Tutorial Week 4: Matplotlib for Visualization (Solutions)

Click to download the Jupyter Notebook files (.ipynb)

Example: Read total sale units for all months and show it using a line plot

Total sale units data provided for each month, generated a line plot as shown at the end of this question (directly below the code cell).

Hint:

- The data are available in the files company_sales_data.csv.
- Use `plt.plot()`.

[1]:
```
import pandas as pd
import matplotlib.pyplot as plt

df = pd.read_csv("company_sales_data.csv")
df.head(5)
```

[1]:
```
   month_number  facecream  facewash  toothpaste  bathingsoap  shampoo  \
0             1       2500      1500        5200         9200     1200
1             2       2630      1200        5100         6100     2100
2             3       2140      1340        4550         9550     3550
3             4       3400      1130        5870         8870     1870
4             5       3600      1740        4560         7760     1560

   moisturizer  total_units
0         1500        21100
1         1200        18330
2         1340        22470
3         1130        22270
4         1740        20960
```

[2]:
```
### Start your code here ###

plt.plot(df['month_number'], df['total_units'])
plt.xlabel('Month number')
plt.ylabel('Total units')
plt.title('Total units sold per month')

plt.show()

### End your code here ###
```

1 Questions

Hint:

- The data are available in the files company_sales_data.csv.
- Write **your code** between two comment lines: `### Start/End your code here ###`.
- **Expected output** is shown at the end of each question (directly below the code cell).
- See the Matplotlib gallery that shows hundreds of charts made with Matplotlib.

1.1 Read all product sales data and show it using a multiline plot

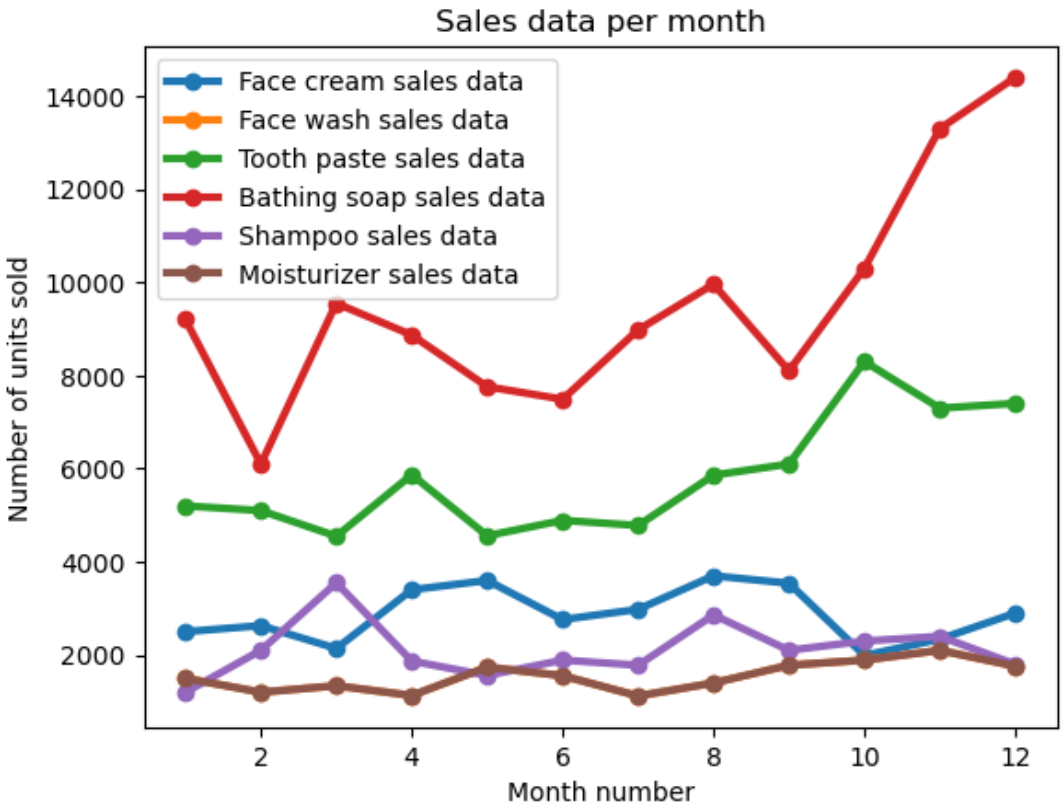
Display the number of units sold per month for each product using multiline plots (i.e., separate line for each product).

Hint:

- Use `plt.plot()`.

[3]:
```
import pandas as pd
import matplotlib.pyplot as plt

df = pd.read_csv("company_sales_data.csv")

### Start your code here ###
```

```
plt.plot(df['month_number'], df['facecream'], label = 'Face cream sales␣
 ↪data', marker='o', linewidth=3)
plt.plot(df['month_number'], df['facewash'], label = 'Face wash sales data', ␣
 ↪marker='o', linewidth=3)
plt.plot(df['month_number'], df['toothpaste'], label = 'Tooth paste sales␣
 ↪data', marker='o', linewidth=3)
plt.plot(df['month_number'], df['bathingsoap'], label = 'Bathing soap sales␣
 ↪data', marker='o', linewidth=3)
plt.plot(df['month_number'], df['shampoo'], label = 'Shampoo sales data',␣
 ↪marker='o', linewidth=3)
plt.plot(df['month_number'], df['moisturizer'], label = 'Moisturizer sales␣
 ↪data', marker='o', linewidth=3)

plt.xlabel('Month number')
plt.ylabel('Number of units sold')
plt.legend(loc='upper left')
plt.title('Sales data per month')

plt.show()

### End your code here ###
```

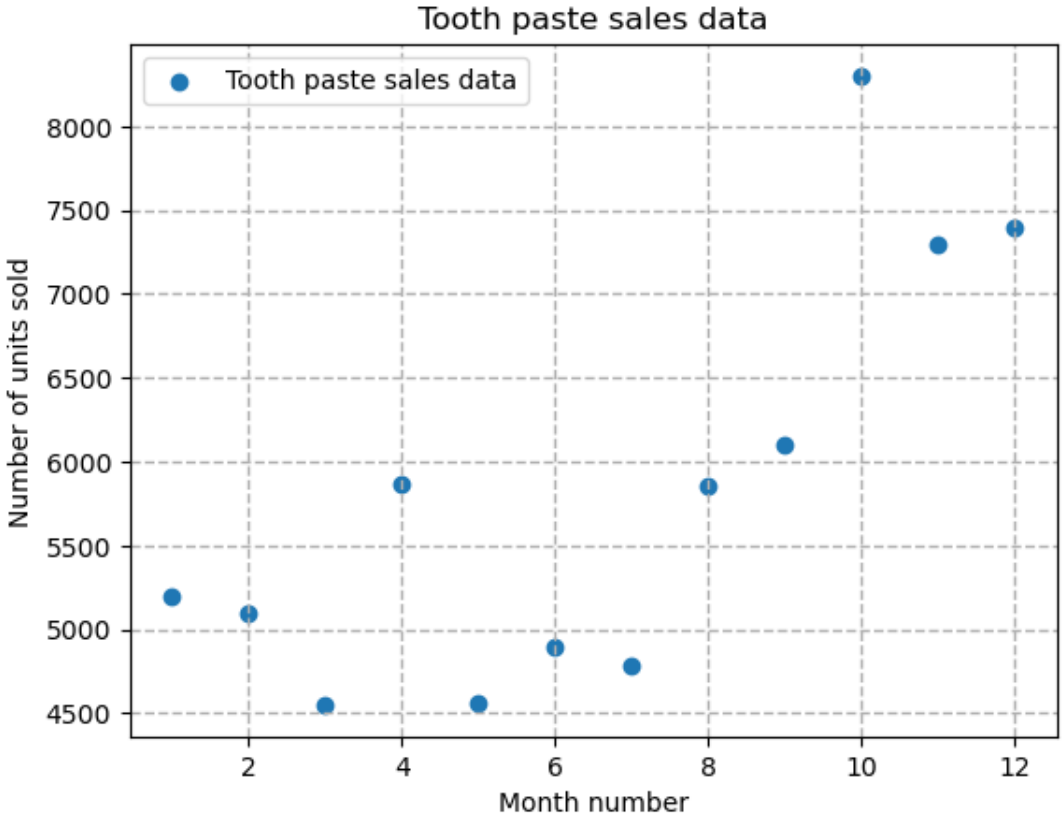


1.2 Read toothpaste sales data of each month and show it using a scatter plot

Hint:

- Use `plt.scatter()`.

[4]:
```
import pandas as pd
import matplotlib.pyplot as plt

df = pd.read_csv("company_sales_data.csv")

### Start your code here ###

plt.scatter(df['month_number'], df['toothpaste'], label = 'Tooth paste sales␣
↪data')

plt.xlabel('Month number')
plt.ylabel('Number of units sold')
plt.legend(loc='upper left')
plt.title('Tooth paste sales data')
plt.grid(True, linewidth= 1, linestyle="--")

plt.show()

### End your code here ###
```

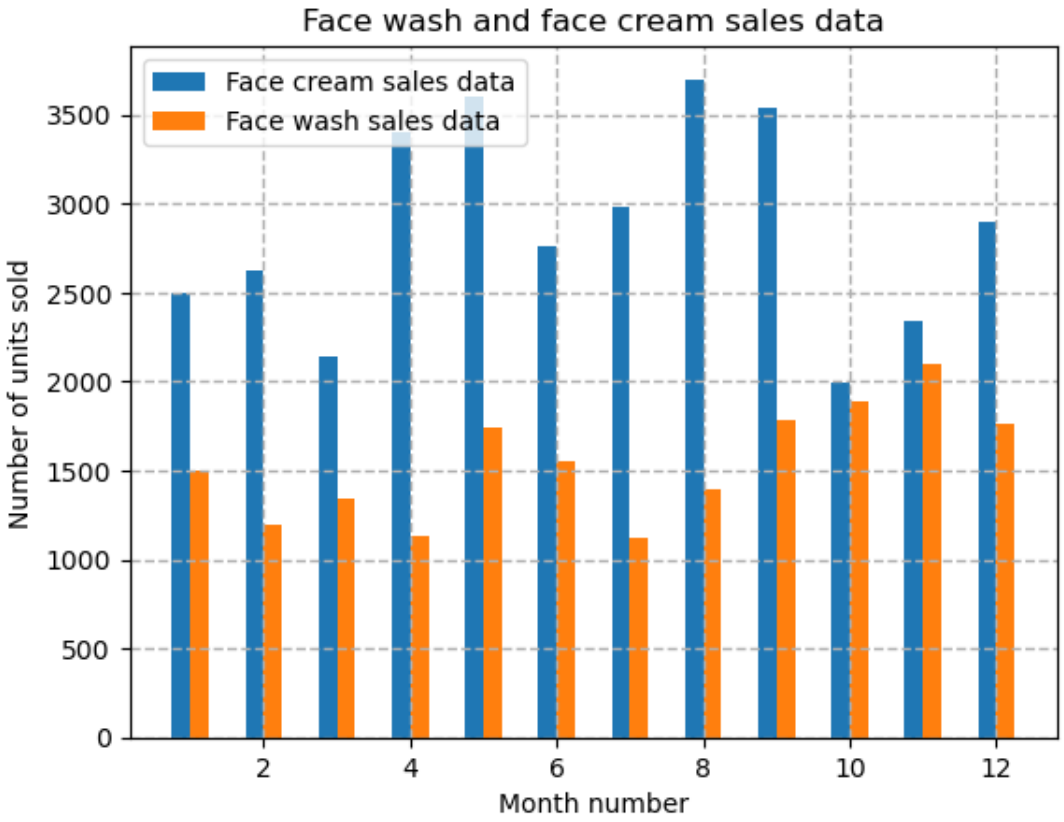


1.3 Read face cream and face wash product sales data and show it using the bar chart

Hint:

- Use `plt.bar()`.

[5]:
```
import pandas as pd
import matplotlib.pyplot as plt

df = pd.read_csv("company_sales_data.csv")

### Start your code here ###

monthList  = df ['month_number'].tolist()

plt.bar([a-0.25 for a in monthList], df['facecream'], width= 0.25, label =␣
↪'Face cream sales data', align='edge')
plt.bar([a+0.25 for a in monthList], df['facewash'], width= -0.25, label =␣
↪'Face wash sales data',  align='edge')

plt.xlabel('Month number')
plt.ylabel('Number of units sold')
plt.legend(loc='upper left')
plt.title('Face wash and face cream sales data')
plt.grid(True, linewidth= 1, linestyle="--")

plt.show()

### End your code here ###
```



1.4 Calculate total sale data for the whole year for each product and show it using a Pie chart

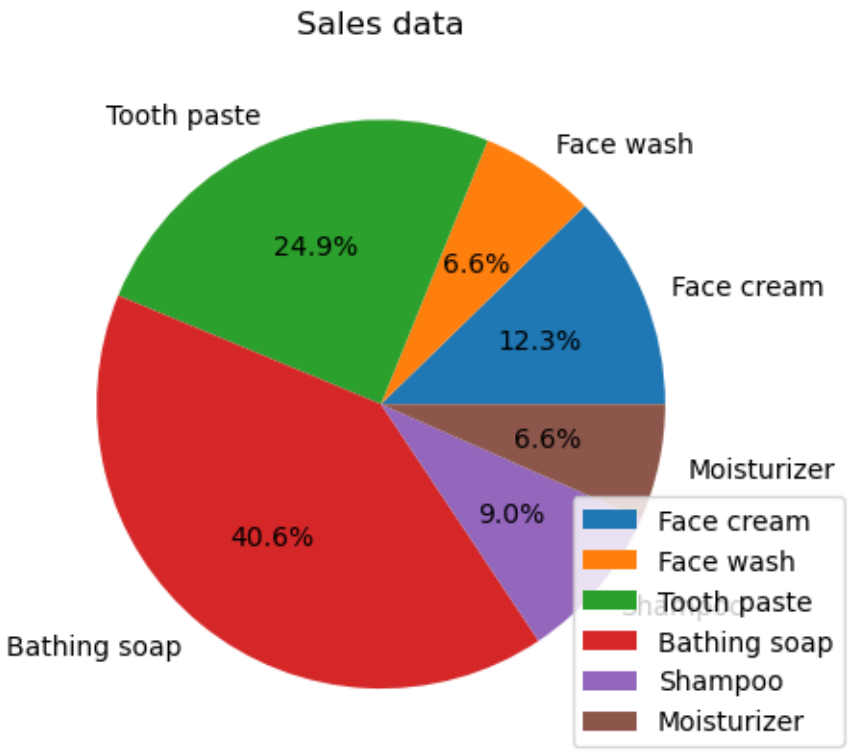
The Pie chart displays number of units sold per year for each product in percentage.

Hint:

- Use `plt.pie()`.

[6]:
```
import pandas as pd
import matplotlib.pyplot as plt

df = pd.read_csv("company_sales_data.csv")

### Start your code here ###

labels = ['Face cream', 'Face wash', 'Tooth paste', 'Bathing soap',␣
↪'Shampoo', 'Moisturizer']
salesData = [df['facecream'].sum(), df['facewash'].sum(), df['toothpaste'].
↪sum(),
             df['bathingsoap'].sum(), df['shampoo'].sum(), df['moisturizer'].
↪sum()]

plt.axis("equal")
plt.pie(salesData, labels=labels, autopct='%1.1f%%')
```

```
plt.legend(loc='lower right')
plt.title('Sales data')

plt.show()

### End your code here ###
```



1.5 Read the data of bathing soap and face wash and display them in separate subplots

Hint:

- Use `plt.subplots()`.

[7]:
```
import pandas as pd
import matplotlib.pyplot as plt

df = pd.read_csv("company_sales_data.csv")

### Start your code here ###

f, ax = plt.subplots(2, sharex=True)
ax[0].plot(df['month_number'], df['bathingsoap'], label = 'Bathingsoap sales␣
 ↪data', color='k', marker='o', linewidth=3)
ax[0].set_ylabel('Sold units of bathing soap')
```

```
ax[1].plot(df['month_number'], df['facewash'], label = 'Face wash sales␣
 ↪data', color='r', marker='o', linewidth=3)
ax[1].set_ylabel('Sold units of face wash')


plt.xlabel('Month number')

plt.show()

### End your code here ###
```