Matplotlib 21_ Pandas Data Frame dan Matplotlib (part 2)

June 9, 2022

1 Visualisasi Data yang tersimpan pada Pandas Data Frame (bagian 2)

Dalam sesi ini kita akan mempelajari cara untuk melakukan visualisasi data yang tersimpan pada Pandas Data Frame dengan Matplotlib.

1.1 1. Import Modules

[1]:
```
%matplotlib inline
```

[2]:
```
import matplotlib
import matplotlib.pyplot as plt
import pandas as pd

print(matplotlib.__version__)
print(pd.__version__)
```

```
3.3.4
1.2.4
```

1.2 2. Line Plot

1.2.1 Sample Dataset

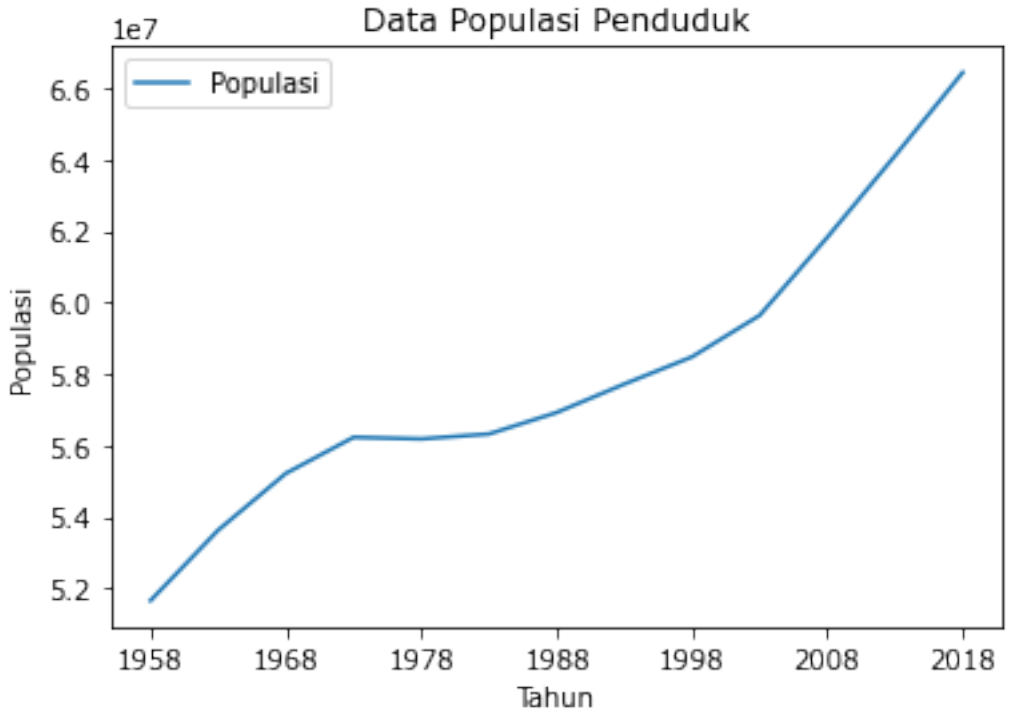
[3]:
```
data = {
    'Tahun': ['1958', '1963', '1968', '1973', '1978', '1983', '1988', '1993',␣
 ↪'1998', '2003', '2008', '2013', '2018'],
    'Populasi': [51652500, 53624900, 55213500, 56223000, 56178000, 56315000,␣
 ↪56916000, 57713000, 58474000, 59636000, 61823000, 64105000, 66436000]
}

df = pd.DataFrame(data)
df
```

[3]:

	Tahun	Populasi
0	1958	51652500
1	1963	53624900

```
2   1968   55213500
3   1973   56223000
4   1978   56178000
5   1983   56315000
6   1988   56916000
7   1993   57713000
8   1998   58474000
9   2003   59636000
10  2008   61823000
11  2013   64105000
12  2018   66436000
```

[4]:
```
df.plot(x='Tahun', y='Populasi', kind='line')

plt.title('Data Populasi Penduduk')
plt.ylabel('Populasi')
plt.xlabel('Tahun')

plt.show()
```



1.3 3.Bar Plot

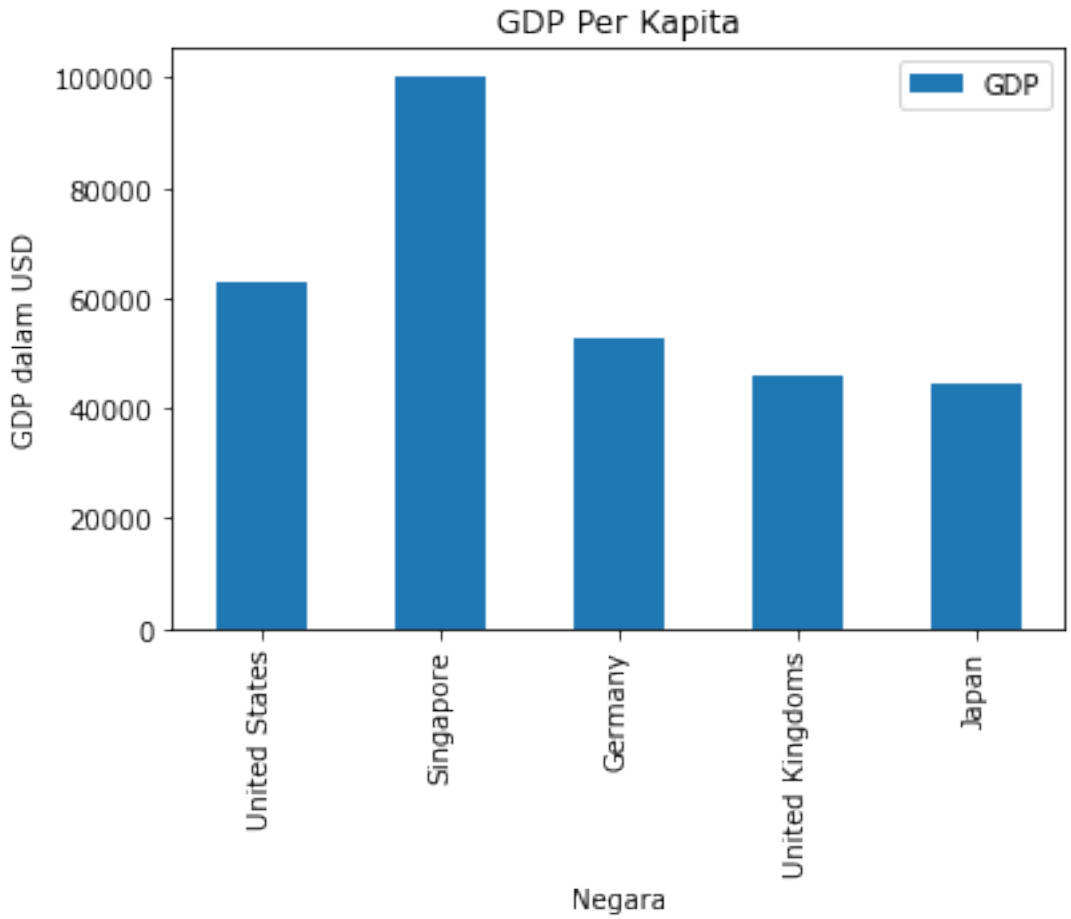
1.3.1 Sample Dataset

[5]:
```
Data = {'Negara': ['United States', 'Singapore', 'Germany', 'United Kingdoms',␣
↪'Japan'],
       'GDP': [62606, 100345, 52559, 45705, 44227]
       }

df = pd.DataFrame(Data, columns=['Negara', 'GDP'])
df
```

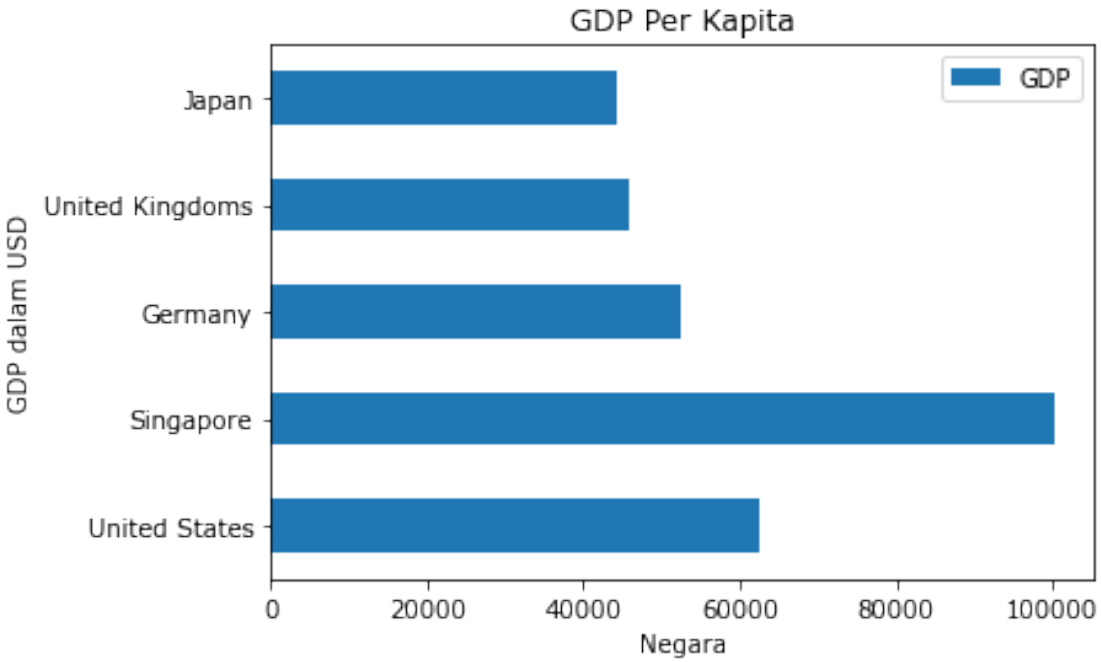
[5]:

	Negara	GDP
0	United States	62606
1	Singapore	100345
2	Germany	52559
3	United Kingdoms	45705
4	Japan	44227

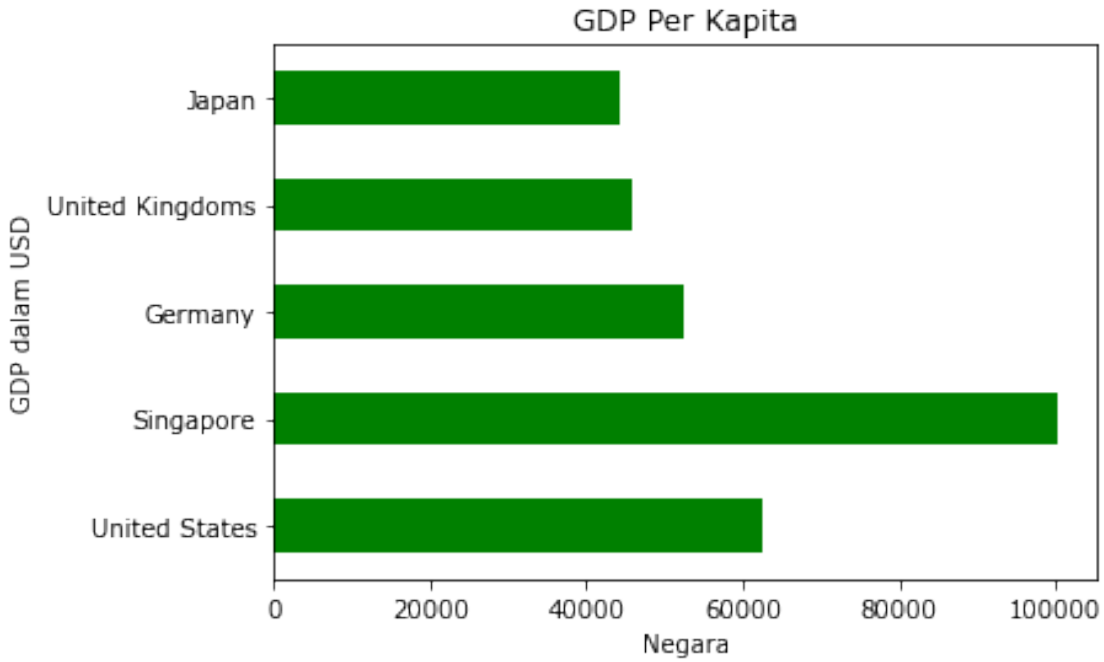
[6]:
```
df.plot(x='Negara', y='GDP', kind='bar')

plt.title('GDP Per Kapita')
plt.ylabel('GDP dalam USD')
plt.xlabel('Negara')

plt.show()
```



[7]:
```
df.plot(x='Negara', y='GDP', kind='barh')

plt.title('GDP Per Kapita')
plt.ylabel('GDP dalam USD')
plt.xlabel('Negara')

plt.show()
```



[8]:
```
df.plot(x='Negara', y='GDP', kind='barh', color='green', legend=False)

plt.title('GDP Per Kapita')
plt.ylabel('GDP dalam USD')
plt.xlabel('Negara')

plt.show()
```



1.4 4. Scatter Plot

1.4.1 Sample Dataset

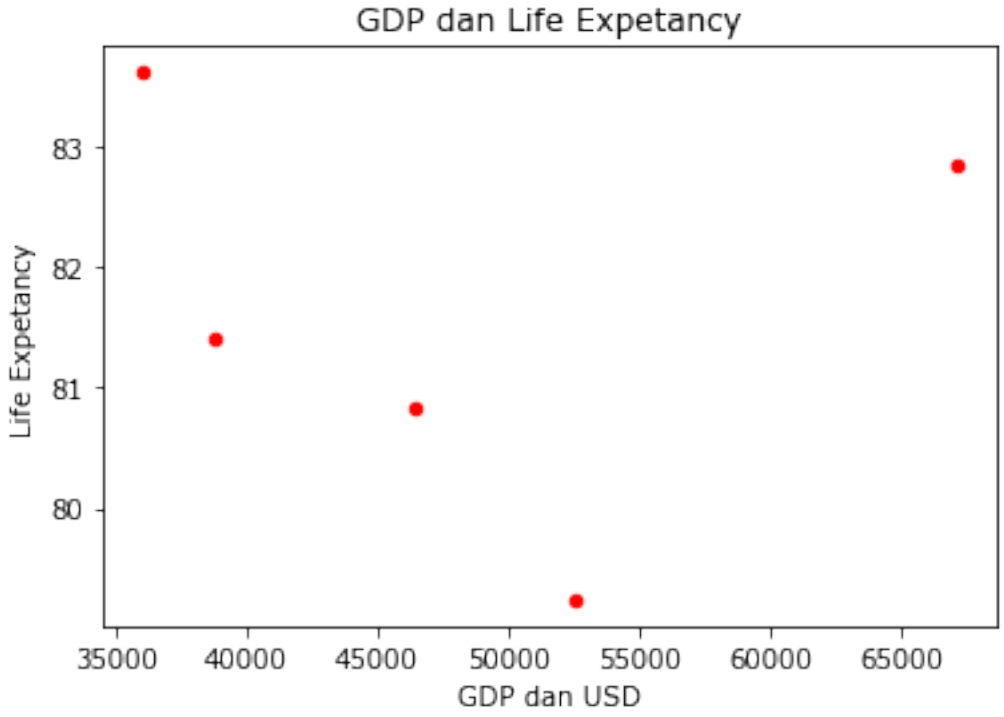
[9]:
```
data = {'Negara': ['United States', 'Singapore', 'Germany', 'United Kingdom',␣
 ↪'Japan'],
       'GDP': [52591, 67110, 46426, 38749, 36030],
       'Life_Expetancy': [79.24, 82.84, 80.84, 81.40, 83.62]
       }

df = pd.DataFrame(data, columns=['Negara', 'GDP', 'Life_Expetancy'])
df
```

[9]:

	Negara	GDP	Life_Expetancy
0	United States	52591	79.24
1	Singapore	67110	82.84
2	Germany	46426	80.84
3	United Kingdom	38749	81.40
4	Japan	36030	83.62

[10]:
```
df.plot(kind='scatter', x='GDP', y='Life_Expetancy', color='red')

plt.title('GDP dan Life Expetancy')
plt.ylabel('Life Expetancy')
plt.xlabel('GDP dan USD')
```

```
plt.show()
```



1.5 5. Pie Plot

1.5.1 Sample Dataset

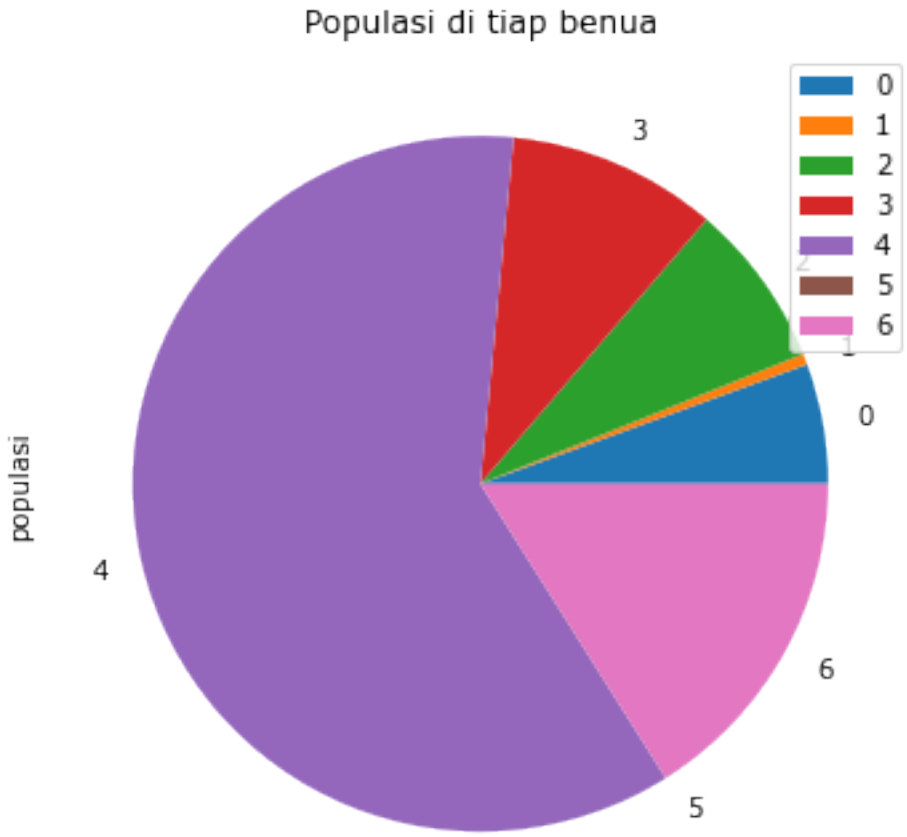
[11]:
```
data = {'benua': ['South America', 'Ocenia', 'North America', 'Europe', 'Asia',
 ↪'Antartica', 'Africa'],
       'populasi': [422535000, 38384000, 579024000, 738849000, 4581757408,
 ↪1106, 1216130000]
       }

df = pd.DataFrame(data)
df
```

[11]:

	benua	populasi
0	South America	422535000
1	Ocenia	38384000
2	North America	579024000
3	Europe	738849000
4	Asia	4581757408
5	Antartica	1106

```
6          Africa  1216130000
```

[12]:
```
df.plot(kind='pie', y='populasi', figsize=(6, 6))

plt.title('Populasi di tiap benua')
plt.show()
```

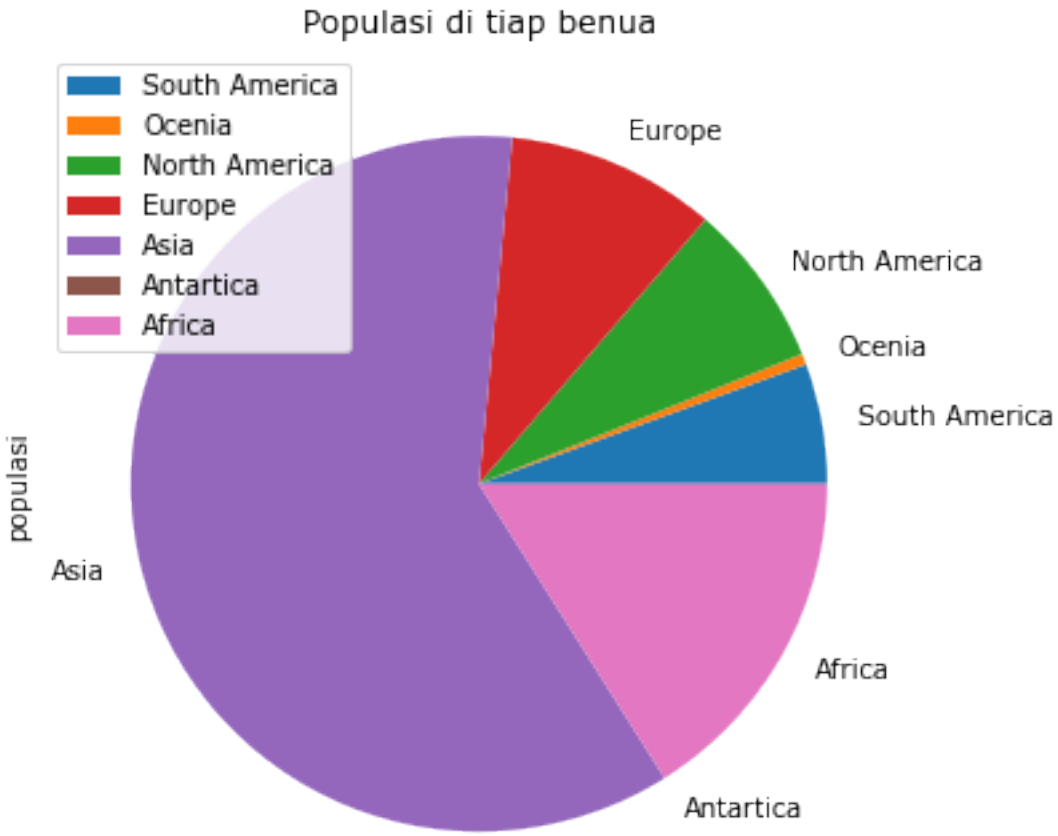


[13]:
```
df = df.set_index('benua')
df
```

[13]:
```
                populasi
benua
South America   422535000
Ocenia           38384000
North America   579024000
Europe          738849000
Asia           4581757408
Antartica            1106
```

```
Africa          1216130000
```

[14]:
```
df.plot(kind='pie', y='populasi', figsize=(6,6))

plt.title('Populasi di tiap benua')

plt.show()
```



1.6 6. Box Plot

1.6.1 sample dataset

[15]:
```
data = {'benua': ['South America', 'Ocenia', 'North America', 'Europe', 'Asia',␣
↪'Antartica', 'Africa'],
      'populasi': [422535000, 38384000, 579024000, 738849000, 4581757408,␣
↪1106, 1216130000]
      }
```

```
df = pd.DataFrame(data)
df
```

[15]:

	benua	populasi
0	South America	422535000
1	Ocenia	38384000
2	North America	579024000
3	Europe	738849000
4	Asia	4581757408
5	Antartica	1106
6	Africa	1216130000

[16]:
```
df['populasi'].plot(kind='box')

plt.title('Sebaran Populasi')
plt.ylabel('Jumlah')

plt.show()
```