Sumber Gambar: https://www.dreamstime.com/royalty-free-stock-photos-statistics-background-concept-wordcloud-illustration-print-concept-word-cloud-graphic-collage-image35558718

STATISTIKA DASAR

< PERTEMUAN 02 >

PROGRAM PRAKTISI MENGAJAR

BATCH 3 TAHUN 2023

TIM DOSEN

1. **Putri** Sukma Dewi, S.Pd., M.Pd. (*Dosen Pengampu*)
2. Eko **Teguh** Widodo, SST., M.Sc. (*Dosen Praktisi*)

EKO TEGUH WIDODO

Working Experience

- *Senior IT Specialist*, BPS-Statistics Indonesia (Provincial Branch Office)
- *IT Specialist*, BPS-Statistics Indonesia (Provincial Branch Office)
- *Statistician*, BPS-Statistics Indonesia (Municipal and Provincial Branch Office)
- *IT Engineering Intern*, Sekolah Tinggi Ilmu Statistik (Institute of Statistics)
- *Freelancer* as Private Instructor/Tutor (Statistics, Mathematics, Information Technology), Software Engineer (Web, Desktop)

Education

- *Master of Science (M.Sc.)* in Information and Telecommunication Technology, Korea Advanced Institute of Science and Technology (KAIST), Republic of Korea
- *Bachelor of Applied Science (B.A.Sc., SST.)* in Computational Statistics, Sekolah Tinggi Ilmu Statistik (Institute of Statistics), Jakarta

Core Competency

- Statistical Data Analysis, Big Data and Software Engineering, Natural Language Processing, Knowledge Graph, Digital Transformation, Information Technology Management and Policy

CONCEPTS

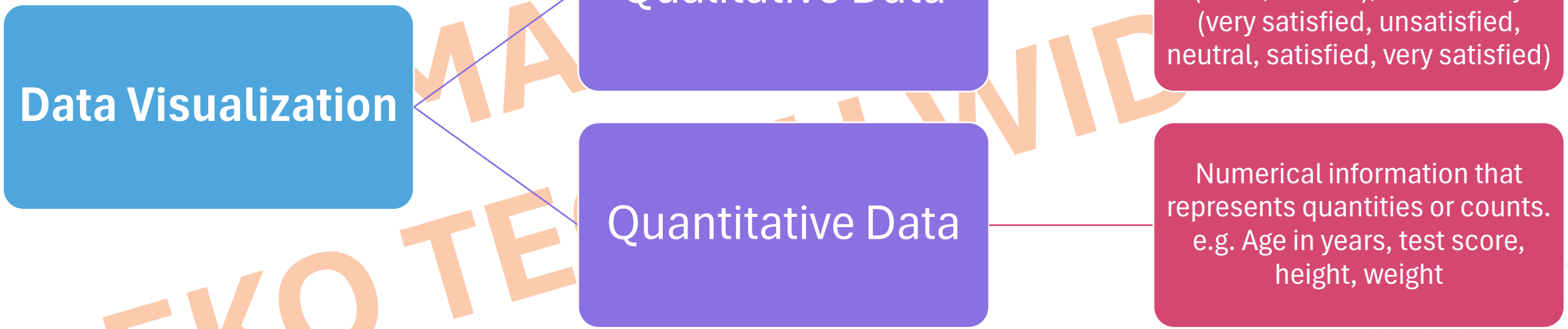
- Visualizing Qualitative and Quantitative Data
- Data Visualization Practices in BPS-Statistics Indonesia

PRACTICES

- Data Visualization in R
- Demonstration and Coding Practice

STATISTIKA DASAR

DATA VISUALIZATION



✓Data Processing:

- Data Processing → transforming data from documents
- Objectives:
  - Transforming data scattered in the documents to organized in terms of structure and format
  - Producing a brief/simple information
  - Simplifying for statistical calculations

IDENTITAS WILAYAH			IDENTITAS PRIBADI	VARIABEL/KARAKTERISTIK				
/kota	urut responden	Alamat	Nama	Sex	Umur	Status kawin	Pendidikan tertinggi	Kegiatan
2	1	Jalan Abdullah	Nasrullah Abidin	L	32	2	7	1
2	2	Jalan Teuku Umar	Cut Nasyilah	P	21	1	5	2
2	3	Jalan Iskandar Muda	Miriam Abdillah	P	33	2	6	3
71	1							

✓Objectives of Data Visualization

- Delivering the information in interesting way
- Delivering the information which is easy to understand
- Readers are easy to capture message conveyed

✓Forms of Data Visualization

- **TABLE**: organizing observation unit into (some) groups corresponding with cohort and individual characteristics
- **DIAGRAM/CHART**: serving group of data into visual forms, such as line, bar, pie, radar, map, or its combination, therefore the direction, pattern, and disparity can be seen clearly

✓TABLE

- Collections of information arranged in rows and columns based on certain categories.

- **Structure of Table**
  1. *Table Number* → particularly when more than one table is served
  2. *Table Title* → shows detail/content, location, time descriptions
  3. *Table Body Table* → consists of rows and columns
  4. *Table Source* → listed below table

- **Two types of table:**
  1. **Reference Table,** source of all detailed information → ***Raw Data*, originated from data collection process**
  2. **Frequency Distribution Table,** Raw data → grouping (interval classes/categories, etc.) → ***Tabulation***

✓Examples of Reference Table and Frequency Distribution Table

nama_kab	kode_kec	nama_kec	kode_des	nama_des	krt_dsrt	b4_k5	b4_k6	b4_k7_bl	b4_k7_th	b4_k8	b4_k9	b4_k10	b5_rinfo
JAKARTA UTARA	010	PENJARINGAN	005	PLUIT	KAHEM JAPUTRA JAP	3	1	3	1952	67	3	2	1
JAKARTA UTARA	010	PENJARINGAN	005	PLUIT	KAHEM JAPUTRA JAP	3	2	12	1963	55	3	2	1
JAKARTA UTARA	010	PENJARINGAN	005	PLUIT	KAHEM JAPUTRA JAP	3	2	9	1982	36	3	1	1
JAKARTA UTARA	010	PENJARINGAN	005	PLUIT	KAHEM JAPUTRA JAP	3	1	1	1991	28	3	1	1
JAKARTA UTARA	010	PENJARINGAN	005	PLUIT	KAHEM JAPUTRA JAP	3	2	1	1984	35	3	2	5
JAKARTA UTARA	010	PENJARINGAN	005	PLUIT	TURINO HERRY	3	1	10	1938	80	3	2	2
JAKARTA UTARA	010	PENJARINGAN	005	PLUIT	TURINO HERRY	3	2	11	1945	73	3	2	2
JAKARTA UTARA	010	PENJARINGAN	005	PLUIT	TURINO HERRY	3	2	07	1975	44	3	1	2
JAKARTA UTARA	010	PENJARINGAN	005	PLUIT	TURINO HERRY	3	1	03	1992	27	2	1	2
JAKARTA UTARA	010	PENJARINGAN	005	PLUIT	AMINOTO CHONG	3	1	07	1949	70	3	2	1
JAKARTA UTARA	010	PENJARINGAN	005	PLUIT	AMINOTO CHONG	3	2	11	1950	68	3	2	1
JAKARTA UTARA	010	PENJARINGAN	005	PLUIT	AMINOTO CHONG		2	04	1963	56	3	1	3
JAKARTA UTARA	010	PENJARINGAN	005	PLUIT	AMINOTO CHONG		2	12	1984	34	3	1	3
JAKARTA UTARA	010	PENJARINGAN	005	PLUIT	AMINOTO CHONG		2	09	1988	30	3	2	3
JAKARTA UTARA	010	PENJARINGAN	005	PLUIT	TJUNG INDRADI PUTRA	3	2	07	1958	61	3	4	2
JAKARTA UTARA	010	PENJARINGAN	005	PLUIT	TJUNG INDRADI PUTRA	3	2	07	1982	37	3	1	2
JAKARTA UTARA	010	PENJARINGAN	005	PLUIT	TJUNG INDRADI PUTRA	3	1	04	1986	33	3	1	2
JAKARTA UTARA	010	PENJARINGAN	005	PLUIT	SUDIN								
JAKARTA UTARA	010	PENJARINGAN	005	PLUIT	SUDIN								
JAKARTA UTARA	010	PENJARINGAN	005	PLUIT	SUDIN								
JAKARTA UTARA	010	PENJARINGAN	005	PLUIT	SUDIN								
JAKARTA UTARA	010	PENJARINGAN	005	PLUIT	SUDIN								

Reference Table

Frequency Distribution Table

Tabel 5.4

Jumlah Penduduk Miskin menurut Kabupaten/Kota (Ribu Jiwa)

Tahun 2005 - 2017

Kota/Kabupaten Administratif	Jumlah Penduduk Miskin (Ribu Jiwa)												
	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Kepulauan Seribu	3,40	3,20	2,93	2,56	2,40	2,75	2,47	2,60	2,50	2,70	2,70	2,96	3,09
Jakarta Selatan	64,00	76,31	76,30	71,05	73,71	78,65	71,84	74,10	74,60	80,80	74,50	71,96	69,82
Jakarta Timur	71,20	85,14	94,56	79,77	81,19	91,74	83,82	86,50	86,80	96,50	91,40	91,37	95,67
Jakarta Pusat	28,50	43,59	34,46	31,04	32,15	35,69	32,63	33,70	33,60	37,70	38,00	35,82	34,83
Jakarta Barat	57,40	89,46	85,23	72,85	74,03	87,19	79,71	82,30	83,20	91,00	89,40	84,08	86,96
Jakarta Utara	91,70	109,39	112,21	85,20	76,17	92,66	84,73	87,50	90,90	104,20	103,00	98,11	99,31
DKI Jakarta	316,20	407,10	405,70	379,61	323,17	388,68	355,20	366,70	371,70	412,80	398,92	384,30	389,69

Sumber : Survei Sosial Ekonomi Nasional (Susenas), BPS

✓ Frequency Distribution Table for Qualitative Data

TABEL 4-1 Data mengenai 50 Orang Pembeli Komputer dari Beberapa Jenis Perusahaan Komputer

IBM	Compaq	Compaq	IBM	IBM
Compaq	Compaq	Packard Bell	Gateway 2000	Packard Bell
Apple	Apple	IBM	Apple	Compaq
Packard Bell	Compaq	Compaq	IBM	Packard Bell
Gateway 2000	Apple	Apple	Packard Bell	Compaq
IBM	Apple	Apple	Packard Bell	Packard Bell
Apple	Apple	Compaq	Gateway 2000	Compaq
Packard Bell	IBM	Gateway 200	Compaq	Apple
Packard Bell	IBM	Packard Bell	Compaq	Packard Bell
Gateway 2000	Apple	IBM	Apple	Apple

RAW DATA

TABEL 4-2 Distribusi Frekuensi Pembelian Komputer dari 5 Merek

Perusahaan	Frekuensi
Apple	13
Compaq	12
Gateway 2000	5
IBM	9
Packard Bell	11
Jumlah	50

n

Perusahaan	Frekuensi Relatif	Frekuensi Persentase
Apple	0,26	26
Compaq	0,24	24
Gateway 2000	0,10	10
IBM	0,18	18
Packard Bell	0,22	22
Total	1,00	100

relative frequency x 100

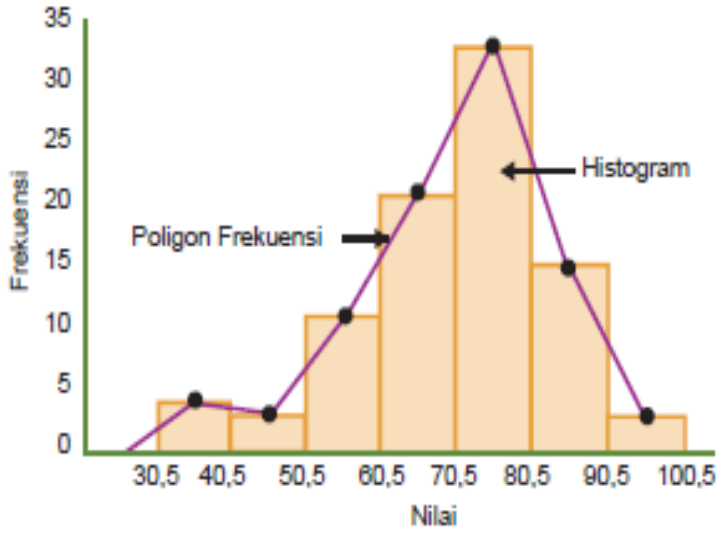
Proportion of the total number of observations associated with each value or class of values

$$\frac{\text{class frequency}}{n}$$

✓ Frequency Distribution Table for Quantitative Data

- Sort the data
- Find the range = largest value – smallest value
- Pick the number of classes to use
- Find the class width (interval)
- Find the class limits
- Find the class boundaries
- Find the class midpoints
- Tally and find the frequency of the data

✓ *Histogram* and *Frequency Polygon*



Number of Classes

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

k = number of classes

n = number of observations

Class Width (Interval)

$$c = (X_n - X_1)/k$$

c = class width

k = number of classes

X_n = largest observation value

X_1 = smallest observation value

Source:
https://stats.libretexts.org/Courses/Highline_College/Book%3A_Statistics_Using_Technology_(Kozak)/02%3A_Graphical_Descriptions_of_Data/2.02%3A_Quantitative_Data

Tabel 1.3 — Table Number

Garis Kemiskinan Daerah Perkotaan Menurut Provinsi, Tahun 2018 (Rupiah) — Table Title

Provinsi	GKM			GKNM			GK		
	Maret 2018	September 2018	Growth	Maret 2018	September 2018	Growth	Maret 2018	September 2018	Growth
11 Aceh	351 777	360 318	2,43	134 561	136 434	1,39	486 338	496 752	2,14
12 Sumatera Utara	328 225	345 167	5,16	120 138	120 623	0,40	448 363	465 790	3,89
13 Sumatera Barat	356 907	367 319	2,92	139 235	140 238	0,72	496 142	507 557	2,30
14 Riau	342 425	350 004	2,21	148 954	149 398	0,30	491 379	499 402	1,63
15 Jambi	358 637	363 153	1,26	129 154	129 211	0,04	487 791	492 364	0,94
16 Sumatera Selatan	313 669	314 492	0,26	124 314	126 557	1,80	437 983	441 049	0,70
17 Bengkulu	380 731	388 651	2,08	136 223	142 004	4,24	516 954	530 655	2,65
18 Lampung	321 620	327 530	1,84	122 626	125 523	2,36	444 246	453 053	1,98
19 Bangka Belitung	442 588	469 430	6,06	180 347	186 719	3,53	622 935	656 148	5,33
21 Kepulauan Riau	373 829	376 277	0,65	190 164	194 633	2,35	563 993	570 910	1,23
31 DKI Jakarta	394 158	408 791	3,71	198 949	198 987	0,02	593 108	607 778	2,47
32 Jawa Barat	263 583	266 340	1,05	105 097	105 920	0,78	368 680	372 260	0,97
33 Jawa Tengah	254 959	259 616	1,83	98 281	99 910	1,66	353 240	359 526	1,78
34 DI Yogyakarta	301 252	305 495	1,41	125 328	126 523	0,95	426 580	432 018	1,27
35 Jawa Timur	282 265	296 062	4,89	102 744	103 272	0,51	385 009	399 334	3,72
36 Banten	316 089	331 096	4,75	132 429	137 477	3,81	448 518	468 572	4,47

Columns

Table Body

Rows

Sumber: Susenas Maret dan September 2018 — Table Source

Type of Data Visualization: TABLE FORMS

1 ONE-WAY
One characteristic

Provinsi	Penduduk Miskin (000)
Aceh	819,44
Sumatera Utara	1282,04
Sumatera Barat	348.22
Riau	490.72
Jambi	274.32
Sumatera Selatan	1073.74
Bengkulu	302.30
Lampung	1063.66
Kep. Bangka Belitung	68.38
Kep. Riau	128.46
DKI Jakarta	365.55

2 TWO-WAY
Two characteristics

CROSS TABULATION

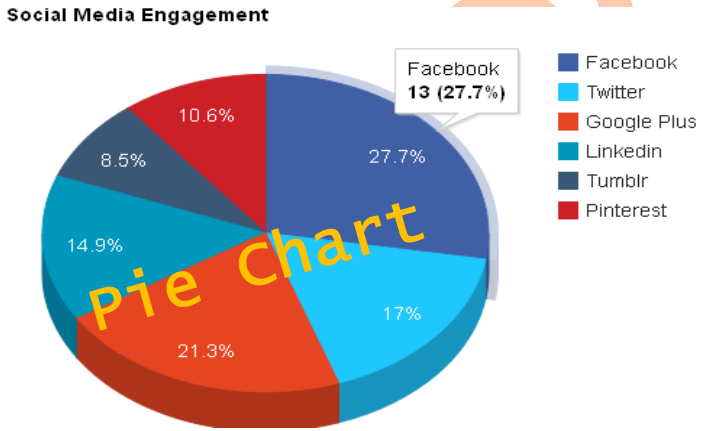
Provinsi	Penduduk Miskin (000)	
	Perkotaan	Perdesaan
Aceh	168.11	651.33
Sumatera Utara	675.74	606.30
Sumatera Barat	121.35	226.87
Riau	175.93	314.79
Jambi	115.08	159.24
Sumatera Selatan	384.53	689.22
Bengkulu	96.52	205.78
Lampung	231.86	831.80
Kep. Bangka Belitung	23.31	45.07
Kep. Riau	104.21	24.25
DKI Jakarta	365.55	-

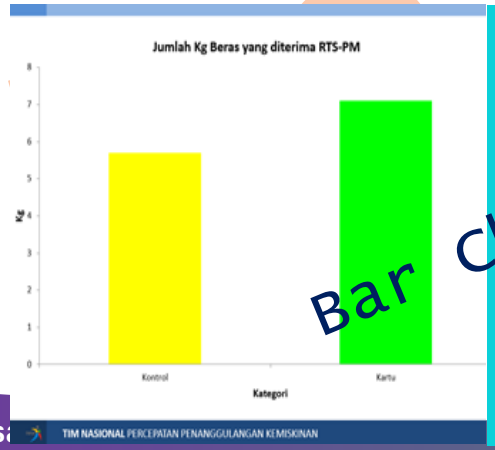
3 THREE-WAY
Three characteristics

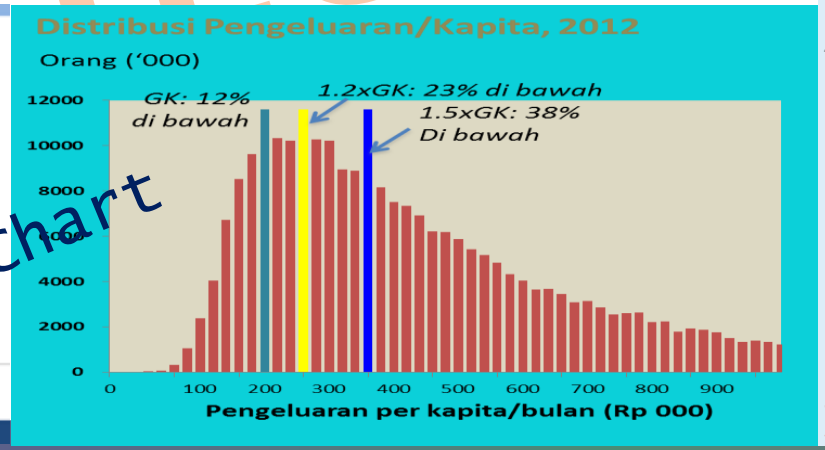
Provinsi	Penduduk Miskin (000)					
	2017		2018		2019	
	K	D	K	D	K	D
Aceh	172.35	700.26	172.09	667.40	168.11	651.33
Sumatera Utara	710.71	743.17	694.85	630.13	675.74	606.30
Sumatera Barat	113.01	251.50	114.84	242.29	121.35	226.87
Riau	178.58	336.03	173.57	326.86	175.93	314.79
Jambi	120.62	165.93	118.62	163.07	115.08	159.24
Sumatera Selatan	375.25	711.67	378.55	689.71	384.53	689.22
Bengkulu	100.84	216.14	97.47	204.34	96.52	205.78
Lampung	228.32	903.41	228.82	868.22	231.86	831.80
Kep. Bangka Belitung	21.47	52.61	24.09	52.18	23.31	45.07
Kep. Riau	91.49	33.88	99.20	32.48	104.21	24.25
DKI Jakarta	389.69	-	373.12	-	365.55	-

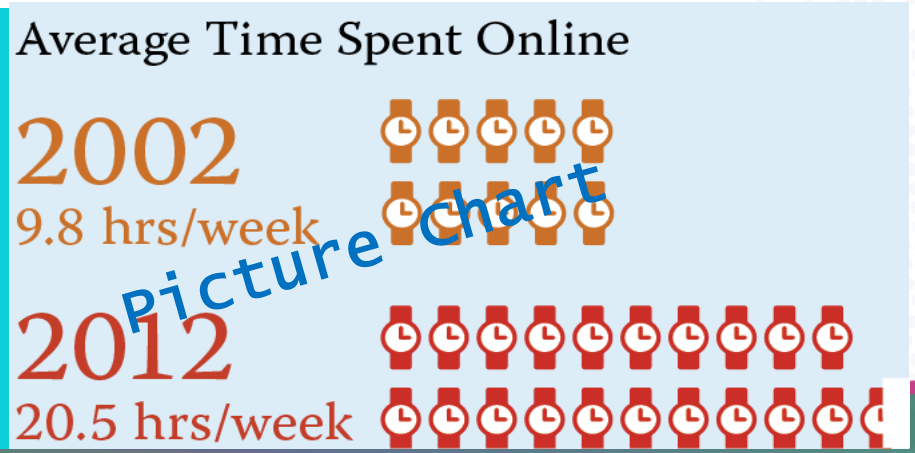
✓DIAGRAM/CHART

- **LINE**, shows the change in a relatively long period of time
- **BAR**, shows the differences between two or more aspects
- **PICTURE/IMAGE**, gives the nuances and reminds the readers about the problems discussed
- **MAP**, gives the description of disparity between region (country, province, city/municipality, village, etc.)



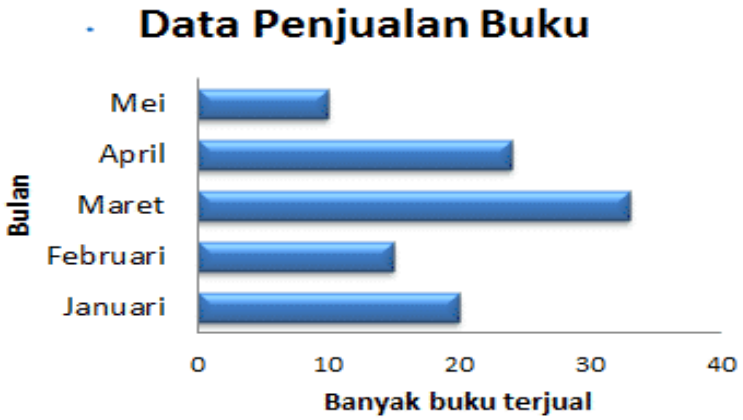


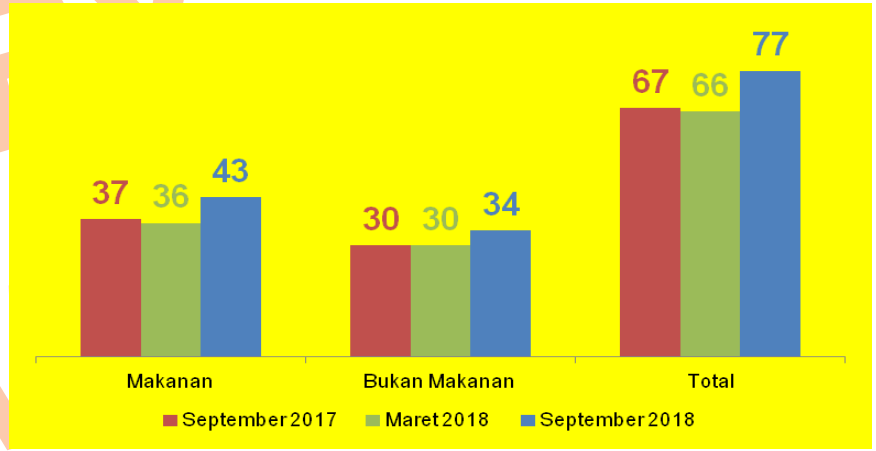


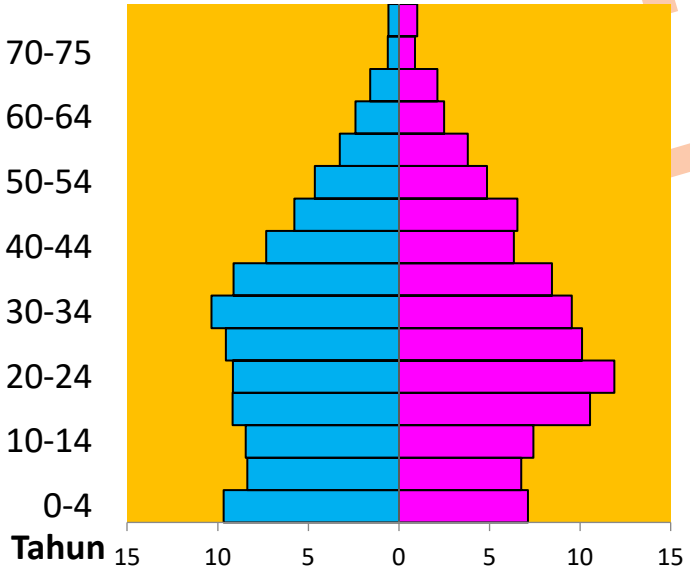


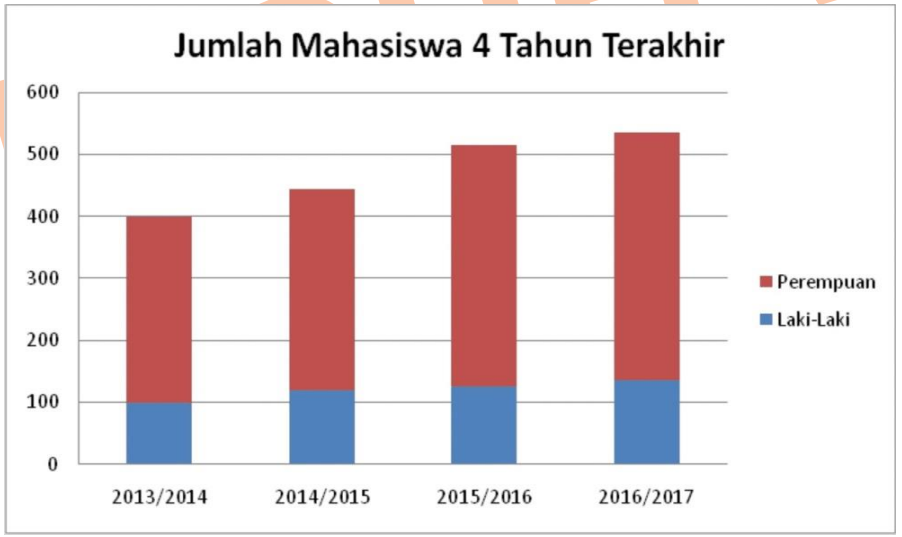
✓ OTHER TYPES OF DIAGRAM/CHART

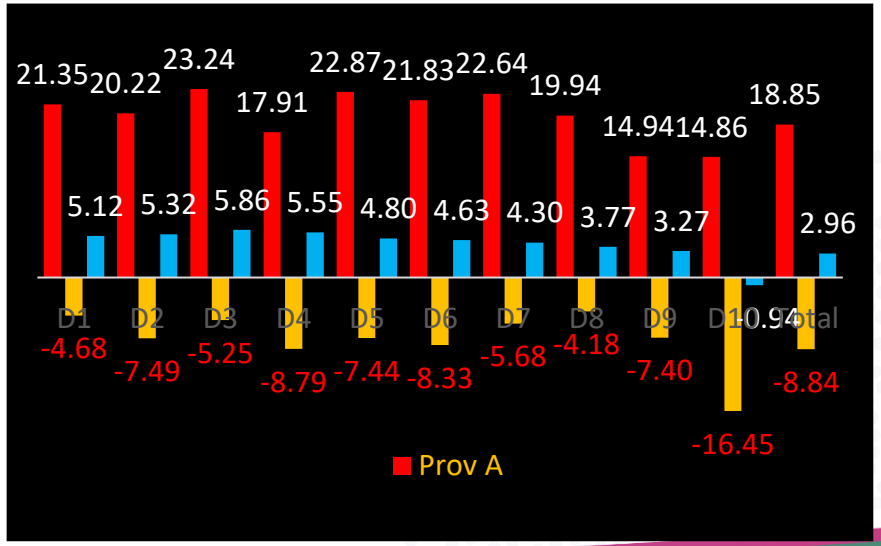
- Group (line/bar) chart
- Single (line/bar) chart
- Stacked (bar) chart
- Net balanced (bar) chart
- Pyramid chart
- ...



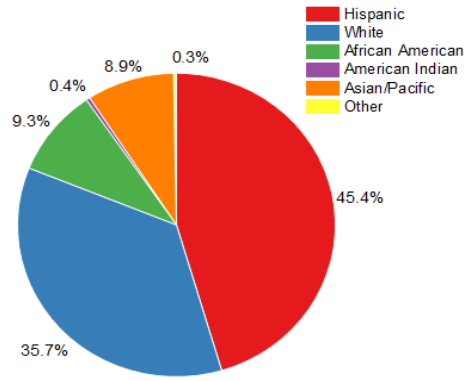


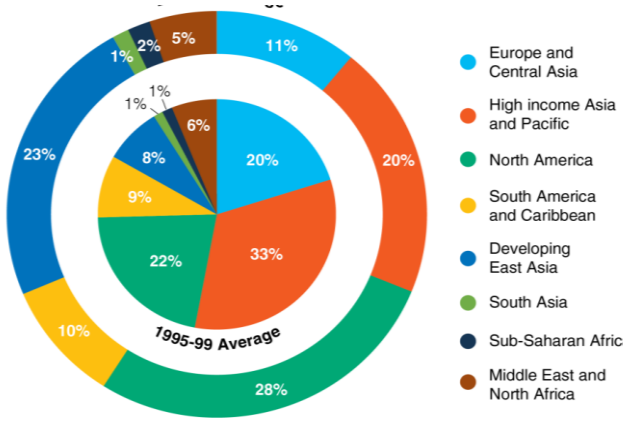






Type of Data Visualization: DIAGRAM/CHART



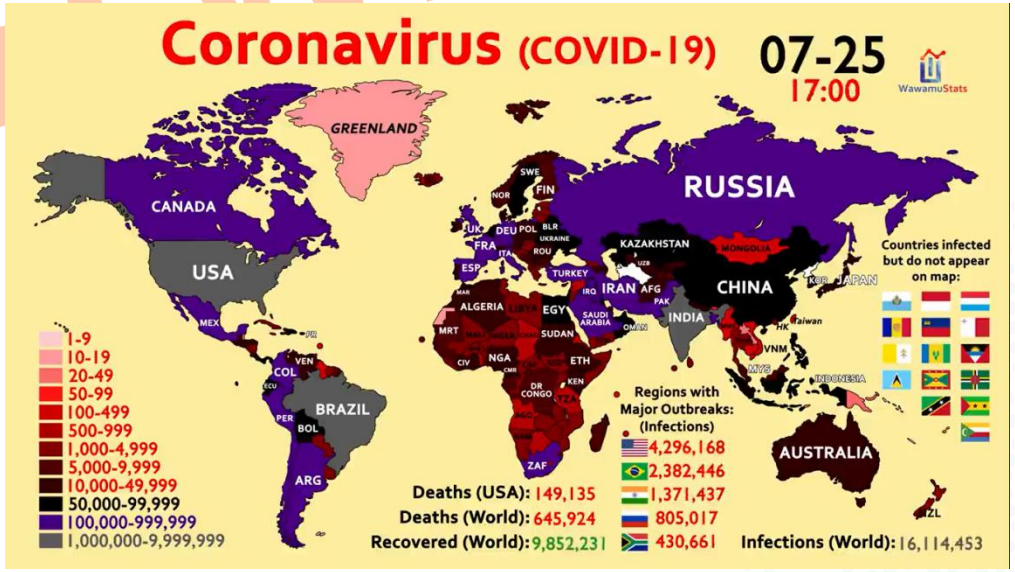


Access to Electricity in Africa

25%
of African homes have access to electricity that **always** works

21%
of African homes have access to electricity that works most or **half of the time.**

54%
of African homes either don't have electricity or have electricity that works **occasionally.**



Type of Data Visualization: INFOGRAPHICS

✓ QUESTIONS

- What kinds of data visualization that the infographics serve?
- What information could we absorb from the infographics?
- **What is infographic?**

Source: https://lampung.bps.go.id/galery.html

- Understand the substance of data
- Depend on the table format: *one-way* or *two-way*, even more; types of chart; which make some information are interrelated
- *Simple*, only describing direction and pattern of the data
- *Advanced* → *simple* + prediction/forecasting (trends/series data) or recommendations to solve the problems

Indikator	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Gini ratio	0,32	0,33	0,33	0,36	0,35	0,37	0,38	0,41	0,41	0,41
Angka pengangguran	9,90	10,26	10,45	9,75	8,46	8,14	7,41	6,80	6,32	5,92
% miskin	16,70	16,00	17,80	16,60	15,40	14,20	13,33	12,49	11,96	11,37
Pertumbuhan ekonomi (%)	5,0	5,7	5,5	6,3	6,0	4,6	6,2	6,5	6,2	5,6

The poverty indicator of province X demonstrates dropped gradually in the last 8 years

R Programming for Statistical Data Analysis

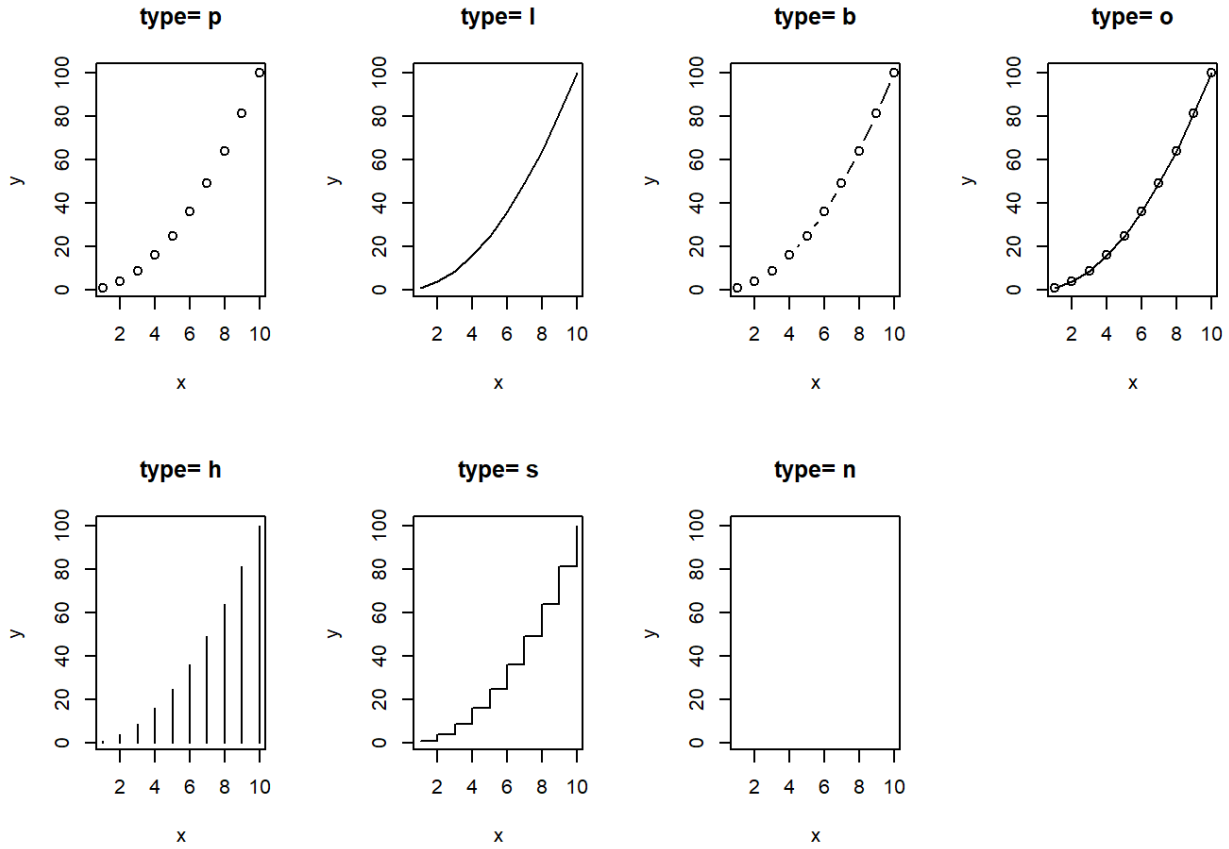
DATA VISUALIZATION

✓Packages for Data Visualization?

- ***plot***

```
# membuat vektor data
x <- c(1:10); y <- x^2
```

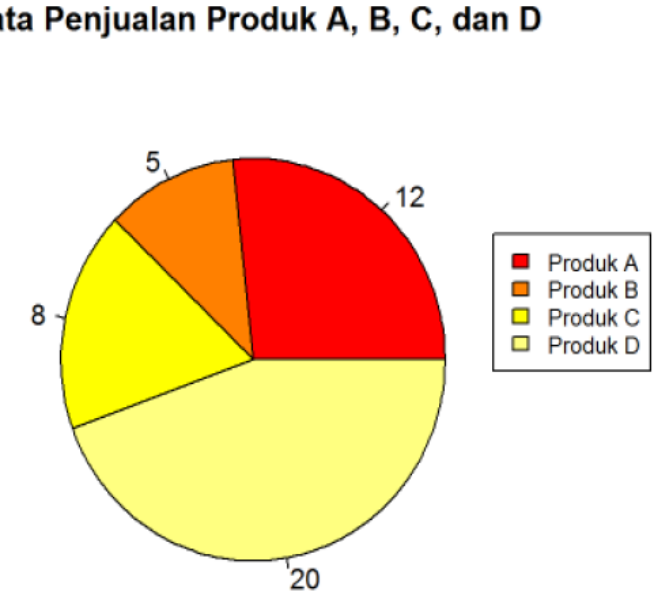
```
# membagi jendela grafik menajdi 2 baris dan 4 kolom
par(mfrow=c(2,4))

# loop
type <- c("p","l","b","o","h","s","n")
for (i in type){
  plot(x,y, type= i,
       main= paste("type=", i))
}
```

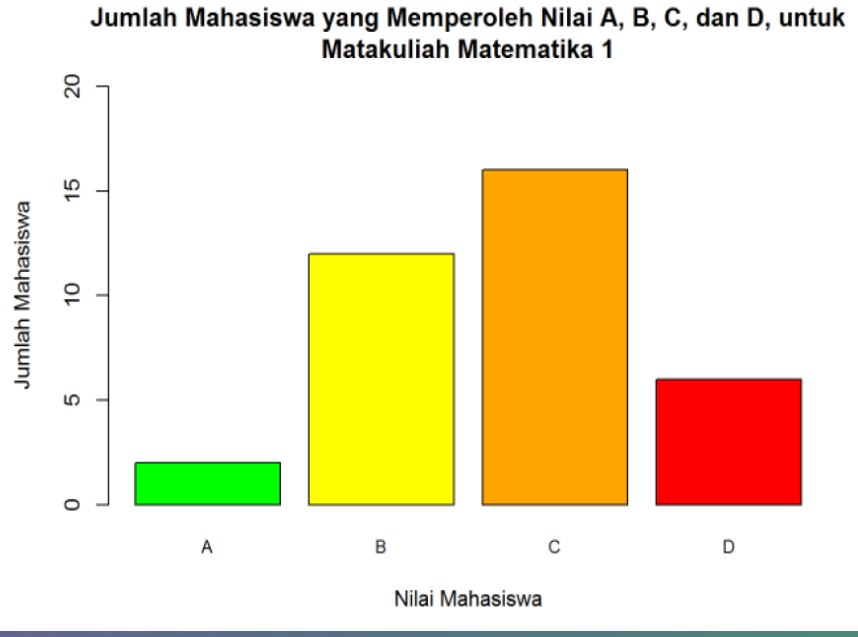


✓Packages for Data Visualization?

- *Ggplot2 → plot, data, aesthetic, layer*

```
library(ggplot2)

# Menggunakan package openxlsx
library(openxlsx)

# Membaca file mahasiswa.xlsx
mahasiswa <- read.xlsx("https://academy.dqlab.id/dataset/mahasiswa.xlsx",sheet = "Sheet 1")

# Membuat kanvas
gambar <- ggplot(mahasiswa, aes(x=Fakultas, y=JUMLAH, fill=Fakultas))

# Menambahkan objek bar chart, simpan kembali sebagai variable gambar
gambar <- gambar + geom_bar(width=1, stat="identity")

# Menggambar grafik
gambar
```





Open the Link:

https://github.com/ekotwidodo/UTI-Statistika-Dasar

Download File:

Penyajian Data R.Rmd

References

- Douglass A, Lind, William G, Marchal and Samuel A, Wathen. *Statistical Techniques in Business & Economics*. Mc. Graw Hill, NY, 2005.
- Supranto, J, MA. *Statistik-Teori dan Aplikasi*, Jilid I, edisi kedelapan. Erlangga. 2016. Jakarta
- Anderson, T.W, and Sclove, L. Stanley, *The Statistical Analysis of Data*, Second Edition, Houghton Mifflin Company, 1986, USA.
- Johnson R, and Bhattacharyya G. *Statistics, Principles and Methods*. John Wiley & Sons Inc. 1985. Canada.
- Asra A dan Sutomo S. *Pengantar Statistika I*. Rajawali Pers. 2016. Jakarta
- Asra A dan N.B. Parwanto. 2018. *1001 Soal Jawab Statistika Deskriptif*. Jakarta. In Media.
- ...

THANK YOU

Link Materials:
https://github.com/ekotwidodo/UTI-Statistika-Dasar

Requests for Collaboration:

1. Terdapat kumpulan data hasil timbangan berat badan mahasiswa Prodi Pendidikan Matematika Universitas Teknokrat Indonesia sebagai berikut:

41	45	49	51	52	53	55	56	56	57
57	58	59	59	60	61	61	62	63	63
65	65	65	67	67	67	67	69	69	69
69	70	71	71	71	73	73	73	73	73
75	75	77	77	77	77	79	81	83	83
87	89	89	91	91	91	92	93	94	96

a. Sajikan data tersebut dengan jenis penyajian yang sesuai (tabel)
b. Buatlah histogram dan polygon dari tabel tersebut

1. Terdapat Kumpulan indikator, seperti Gini Ratio, Angka Pengangguran, Persentase Kemiskinan, dan Pertumbuhan Ekonomi suatu wilayah dari tahun 2004-2013:

Indikator	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Gini ratio	0,32	0,33	0,33	0,36	0,35	0,37	0,38	0,41	0,41	0,41
Angka pengangguran	9,90	10,26	10,45	9,75	8,46	8,14	7,41	6,80	6,32	5,92
% miskin	16,70	16,00	17,80	16,60	15,40	14,20	13,33	12,49	11,96	11,37
Pertumbuhan ekonomi (%)	5,0	5,7	5,5	6,3	6,0	4,6	6,2	6,5	6,2	5,6

Dengan aplikasi R

a. Buat grafik yang paling tepat untuk informasi di atas!
b. Buat grafik lain (batang misalnya), kemudian bandingkan!