

# Aussa Tris Rahmatullah

## Education Background



**2010-2014**  
**Bachelor Degree**  
Electrical Engineering



**2018-2019**  
**Master Degree**  
Data Science and Analytics



Aussa Tris Rahmatullah

<https://id.linkedin.com/in/aussa-tris-rahmatullah>

**Aussa Rahmatullah**

Data Scientist

Telco Company

# Data Visualization

Basic Plots



# Objektif



Memahami apa **urgensi proses** memvisualisasikan data dan apa **prinsip yang harus dipegang** dalam melakukan visualisasi data, dan mampu membuat dan memilih jenis chart sesuai dengan kegunaan masing masing.

# Expected Output



1. Memahami apa **urgensi proses** memvisualisasikan data dan apa **prinsip** yang harus dipegang dalam melakukan visualisasi data
2. Memahami **perbedaan kegunaan** masing masing jenis chart dan tabel dalam merepresentasikan data
3. Mampu melakukan operasi pembuatan chart dan tabel untuk menggambarkan **suatu perbandingan data**
4. Mampu melakukan operasi pembuatan chart dan tabel untuk menggambarkan **hubungan dalam data**
5. Mampu melakukan operasi pembuatan chart dan tabel untuk menggambarkan **distribusi pada data**
6. Mampu melakukan operasi pembuatan chart dan tabel untuk menggambarkan **komposisi pada data**



# Topik Data Visualization

## Sesi : Basic

 Pengenalan Data Visualization  
(Apa dan Mengapa)

 Tipe Plot dan Penggunaannya - Part 1

 Tipe Plot dan Penggunaannya - Part 2

 Pengenalan pada Matplotlib  
dan Seaborn

 Data Visualization Dasar (Praktik) - Part 1

 Data Visualization Dasar (Praktik) - Part 2

## Sesi : Advanced

 Customize Plot (Praktik) - Part 1

 Customize Plot (Praktik) - Part 2

 Data Storytelling - Part 1

 Data Storytelling - Part 2

# Topik Data Visualization

## Sesi : Basic



**Pengenalan Data Visualization  
(Apa dan Mengapa)**

**Tipe Plot dan Penggunaannya - Part 1**

**Tipe Plot dan Penggunaannya - Part 2**

**Pengenalan pada Matplotlib  
dan Seaborn**

**Data Visualization Dasar (Praktik) - Part 1**

**Data Visualization Dasar (Praktik) - Part 2**

## Sesi : Advanced

Customize Plot (Praktik) - Part 1

Customize Plot (Praktik) - Part 2

Data Storytelling - Part 1

Data Storytelling - Part 2

# Pengenalan **Data Visualization**

## Indonesia Situation

# 3,892,479

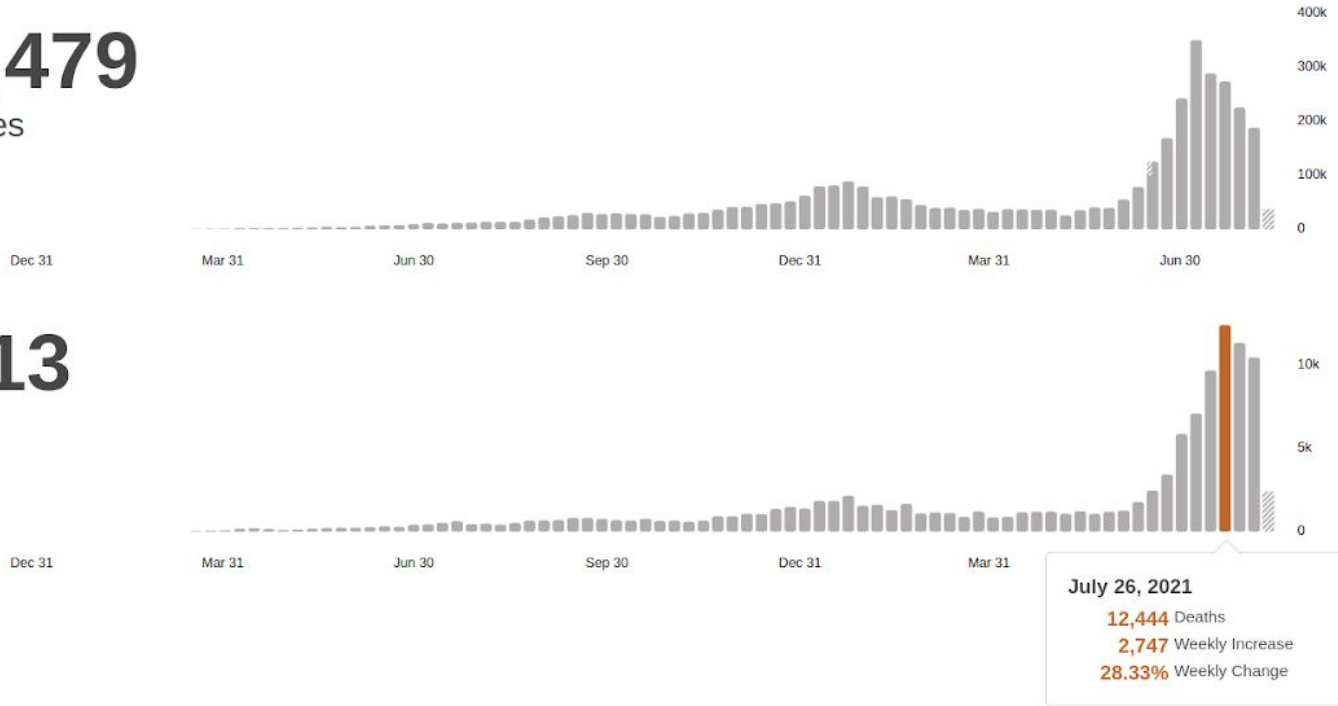
confirmed cases

# 120,013

deaths

Source: World Health Organization

Data may be incomplete for the current day or week.

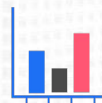


Jumlah kasus baru COVID-19 yang terkonfirmasi per minggunya di Indonesia

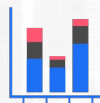


# Apa itu *Data Visualization*?

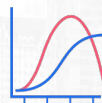
“Teknik yang digunakan untuk **mengkomunikasikan data** atau informasi dengan mengkodekannya sebagai **objek visual** (titik, garis atau batang) yang terdapat dalam grafik.”



Bar chart



Stacked bar chart



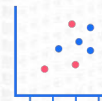
Line graph



Gantt chart



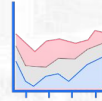
Polar area diagram



Scatter plot



Calendar heatmap



Stacked area chart



Sparkline

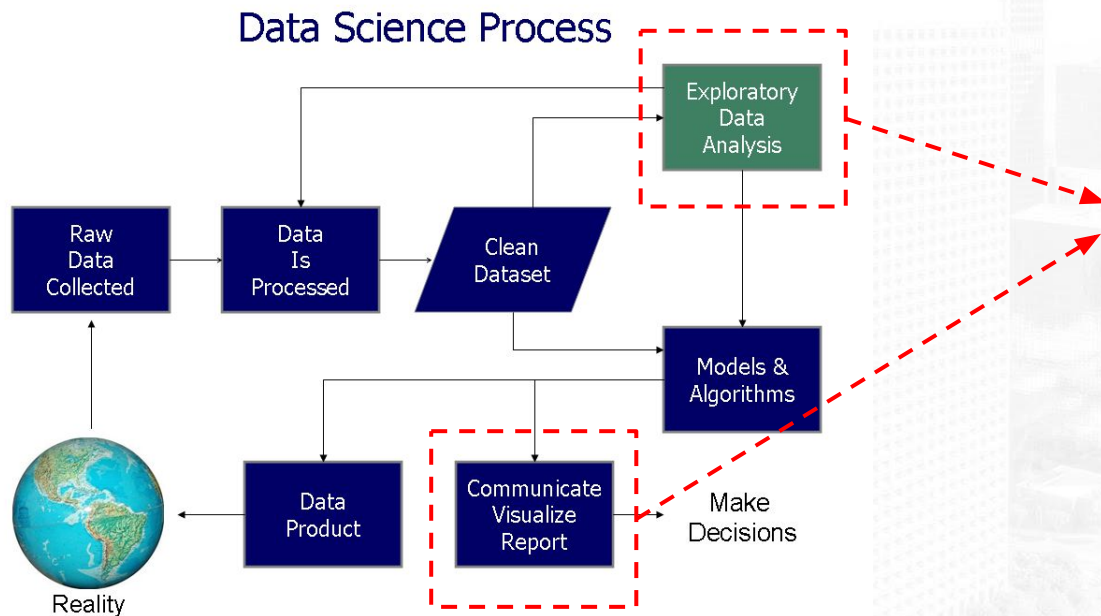


Column sparkline

*“An ideal visualization should not only communicate clearly, but stimulate **viewer engagement** and **attention**”*

Fernanda Viegas and Martin Wattenberg (April 19, 2011).  
"How To Make Data Look Sexy". CNN.com

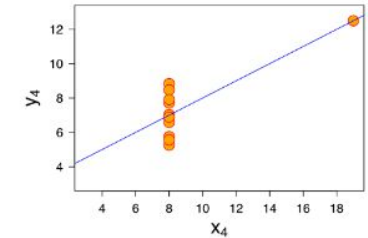
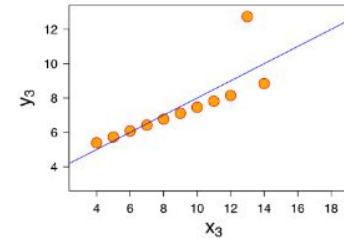
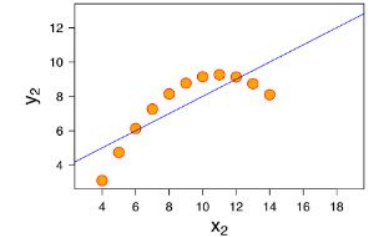
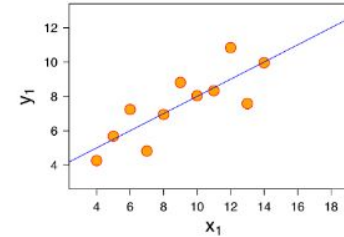
# Mengapa *Data Visualization*?



*Data Visualization* dibutuhkan untuk melakukan dua tahap ini

# Mengapa *Data Visualization*?

	I		II		III		IV	
	x	y	x	y	x	y	x	y
	10	8,04	10	9,14	10	7,46	8	6,58
	8	6,95	8	8,14	8	6,77	8	5,76
	13	7,58	13	8,74	13	12,74	8	7,71
	9	8,81	9	8,77	9	7,11	8	8,84
	11	8,33	11	9,26	11	7,81	8	8,47
	14	9,96	14	8,1	14	8,84	8	7,04
	6	7,24	6	6,13	6	6,08	8	5,25
	4	4,26	4	3,1	4	5,39	19	12,5
	12	10,84	12	9,13	12	8,15	8	5,56
	7	4,82	7	7,26	7	6,42	8	7,91
	5	5,68	5	4,74	5	5,73	8	6,89
SUM	99,00	82,51	99,00	82,51	99,00	82,50	99,00	82,51
AVG	9,00	7,50	9,00	7,50	9,00	7,50	9,00	7,50
STDEV	3,32	2,03	3,32	2,03	3,32	2,03	3,32	2,03



## Anscombe's Quartet

Terdapat 4 dataset yang mempunyai statistik deskriptif yang sama, namun pola visualisasi yang jauh berbeda

# Topik Data Visualization

## Sesi : Basic

 **Pengenalan Data Visualization  
(Apa dan Mengapa)**

 **Tipe Plot dan Penggunaannya - Part 1**

 **Tipe Plot dan Penggunaannya - Part 2**

 **Pengenalan pada Matplotlib  
dan Seaborn**

 **Data Visualization Dasar (Praktik) - Part 1**

 **Data Visualization Dasar (Praktik) - Part 2**

## Sesi : Advanced

 **Customize Plot (Praktik) - Part 1**

 **Customize Plot (Praktik) - Part 2**

 **Data Storytelling - Part 1**

 **Data Storytelling - Part 2**



# Topik Data Visualization

## Sesi : Basic

 **Pengenalan Data Visualization  
(Apa dan Mengapa)**

  **Tipe Plot dan Penggunaannya - Part 1**

 **Tipe Plot dan Penggunaannya - Part 2**

 **Pengenalan pada Matplotlib  
dan Seaborn**

 **Data Visualization Dasar (Praktik) - Part 1**

 **Data Visualization Dasar (Praktik) - Part 2**

## Sesi : Advanced

 **Customize Plot (Praktik) - Part 1**

 **Customize Plot (Praktik) - Part 2**

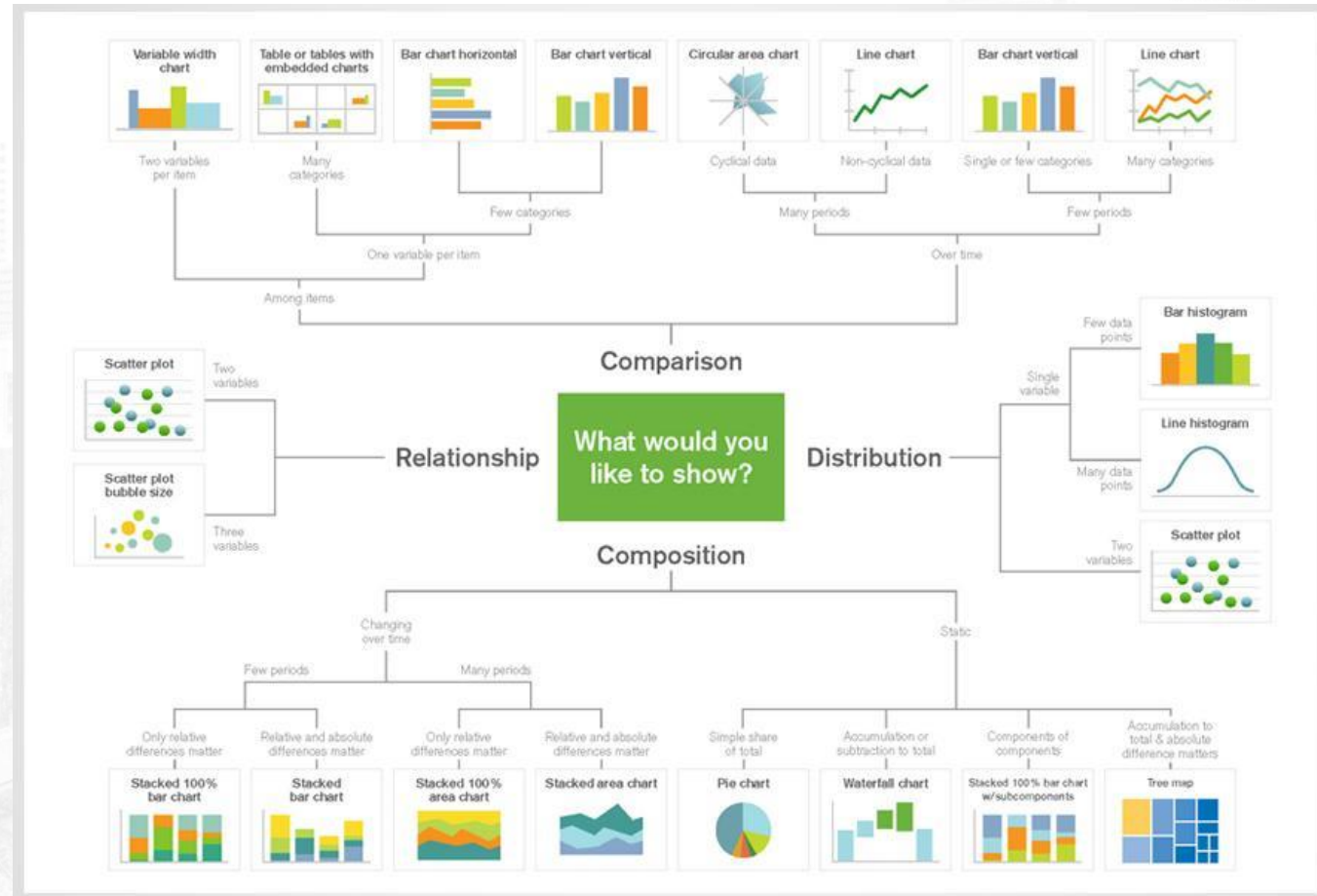
 **Data Storytelling - Part 1**

 **Data Storytelling - Part 2**

# Tipe **Plot** dan **Penggunaannya** Part 1

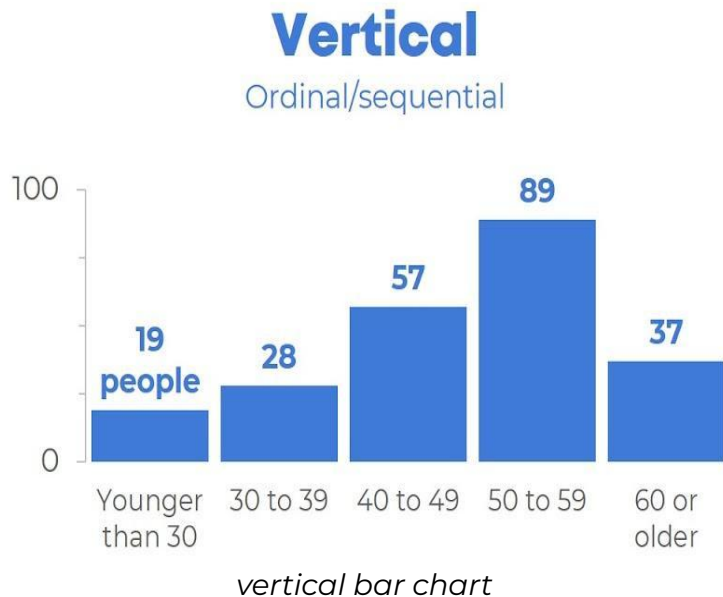
# Penggunaan

- Perbandingan
- Distribusi
- Hubungan
- Komposisi



# Perbandingan antar Item

Digunakan untuk membandingkan item yang **lebih sedikit** ( kurang dari 6 item)



Dapat dilihat dari grafik di samping bahwa jumlah orang terbanyak ada di rentang usia 50 sampai 59 dengan jumlah 89 orang

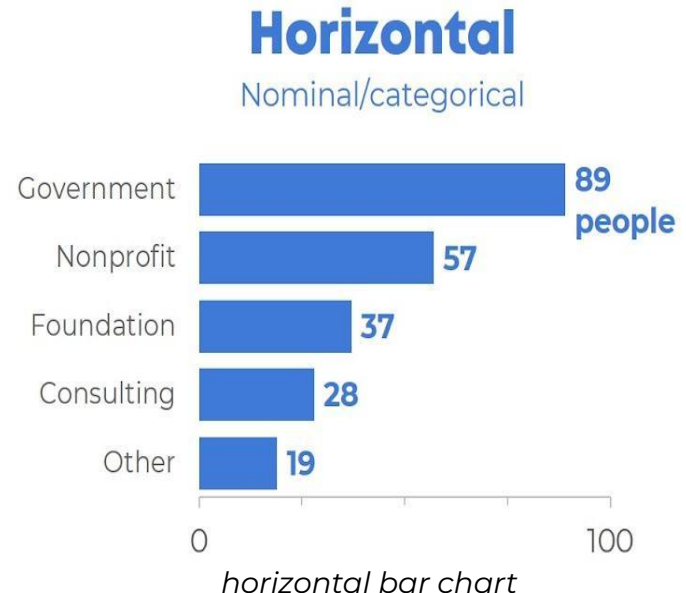
Sedangkan untuk jumlah orang terendah ada di rentang usia kurang dari 30 tahun dengan jumlah 19 orang

# Perbandingan antar Item

Digunakan untuk membandingkan item yang **cukup banyak** ( lebih dari 6 item)

Dapat dilihat dari grafik di samping bahwa jumlah orang terbanyak ada di sektor Government dengan jumlah 89 orang

Sedangkan untuk jumlah orang terendah ada di sektor Other dengan jumlah 19 orang

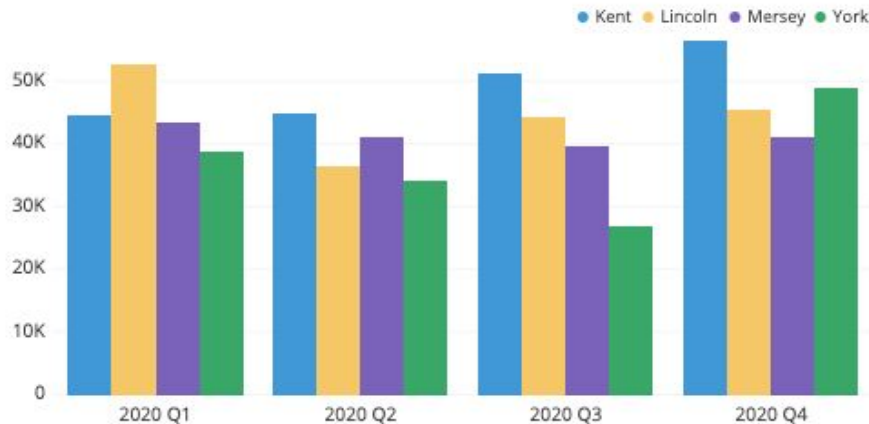




# Perbandingan antar Item terhadap waktu

Digunakan untuk membandingkan item terhadap waktu yang **lebih sedikit** (kurang dari 10 periode waktu)

New Revenue



*multi-bar chart*

Grafik di samping menunjukkan new revenue dari berbagai Kota di US dari tahun 2020 kuartal 1 hingga kuartal 4

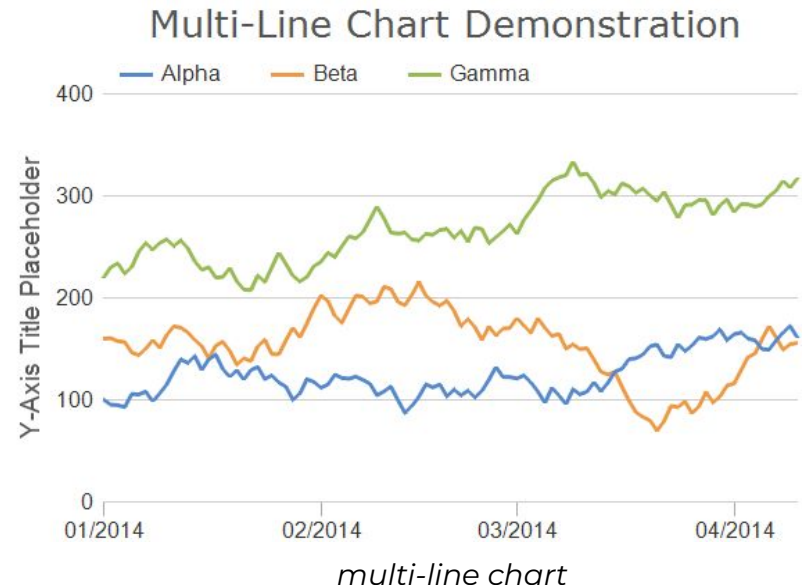
Pada Kuartal 1, revenue tertinggi ada di kota Lincoln sedangkan pada kuartal 4, kota kent merupakan kota dengan revenue terbesar

# Perbandingan antar Item terhadap waktu

Grafik di samping menunjukkan bahwa Gamma mempunyai nilai yang tertinggi dari bulan Januari 2014 hingga April 2014

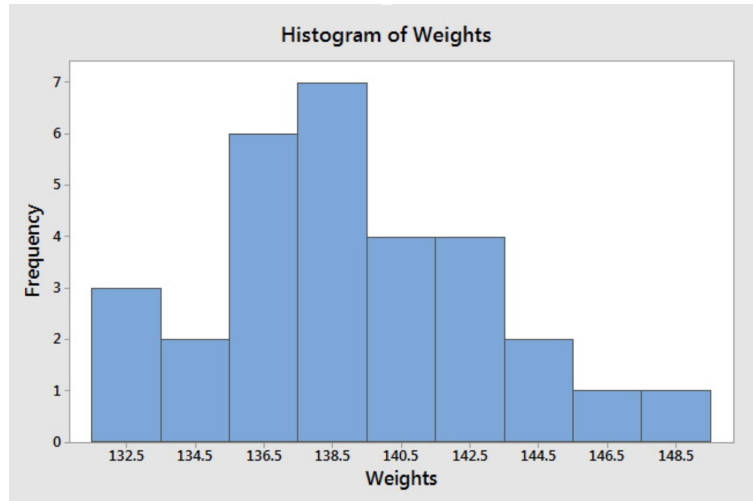
Sedangkan untuk Beta sempat menjadi tertinggi kedua, namun pada pertengahan bulan Maret 2014 turun menjadi urutan ke 3, dan sempat naik sedikit sebentar pada awal April, 2014

Digunakan untuk membandingkan item terhadap waktu yang **cukup banyak** (lebih dari 10 periode waktu)



# Distribusi suatu Item

Plot distribusi konvensional yang bergantung pada jumlah *bin* (banyaknya kotak yang menggambarkan *range* data)



*histogram*

Grafik di samping menunjukkan distribusi dari Weights, dimana untuk nilai tertinggi terdapat pada rentang 137-139 dengan total 7 data.

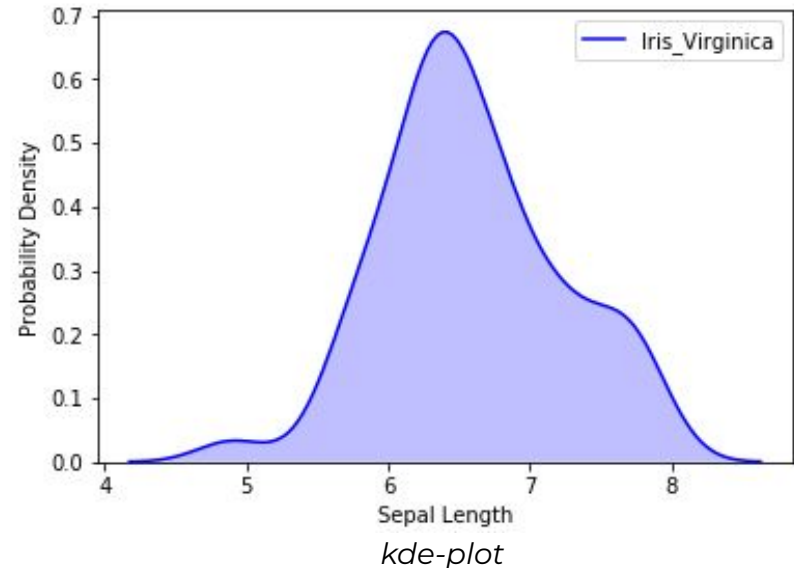
Sebaliknya, nilai terendah terdapat pada rentang 145-149 dengan total 2 data.

# Distribusi suatu Item

Grafik di samping menunjukkan distribusi data Sepal Length dari bunga Iris Virginica yang telah ditransformasi menggunakan teknik *kernel-smoothing*.

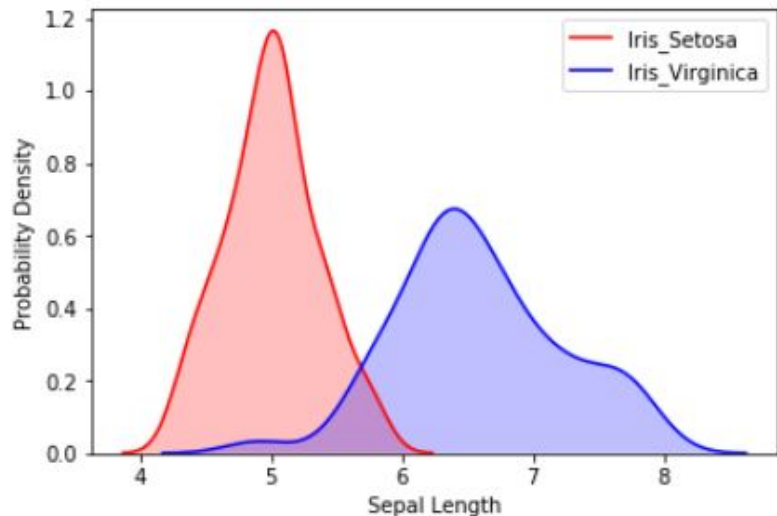
Dapat dilihat bahwa distribusi terbanyak ada di range panjang 6-7 dengan nilai *probability density* 0.7.

Menggunakan [\*kernel-smoothing\*](#), kde-plot lebih baik dalam menentukan bentuk distribusi karena tidak terpengaruh oleh jumlah *bin* yang digunakan



# Distribusi antar Item

Atau kita juga dapat menggunakan **grouped-kdeplot** untuk melihat distribusi antar item



*grouped-kdeplot*

Grafik di samping menunjukkan distribusi data Sepal Length dari bunga Iris Virginica dan bunga Iris Setosa yang telah ditransformasi menggunakan teknik *kernel-smoothing*.

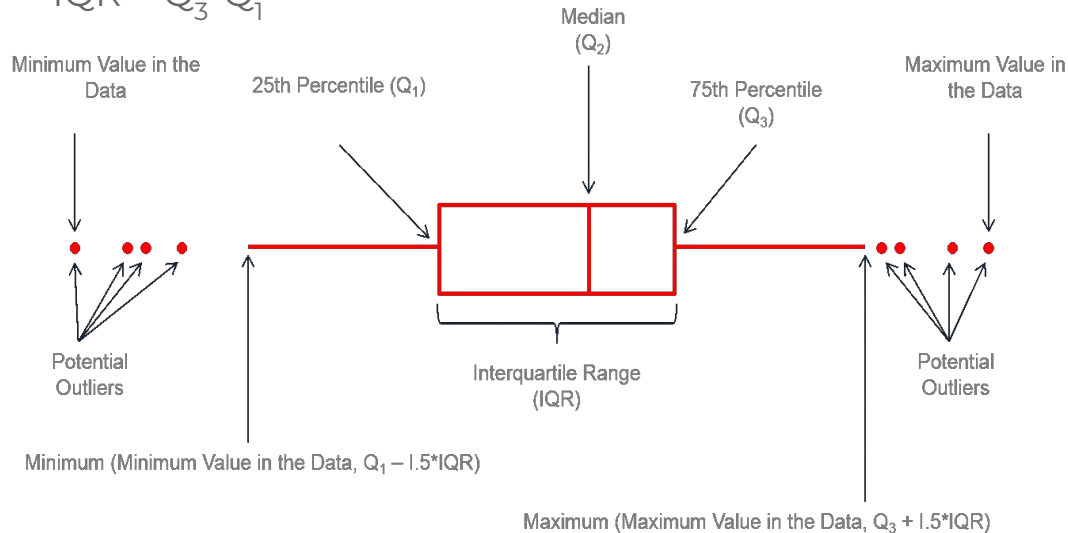
Dapat dilihat bahwa Sepal Length Iris Virginica mayoritas lebih panjang dari Sepal Length Iris Setosa.



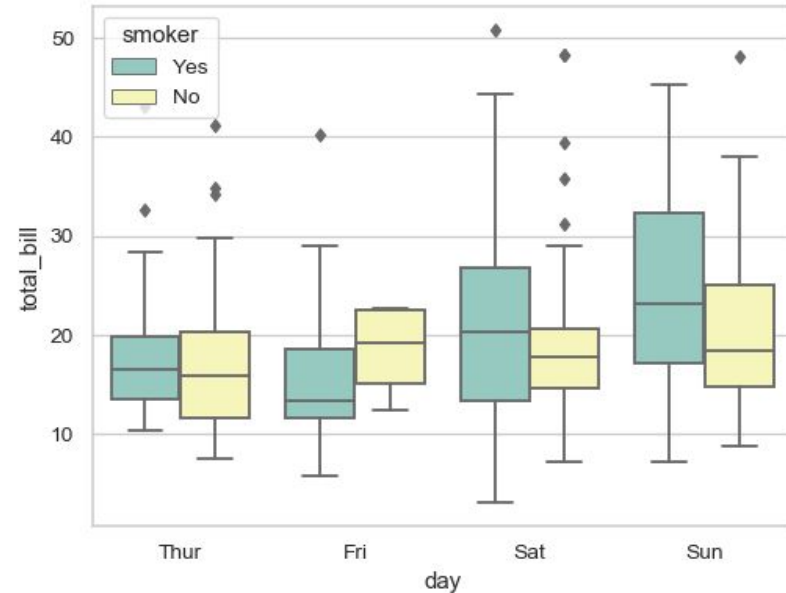
# Distribusi antar Item

Boxplot memberikan deskripsi data yang cukup detail, seperti median, kuantil ke 25 dan 75, serta outlier

$$IQR = Q_3 - Q_1$$



*cara membaca boxplot*

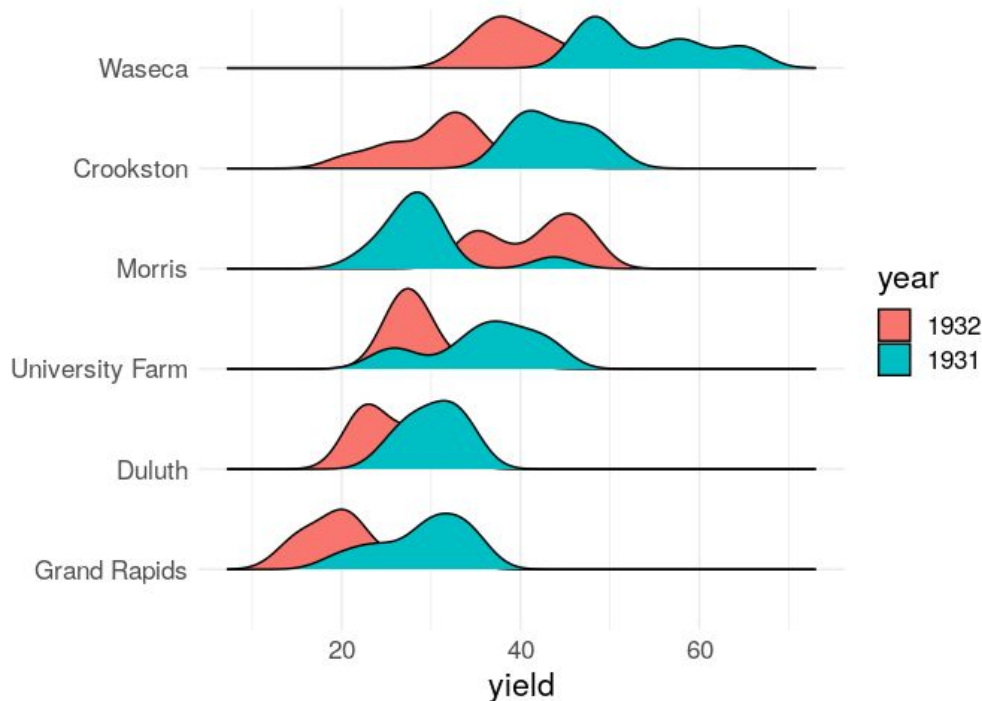


*grouped-boxplot*

# Distribusi antar Item

Grafik di samping menunjukkan distribusi *yield* dari berbagai kota pada tahun 1931 dan 1932.

Dapat dilihat bahwa *yield* terbesar berada di kota Waseca. Sebagian besar *yield* pada 1932 lebih kecil dibanding *yield* pada tahun sebelumnya, kecuali untuk kota Morris.



# Topik Data Visualization

## Sesi : Basic

 **Pengenalan Data Visualization  
(Apa dan Mengapa)**

 **Tipe Plot dan Penggunaannya - Part 1**


 **Tipe Plot dan Penggunaannya - Part 2**

 **Pengenalan pada Matplotlib  
dan Seaborn**

 **Data Visualization Dasar (Praktik) - Part 1**

 **Data Visualization Dasar (Praktik) - Part 2**

## Sesi : Advanced

 Customize Plot (Praktik) - Part 1

 Customize Plot (Praktik) - Part 2

 Data Storytelling - Part 1

 Data Storytelling - Part 2

# Topik Data Visualization

## Sesi : Basic

 **Pengenalan Data Visualization  
(Apa dan Mengapa)**

 **Tipe Plot dan Penggunaannya - Part 1**

  **Tipe Plot dan Penggunaannya - Part 2**

 **Pengenalan pada Matplotlib  
dan Seaborn**

 **Data Visualization Dasar (Praktik) - Part 1**

 **Data Visualization Dasar (Praktik) - Part 2**

## Sesi : Advanced

 **Customize Plot (Praktik) - Part 1**

 **Customize Plot (Praktik) - Part 2**

 **Data Storytelling - Part 1**

 **Data Storytelling - Part 2**

# Tipe **Plot** dan **Penggunaannya** Part 2

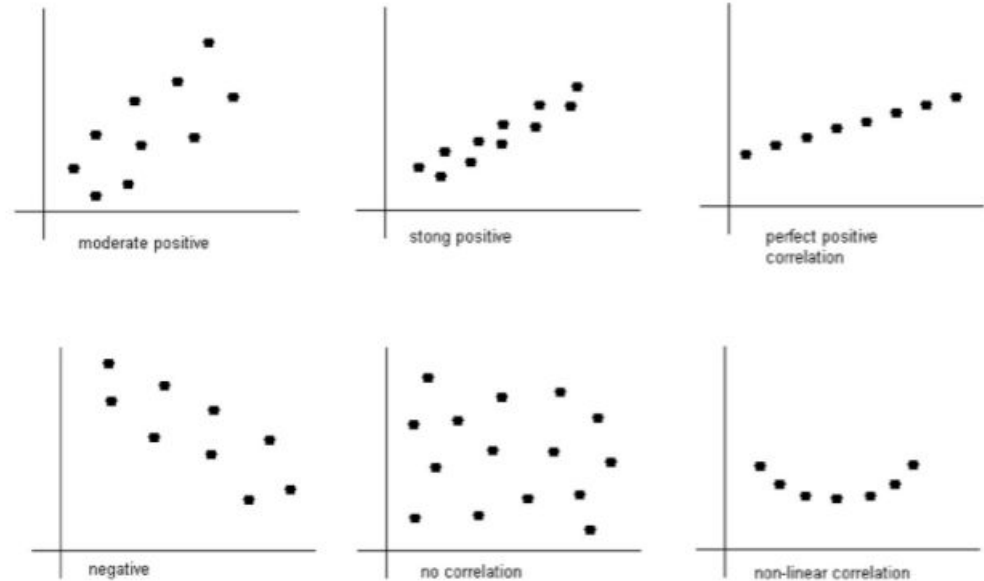


# Hubungan antar Item

Dengan menggunakan scatter plot kita dapat melihat korelasi atau hubungan antar item



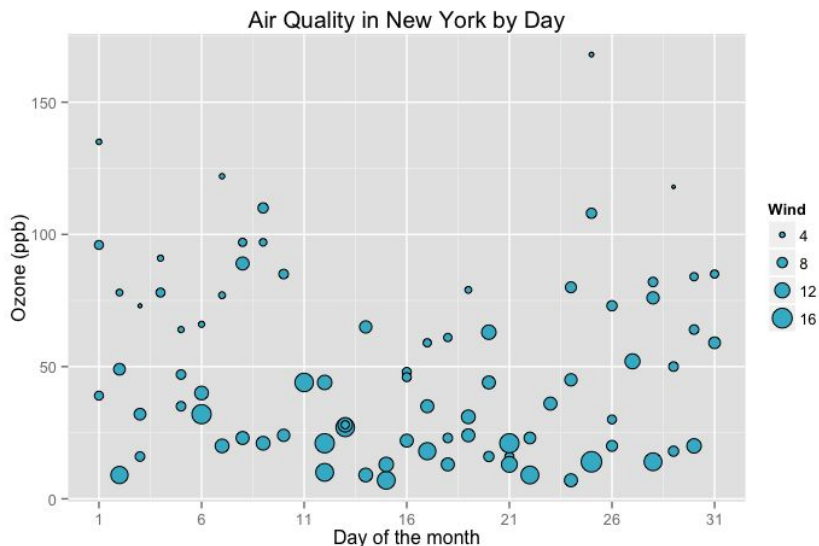
*scatter plot*



*correlation in scatter plot*

# Hubungan antar beberapa Item

Memainkan **ukuran** dari data point untuk mencari hubungan terhadap item lain



*bubble scatter plot*

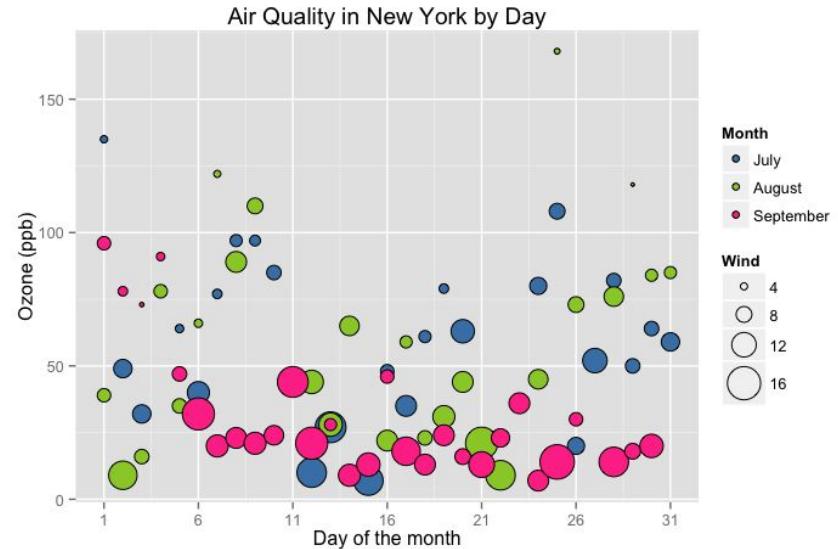
Grafik di samping menunjukkan persebaran tingkat Ozone dari harian dalam satu bulan dibandingkan dengan kekuatan anginnya, dimana semakin besar *point* maka semakin besar kekuatan angin.

Dilihat dari grafik di samping, semakin rendah tingkat Ozone maka semakin tinggi kekuatan angin.

# Hubungan antar **beberapa** Item

Memainkan **warna** dari data point untuk mencari hubungan terhadap item lain

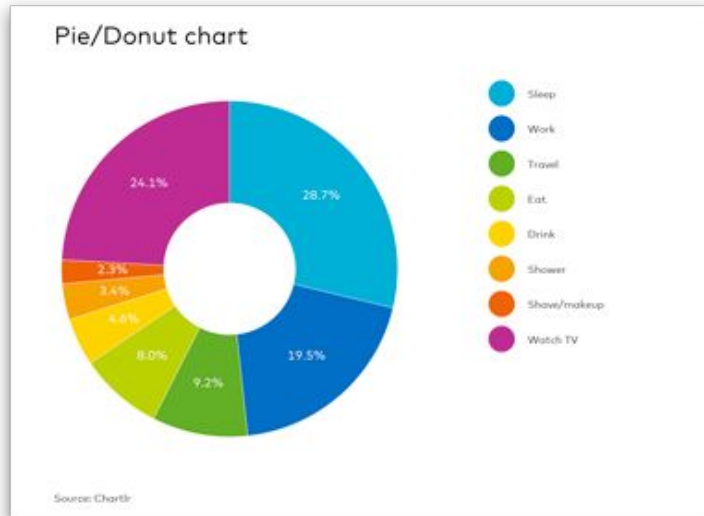
Sama dengan grafik diatas, namun pada grafik ini ditambahkan satu fitur lagi yaitu fitur bulan yang direpresentasikan dengan warna.



*bubble scatter plot*

# Komposisi antar Item

Cocok digunakan untuk menjelaskan kontribusi dari setiap item, namun banyak *expert* yang tidak menyarankan



Grafik di samping menunjukkan aktivitas terbanyak adalah sleep dengan nilai 28.7%, namun *pie chart* ini sebaiknya digunakan untuk data yang tidak lebih dari 5.

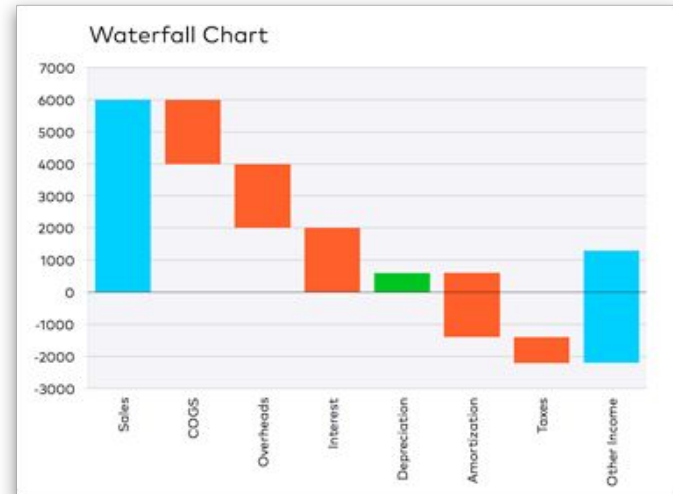
*pie chart*

# Komposisi antar Item

Grafik di samping menunjukkan bagaimana suatu nilai awal terpengaruh dari rangkaian nilai positif ataupun negatif.

Nilai awal dari sales yang bernilai 6000 mengalami pengurangan dari biaya COGS, *overheads*, *interest*, *amortisation*, dan *taxes*. Dan, terdapat penambahan nilai dari *depreciation* dan *other income*.

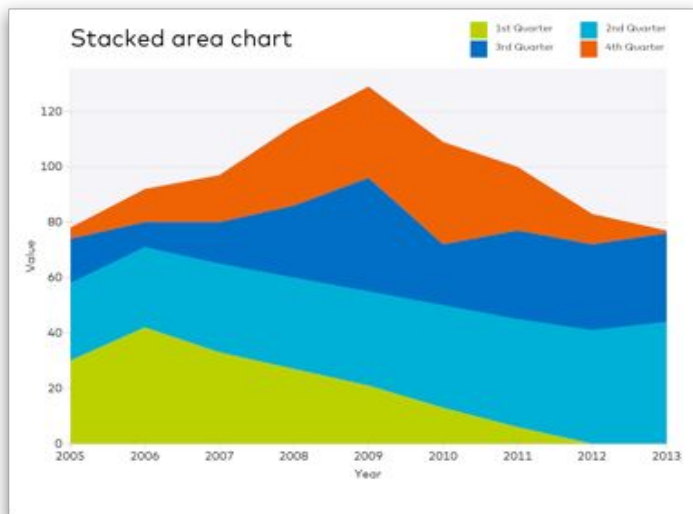
Cocok digunakan untuk melihat perubahan komposisi terhadap suatu event



waterfall chart

# Komposisi antar Item terhadap waktu

Grafik ini cocok digunakan untuk melihat komposisi dari beberapa item terhadap waktu, tergantung preferensi masing-masing



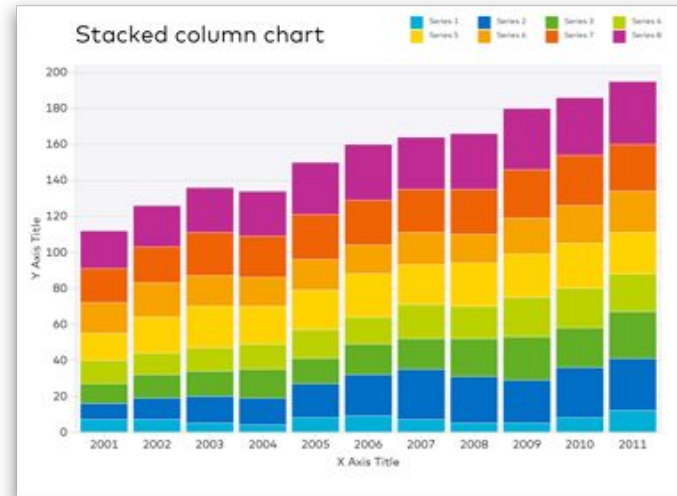
*stacked area chart*

Grafik di samping menunjukkan value dari tahun 2005-2013 yang dibagi menjadi 4 quarter.

Dapat terlihat bahwa pada awal tahun 2005 di quarter pertama memiliki value terbesar. Sedangkan, pada akhir tahun 2013, value dari quarter 1 adalah 0; value terbesar ada pada quarter 2.

# Komposisi antar Item terhadap waktu

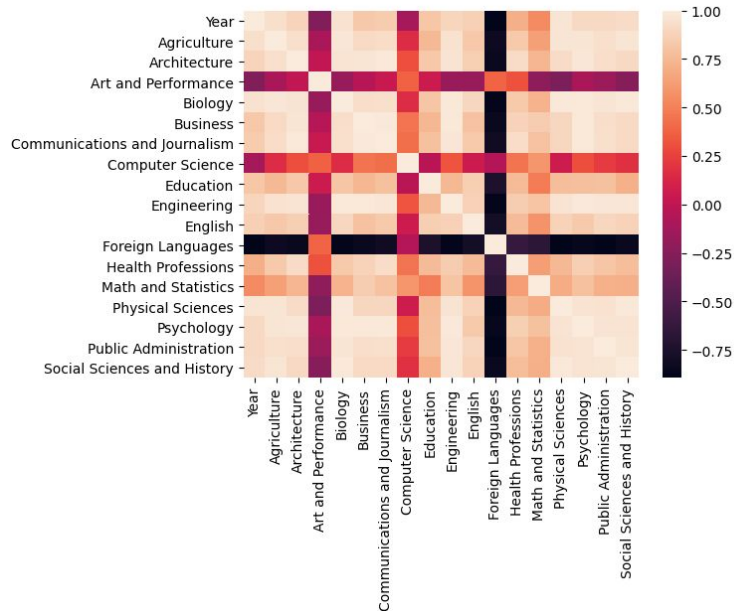
Selain grafik diatas, grafik ini cocok digunakan untuk melihat komposisi dari beberapa item terhadap waktu, namun menggunakan column chart.



*stacked column chart*



# Jenis chart **lainnya**



*heatmap chart*

Biasa digunakan untuk melihat **korelasi** atau **hubungan** antara variabel dengan variabel lainnya.

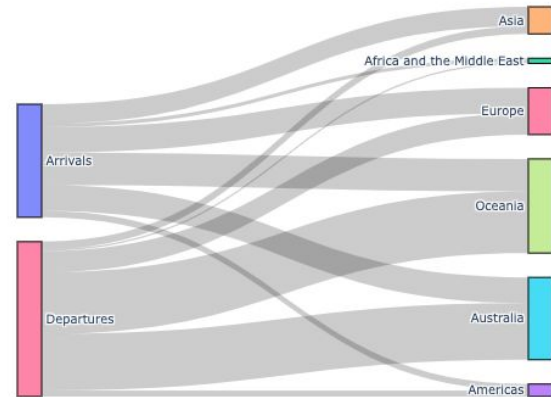
Terlihat bahwa semakin gelap warna chart maka dua fitur tersebut semakin berkorelasi negatif. Sebaliknya, semakin terang warna chart maka semakin berkorelasi positif.

# Jenis chart **lainnya**

Bertujuan untuk menggambarkan **aliran** dari satu set nilai ke nilai yang lain.

Terlihat grafik mengenai migrasi populasi antara New Zealand dan benua lainnya. Sebagian besar kedatangan warga NZ berasal dari Europe, Oceania, dan Australia. Sedangkan, sebagian besar kepergian warga NZ menuju Oceania dan Australia.

Population Migration between New Zealand and Other Continents



sankey chart

# Topik Data Visualization

## Sesi : Basic

 **Pengenalan Data Visualization  
(Apa dan Mengapa)**

 **Tipe Plot dan Penggunaannya - Part 1**


 **Tipe Plot dan Penggunaannya - Part 2**

 **Pengenalan pada Matplotlib  
dan Seaborn**

 **Data Visualization Dasar (Praktik) - Part 1**

 **Data Visualization Dasar (Praktik) - Part 2**

## Sesi : Advanced

 **Customize Plot (Praktik) - Part 1**

 **Customize Plot (Praktik) - Part 2**

 **Data Storytelling - Part 1**

 **Data Storytelling - Part 2**

# Topik Data Visualization

## Sesi : Basic

 **Pengenalan Data Visualization  
(Apa dan Mengapa)**

 **Tipe Plot dan Penggunaannya - Part 1**

 **Tipe Plot dan Penggunaannya - Part 2**

  **Pengenalan pada Matplotlib  
dan Seaborn**

 **Data Visualization Dasar (Praktik) - Part 1**

 **Data Visualization Dasar (Praktik) - Part 2**

## Sesi : Advanced

 **Customize Plot (Praktik) - Part 1**

 **Customize Plot (Praktik) - Part 2**

 **Data Storytelling - Part 1**

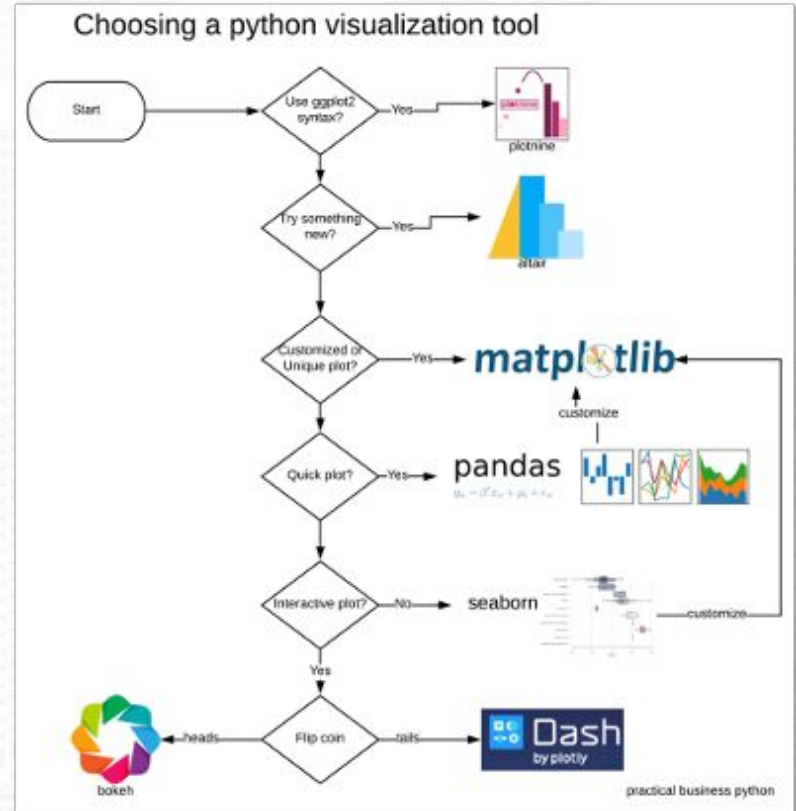
 **Data Storytelling - Part 2**

# **Matplotlib dan Seaborn**

# Data Visualization Library

Ini adalah beberapa *library* visualisasi di python yang sering digunakan

- **Matplotlib:** paling basic, punya tingkat kebebasan yang tinggi
- **Pandas Visualization:** interface yang sederhana, terintegrasi dengan matplotlib
- **Seaborn:** high-level interface, plot-plot statistika
- **ggplot:** berbasis dari ggplot2 bahasa pemrograman R
- **Plotly / Bokeh:** bisa membuat plot interaktif

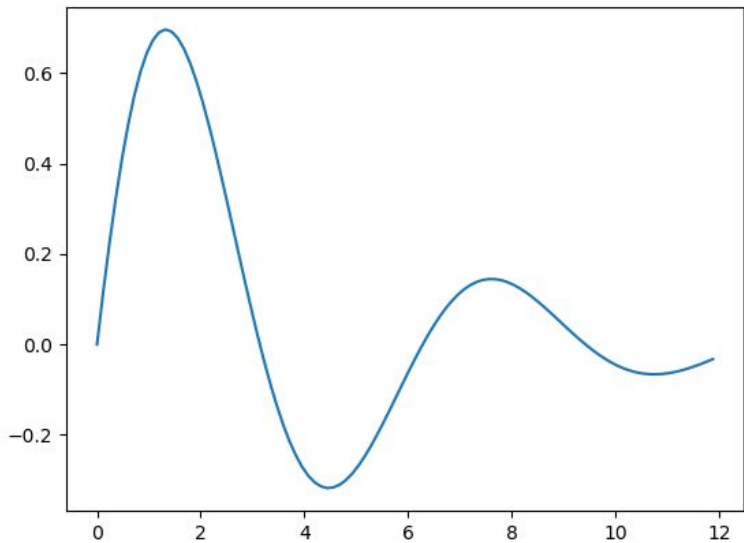


# Matplotlib vs Seaborn

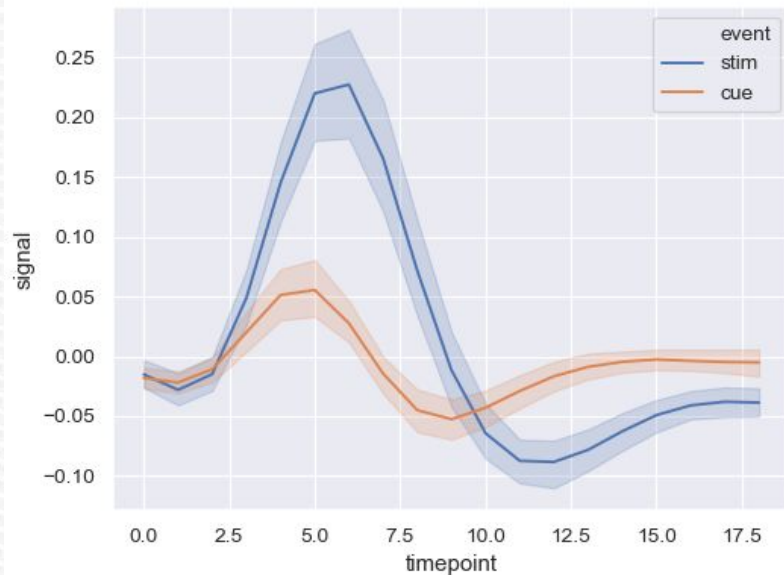
Fitur	Matplotlib	Seaborn
Fungsionalitas	Dibuat untuk plot-plot dasar, berbasis fitur dan sintaks dari MATLAB	Terdapat banyak tipe plot, sangat baik terutama untuk statistika.
Sintaks	Lebih kompleks dan panjang	Sederhana dan lebih mudah dipelajari
Dataframe	Terintegrasi baik dengan Numpy dan Pandas	Sangat baik digunakan dengan Pandas
Fleksibilitas	Sangat bebas (fleksibel)	Dalam beberapa kasus, cukup terbatas
Penggunaan	Digunakan untuk membuat atau memodifikasi plot-plot dasar	Extended version dari Matplotlib, memiliki banyak tipe plot dan tema



# Matplotlib vs Seaborn

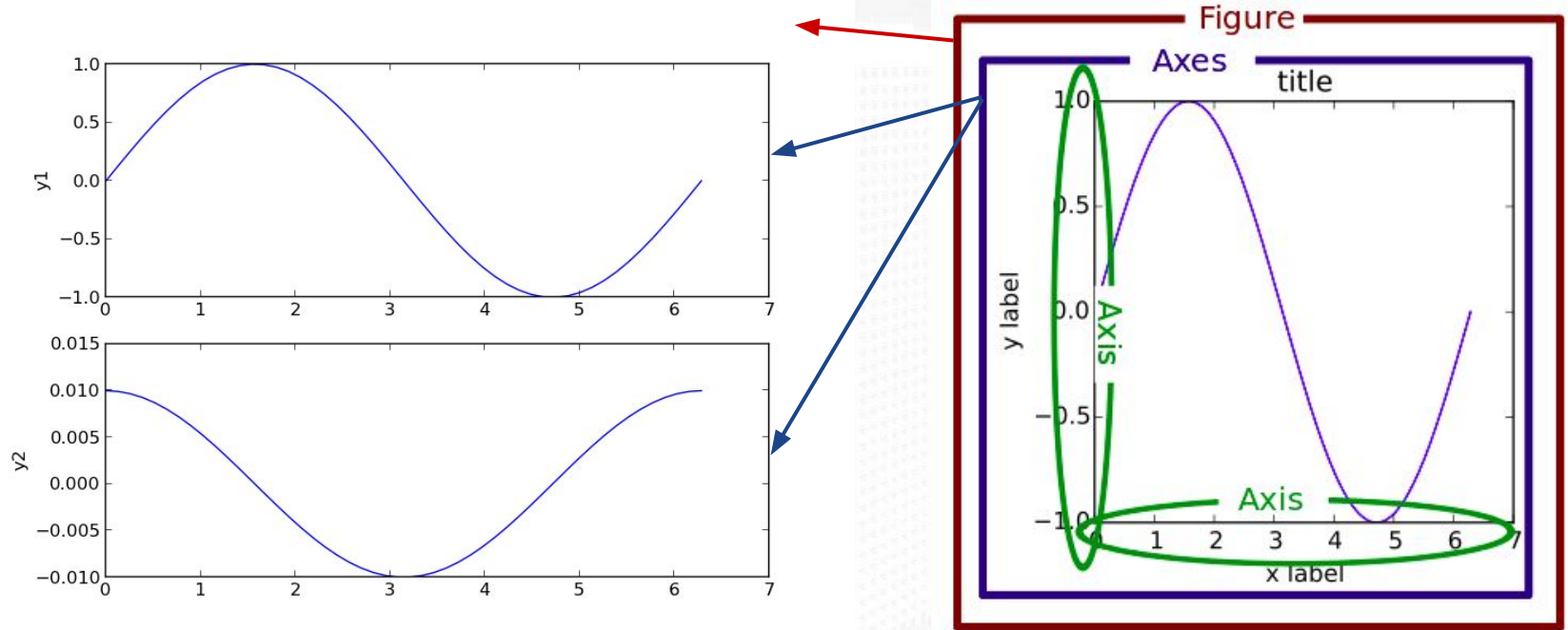


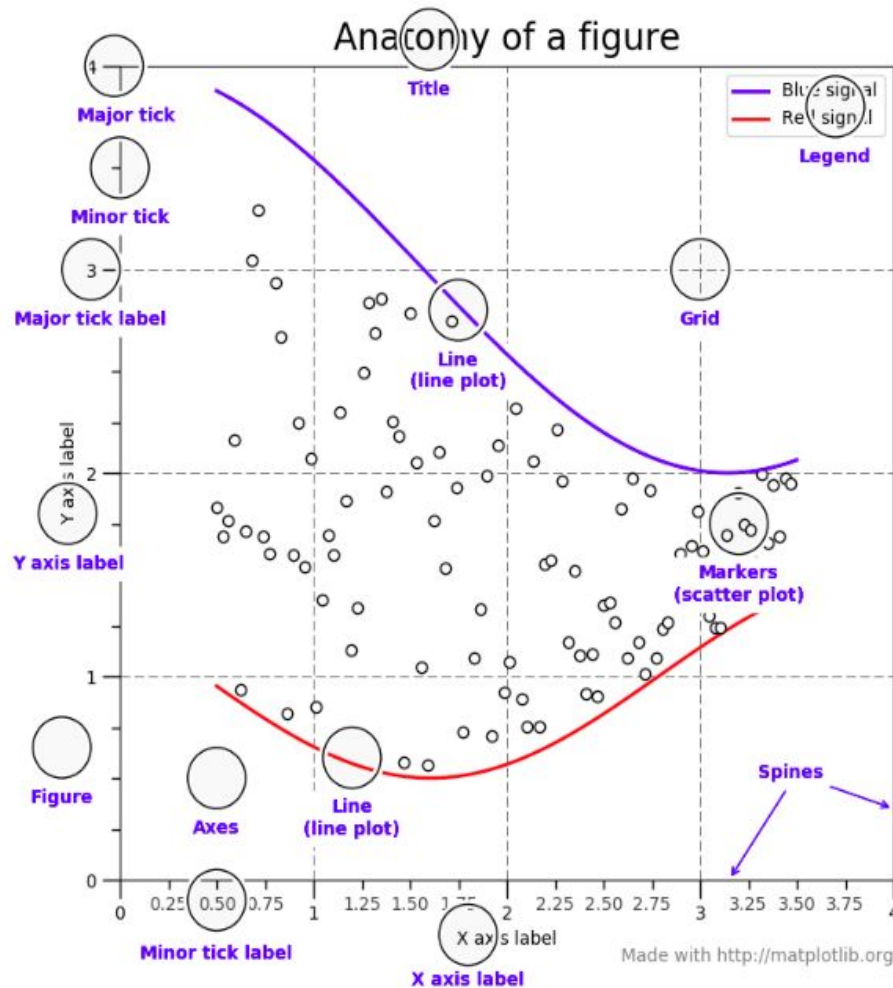
*matplotlib*



*seaborn*

# Anatomi Matplotlib





Matplotlib actually create this figure  
with python code!

# Topik Data Visualization

## Sesi : Basic

 **Pengenalan Data Visualization  
(Apa dan Mengapa)**

 **Tipe Plot dan Penggunaannya - Part 1**


 **Tipe Plot dan Penggunaannya - Part 2**

 **Pengenalan pada Matplotlib  
dan Seaborn**

 **Data Visualization Dasar (Praktik) - Part 1**

 **Data Visualization Dasar (Praktik) - Part 2**

## Sesi : Advanced

 **Customize Plot (Praktik) - Part 1**

 **Customize Plot (Praktik) - Part 2**

 **Data Storytelling - Part 1**

 **Data Storytelling - Part 2**

# Topik Data Visualization

## Sesi : Basic

 **Pengenalan Data Visualization  
(Apa dan Mengapa)**

 **Tipe Plot dan Penggunaannya - Part 1**

 **Tipe Plot dan Penggunaannya - Part 2**

 **Pengenalan pada Matplotlib  
dan Seaborn**

  **Data Visualization Dasar (Praktik) - Part 1**

 **Data Visualization Dasar (Praktik) - Part 2**

## Sesi : Advanced

 **Customize Plot (Praktik) - Part 1**

 **Customize Plot (Praktik) - Part 2**

 **Data Storytelling - Part 1**

 **Data Storytelling - Part 2**

# **Visualisasi Dasar**

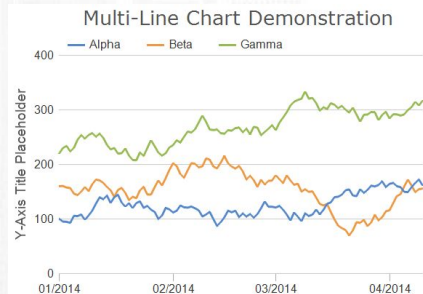
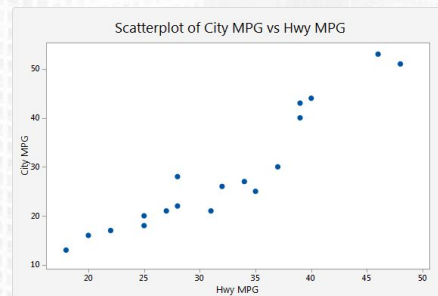
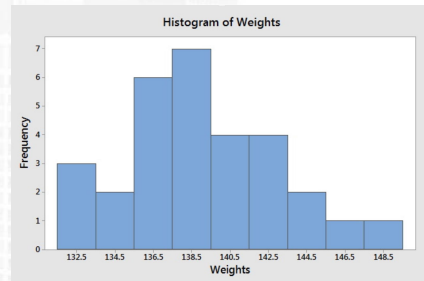
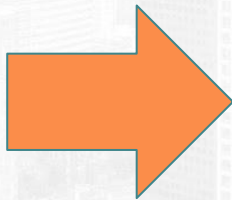
## Saatnya Praktik!



# Saatnya Praktik!

matplotlib  
Seaborn

pandas



Membandingkan 3 *library* yang berbeda untuk berbagai tipe plot!



# Dataset yang digunakan

## Telco Customer Churn

- **Deskripsi:**

Dataset ini menggambarkan perilaku dan profil pelanggan di perusahaan Telco yang digunakan guna menganalisis dan memprediksi retensi pelanggan.

- **Data:**

Setiap baris mewakili pelanggan, setiap kolom berisi atribut pelanggan.

- **Link download [disini](#)**

# Import Library

Pandas	Matplotlib	Seaborn
import pandas as pd	import matplotlib.pyplot as plt	import seaborn as sns
<pre> 1 import pandas as pd 2 pd.__version__ </pre> <p>executed in 6ms, finished 20:16:57 2020-08</p> <p>'0.23.4'</p>	<pre> 1 import matplotlib 2 import matplotlib.pyplot as plt 3 4 matplotlib.__version__ </pre> <p>executed in 5ms, finished 20:17:00 2020-08-08</p> <p>'3.0.3'</p>	<pre> 1 import seaborn as sns 2 sns.__version__ </pre> <p>executed in 455ms, finished 20:17:15 2020-08-08</p> <p>'0.9.0'</p>

# Membaca (*Read*) Dataset

```
1 # read csv
2 df = pd.read_csv('csv/telco_customer.csv')
3 df.head()
```

executed in 87ms, finished 20:14:05 2020-08-08

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	DeviceProtection	Tech
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	...	No	
1	5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	...	Yes	
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	...	No	
3	7795-CFOCW	Male	0	No	No	45	No	No phone service	DSL	Yes	...	Yes	
4	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	...	No	

5 rows x 21 columns

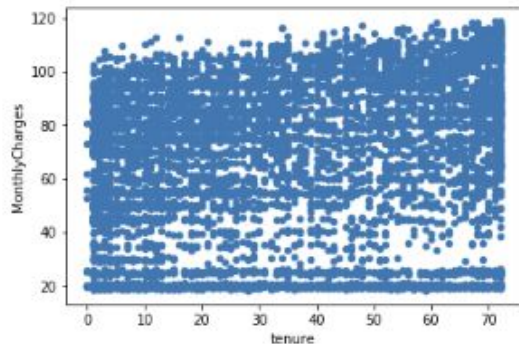
# Scatter Plot: Mencari hubungan antara masa berlangganan dan biaya bulanan

Pandas

`df.plot(x,y, kind='scatter')`

```
1 df.plot(x = 'tenure', y= 'MonthlyCharges',
2         kind='scatter');
```

executed in 262ms, finished 20:14:58 2020-08-08

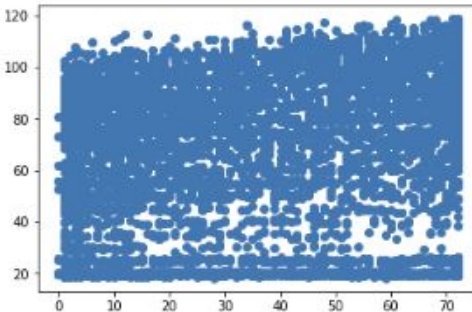


Matplotlib

`plt.scatter(x,y, data=df)`

```
1 plt.scatter(x = 'tenure', y= 'MonthlyCharges',
2            data=df);
```

executed in 315ms, finished 20:18:31 2020-08-08

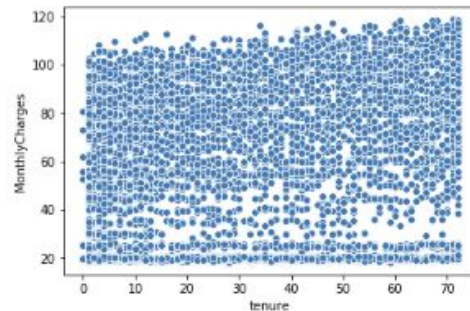


Seaborn

`sns.scatterplot(x, y, data=df)`

```
1 sns.scatterplot(x = 'tenure', y= 'MonthlyCharges',
2                data=df);
```

executed in 284ms, finished 20:18:54 2020-08-08



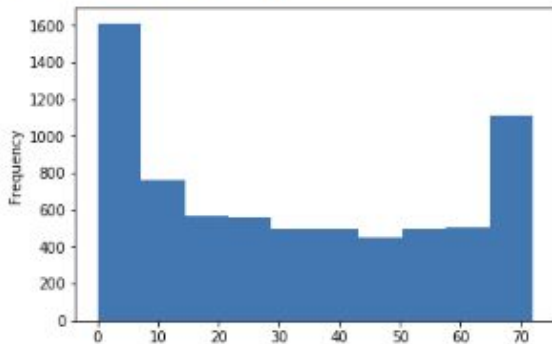
# Histogram: Mencari distribusi dari masa berlangganan

## Pandas

`df[x].plot(kind='hist')`

```
1 df['tenure'].plot(kind='hist');
```

executed in 247ms, finished 20:23:23 2020-08-08

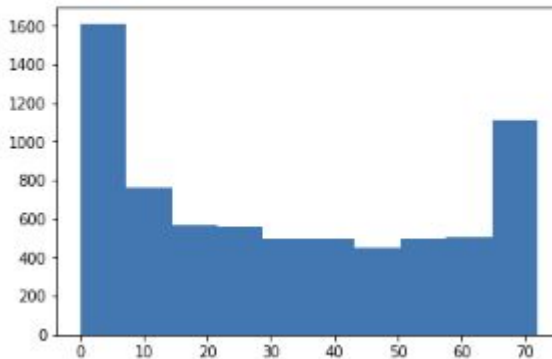


## Matplotlib

`plt.hist(x, data=df)`

```
1 plt.hist(x = 'tenure', data=df);
```

executed in 242ms, finished 20:24:08 2020-08-08

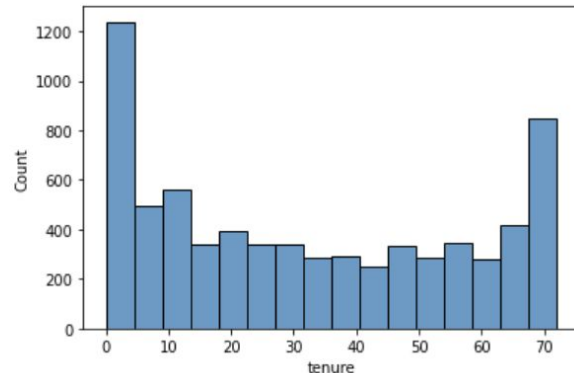


## Seaborn

`sns.histplot(x, data=df)`

```
1 sns.histplot(x='tenure', data=df);
```

executed in 382ms, finished 16:42:54 2021-06-26





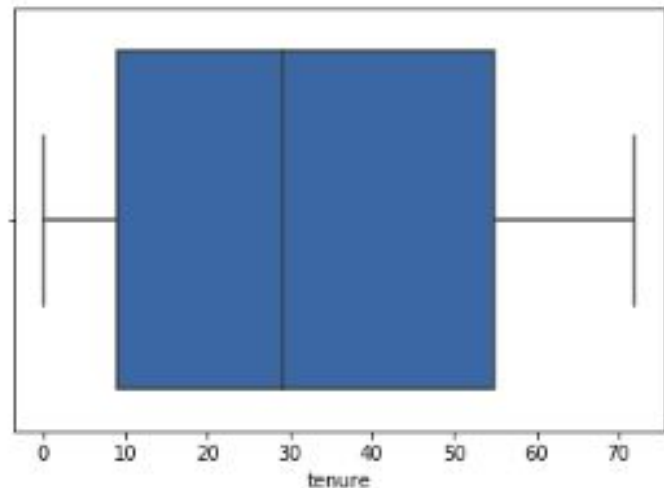
# Boxplot: Mencari distribusi dari masa berlangganan

Seaborn one variable

`sns.boxplot(x, data=df)`

```
1 sns.boxplot(x='tenure', data=df);
```

executed in 163ms, finished 22:15:39 2020-08-11

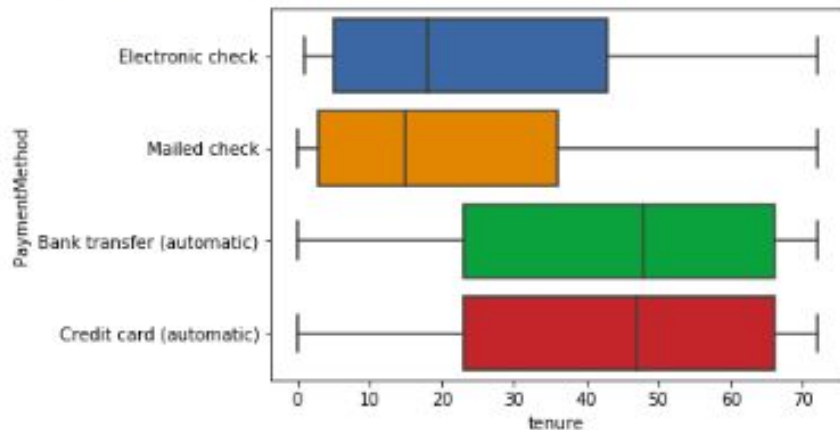


Seaborn two variable

`sns.boxplot(x, y, data=df)`

```
1 sns.boxplot(x='tenure', y='PaymentMethod', data=df);
```

executed in 209ms, finished 22:12:44 2020-08-11



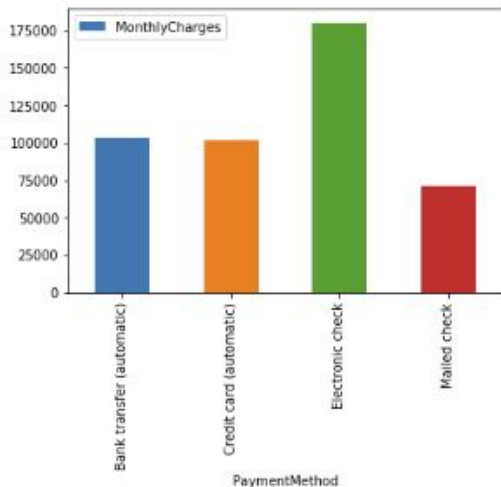
# Barplot: Mencari rata-rata biaya bulanan untuk setiap metode pembayaran

## Pandas

`df.plot(x=y, kind='bar')`

```
1 df.plot(x='PaymentMethod', y='MonthlyCharges',
2        kind='bar');
```

executed in 190ms, finished 20:31:06 2020-08-08

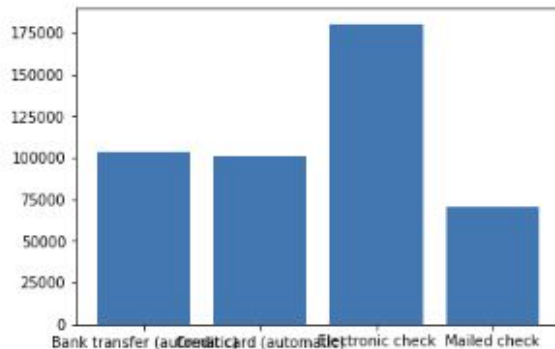


## Matplotlib

`plt.bar(x,height, data=df)`

```
1 plt.bar(x = 'PaymentMethod',
2        height= 'MonthlyCharges',
3        data=dfg);
```

executed in 205ms, finished 20:34:18 2020-08-08

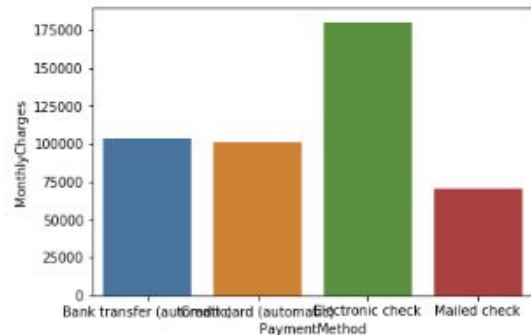


## Seaborn

`sns.barplot(x, y, data=df)`

```
1 sns.barplot(x = 'PaymentMethod',
2            y= 'MonthlyCharges',
3            data=dfg);
```

executed in 228ms, finished 21:37:34 2020-08-08





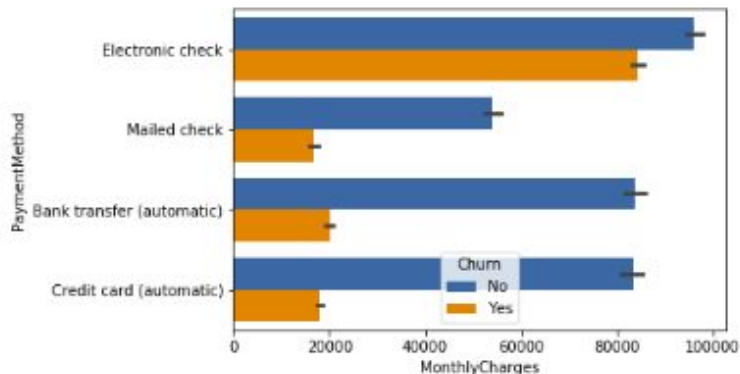
# Multi-barplot: Mencari rata-rata biaya bulanan untuk setiap metode pembayaran dan berdasarkan *churn*

Multi-Bar plot

`sns.barplot(df[x], df[y], hue=df[z])`

```
1 sns.barplot(x='MonthlyCharges', y='PaymentMethod', data=df,
2             estimator=sum, hue='Churn');
```

executed in 954ms, finished 22:33:44 2020-08-11

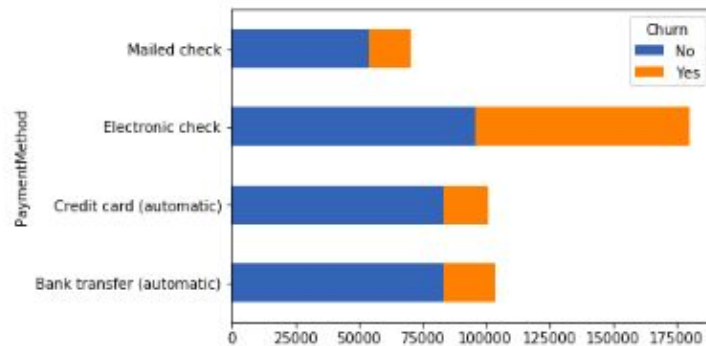


Stacked Multi-Bar plot

`df.plot(kind='bar',stacked=True)`

```
1 dfgc = df.groupby(['PaymentMethod', 'Churn'])['MonthlyCharges']\
2       .sum().reset_index()
3 dfp = dfgc.pivot_table(index='PaymentMethod', columns='Churn',
4                         values='MonthlyCharges')
5 dfp.plot(kind='barh', stacked=True);
```

executed in 203ms, finished 22:51:25 2020-08-11



# Topik Data Visualization

## Sesi : Basic

 **Pengenalan Data Visualization  
(Apa dan Mengapa)**

 **Tipe Plot dan Penggunaannya - Part 1**

 **Tipe Plot dan Penggunaannya - Part 2**

 **Pengenalan pada Matplotlib  
dan Seaborn**

 **Data Visualization Dasar (Praktik) - Part 1**

 **Data Visualization Dasar (Praktik) - Part 2**

## Sesi : Advanced

 **Customize Plot (Praktik) - Part 1**

 **Customize Plot (Praktik) - Part 2**

 **Data Storytelling - Part 1**

 **Data Storytelling - Part 2**

# Topik Data Visualization

## Sesi : Basic

 **Pengenalan Data Visualization  
(Apa dan Mengapa)**

 **Tipe Plot dan Penggunaannya - Part 1**


 **Tipe Plot dan Penggunaannya - Part 2**

 **Pengenalan pada Matplotlib  
dan Seaborn**

 **Data Visualization Dasar (Praktik) - Part 1**

  **Data Visualization Dasar (Praktik) - Part 2**

## Sesi : Advanced

 Customize Plot (Praktik) - Part 1

 Customize Plot (Praktik) - Part 2

 Data Storytelling - Part 1

 Data Storytelling - Part 2

# Topik Data Visualization

## Sesi : Basic

 **Pengenalan Data Visualization  
(Apa dan Mengapa)**

 **Tipe Plot dan Penggunaannya - Part 1**

 **Tipe Plot dan Penggunaannya - Part 2**

 **Pengenalan pada Matplotlib  
dan Seaborn**

 **Data Visualization Dasar (Praktik) - Part 1**

 **Data Visualization Dasar (Praktik) - Part 2**

## Sesi : Advanced

 Customize Plot (Praktik) - Part 1

 Customize Plot (Praktik) - Part 2

 Data Storytelling - Part 1

 Data Storytelling - Part 2



# Terima Kasih!



Aussa Tris Rahmatullah

<https://id.linkedin.com/in/aussa-tris-rahmatullah>

**Aussa Rahmatullah**

Data Scientist

Telco Company