



UNIVERSITAS AIRLANGGA  
Excellence with Morality

# JENIS DATA DAN VISUALISASINYA



# Materi Pertemuan ke-3

1. Skala data
2. Visualisasi data berdasarkan skala data
3. Interpretasi hasil visualisasi data





# Data Pengukuran

Nominal

Ordinal

Interval

Ratio



# Skala Data

	Nominal	Ordinal	Interval	Rasio
Dapat diklasifikasikan	Ö	Ö	Ö	Ö
Dapat diurutkan		Ö	Ö	Ö
Dapat dihitung selisihnya			Ö	Ö
Mempunyai titik nol mutlak				Ö





## Nominal & Ordinal (KATEGORIK)

- Data nominal & ordinal : mempunyai kategori variabel.
- Contoh I :
  - Jenis Kelamin ----- variabel
    - \* Pria & Wanita ----- kategori data



- Contoh II :

- Kadar Kolesterol -----variabel

- Rendah -----
    - Normal -----
    - Tinggi -----

kategori data

- Pertanyaan :

1. Apakah Anda melihat perbedaan contoh I & II ?
2. Contoh mana yang mempunyai skala nominal & mana yang skala ordinal ?
3. Apa yang membedakan skala nominal & skala ordinal ?
4. Diskusikan ! Buat contoh-contoh yang lain !





- Berdasarkan kategori itulah dapat dibedakan data nominal & variabel ordinal;
- Data nominal : mempunyai kategori yang **"sederajat"** (pria & wanita), data ini sifatnya hanya bisa diklasifikasikan berdasarkan namanya saja
- Data ordinal : mempunyai kategori yang **"tidak sederajat" atau "kategori yang bertingkat"**



# Penyajian Data

- Dalam penelitian, data yang dikumpulkan untuk keperluan analisis **harus disajikan sedemikian rupa** sehingga mudah **dipelajari dan diambil informasi** yang diperlukan.
- Tujuan : memberikan gambaran yang baik dan menyeluruh dari objek yang diteliti.
- Penyajian data menjadi penting karena peneliti harus bisa memilih cara penyajian data yang tepat agar data yang dimiliki lebih dapat dipahami oleh pembacanya.





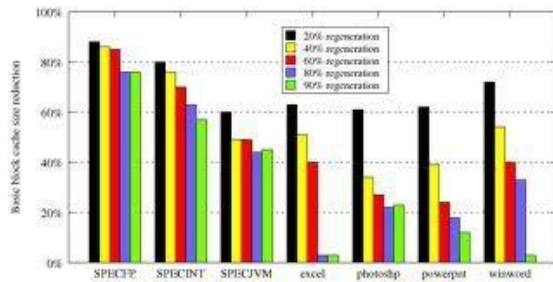
# Visualisasi Data

- Bentuk data nominal dan ordinal dapat divisualisasikan menggunakan beberapa cara:
  1. Bar graph
  2. Pie chart
  3. Line graph

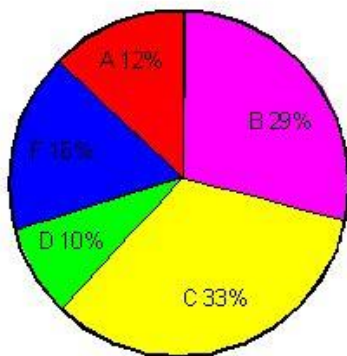


# GRAPHING- 3 Types of Graphs

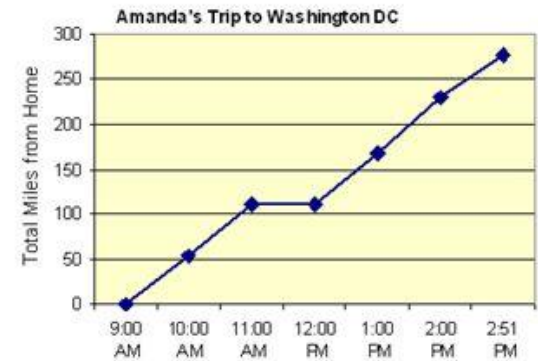
## -BAR GRAPHS



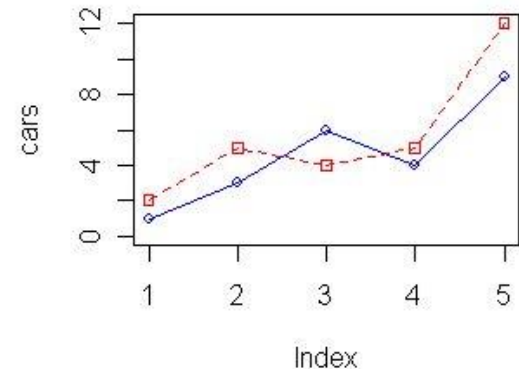
## -PIE CHARTS



## -LINE GRAPHS



### **Autos**







# Bar Graph

- Bar graph atau grafik batang merupakan grafik untuk menampilkan data kategorik. Bar chart digunakan jika kita ingin membandingkan data antar kategori dan melihat frekuensi atau besaran dari kategori/variabel tersebut. Bar chart dapat disajikan dalam bentuk vertikal maupun horizontal.



# Pie Chart

- Pie chart atau diagram lingkaran merupakan grafik yang paling sederhana dan paling sering digunakan dalam memvisualisasikan komposisi dari data kita. Pie chart dapat digunakan untuk data nominal dan ordinal (data kategorik). Diagram lingkaran mewakili angka dalam persentase, dan jumlah total semua segmen harus sama dengan 100%. Yang perlu diperhatikan dalam pembuatan pie chart adalah jangan memasukkan terlalu banyak kategori agar dapat terlihat perbedaan antar komposisinya.





# Line Chart

- Line chart atau grafik garis merupakan chart yang pada umumnya digunakan untuk mengilustrasikan trend atau menggambarkan hubungan dalam bagaimana data berubah dalam periode waktu tertentu. Kita dapat melihat pola dan fluktuasi dalam data, membandingkan variabel, dan membuat proyeksi dari data kita. Tipe data yang dapat digunakan dalam line chart umumnya adalah kontinum, namun line chart juga dapat digunakan untuk data diskrit (misalnya jumlah data per kategori variable dalam beberapa periode waktu)



UNIVERSITAS AIRLANGGA  
Excellence with Morality

# Grafik yang menyesatkan → jangan ditiru

Terlihat seperti terjadi peningkatan yang luar biasa.

Tetapi perhatikan sumbu y tidak dimulai dari 0

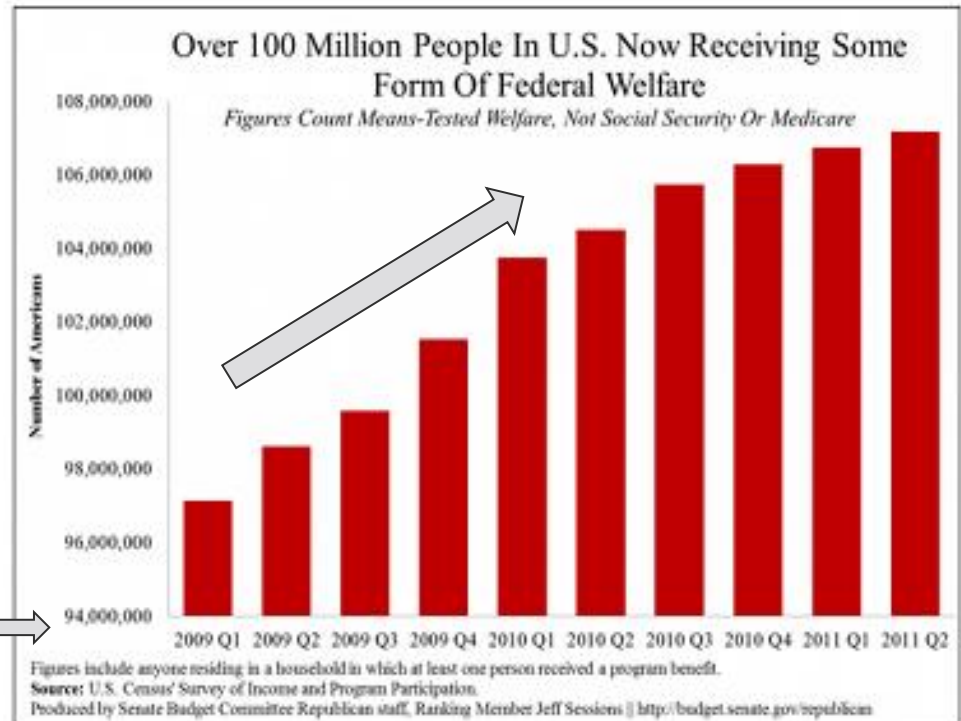
## THE BLOG

### Over 100 Million Now Receiving Federal Welfare

2:40 PM, AUG 8, 2012 - BY DANIEL HALPER

ISSUE PAGE PRINT LARGER TEXT SMALLER TEXT ALERTS

A new chart set to be released later today by the Republican side of the Senate Budget Committee details a startling statistic: "Over 100 Million People in U.S. Now Receiving Some Form Of Federal Welfare."



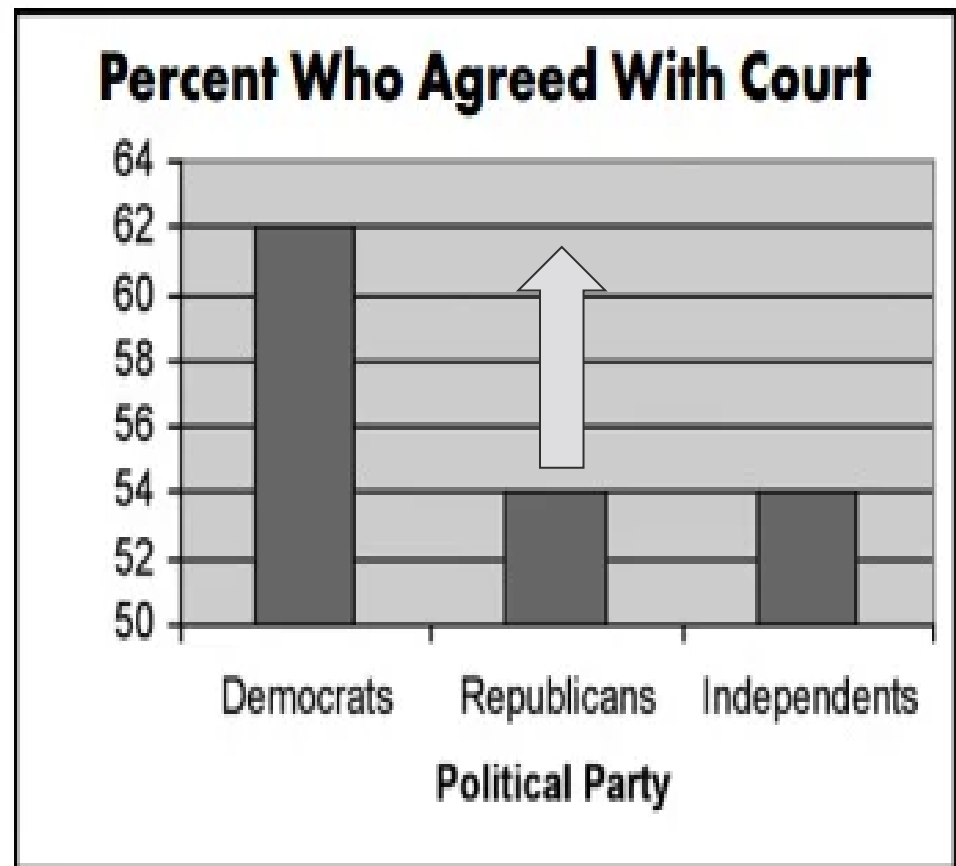




# Grafik yang menyesatkan → jangan ditiru

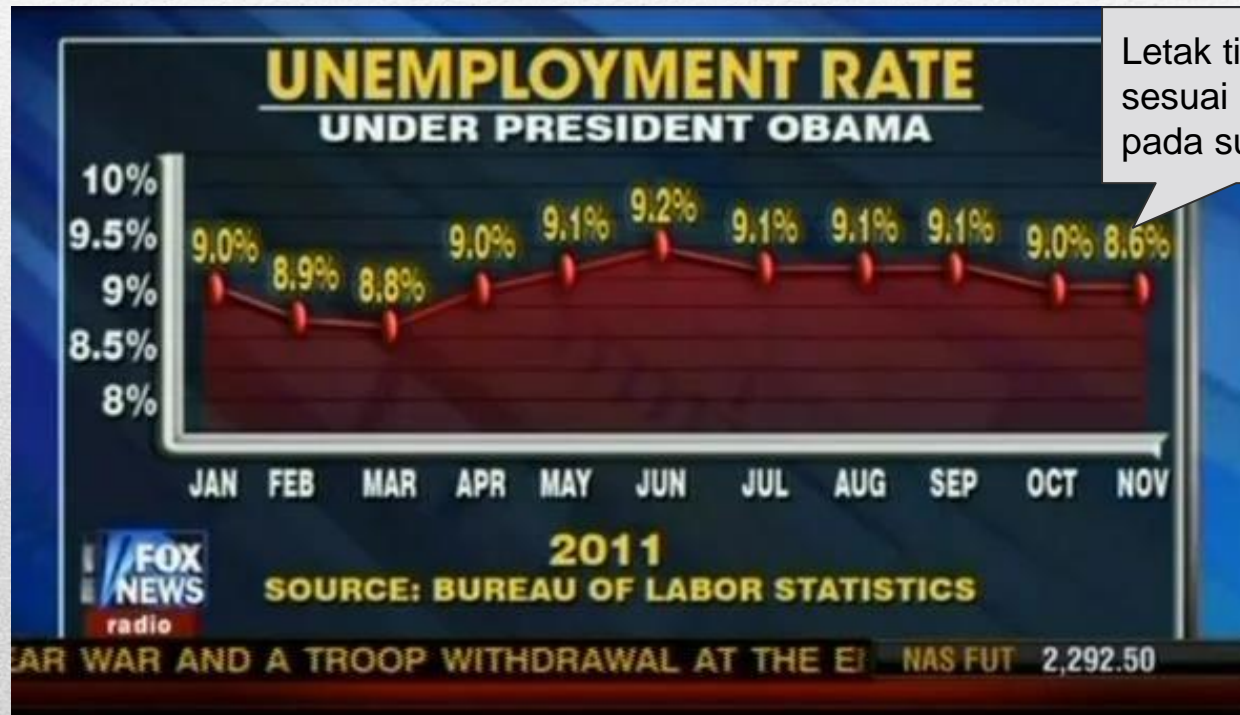
Terlihat seperti ada  
perbedaan 3 kali lipat  
antara demokrat dengan  
republikan dan  
independen

Tetapi perhatikan sumbu  
y tidak dimulai dari 0





# Grafik yang menyesatkan → jangan ditiru



Nilai 8.6% tidak diletakkan sesuai skala pada sumbu Y sehingga terlihat tidak terjadi penurunan

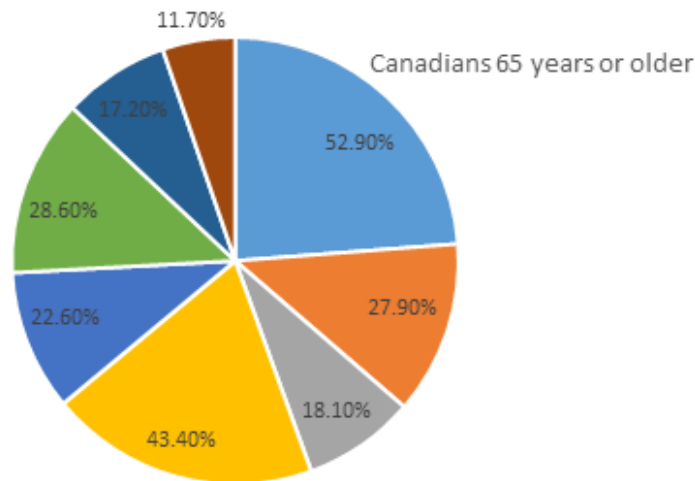




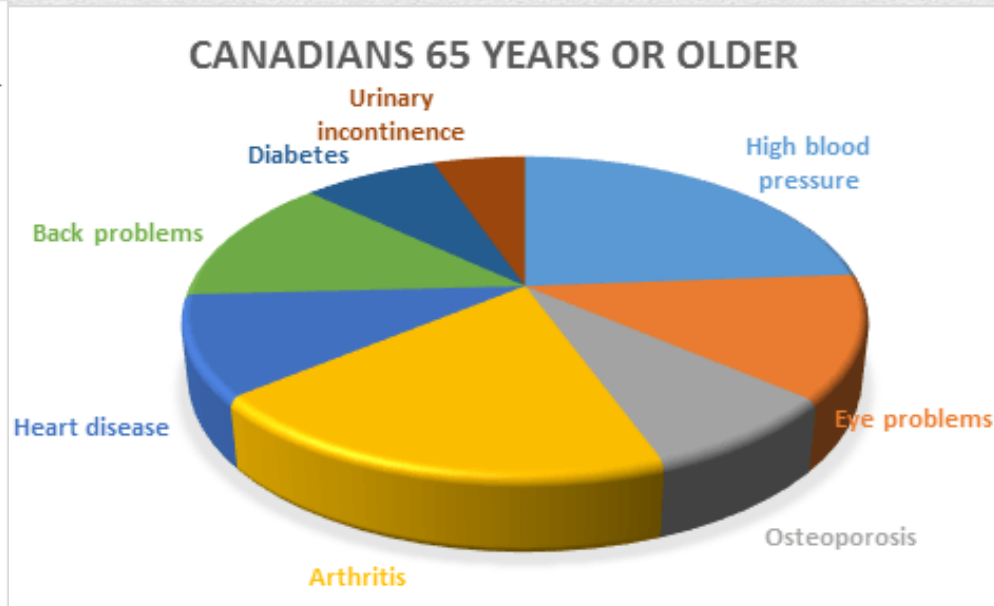
# Grafik yang menyesatkan → jangan ditiru

Pie chart tidak berjumlah 100%

Grafik 3 dimensi membuat arthritis terlihat lebih banyak daripada high blood pressure



■ High blood pressure ■ Eye problems ■ Osteoporosis  
■ Arthritis ■ Heart disease ■ Back problems  
■ Diabetes ■ Urinary incontinence





## RATIO & INTERVAL

- Data interval & ratio : tidak berupa kategori
- Data interval adalah, data yang didapatkan dengan cara pengukuran. Dimana jarak antara dua titik di dalam skala sudah diketahui.
- Ukuran rasio atau data rasio yaitu ukuran yang memberikan keterangan mengenai nilai absolut dari objek yang diukur.





- Contoh III :
  - Suhu
  - Kecerdasan (IQ)
  
- Contoh IV :
  - Berat,
  - tinggi,
  - jarak,
  - waktu,
  - volume



- **Pertanyaan :**

1. Apakah anda melihat perbedaan contoh III & IV ?
2. Contoh mana yang merupakan data interval & contoh mana yang data ratio ?
3. Apa yang membedakan data interval & data ratio ?
4. Diskusikan ! Buat contoh-contoh yang lain !





- Anda dapat membedakan kedua variabel tsb berdasarkan **"nilai nolnya"**;
- Apabila data pengukuran mempunyai nilai nol alami/mutlak, maka disebutnya sebagai data ratio (misalnya berat & tinggi badan, jarak);
- Apabila data tidak mempunyai nilai nol alami/mutlak, disebut data Interval (misalnya suhu, IQ)



## Case Notes

Name: Alice Nightingale

Sex: Female

Occupation: Scientist

Closest Friend : 67ch

Age: 25

Height: 1.72 m (5ft 8)

Weight: 65kg (143 lb)

College Grades: A, A, A, A



## Initial Scores

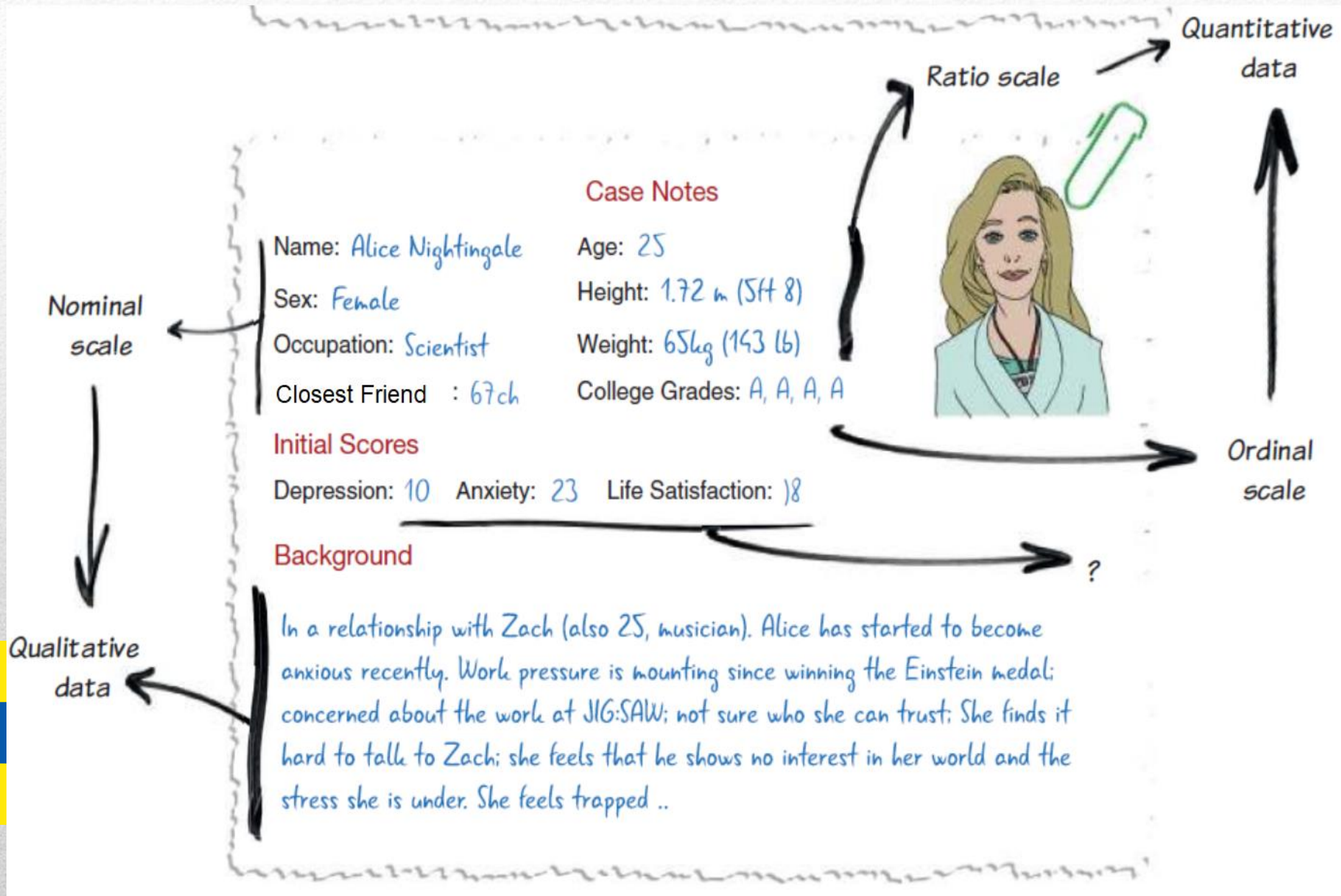
Depression: 10   Anxiety: 23   Life Satisfaction: 18

## Background

In a relationship with Zach (also 25, musician). Alice has started to become anxious recently. Work pressure is mounting since winning the Einstein medal; concerned about the work at JIG:SAW; not sure who she can trust; She finds it hard to talk to Zach; she feels that he shows no interest in her world and the stress she is under. She feels trapped ..

Skala data apa saja yang ada dalam contoh diatas?

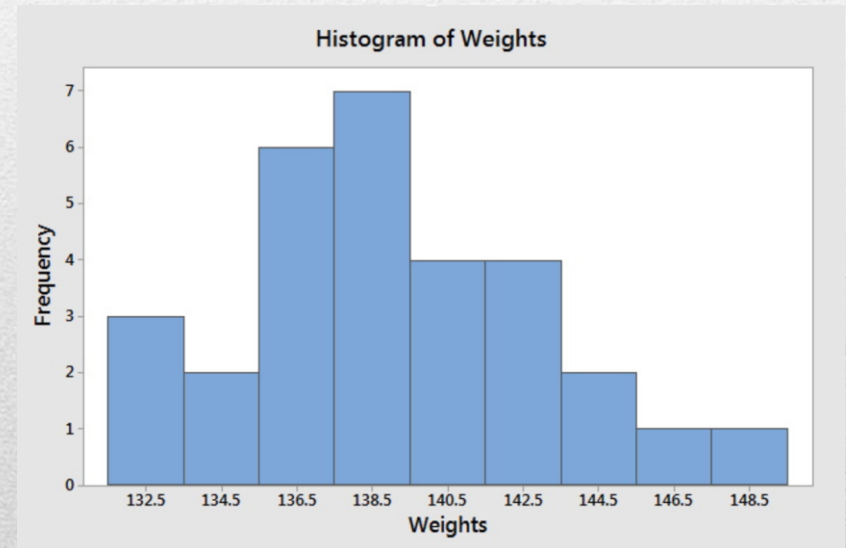






# Visualisasi Data

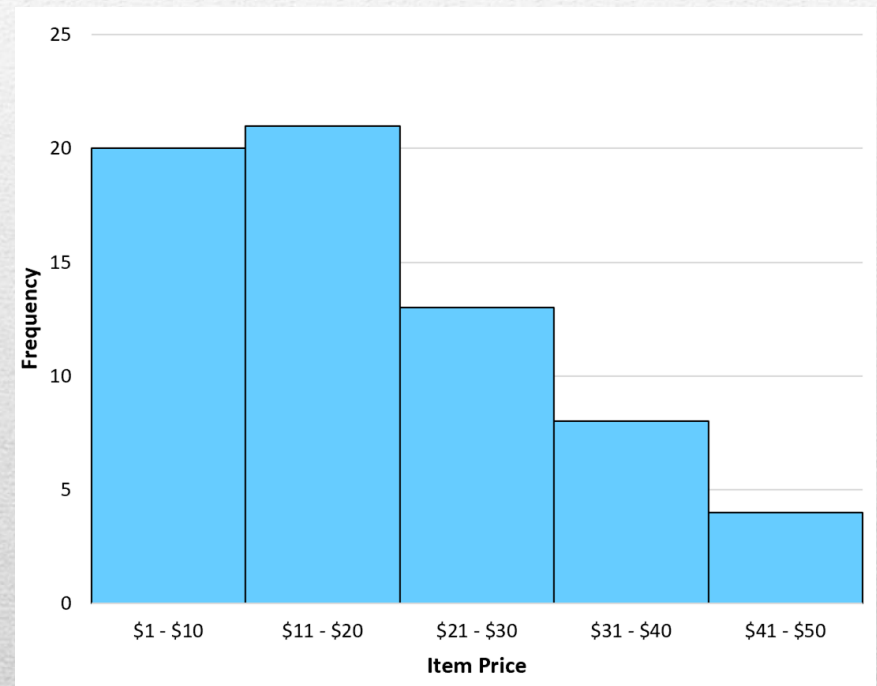
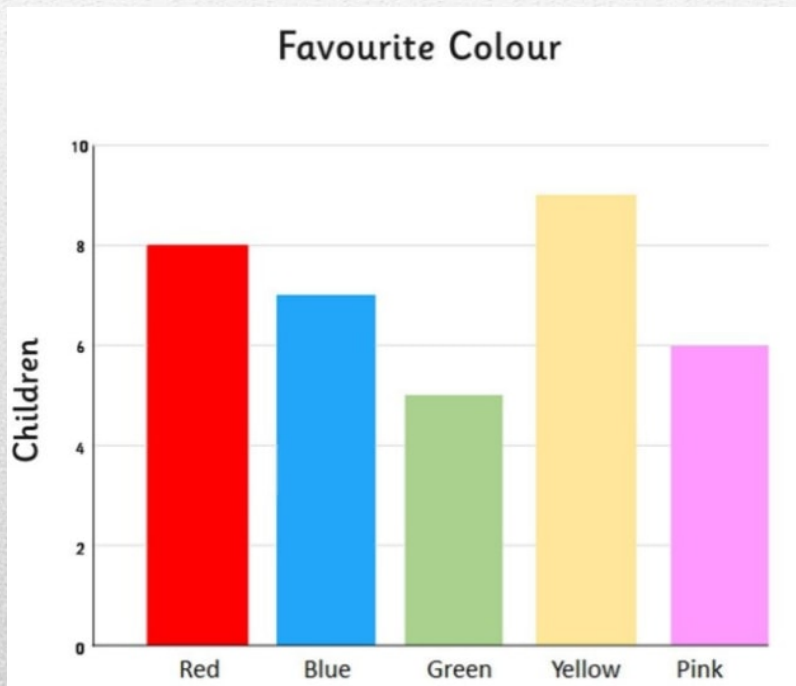
- Histogram, grafik ini ditampilkan dalam bentuk beberapa bar atau batang yang mewakili rentang-rentang data. Tinggi setiap bar menunjukkan seberapa banyak data ada di rentang tersebut.







# Beda Histogram dengan bar chart

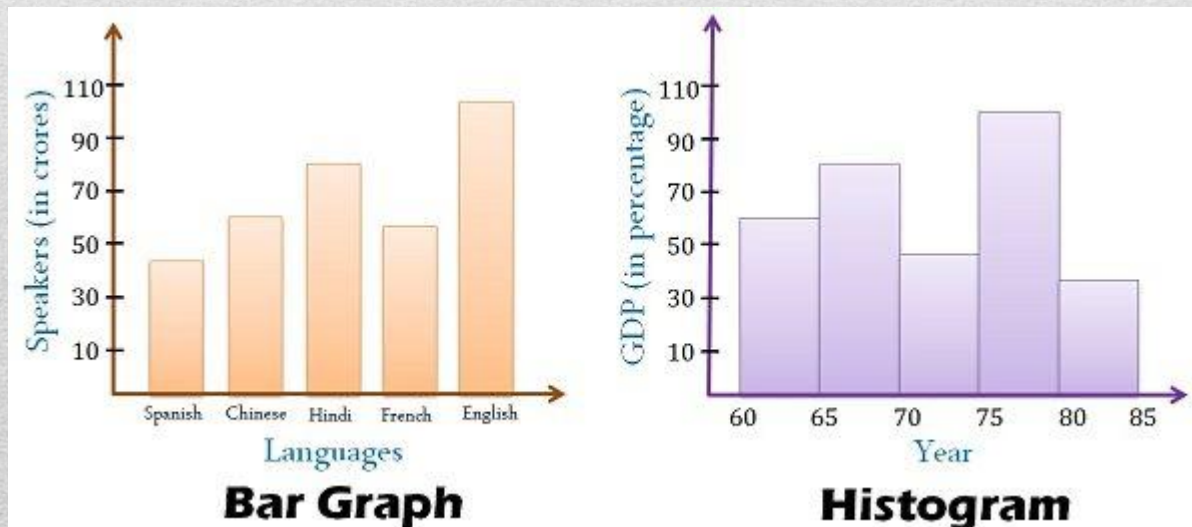


Diskusikan : Mana yang merupakan bar chart, mana yang histogram? Kenapa?



# Beda histogram dengan bar chart

- Histogram untuk variabel numerik, bar chart untuk variabel kategorik
- Bentuk histogram lebih padat dan rapat.







# Tahapan Menyusun Histogram

1. Menyusun distribusi frekuensi. Langkah menyusun distribusi frekuensi
  - a. Menentukan jumlah kelas
  - b. Memasukkan angka-angka dalam kelas yang sesuai kemudian menghitung frekuensinya
  - c. Membuat tabel distribusi frekuensi
2. Menyusun histogram berdasarkan distribusi frekuensi



# Cara Penentuan jumlah kelas

1. Jumlah kelas hendaknya jangan terlalu kecil dan terlalu besar
2. Besar interval dalam tiap kelas hendaknya sama serta dalam bilangan yang praktis
3. Penentuan batas kelas sebaiknya diusahakan tidak ada satupun data yang tidak dapat dimasukkan ke dalam kelas-kelas dan tidak terdapat keragu-raguan dalam memasukkan data kedalam kelas-kelas





# Penentuan Jumlah Kelas

Penentuan jumlah kelas menurut Sturges (1926)

$$k = 1 + 3,322 \log n$$

Dimana :

$k$  = jumlah kelas

$n$  = jumlah data

Rumus besar interval kelas :  $i = \frac{\text{jarak}}{k}$

Jarak = nilai terbesar – nilai terkecil



# Penentuan batas kelas

1. Penentuan jarak untuk menentukan jumlah kelas dan interval kelas sebaiknya dilakukan atas dasar perbedaan angka terendah yang telah mengalami pembulatan ke bawah dan angka tertinggi yang telah mengalami pembulatan ke atas
2. Pengulangan penggunaan batas kelas atas bagi batas kelas bawah berikutnya dihindari
3. Batas kelas sebaiknya menggunakan bilangan bulat





# Contoh penggunaan aturan Sturges.

Diketahui data-data sebagai berikut:

42 62 59 46 53 60 65 72 51 53 61 63 69 64 65 67 72 91 52 59  
69 72 88 65 65 75 67 78 81 47 79 55 62 32 71 67 62 54 66 69

- $N = 40$  data
- Data terkecil = 32, data terbesar = 91
- Range =  $91 - 32 = 59$
- Banyak kelas  $k = 1 + 3,322 \log 40 = 6,3$ . Dibulatkan  $k = 6$  kelas
- Panjang interval kelas =  $59/6 = 9,8$  dibulatkan menjadi 10
- Batas bawah kelas pertama = 32, batas atas kelas pertama = 41



Rumus Sturges tidak sesuai jika digunakan dalam data yang jumlahnya terlalu sedikit atau terlalu banyak

Penentuan jumlah kelas, interval kelas dan batas kelas sangat tergantung dengan peneliti





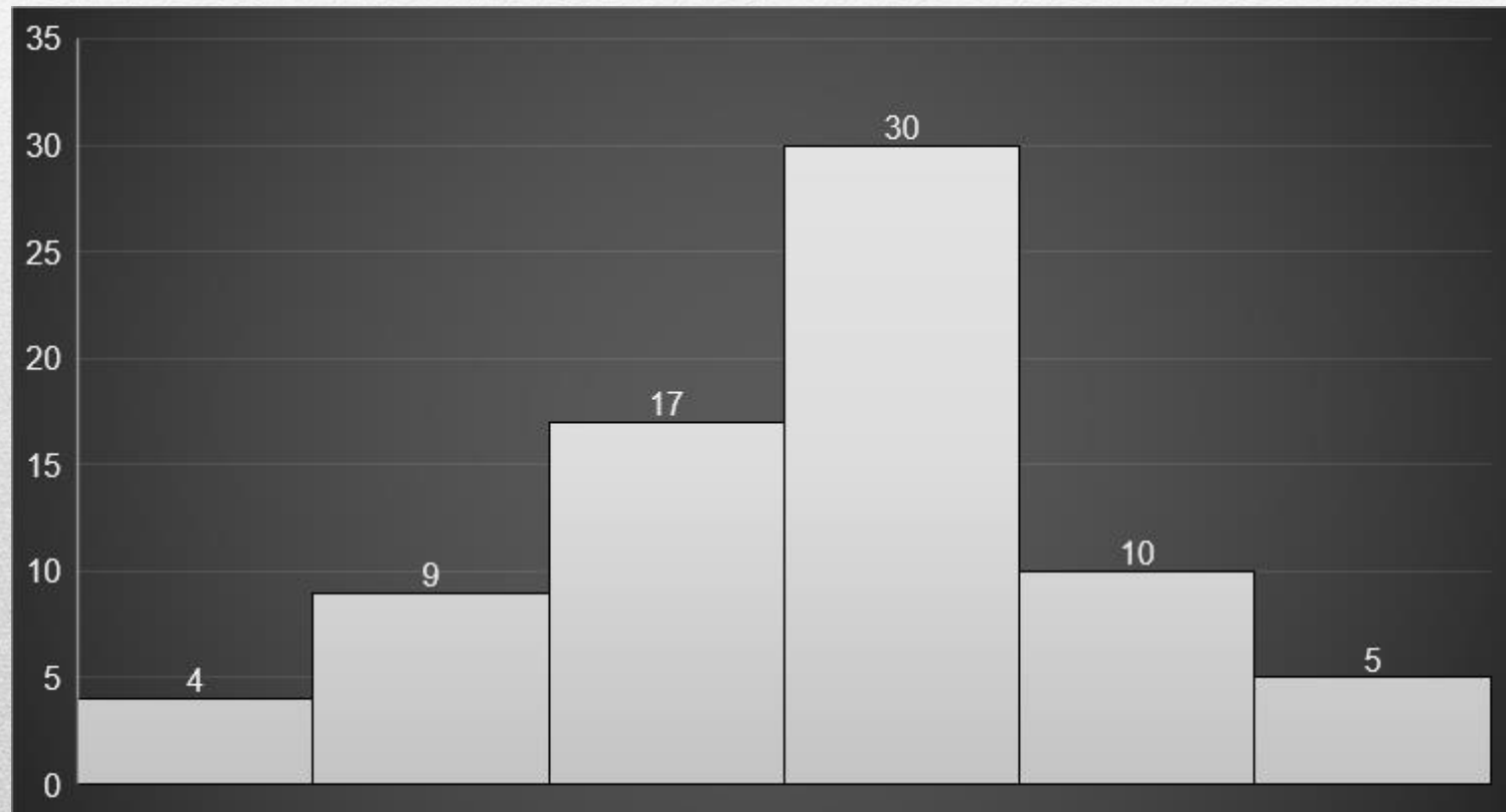


## Tabel Distribusi Frekuensi

Nilai Ujian		Jumlah Mahasiswa
20,00 – 29,99	IIII	4
30,00 – 39,99	IIII IIII	9
40,00 – 49,99	IIII IIII IIII II	17
50,00 – 59,99	IIII IIII IIII IIII IIII IIII	30
60,00 – 69,99	IIII IIII	10
70,00 – 79,99	IIII	5
Jumlah		75



# Histogram Frekuensi





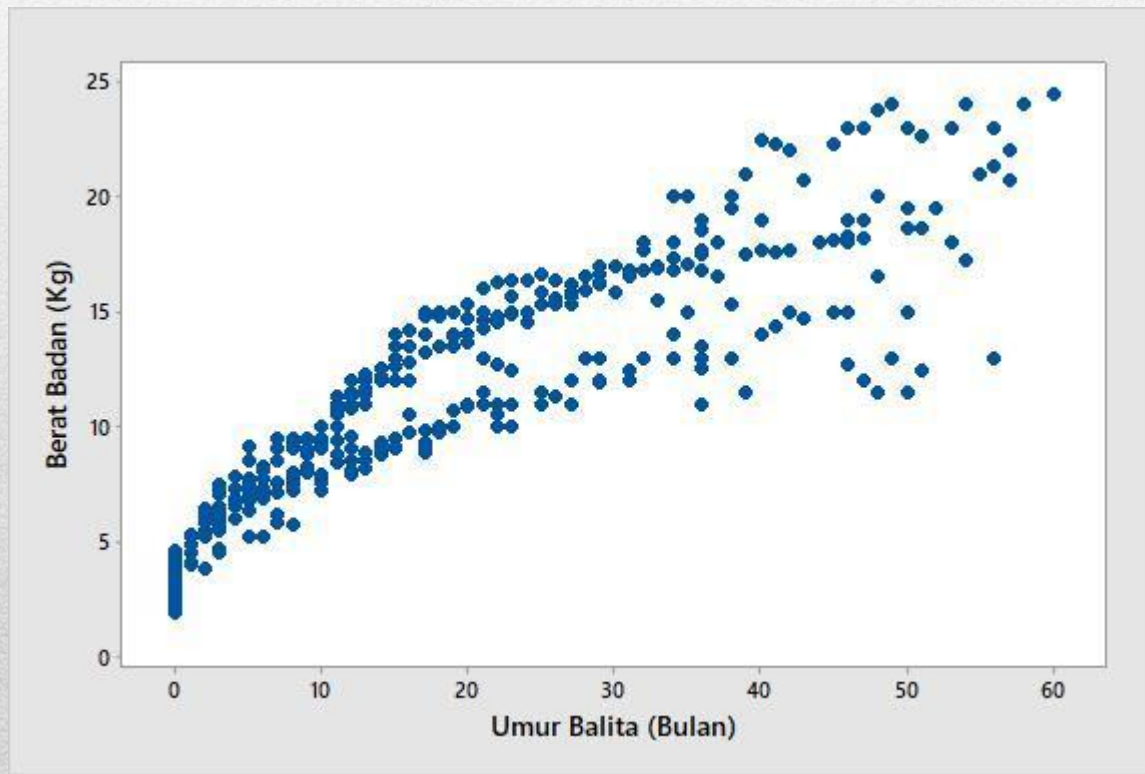


# Visualisasi Data : Scatterplot

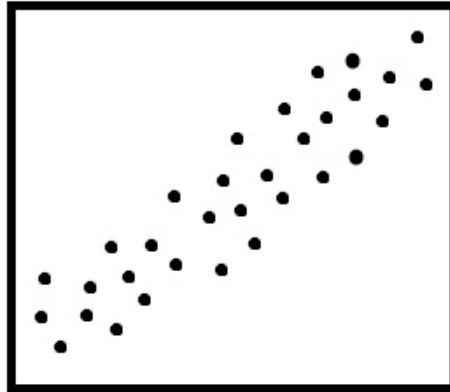
- Scatterplot adalah salah satu tipe grafik yang menampilkan titik-titik dalam koordinat Cartesian untuk menunjukkan hubungan antar dua dataset.
- Misalnya, pada suatu kecamatan dikumpulkan data berat badan dan usia sejumlah balita. Data ini digambarkan di scatterplot berikut. Sumbu X menunjukkan umur balita, dan sumbu Y menunjukkan berat badan. Terlihat ada banyak balita yang berusia 10 bulan ternyata memiliki berat badan antara 7 hingga 10 kg



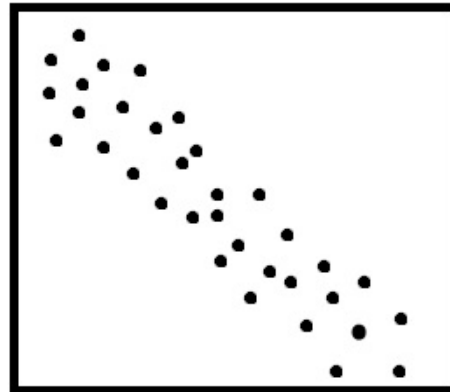
# Contoh Scatter Plot



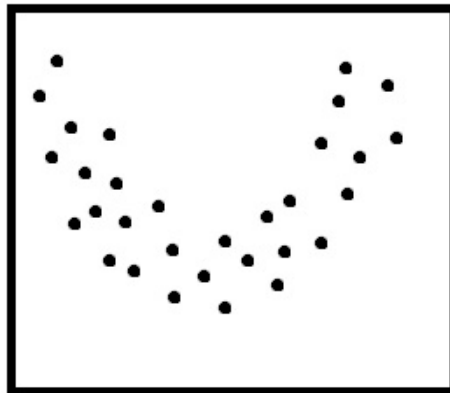




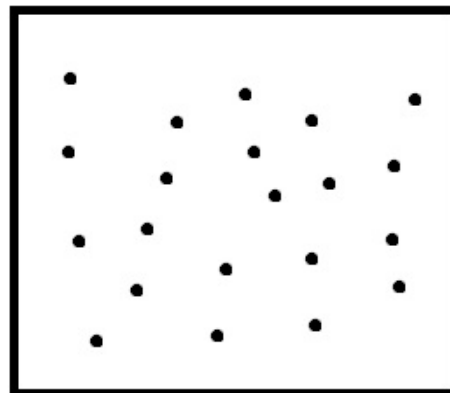
**positive linear  
association**



**negative linear  
association**



**nonlinear  
association**



**no association**



# Visualisasi Data : Stem and Leaf Plot

- Stem-leaf plot merupakan alat untuk menyajikan data kuantitatif dalam format grafis, mirip dengan histogram, yaitu untuk **membantu dalam memvisualisasikan bentuk distribusi data yang sering digunakan dalam analisis eksplorasi.**
- Stem-and-leaf plot memberikan informasi lebih banyak tentang nilai yang sebenarnya dibanding histogram. Seperti dalam histogram, panjang setiap batang sesuai dengan jumlah kejadian yang jatuh ke dalam interval tertentu. Pada Histogram, kita hanya bisa melihat nilai frekuensi dari data namun kita tidak tahu berapa nilai angka sebenarnya. Berbeda dengan histogram, pada SLP selain kita bisa mengetahui nilai frekuensinya, kita pun bisa tau berapa nilai data sebenarnya. Hal ini dilakukan dengan membagi nilai-nilai yang diamati menjadi dua komponen, stem dan leaf.





- **Stem-and-leaf plot** menggambarkan/menyajikan data dengan cara memisahkan setiap nilai menjadi dua bagian: bagian batang (*stem*) yaitu digit angka paling kiri dan diikuti dengan angka berikutnya, yaitu daun (*leaf*), digit angka paling kanan.
- Tujuan utama Stem-and-leaf plot:
  - Apakah pola pengamatan simetris?
  - Penyebaran atau variasi dari data pengamatan.
- Apakah terdapat pencilan (outlier, nilai-nilai yang berada jauh dari yang lainnya).
- Titik pemusatan data.
- Ada Lokasi yang merupakan gap (kesenjangan dalam data)



# Stem and Leaf Plots

These are Mary's test scores:

72, 49, 62, 58, 73, 55, 78, 83, 57, 63, 73, 73, 75, 85, 85,  
64, 61, 67, 75, 91

The stem-and-leaf plot for her scores is shown below.

Stem	Leaf
4	9
5	5 7 8
6	1 2 3 4 7
7	2 3 3 3 5 5 8
8	3 5 5
9	1





# Keterangan Stem and Leaf Plot

UNIVERSITAS AIRLANGGA  
Excellence with Morality

<b>Diagram Dahan Daun</b>	<b>Keterangan</b>
4    9	Memuat data 49
5    578	Memuat data 55,57,58
6    12347	Memuat data 61,61,63,64,67
7    2333558	Memuat data 72,73,73,73,73,75,75,78
8    355	Memuat data 83,85,85
9    1	Memuat data 91



21	30	34	41	43	46	50	57	61
23	31	35	41	44	47	52	58	67
23	32	35	42	45	48	52	59	70
25	33	35	42	45	48	55	59	77
26	33	38	43	46	50	57	60	78

2	133
2	56
3	012334
3	5558
4	1122334
4	5566788
5	0022
5	577899

Contoh *stem and leaf plot* dengan 2 baris stem yang digunakan. Banyaknya stem bergantung dengan jumlah data. Peneliti yang memutuskan berapa banyak stem yang digunakan.





# Diskusi

Apa perbedaan visualisasi data dengan skala data kategorikal (nominal & ordinal) dan skala data kuantitatif (interval & ratio)?



# Jenis Data & Fungsinya

JENIS DATA	FUNGSI
Nominal	Untuk identifikasi, klasifikasi & kategorisasi.
Ordinal	Penjenjangan kualitatif.
Interval	Dapat dihitung jaraknya.
Ratio	Memiliki angka nol mutlak





## ***SOME ASPECTS OF SCALES OF MEASUREMENT***

	<b><i>Nominal</i></b>	<b><i>Ordinal</i></b>	<b><i>Interval</i></b>	<b><i>Ratio</i></b>
Mathematical properties	Identity	Identity Magnitude	Identity Magnitude Equal interval	Identity Magnitude Equal interval True zero point
Mathematical operations	None	Rank order	Add Subtract	Add Subtract Multiply Divide
Type of data	Nominal	Ordered	Score	Score
Typical statistics used	Chi-square	Sign test Mann-Whitney U-test	t-test Anova	t-test Anova
Examples	Diagnostic categories Brand name	Socioeconomic class Rank	IQ test scores Personality & attitude scales	Weight Length Reaction time <sub>45</sub>



UNIVERSITAS AIRLANGGA  
Excellence with Morality

TERIMA KASIH