Homework #2

CSCI 580, Fall 2023 Section 1

Liam Smith

▾ **Instructions**

In this assignment, you will be using the *ACME Sales* dataset to generate visualizations. Each question below asks you to provide a code section that will generate the requested chart type.

For full credit, make sure that each chart you generate

- has an appropriate title that clearly describes the information presented
- uses relevant labels for each axis
- deviates from any default style setting by changing the default (line/bar) color, marker shape, marker color, line width, or other feature
- uses any additional chart feature that will enhance the information your visualization is attempting to convey.

You are free to choose using `matplotlib` or `seaborn` or any other library to generate your visualization.

To get you started, the following code loads the *ACME Sales* dataset into a `pandas` `DataFrame` object and uses *Google Colab*'s interactive table feature to show you what the data looks like:

```
import pandas as pd

df = pd.read_csv('https://www.ecst.csuchico.edu/~bjuliano/csci581/datasets/acme_sa
df
```

	month_number	facecream	facewash	toothpaste	bathsoap	shampoo	moisturiz
0	1	2500	1500	5200	9200	1200	15(
1	2	2630	1200	5100	6100	2100	12(
2	3	2140	1340	4550	9550	3550	13
3	4	3400	1130	5870	8870	1870	11:
4	5	3600	1740	4560	7760	1560	17
5	6	2760	1555	4890	7490	1890	15!
6	7	2980	1120	4780	8980	1780	11:
7	8	3700	1400	5860	9960	2860	14(
8	9	3540	1780	6100	8100	2100	17
9	10	1990	1890	8300	10300	2300	18!
10	11	2340	2100	7300	13300	2400	21(
11	12	2900	1760	7400	14400	1800	17(

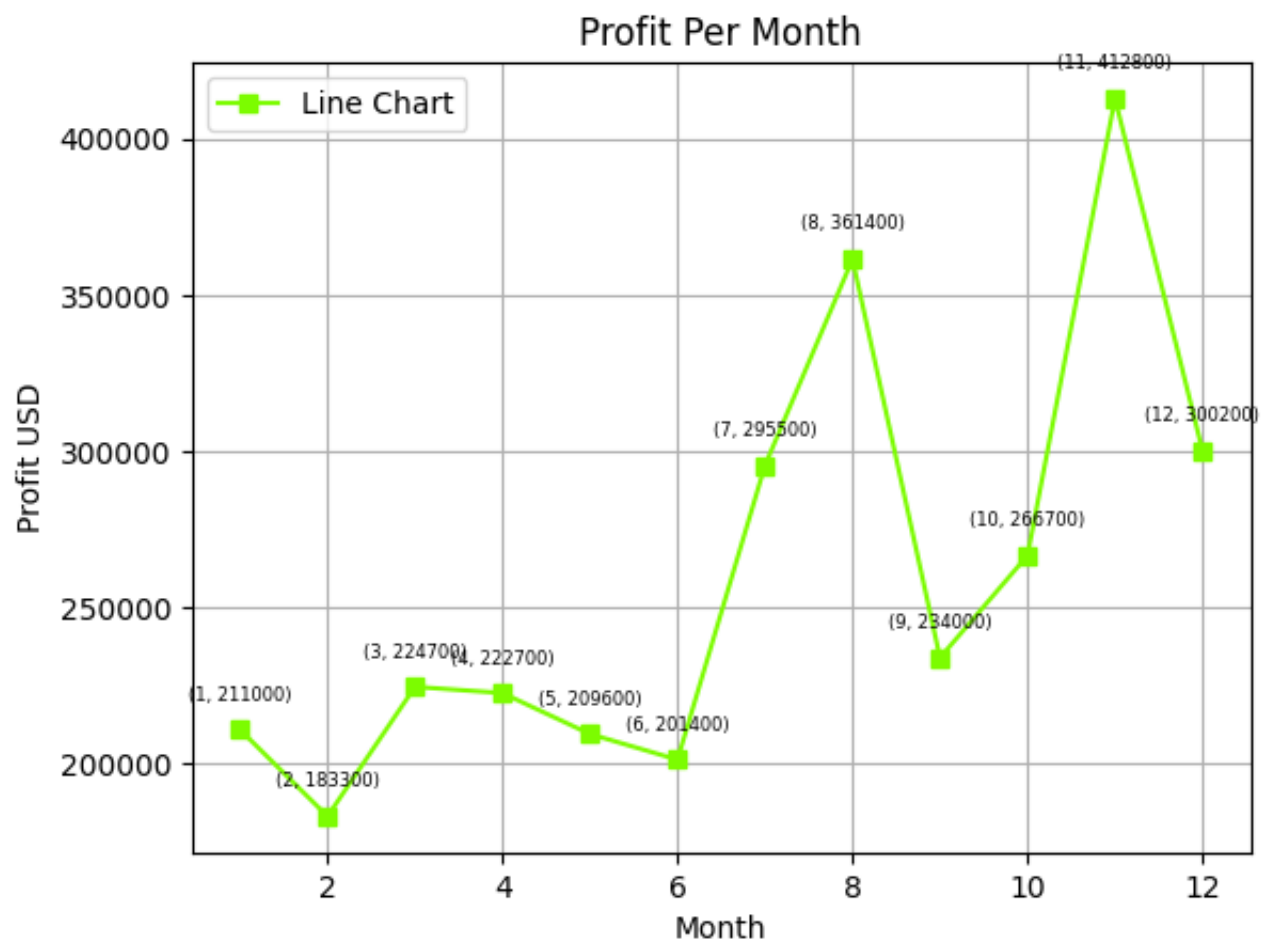# 1. Generate a line chart showing the company's total profit per month.

Solution:

```
import pandas as pd
import matplotlib.pyplot as plt
df = pd.read_csv('https://www.ecst.csuchico.edu/~bjuliano/csci581/datasets/acme_sa
df
x_vals = df['month_number']
y_vals = df['total_profit']
```

```python
plt.plot(x_vals, y_vals, label='Line Chart', marker='s', linestyle='-', color='Lawr
for x, y in zip(x_vals, y_vals):
    plt.annotate(f'({x}, {y})', (x, y), xytext=(0, 10), textcoords="offset points",

plt.xlabel('Month')
plt.ylabel('Profit USD')
plt.title('Profit Per Month')
plt.legend()
plt.grid(True)



plt.show()
```

2. Generate a multiline plot (*i.e.*, use a separate plot line for each of the six products) showing the total units sold per month for each product.

Solution(s)

```
import pandas as pd
import matplotlib.pyplot as plt
df = pd.read_csv('https://www.ecst.csuchico.edu/~bjuliano/csci581/datasets/acme_sal
df
x_vals = df['month_number']
creamY = df['facecream']
washY = df['facewash']
moistY = df['moisturizer']
shampY = df['shampoo']
soapY = df['bathsoap']
toothY = df['toothpaste']
plt.plot(x_vals, creamY, label='Shaving Cream', marker='s', linestyle='-', color='S
plt.plot(x_vals, washY, label='Face Cream', marker='h', linestyle='-', color='orang
plt.plot(x_vals, moistY, label='Moisturizer', marker='p', linestyle='-', color='pur
plt.plot(x_vals, shampY, label='Shampoo', marker='v', linestyle='-', color='red')
plt.plot(x_vals, toothY, label='Bathsoap', marker='^', linestyle='-', color='green'
plt.plot(x_vals, toothY, label='Toothpaste', marker='o', linestyle='-', color='blue
plt.xlabel('Month Number')
plt.ylabel('Units Sold')
plt.legend()
plt.show()
```
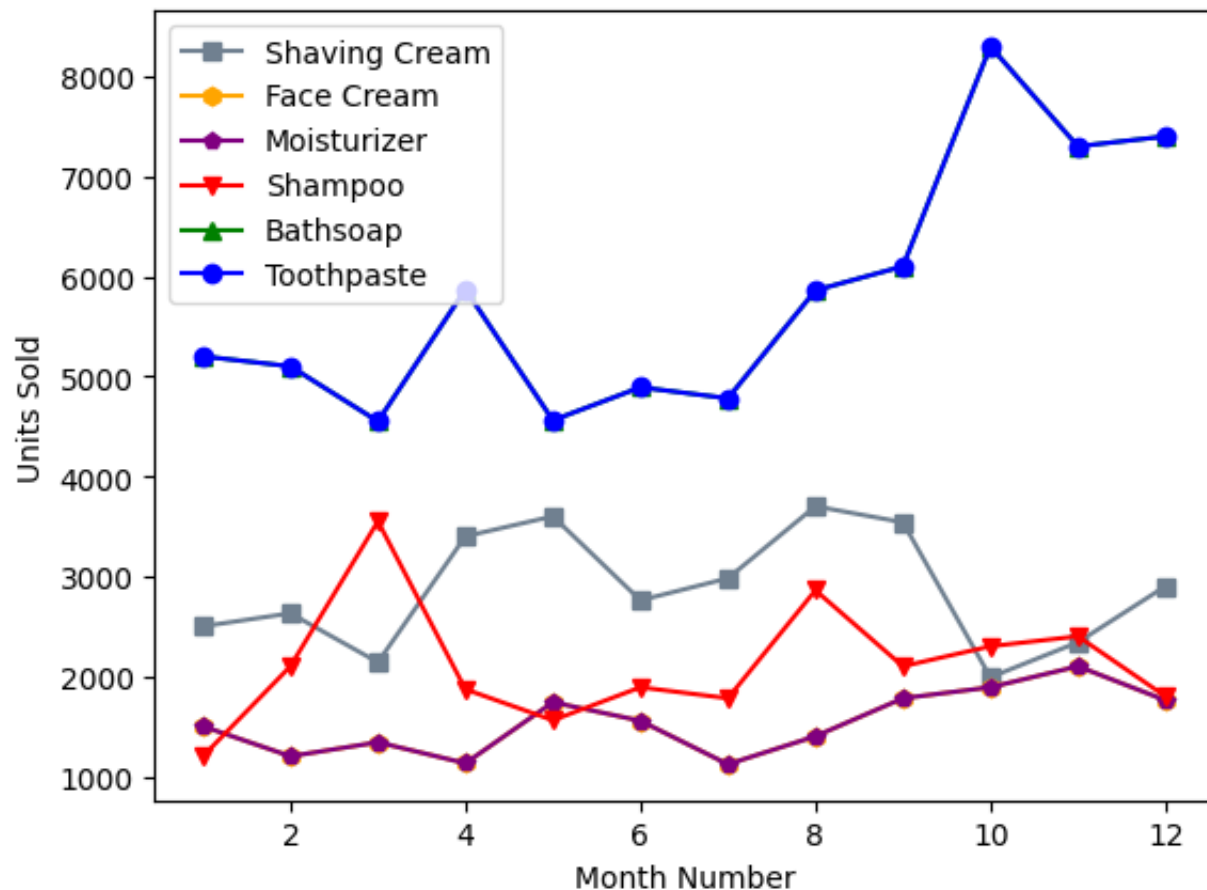
## 3. Generate a scatter plot that shows the toothpaste sales data for each month.

Solution(s)

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
df = pd.read_csv('https://www.ecst.csuchico.edu/~bjuliano/csci581/datasets/acme_sal
df
x_vals = df['month_number']
# x_vals = df['facecream']
y_vals = df['facecream']
plt.scatter(x_vals, y_vals, marker='o', linestyle='-', color='purple')
```

```python
for x, y in zip(x_vals, y_vals):
    plt.annotate(f'({x}, {y})', (x, y), xytext=(0, 10), textcoords="offset points",


# plt.xticks(df['month_number'].tolist())
plt.xlabel('Month')
plt.ylabel('Profit USD')
plt.title('Monthly toothpaste sales data')
plt.legend()
plt.grid(True)




plt.show()
```

/var/folders/dq/hsqrwhy96kv60x8zvgbhx3800000gn/T/ipykernel_12023/2183826115.py
See: https://matplotlib.org/stable/tutorials/intermediate/legend_guide.html#in
  plt.legend()
No artists with labels found to put in legend. Note that artists whose label

Monthly toothpaste sales data
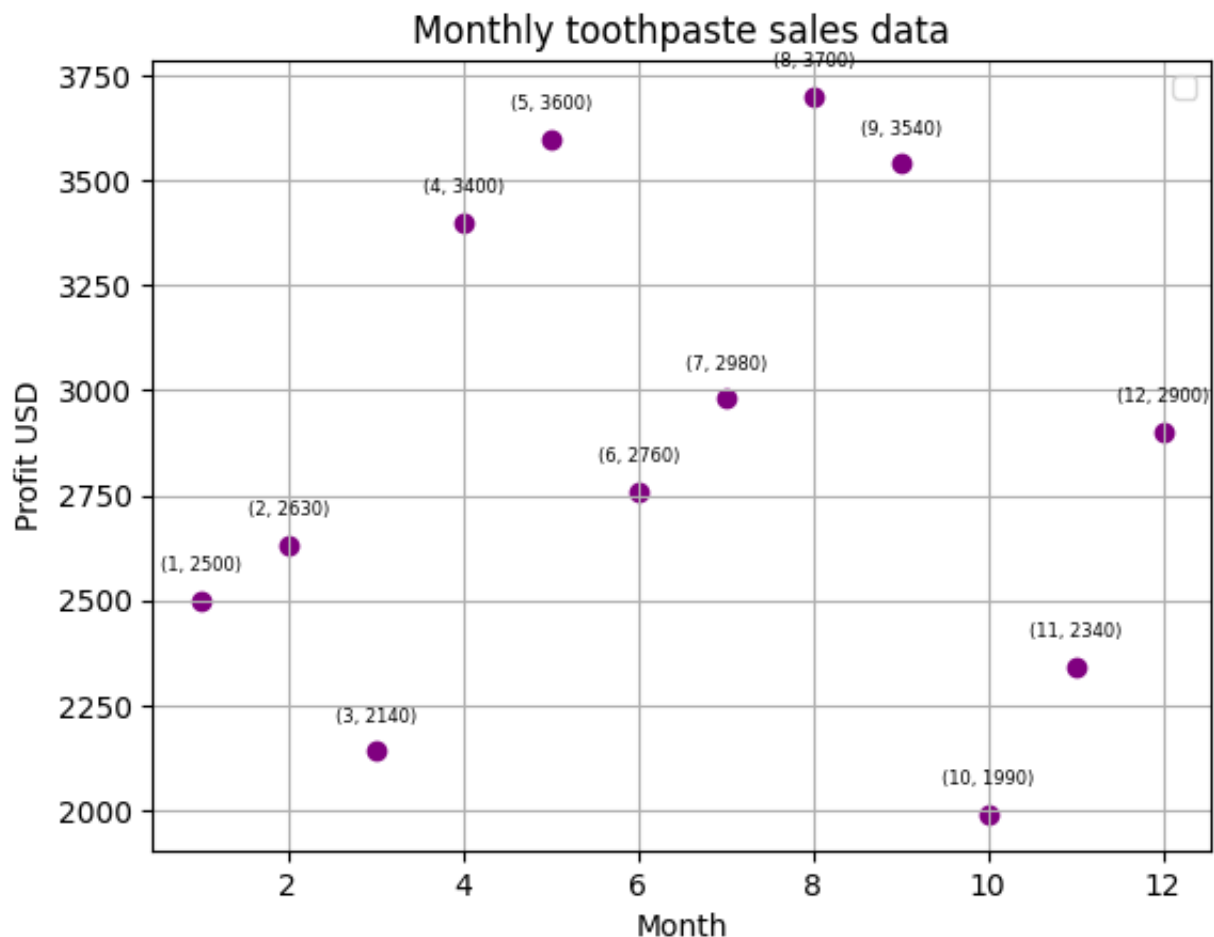
4. Generate a bar chart showing the face cream and face
▾ wash sales per month. Use a separate bar for each product
in the same chart.

▾ Solution(s)

```
import pandas as pd
import matplotlib.pyplot as plt
import matplotlib.pylab as plb

df = pd.read_csv('https://www.ecst.csuchico.edu/~bjuliano/csci581/datasets/acme_sal
df

barWidth = 2000


cream = df['facecream'].tolist()
wash = df['facewash'].tolist()
months = df['month_number'].tolist()

fig, axis = plt.subplots()
print(fig)
# cream = [ -x for x in cream]
print(cream)
bars1 = axis.bar(x=months, height=cream, align='edge', label='Facecream', width=-0.
for bar in bars1:
    y = bar.get_height()
    plt.text(bar.get_x() + bar.get_width() / 2, y+20, round(y, 2), va='bottom', ha=

bars2 = axis.bar(x=months, height=wash, align='edge', label='Facewash', width=0.4,
for bar in bars2:
    y = bar.get_height()
    plt.text(bar.get_x() + bar.get_width() / 2, y, round(y, 2), va='bottom', ha='ce
plt.ylabel('Units sold')
plt.xlabel("Month Number")
plt.title('Sales Data of Acme Facewash and Facecream')
# plt.xlabel(df['month_number'].tolist())
```
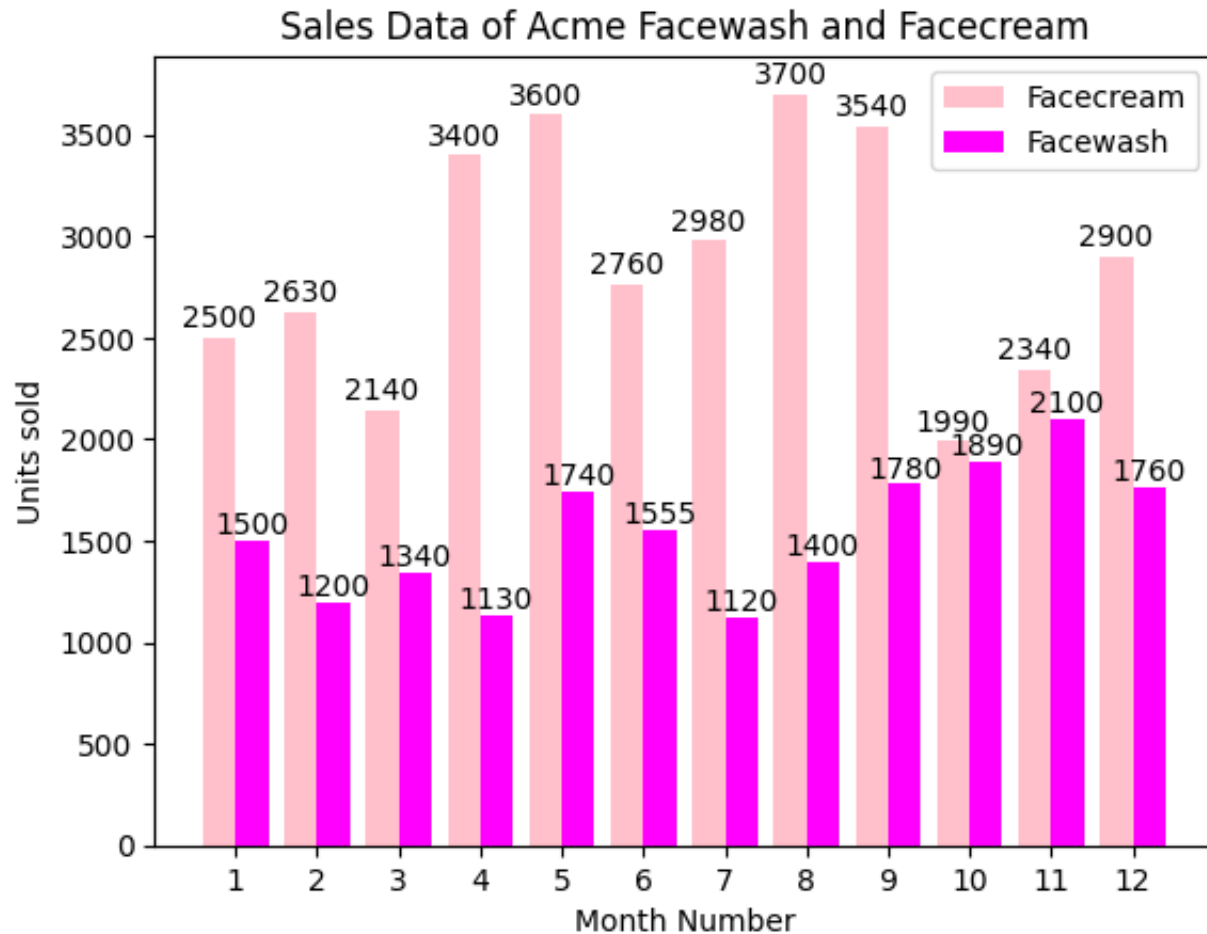
```
plt.xticks(months)
plt.legend()
plt.show()
```

```
Figure(640x480)
[2500, 2630, 2140, 3400, 3600, 2760, 2980, 3700, 3540, 1990, 2340, 2900]
```


5. Generate a pie chart showing the sales proportion of each product for the total sale for the year. Explode the slice that has the highest proportion.

▾ Solution(s)

```
import pandas as pd
```

```python
import matplotlib.pyplot as plt
df = pd.read_csv('https://www.ecst.csuchico.edu/~bjuliano/csci581/datasets/acme_sal
df

cream = df['facecream'].tolist()
wash = df['facewash'].tolist()
moist = df['moisturizer'].tolist()
shamp = df['shampoo'].tolist()
soap = df['bathsoap'].tolist()
tooth = df['toothpaste'].tolist()

labels = ['Facecream', 'Facewash', 'Moisturizer', 'Shampoo', 'Bathsoap', 'Toothpast
# labels = list(labels)
plt.legend()
plt.title('Percentage of Sales per Product for the Year')
percentages = [sum(x) for x in zip(cream, wash, moist, shamp, soap, tooth)]
maxPos = percentages.index(max(percentages))

seperate = [0.1 if i == maxPos else 0 for i in range(len(percentages))]
plt.pie(percentages, explode=seperate, autopct='%1.1f%%')
```

```
    No artists with labels found to put in legend.  Note that artists whose label
    ([<matplotlib.patches.Wedge at 0x129871e90>,
      <matplotlib.patches.Wedge at 0x1298a4c50>,
      <matplotlib.patches.Wedge at 0x1298a6410>,
      <matplotlib.patches.Wedge at 0x1298a7b10>,
      <matplotlib.patches.Wedge at 0x1298b5150>,
      <matplotlib.patches.Wedge at 0x1298b6950>,
      <matplotlib.patches.Wedge at 0x1298c0210>,
      <matplotlib.patches.Wedge at 0x1298c1b50>,
      <matplotlib.patches.Wedge at 0x1298b5250>,
      <matplotlib.patches.Wedge at 0x1298c48d0>,
      <matplotlib.patches.Wedge at 0x1298c6010>,
      <matplotlib.patches.Wedge at 0x1298c7710>],
     [Text(1.0695005389693513, 0.25723257403421385, ''),
      Text(0.8573228361733413, 0.6892006635049756, ''),
      Text(0.4658216562783232, 0.9964989636432744, ''),
      Text(-0.06947722811300897, 1.0978036777009508, ''),
      Text(-0.5719444887538202, 0.9396166781108835, ''),
      Text(-0.9294931397842136, 0.5882537743985028, ''),
      Text(-1.0937420851278845, 0.11716762018624347, ''),
      Text(-1.0100630208985955, -0.43562907824570635, ''),
      Text(-0.6392952352401522, -0.895154513030705, ''),
      Text(-0.06603846654573663, -1.0980159019505535, ''),
      Text(0.5924068388935678, -0.9268517342229718, ''),
```

```
  Text(1.1329699782371347, -0.3954478833087194, '')],
 [Text(0.5833639303469188, 0.1403086767459348, '7.5%'),
  Text(0.4676306379127315, 0.3759276346390776, '6.5%'),
  Text(0.25408453978817624, 0.5435448892599678, '8.0%'),
  Text(-0.03789666987982306, 0.5988020060187004, '7.9%'),
  Text(-0.31196972113844734, 0.5125181880604818, '7.5%'),
  Text(-0.5069962580641164, 0.3208656951264561, '7.2%'),
  Text(-0.596586591887937, 0.06390961101067826, '7.4%'),
  Text(-0.5509434659446883, -0.23761586086129435, '9.0%'),
  Text(-0.34870649194917386, -0.48826609801674814, '8.3%'),
  Text(-0.03602098175221 9974, -0.5989177647003019, '9.5%'),
  Text(0.32313100303285514, -0.5055554913943482, '10.5%'),
  Text(0.660899153971662, -0.2306779319300863, '10.7%')])
```

Percentage of Sales per Product for the Year
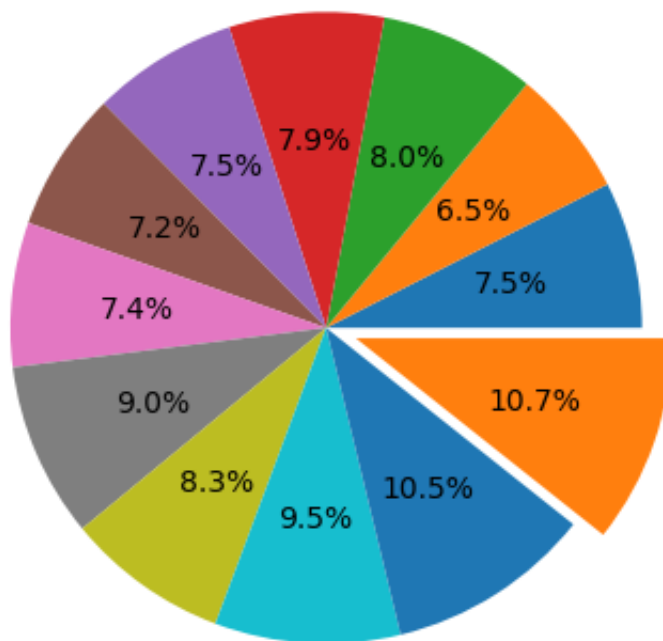
Notes

(Optional) Include any final thoughts, comments, or observations here, if applicable.