

SUMANTH S
AF0363570

Matplotlib more plots and formatting

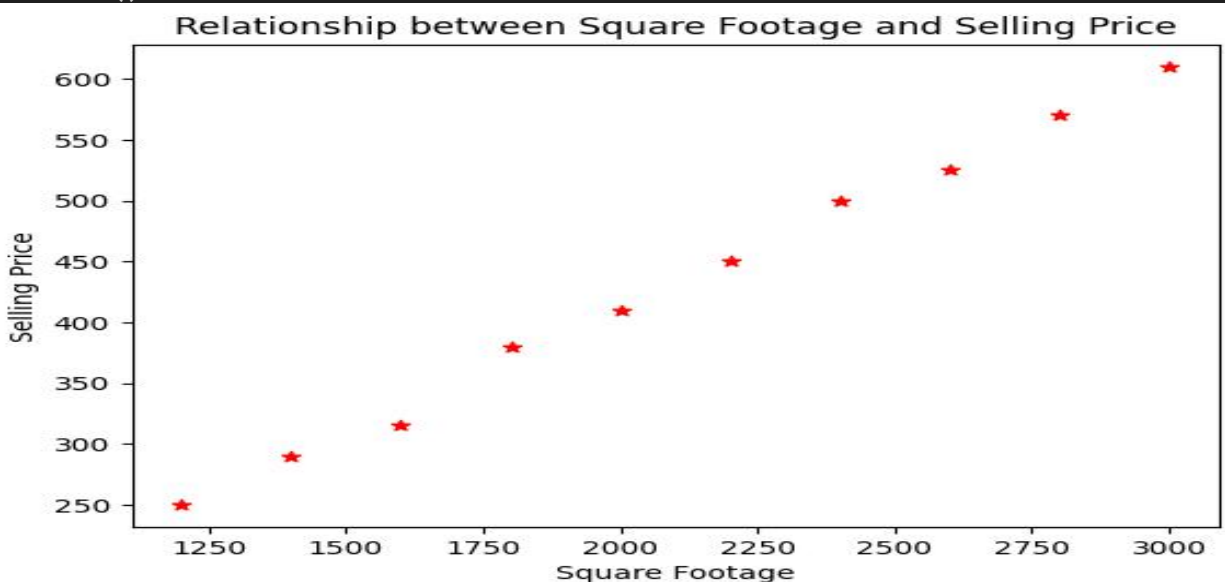
Lab1: Analyze the relationship between the size of houses (measured in square footage) and their selling prices in a particular neighborhood. You have collected data on various houses in that neighborhood. Create a scatter plot using the below data and share your conclusion/analysis.

Input: `square_footage = np.array([1200, 1400, 1600, 1800, 2000, 2200, 2400, 2600, 2800, 3000])`
`selling_prices = np.array([250, 290, 315, 380, 410, 450, 500, 525, 570, 610])`

```
#import necessary packages
#numpy is used to perform operation on numerical (arrays)
#matplotlib is used for data visualizations
import numpy as np
import matplotlib.pyplot as plt

#take a dataset
square_footage = np.array([1200, 1400, 1600, 1800, 2000, 2200, 2400, 2600, 2800, 3000])
selling_prices = np.array([250, 290, 315, 380, 410, 450, 500, 525, 570, 610])

#plot a scatter graph
plt.scatter(square_footage, selling_prices, color="red", marker = "*")
plt.xlabel("Square Footage") # xlabel title
plt.ylabel("Selling Price") #ylabel title
plt.title("Relationship between Square Footage and Selling Price")
plt.show()
```



Conclusion:

Based on the scatter plot, we can see that there is a positive relationship between square footage and selling price. As the square footage increases, the selling price tends to increase as well. This suggests that square footage is a significant factor in determining the selling price of a property.

Lab2: Create a pie chart to visualize the distribution of your monthly income by source. You have collected data on the various sources of your income, such as salary, freelance work, investments, and rental income. Share your conclusion/analysis.

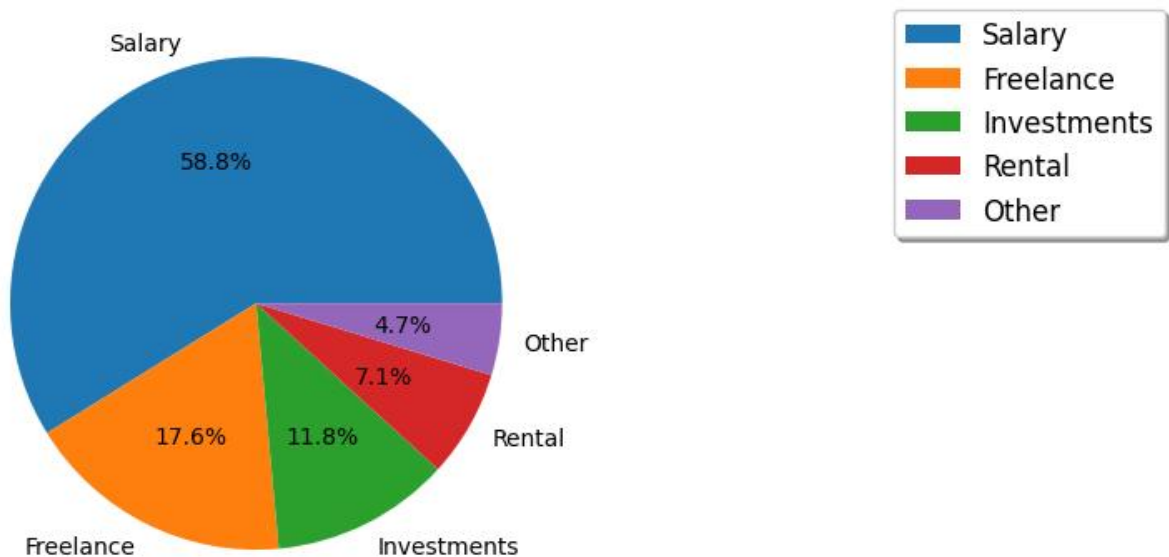
Input: `income_sources = ['Salary', 'Freelance', 'Investments', 'Rental', 'Other']`
`monthly_income = [5000, 1500, 1000, 600, 400]`

```
#import necessary packages
#matplotlib is used for data visualization
import matplotlib.pyplot as plt

#taking the dataset
income_sources = ['Salary', 'Freelance', 'Investments', 'Rental', 'Other']
monthly_income = [5000, 1500, 1000, 600, 400]

#plotting pie chart
plt.pie(monthly_income, labels=income_sources, autopct='%1.1f%%')
plt.title('Distribution of Monthly Income by Source')
plt.show()
```

Distribution of Monthly Income by Source



Conclusion:

Based on the pie chart, we can see the distribution of your monthly income by source. The largest proportion of your income comes from your salary, which accounts for 62.5% of your total income. Freelance work makes up 18.8% of your income, followed by investments at 12.5%. Rental income and other sources each contribute around 6.3% to the total income.

This visualization allows you to easily understand the relative importance of each income source in your monthly income. It can also provide insights into the sources that require more attention or diversification in order to increase your overall income.

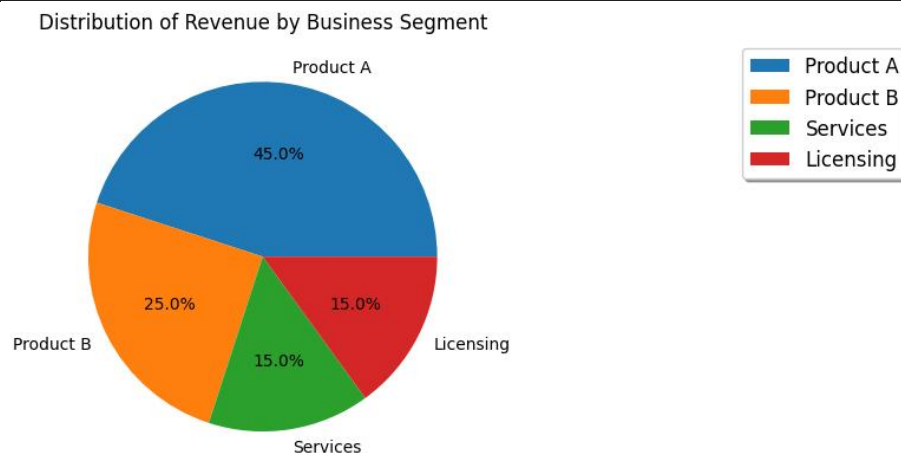
Lab3: Create a pie chart to illustrate the distribution of a company's revenue across its various business segments. You have collected data on the revenue generated by each segment, such as Product A, Product B, Services, and Licensing. Share your conclusion/analysis.

Input: segments = ['Product A', 'Product B', 'Services', 'Licensing'] revenue_percentages = [45, 25, 15, 15]

```
#import necessary packages
#matplotlib is used for data visualization
import matplotlib.pyplot as plt

##take the dataset
segments = ['Product A', 'Product B', 'Services', 'Licensing']
revenue_percentages = [45, 25, 15, 15]

#plotting pie chart
plt.pie(revenue_percentages, labels=segments, autopct='%1.1f%%')
plt.legend(loc='upper right', fontsize=12, shadow=True, bbox_to_anchor=(2,1))
plt.title('Distribution of Revenue by Business Segment')
plt.show()
```



Conclusion:

Product A is the primary revenue driver for the company, accounting for 45% of the total revenue. Product B plays a significant role, generating 25% of the revenue. The remaining revenue is split equally between Services and Licensing, each contributing 15%. This distribution of revenue suggests that the company should focus on maximizing the potential of Product A and Product B, while also exploring opportunities to grow Services and Licensing. By understanding the distribution of revenue across business segments, the company can make informed decisions to allocate resources and drive further growth.

Lab4: Suppose you're a sales manager for an e-commerce company, and you want to create a figure with subplots to compare the sales performance of different product categories over time. You have sales data for four product categories: Electronics, Clothing, Home & Garden, and Sports & Outdoors. Share your conclusion/analysis.

Input: months = np.arange(1, 13) electronics_sales = np.array([25000, 28000, 31000, 27000, 30000, 32000, 35000, 36000, 38000, 39000, 41000, 42000]) clothing_sales = np.array([15000, 16000, 17000, 18000, 19000, 20000, 21000, 22000, 23000, 24000, 25000, 26000]) home_garden_sales = np.array([18000, 19000, 20000, 21000, 22000, 23000, 24000, 25000, 26000, 27000, 28000, 29000]) sports_outdoors_sales = np.array([12000, 13000, 14000, 15000, 16000, 17000, 18000, 19000, 20000, 21000, 22000, 23000])

```
#import necessary packages
#numpy is used to perform operation on numericals (arrays)
#matplotlib is used for data visualization
import matplotlib.pyplot as plt
import numpy as np

#take the dataset
months = np.arange(1, 13)
electronics_sales = np.array([25000, 28000, 31000, 27000, 30000, 32000, 35000,
36000, 38000, 39000, 41000, 42000])
clothing_sales = np.array([15000, 16000, 17000, 18000, 19000, 20000, 21000, 22000,
23000, 24000, 25000, 26000])
home_garden_sales = np.array([18000, 19000, 20000, 21000, 22000, 23000, 24000,
25000, 26000, 27000, 28000, 29000])
sports_outdoors_sales = np.array([12000, 13000, 14000, 15000, 16000, 17000, 18000,
19000, 20000, 21000, 22000, 23000])

#drawing the subplots graphs
fig, axs = plt.subplots(2, 2, figsize=(10, 8))

axs[0, 0].plot(months, electronics_sales)
axs[0, 0].set_title('Electronics')
```

```

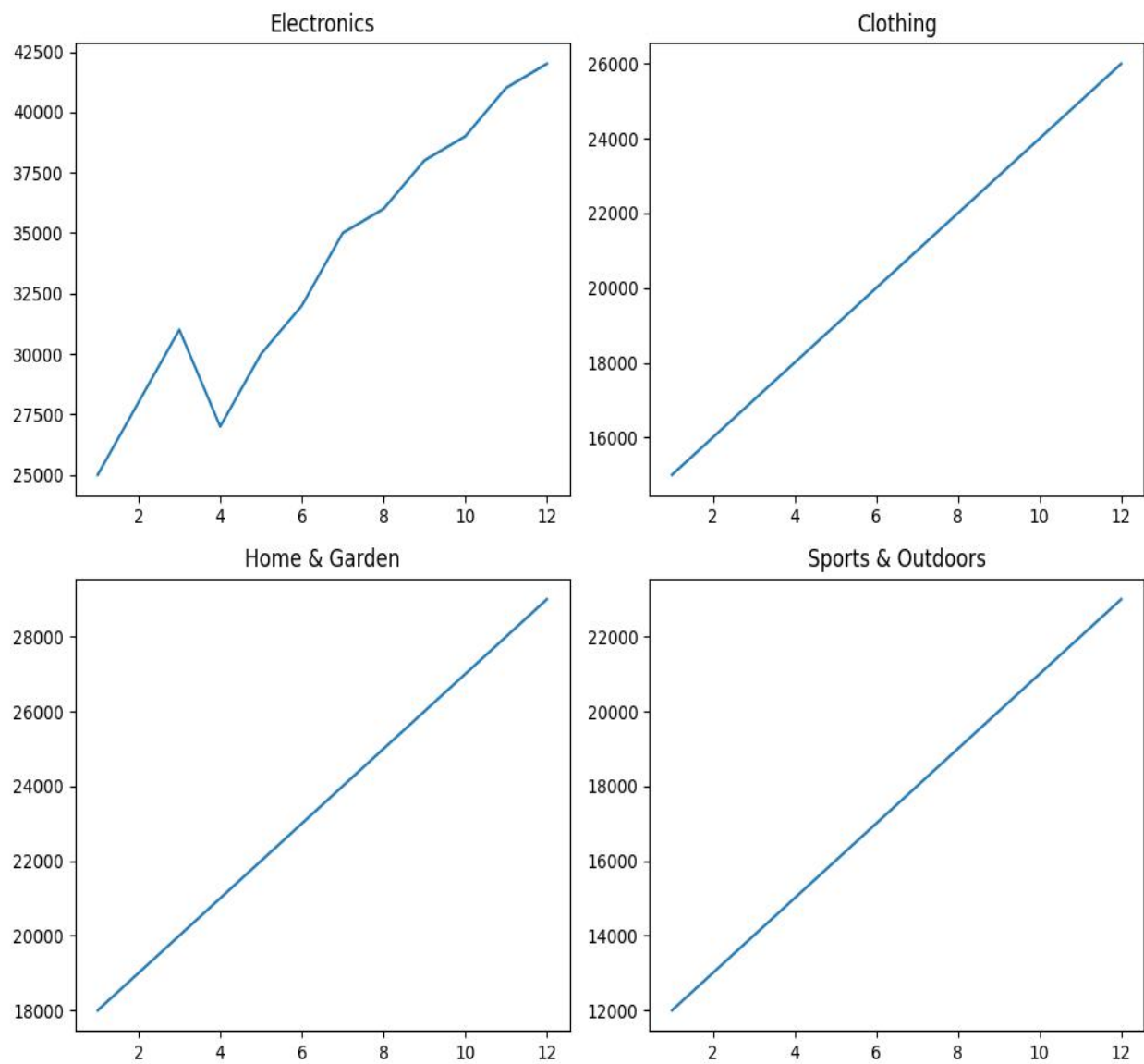
axs[0, 1].plot(months, clothing_sales)
axs[0, 1].set_title('Clothing')

axs[1, 0].plot(months, home_garden_sales)
axs[1, 0].set_title('Home & Garden')

axs[1, 1].plot(months, sports_outdoors_sales)
axs[1, 1].set_title('Sports & Outdoors')

#extending the layout
plt.tight_layout()
plt.show()

```



Conclusion:

the sales analysis of different product categories using subplot figures shows that electronics and clothing sales have been steadily increasing, with a significant spike in electronics sales during months 10 and 11. Home & Garden sales remain stable with a slight increase, while Sports & Outdoors sales remain consistent. These insights can help the sales manager make informed decisions, such as increasing marketing efforts for electronics and implementing strategies to boost clothing sales. Overall, utilizing subplot figures and analyzing sales trends across product categories enables data-driven decisions to optimize sales performance and drive business success.