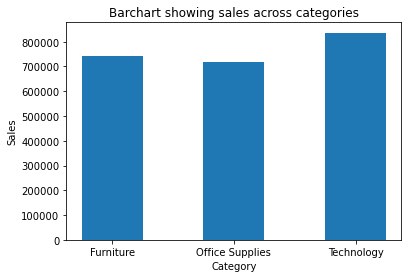
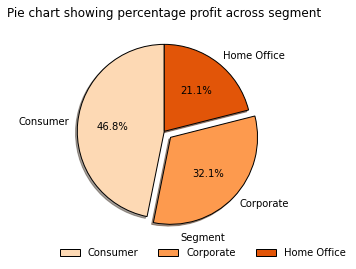
7PAM 2000 APPLIED DATA SCIENCE 1: ASSIGNMENT 1-VISUALIZATION.

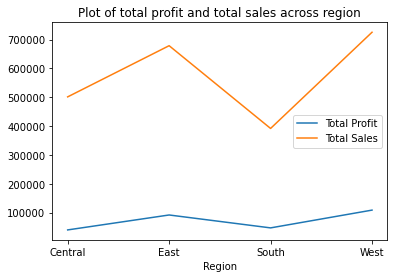


I chose the bar chart for the above presentation because it displays the data in a very understanding way and thereby making it easy to interpret.

The bar chart shows the sales across categories in a retail shop in the USA. From the chart, bulk of the sales (more than 800,000 units) came from technology category. The business owner can do more of tech products because much revenue will be generated from there.



I chose the pie chart for the above, to enable me have a pictorial and easy to quantify graph. The graph shows a percentage profit across the segment of the retail shop. Going by the presentation, 46.8% of the total profits, comes from the consumer segment. Hence, more resources should be channeled to this segment of the retail store.



The line plot above, shows the total sales and total profits made across different regions where the retail shop is located. The blue line represents the total profits made while the orange line represents total sales. In summary, it can be said that the West region generated the most sales and the highest profits. Also, the profit margins across the regions are not proportional to the volume of sales across the regions.

The central and south regions with sales of $500,000 and $400,000 respectively, generated about the same profit of below $100,000.

Link to data source: [Retail Supermarket | Kaggle](https://www.kaggle.com/datasets/roopacalistus/superstore)

2. **My Code**

#import relevant libraries

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

#load csv file via pandas read\_csv() function

df = pd.read\_csv('SampleSuperstore.csv')

#using pandas pivot table function to find total sales per category

table = pd.pivot\_table(data=df, index ='Category',values ='Sales', aggfunc=np.sum)

#Plotting barchart of aggregate sales across product categories

fig, ax = plt.subplots()

category = table.index

sales = table['Sales']

bar\_labels =['Furniture','Office Supplies', 'Technology']

ax.bar(category,sales,label=bar\_labels,width=0.5)

ax.set\_ylabel('Sales')

ax.set\_xlabel('Category')

ax.set\_title('Barchart showing sales across categories')

plt.show()

#Plotting pie chart to percentage of profit across segment

#using pandas pivot table function to find total sales across segment

table\_1 = pd.pivot\_table(data=df, index='Segment', values='Profit', aggfunc=np.sum)

fig, ax = plt.subplots()

Segment = table\_1.index

profit = table\_1['Profit']

colors = plt.get\_cmap('Oranges')(np.linspace(0.2, 0.7, len(profit)))

explode = (0,0.1,0)

pie\_labels =['Consumer', 'Corporate', 'Home Office']

ax.pie(profit,explode=explode, labels=pie\_labels, colors=colors, autopct='%1.1f%%', shadow=True,startangle=90,wedgeprops={"linewidth": 1, "edgecolor": "black"}, frame=False)

ax.set\_title('Pie chart showing percentage profit across segment')

ax.legend(title = 'Segment', loc='lower left', bbox\_to\_anchor= (0.0, -0.1), ncol=3, borderaxespad=0, frameon=False)

plt.show()

#making line plots from aggregate of sales made across categories and profit across segment

pivot = pd.pivot\_table(df,index='Region', values='Profit', aggfunc=np.sum)

pivot1 = pd.pivot\_table(df, index='Region', values='Sales', aggfunc=np.sum)

#pivot2 = pd.pivot\_table(df, index='Sub-Category', values='Sales', aggfunc=np.sum)

Region = pivot.index

profit\_region = pivot['Profit']

sales\_region = pivot1['Sales']

fig, ax = plt.subplots()

ax.plot(Region,profit\_region, Region, sales\_region)

ax.set\_xlabel('Region')

ax.legend(["Total Profit", "Total Sales"], loc ="best")

ax.set\_title('Plot of total profit and total sales across region')

plt.show()