Lab2

July 31, 2025

Machine Learning Lab #2

Data Analysis and Visualisation with Python

[31]:
```
#Generic Imports

import numpy as np
import pandas as pd
import matplotlib
from matplotlib import pyplot as plt
from scipy.sparse import csr_matrix
```

Questions

1. Follow along with these steps:

a) Create a figure object called fig using plt.figure()

[38]:
```
fig = plt.figure()
```

```
<Figure size 640x480 with 0 Axes>
```

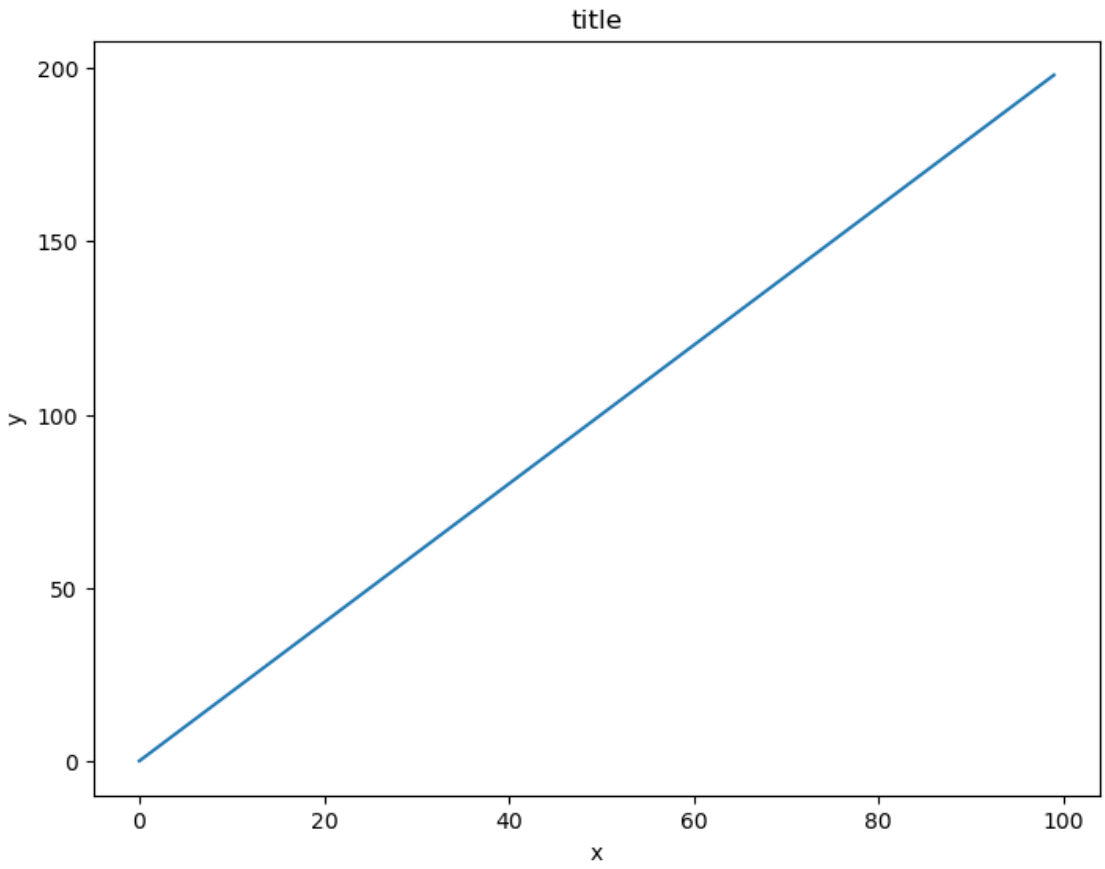
b) Use add_axes to add an axis to the figure canvas at [0,0,1,1]. Call this new axis ax.

[39]:
```
x = np.arange(0,100)
y = x*2
```

c) Plot (x,y) on that axes and set the labels and titles to match the plot given in the lab manual

[40]:
```
ax = fig.add_axes([0,0,1,1])
ax.plot(x,y)
ax.set_xlabel('x')
ax.set_ylabel('y')
ax.set_title('title')
ax.set_yticks([0,50,100,150,200])
fig
```

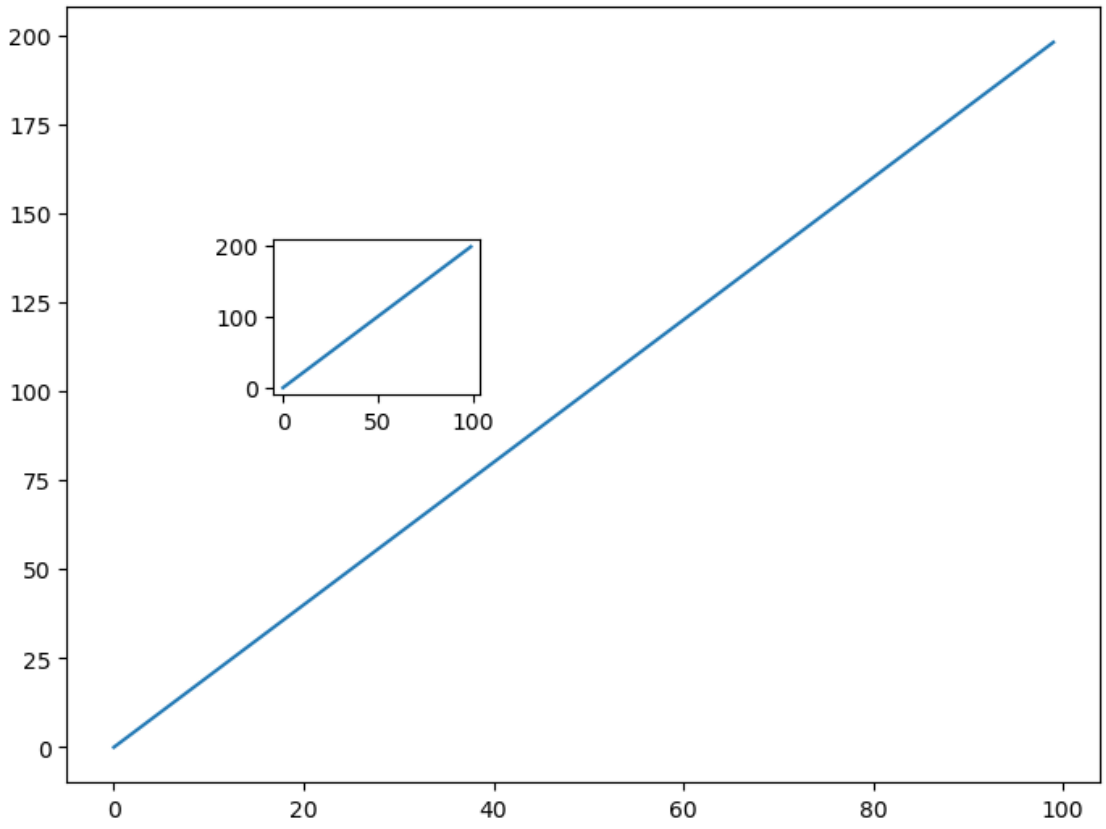
[40]:



2. Create a figure object and put two axes on it, ax1 and ax2. Located at [0,0,1,1] and [0.2,0.5,.2,.2] respectively. Now plot (x,y) on both axes. And call your figure object to show it.

[41]:
```
fig = plt.figure()
ax1 = fig.add_axes([0,0,1,1])
ax2 = fig.add_axes([0.2,0.5,.2,.2])
ax1.plot(x,y)
ax2.plot(x,y)
```

[41]:
```
[<matplotlib.lines.Line2D at 0x7f8e77d48d10>]
```



3. Use the company sales dataset csv file, read Total profit of all months and show it using a line plot. Total profit data provided for each month. Generated line plot must include the following properties:
4. X label name = Month Number
5. Y label name = Total profit

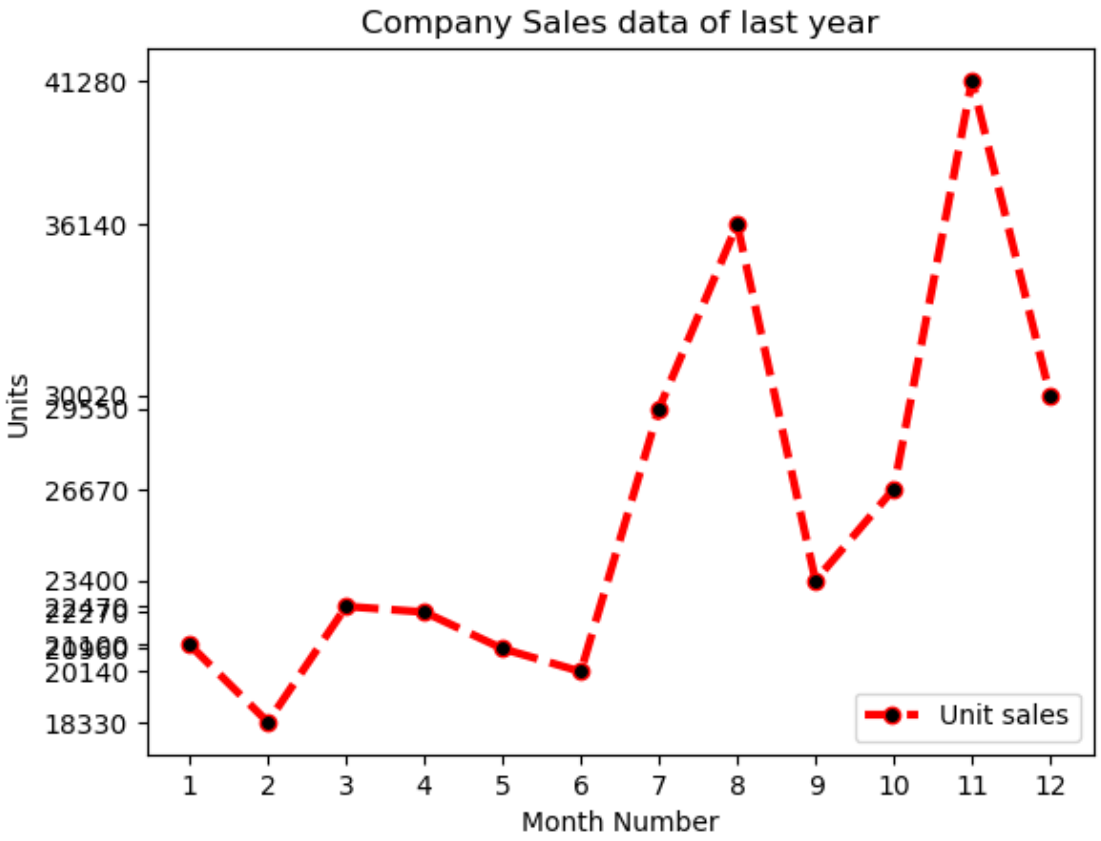
[44]:
```
df = pd.read_csv("/home/cvl-5aiml-a2/Downloads/company_sales_data.csv")
profitList = df ['total_profit'].tolist()
monthList  = df ['month_number'].tolist()
plt.plot(monthList, profitList, label = 'Month-wise Profit data of last year')
plt.xlabel('Month number')
plt.ylabel('Profit')
plt.xticks(monthList)
plt.title('Company profit per month')
plt.yticks([100000, 200000, 300000, 400000, 500000])
plt.show()
```

4. Use the company sales dataset csv file, get total profit of all months and show line plot with the following Style properties. Generated line plot must include following Style properties: –
5. Line Style dotted and Line-color should be red
6. Show legend at the lower right location.
7. X label name = Month Number
8. Y label name = Sold units number
9. Add a circle marker.
10. Line marker color as read
11. Line width should be 3

[43]:
```
unitList = df ['total_units'].tolist()
monthList  = df ['month_number'].tolist()

plt.plot(monthList, unitList, label = 'Unit sales',
      color='r', marker='o', markerfacecolor='k',
      linestyle='--', linewidth=3)
```

```
plt.xlabel('Month Number')
plt.ylabel('Units')
plt.legend(loc='lower right')
plt.title('Company Sales data of last year')
plt.xticks(monthList)
plt.yticks(unitList)
plt.show()
```



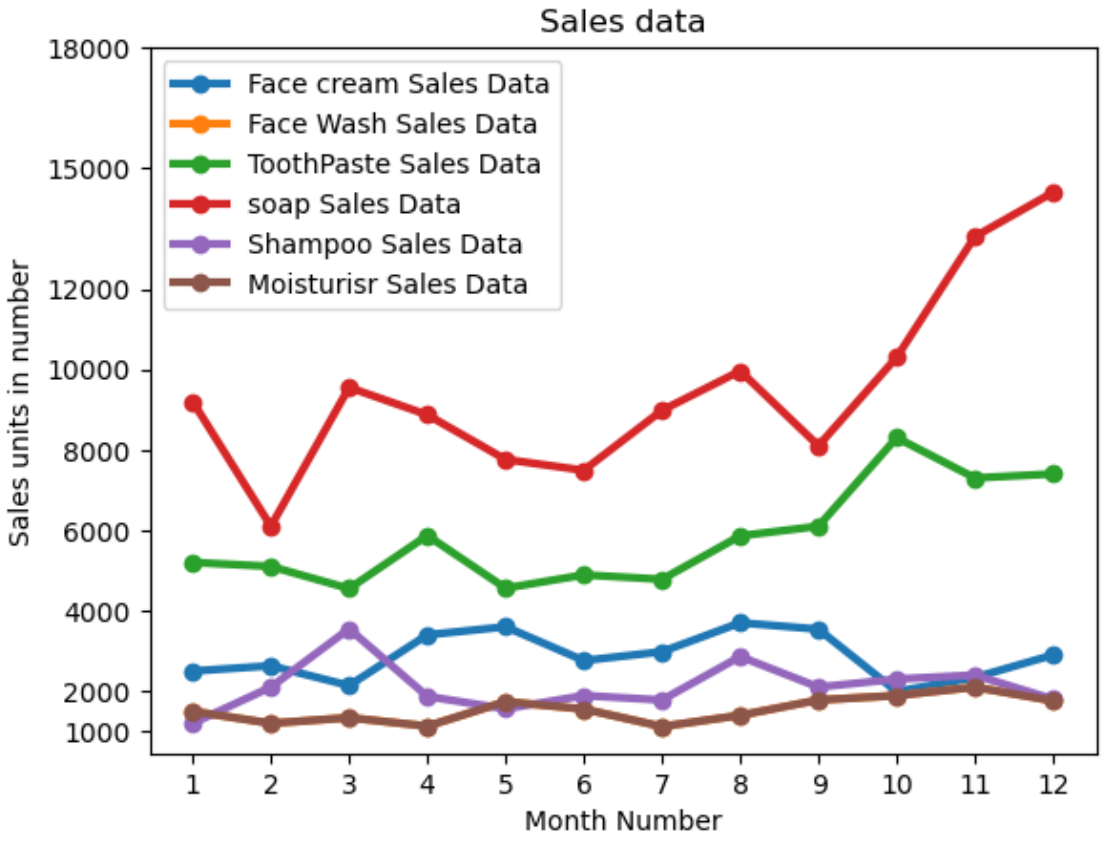
Additional Questions

Use the company sales dataset csv file, read all product sales data and show it using a multiline plot. Display the number of units sold per month or each product using multiline plots.

[50]:
```
monthList  = df ['month_number'].tolist()
faceCremSalesData   = df ['facecream'].tolist()
faceWashSalesData   = df ['facewash'].tolist()
toothPasteSalesData = df ['toothpaste'].tolist()
bathingsoapSalesData   = df ['bathingsoap'].tolist()
shampooSalesData   = df ['shampoo'].tolist()
moisturizerSalesData = df ['moisturizer'].tolist()
```

```
plt.plot(monthList, faceCremSalesData,   label = 'Face cream Sales Data',␣
 ↪marker='o', linewidth=3)
plt.plot(monthList, faceWashSalesData,   label = 'Face Wash Sales Data', ␣
 ↪marker='o', linewidth=3)
plt.plot(monthList, toothPasteSalesData, label = 'ToothPaste Sales Data',␣
 ↪marker='o', linewidth=3)
plt.plot(monthList, bathingsoapSalesData, label = 'soap Sales Data',␣
 ↪marker='o', linewidth=3)
plt.plot(monthList, shampooSalesData, label = 'Shampoo Sales Data', marker='o',␣
 ↪linewidth=3)
plt.plot(monthList, moisturizerSalesData, label = 'Moisturisr Sales Data',␣
 ↪marker='o', linewidth=3)

plt.xlabel('Month Number')
plt.ylabel('Sales units in number')
plt.legend(loc='upper left')
plt.xticks(monthList)
plt.yticks([1000, 2000, 4000, 6000, 8000, 10000, 12000, 15000, 18000])
plt.title('Sales data')
plt.show()
```



Use the company sales dataset csv file, calculate total sale data for last year for each product and show it using a Pie chart.

[51]:
```
labels = ['FaceCream', 'FaseWash', 'ToothPaste', '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()
```