ship Date       Ship Mode Customer ID  \
0       1  CA-2016-152156   11/8/2016  11/11/2016    Second Class    CG-12520
```

```
1       2  CA-2016-152156   11/8/2016  11/11/2016    Second Class    CG-12520
2       3  CA-2016-138688   6/12/2016   6/16/2016    Second Class    DV-13045
3       4  US-2015-108966  10/11/2015  10/18/2015  Standard Class    SO-20335
4       5  US-2015-108966  10/11/2015  10/18/2015  Standard Class    SO-20335

     Customer Name    Segment        Country             City  …  \
0      Claire Gute   Consumer  United States        Henderson  …
1      Claire Gute   Consumer  United States        Henderson  …
2  Darrin Van Huff  Corporate  United States      Los Angeles  …
3   Sean O'Donnell   Consumer  United States  Fort Lauderdale  …
4   Sean O'Donnell   Consumer  United States  Fort Lauderdale  …

  Postal Code  Region       Product ID         Category Sub-Category  \
0       42420   South  FUR-BO-10001798        Furniture    Bookcases
1       42420   South  FUR-CH-10000454        Furniture       Chairs
2       90036    West  OFF-LA-10000240  Office Supplies       Labels
3       33311   South  FUR-TA-10000577        Furniture       Tables
4       33311   South  OFF-ST-10000760  Office Supplies      Storage

                                        Product Name     Sales  Quantity  \
0                 Bush Somerset Collection Bookcase  261.9600         2
1  Hon Deluxe Fabric Upholstered Stacking Chairs,…  731.9400         3
2  Self-Adhesive Address Labels for Typewriters b…   14.6200         2
3      Bretford CR4500 Series Slim Rectangular Table  957.5775         5
4                    Eldon Fold 'N Roll Cart System   22.3680         2

   Discount    Profit
0      0.00   41.9136
1      0.00  219.5820
2      0.00    6.8714
3      0.45 -383.0310
4      0.20    2.5164

[5 rows x 21 columns]
```

[14]:
```
data[['Sales','Discount','Profit' ]].describe()
```

[14]:
```
              Sales     Discount       Profit
count   9994.000000  9994.000000  9994.000000
mean     229.858001     0.156203    28.656896
std      623.245101     0.206452   234.260108
min        0.444000     0.000000 -6599.978000
25%       17.280000     0.000000     1.728750
50%       54.490000     0.200000     8.666500
75%      209.940000     0.200000    29.364000
max    22638.480000     0.800000  8399.976000
```

[11]:
```
pd.isnull(data).sum()
```

[11]:
```
Row ID           0
Order ID         0
Order Date       0
Ship Date        0
Ship Mode        0
Customer ID      0
Customer Name    0
Segment          0
Country          0
City             0
State            0
Postal Code      0
Region           0
Product ID       0
Category         0
Sub-Category     0
Product Name     0
Sales            0
Quantity         0
Discount         0
Profit           0
dtype: int64
```

[15]:
```
data.columns
```

[15]:
```
Index(['Row ID', 'Order ID', 'Order Date', 'Ship Date', 'Ship Mode',
       'Customer ID', 'Customer Name', 'Segment', 'Country', 'City', 'State',
       'Postal Code', 'Region', 'Product ID', 'Category', 'Sub-Category',
       'Product Name', 'Sales', 'Quantity', 'Discount', 'Profit'],
      dtype='object')
```

0.1 Converting Order date in new column as order month,year,day

[19]:
```
data['Order Date'] = pd.to_datetime(data['Order Date'])   # the date available␣
 ↪in data set is of type string so we had to conver into date format by using␣
 ↪to_datetime
data['Ship Date'] = pd.to_datetime(data['Ship Date'])

data['Order Month'] = data['Order Date'].dt.month
data['Order Year'] = data['Order Date'].dt.year
data['Order day of week'] = data['Order Date'].dt.dayofweek
```

[20]:
```
data.head(3)
```

[20]:
```
   Row ID        Order ID Order Date  Ship Date     Ship Mode Customer ID  \
0       1  CA-2016-152156 2016-11-08 2016-11-11  Second Class    CG-12520
1       2  CA-2016-152156 2016-11-08 2016-11-11  Second Class    CG-12520
2       3  CA-2016-138688 2016-06-12 2016-06-16  Second Class    DV-13045

     Customer Name    Segment        Country         City  …  \
0      Claire Gute   Consumer  United States    Henderson  …
1      Claire Gute   Consumer  United States    Henderson  …
2  Darrin Van Huff  Corporate  United States  Los Angeles  …

          Category  Sub-Category  \
0        Furniture     Bookcases
1        Furniture        Chairs
2  Office Supplies        Labels

                                        Product Name   Sales Quantity  \
0                 Bush Somerset Collection Bookcase  261.96        2
1  Hon Deluxe Fabric Upholstered Stacking Chairs,…  731.94        3
2  Self-Adhesive Address Labels for Typewriters b…   14.62        2

  Discount    Profit  Order Month  Order Year  Order day of week
0      0.0   41.9136           11        2016                  1
1      0.0  219.5820           11        2016                  1
2      0.0    6.8714            6        2016                  6

[3 rows x 24 columns]
```

1 Month sales Analysis

[28]:
```
sales_month= data.groupby('Order Month')['Sales'].sum().reset_index()
fig = px.line(sales_month, x= 'Order Month', y= 'Sales',title='Monthly sales␣
 ↪Analysis')
fig.show()
```

1.0.1 From the above graph it can be viewed that the maximum sales(352.4611k) was on the month of November

2 Sales by Category

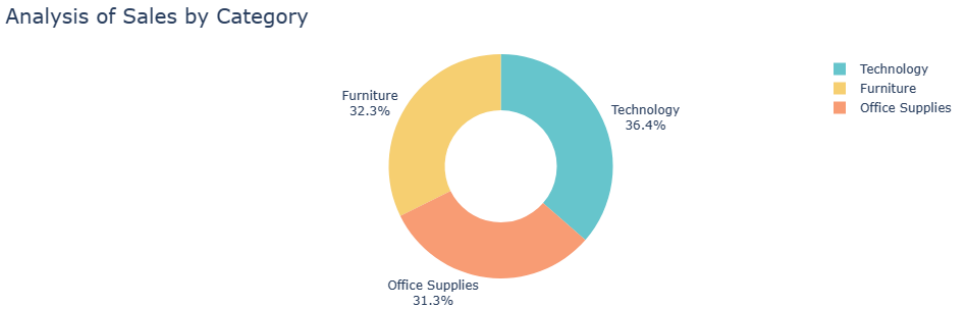
[31]:
```
data.columns
```

[31]:
```
Index(['Row ID', 'Order ID', 'Order Date', 'Ship Date', 'Ship Mode',
       'Customer ID', 'Customer Name', 'Segment', 'Country', 'City', 'State',
       'Postal Code', 'Region', 'Product ID', 'Category', 'Sub-Category',
       'Product Name', 'Sales', 'Quantity', 'Discount', 'Profit',
       'Order Month', 'Order Year', 'Order day of week'],
      dtype='object')
```

[45]:
```
sales_category= data.groupby('Category')['Sales'].sum().reset_index()

fig = px.pie(sales_category, values= 'Sales', names= 'Category',hole=0.
 ↪5,color_discrete_sequence= px.colors.qualitative.Pastel)
# Donut chart so there is a hole
fig.update_traces (textposition = 'outside', textinfo= 'percent+label' )
fig.update_layout (title_text= 'Analysis of Sales by Category',title_font=␣
 ↪dict(size=20))

fig.show()
```



3 Monthly Profit

[46]:
```
profit_month= data.groupby('Order Month')['Profit'].sum().reset_index()
```

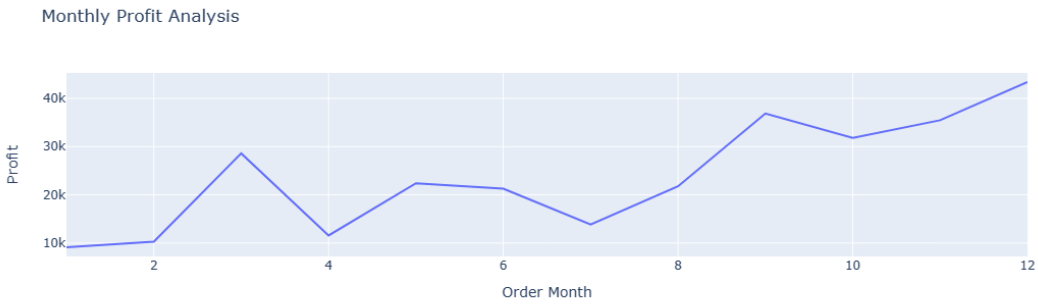
[47]:
```
profit_month
```

[47]:

	Order Month	Profit
0	1	9134.4461

```
1              2  10294.6107
2              3  28594.6872
3              4  11587.4363
4              5  22411.3078
5              6  21285.7954
6              7  13832.6648
7              8  21776.9384
8              9  36857.4753
9             10  31784.0413
10            11  35468.4265
11            12  43369.1919
```

[49]:
```
profit_month= data.groupby('Order Month')['Profit'].sum().reset_index()
fig = px.line(profit_month, x= 'Order Month', y= 'Profit',title='Monthly Profit␣
↪Analysis')
fig.show()
```



From the above graph we have come to the conclusion that the Maximum Profit was in the month of December of 43369.1919

4 Montly Profit by category

[51]:
```
sales_Profit= data.groupby('Category')['Profit'].sum().reset_index()
```

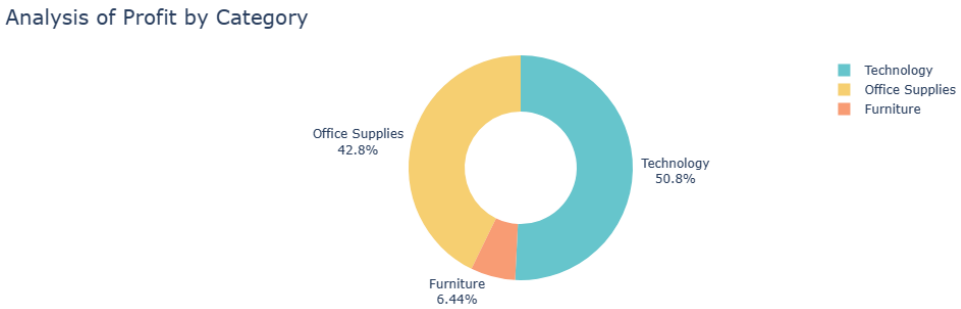
[52]:
```
sales_Profit
```

[52]:
```
          Category       Profit
0        Furniture   18451.2728
1  Office Supplies  122490.8008
2       Technology  145454.9481
```

[50]:
```
sales_Profit= data.groupby('Category')['Profit'].sum().reset_index()

fig = px.pie(sales_Profit, values= 'Profit', names= 'Category',hole=0.
 ↪5,color_discrete_sequence= px.colors.qualitative.Pastel)
# Donut chart so there is a hole
fig.update_traces (textposition = 'outside', textinfo= 'percent+label' )
fig.update_layout (title_text= 'Analysis of Profit by Category',title_font=␣
 ↪dict(size=20))

fig.show()
```



From the above graph we have come to the conclusion that the Maximum category that was sold is technology

5 Sales and Profit by customers segment

[57]:
```
data.columns
```

[57]:
```
Index(['Row ID', 'Order ID', 'Order Date', 'Ship Date', 'Ship Mode',
       'Customer ID', 'Customer Name', 'Segment', 'Country', 'City', 'State',
       'Postal Code', 'Region', 'Product ID', 'Category', 'Sub-Category',
       'Product Name', 'Sales', 'Quantity', 'Discount', 'Profit',
       'Order Month', 'Order Year', 'Order day of week'],
      dtype='object')
```

[61]:
```
sales_Profit_segment= data.groupby('Segment').agg({'Sales': 'sum','Profit':
 ↪'sum'}).reset_index()
```

[62]:
```
sales_Profit_segment
```

[62]:

	Segment	Sales	Profit
0	Consumer	1.161401e+06	134119.2092

```
1    Corporate  7.061464e+05   91979.1340
2  Home Office  4.296531e+05   60298.6785
```

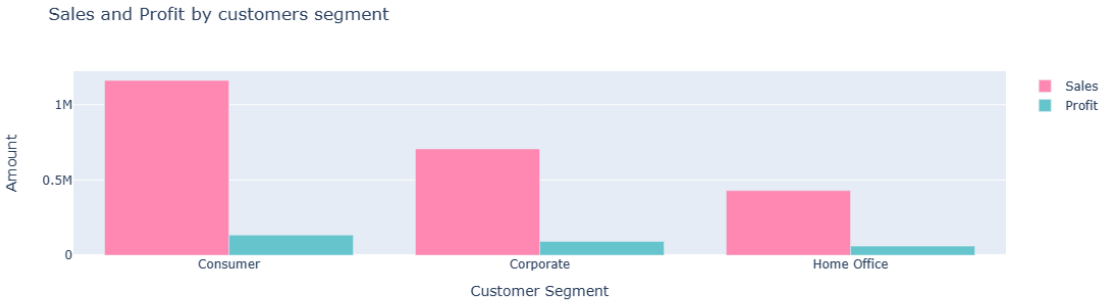
[60]:
```
sales_Profit_segment= data.groupby('Segment').agg({'Sales': 'sum','Profit':
 ↪'sum'}).reset_index()
color_palette= colors.qualitative.Pastel

fig=go.Figure()

fig.add_trace (go.Bar(x=␣
 ↪sales_Profit_segment['Segment'],y=sales_Profit_segment['Sales'], name=␣
 ↪'Sales',marker_color= color_palette[6]))
fig.add_trace (go.Bar(x=␣
 ↪sales_Profit_segment['Segment'],y=sales_Profit_segment['Profit'], name=␣
 ↪'Profit',marker_color= color_palette[0]))

fig.update_layout (title_text= 'Sales and Profit by customers␣
 ↪segment',xaxis_title= 'Customer Segment', yaxis_title= 'Amount')

fig.show()
```



[ ]: