



Lecture 30 Visualisasi (Part 1)

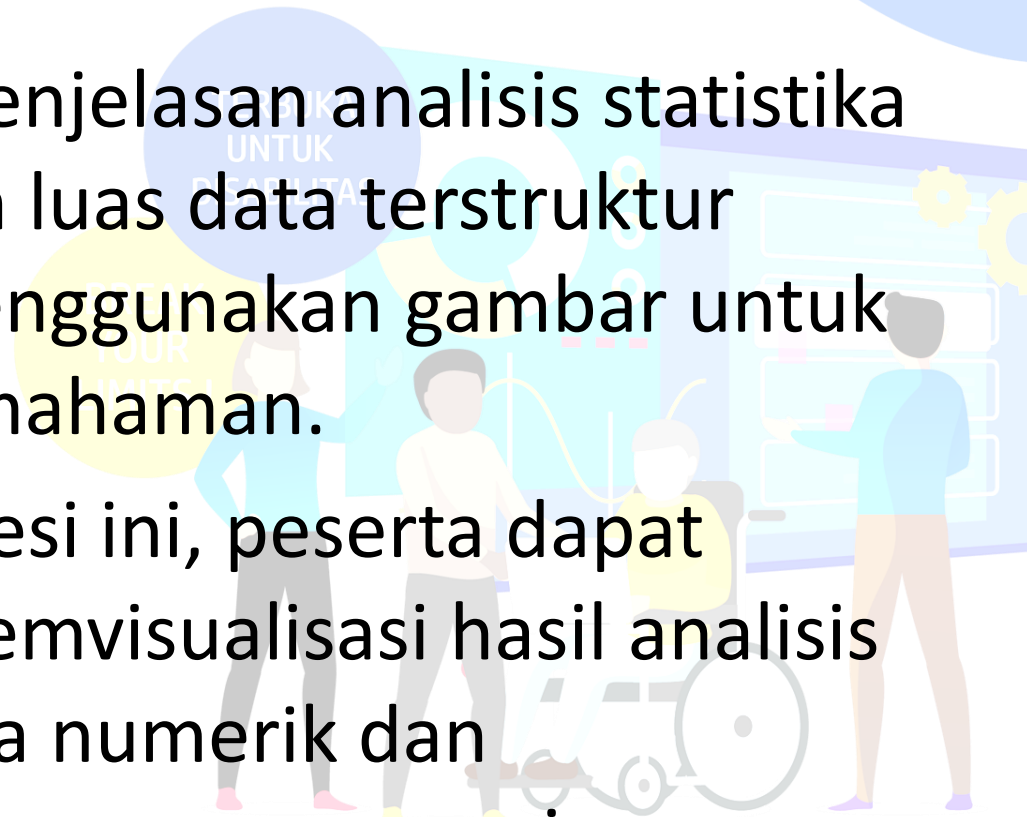




DIGITAL
TALENT
SCHOLARSHIP

Lecture's Objective

- Mempelajari cara penjelasan analisis statistika kuantitatif dan lebih luas data terstruktur (berbasis angka) menggunakan gambar untuk mempermudah pemahaman.
- Setelah mengikuti sesi ini, peserta dapat mengetahui cara memvisualisasi hasil analisis atau model dari data numerik dan menerapkan pada kasus yang sesuai.





DIGITAL
TALENT
SCHOLARSHIP

Outline

- 60 menit:
 - Data visualization
 - Use case for data visualization
- 140 menit:
 - Mencoba source code untuk suatu contoh kasus.

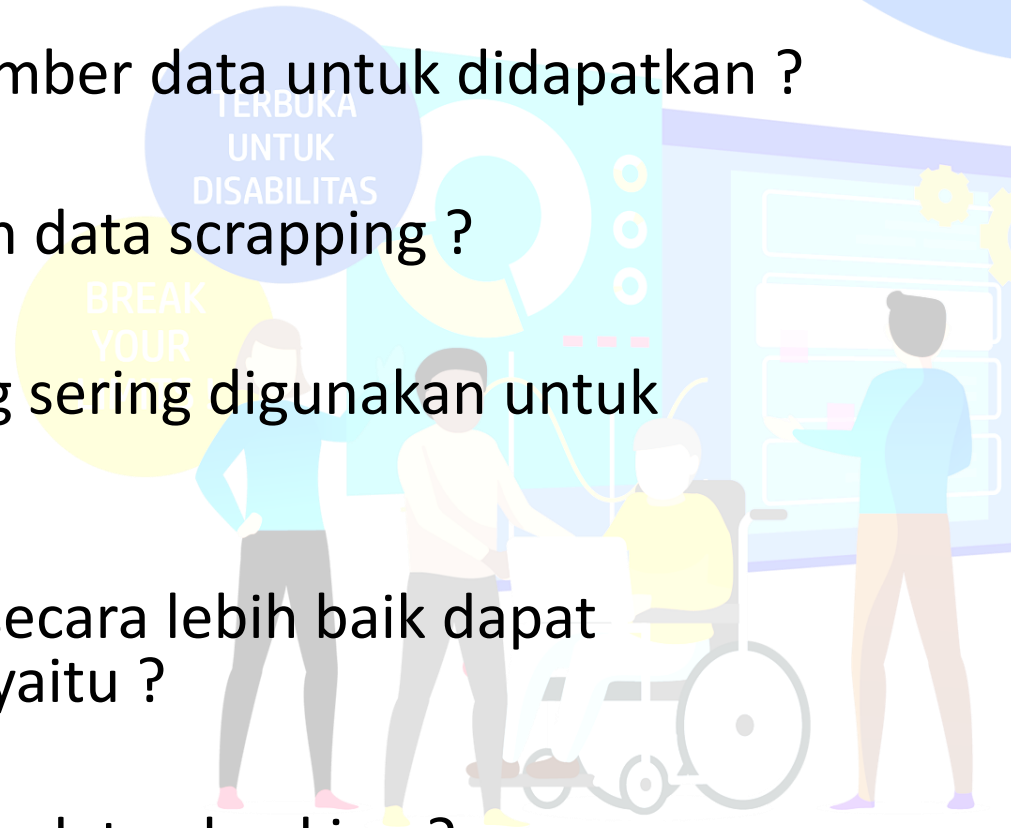




DIGITAL
TALENT
SCHOLARSHIP

Evaluasi Persiapan Kelas

- Berikan contoh sumber-sumber data untuk didapatkan ?
- Apa yang dimaksud dengan data scrapping ?
- Sebutkan tools umum yang sering digunakan untuk mengolah data ?
- Untuk menampilkan data secara lebih baik dapat menggunakan suatu cara, yaitu ?
- Apa yang dimaksud dengan data-checking ?



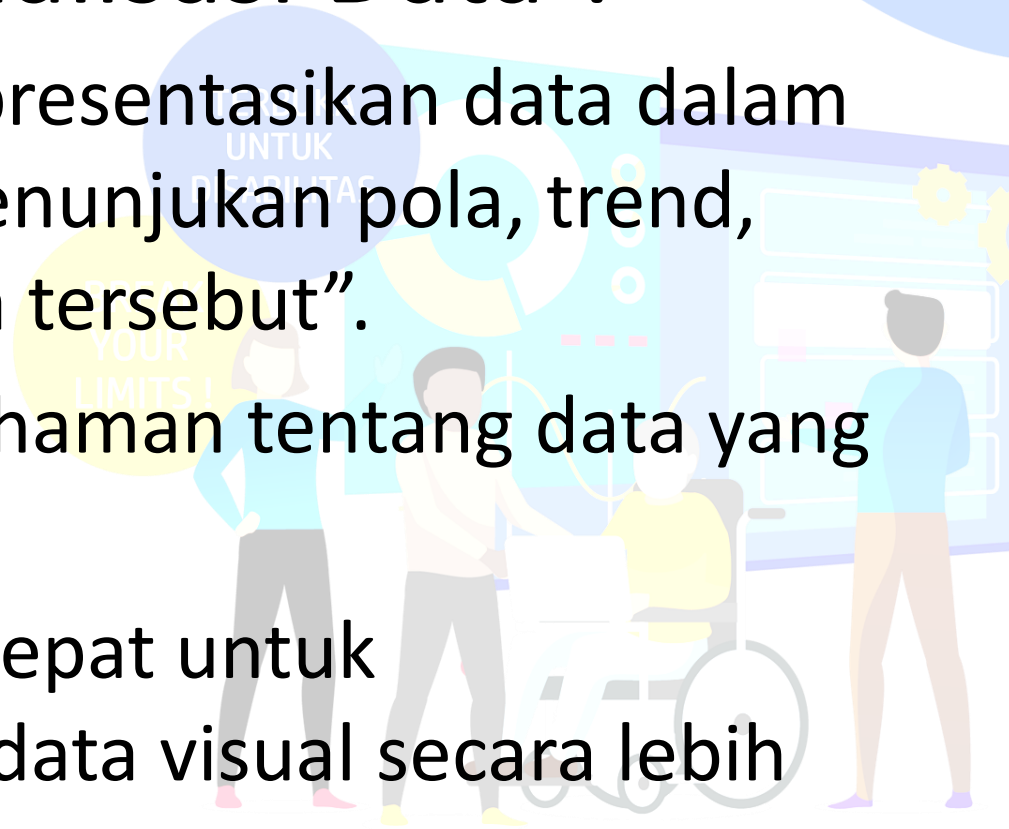


DIGITAL
TALENT
SCHOLARSHIP

Apa itu Visualisasi Data ?

“ Sebuah proses mempresentasikan data dalam bentuk grafik untuk menunjukkan pola, trend, dan informasi dari data tersebut”.

- Meningkatkan pemahaman tentang data yang diberikan.
- Otak manusia lebih cepat untuk menginterpretasikan data visual secara lebih cepat.



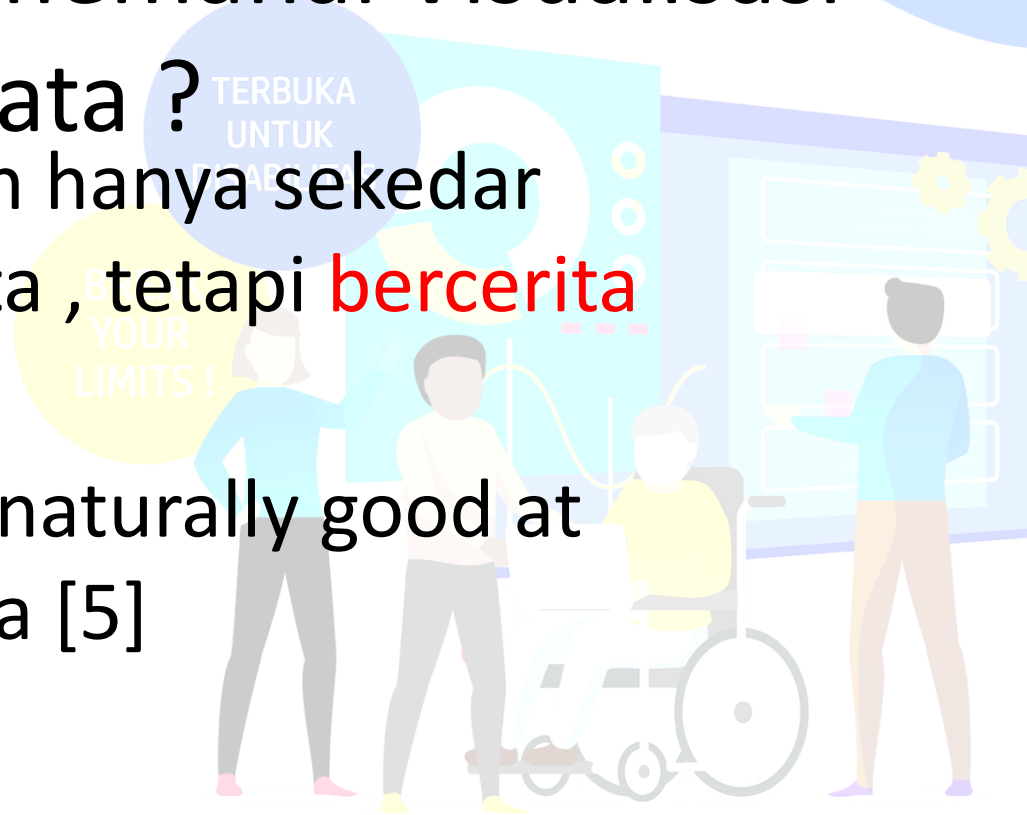


DIGITAL
TALENT
SCHOLARSHIP

Mengapa perlu memahami Visualisasi

Data ?

- Visualisasi data bukan hanya sekedar memperlihatkan data , tetapi **bercerita tentang Data**.
- We (human) aren't naturally good at storytelling with data [5]

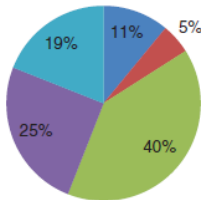


Contoh visualisasi data (showing vs story telling)

Survey Results

PRE: How do you feel about doing science?

■ Bored ■ Not great ■ OK ■ Kind of interested ■ Excited



POST: How do you feel about doing science?

■ Bored ■ Not great ■ OK ■ Kind of interested ■ Excited

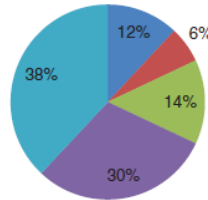
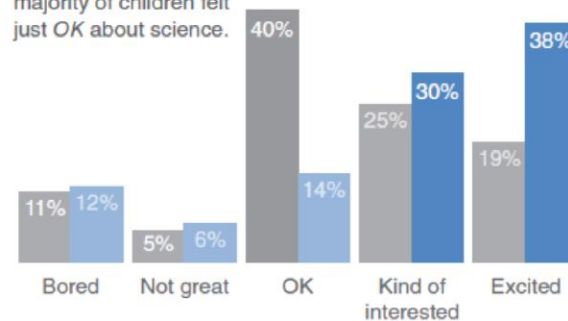


FIGURE 0.4 Example 2 (before): showing data

How do you feel about science?

BEFORE program, the majority of children felt just OK about science.

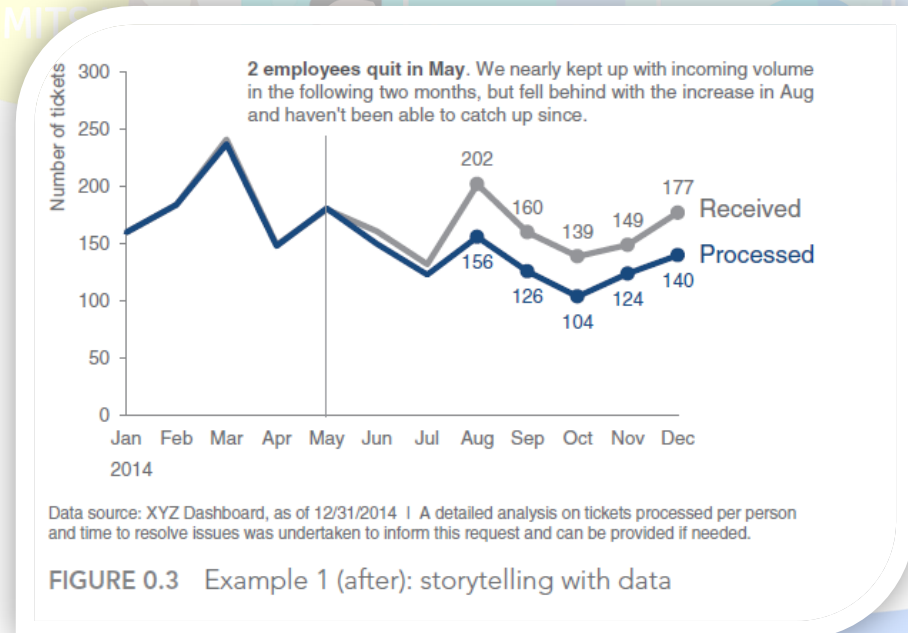
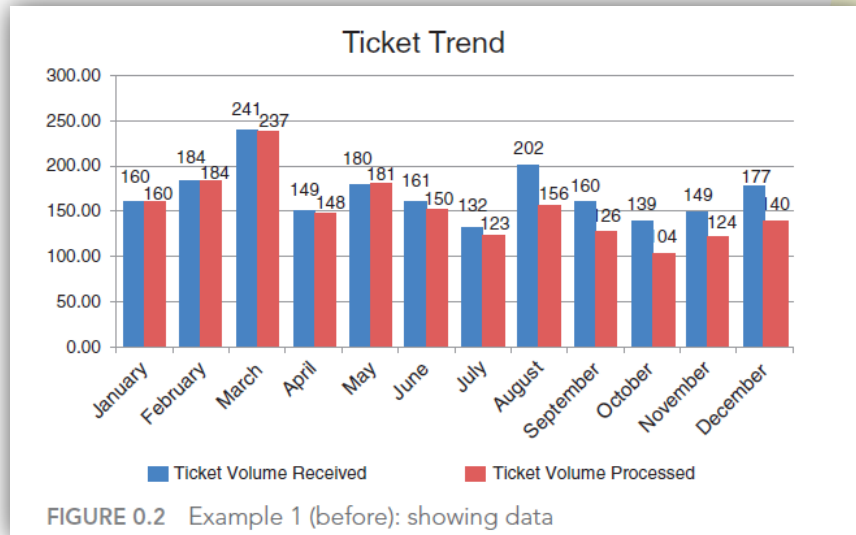


AFTER program, more children were *Kind of interested* & *Excited* about science.

Based on survey of 100 students conducted before and after pilot program (100% response rate on both surveys).

FIGURE 0.5 Example 2 (after): storytelling with data

Contoh visualisasi data (showing vs story telling)



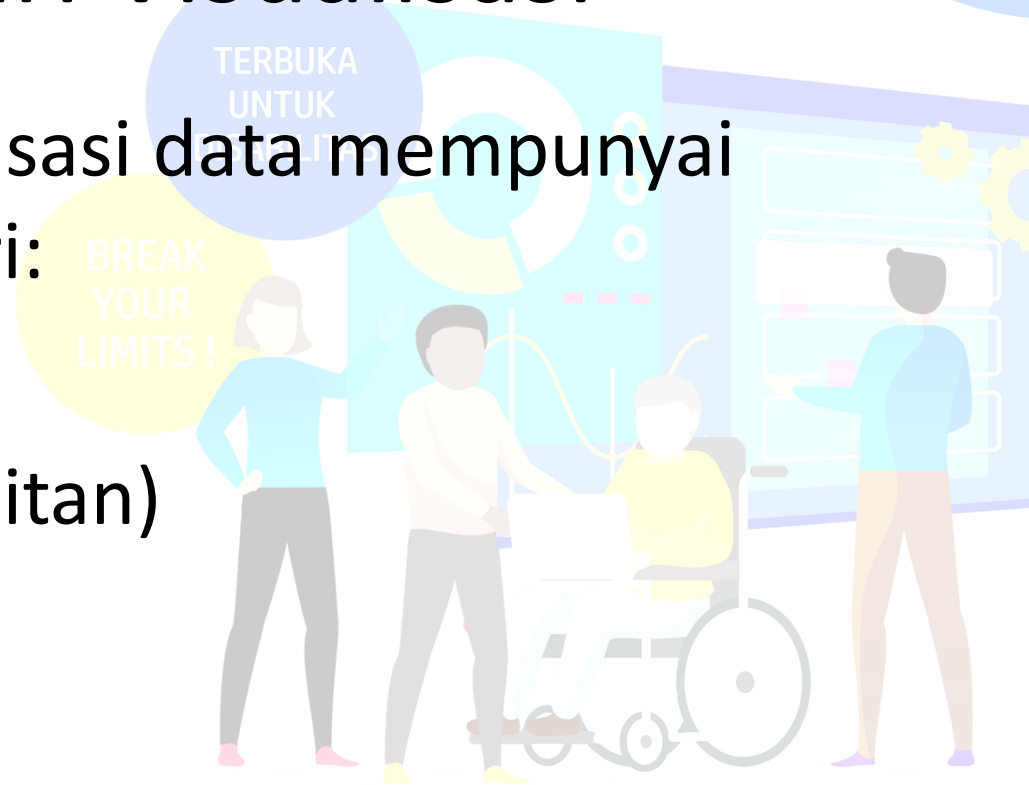


DIGITAL
TALENT
SCHOLARSHIP

Tujuan dari Visualisasi

Secara umum, visualisasi data mempunyai tujuan untuk mencari:

1. Pattern (Pola)
2. Relations (Keterkaitan)

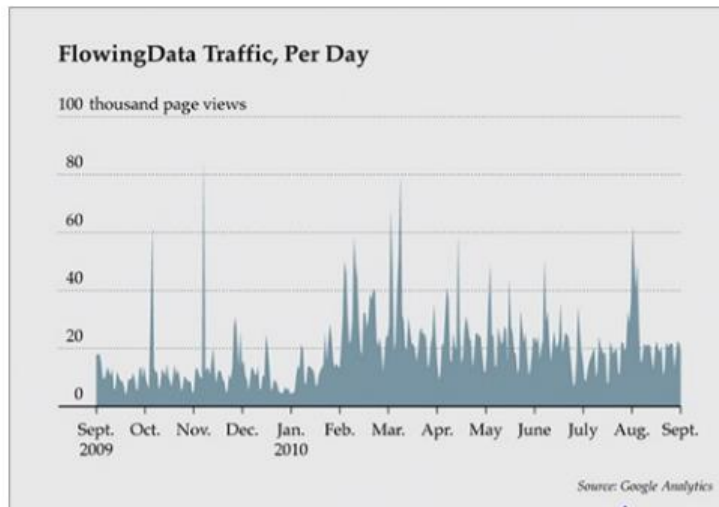




DIGITAL
TALENT
SCHOLARSHIP

Tujuan dari Visualisasi Data : Pattern

Pola dapat muncul pada sebuah data dan dapat di analisis untuk mendeteksi pemicu perubahan data.

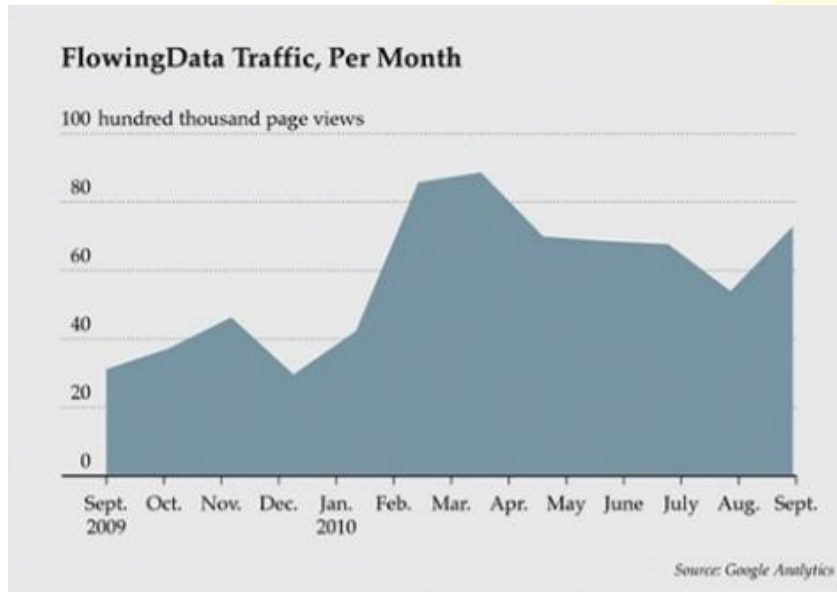




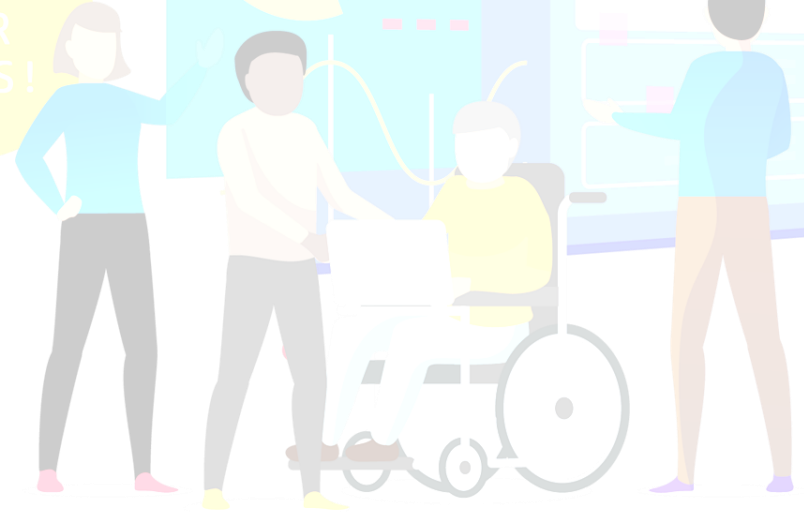
DIGITAL
TALENT
SCHOLARSHIP

Tujuan dari Visualisasi Data : Relation

Banyak variable mungkin membentuk sebuah keterkaitan antara satu dengan yang lainnya.



BREAK
YOUR
LIMITS!

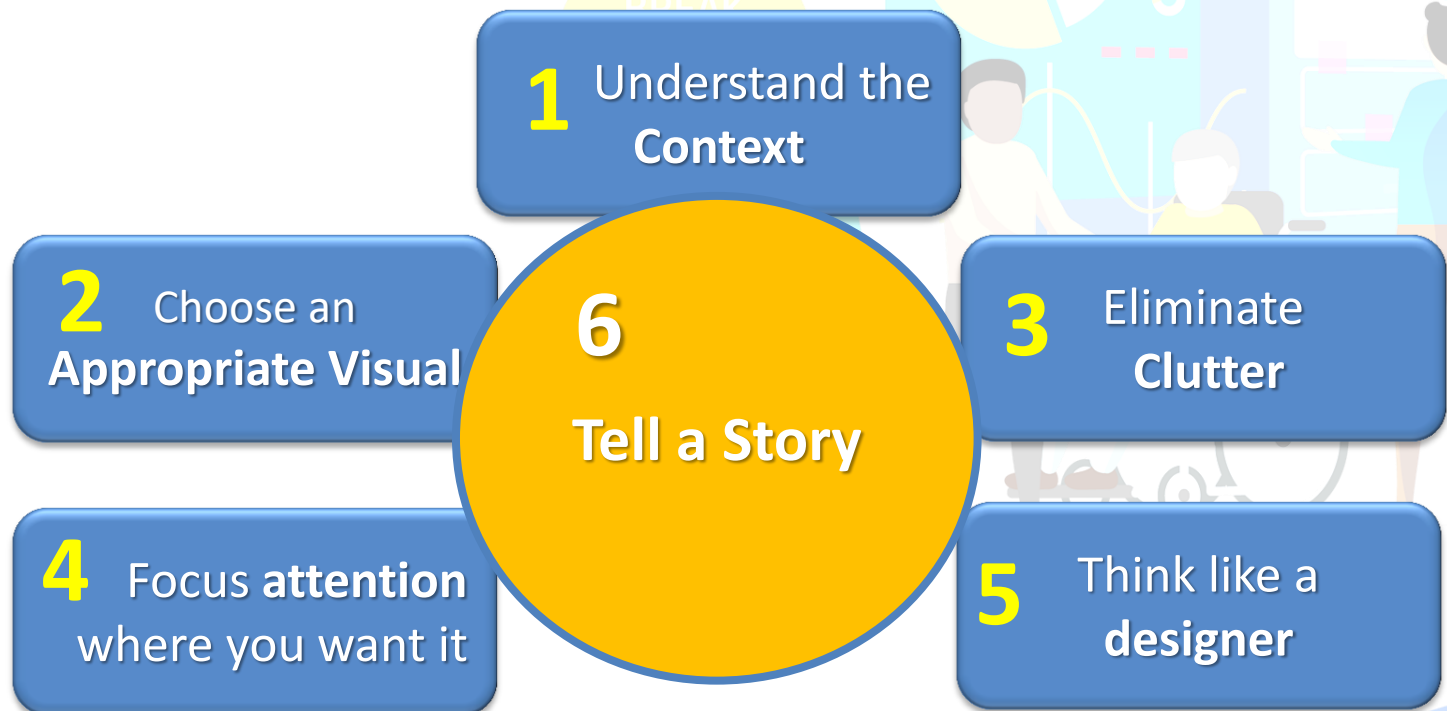




DIGITAL
TALENT
SCHOLARSHIP

Cara membangun sebuah data visual

Terdapat 6 tahapan dalam membentuk data visual:





DIGITAL
TALENT
SCHOLARSHIP

Understand the context

Untuk memahami konteks, dapat dimulai dari menjawab beberapa pertanyaan berikut :

1. *Who is your audience?*
2. *What do you need your audience to know or do?*
3. *(how) What data is available that will help make mypoint?*





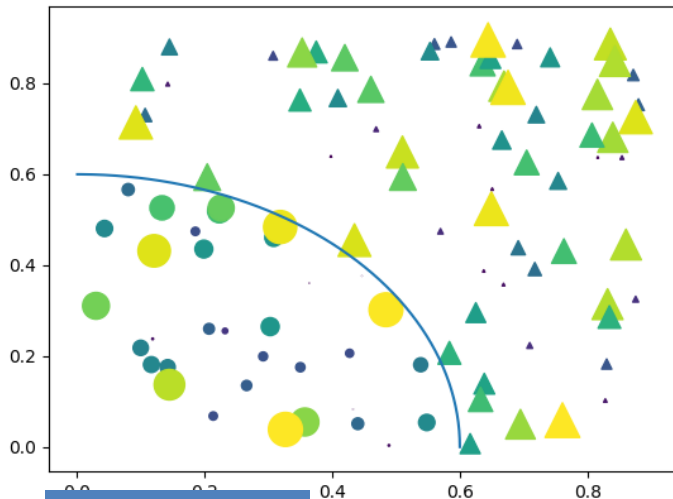
DIGITAL
TALENT
SCHOLARSHIP

Memilih visualisasi yang tepat

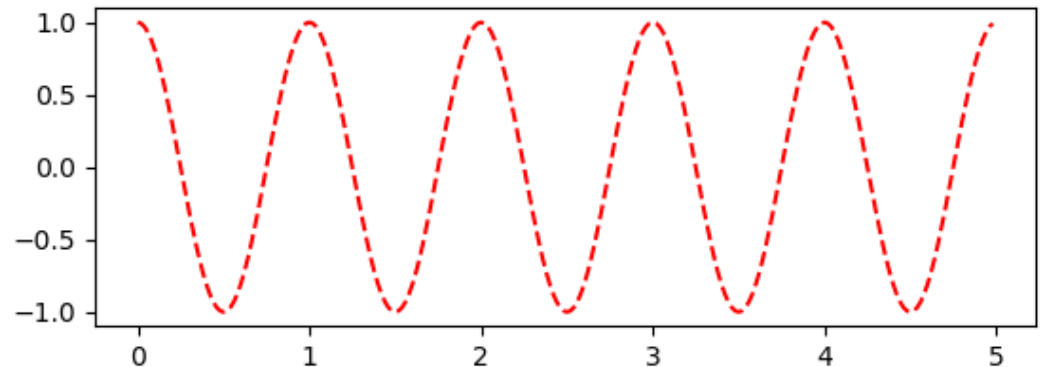
Ada banyak cara untuk memvisualisasikan data, antara lain adalah menggunakan chart (paling umum).

91%

Simple text



Scatter Plot



Line Graphs

EXPENSES	Q1	Q2
Salaries	1200000	1300000
Office	20000	20000
Merchandise	80000	70000
Legal	2000	2000
TOTAL	12120000	130902000

Table



DIGITAL
TALENT
SCHOLARSHIP

Eliminate Clutter

- Merujuk pada Gestalt Principles of Visual Perception. Pada Gestalt Principles of Visual Perception mempunyai 6 prinsip, yaitu:

Proximity

Similarity

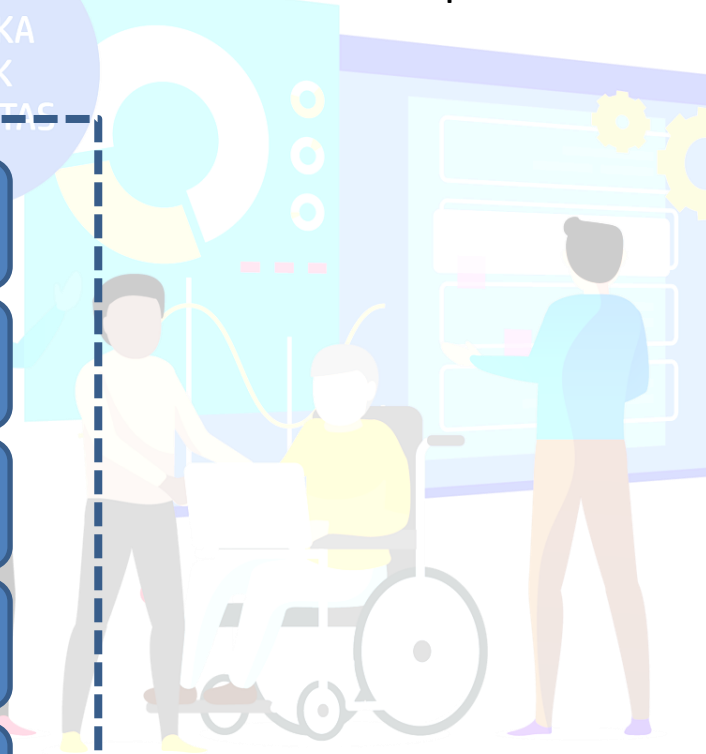
Enclosure

Closure

Continuity

Connection

TERBUKA
UNTUK
DICABILITAS





DIGITAL
TALENT
SCHOLARSHIP

Focus attention where you want it

- Pada pembuatan visualisasi data perlu memerhatikan beberapa hal berikut:

1. Ukuran
2. Warna
3. Tata Letak



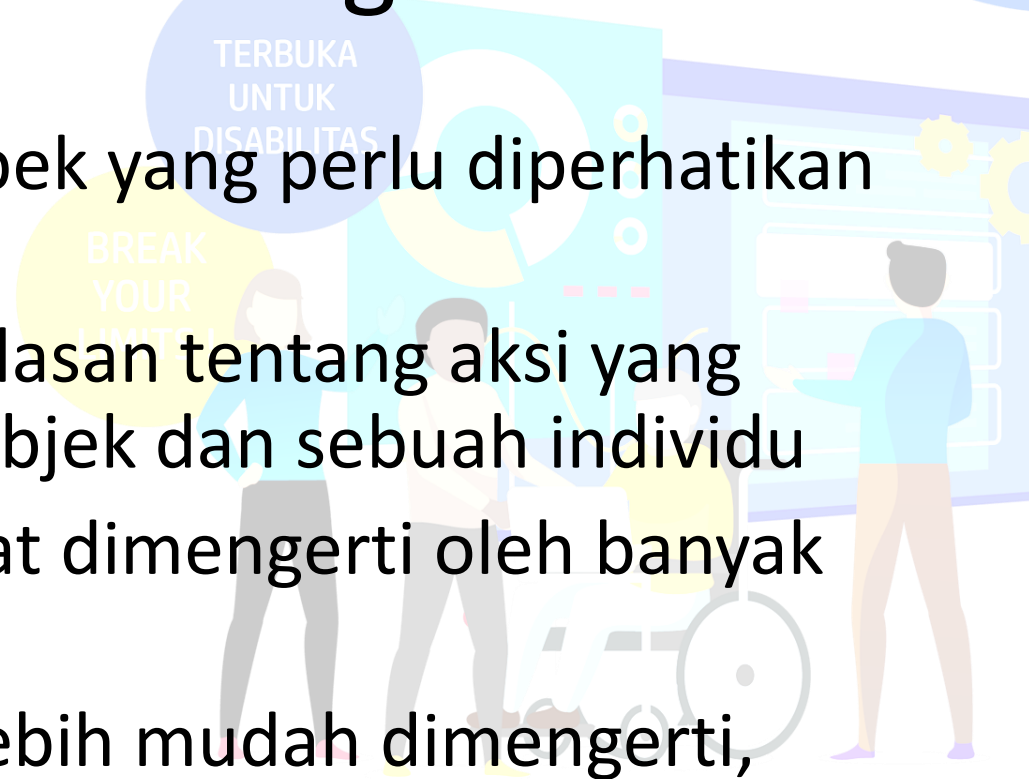


DIGITAL
TALENT
SCHOLARSHIP

Think like a Designer

Terdapat beberapa aspek yang perlu diperhatikan , yaitu:

1. *Affordances* → Kejelasan tentang aksi yang mungkin di antara objek dan sebuah individu
2. *Accessibility* → dapat dimengerti oleh banyak kalangan
3. *Aesthetics* → agar lebih mudah dimengerti,







DIGITAL
TALENT
SCHOLARSHIP

Visualizing Pattern over Time

- Discrete Temporal Data = Statis, tidak berkelanjutan dan dalam satu kurun waktu saja. Data dapat digambarkan dengan: Bar Chart, dan Point Chart
- Continuous Temporal Data = Data berkelanjutan, dapat dilihat progress data dalam jangka waktu tertentu (perubahan lebih terlihat dan tidak konstan seperti diskrit). Data dapat digambarkan dengan: Line Chart, Step Chart,



DIGITAL
TALENT
SCHOLARSHIP

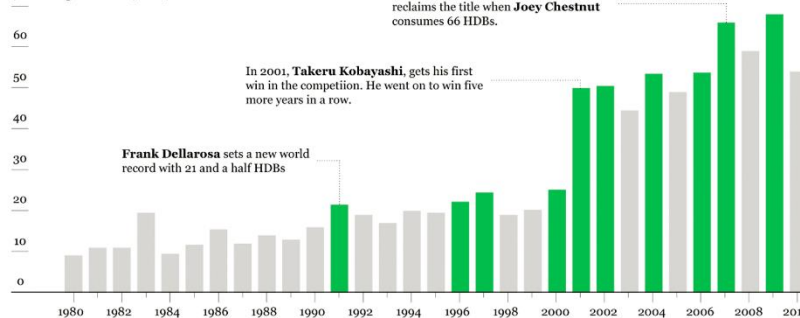
Bar Chart

- Untuk menunjukkan perbandingan dari dua aspek (X dan Y)
- Umumnya digunakan untuk perbandingan waktu atau besaran konstan lainnya

HOT DOG EATING

Nathan's hot dog eating contest every July 4th has been going on since the early 1900s, but it wasn't until 2001 when things got serious. Takeru Kobayashi from Japan raised the bar, more than doubling the previous world record. Highlighted bars indicate new records.

70 Hot dogs and buns (HDBs)



Source: Wikipedia | Nathan Yau



DIGITAL
TALENT
SCHOLARSHIP

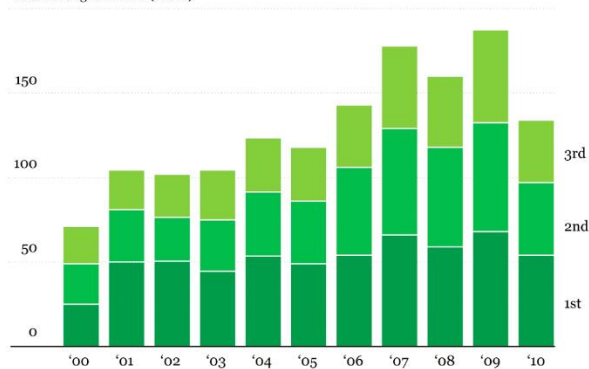
Stacked Bar Chart

- Untuk menunjukkan perbandingan dari dua aspek (X dan Y)
- Serupa dengan bar chart, namun bar memiliki lapisan yang menjelaskan sub kategori

TOP THREE HOT DOG EATERS

The year before Takeru Kobayashi started to compete in Nathan's Hot Dog Eating Contest, the top three eaters were close in skills. However, from 2001 to 2005, Kobayashi always had a substantial lead. That changed in 2006 when Joey Chestnut started competing.

200 Hot dogs and buns (HDBs)



Source: Wikipedia | Nathan Yau



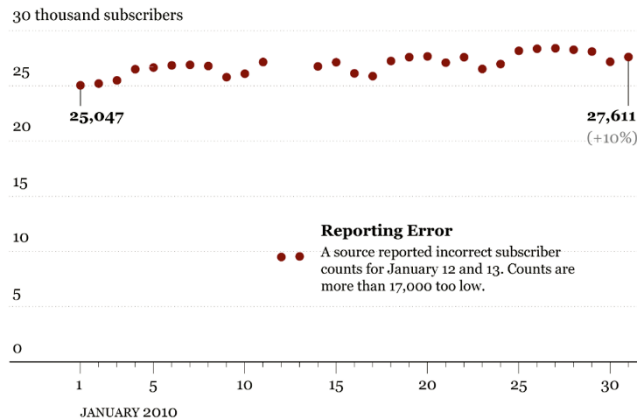
DIGITAL
TALENT
SCHOLARSHIP

Titik (Points)

- Untuk menunjukkan perbandingan dari dua aspek (X dan Y)
- Data bersifat padat, sehingga bar digantikan dengan titik

INCREASE IN SUBSCRIBERS

In January 2010, the number of subscribers via RSS and email increase to 27,611, making it the tenth month in a row with at least a ten percent increase.



Source: Feedburner | Nathan Yau



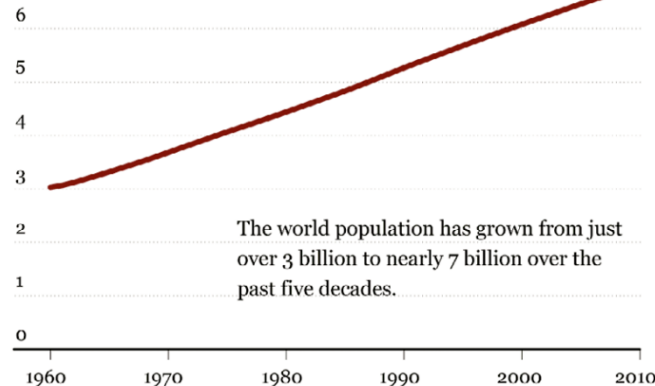
DIGITAL
TALENT
SCHOLARSHIP

Line Chart

- Pengembangan dari point chart yang dapat menunjukkan perubahan secara jelas
- Setiap titik dihubungkan dengan garis untuk menjelaskan proses perubahan

WORLD POPULATION

7 billion people

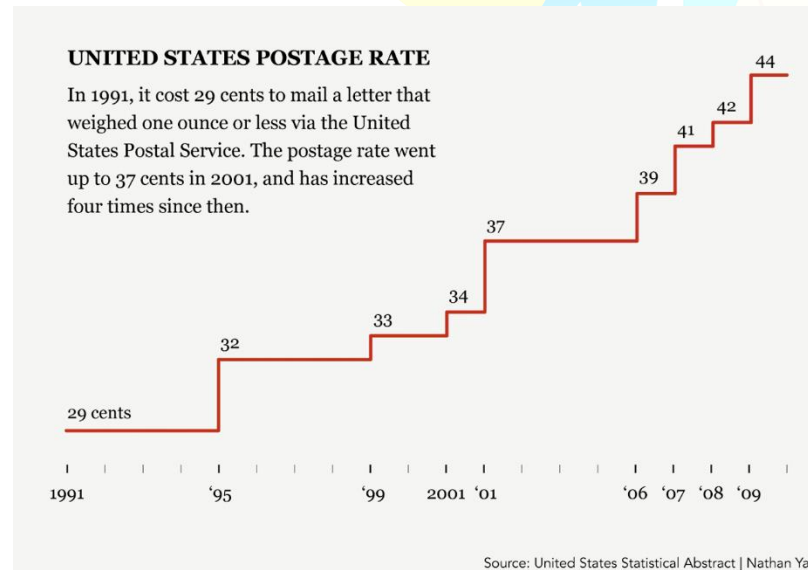




DIGITAL
TALENT
SCHOLARSHIP

Step Chart

- Digunakan untuk data yang statis naik/turun perkembangannya.
- Modifikasi dari line chart

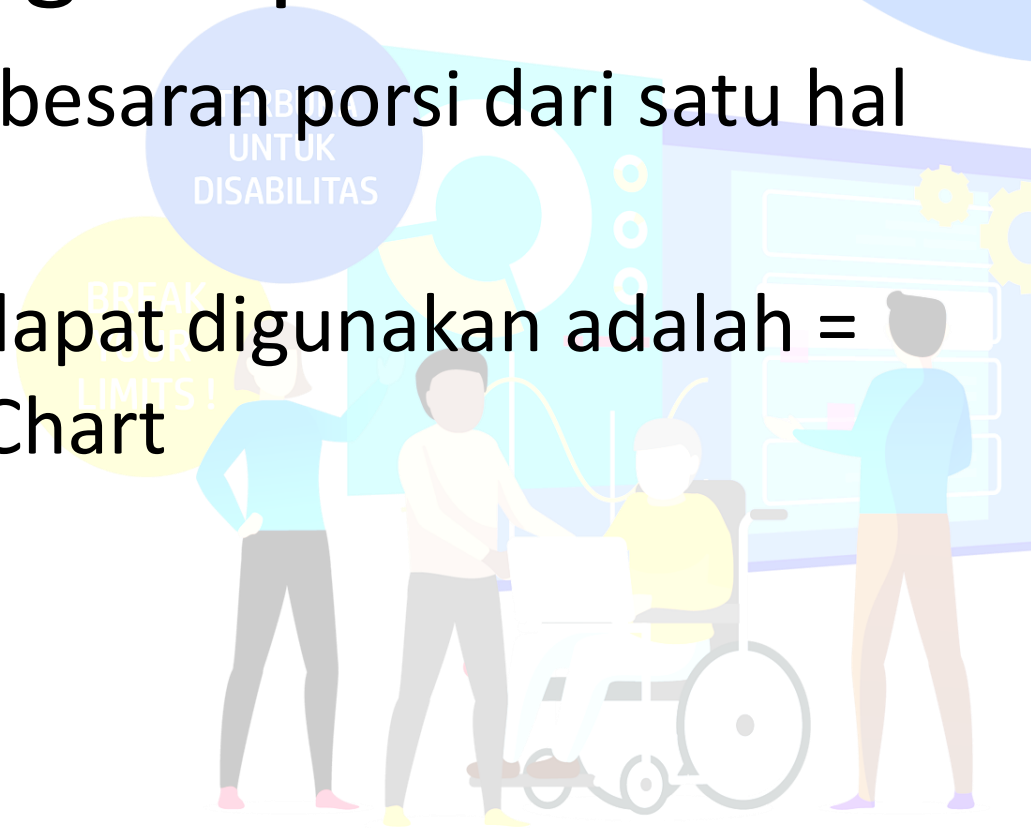




DIGITAL
TALENT
SCHOLARSHIP

Visualizing Proportion

- Untuk menjelaskan besaran porsi dari satu hal mutlak
- Contoh chart yang dapat digunakan adalah = Pie Chart dan Area Chart





DIGITAL
TALENT
SCHOLARSHIP

Pie Chart

- Untuk data yang menunjukkan persentase atau besaran yang dimiliki suatu hal
- Cocok digunakan untuk menggambarkan perbandingan dari suatu hal yang memiliki proporsi numerik

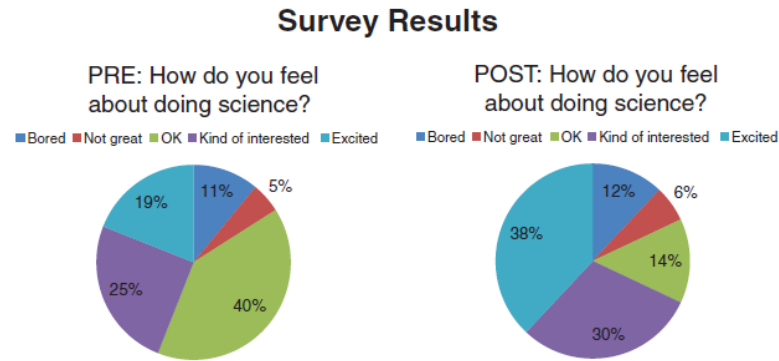


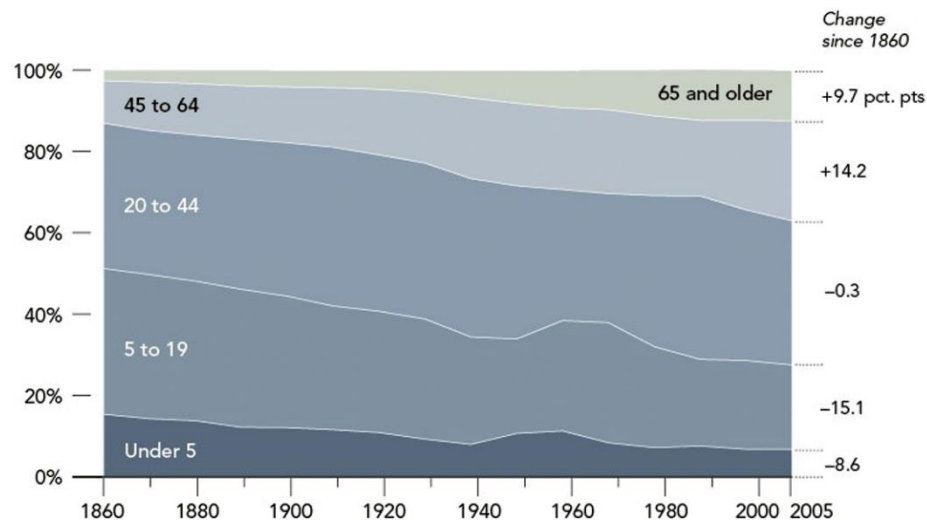
FIGURE 0.4 Example 2 (before): showing data



DIGITAL
TALENT
SCHOLARSHIP

Area Chart

- Menunjukkan proporsi dalam satuan waktu
- Modifikasi dari line chart yang menggambarkan proporsi data

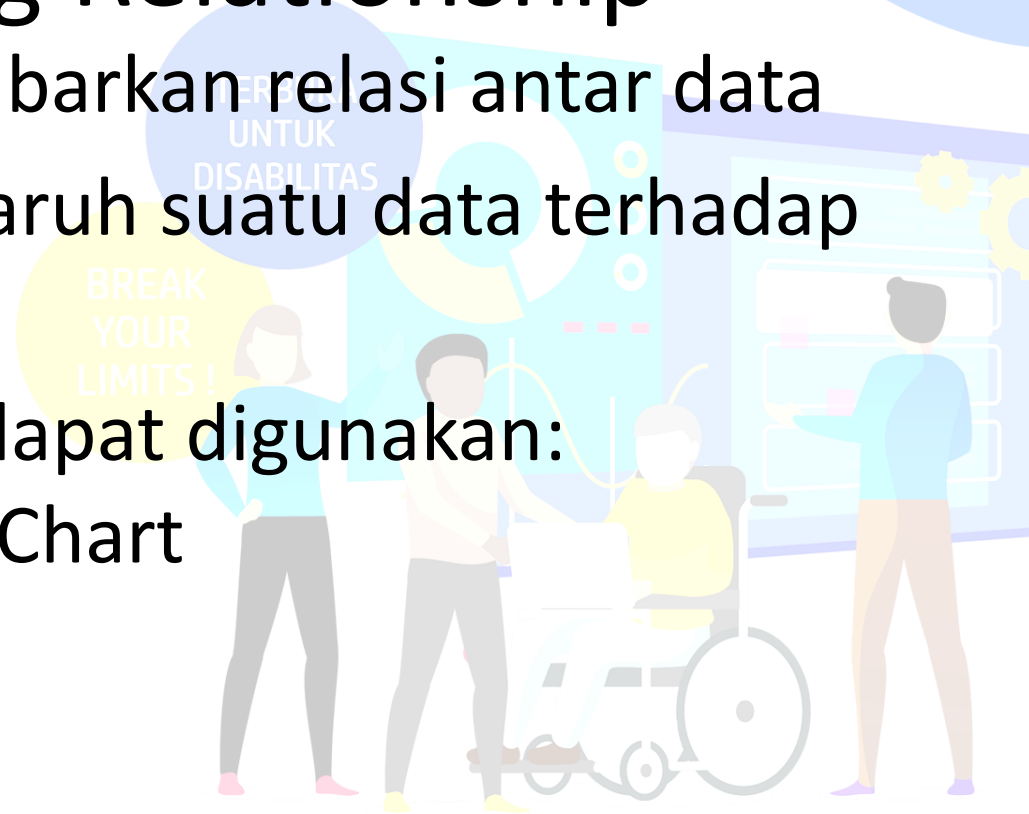




DIGITAL
TALENT
SCHOLARSHIP

Visualizing Relationship

- Pola yang menggambarkan relasi antar data
- Menunjukkan pengaruh suatu data terhadap data lainnya
- Contoh chart yang dapat digunakan:
Scatterplot, Bubble Chart





DIGITAL
TALENT
SCHOLARSHIP

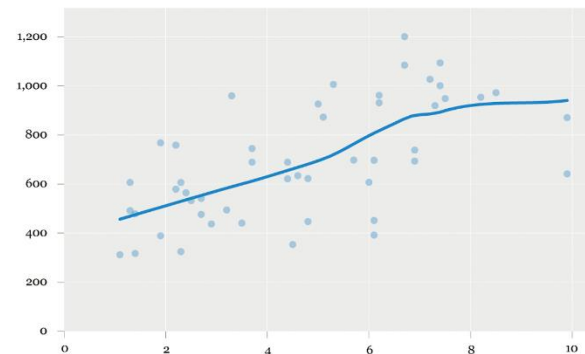
Scatterplot

- Implementasi dari dua data pada line chart dan point chart
- Umumnya kedua jenis data saling berkaitan untuk mengetahui dampak data satu sama lain

MURDERS VERSUS BURGLARIES IN THE UNITED STATES

States with higher murder rates tend to have higher burglary rates.

Burglaries
per 100,000 population



Murders
per 100,000 population

Source: U.S. Census Bureau | Nathan Yau

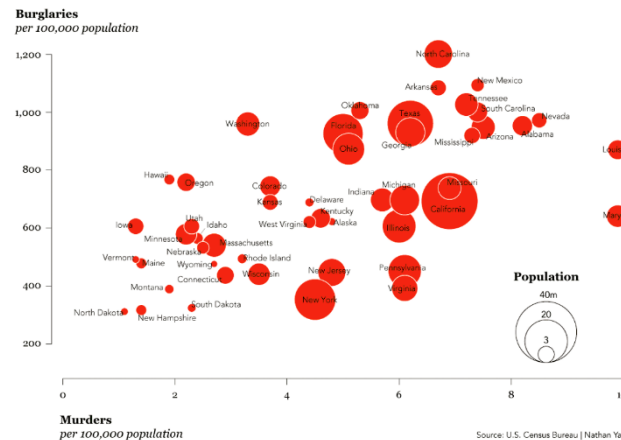


DIGITAL
TALENT
SCHOLARSHIP

Bubble Chart

- Modifikasi dari point chart
- Dapat digunakan untuk komparasi 3 hal sekaligus, dengan sumbu X, sumbu Y, dan besarnya titik

MURDERS VERSUS BURGLARIES IN THE UNITED STATES

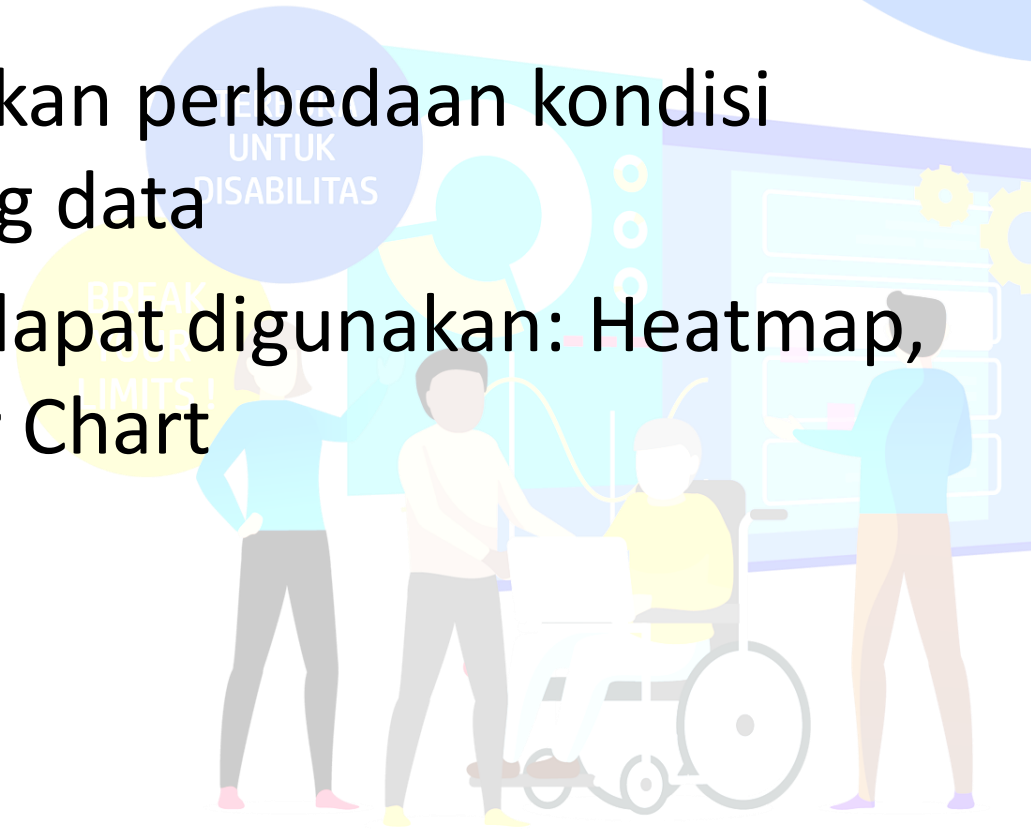




DIGITAL
TALENT
SCHOLARSHIP

Spotting Differences

- Untuk memperlihatkan perbedaan kondisi antar masing-masing data
- Contoh chart yang dapat digunakan: Heatmap, Chernoff Faces, Star Chart

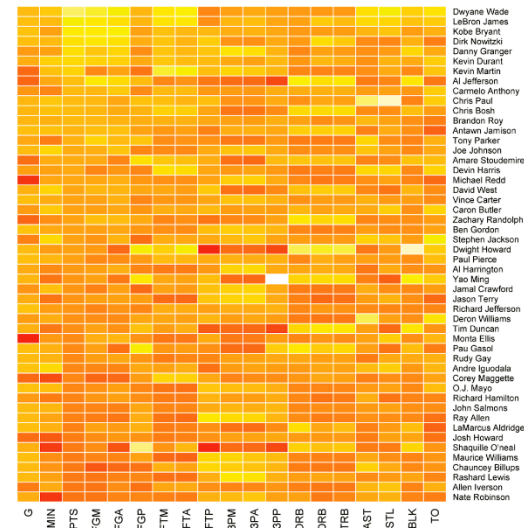




DIGITAL
TALENT
SCHOLARSHIP

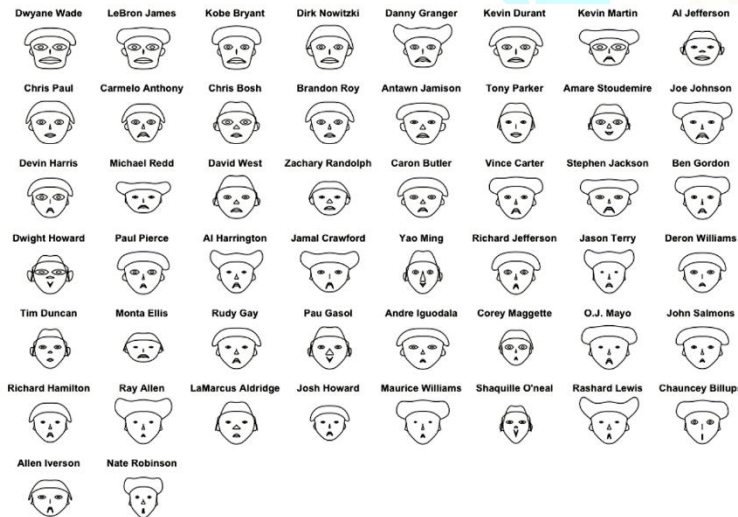
Heatmap

- Menggambarkan kondisi perbedaan antar masing-masing data dengan warna
- Data digambarkan dari warna, dimana warna dapat ditinjau dari kecerahan maupun kegelapan



Chernoff Faces

- Menggambarkan variasi data dari bentuk wajah yang dihasilkan
- Beberapa faktor pembeda dapat dilihat dari atribut wajah, seperti mata, hidung, mulut, dsb

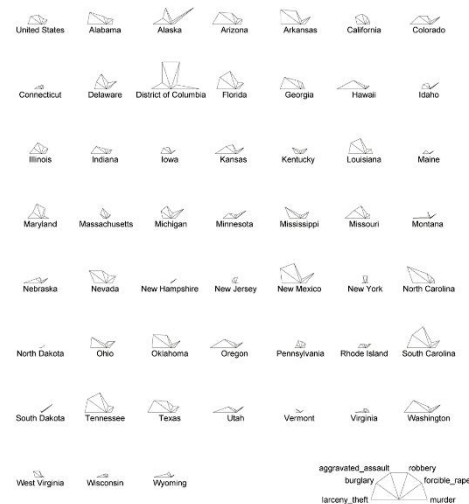




DIGITAL
TALENT
SCHOLARSHIP

Star Chart

- Serupa dengan Chernoff Faces, namun menggunakan bentuk bangun datar yang menggambarkan data
- Bangun datar dapat ditinjau dari ukuran, sudut, bentuk, dsb

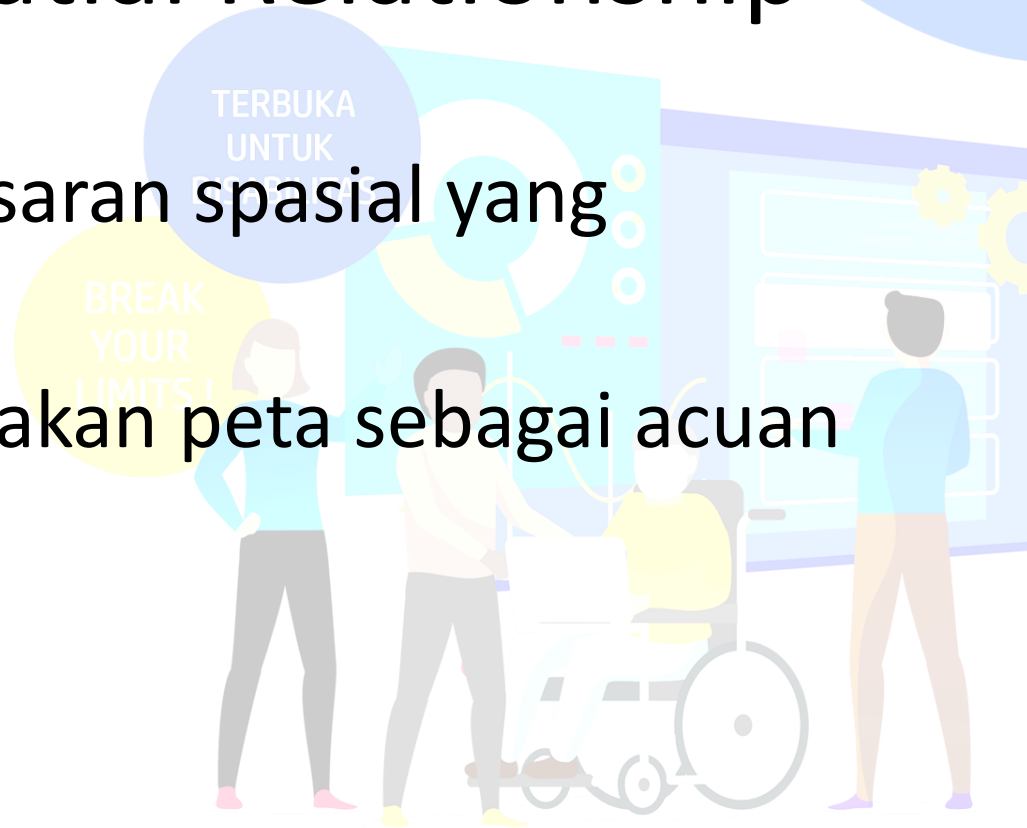




DIGITAL
TALENT
SCHOLARSHIP

Visualizing Spatial Relationship

- Menggambarkan besaran spasial yang berkaitan antar data
- Umumnya menggunakan peta sebagai acuan data





DIGITAL
TALENT
SCHOLARSHIP

Map with Dots

- Penggunaan titik pada peta yang dapat digunakan untuk menjelaskan lokasi atau keberadaan suatu data pada peta
- Dapat menunjukkan keberadaan serta kepadatan berdasarkan titik

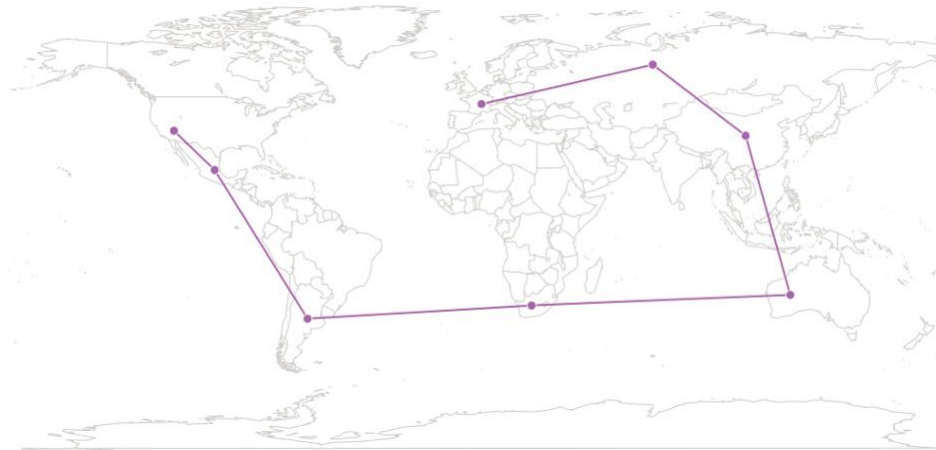




DIGITAL
TALENT
SCHOLARSHIP

Map with Lines

- Dapat menggambarkan jarak antar titik dari Map with Dots
- Digunakan untuk identifikasi jarak, rute, koneksi, ' '



TERBUKA
UNTUK
PROBLEM

BREAK
YOUR
LIMITS!

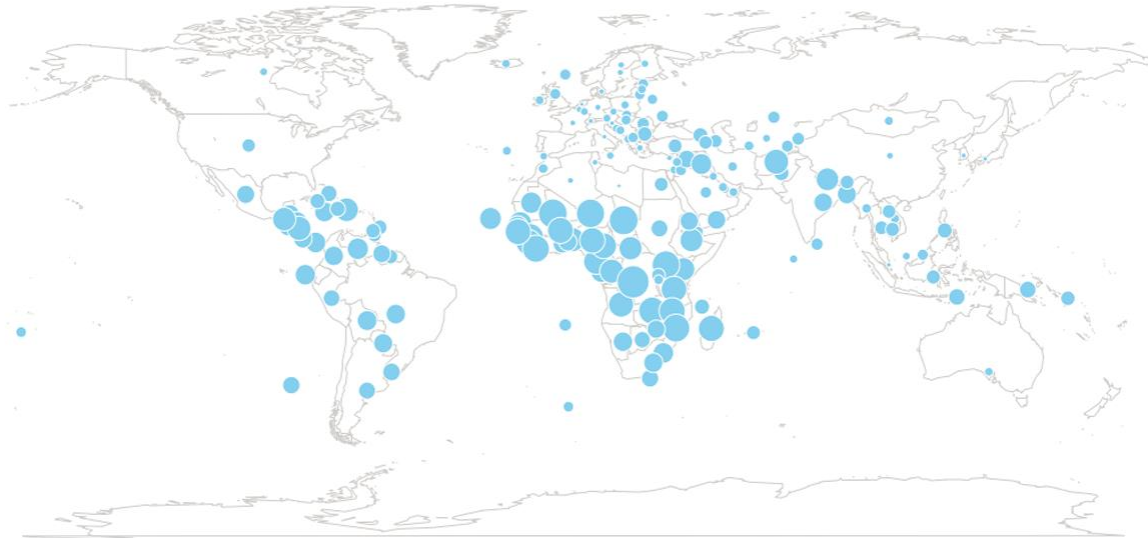




DIGITAL
TALENT
SCHOLARSHIP

Map with Bubbles

- Sama seperti Map with Dots, namun titik menggambarkan besaran data pada lokasi tersebut





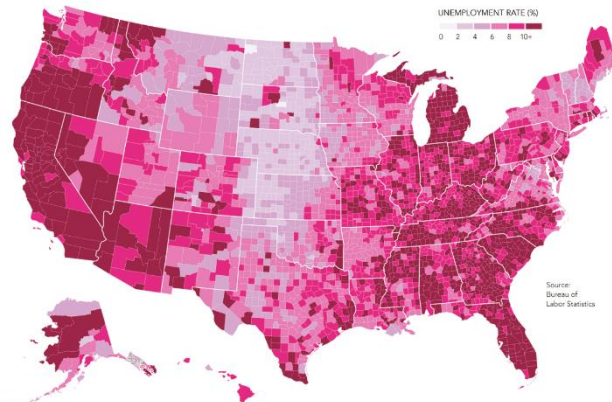
DIGITAL
TALENT
SCHOLARSHIP

Map with Regions

- Menggambarkan data sebagai bagian dari peta
- Data digambarkan dari pemecahan peta menggunakan garis/warna yang menggambarkan

UNEMPLOYMENT RATES, AUGUST 2010

The national unemployment rate in August 2009 was 9.7 percent. One year later, the rate has not changed much with an average of 9.6 percent.

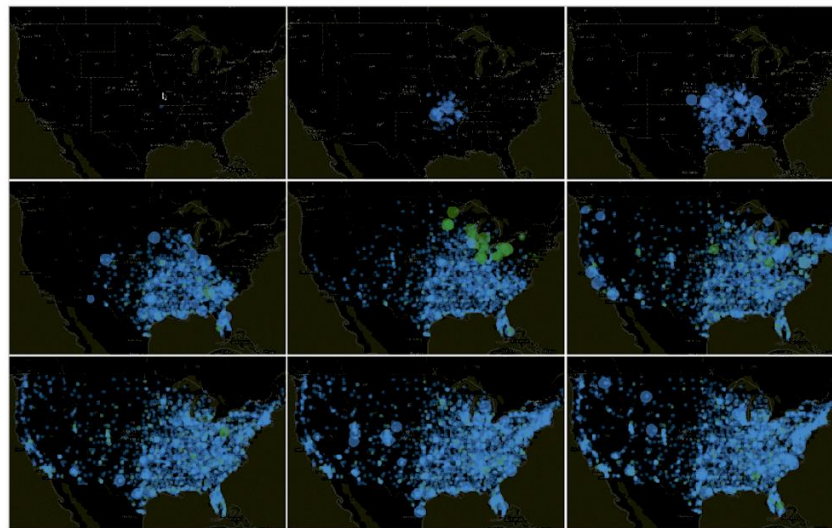




DIGITAL
TALENT
SCHOLARSHIP

Map over Space and Time

- Menunjukkan perkembangan data dalam kurun waktu tertentu
- Umumnya menggunakan Map with Region dengan jumlah yang lebih dari satu



TERBUKA
KURANG
DISABILITAS

BREAK
YOUR
LIMITS!

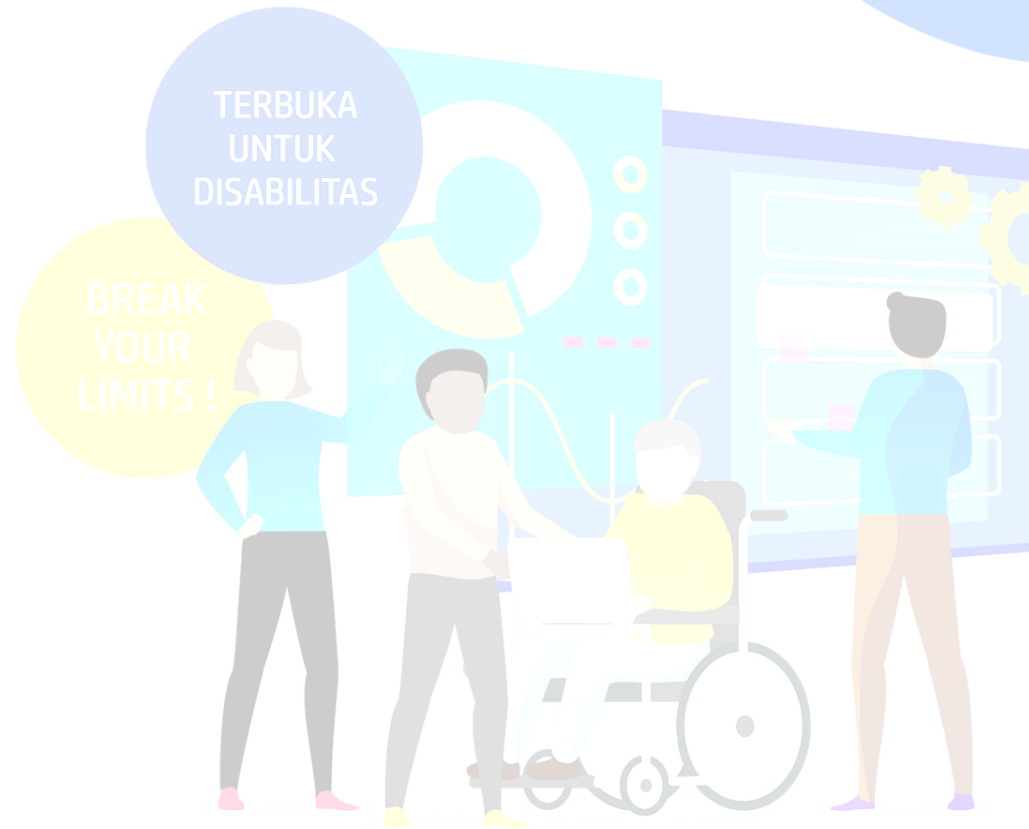




DIGITAL
TALENT
SCHOLARSHIP

Designing with a Purpose

- Prepare yourself
- Prepare for readers
- Visual Cues
- Good Visualization





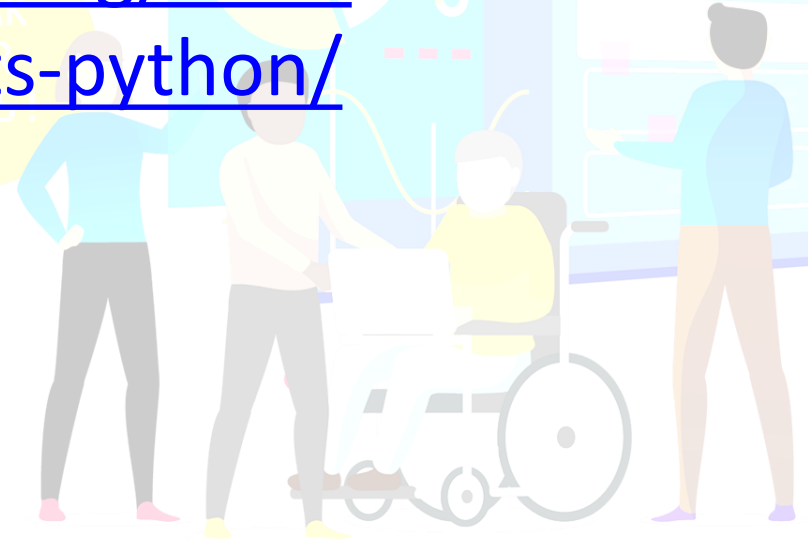
DIGITAL
TALENT
SCHOLARSHIP

Tugas 1: Aneka Visualisasi

Ikuti tutorial di:

<https://www.geeksforgeeks.org/data-visualization-different-charts-python/>

TERBUKA
UNTUK
DISABILITAS





DIGITAL
TALENT
SCHOLARSHIP

Sample Code To Visualize Data

- Membuat histogram

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

# create 2D array of table given above
data = [['E001', 'M', 34, 123, 'Normal', 350],
        ['E002', 'F', 40, 114, 'Overweight', 450],
        ['E003', 'F', 37, 135, 'Obesity', 169],
        ['E004', 'M', 30, 139, 'Underweight', 189],
        ['E005', 'F', 44, 117, 'Underweight', 183],
        ['E006', 'M', 36, 121, 'Normal', 80],
        ['E007', 'M', 32, 133, 'Obesity', 166],
        ['E008', 'F', 26, 140, 'Normal', 120],
        ['E009', 'M', 32, 133, 'Normal', 75],
        ['E010', 'M', 36, 133, 'Underweight', 40] ]

# dataframe created with
# the above data array
df = pd.DataFrame(data, columns = ['EMPID', 'Gender',
                                   'Age', 'Sales',
                                   'BMI', 'Income' ] )

# create histogram for numeric data
df.hist()

# show plot
plt.show()
```

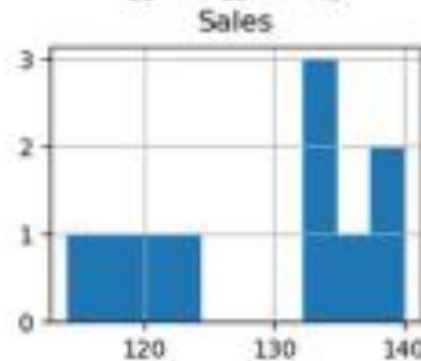
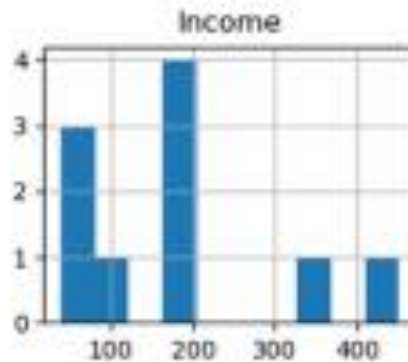
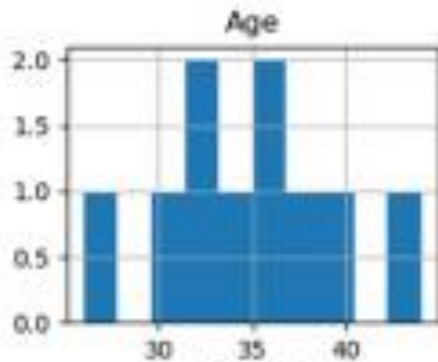




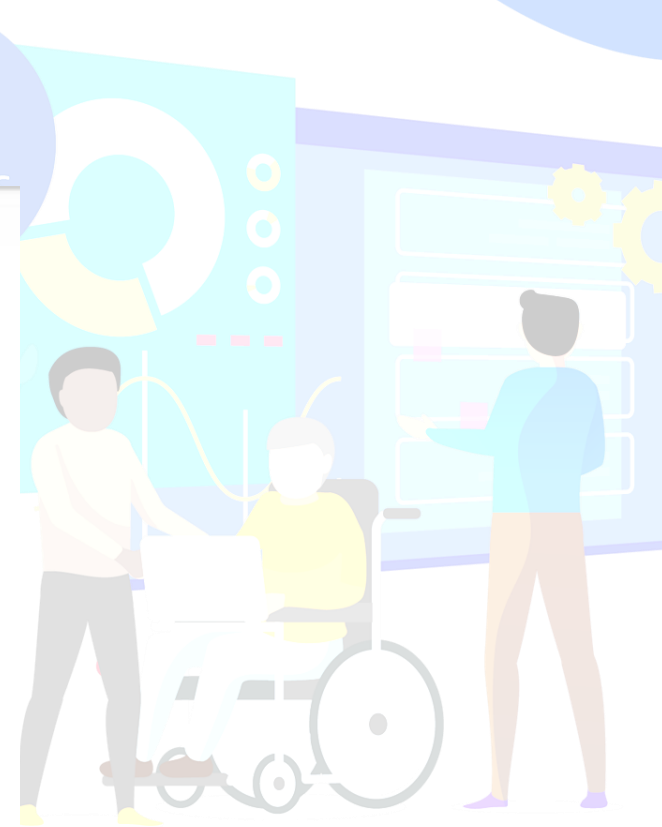
DIGITAL
TALENT
SCHOLARSHIP

Sample Code To Visualize Data

- Membuat histogram



TERBUKA
UNTUK
DISABILITAS





DIGITAL
TALENT
SCHOLARSHIP

Sample Code To Visualize Data

- Membuat column chart

```
# Dataframe of previous code is used here
```

```
# Plot the bar chart for numeric values  
# a comparison will be shown between  
# all 3 age, income, sales  
df.plot.bar()
```

```
# plot between 2 attributes  
plt.bar(df['Age'], df['Sales'])  
plt.xlabel("Age")  
plt.ylabel("Sales")  
plt.show()
```



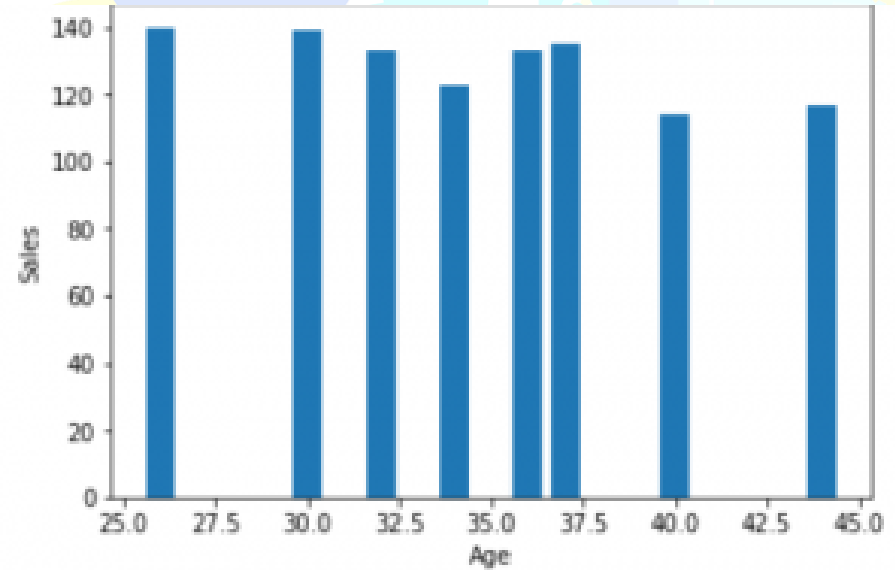
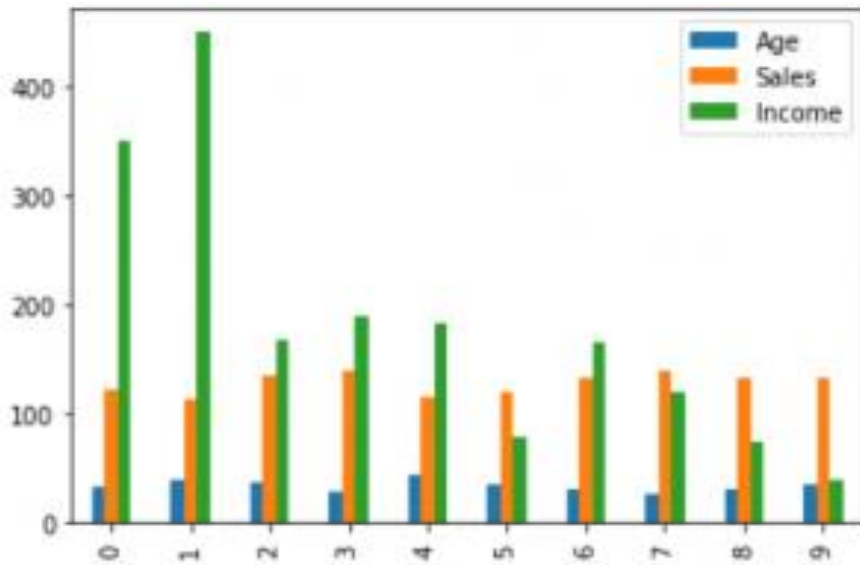


DIGITAL
TALENT
SCHOLARSHIP

Sample Code To Visualize Data

- Membuat column chart

TERBUKA
UNTUK
DISABILITAS





DIGITAL
TALENT
SCHOLARSHIP

Sample Code To Visualize Data

- Membuat pie chart

```
plt.pie(df['Age'], labels = {"A", "B", "C",  
                             "D", "E", "F",  
                             "G", "H", "I", "J"},
```

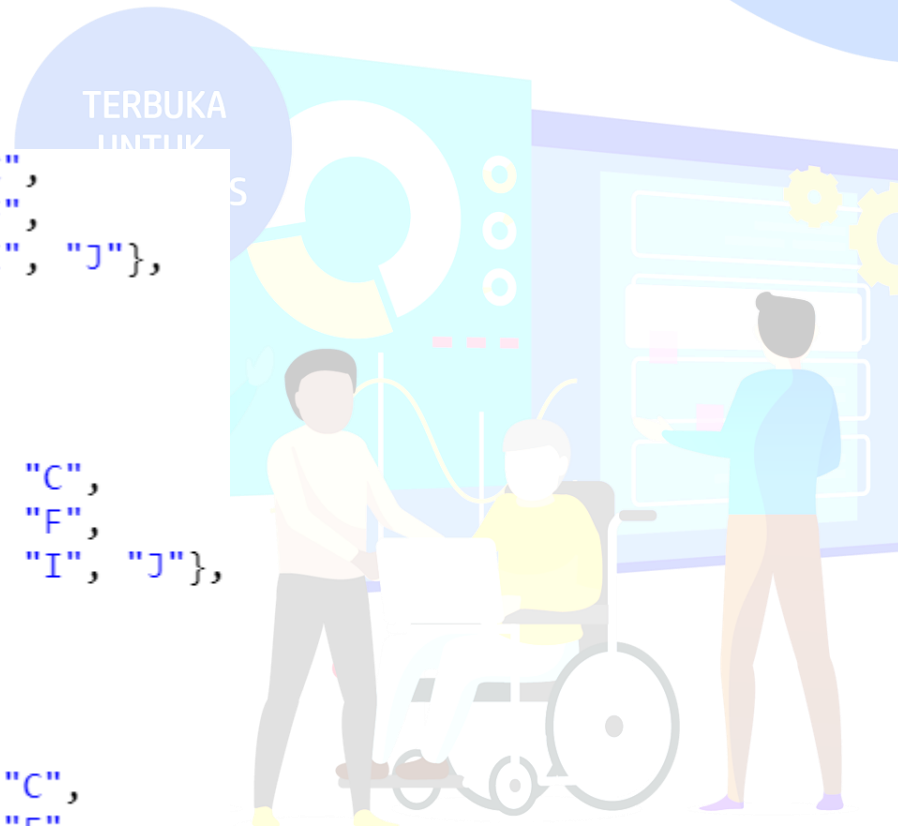
```
autopct = '% 1.1f %%', shadow = True)  
plt.show()
```

```
plt.pie(df['Income'], labels = {"A", "B", "C",  
                                "D", "E", "F",  
                                "G", "H", "I", "J"},
```

```
autopct = '% 1.1f %%', shadow = True)  
plt.show()
```

```
plt.pie(df['Sales'], labels = {"A", "B", "C",  
                               "D", "E", "F",  
                               "G", "H", "I", "J"},
```

```
autopct = '% 1.1f %%', shadow = True)  
plt.show()
```

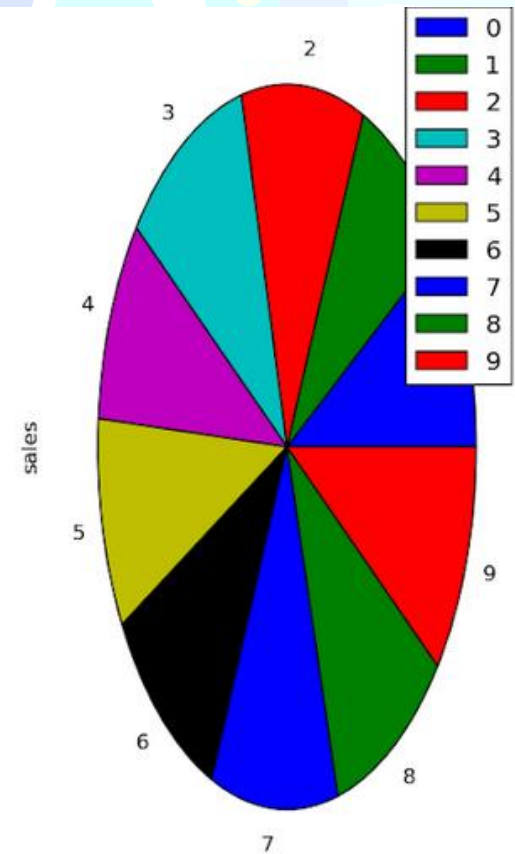
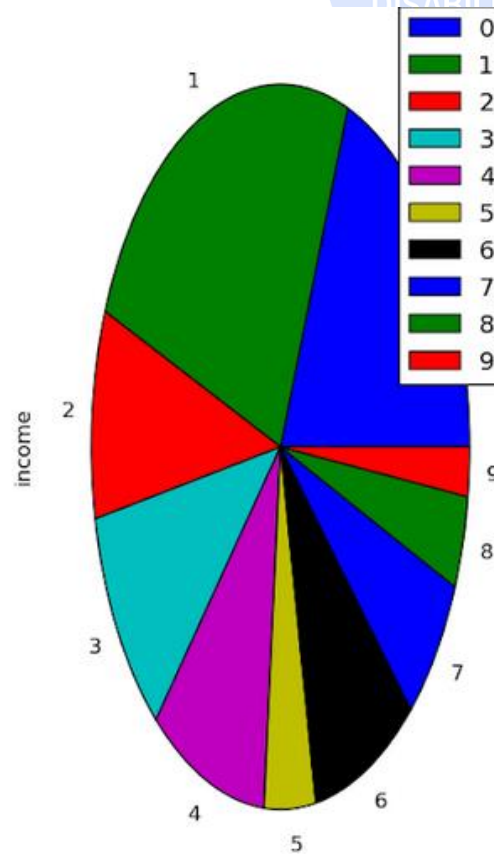
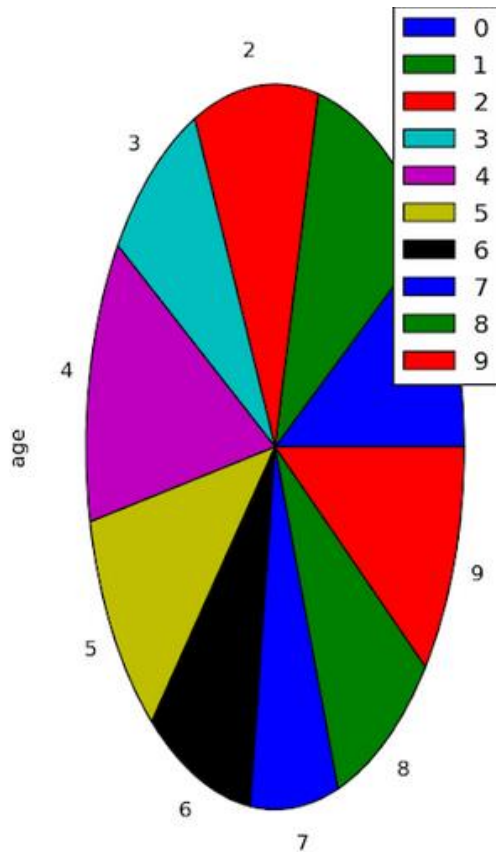




DIGITAL
TALENT
SCHOLARSHIP

Sample Code To Visualize Data

- Membuat pie chart



TERBUKA
UNTUK
DISABILITAS



DIGITAL
TALENT
SCHOLARSHIP

Sample Code To Visualize Data

- Membuat scatter plot

```
# scatter plot between income and age  
plt.scatter(df['income'], df['age'])  
plt.show()
```

```
# scatter plot between income and sales  
plt.scatter(df['income'], df['sales'])  
plt.show()
```

```
# scatter plot between sales and age  
plt.scatter(df['sales'], df['age'])  
plt.show()
```

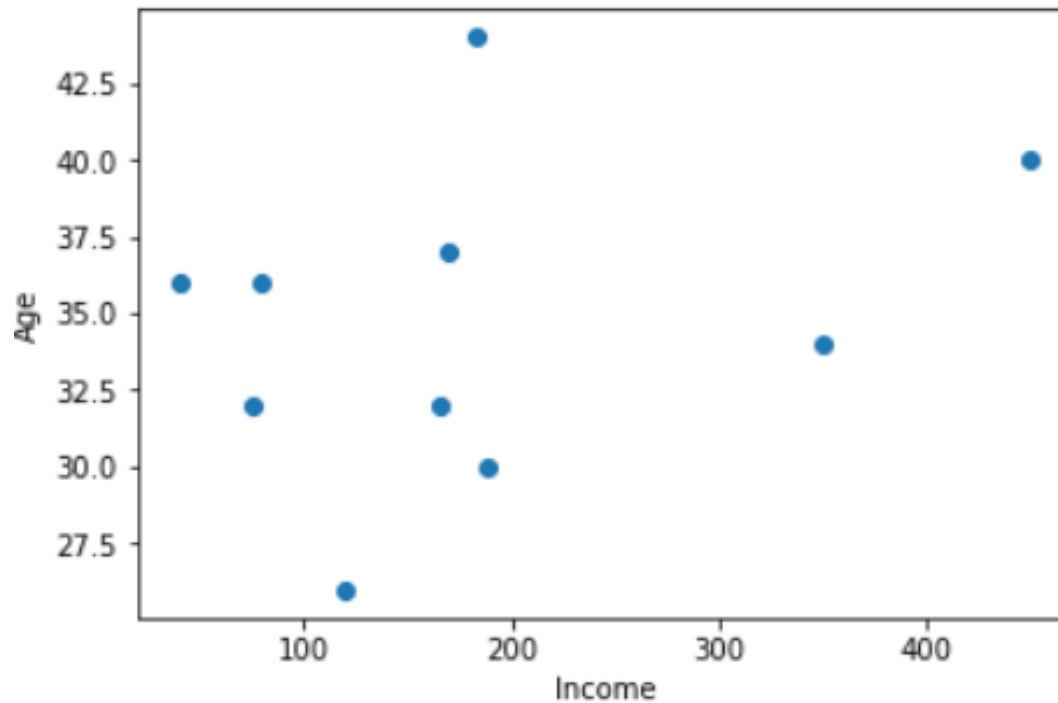




DIGITAL
TALENT
SCHOLARSHIP

Sample Code To Visualize Data

- Membuat scatter plot



TERBUKA
UNTUK
DISABILITAS









DIGITAL
TALENT
SCHOLARSHIP

Referensi

- Yau, N, *Visualize This: The Flowing Data Guideto Design, Visualization and Statistics*
- Knafllic, CN, 2015, *Story Telling with Data*, Wiley
- <http://had.co.nz/stat645/model-vis.pdf>
- National Forum on Education Statistics. (2016). *Forum Guide to Data Visualization: A Resource for Education Agencies*. (NFES 2017-016). U.S. Department of Education. Washington, DC: National Center for Education Statistics.
- Treisman, A. (1985). Preattentive Processing in Vision. *Computer Vision, Graphics, and Image Processing*, 31(2):156-177.
- Treisman, A. (1986). Features and Objects in Visual Processing. *Scientific American*, 255(5):114-125.
- Wolfe, J. and Robertson, L (Eds). 2012, *From Perception to Consciousness: Searching with Anne Triesman*. Oxford University Press: Oxford, England.
- <https://www.geeksforgeeks.org/data-visualization-different-charts-python/>

IKUTI KAMI



-  [digitalent.kominfo](https://www.facebook.com/digitalent.kominfo)
-  [digitalent.kominfo](https://www.instagram.com/digitalent.kominfo)
-  [DTS_kominfo](https://twitter.com/DTS_kominfo)
-  [Digital Talent Scholarship 2019](https://www.telegram.com/join/digitalent.kominfo)

Pusat Pengembangan Profesi dan Sertifikasi
Badan Penelitian dan Pengembangan SDM
Kementerian Komunikasi dan Informatika
Jl. Medan Merdeka Barat No. 9
(Gd. Belakang Lt. 4 - 5)
Jakarta Pusat, 10110

