

```
#Georgia Sugisandhea_535230080_Kelas C
```

```
import numpy as np #mengimport library numpy dengan variable np
import pandas as pd #mengimport library pandas dengan variable pd
from sklearn import datasets #mengimport module datasets dari sklearn
import matplotlib.pyplot as plt #mengimport library matplotlib.pyplot
dengan variable plt
```

```
iris = pd.read_csv('Iris.csv') #Membaca file Iris.csv dengan pd dan
memasukkannya ke variable iris
```

```
iris.head() #Mengeluarkan 5 baris teratas dan nama kolomnya dari iris
```

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm
Species					
0	1	5.1	3.5	1.4	0.2
setosa					
1	2	4.9	3.0	1.4	0.2
setosa					
2	3	4.7	3.2	1.3	0.2
setosa					
3	4	4.6	3.1	1.5	0.2
setosa					
4	5	5.0	3.6	1.4	0.2
setosa					

```
df = iris.drop(['Id', 'Species'], axis=1) #membuang kolom id dan
species dan memasukkannya ke variable df
```

```
df.head() #menampilkan 5 baris teratas dan nama kolomnya dari df,
hasil pembuangan id dan species dari data
```

	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm
0	5.1	3.5	1.4	0.2
1	4.9	3.0	1.4	0.2
2	4.7	3.2	1.3	0.2
3	4.6	3.1	1.5	0.2
4	5.0	3.6	1.4	0.2

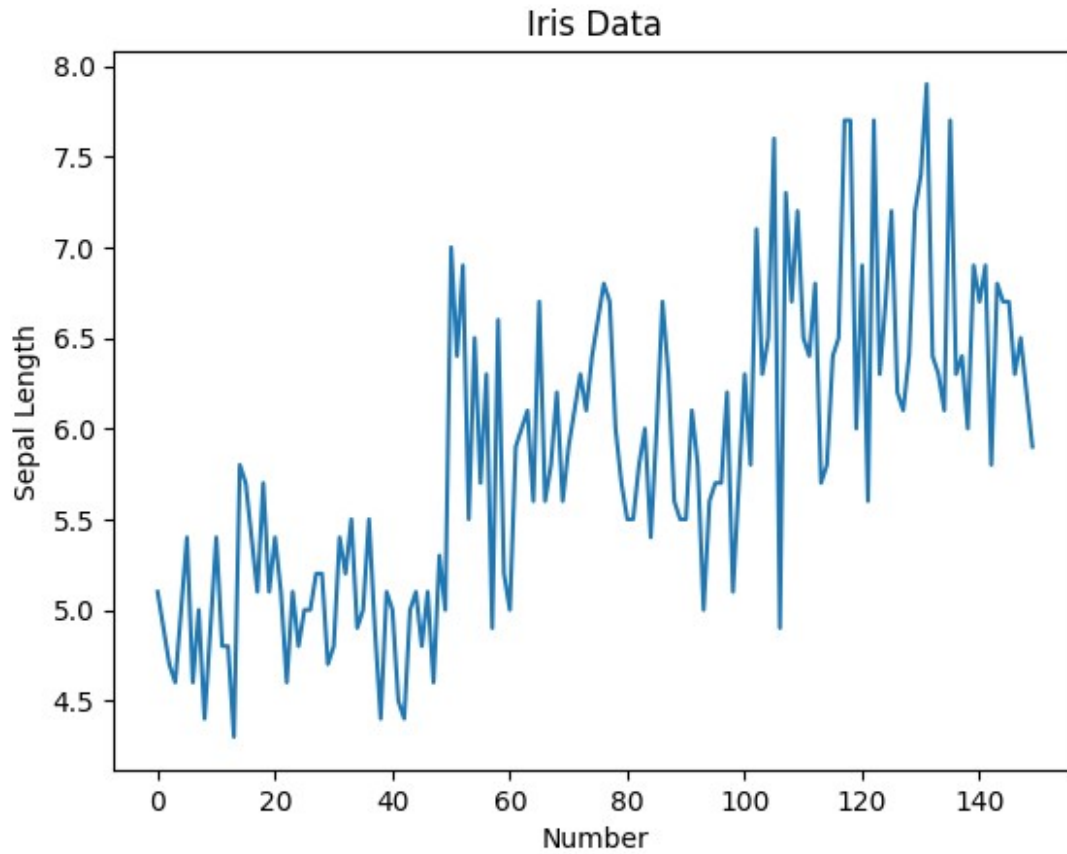
```
plt.plot(df['SepalLengthCm']) #Menampilkan grafik garis dengan semua
data di kolom SepalLengthCm di df
```

```
plt.xlabel('Number') #Membuat label sumbu x bernama Number
```

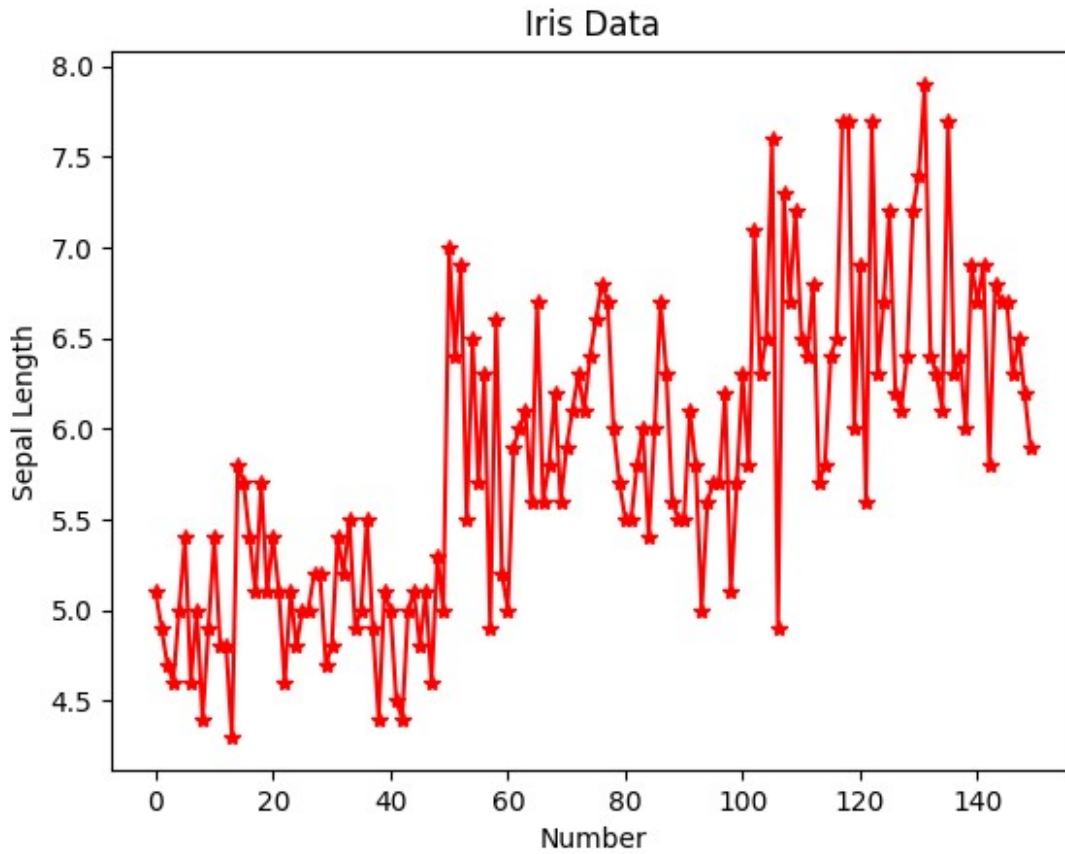
```
plt.ylabel('Sepal Length') #Membuat label sumbu y bernama Sepal Length
```

```
plt.title('Iris Data') #Membuat judul dari grafiknya Iris Data
```

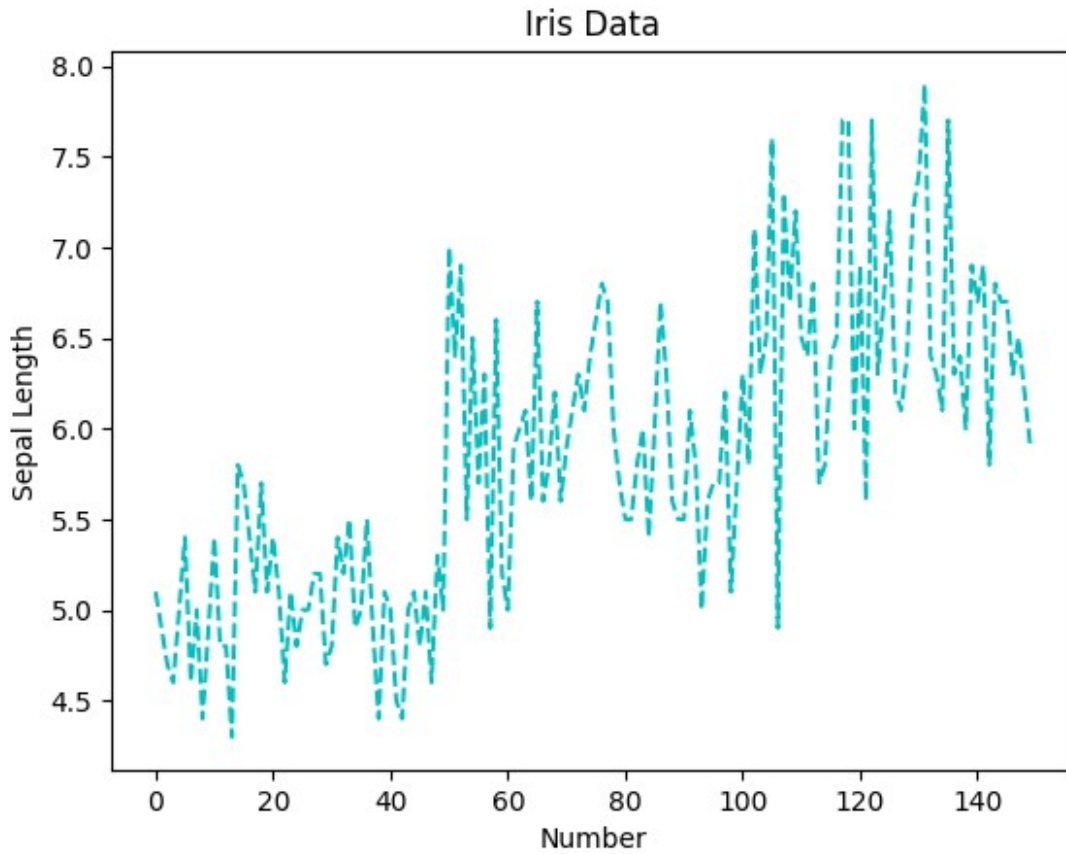
```
plt.show() #Menampilkan hasil grafik garis yang dibuat
```



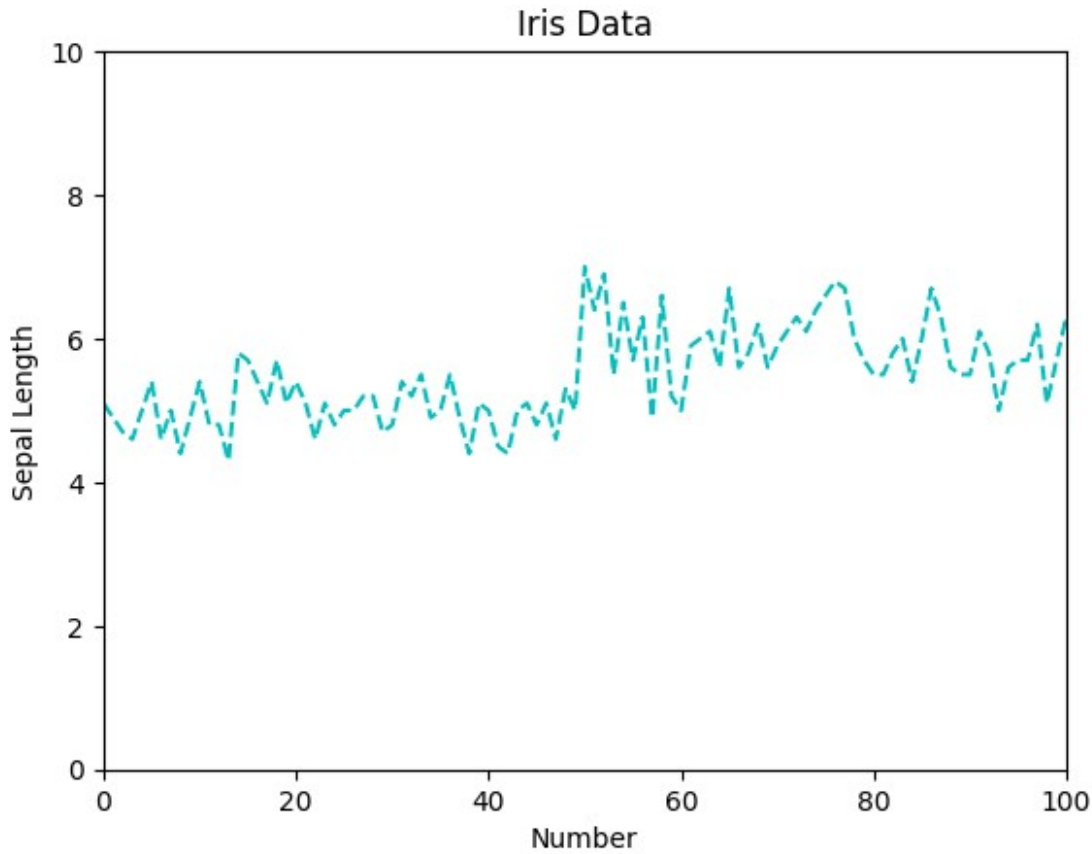
```
plt.plot(df['SepalLengthCm'], color='red', marker='*')  
#Membuat grafik garis dari semua data di kolom SepalLengthCm di df  
#color='red' => Membuat garis pada grafiknya berwarna merah  
#marker='*' => Membuat titik data pada grafik garis *  
plt.xlabel('Number') #Membuat label sumbu x bernama Number  
plt.ylabel('Sepal Length') #Membuat label sumbu y bernama Sepal Length  
plt.title('Iris Data') #Membuat judul dari grafiknya Iris Data  
plt.show() #Menampilkan hasil grafik garis yang dibuat
```



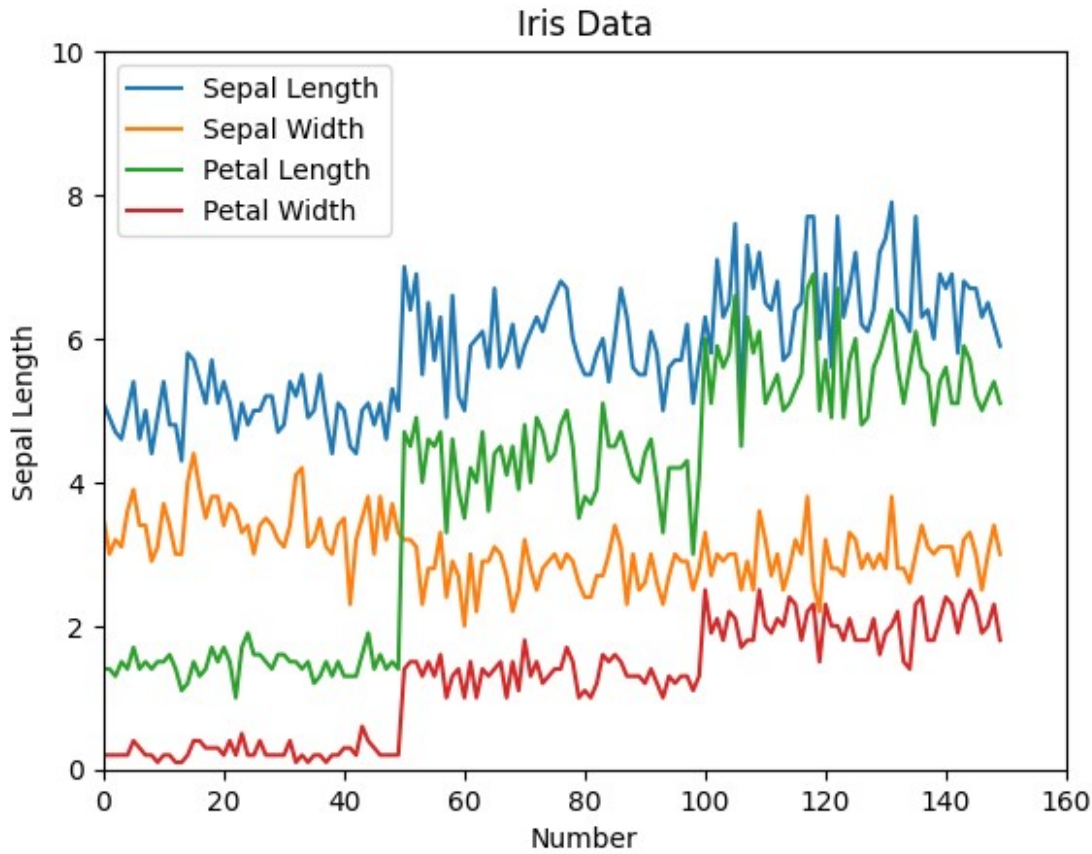
```
plt.plot(df['SepalLengthCm'], '--c')  
#Membuat grafik garis dengan semua data di kolom SepalLengthCm di df  
#Membuat garisnya putus putus dengan '--' dan warnanya cyan dengan c  
plt.xlabel('Number') #Membuat label sumbu x bernama Number  
plt.ylabel('Sepal Length') #Membuat label sumbu y bernama Sepal Length  
plt.title('Iris Data') #Membuat judul dari grafiknya Iris Data  
plt.show() #Menampilkan hasil grafik garis yang dibuat
```



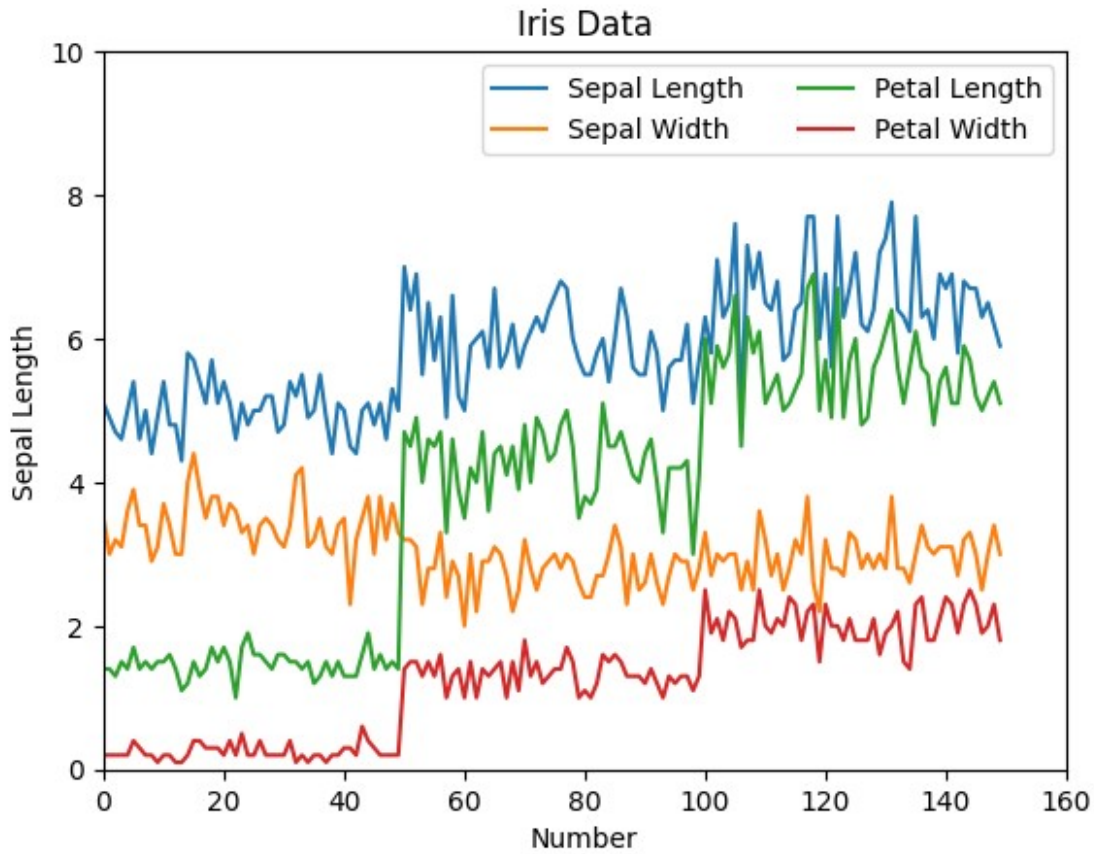
```
plt.plot(df['SepalLengthCm'], '--c')  
#Membuat grafik garis dengan semua data di kolom SepalLengthCm di df  
#Membuat garisnya putus putus dengan '--' dan warnanya cyan dengan c  
plt.xlabel('Number') #Membuat label sumbu x bernama Number  
plt.ylabel('Sepal Length') #Membuat label sumbu y bernama Sepal Length  
plt.title('Iris Data') #Membuat judul dari grafiknya Iris Data  
limits = [0, 100, 0, 10] #Membuat array bernama limits  
plt.axis(limits) #Membuat sumbu x dari 0 sampai 100 dan sumbu y 0  
sampai 10 sesuai dengan array  
plt.show() #Menampilkan hasil grafik garis yang dibuat
```



```
plt.plot(df['SepalLengthCm']) #Membuat grafik garis dengan semua data
                               di kolom SepalLengthCm di df
plt.plot(df['SepalWidthCm']) #Menambahkan grafik data dengan semua
                              data di kolom SepalWidthCm di df
plt.plot(df['PetalLengthCm']) #Menambahkan grafik data dengan semua
                              data di kolom PetalLengthCm di df
plt.plot(df['PetalWidthCm']) #Menambahkan grafik data dengan semua
                              data di kolom PetalWidthCm di df
plt.xlabel('Number') #Membuat label sumbu x bernama Number
plt.ylabel('Sepal Length') #Membuat label sumbu y bernama Sepal Length
plt.title('Iris Data') #Membuat judul dari grafiknya Iris Data
plt.legend(['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal
Width'])
#Membuat penanda akan garis warna apa menandakan data dari mana dengan
urutan sesuai dengan urutan plt.plot yang dimasukkan
#Menamakan tiap garis sesuai dengan nama data yang dimasukkan sesuai
urutannya
limits = [0, 160, 0, 10] #Membuat array bernama limits
plt.axis(limits) #Membuat sumbu x dari 0 sampai 160 dan sumbu y 0
sampai 10 sesuai dengan array
plt.show() #Menampilkan hasil grafik grais yang dibuat
```



```
plt.plot(df['SepalLengthCm']) #Membuat grafik garis dengan semua data
di kolom SepalLengthCm di df
plt.plot(df['SepalWidthCm']) #Menambahkan grafik data dengan semua
data di kolom SepalWidthCm di df
plt.plot(df['PetalLengthCm']) #Menambahkan grafik data dengan semua
data di kolom PetalLengthCm di df
plt.plot(df['PetalWidthCm']) #Menambahkan grafik data dengan semua
data di kolom PetalWidthCm di df
plt.xlabel('Number') #Membuat label sumbu x bernama Number
plt.ylabel('Sepal Length') #Membuat label sumbu y bernama Sepal Length
plt.title('Iris Data') #Membuat judul dari grafiknya Iris Data
plt.legend(['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal
Width'], loc='upper right', ncol=2)
#Membuat penanda akan garis warna apa menandakan data dari mana dengan
urutan sesuai dengan urutan plt.plot yang dimasukkan
#Menamakan tiap garis sesuai dengan nama data yang dimasukkan sesuai
urutannya
#Mengganti lokasi penanda jadi di atas kanan dengan loc='upper right'
#Mengganti barisan menjadi 2 kolom dengan ncol=2
limits = [0, 160, 0, 10] #Membuat array bernama limits
plt.axis(limits) #Membuat sumbu x dari 0 sampai 160 dan sumbu y 0
sampai 10 sesuai dengan array
plt.show() #Menampilkan hasil grafik grais yang dibuat
```

```
data = df.mean().tolist() #Membuat list hasil perhitungan mean dari
semua data di masing masing kolom di df dan memasukkan ke variable
data
data #Menampilkan isi data
```

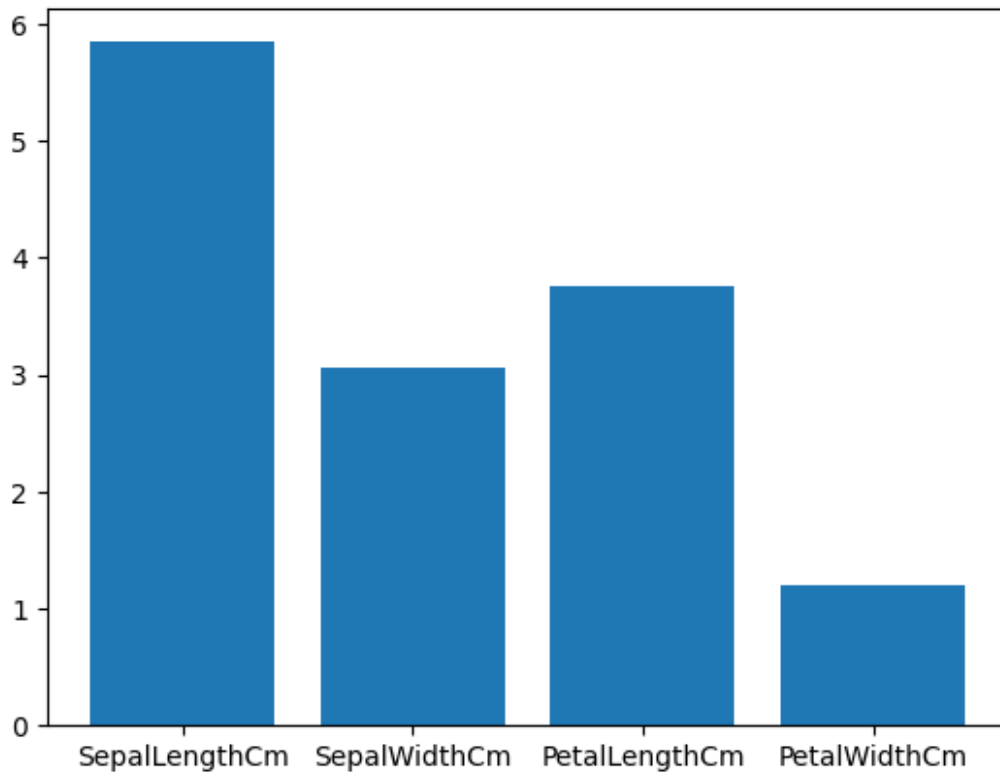
```
[5.843333333333334, 3.0540000000000003, 3.7586666666666666,
1.1986666666666668]
```

```
labels=df.columns.tolist() #Membuat list nama nama kolom yang ada di
df dan memasukkan ke variable labels
labels #Menampilkan isi labels
```

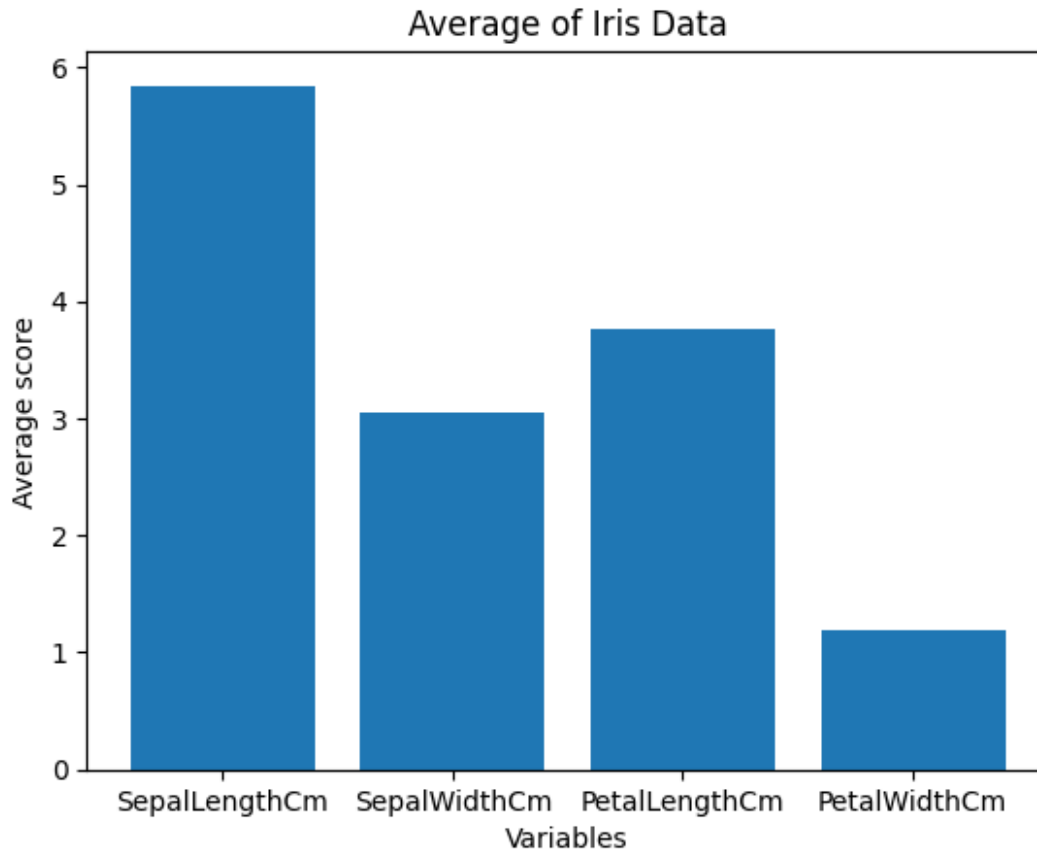
```
['SepalLengthCm', 'SepalWidthCm', 'PetalLengthCm', 'PetalWidthCm']
```

```
plt.bar(labels, data) #Membuat grafik batang dengan list labels dan
data yang sudah kita buat, menampilkan mean dari masing masing kolom
```

```
<BarContainer object of 4 artists>
```



```
plt.bar(labels, data) #Membuat grafik batang dengan list labels dan
data yang sudah kita buat, menampilkan mean dari masing masing kolom
plt.title('Average of Iris Data') #Membuat judul grafik menjadi
'Average of Iris Data'
plt.xlabel('Variables') #Membuat judul dari sumbu x menjadi
'Variables'
plt.ylabel('Average score') #Membuat judul dari sumbu y menjadi
'Average score'
Text(0, 0.5, 'Average score')
```

```

labels = df.columns.tolist() #Membuat list nama nama kolom yang ada di
df dan memasukkan ke variable labels
plt.title('Basic Plot') #Membuat judul dari grafiknya menjadi 'Basic
Plot'
plt.boxplot(df, patch_artist=True, labels = labels)
#Membuat grafik boxplot dengan data yang ada di df
#patch_artist=True -> mengisi kotaknya dengan warna
#labels = labels => Untuk membuat label sesuai dengan list yang sudah
kita buat

#hasil=
#kotak = kuartil 1 sampai 3
#garis = median
#garis atas = sebaran batas atas
#garis bawah = sebaran batas bawah
#bulet bulet = yang offside datanya

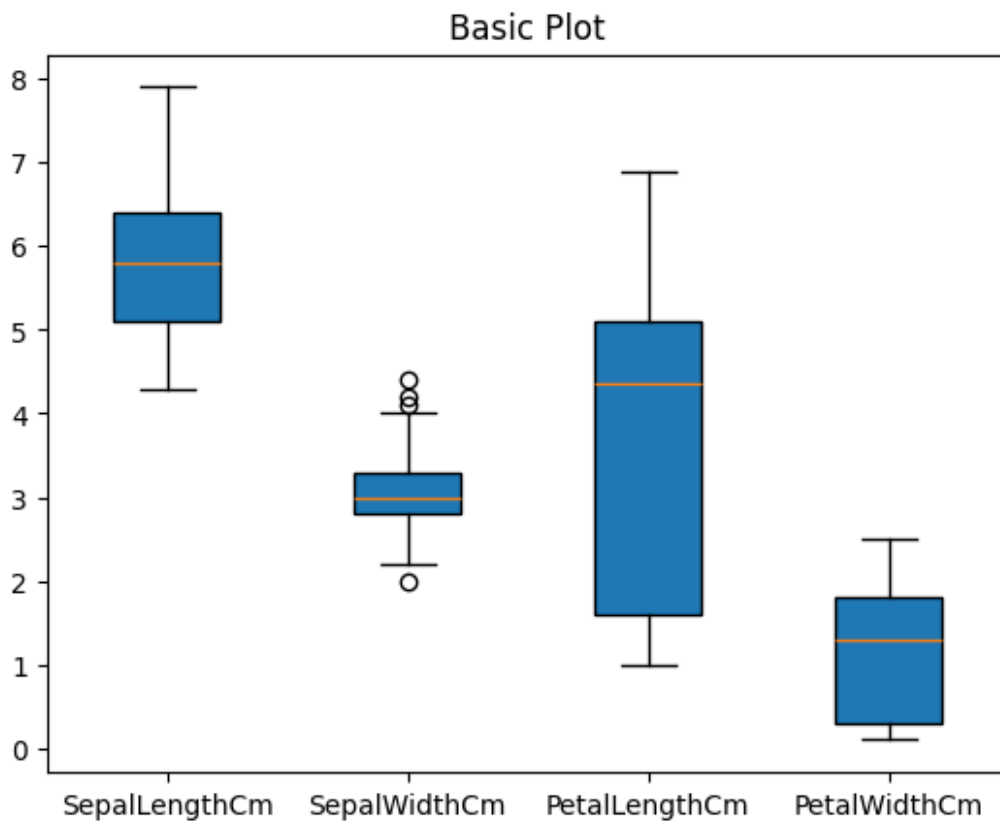
{'whiskers': [<matplotlib.lines.Line2D at 0x1fe295f4470>,
<matplotlib.lines.Line2D at 0x1fe295f6b10>,
<matplotlib.lines.Line2D at 0x1fe295a3f50>,
<matplotlib.lines.Line2D at 0x1fe295a0290>,
<matplotlib.lines.Line2D at 0x1fe295a39b0>,
<matplotlib.lines.Line2D at 0x1fe295a0da0>],

```

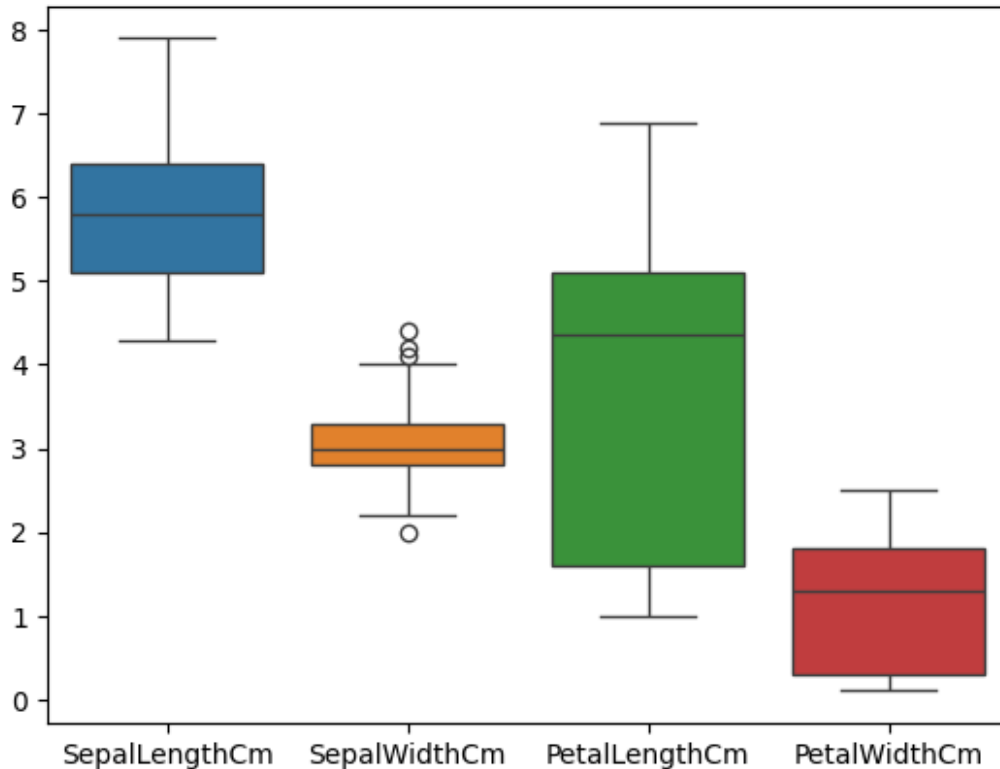
```

<matplotlib.lines.Line2D at 0x1fe295a2930>,
<matplotlib.lines.Line2D at 0x1fe295a3aa0>],
'caps': [<matplotlib.lines.Line2D at 0x1fe295f5550>,
<matplotlib.lines.Line2D at 0x1fe295a3200>,
<matplotlib.lines.Line2D at 0x1fe295a20f0>,
<matplotlib.lines.Line2D at 0x1fe295a0b90>,
<matplotlib.lines.Line2D at 0x1fe295a15e0>,
<matplotlib.lines.Line2D at 0x1fe295a1220>,
<matplotlib.lines.Line2D at 0x1fe295a02c0>,
<matplotlib.lines.Line2D at 0x1fe294b85c0>],
'boxes': [<matplotlib.patches.PathPatch at 0x1fe295f53d0>,
<matplotlib.patches.PathPatch at 0x1fe295alc40>,
<matplotlib.patches.PathPatch at 0x1fe295a0740>,
<matplotlib.patches.PathPatch at 0x1fe295ald60>],
'medians': [<matplotlib.lines.Line2D at 0x1fe295a11f0>,
<matplotlib.lines.Line2D at 0x1fe295a0bf0>,
<matplotlib.lines.Line2D at 0x1fe295a2630>,
<matplotlib.lines.Line2D at 0x1fe295650d0>],
'fliers': [<matplotlib.lines.Line2D at 0x1fe295a2d20>,
<matplotlib.lines.Line2D at 0x1fe295a2fc0>,
<matplotlib.lines.Line2D at 0x1fe295a1d30>,
<matplotlib.lines.Line2D at 0x1fe29566d20>],
'means': []}

```



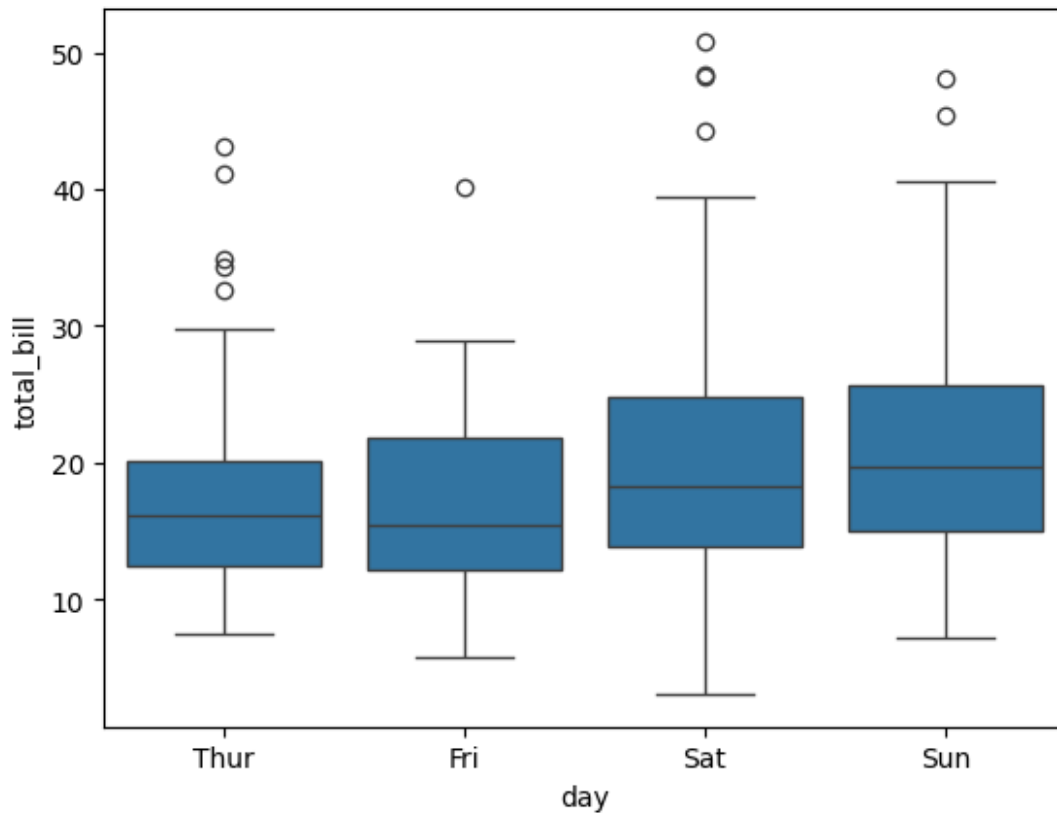
```
import seaborn as sns #import library seaborn sebagai sns
ax = sns.boxplot(data=df) #Membuat grafik boxplot menggunakan seaborn
dengan data yang ada di setiap kolom di df
```



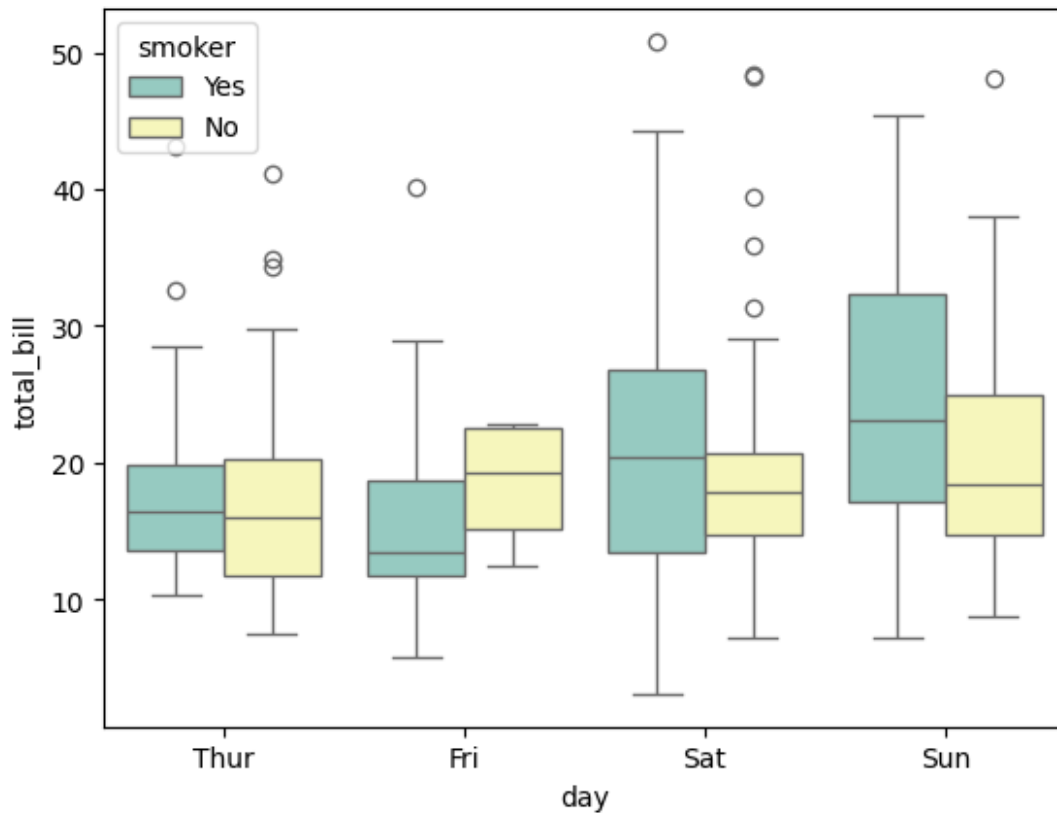
```
tips = sns.load_dataset("tips") #mengambil dataset dari library
seaborn bernama tips dan memasukkan ke variable tips
tips.head() #Menampilkan 5 baris teratas dari data dalam tips
```

	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4

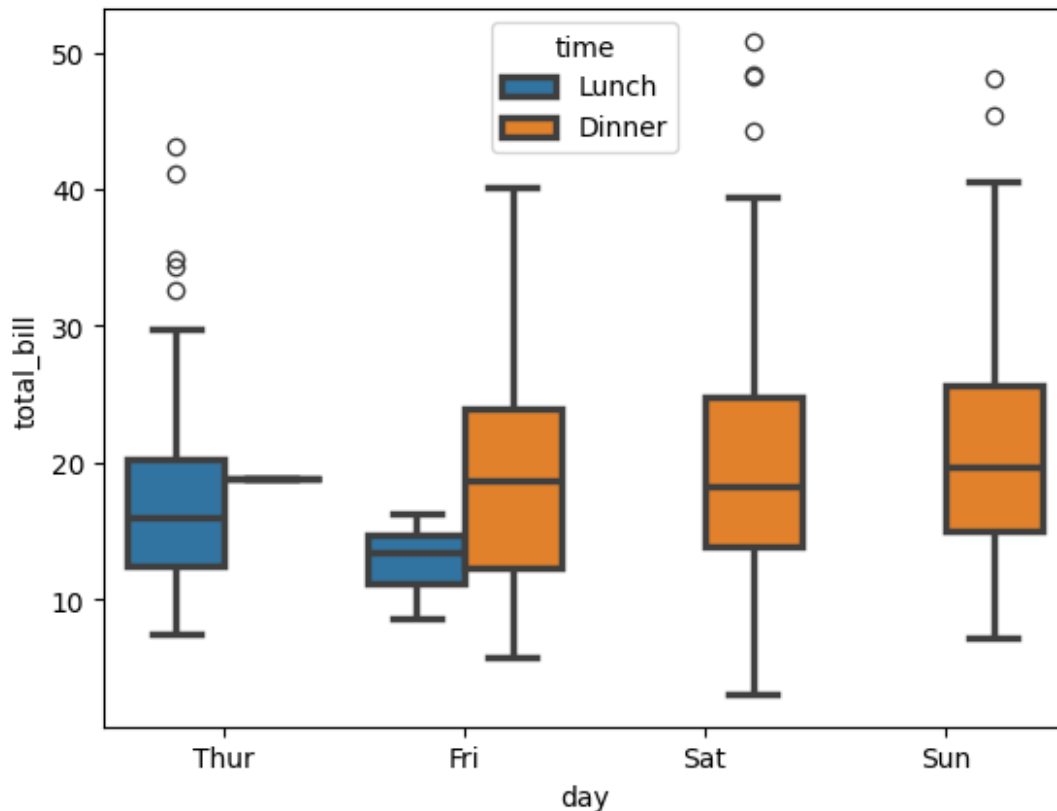
```
ax = sns.boxplot(x="day", y="total_bill", data=tips)
#Membuat boxplot dengan library seaborn dengan data dari variable tips
#Membuat judul sumbu x berupa "day"
#Membuat judul sumbu y berupa "total_bill"
```



```
ax = sns.boxplot(x="day", y="total_bill", hue="smoker", data=tips,
palette="Set3")
#Membuat boxplot dengan data dari variable tips, dikhususkan dari
kolom "smoker" dalam data tersebut
#Membuat judul sumbu x berupa "day"
#Membuat judul sumbu y berupa "total_bill"
#Menggunakan set warna dengan palette = "Set3"
```



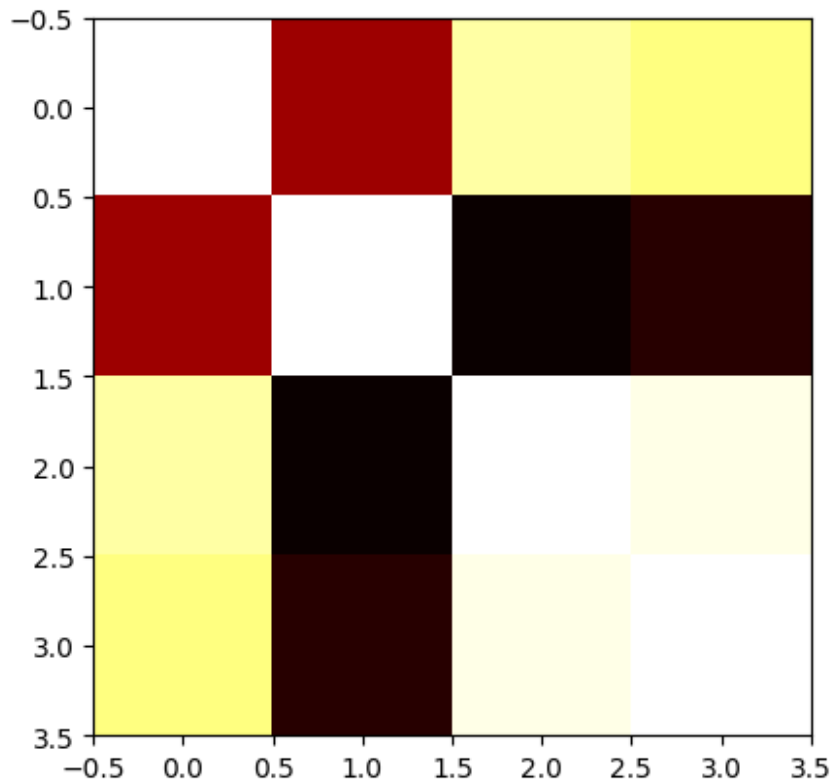
```
ax = sns.boxplot(x="day", y="total_bill", hue="smoker", data=tips,
linewidth=2.5)
#Membuat boxplot dengan data dari variable tips, dikhususkan dari
kolom "smoker" dalam data tersebut
#Membuat judul sumbu x berupa "day"
#Membuat judul sumbu y berupa "total_bill"
#Membuat ketebalan garis pada grafik menjadi 2.5 dengan linewidth=2.5
#Menggunakan set warna dengan palette = "Set3"
```



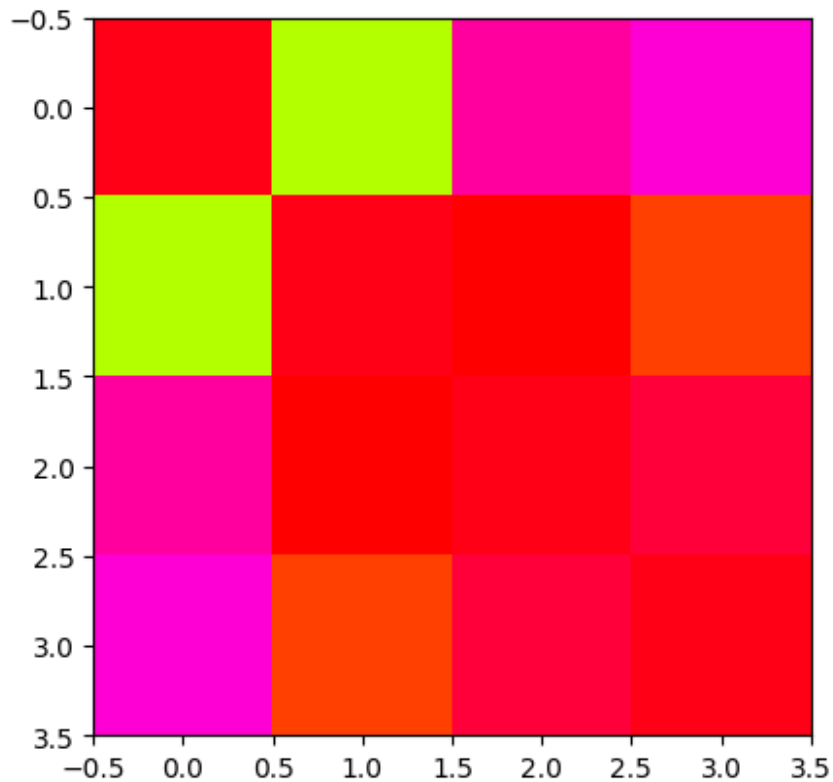
```
c = df.corr() #Menghitung korelasi antara satu kolom dengan semua
kolom dalam data df dan memasukkan ke variable c
c #Menampilkan hasil perhitungan
```

	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm
PetalWidthCm				
SepalLengthCm	1.000000	-0.109369	0.871754	0.817954
SepalWidthCm	-0.109369	1.000000	-0.420516	0.356544
PetalLengthCm	0.871754	-0.420516	1.000000	0.962757
PetalWidthCm	0.817954	-0.356544	0.962757	1.000000

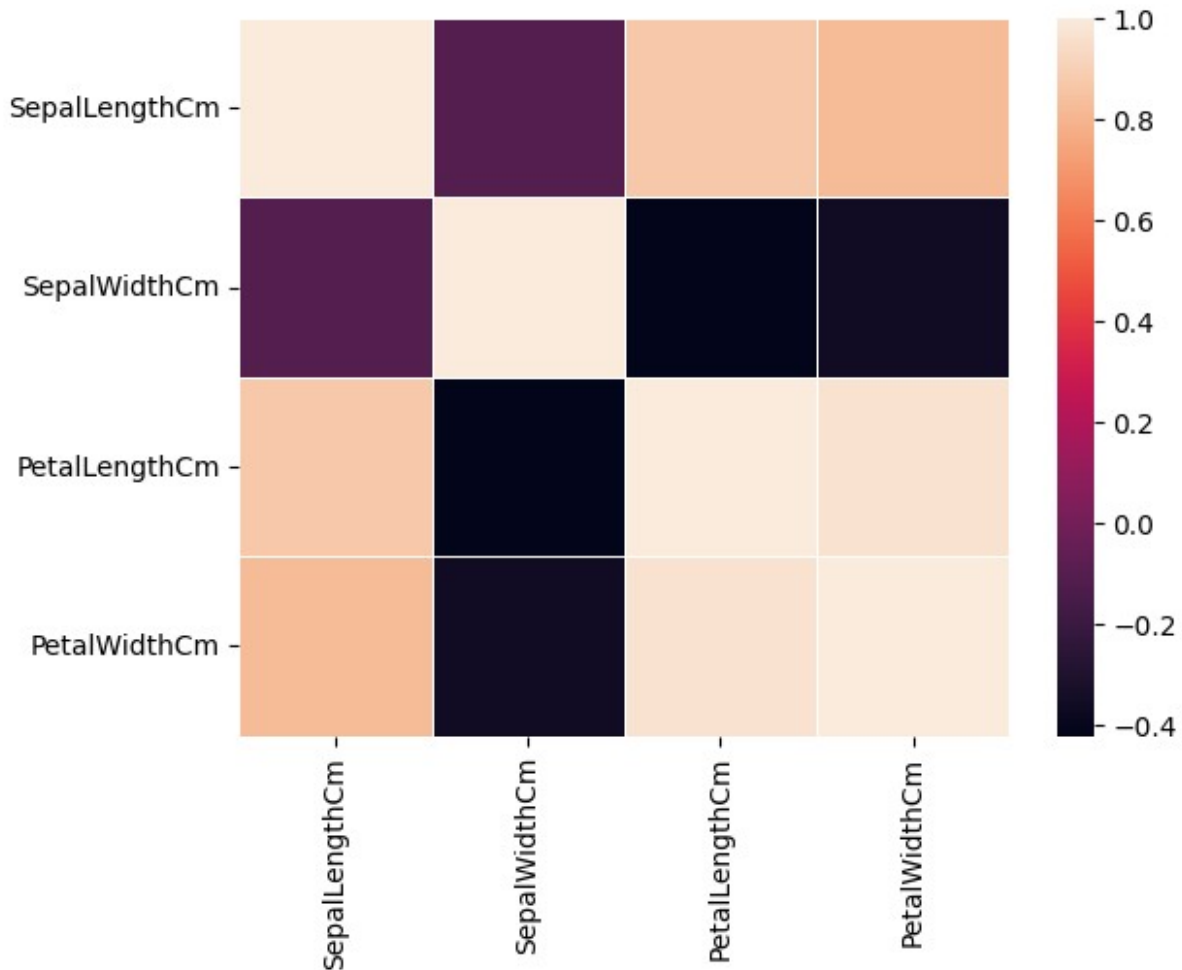
```
plt.imshow(c, cmap='hot')
#Membuat heatmap dengan data c yang telah dihitung, dan menggunakan
set warna hot dengan cmap='hot'
plt.show() #Menampilkan heatmap yang telah dibuat
```



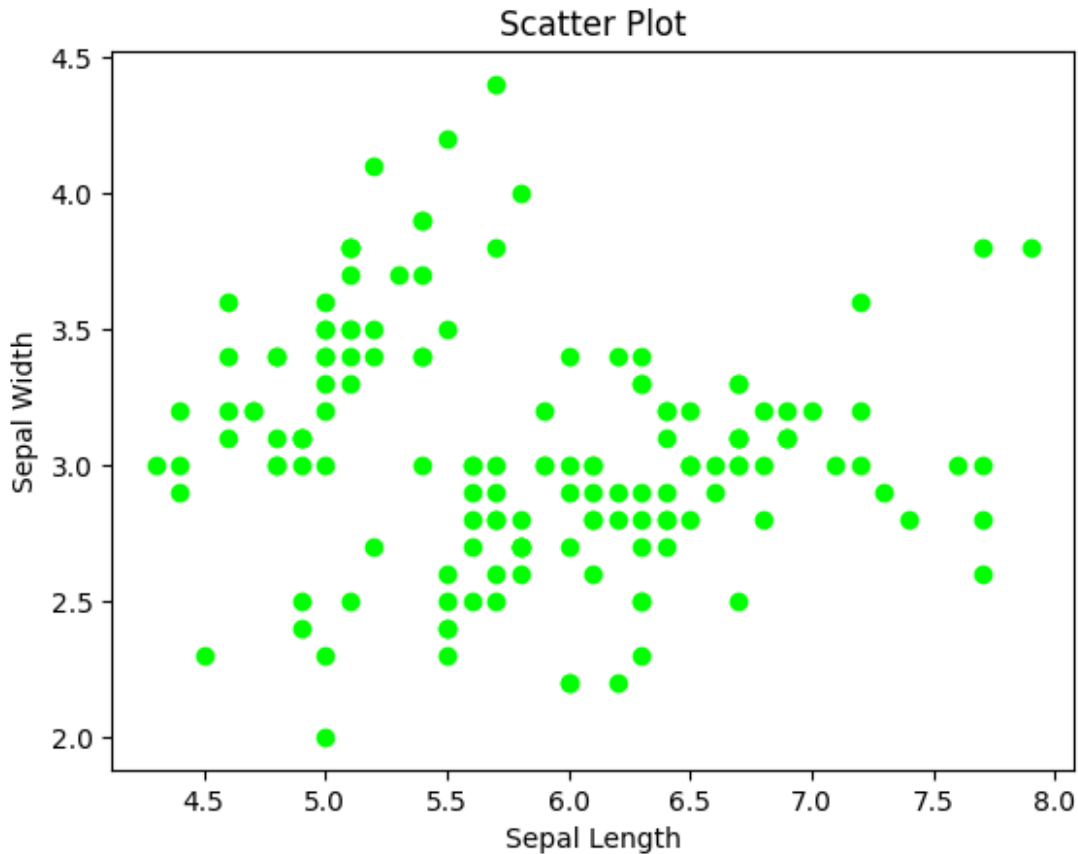
```
plt.imshow(c, cmap='hsv')  
#Membuat heatmap dengan data c yang telah dihitung, dan menggunakan  
#set warna hsv dengan cmap='hsv'  
plt.show() #Menampilkan heatmap yang telah dibuat
```

```
ax = sns.heatmap(c, linewidth=0.5)  
#Membuat heatmap menggunakan library seaborn, dengan data c yang telah  
dihitung sebelumnya  
#Membuat ketebalan garis pada heatmap menjadi 0.5  
plt.show() #Menunjukkan heatmap yang telah dibuat
```



```
plt.scatter(df['SepalLengthCm'], df['SepalWidthCm'], color='lime')
#Membuat Scatter plot antara data SepalLengthCm dan SepalWidthCm dari
data df, dan menggunakan warna lime dengan color='lime'
plt.xlabel('Sepal Length') #Membuat judul sumbu x 'Sepal Length'
plt.ylabel('Sepal Width') #Membuat judul sumbu y 'Sepal Width'
plt.title('Scatter Plot') #Membuat judul grafiknya menjadi 'Scatter
Plot'
Text(0.5, 1.0, 'Scatter Plot')
```



```
fig, ax = plt.subplots(nrows=2, ncols=3, figsize=(15,8))
#Membuat 6 grafik dalam satu tampilan sekaligus, 2 baris menggunakan
nrows=2, dan 3 kolom menggunakan ncols=3
#figsize(15,8) => memspesifikasi ukuran figur menjadi 15 inch kali 8
inch
#Menspesifikasi lokasi grafik setiap data dengan ax[n,n], misalnya
ax[0,0] berarti dia di kolom 0 dan baris 0
#Data yang digunakan sesuai dengan yang disebutkan di setiap baris,
dan warna yang digunakan sesuai dengan color = ' ' warna yang
disebutkan disitu
#{:.2f}".format(c.iloc[0]['SepalWidthCm']) => memspesifikasi format
dari korelasi yang dihitung di data c dengan 2 angka dibelakang
desimal
#Membuat judul sumbu x dengan set_xlabel dan judul sumbu y dengan
set_ylabel
ax[0,0].scatter(df['SepalLengthCm'], df['SepalWidthCm'],
color='magenta')
ax[0,0].title.set_text('SepalLength and SepalWidth\n'+ 'Correlation =
'+ "{:.2f}".format(c.iloc[0]['SepalWidthCm']))
ax[0,0].set_xlabel('Sepal Length')
ax[0,0].set_ylabel('Sepal Width')

ax[0,1].scatter(df['SepalLengthCm'], df['PetalLengthCm'],
```

```

color='green')
ax[0,1].title.set_text('SepalLength and PetalLength\n'+ 'Correlation =
'+ "{:.2f}".format(c.iloc[0]['PetalLengthCm']))
ax[0,1].set_xlabel('Sepal Length')
ax[0,1].set_ylabel('Petal Length')

ax[0,2].scatter(df['SepalLengthCm'], df['PetalWidthCm'], color='blue')
ax[0,2].title.set_text('SepalLength and PetalWidth\n'+ 'Correlation =
'+ "{:.2f}".format(c.iloc[0]['PetalWidthCm']))
ax[0,2].set_xlabel('Sepal Length')
ax[0,2].set_ylabel('Petal Width')

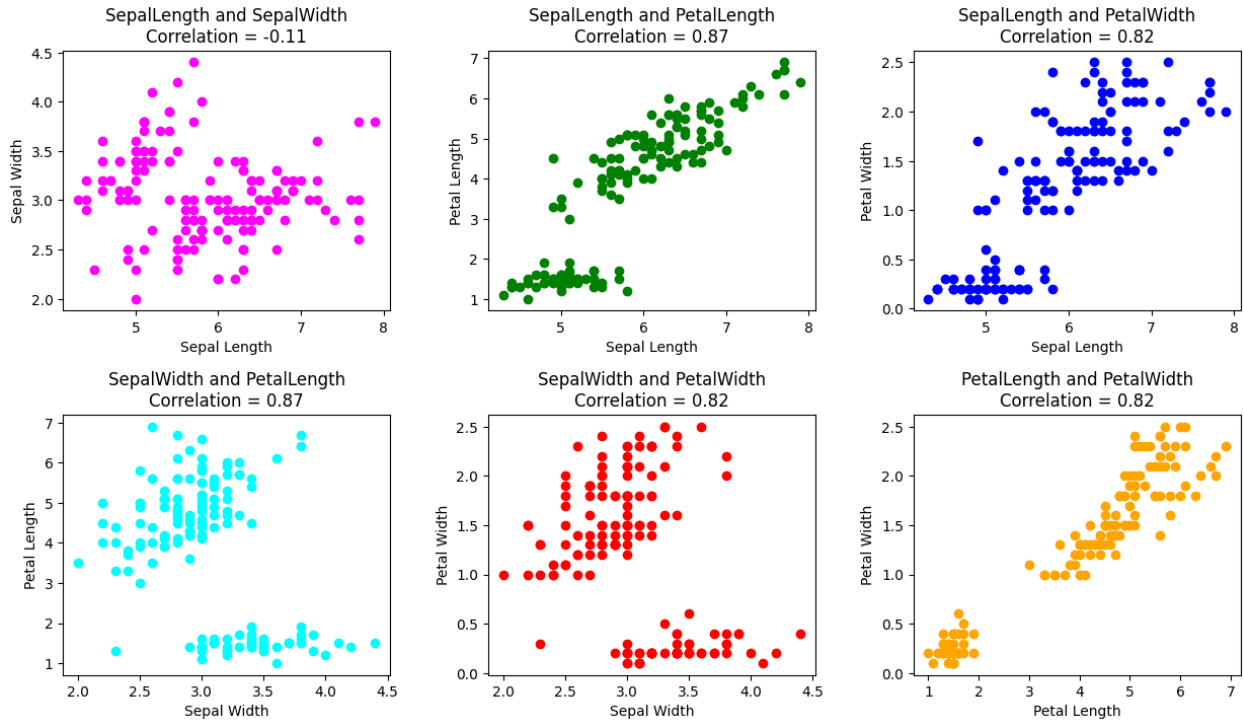
ax[1,0].scatter(df['SepalWidthCm'], df['PetalLengthCm'], color='cyan')
ax[1,0].title.set_text('SepalWidth and PetalLength\n'+ 'Correlation =
'+ "{:.2f}".format(c.iloc[0]['PetalLengthCm']))
ax[1,0].set_xlabel('Sepal Width')
ax[1,0].set_ylabel('Petal Length')

ax[1,1].scatter(df['SepalWidthCm'], df['PetalWidthCm'], color='red')
ax[1,1].title.set_text('SepalWidth and PetalWidth\n'+ 'Correlation =
'+ "{:.2f}".format(c.iloc[0]['PetalWidthCm']))
ax[1,1].set_xlabel('Sepal Width')
ax[1,1].set_ylabel('Petal Width')

ax[1,2].scatter(df['PetalLengthCm'], df['PetalWidthCm'],
color='orange')
ax[1,2].title.set_text('PetalLength and PetalWidth\n'+ 'Correlation =
'+ "{:.2f}".format(c.iloc[0]['PetalWidthCm']))
ax[1,2].set_xlabel('Petal Length')
ax[1,2].set_ylabel('Petal Width')

fig.subplots_adjust(wspace=0.3, hspace=0.4) #Mengatur setiap grafik
dengan spasi 0.3 di lebar dan 0.4 di tinggi
plt.show() #Menampilkan tampilan grafik yang telah dibuat

```



```
fig, ax = plt.subplots(nrows=2, ncols=2, figsize=(15,7))
#Membuat 4 grafik dengan 2 baris (nrows=2) dan 2 kolom dengan
(ncols=2)
#figsize(15,8) => memspesifikasi ukuran figur menjadi 15 inch kali 8
inch
#Menspesifikasi lokasi grafik setiap data dengan ax[n,n], misalnya
ax[0,0] berarti dia di kolom 0 dan baris 0
#Membuat jenis grafik yang berbeda di setiap kotak dengan coding yang
telah dijelaskan sebelumnya

#Membuat grafik scatter plot
ax[0,0].scatter(df['SepalLengthCm'], df['SepalWidthCm'],
color='magenta')
ax[0,0].title.set_text('Scatter plot')
ax[0,0].set_xlabel('Sepal Length')
ax[0,0].set_ylabel('Sepal Width')

#Membuat grafik batang
ax[0,1].bar(labels,data)
ax[0,1].title.set_text('Bar plot')
ax[0,1].set_xlabel('Variables')
ax[0,1].set_ylabel('Average Scores')

#Membuat grafik garis
ax[1,0].plot(df['SepalLengthCm'])
ax[1,0].plot(df['PetalLengthCm'])
ax[1,0].plot(df['SepalWidthCm'])
```

```

ax[1,0].plot(df['PetalWidthCm'])
ax[1,0].title.set_text('Line Plot')
ax[1,0].set_xlabel('Sepal Width')
ax[1,0].set_ylabel('Petal Length')
ax[1,0].legend(['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width'], loc='upper right', ncol=2)
ax[1,0].set_xlim([0,160])
ax[1,0].set_ylim([1,10])

```

#Membuat grafik boxplot

```

ax[1,1].boxplot(df, patch_artist=True, labels=labels)
ax[1,1].title.set_text('BoxPlot')

```

fig.subplots_adjust(wspace=0.3, hspace=0.4) #Mengatur setiap grafik dengan spasi 0.3 di lebar dan 0.4 di tinggi
plt.show() #Menampilkan tampilan grafik yang telah dibuat

