Nama : Dimas Ramadhani

NIM : 4611422126

Mentor : Rusnandi Farhan

Section 1 : Library Matplotlib

Buatlah diagram yang diperintahkan pada masing-masing soal menggunakan library matplotlib.pyplot. Akan mendapat nilai lebih apabila melengkapi syntax python dengan menambahkan judul chart, penamaan sumbu x dan y, serta memunculkan legend pada chart apabila diperlukan.

```
import matplotlib.pyplot as plt
```

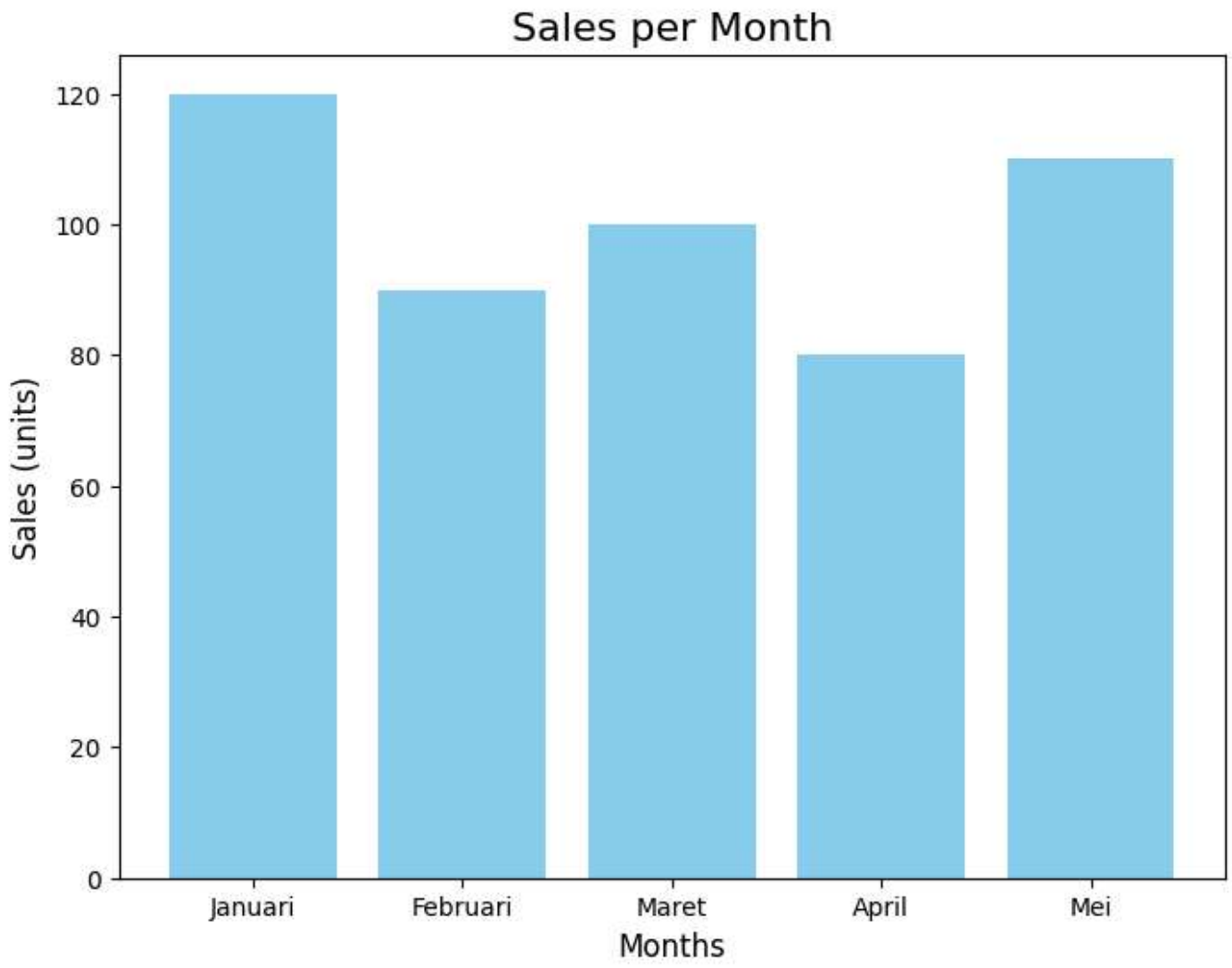
Tugas 1

```
# Data for the bar chart
months = ['Januari', 'Februari', 'Maret', 'April', 'Mei']
sales = [120, 90, 100, 80, 110]


# Creating the bar chart
plt.figure(figsize=(8, 6))
plt.bar(months, sales, color='skyblue')

# Adding title and labels
plt.title('Sales per Month', fontsize=16)
plt.xlabel('Months', fontsize=12)
plt.ylabel('Sales (units)', fontsize=12)

# Display the chart
plt.show()
```



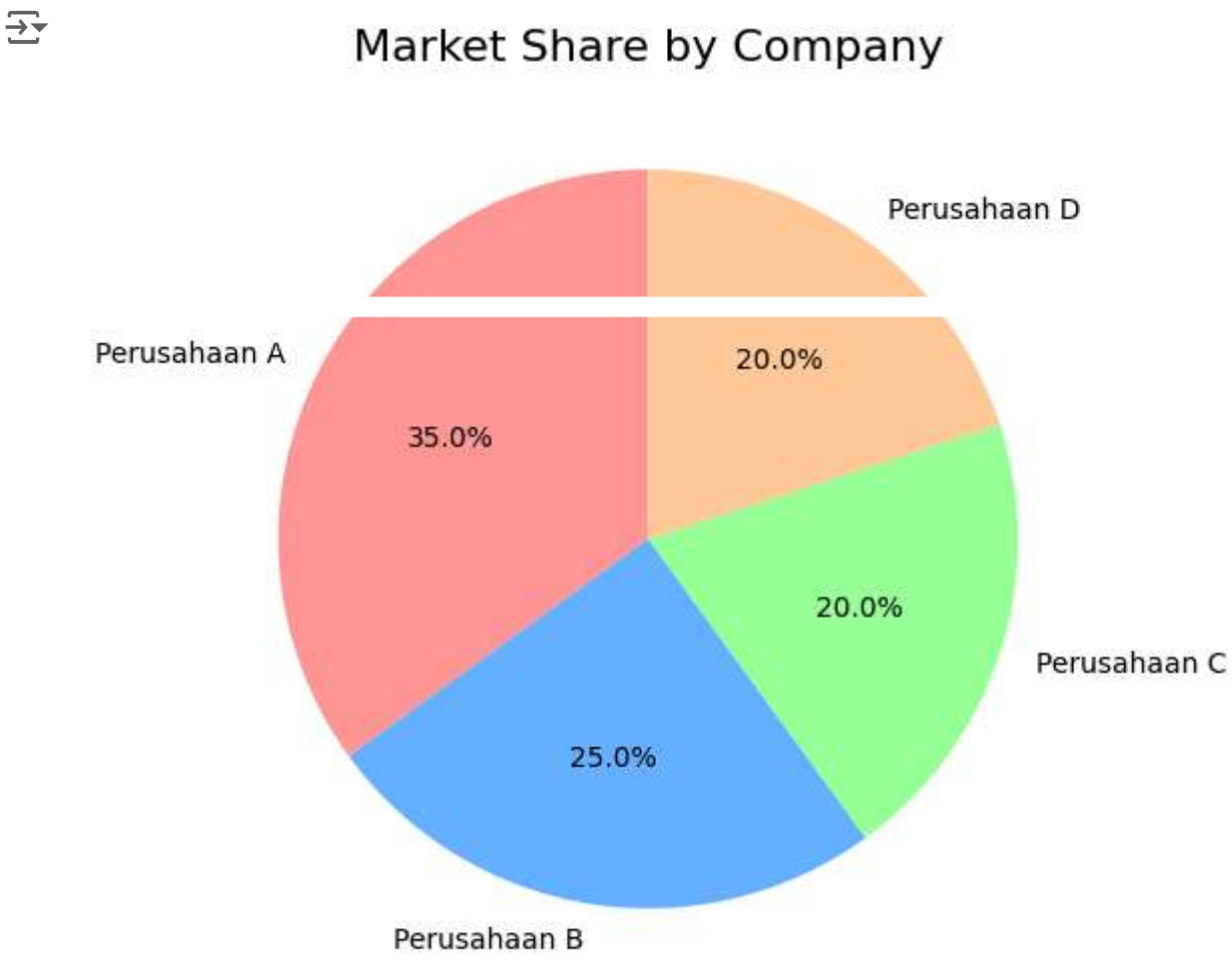
Tugas 2

```
# Data for the pie chart
companies = ['Perusahaan A', 'Perusahaan B', 'Perusahaan C', 'Perusahaan D']
market_share = [35, 25, 20, 20]


# Creating the pie chart
plt.figure(figsize=(8, 6))
plt.pie(market_share, labels=companies, autopct='%1.1f%%', startangle=90, colors=['#ff9999',

# Adding title
plt.title('Market Share by Company', fontsize=16)
```

```
# Display the chart
plt.show()
```



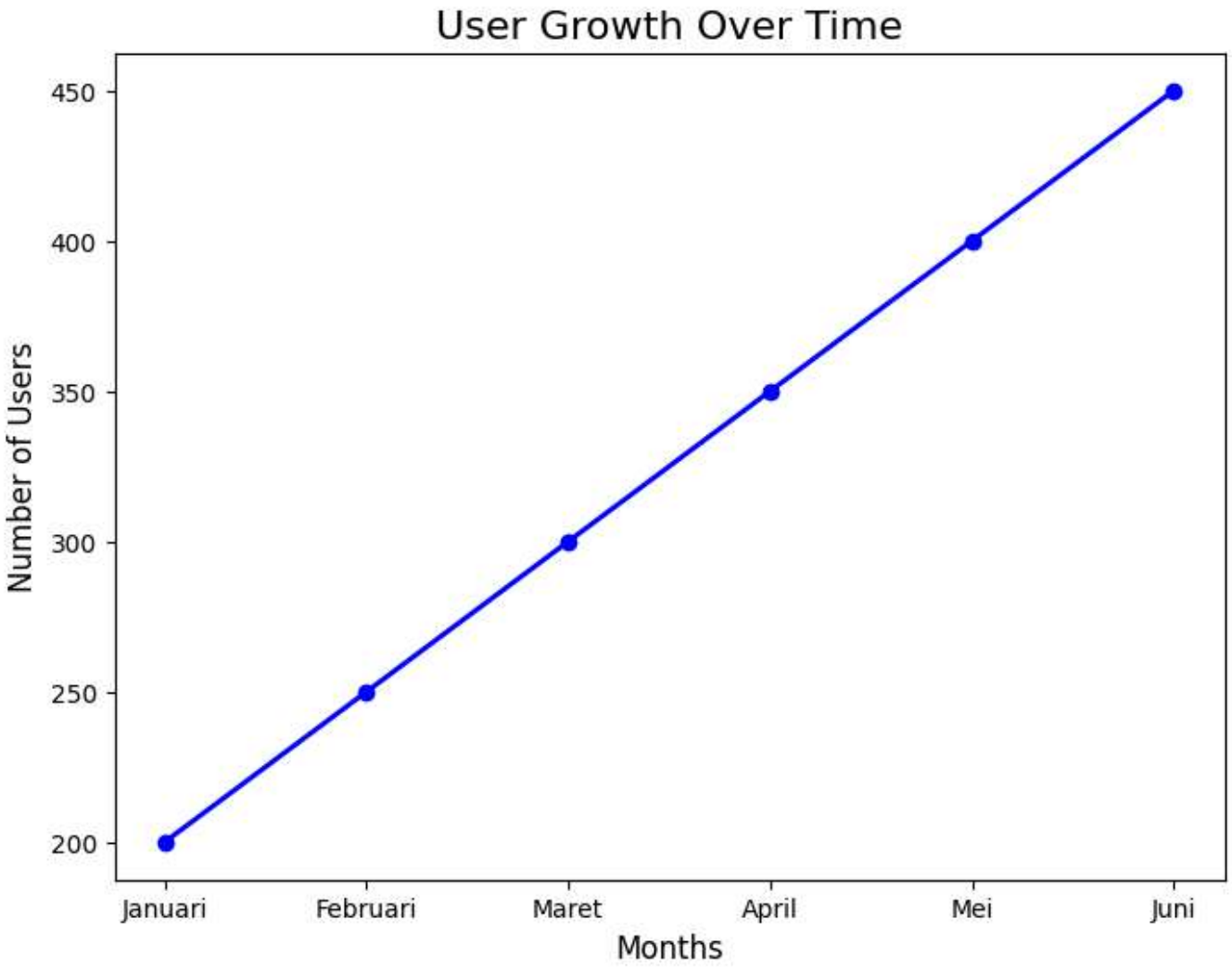
Tugas 3

```
# Data for the line chart
months_line = ['Januari', 'Februari', 'Maret', 'April', 'Mei', 'Juni']
users = [200, 250, 300, 350, 400, 450]


# Creating the line chart
plt.figure(figsize=(8, 6))
plt.plot(months_line, users, marker='o', color='b', linestyle='-', linewidth=2)

# Adding title and labels
plt.title('User Growth Over Time', fontsize=16)
plt.xlabel('Months', fontsize=12)
plt.ylabel('Number of Users', fontsize=12)
```

```
# Display the chart
plt.show()
```



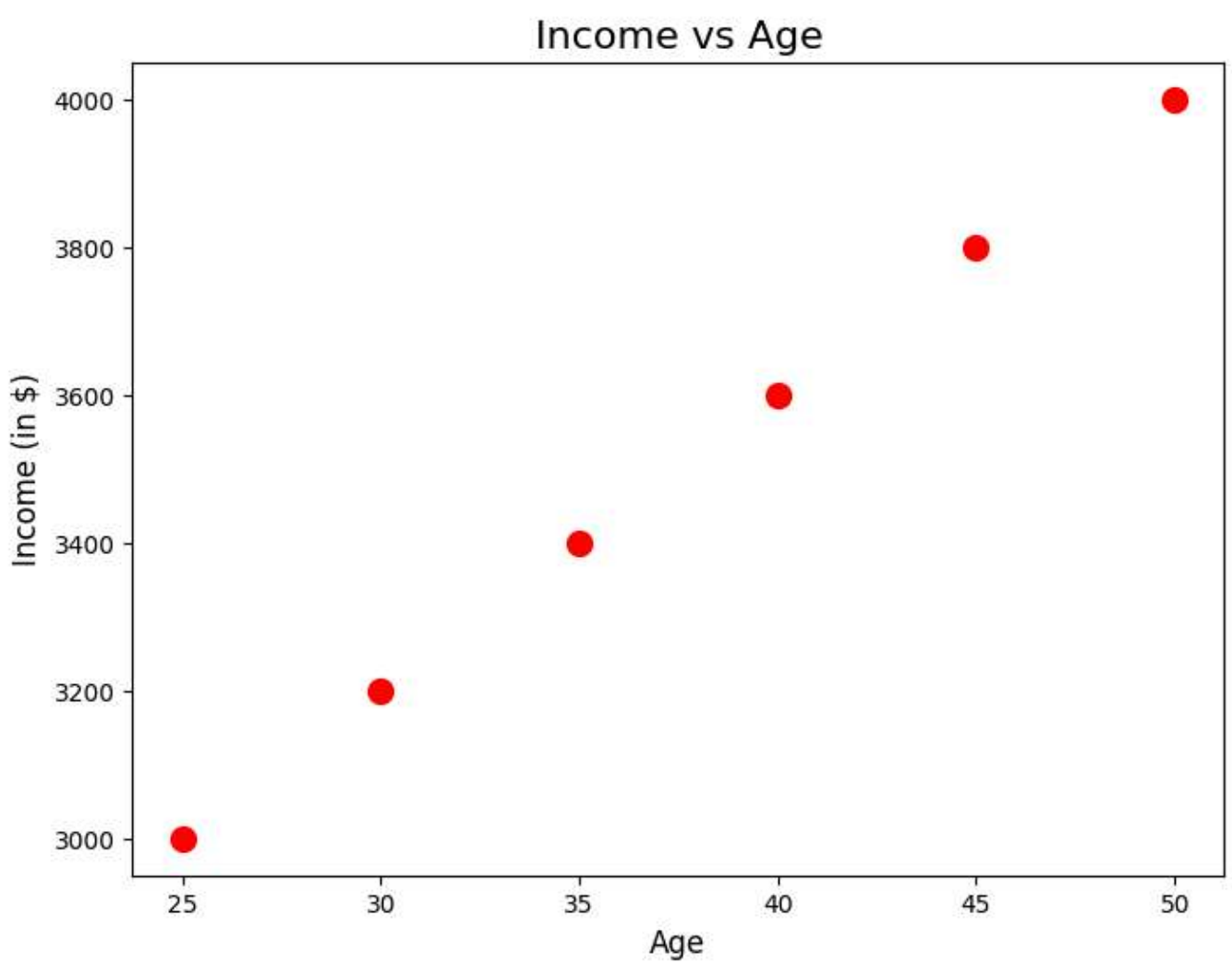
Tugas 4

```
# Data for the scatter plot
ages = [25, 30, 35, 40, 45, 50]
incomes = [3000, 3200, 3400, 3600, 3800, 4000]


# Creating the scatter plot
plt.figure(figsize=(8, 6))
plt.scatter(ages, incomes, color='red', s=100)
```

```
# Adding title and labels
plt.title('Income vs Age', fontsize=16)
plt.xlabel('Age', fontsize=12)
plt.ylabel('Income (in $)', fontsize=12)

# Display the chart
plt.show()
```



Tugas 5

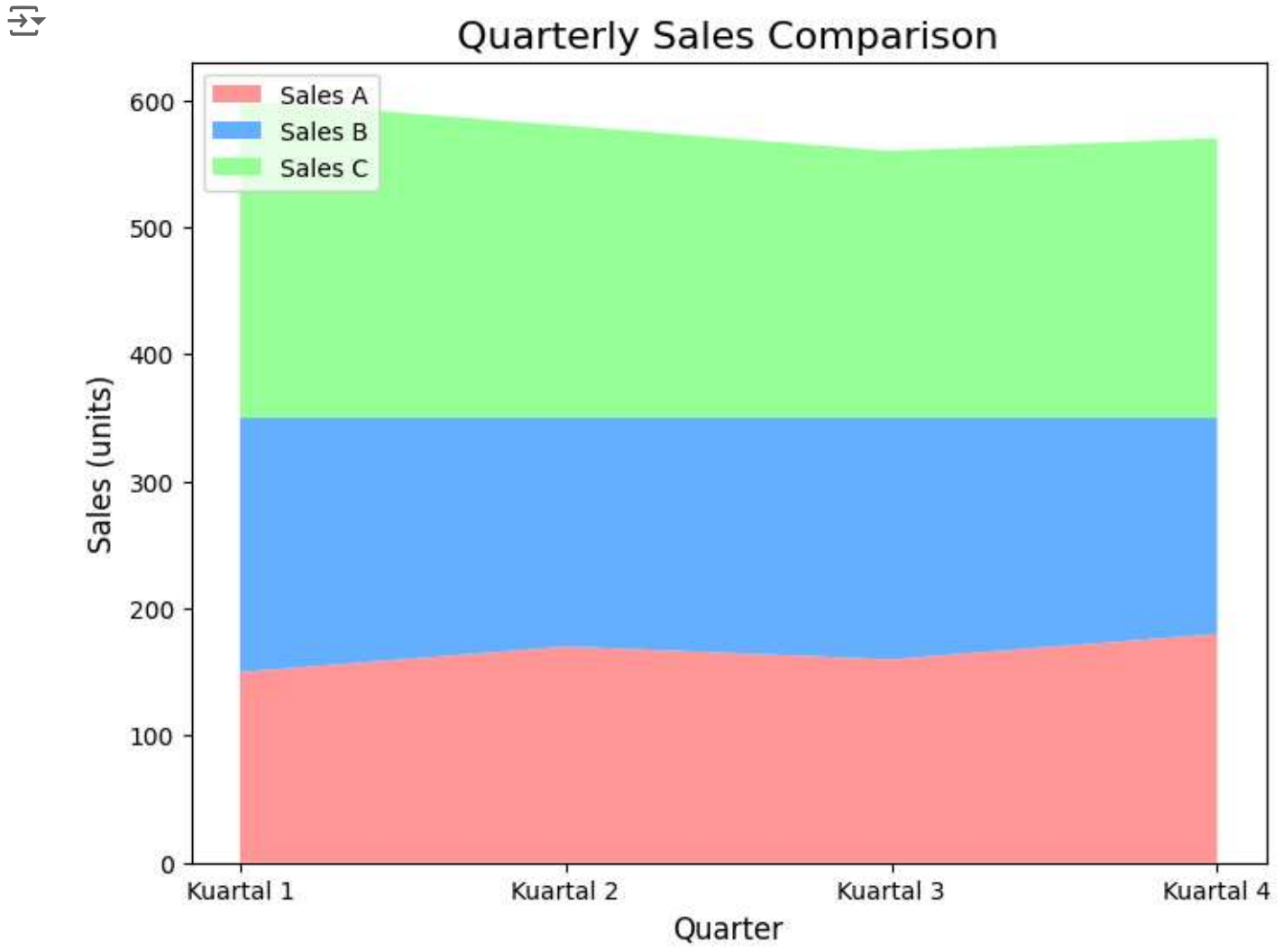
```
# Data for the stacked area chart
quarters = ['Kuartal 1', 'Kuartal 2', 'Kuartal 3', 'Kuartal 4']
sales_A = [150, 170, 160, 180]
sales_B = [200, 180, 190, 170]
sales_C = [250, 230, 210, 220]
```

```
# Creating the stacked area chart
plt.figure(figsize=(8, 6))
plt.stackplot(quarters, sales_A, sales_B, sales_C, labels=['Sales A', 'Sales B', 'Sales C'],

# Adding title and labels
plt.title('Quarterly Sales Comparison', fontsize=16)
plt.xlabel('Quarter', fontsize=12)
plt.ylabel('Sales (units)', fontsize=12)

# Adding legend
plt.legend(loc='upper left')

# Display the chart
plt.show()
```



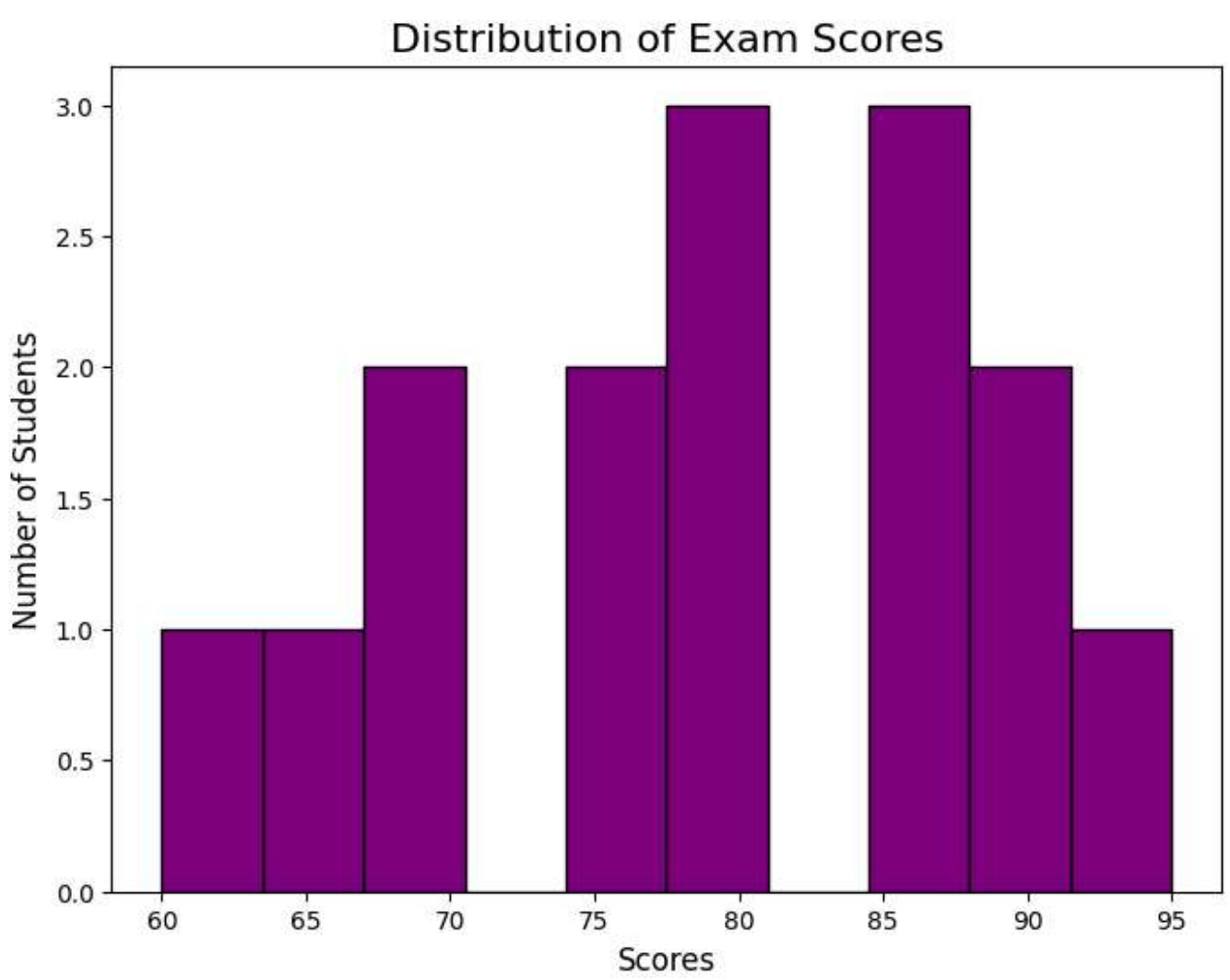
Tugas 6

```
# Data for the histogram
exam_scores = [70, 80, 75, 85, 90, 95, 65, 60, 80, 70, 75, 85, 90, 80, 85]


# Creating the histogram
plt.figure(figsize=(8, 6))
plt.hist(exam_scores, bins=10, color='purple', edgecolor='black')

# Adding title and labels
plt.title('Distribution of Exam Scores', fontsize=16)
plt.xlabel('Scores', fontsize=12)
plt.ylabel('Number of Students', fontsize=12)

# Display the chart
plt.show()
```



Section 2 : Library Pandas

Gunakan library pandas dan juga data yang disediakan untuk menjawab pertanyaan-pertanyaan berikut. Link dataset : https://drive.google.com/file/d/1K7nD0K2DKCK3YIFYgWJ7vvunwP7zSA0D/view?usp=sharing

```
import pandas as pd


# Load dataset dari URL
url = '/content/drive/MyDrive/Celerates Python/Tugas Python 1/Mall_Customers.csv'
data = pd.read_csv(url)
```

No 1

Load data yanng disediakan menggunakan pandas, kemudian print shape dari data tersebut serta munculkan data.head() dari dataframe tersebut

```
# Tampilkan shape dan 5 baris pertama dari dataframe
print("Shape of the data:", data.shape)
print(data.head())
```

```
Shape of the data: (200, 5)
   CustomerID  Gender  Age  Annual Income (k$)  Spending Score (1-100)
0           1    Male   19                  15                      39
1           2    Male   21                  15                      81
2           3  Female   20                  16                       6
3           4  Female   23                  16                      77
4           5  Female   31                  17                      40
```

No 2

Keluarkan informasi dari dataframe melalui fungis .info()

```
# Tampilkan informasi mengenai dataframe
print(data.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 200 entries, 0 to 199
Data columns (total 5 columns):
 #   Column                  Non-Null Count  Dtype
---  ------                  --------------  -----
 0   CustomerID              200 non-null    int64
 1   Gender                  200 non-null    object
 2   Age                     200 non-null    int64
 3   Annual Income (k$)      200 non-null    int64
 4   Spending Score (1-100)  200 non-null    int64
dtypes: int64(4), object(1)
memory usage: 7.9+ KB
None
```

No 3

Munculkan informasi statistik deskriptif dari dataframe

```
# Tampilkan statistik deskriptif
print(data.describe())
```

```
       CustomerID         Age  Annual Income (k$)  Spending Score (1-100)
count  200.000000  200.000000          200.000000              200.000000
mean   100.500000   38.850000           60.560000               50.200000
std     57.879185   13.969007           26.264721               25.823522
min      1.000000   18.000000           15.000000                1.000000
25%     50.750000   28.750000           41.500000               34.750000
50%    100.500000   36.000000           61.500000               50.000000
75%    150.250000   49.000000           78.000000               73.000000
max    200.000000   70.000000          137.000000               99.000000
```

No 4

Gunakan fungsi “group by” untuk menunjukkan rata-rata spending score berdasarkan gender

```
# Tampilkan rata-rata spending score berdasarkan gender
print(data.groupby('Gender')['Spending Score (1-100)'].mean())
```

```
Gender
Female    51.526786
Male      48.511364
Name: Spending Score (1-100), dtype: float64
```

No 5

Buat kolom baru bernama “annual_income_idr” yang isinya adalah nilai pada kolom “Annual Income (k$)” dikali dengan 17000 sebagai representasi konversi nilai Annual Income dalam bentuk rupiah.

```
# Tambahkan kolom baru untuk Annual Income dalam IDR (Rupiah)
data['annual_income_idr'] = data['Annual Income (k$)'] * 17000

# Tampilkan beberapa baris untuk memastikan kolom baru berhasil ditambahkan
print(data[['Annual Income (k$)', 'annual_income_idr']].head())
```

```
   Annual Income (k$)  annual_income_idr
0                  15             255000
1                  15             255000
2                  16             272000
3                  16             272000
4                  17             289000
```