

LEMBAR KERJA PRAKTIKUM
TOPIK DALAM SISTEM PAKAR DAN SISTEM PENUNJANG KEPUTUSAN
LKP PERTEMUAN 3

VISUALISASI DATA

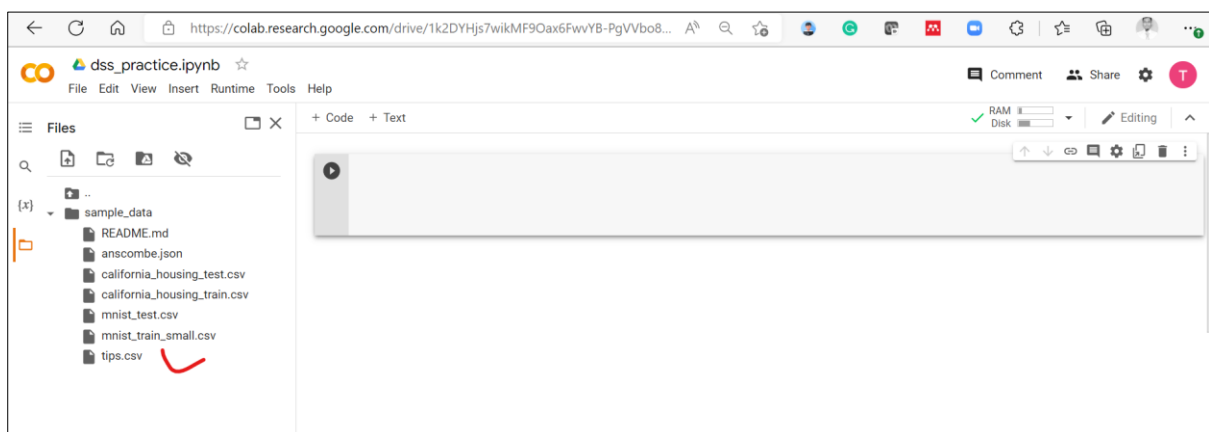
Rangga Pebrianto G6601231006

https://github.com/rangapebrianto/LKP_3_Rangga_Pebrianto_G6601231006

Melakukan visualisasi data merupakan salah satu bagian yang dapat digunakan untuk menganalisis. Untuk dapat melakukan analisis data, Python sebagai salah satu bahasa pemrograman menyediakan beberapa library yang dapat digunakan. Diantara library tersebut antara lain :

- a. Matplotlib
- b. Seaborn
- c. Plotly

Sebagai salah editor, google colab sebagai salah satu pilihan yang dapat digunakan dengan mengakes pada laman : <https://colab.research.google.com/>



Berikut ini adalah beberapa contoh penggunaan library tersebut.

- a. Matplotlib

- a.1 Melakukan pembacaan data.

Data yang digunakan adalah adalah Tips Database yang merupakan rekaman dataset tip yang diberikan oleh pengunjung suatu restoran dua setengah bulan di awal tahun 90 an. Data ini memiliki enam kolom mencakup : total_bill, tip, sex, smoker, day, time, size. Data dapat diunduh di newlms atau pada link berikut : <https://ipb.link/tips-database>. Melakukan pembacaan data pada python dapat dilakukan dengan menggunakan library pandas. Pada google colab library tersebut sudah disertakan sehingga kita tidak harus melakukan instalasi.

```
import pandas as pd

# reading the database
data = pd.read_csv("tips.csv") #lokasi menyesuaikan

# printing the top 10 rows
display(data.head(10))
```

Hasilnya adalah

	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
5	25.29	4.71	Male	No	Sun	Dinner	4
6	8.77	2.00	Male	No	Sun	Dinner	2
7	26.88	3.12	Male	No	Sun	Dinner	4
8	15.04	1.96	Male	No	Sun	Dinner	2
9	14.78	3.23	Male	No	Sun	Dinner	2

a.2 Scatter plot

Fungsi yang digunakan untuk menampilkan scatter plot adalah fungsi `scatter()`.

```
#Matplotlib Library

import pandas as pd
import matplotlib.pyplot as plt

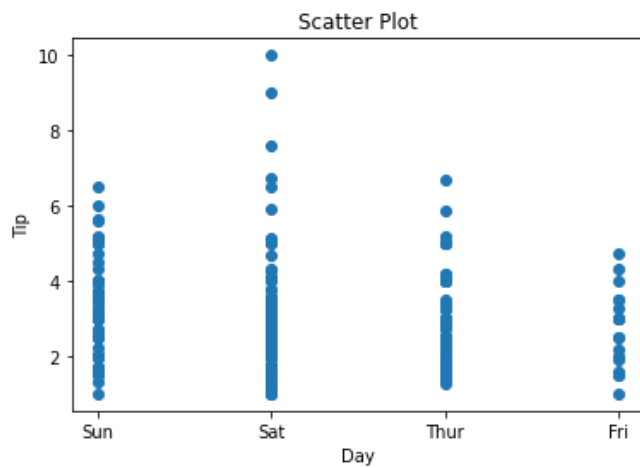
# reading the database
data = pd.read_csv("/content/sample_data/tips.csv")

# Scatter plot with day against tip
plt.scatter(data['day'], data['tip'])

# Adding Title to the Plot
plt.title("Scatter Plot")

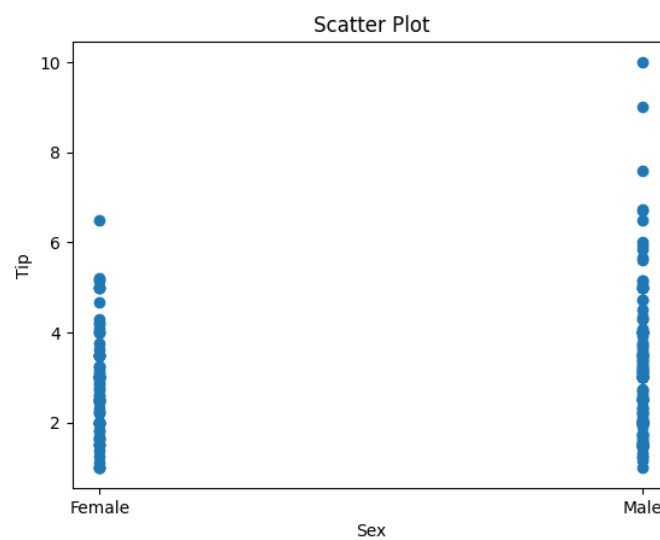
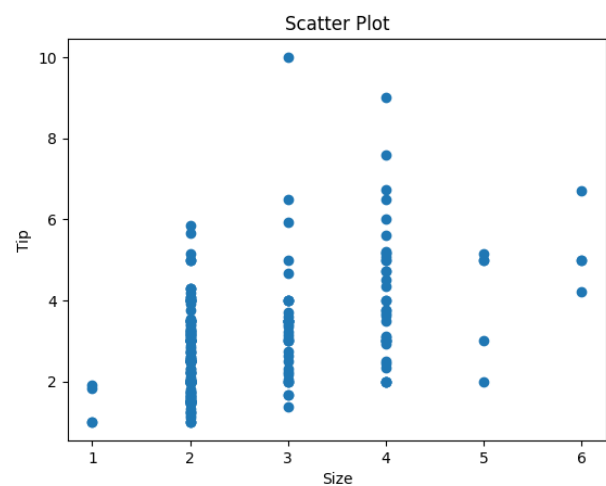
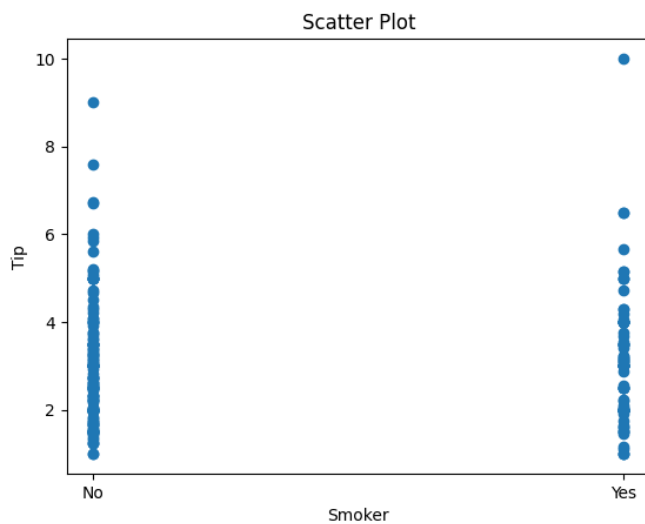
# Setting the X and Y labels
plt.xlabel('Day')
plt.ylabel('Tip')

plt.show()
```



Tugas

Tampilkan pada sumbu y adalah Tip dan sumbu x adalah smoker, size atau sex !



Grafik akan lebih menarik dan informatif jika ditampilkan menggunakan colobar dengan parameter c dan s yang dapat dilihat pada sintaks berikut.

```
import pandas as pd
import matplotlib.pyplot as plt

# reading the database
data = pd.read_csv("/content/sample_data/tips.csv")

# Scatter plot with day against tip
#plt.scatter(data['sex'], data['tip'])
# Scatter plot with day against tip
plt.scatter(data['day'], data['tip'], c=data['size'],
            s=data['total_bill'])

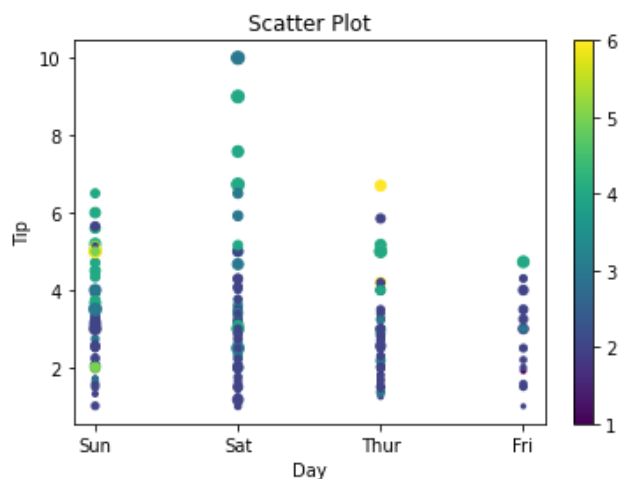
# Adding Title to the Plot
plt.title("Scatter Plot")

# Setting the X and Y labels
plt.xlabel('Day')
plt.ylabel('Tip')

plt.colorbar()

plt.show()
```

Visualisasi hasilnya dapat dilihat pada Gambar di bawah ini.



a.3 line Chart

Untuk membuat line chart, dapat dilakukan dengan menggunakan fungsi plot

Visualisasi hasilnya adalah sebagai berikut

```
import pandas as pd
import matplotlib.pyplot as plt

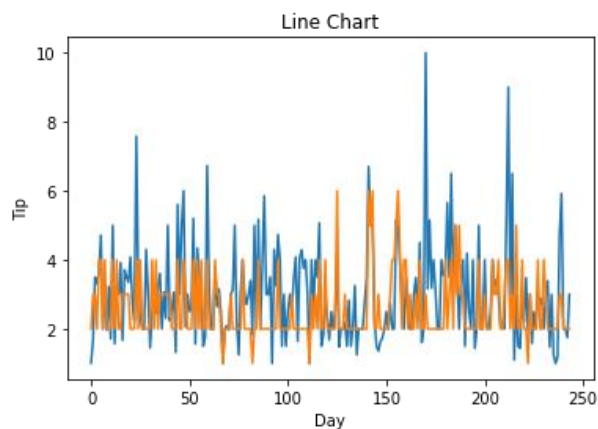
# reading the database
data = pd.read_csv("/content/sample_data/tips.csv")

# Scatter plot with day against tip
plt.plot(data['tip'])
plt.plot(data['size'])

# Adding Title to the Plot
plt.title("Scatter Plot")

# Setting the X and Y labels
plt.xlabel('Day')
plt.ylabel('Tip')

plt.show()
```



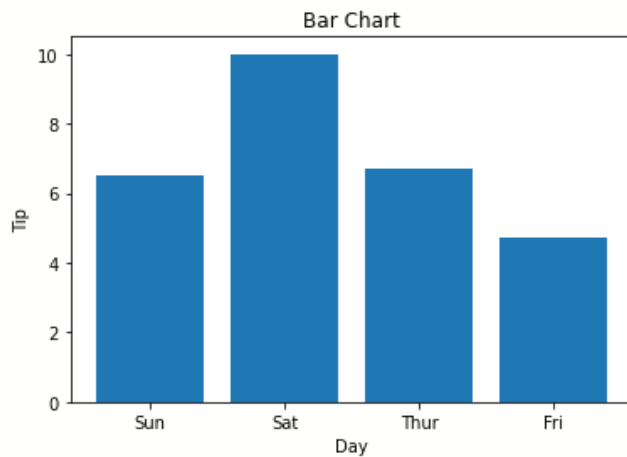
a.4 Bar Chart

Untuk membuat bar chart, dapat dilakukan dengan menggunakan fungsi bar(). Fungsi ini akan menyajikan data kategori dalam suatu diagram batang dengan tinggi dan lebar tertentu. Modifikasi yang dilakukan di bagian ini.

```
# Bar chart
plt.bar(data['day'], data['tip'])

# Adding Title to the Plot
plt.title("Bar Chart")
```

Visualisasi dari sintaks tersebut adalah sebagai berikut :



a.5 Histogram

Histogram pada dasarnya digunakan untuk merepresentasikan data dalam bentuk beberapa kelompok. Ini adalah jenis plot batang di mana sumbu X mewakili rentang bin sementara sumbu Y memberikan informasi tentang frekuensi. Fungsi hist() digunakan untuk menghitung dan membuat histogram. Berikut adalah sintaks untuk membuat histogram

```
import pandas as pd
import matplotlib.pyplot as plt

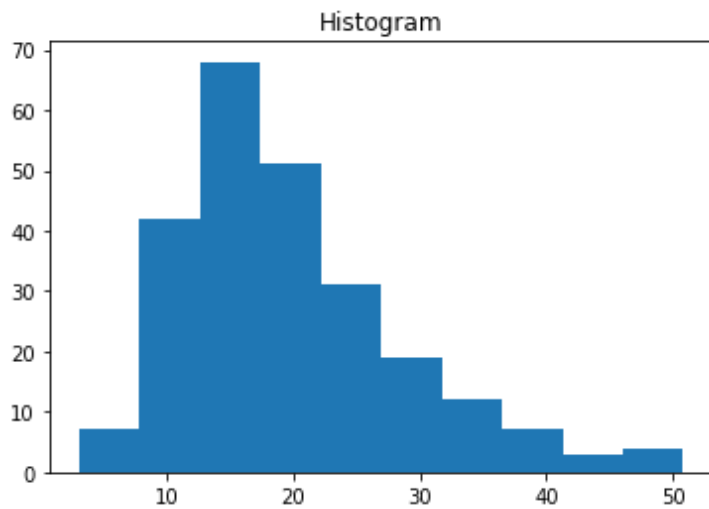
# reading the database
data = pd.read_csv("/content/sample_data/tips.csv")

# histogram of total_bills
plt.hist(data['total_bill'])

plt.title("Histogram")

# Adding the legends
plt.show()
```

Hasil visualisasinya adalah sebagai berikut :



b. Seaborn

Seaborn adalah library dengan tampilan antarmuka style yang leboh baik. Untuk dapat menggunakan library ini dapat melakukan pemanggilan paket seaborn dengan perintah import. Sintakn di bawah ini memperlihatkan penggunaan seaborn

```
import seaborn as sns
import matplotlib.pyplot as plt
import pandas as pd

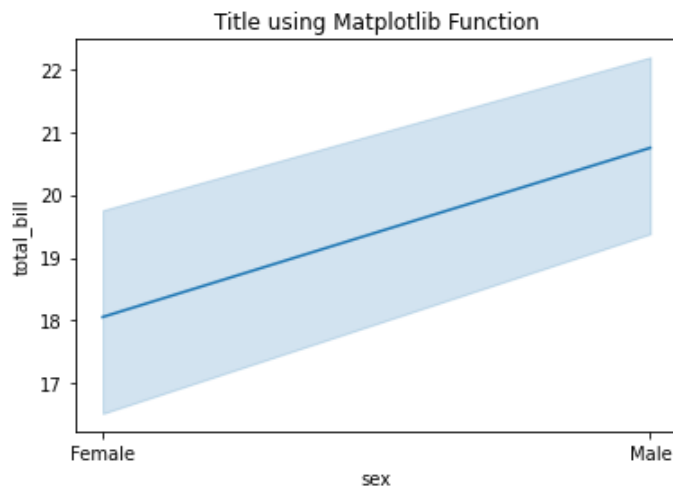
# reading the database
data = pd.read_csv("/content/sample_data/tips.csv")

# draw lineplot
sns.lineplot(x="sex", y="total_bill", data=data)

# setting the title using Matplotlib
plt.title('Title using Matplotlib Function')

plt.show()
```

Hasil visualiasi dari sintaks tersebut adalah



Tugas

Lakukan beberapa percobaan pada google colab dan sajikan visualiasinya pada lembar kerja di bawah ini !

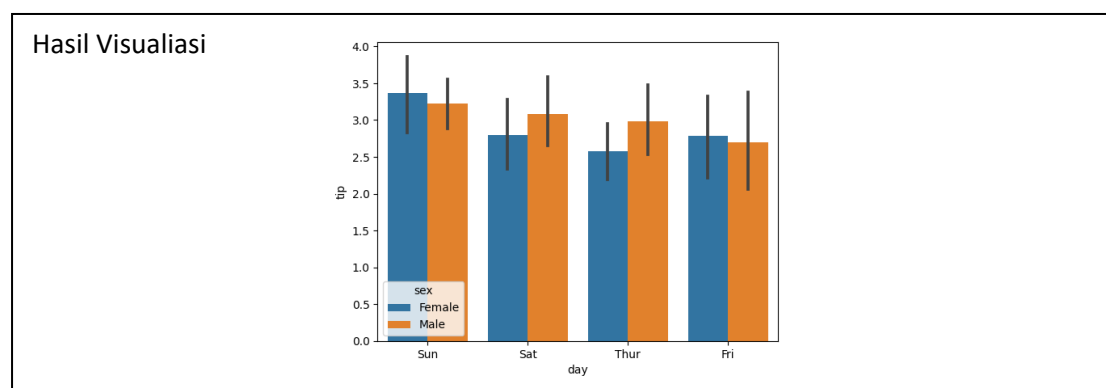
```
import seaborn as sns

import matplotlib.pyplot as plt
import pandas as pd

# reading the database
data = pd.read_csv("/content/sample_data/tips.csv")

sns.barplot(x='day', y='tip', data=data,
            hue='sex')

plt.show()
```



```
# importing packages
import seaborn as sns
import matplotlib.pyplot as plt
import pandas as pd

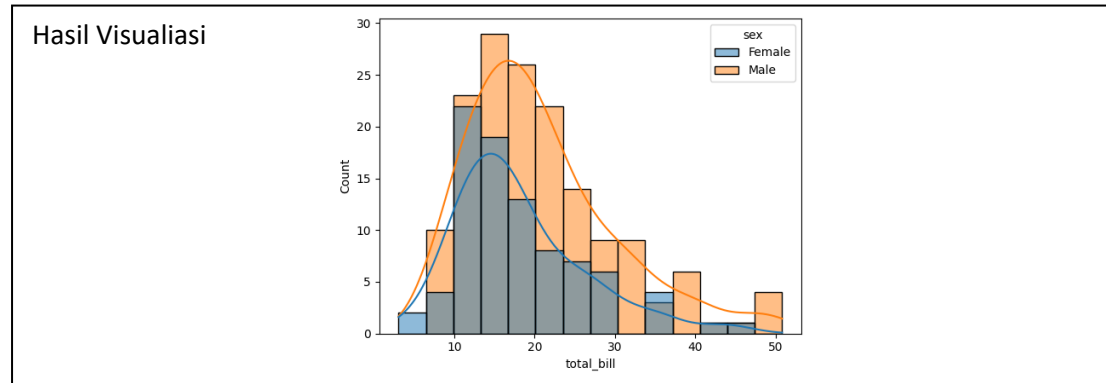
# reading the database
```



```
data = pd.read_csv("/content/sample_data/tips.csv ")

sns.histplot(x='total_bill', data=data, kde=True, hue='sex')

plt.show()
```



c. Plotly

Plotly menjadikan grafik lebih atraktif dan dapat dilakukan kustomisasi. Untuk dapat melakukan visualisasi ini, dapat menggunakan library plotly pada modul express.

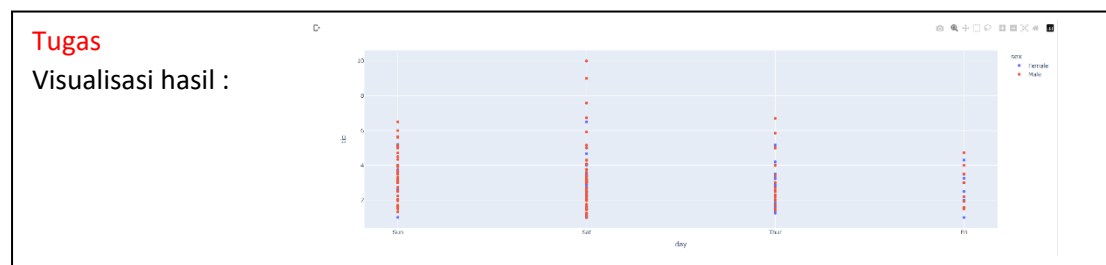
c.1 scatter plot

```
import plotly.express as px
import pandas as pd

# reading the database
data = pd.read_csv("/content/sample_data/tips.csv")

# plotting the scatter chart
fig = px.scatter(data, x="day", y="tip", color='sex')

# showing the plot
fig.show()
```



c.2 Line Chart

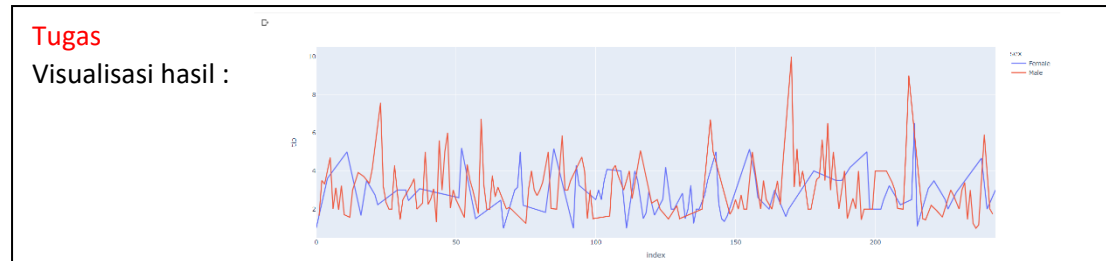
```
import plotly.express as px
import pandas as pd

# reading the database
```

```
data = pd.read_csv("/content/sample_data/tips.csv")

# plotting the scatter chart
fig = px.line(data, y='tip', color='sex')

# showing the plot
fig.show()
```



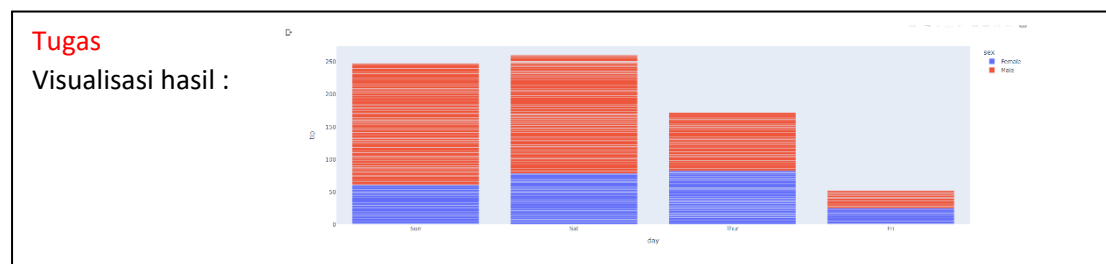
c.3 Bar Chart

```
import plotly.express as px
import pandas as pd

# reading the database
data = pd.read_csv("/content/sample_data/tips.csv")

# plotting the scatter chart
fig = px.bar(data, x='day', y='tip', color='sex')

# showing the plot
fig.show()
```



c.4 Histogram

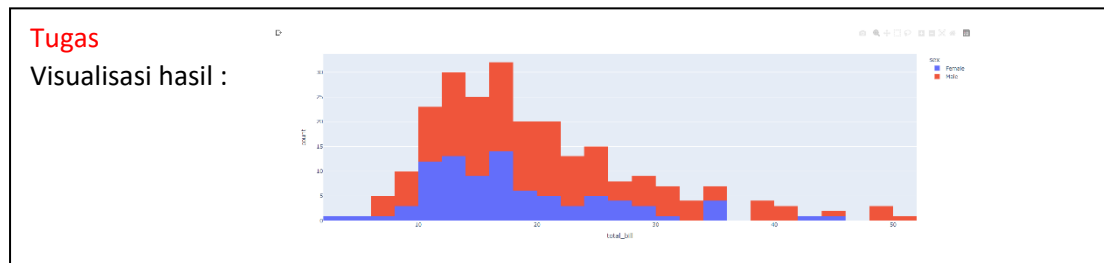
```
import plotly.express as px
import pandas as pd

# reading the database
```

```
data = pd.read_csv("/content/sample_data/tips.csv")

# plotting the scatter chart
fig = px.histogram(data, x='total_bill', color='sex')

# showing the plot
fig.show()
```



REFLEKSI PRAKTIKUM

Tugas

Refleksi Praktikum

Silahkan menuliskan apa yang menjadi bahan pembelajaran Anda

Tugas

Refleksi Praktikum

Silahkan menuliskan apa yang menjadi bahan pembelajaran Anda

"Kodingan ini telah memberikan saya wawasan yang berharga tentang penggunaan beberapa library populer dalam analisis dan visualisasi data di Python. Saya belajar cara mengambil dan memanipulasi data dari file CSV menggunakan Pandas, menciptakan berbagai jenis visualisasi seperti scatter plot, bar chart, dan histogram dengan Matplotlib, mengaplikasikan Seaborn untuk visualisasi data statistik yang lebih kompleks, dan menggunakan Plotly Express untuk membuat visualisasi data interaktif yang menarik. Berikut adalah beberapa poin penting yang saya pelajari dari setiap bagian kodingan:

- Penggunaan Pandas dalam membaca dan menampilkan data dari file CSV, termasuk teknik menampilkan 10 baris pertama data dengan `display()`.
- Pembuatan scatter plot dengan Matplotlib, termasuk penambahan judul dan label sumbu untuk memperjelas plot.
- Penggunaan warna dan ukuran tambahan pada scatter plot dengan Matplotlib berdasarkan kolom "size" dan "total_bill".
- Membuat plot garis dengan Matplotlib dari kolom "tip" dan "size", serta menambahkan elemen judul dan label sumbu.
- Membuat bar chart dengan Matplotlib berdasarkan data kolom "day" dan "tip", serta menambahkan judul yang sesuai.

- Membuat histogram dari kolom "total_bill" dengan Matplotlib, dan menambahkan judul untuk memberikan konteks pada visualisasi.
- Penggunaan Seaborn untuk membuat bar plot yang kompleks dengan pengelompokan berdasarkan kolom "day" dan "tip", dengan tambahan pemisahan berdasarkan kolom "sex".
- Membuat histogram yang informatif dari kolom "total_bill" dengan Seaborn, termasuk estimasi densitas kernel (KDE) dan pemilihan warna berdasarkan kolom "sex".
- Membuat scatter plot, plot garis, bar chart, dan histogram interaktif dengan Plotly Express untuk eksplorasi data yang lebih mendalam.

Semua pembelajaran ini memberikan fondasi yang kokoh dalam analisis dan visualisasi data, yang dapat saya terapkan dalam berbagai proyek analisis data saya di masa mendatang."