

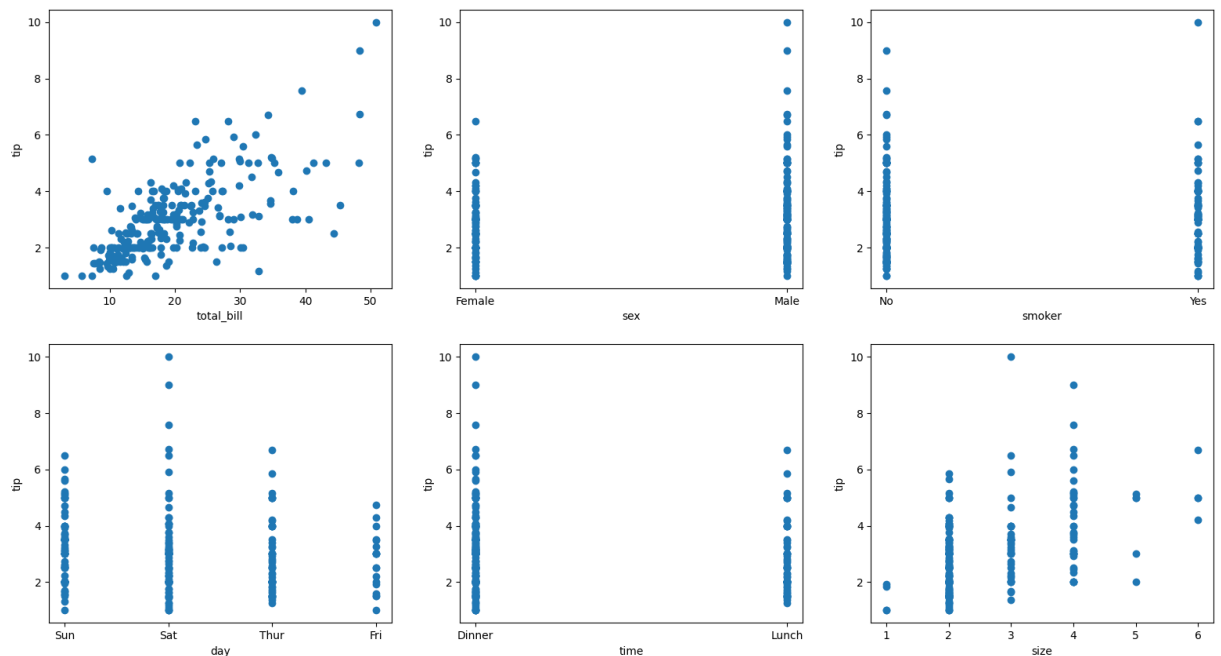
Visualisasi Data Menggunakan Python dan *library* Matplotlib Sebagai Pustaka Visualisasi dan *library* Pandas Sebagai Pustaka Penyimpanan Dataframe

Wildan Fajri Alfarabi G64190060

Data yang digunakan merupakan data nominal dan numerik, serta visualisasi yang telah dilakukan adalah scatterplot, linechart, barchart, dan histogram. Data yang digunakan dalam pembelajaran ini adalah data dari LKP yaitu data <https://ipb.link/tips-database>.

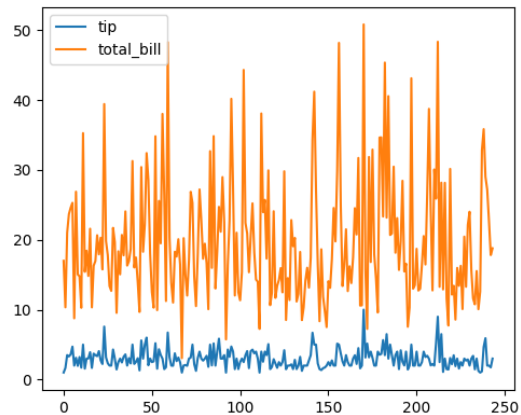
Visualisasi Data :

- Scatter Plot



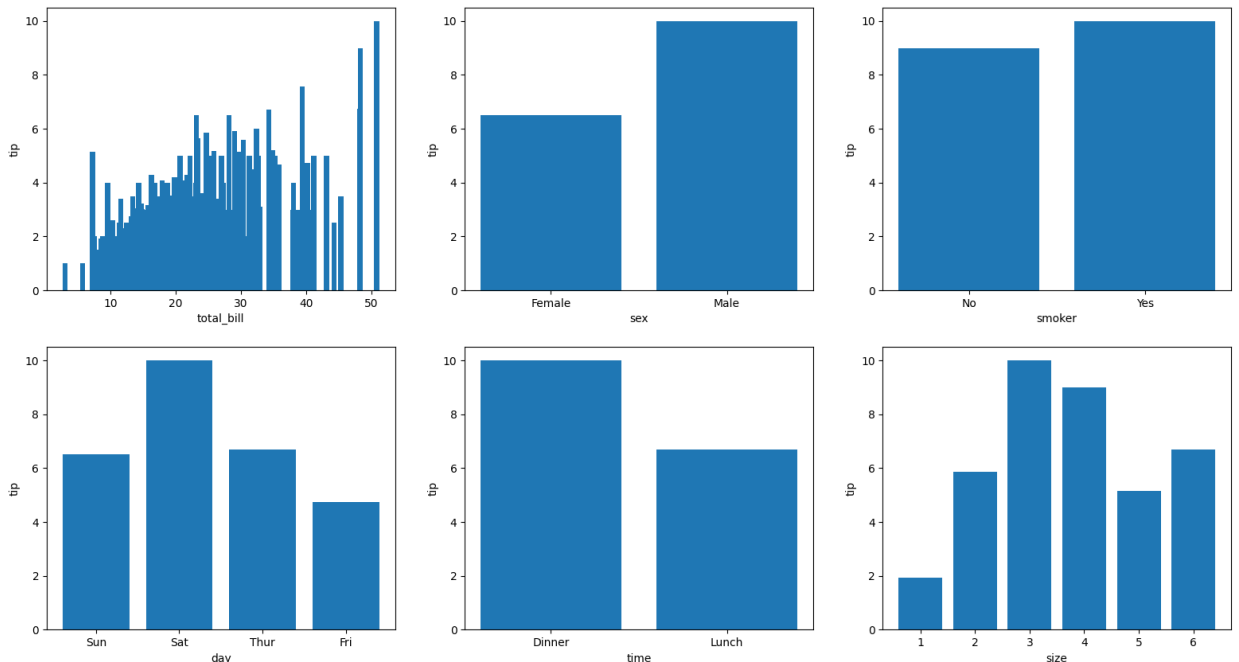
Scatterplot dapat menunjukkan pengaruh atribut terhadap atribut lain maupun amatan suatu data. Pada data tips.csv, scatterplot dapat menunjukkan pengaruh tiap variabel terhadap jumlah tips yang diberikan pelanggan. Dapat dilihat bahwa total bill berbanding lurus dengan tips yang diberikan, menunjukkan bahwa semakin tinggi bill pelanggan, semakin tinggi juga tip yang diberikan.

- Line Chart



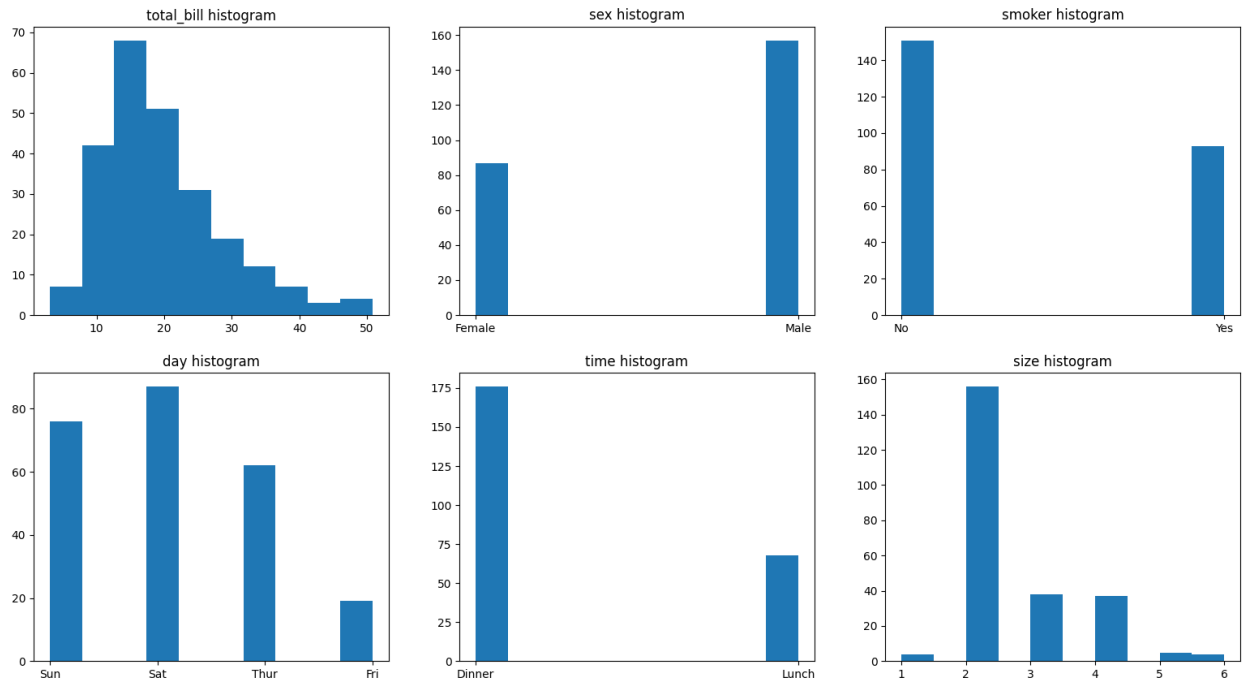
Line Chart dapat menunjukkan pergerakan data terhadap suatu satuan waktu, maupun dapat digunakan sebagai perbandingan 2 buah atribut yang memiliki nilai satuan yang sama. Dalam data tips.csv, variabel yang dapat dibandingkan adalah tip dengan total_bill, dimana axis-x adalah id data dan axis-y adalah jumlah uang, dari line chart diketahui tip semakin tinggi mengikuti total_bill yang semakin tinggi.

- Bar Chart



Bar Chart dapat mengetahui nilai kelas atau amatan pada nilai tertentu suatu atribut. Untuk atribut total_bill, diketahui nilai tips minimum ada pada nilai < 2 ketika total_bill dibawah < 10 dan tips maksimum ketika total_bill bernilai maksimum > 50 . Selain itu, apda barchart atribut gender sex dengan tips, diketahui rata-rata pelanggan laki-laki memberikan tips yang lebih banyak dibandingkan perempuan.

- Histogram



Histogram menunjukkan frekuensi kemunculan nilai pada suatu atribut. Histogram dapat digunakan untuk menganalisa sebaran data. Pada histogram tip.csv diatas, sebaran atribut size memiliki modus pada size 2 dan pada atribut total_bill, rata-rata pelanggan memesan dengan total_bill dengan rentang 10-30. Pelanggan juga paling ramai pada hari saturday apabila mengacu pada histogram diatas.

Code yang telah dibuat :

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

def scatter_plots(ax, row, col, data, var_title, data_classes):
    ax[row, col].scatter(data[var_title], data[data_classes])
    ax[row, col].set(xlabel=var_title, ylabel=data_classes)

def line_chart(ax, row, col, data, var_title, data_classes):
    ax[row, col].plot(data[data_classes], label=data_classes)
```

```

ax[row, col].plot(data[var_title], label=var_title)
ax[row, col].legend(loc="upper left")

def bar_chart(ax, row, col, data, var_title, data_classes):
    ax[row, col].bar(data[var_title], data[data_classes])
    ax[row, col].set(xlabel=var_title, ylabel=data_classes)

def histogram(ax, row, col, data, var_title):
    ax[row, col].hist(data[var_title])
    ax[row, col].set_title(var_title + " histogram")

def subplots(data, type_visual, data_classes):
    var = data.drop(columns=[data_classes])
    var_cnt = len(data.axes[1]) - 1
    fig, ax = plt.subplots(2, var_cnt // 2)
    for idx, var_title in enumerate(var.axes[1]):
        if idx >= var_cnt // 2:
            if type_visual == "scatterplot":
                scatter_plots(ax, 1, idx % (var_cnt // 2),
data, var_title, data_classes)
            elif type_visual == "linechart":
                line_chart(ax, 1, idx % (var_cnt // 2), data,
var_title, data_classes)
            elif type_visual == "barchart":
                bar_chart(ax, 1, idx % (var_cnt // 2), data,
var_title, data_classes)
            elif type_visual == "histogram":
                histogram(ax, 1, idx % (var_cnt // 2), data,
var_title)
        else:
            if type_visual == "scatterplot":

```

```

        scatter_plots(ax, 0, idx, data, var_title,
data_classes)
        elif type_visual == "linechart":
            line_chart(ax, 0, idx, data, var_title,
data_classes)
        elif type_visual == "barchart":
            bar_chart(ax, 0, idx, data, var_title,
data_classes)
        elif type_visual == "histogram":
            histogram(ax, 0, idx, data, var_title)
plt.show()

if __name__ == "__main__":
    data_contoh_tipsdb = pd.read_csv("tips.csv")
    print(len(data_contoh_tipsdb.axes[0]),
len(data_contoh_tipsdb.axes[1]))
    print(data_contoh_tipsdb.axes[1][0])

    subplots(data_contoh_tipsdb, "scatterplot", "tip")
    subplots(data_contoh_tipsdb, "linechart", "tip")
    subplots(data_contoh_tipsdb, "barchart", "tip")
    subplots(data_contoh_tipsdb, "histogram", "tip")

    print(data_contoh_tipsdb.head(10))

```