

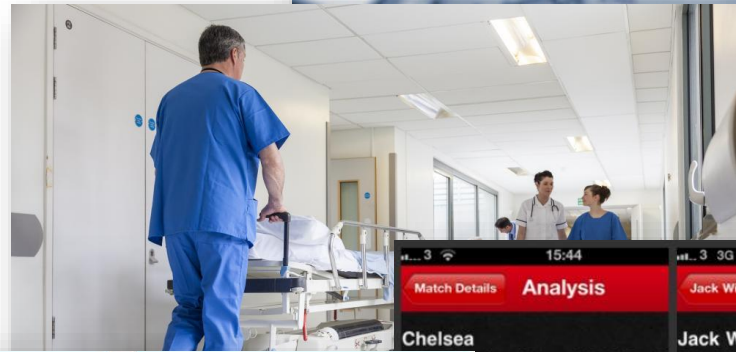
# Data dan Informasi

Informatika Lingkungan

# Manusia Memproduksi Data

Manusia memproduksi beragam data yang **jumlah dan ukurannya sangat besar**

- Astronomi
- Bisnis
- Kedokteran
- Ekonomi
- Olahraga
- Cuaca
- Financial
- ...



# Winning An Unfair Game



## The Saber Revolution

ASSESSING  
THE  
GROWTH  
OF  
ANALYTICS  
IN  
BASEBALL

### SABERMETRICS, SCOUTING AND THE SCIENCE OF BASEBALL

<b>wOBA</b> <small>Weighted on-base average</small>								
<b>Ops+</b> <small>Adjusted on-base plus slugging</small>	<b>WAR</b> <small>Wins Above Replacement</small>							
<b>Lwts</b> <small>Linear weights</small>	<b>Rx</b> <small>Run Expectancy</small>							
<b>Ops</b> <small>On-base plus slugging</small>	<b>H</b> <small>Hits</small>	<b>HR</b> <small>Home Runs</small>	<b>RBI</b> <small>Runs Batted In</small>	<b>Gp</b> <small>Games Played</small>	<b>Sb</b> <small>Stolen Bases</small>	<b>Br+</b> <small>Base Running</small>	<b>Dw%</b> <small>Defensive Win Percentage</small>	<b>Sec</b> <small>Runs Saved</small>
<b>Ave</b> <small>Adjusted Average</small>	<b>Tb</b> <small>Total Bases</small>	<b>Jb</b> <small>Jobs</small>	<b>Li</b> <small>Linear Weights</small>	<b>Ab</b> <small>Adjusted Batting Average</small>	<b>Cs</b> <small>Caught Stealing</small>	<b>Bw+</b> <small>Baseball Win Percentage</small>	<b>Xbb</b> <small>Extra Bases</small>	<b>Rza</b> <small>Runs Above Replacement</small>
<b>Obp</b> <small>On-base Percentage</small>	<b>Rc</b> <small>Runs Created</small>	<b>Zb</b> <small>Zips</small>	<b>Wpa</b> <small>Win Probability Added</small>	<b>Pa</b> <small>Plate Appearances</small>	<b>Sh</b> <small>Strikeouts</small>	<b>K</b> <small>Errors</small>	<b>Tob</b> <small>Total On Base</small>	<b>Lob</b> <small>Left On Base</small>
<b>Slg</b> <small>Slugging Percentage</small>	<b>Rp</b> <small>Runs Produced</small>	<b>Ib</b> <small>Innings</small>	<b>Iso</b> <small>Isolated Power</small>	<b>Ppa</b> <small>Plate Plate Appearances</small>	<b>Sl</b> <small>Strikeouts</small>	<b>Bb</b> <small>Bases on Balls</small>	<b>Ibb</b> <small>Innings on Base</small>	<b>Wpa</b> <small>Win Probability Added</small>
		<b>A</b> <small>ERA</small>	<b>P</b> <small>Pitches</small>	<b>S</b> <small>Speed</small>	<b>F</b> <small>Fatigue</small>	<b>20/80</b> <small>20/80 Ratio</small>	<b>Pt</b> <small>Pitches</small>	<b>Bs</b> <small>Bases on Balls</small>

BRAD PITT

MONEYBALL

JONAH HILL PHILIP SEYMOUR HOFFMAN

BASED ON A TRUE STORY

COMING SOON

# How Do The Oakland A's Win With a Poor Team?

# Pertumbuhan Data

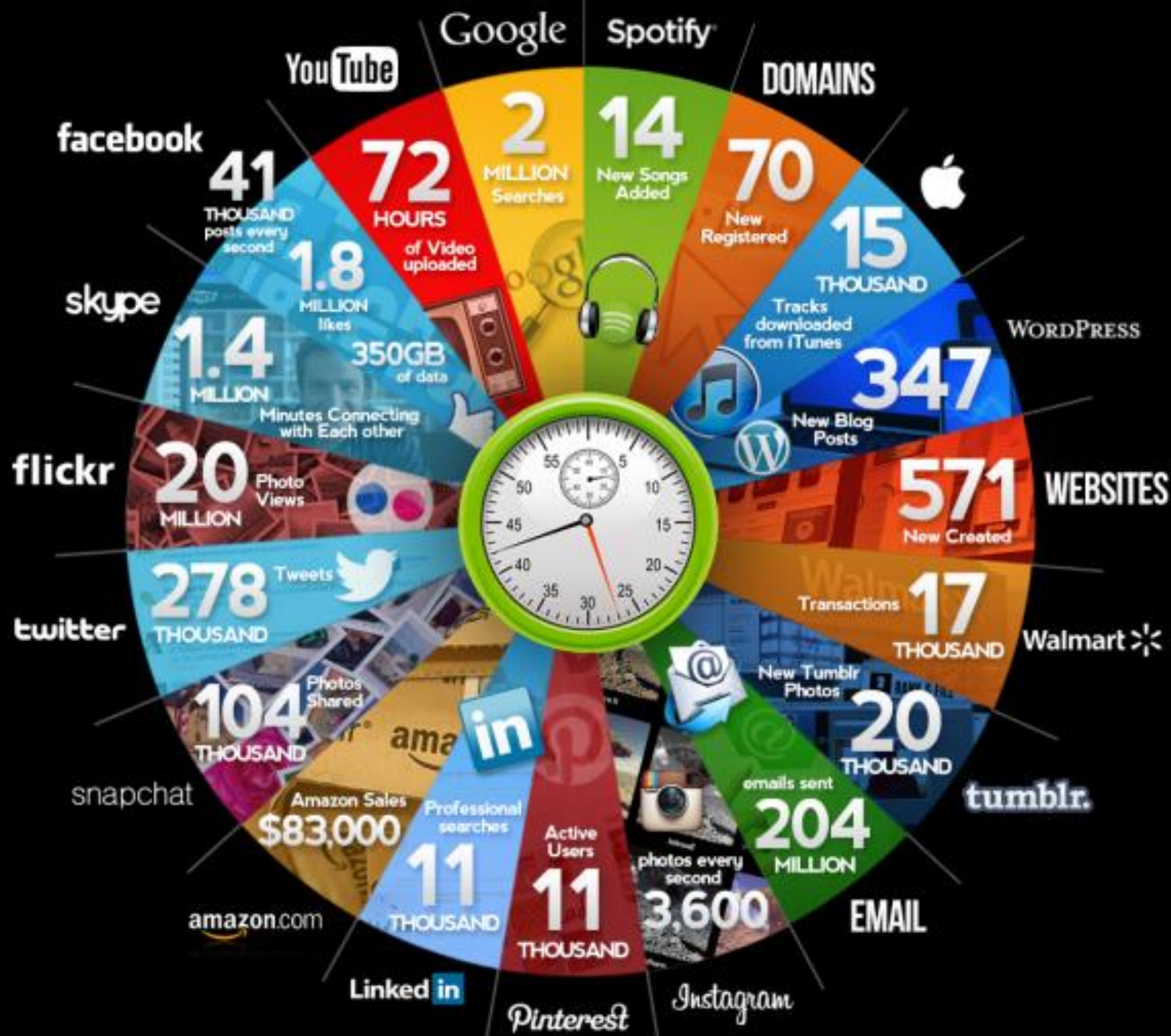
## Astronomi

- **Sloan Digital Sky Survey**
  - New Mexico, 2000
  - **140TB** over 10 years
- **Large Synoptic Survey Telescope**
  - Chile, 2016
  - Will acquire **140TB every five days**

kilobyte (kB)	$10^3$
megabyte (MB)	$10^6$
gigabyte (GB)	$10^9$
terabyte (TB)	$10^{12}$
petabyte (PB)	$10^{15}$
exabyte (EB)	$10^{18}$
zettabyte (ZB)	$10^{21}$
yottabyte (YB)	$10^{24}$

## Biologi dan Kedokteran

- European Bioinformatics Institute (**EBI**)
  - **20PB of data** (genomic data doubles in size each year)
  - A single sequenced human genome can be around **140GB** in size



# Datangnya Tsunami Data

- **Mobile Electronics** market
  - 4.43B mobile phone users in 2015
  - 7B mobile phone subscriptions in 2015
- **Web and Social Networks** generates amount of data
  - Google processes 100 PB per day, 3 million servers
  - Facebook has 300 PB of user data per day
  - Youtube has 1000PB video storage
  - 235 TBs data collected by the US Library of Congress
  - 15 out of 17 sectors in the US have more data stored per company than the US Library of Congress

kilobyte (kB)	$10^3$
megabyte (MB)	$10^6$
gigabyte (GB)	$10^9$
terabyte (TB)	$10^{12}$
petabyte (PB)	$10^{15}$
exabyte (EB)	$10^{18}$
zettabyte (ZB)	$10^{21}$
yottabyte (YB)	$10^{24}$

Kebanjiran Data tapi Miskin Pengetahuan

We are **drowning in data**, but  
**starving for knowledge!**

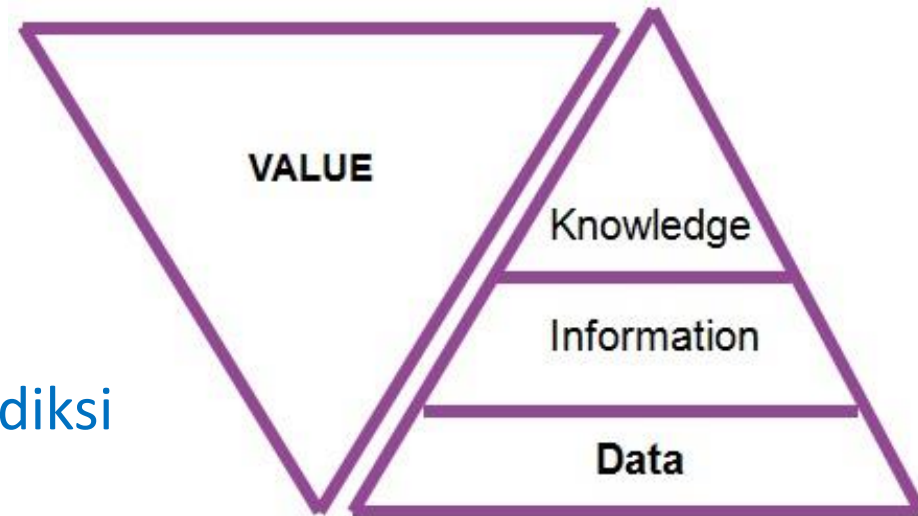
*(John Naisbitt, Megatrends, 1988)*



# Mengubah Data Menjadi Pengetahuan

- Data harus kita olah menjadi **pengetahuan** supaya bisa **bermanfaat** bagi manusia

- Dengan **pengetahuan** tersebut, manusia dapat:
  - Melakukan **estimasi** dan **prediksi** apa yang terjadi di depan
  - Melakukan analisis tentang **asosiasi**, **korelasi** dan **pengelompokan** antar data dan atribut
  - Membantu **pengambilan keputusan** dan **pembuatan kebijakan**





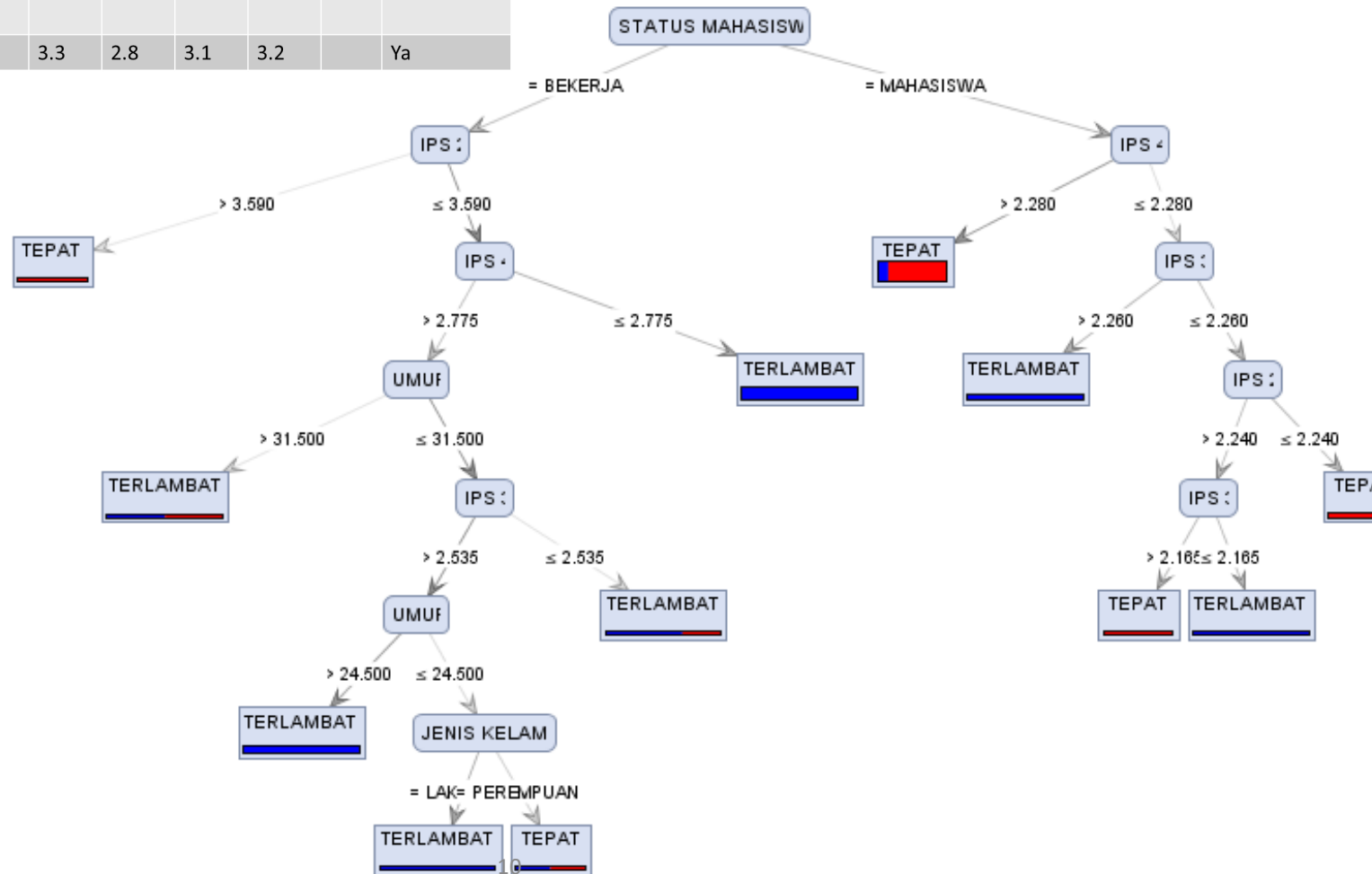
# Contoh Data di Kampus

- **Puluhan ribu data** mahasiswa di kampus yang diambil dari sistem informasi akademik
- Apakah **pernah kita ubah menjadi pengetahuan** yang lebih bermanfaat? TIDAK!
- Seperti apa pengetahuan itu? **Rumus, Pola, Aturan**

NIM	Gender	Nilai UN	Asal Sekolah	IPS1	IPS2	IPS3	IPS 4	...	Lulus Tepat Waktu
10001	L	28	SMAN 2	3.3	3.6	2.89	2.9		Ya
10002	P	27	SMA DK	4.0	3.2	3.8	3.7		Tidak
10003	P	24	SMAN 1	2.7	3.4	4.0	3.5		Tidak
10004	L	26.4	SMAN 3	3.2	2.7	3.6	3.4		Ya
...									
...									
11000	L	23.4	SMAN 5	3.3	2.8	3.1	3.2		Ya

# Prediksi Kelulusan Mahasiswa

NIM	Gender	Nilai UN	Asal Sekolah	IPS1	IPS2	IPS3	IPS 4	...	Lulus Tepat Waktu
10001	L	28	SMAN 2	3.3	3.6	2.89	2.9		Ya
10002	P	27	SMA DK	4.0	3.2	3.8	3.7		Tidak
10003	P	24	SMAN 1	2.7	3.4	4.0	3.5		Tidak
10004	L	26.4	SMAN 3	3.2	2.7	3.6	3.4		Ya
...									
...									
11000	L	23.4	SMAN 5	3.3	2.8	3.1	3.2		Ya



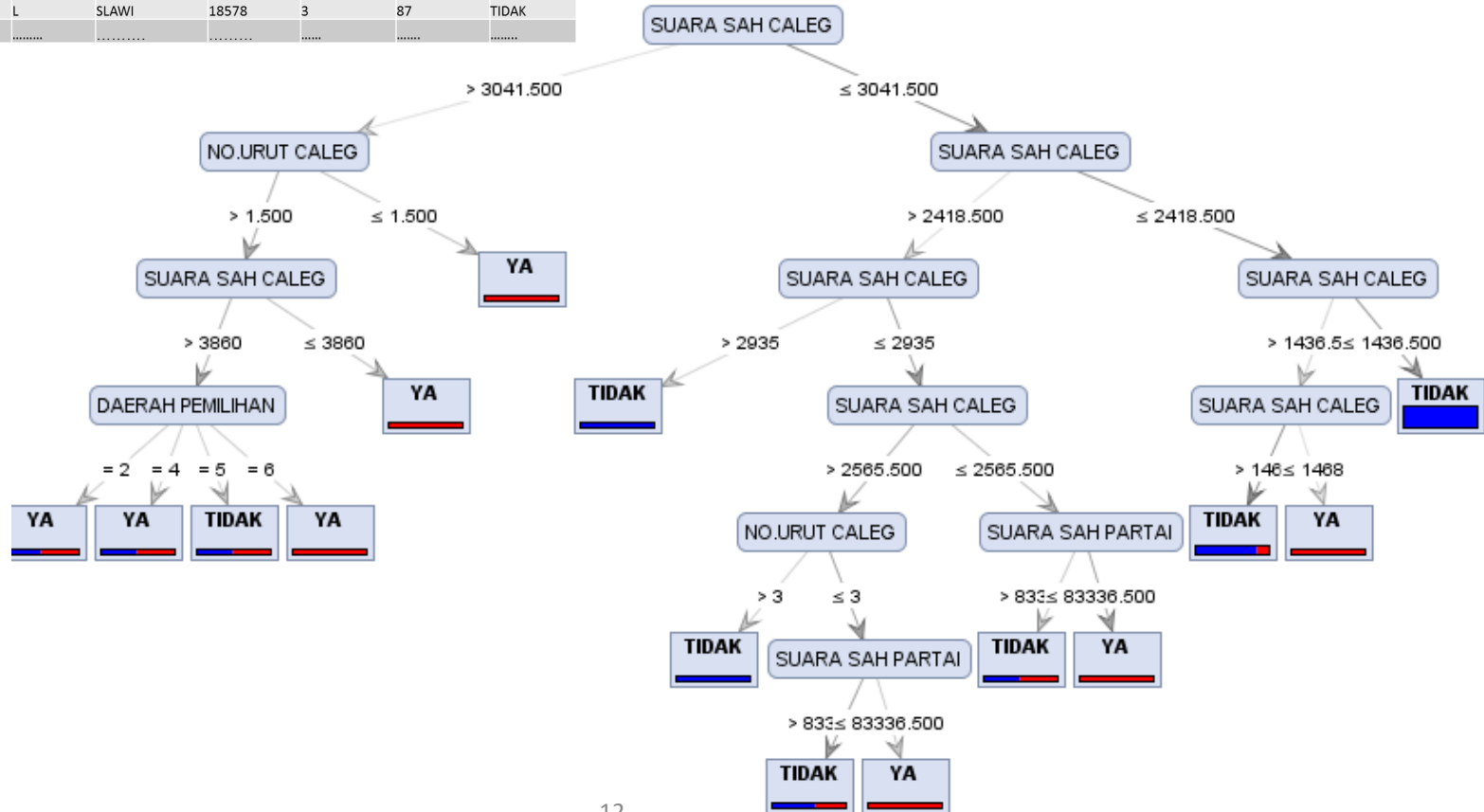
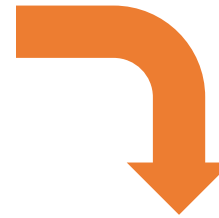
# Contoh Data di Komisi Pemilihan Umum

- Puluhan ribu data calon anggota legislatif di KPU
- Apakah pernah kita ubah menjadi pengetahuan yang lebih bermanfaat? TIDAK!

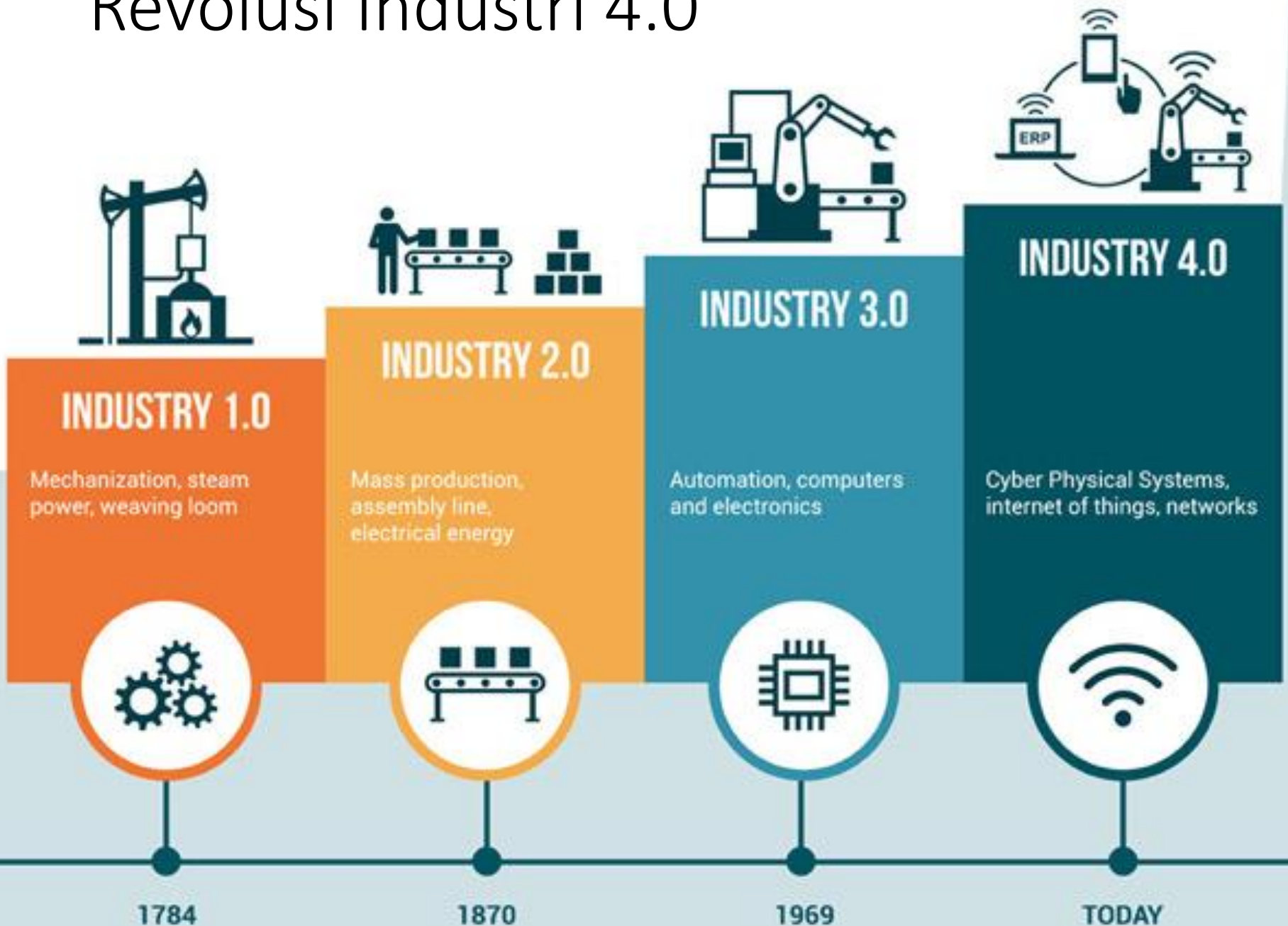
NAMA PARTAI POLITIK	NAMA CALON LEGESLATIF	JENIS KELAMIN	KECAMATAN	SUARA SAH PARTAI	DAERAH PEMILIHAN	SUARA SAH CALEG	TERPILIH ATAU TIDAK
HANURA	TOTO SUKISNO,BSc	L	LEBAKSIU	18578	1	594	TIDAK
HANURA	EDI PURYANTO,SH	L	SLAWI	18578	1	943	TIDAK
PKB	ELI RETNOWATI,SH	P	SLAWI	18578	1	1730	TIDAK
PKB	SAHYUDIN	L	DUKUHWARU	18578	1	2508	YA
GOLKAR	H.FAJAR SIGIT KUSUMAJAYA,SH	L	SLAWI	18578	2	923	TIDAK
GOLKAR	SUMIRAH	P	TARUB	18578	2	308	TIDAK
GOLKAR	DARYOTO	L	TARUB	18578	2	54	TIDAK
PKS	KHAPIP APRONI,S.Pdi	L	BOJONG	18578	3	1682	TIDAK
PKS	ENDANG SUCI RAHAYU	P	JATINEGARA	18578	3	918	TIDAK
PDI-P	KH.CHAFIDZ ISA MUFTI ,LC	L	SLAWI	18578	3	87	TIDAK
.....	.....	.....	.....	.....	.....	.....	.....

# Prediksi Calon Legislatif DKI Jakarta

NAMA PARTAI POLITIK	NAMA CALON LEGESLATIF	JENIS KELAMIN	KECAMATAN	SUARA SAH PARTAI	DAERAH PEMILIHAN	SUARA SAH CALEG	TERPILIH ATAU TIDAK
HANURA	TOTO SUKISNO,BSc	L	LEBAKSIU	18578	1	594	TIDAK
HANURA	EDI PURYANTO,SH	L	SLAWI	18578	1	943	TIDAK
PKB	ELI RETNOWATI,SH	P	SLAWI	18578	1	1730	TIDAK
PKB	SAHYUDIN	L	DUKUHWARU	18578	1	2508	YA
GOLKAR	H.FAJAR SIGIT KUSUMAJAYA,SH	L	SLAWI	18578	2	923	TIDAK
GOLKAR	SUMIRAH	P	TARUB	18578	2	308	TIDAK
GOLKAR	DARYOTO	L	TARUB	18578	2	54	TIDAK
PKS	KHAPIP APRONI,S.Pdi	L	BOJONG	18578	3	1682	TIDAK
PKS	ENDANG SUCI RAHAYU	P	JATINEGARA	18578	3	918	TIDAK
PDI-P	KH.CHAFIDZ ISA MUFTI ,LC	L	SLAWI	18578	3	87	TIDAK
.....	.....	.....	.....	.....	.....	.....	.....

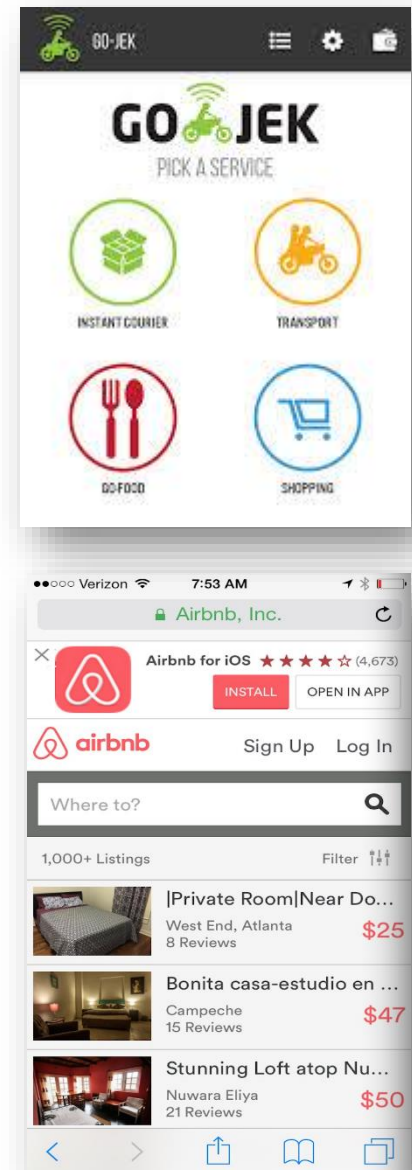


# Revolusi Industri 4.0



# Perusahaan Pengolah Pengetahuan

- **Uber** - the world's largest taxi company, owns no vehicles
- **Google** - world's largest media/advertising company, creates no content
- **Alibaba** - the most valuable retailer, has no inventory
- **Airbnb** - the world's largest accommodation provider, owns no real estate
- **Gojek** - perusahaan angkutan umum, tanpa memiliki kendaraan



# Data - Informasi – Pengetahuan

NIP	TGL	DATANG	PULANG
1103	02/12/2004	07:20	15:40
1142	02/12/2004	07:45	15:33
1156	02/12/2004	07:51	16:00
1173	02/12/2004	08:00	15:15
1180	02/12/2004	07:01	16:31
1183	02/12/2004	07:49	17:00

Data Kehadiran Pegawai

# Data - Informasi – Pengetahuan

NIP	Masuk	Alpa	Cuti	Sakit	Telat
1103	22				
1142	18	2		2	
1156	10	1	11		
1173	12	5			5
1180	10			12	

Informasi Akumulasi Bulanan Kehadiran Pegawai



# Data - Informasi – Pengetahuan

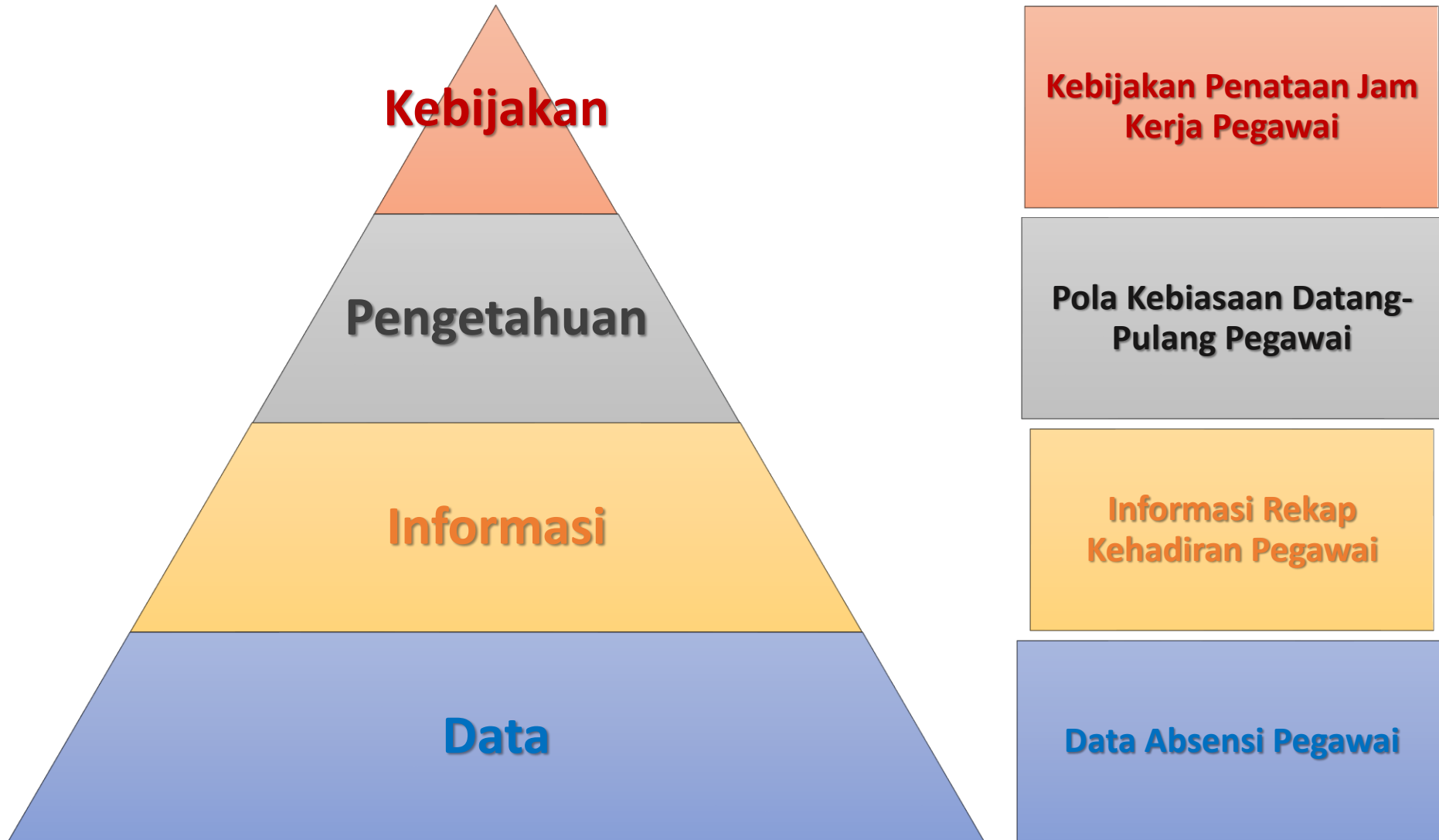
	Senin	Selasa	Rabu	Kamis	Jumat
Terlambat	7	0	1	0	5
Pulang Cepat	0	1	1	1	8
Izin	3	0	0	1	4
Alpa	1	0	2	0	2

Pola Kebiasaan Kehadiran Mingguan Pegawai

# Data - Informasi – Pengetahuan - Kebijakan

- Kebijakan **penataan jam kerja karyawan** khusus untuk hari senin dan jumat
- Peraturan jam kerja:
  - Hari **Senin** dimulai jam 10:00
  - Hari **Jumat** diakhiri jam 14:00
  - Sisa jam kerja **dikompensasi ke hari lain**

# Mengubah Pengetahuan Menjadi Kebijakan



# Data Mining Tasks and Roles

