

COVID-19 ANALYSIS AND DATA VISUALIZATION

by Marcel Siswanto – 07 August 2020

Table of Contents

TABLE OF CONTENTS.....	1
INTRODUCTION	3
LATAR BELAKANG	3
TUGAS DAN LANGKAH	3
ANALISIS DAN VISUALISASI DATA DI INDONESIA	5
MENGAKSES API COVID.GO.ID.....	5
MENGEVALUASI RESPON.....	5
STATUS CODE	5
HEADERS API.....	6
MENGEKSTRAK ISI RESPON.....	7
ANALISIS DATA.....	7
MEMPEROLEH INFORMASI YANG LEBIH LENGKAP	9
PERBAIKAN DATA FRAME KEY_AS_STRING SEBAGAI FORMAT TANGGAL	10
MENGGABUNGKAN BEBERAPA DATA FRAME MENJADI SATU	11
VISUALISASI DATA DENGAN GRAFIK.....	12
POLA DAN DINAMIKA	15
GRAFIK KOMPARASI DATA.....	16
GRAFIK PERKEMBANGAN KASUS PEKANAN	18
SEVENTH-DAY AMPLIFICATION FACTOR.....	20
KESIMPULAN	24
ANALISIS DAN VISUALISASI DATA DI DKI JAKARTA	26
INTRODUCTION	26
MENGAKSES API COVID.GO.ID.....	26
MENGEKSTRAK ISI RESPON.....	26
ANALISIS DATA.....	26
MEMPEROLEH INFORMASI YANG LEBIH LENGKAP	28
MENJINAKKAN DATA	28
VISUALISASI DATA DENGAN GRAFIK.....	39
POLA DAN DINAMIKA	32
GRAFIK KOMPARASI DATA.....	33
GRAFIK PERKEMBANGAN KASUS PEKANAN	34
SEVENTH-DAY AMPLIFICATION FACTOR.....	36
KESIMPULAN	38
ANALISIS DAN VISUALISASI DATA DI JAWA BARAT	41
INTRODUCTION	41
MENGAKSES API COVID.GO.ID.....	41
MENGEKSTRAK ISI RESPON.....	41
ANALISIS DATA.....	41
MEMPEROLEH INFORMASI YANG LEBIH LENGKAP	43

MENJINAKKAN DATA	43
VISUALISASI DATA DENGAN GRAFIK	44
POLA DAN DINAMIKA	47
GRAFIK KOMPARASI DATA.....	48
GRAFIK PERKEMBANGAN KASUS PEKANAN	49
SEVENTH-DAY AMPLIFICATION FACTOR.....	51
KESIMPULAN	53
KOMPARASI PROVINSI DKI JAKARTA DAN JAWA BARAT.....	56
INTRODUCTION	56
POLA DAN DINAMIKA	58
GRAFIK KOMPARASI DATA.....	59
KESIMPULAN	65
KESIMPULAN	66
REFERENCES.....	69

I. Introduction

Latar Belakang

COVID-19 merupakan penyakit yang saat ini telah menjadi pandemi secara global. Kondisi menjadi semakin mengkhawatirkan karena hingga detik ini masih belum ditemukan vaksin yang efektif untuk virus penyebab COVID-19. Pemerintah di berbagai negara umumnya dengan sigap membentuk gugus tugas (*task force unit*) untuk menangani penyebaran COVID-19 di masyarakat, termasuk pemerintah di Indonesia.

Salah satu bentuk aksi yang dilakukan oleh pemerintah adalah dengan mengumpulkan dan menyediakan data pertumbuhan kasus COVID-19 kepada publik. Data pertumbuhan kasus tersebut tidak jarang juga dilengkapi dengan dasbor dan grafik visualisasi pendukung dengan harapan masyarakat dapat memahami informasi dengan lebih mudah. Sebagai contoh adalah portal covid19.go.id besutan Gugus Tugas Penanganan COVID-19 Nasional dan portal PIKOBAR milik pemerintah Provinsi Jawa Barat. Serta banyak portal data COVID-19 lainnya yang disediakan oleh masing-masing pemerintah daerah.

Tugas dan Langkah

Dalam proyek ini, kita akan diminta untuk melakukan eksplorasi dan analisis mengenai COVID-19 di Indonesia. Data yang kita gunakan diambil langsung dari API (*Application Programming Interface*) yang tersedia di covid19.go.id. Sehingga dalam proyek ini kita akan mempelajari teknik mengambil data dari API, teknik mempersiapkan data, serta analisis dan visualisasi data.

Fokus kita pada *project* kali ini adalah data di Indonesia (nasional), DKI Jakarta, dan Jawa Barat. Kita akan melakukan analisis terhadap asumsi dan visualisasi data yang dihasilkan selama mengerjakan *project*.

Tugas dan Langkah yang dilakukan:

1. *Data preparation*
2. Membuat grafik kasus positif harian
3. Membuat grafik kasus sembuh harian
4. Membuat grafik kasus meninggal harian
5. Komparasi akumulasi kasus aktif, meninggal, dan sembuh harian
6. Membuat grafik kasus baru pekanan
7. *Seventh-Day Amplification Factor*
8. Membuat kesimpulan

Kita juga akan melakukan komparasi terhadap timbulnya kasus pasien aktif, sembuh, meninggal, dan A7 di DKI Jakarta dan Jawa Barat.

II. Analisis dan Visualisasi Data di Indonesia

Mengakses API covid.go.id

Rekapitulasi data COVID-19 Indonesia tersedia dalam API publik yang beralamat di <https://data.covid19.go.id/public/api/update.json>.

Salah satu cara untuk mengakses API adalah dengan menggunakan fungsi `GET()` dari paket `httr`. Kita akan menyimpan hasil fungsi tersebut dalam obyek bernama `resp`.

```
#Read API
library(httr)
resp <- GET ("https://data.covid19.go.id/public/api/update.json")
```

Mengevaluasi Respon

Saat kita menjalankan fungsi `GET()`, pada dasarnya hal yang terjadi adalah kita membuat sebuah permintaan kepada server penyedia API. Permintaan kita tersebut selanjutnya diproses dan dijawab oleh server sebagai sebuah respon. Obyek `resp` yang telah kita buat memiliki informasi respon oleh server.

Ada tiga informasi utama dalam sebuah respon API, yaitu:

1. **Status** memiliki informasi apakah permintaan kita berhasil atau tidak dan dinyatakan dalam *status code*.
2. **headers** umumnya mengandung informasi metadata.
3. **body** berisikan konten atas permintaan yang telah dibuat.

Status Code

Ada beberapa jenis *status code* yang dikelompokkan menjadi beberapa kategori, antara lain:

- **1xx (informational)** : mengkomunikasikan transfer informasi *protocol-level*.
- **2xx (success)** : permintaan klien berhasil diterima, dipahami, dan dimengerti.
- **3xx (redirection)** : tindakan lebih lanjut perlu dilakukan oleh agen pengguna untuk memenuhi permintaan tersebut.
- **4xx (client error)** : kemungkinan kesalahan *client* dalam permintaan yang mencegah server untuk memprosesnya.
- **5xx (server error)** : server mengalami galat/*error internal* saat mencoba untuk memproses permintaan klien.

Kita dapat menggunakan fungsi `status_code()` untuk mengetahui status atas permintaan kita melalui API.

```
#Cek Status Code 1
status_code(resp)
```

```
200
```

Selain menggunakan fungsi `status_code()`, kita juga dapat mengetahui status permintaan dengan cara mengakses elemen dari `resp`. `Status code` tersebut tersedia dengan nama `status_code` dan dapat diakses dengan menggunakan operator `$` pada obyek `resp`.

```
#Cek Status Code 2
resp$status_code
identical(resp$status_code, status_code(resp))
```

```
200
```

```
TRUE
```

Headers API

Pada praktik sebelumnya, status permintaan kita melalui API sukses dipenuhi. Sekarang kita jalankan fungsi `headers()` pada `resp` untuk mengetahui metadata apa saja yang tersimpan. Apakah isi dari elemen `content-type`? Kapan terakhir kali informasi diperbaharui?

```
#Data apa saja yang tersimpan?
headers(resp)
```

```
$server
```

```
'nginx'
```

```
$date
```

```
'Fri, 07 Aug 2020 11:25:25 GMT'
```

```
$`content-type`
```

```
'application/json'
```

```
$`last-modified`
```

```
'Fri, 07 Aug 2020 11:21:36 GMT'
```

```
$`transfer-encoding`
```

```
'chunked'
```

```
$connection
```

```
'keep-alive'
```

```
$vary
```

```
'Accept-Encoding'
```

```
$etag
```

```
'W/"5f2d3940-ebe2"'
```

```
$`content-encoding`
```

```
'gzip'
```

Mengekstrak Isi Respon

Respon API dengan *status code* 200 menyatakan bahwa permintaan kita berhasil dipenuhi dan konten yang diminta tersedia untuk diekstrak. Selain itu, kita juga telah mengetahui lewat fungsi `headers()` bahwa konten yang diminta tersedia dalam bentuk *application/json*, yaitu berkas JSON.

Selanjutnya kita dapat mengekstrak konten tersebut dengan menggunakan fungsi `content()`. Fungsi `content()` tersebut secara cerdas akan menyesuaikan *output* sesuai dengan jenis berkas yang tersedia, dalam hal ini adalah berkas JSON. Kita akan menjalankan fungsi tersebut pada obyek `resp` dan menambahkan argumen `as = "parsed"` dan `simplifyVector = TRUE`. Hasilnya akan disimpan sebagai `cov_id_raw`.

```
#Ekstrak konten JSON
cov_id_raw <- content(resp, as = "parsed", simplifyVector = TRUE)
```

Dengan menggunakan fungsi `length()` dan `names()` kita dapat mengamati ada berapa komponen serta apa saja nama komponen dalam obyek `cov_id_raw` tersebut. Kemudian kita ekstrak komponen ke-2 dan simpan dengan nama `cov_id_update`.

```
#Ada berapa komponen dan apa saja komponennnya
length(cov_id_raw)
names(cov_id_raw)
cov_id_update <- cov_id_raw$update
```

```
2
```

```
'data' 'update'
```

Analisis Data

Sekarang kita fokus pada obyek `cov_id_update` untuk menjawab pertanyaan-pertanyaan berikut:

- Kapan tanggal pembaruan data penambahan kasus?
- Berapa jumlah penambahan kasus sembuh?
- Berapa jumlah penambahan kasus meninggal?
- Berapa jumlah total kasus positif hingga saat ini?
- Berapa jumlah total kasus meninggal hingga saat ini?

```
#Analisa data
lapply(cov_id_update, names)

cov_id_update$penambahan$tanggal
cov_id_update$penambahan$jumlah_sembuh
cov_id_update$penambahan$jumlah_meninggal
cov_id_update$total$jumlah_positif
cov_id_update$total$jumlah_meninggal
```

\$penambahan

```
'jumlah_positif' · 'jumlah_meninggal' · 'jumlah_sembuh' · 'jumlah_dirawat' · 'tanggal' ·
'created'
```

\$sharian

```
'key_as_string' · 'key' · 'doc_count' · 'jumlah_meninggal' · 'jumlah_sembuh' ·
'jumlah_positif' · 'jumlah_dirawat' · 'jumlah_positif_kum' · 'jumlah_sembuh_kum' ·
'jumlah_meninggal_kum' · 'jumlah_dirawat_kum'
```

\$total

```
'jumlah_positif' · 'jumlah_dirawat' · 'jumlah_sembuh' · 'jumlah_meninggal'
```

```
'2020-08-07'
```

```
1912
```

```
72
```

```
121226
```

```
5593
```

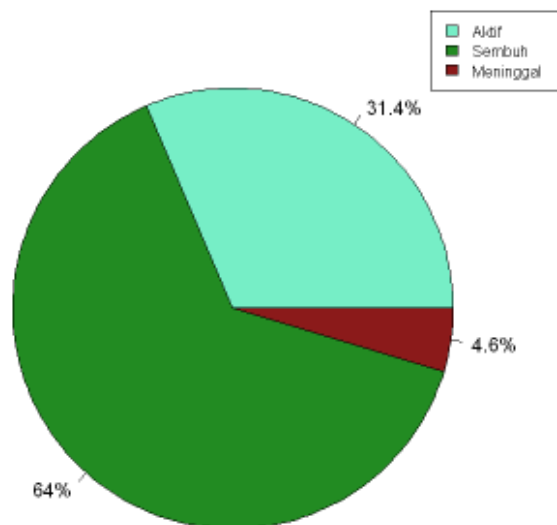
Kita juga dapat membuat visualisasi data untuk persentase jumlah total pasien aktif, sembuh, dan meninggal di Indonesia agar mempermudah pemahaman data dengan membuat *pie chart*.

```
#Membuat visualisasi data pasien aktif, sembuh, dan meninggal
x <- c(cov_id_update$total$jumlah_dirawat,
      cov_id_update$total$jumlah_sembuh,
      cov_id_update$total$jumlah_meninggal)
labels <- c("Aktif", "Sembuh", "Meninggal")

piepercent<- round(100*x/sum(x), 1)
piepercent <- paste(piepercent, "%", sep="")

# Plot the chart.
pie(x, labels = piepercent,
    main = "Persentase Total Pasien Aktif, Sembuh, dan Meninggal di
Indonesia per 7 Agustus 2020",
    col = c("aquamarine2", "forestgreen", "firebrick4"))
legend("topright", c("Aktif", "Sembuh", "Meninggal"),
    cex = 0.8,
    fill = c("aquamarine2", "forestgreen", "firebrick4"))
```


Persentase Total Pasien Aktif, Sembuh, dan Meninggal di Indonesia per 7 Agustus 2020



Memperoleh Informasi yang Lebih Lengkap

Informasi umum mengenai COVID-19 di Indonesia telah kita dapatkan. Namun informasi akan lebih lengkap jika kita memiliki data perkembangan COVID-19 dari waktu ke waktu.

Data historis perkembangan COVID-19 tersebut tersimpan di `update` dan `harian`. Kita akan mengekstrak data tersebut dari `cov_id_raw` dan menyimpan hasilnya sebagai obyek bernama `cov_id`. Lalu kita akan mengamati struktur `cov_id` menggunakan fungsi `str()` dan `head()`.

```
#Kita ambil update harian dan dimasukkan ke variable cov_id
cov_id <- cov_id_raw$update$harian
str(cov_id)
```

```
'data.frame': 159 obs. of 11 variables:
 $ key_as_string      : chr  "2020-03-02T00:00:00.000Z" "2020-03-03T00:00:00.000Z" "2020-03-04T00:00:00.000Z" "2020-03-05T00:00:00.000Z" ...
 $ key                : num  1.58e+12 1.58e+12 1.58e+12 1.58e+12 1.58e+12 ...
 $ doc_count          : int   1 1 1 1 1 1 1 1 1 1 ...
 $ jumlah_meninggal   :'data.frame': 159 obs. of 1 variable:
 ..$ value: int   0 0 0 0 0 0 0 0 1 ...
 $ jumlah_sembuh      :'data.frame': 159 obs. of 1 variable:
 ..$ value: int   0 0 0 0 0 0 0 0 2 0 ...
 $ jumlah_positif      :'data.frame': 159 obs. of 1 variable:
 ..$ value: int   2 0 0 0 2 0 2 13 8 7 ...
 $ jumlah_dirawat      :'data.frame': 159 obs. of 1 variable:
 ..$ value: int   2 0 0 0 2 0 2 13 6 6 ...
 $ jumlah_positif_kum  :'data.frame': 159 obs. of 1 variable:
 ..$ value: int   2 2 2 2 4 4 6 19 27 34 ...
 $ jumlah_sembuh_kum   :'data.frame': 159 obs. of 1 variable:
 ..$ value: int   0 0 0 0 0 0 0 0 2 2 ...
```

```
$ jumlah_meninggal_kum:'data.frame': 159 obs. of 1 variable:
..$ value: int 0 0 0 0 0 0 0 0 1 ...
$ jumlah_dirawat_kum : 'data.frame': 159 obs. of 1 variable:
..$ value: int 2 2 2 2 4 4 6 19 25 31 ...
```

Tampak *data frame* yang ditampilkan memiliki turunan berupa beberapa *data frame* dengan masing-masing terdiri dari satu *variable*. Hal ini akan mempersulit dalam proses pembacaan data. Adapun juga format tanggal yang tersimpan pada *variable* **key_as_string**, kita harus ubah menjadi format tanggal yang sesuai dan mengubah *class* menjadi *date*. Pada praktik berikutnya, kita hanya mengambil kolom **key_as_string**, **jumlah_meninggal**, **jumlah_sembuh**, **jumlah_positif**, dan **jumlah_dirawat**. Hal tersebut guna mempermudah kita dalam menganalisis data sesuai tujuan di awal.

Perbaikan Data Frame **key_as_string** Sebagai Format Tanggal

Format tanggal yang salah atau janggal akan membingungkan pada proses analisis dan visualisasi data. Oleh karena itu, kita akan melakukan perbaikan pada *data frame* **key_as_string** melalui tahapan berikut:

1. Memasukkan **key_as_string** ke dalam *variable* **Tanggal**,
2. Menghapus "T00:00:00.000Z" pada seluruh *cell* di kolom **Tanggal**. Jika tidak dihapus, maka dapat mengganggu proses pembacaan data,
3. Mengecek *class* dari *variable* **Tanggal**,
4. Mengubah *variable* **Tanggal** yang berupa *list* menjadi sebuah *data frame*,
5. Mengubah *class* tipe *character* menjadi *date*,

Kita akan menggunakan fungsi **lapply()** dan **gsub** untuk menghapus "T00:00:00.000Z" pada kolom **Tanggal**

```
#Masukkan update harian key_as_string sebagai Tanggal
Tanggal <- cov_id_raw$update$harian$key_as_string
#Menghapus "T00:00:00.000Z" pada seluruh cell di kolom Tanggal
Tanggal <- lapply(Tanggal, gsub, pattern='T00:00:00.000Z', replacement='')
```

Tampak *variable* **Tanggal** merupakan *list* dan harus diubah menjadi *data frame*. Jika masih berupa *list*, maka tidak dapat terbaca pada proses-proses berikutnya yang memerlukan *variable* berupa *data frame*.

```
#Mengecek class dari Tanggal
class(Tanggal)
```

```
'list'
```

Tanggal yang berupa *data frame* disimpan ke dalam *variable* baru dengan nama **Tanggal_New**. Sekarang kita sudah mengubah **Tanggal** menjadi sebuah *data frame* dengan *class* berupa *date*. Kita akan menggunakan *variable* **Tanggal_New** yang sudah diperbaiki ini dalam proses pengolahan data selanjutnya.

```
#Mengubah dari list menjadi data.frame
Tanggal_New <- data.frame(matrix(unlist(Tanggal), ncol = max(lengths(Tanggal)),
byrow = TRUE))
names(Tanggal_New)[1] <- "Tanggal"

#Mengubah character menjadi format date
Tanggal_New$Tanggal <- as.Date(Tanggal_New$Tanggal)

str(Tanggal_New)

'data.frame': 159 obs. of 1 variable:
 $ Tanggal: Date, format: "2020-03-02" "2020-03-03" ...
```

Menggabungkan Beberapa Data Frame Menjadi Satu

Terlihat pada struktur *variable* **cov_indo**, *data frame*nya memiliki beberapa turunan berupa *data frame* juga. Oleh karena itu, kita akan menggabungkan beberapa turunan *data frame* tersebut menjadi satu *data frame*. Kita hanya mengambil beberapa kolom atau *variable* yang penting untuk digunakan dalam proses analisis dan visualisasi data.

```
#Memasukkan Tanggal_New ke variable cov_id_tanggal
cov_id_tanggal <- Tanggal_New
#Memasukkan jumlah_meninggal ke variable cov_id_meninggal
cov_id_meninggal <- cov_id_raw$update$harian$jumlah_meninggal
names(cov_id_meninggal)[names(cov_id_meninggal) == "value"] <-
"Jumlah_Meninggal"
#Memasukkan jumlah_sembuh ke variable cov_id_sembuh
cov_id_sembuh <- cov_id_raw$update$harian$jumlah_sembuh
names(cov_id_sembuh)[names(cov_id_sembuh) == "value"] <- "Jumlah_Sembuh"
#Memasukkan jumlah_positif ke variable cov_id_positif
cov_id_positif <- cov_id_raw$update$harian$jumlah_positif
names(cov_id_positif)[names(cov_id_positif) == "value"] <- "Jumlah_Positif"
#Memasukkan jumlah_dirawat ke variable cov_id_aktif
cov_id_aktif <- cov_id_raw$update$harian$jumlah_dirawat
names(cov_id_aktif)[names(cov_id_aktif) == "value"] <- "Jumlah_Aktif"
```

Kini data yang akan diolah pada proses selanjutnya sudah kita perbaiki sehingga mudah untuk diproses. Data yang sudah diperbaiki ini telah dimasukkan ke dalam *variable* bernama **cov_id_new**.

```
#Menggabungkan variable di atas menjadi satu data frame
cov_id_new <- cbind(cov_id_tanggal, cov_id_positif, cov_id_sembuh,
                    cov_id_meninggal, cov_id_aktif)

str(cov_id_new)

#Mengambil bagian teratas dari cov_id_new
head(cov_id_new)

'data.frame':  159 obs. of  5 variables:
 $ Tanggal      : Date, format: "2020-03-02" "2020-03-03" ...
 $ Jumlah_Positif : int  2 0 0 0 2 0 2 13 8 7 ...
 $ Jumlah_Sembuh : int  0 0 0 0 0 0 0 0 2 0 ...
 $ Jumlah_Meninggal: int  0 0 0 0 0 0 0 0 0 1 ...
 $ Jumlah_Aktif  : int  2 0 0 0 2 0 2 13 6 6 ...
```

A data.frame: 6 × 5

	Tanggal	Jumlah_Positif	Jumlah_Sembuh	Jumlah_Meninggal	Jumlah_Aktif
	<date>	<int>	<int>	<int>	<int>
1	2020-03-02	2	0	0	2
2	2020-03-03	0	0	0	0
3	2020-03-04	0	0	0	0
4	2020-03-05	0	0	0	0
5	2020-03-06	2	0	0	2
6	2020-03-07	0	0	0	0

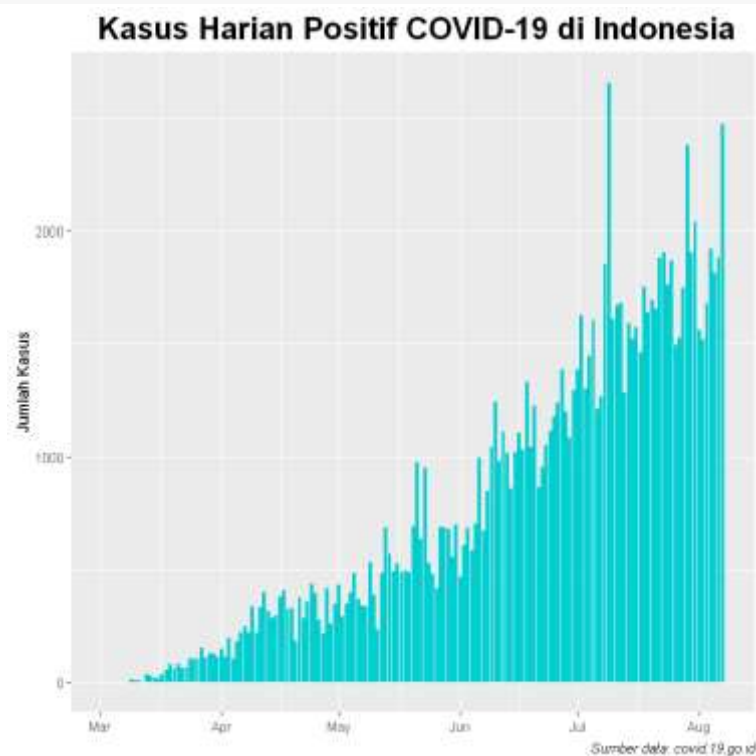
Visualisasi Data Dengan Grafik

sekarang saatnya kita mengekspresikan data tersebut dalam bentuk lain yang harapannya lebih mudah dicerna: grafik. Kita akan memulai merancang visualisasi yang memiliki estetika dengan menggunakan paket **ggplot2** dan paket **hrbrthemes**.

Komponen utama untuk membuat visualisasi antara lain adalah tabel data, kolom data, serta bentuk geometri untuk mempresentasikan data. Sebagai contoh untuk membuat *scatter-plot* yang diperlukan adalah bentuk geometri titik (**geom_col()**), *line-chart* memerlukan geometri garis (**geom_line()**), sedangkan *bar-chart* memerlukan bentuk geometri batang atau kolom (**geom_bar()** atau **geom_col()**).

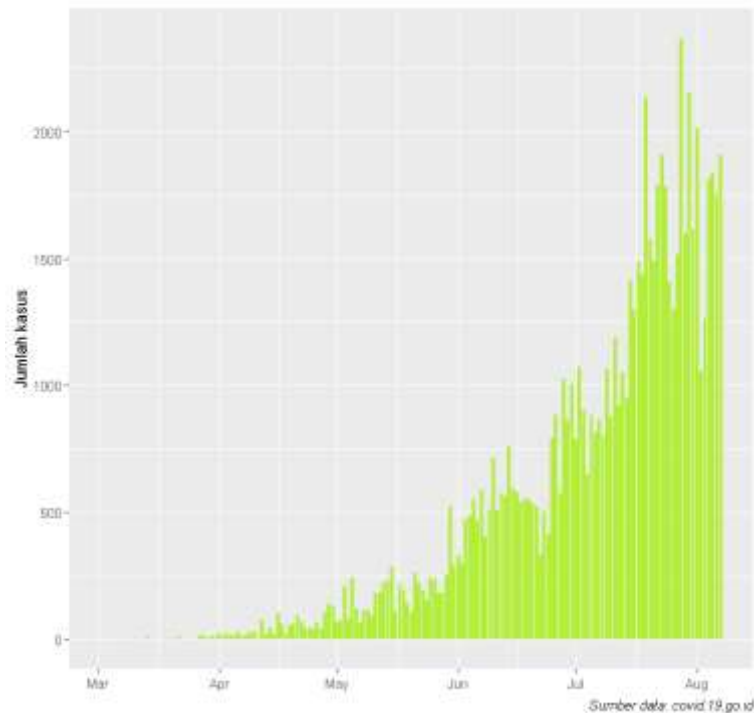
Sekarang kita akan menampilkan grafik kasus harian positif, kasus harian sembuh, dan kasus harian meninggal. Sumbu x berupa tanggal dan sumbu y berupa jumlah kasus.

```
#Menunjukkan melalui gambar grafik positif beserta beberapa keterangan tambahan
library(ggplot2)
library(hrbrthemes)
ggplot(cov_id_new, aes(Tanggal, Jumlah_Positif)) +
  geom_col(fill = "darkturquoise") +
  labs(
    x = NULL,
    y = "Jumlah Kasus",
    title = "Kasus Harian Positif COVID-19 di Indonesia",
    caption = "Sumber data: covid.19.go.id"
  ) + theme(
    plot.title = element_text(hjust = 0.8, size = 21, face = "bold"),
    plot.subtitle = element_text(hjust = 0, size = 12),
    plot.caption = element_text(hjust = 1, face = "italic"),
  ) +
  theme(plot.title.position = "plot")
```



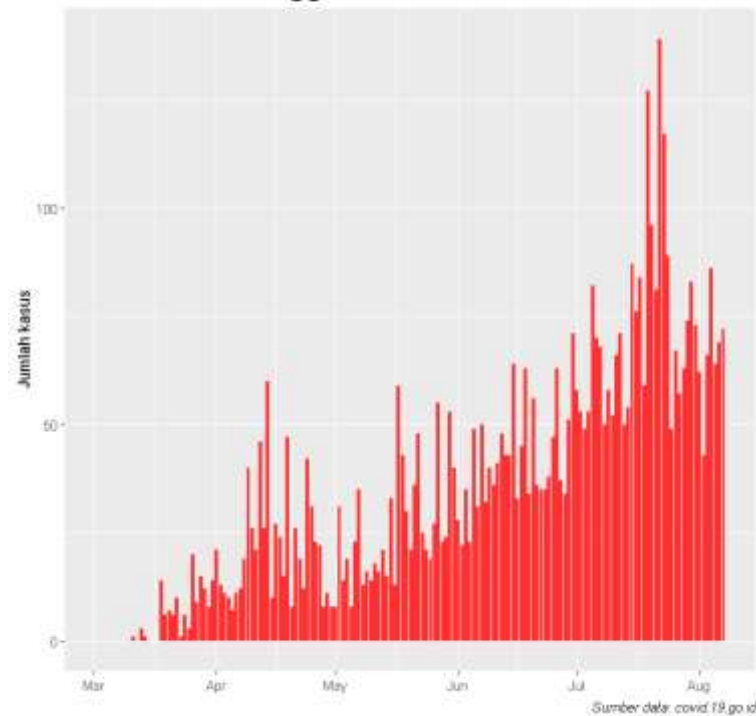
```
#Grafik untuk kasus sembuh
library(ggplot2)
library(hrbrthemes)
ggplot(cov_indo_new, aes(Tanggal, Jumlah_Sembuh)) +
  geom_col(fill = "olivedrab2") +
  labs(
    x = NULL,
    y = "Jumlah kasus",
    title = "Kasus Harian Sembuh Dari COVID-19 di Indonesia",
    caption = "Sumber data: covid.19.go.id"
  ) + theme(
    plot.title = element_text(hjust = 1.2, size = 20, face = "bold"),
    plot.subtitle = element_text(hjust = 0, size = 12),
    plot.caption = element_text(hjust = 1, face = "italic"),
  ) +
  theme(plot.title.position = "plot")
```

Kasus Harian Sembuh Dari COVID-19 di Indonesia



```
#Grafik untuk kasus meninggal
library(ggplot2)
library(hrbrthemes)
ggplot(cov_id_new, aes(Tanggal, Jumlah_Meninggal)) +
  geom_col(fill = "firebrick1") +
  labs(
    x = NULL,
    y = "Jumlah kasus",
    title = "Kasus Harian Meninggal Akibat COVID-19 di Indonesia",
    caption = "Sumber data: covid.19.go.id"
  ) + theme(
    plot.title = element_text(hjust = 0.47, size = 19, face = "bold"),
    plot.subtitle = element_text(hjust = 0, size = 12),
    plot.caption = element_text(hjust = 1, face = "italic"),
  ) +
  theme(plot.title.position = "plot")
```

Kasus Harian Meninggal Akibat COVID-19 di Indonesia



Pola dan Dinamika

Ada yang akhirnya sembuh, namun tak sedikit pula yang meninggal akibat COVID-19. Sementara itu penambahan kasus baru terus terjadi di masyarakat. Hal ini mungkin memicu pertanyaan lain di diri Anda: “Hingga saat ini ada berapa kasus yang masih aktif?”. Aktif dalam artian sedang dalam perawatan atau isolasi. Kita dapat menggunakan fungsi `cumsum()` untuk menghitung nilai akumulasi dari suatu vektor numerik.

```
#Menghitung nilai akumulasi dari suatu vektor numerik
library(dplyr)
cov_id_akumulasi <-
  cov_id_new %>%
  transmute(
    Tanggal,
    akumulasi_aktif = cumsum(Jumlah_Aktif),
    akumulasi_sembuh = cumsum(Jumlah_Sembuh),
    akumulasi_meninggal = cumsum(Jumlah_Meninggal)
  )
tail(cov_id_akumulasi)
```


A data.frame: 6 × 4

	Tanggal	akumulasi_aktif	akumulasi_sembuh	akumulasi_meninggal
	<date>	<int>	<int>	<int>
154	2020-08-02	37244	68975	5236
155	2020-08-03	37595	70237	5302
156	2020-08-04	37618	72050	5388
157	2020-08-05	37530	73889	5452
158	2020-08-06	37587	75645	5521
159	2020-08-07	38076	77557	5593

Grafik Komparasi Data

Kita akan membuat grafik komparasi antara akumulasi kasus aktif, kasus sembuh, dan kasus meninggal.

Ada dua pilihan cara yang dapat kita tempuh untuk membuat grafik tersebut:

1. Menggunakan data `cov_id_akumulasi`, kemudian buat tiga *layer* `geom_line()` untuk setiap kolom satu per satu dengan warna garis yang berbeda; atau
2. Melakukan *pivot* pada data `cov_id_akumulasi`, kemudian buat satu *layer* `geom_line()` dengan menggunakan *colour aesthetic*

Kali ini kita akan menempuh cara nomor 2, yaitu dengan cara melakukan *pivot* pada data terlebih dahulu.

Kita akan menggunakan fungsi `gather()` dari paket `tidyr` untuk mentransformasi data `cov_id_akumulasi`. Data tersebut akan dirubah dari yang semula berformat *wide* menjadi format *long*. Selanjutnya hasil transformasi data tersebut akan disimpan sebagai `cov_id_akumulasi_pivot`.

```
#Melakukan pivot, transformasi data dari wide ke Long
library(dplyr)
library(tidyr)

dim(cov_id_akumulasi)

cov_id_akumulasi_pivot <-
  cov_id_akumulasi %>%
  gather(
    key = "kategori",
    value = "Jumlah",
    -Tanggal
  ) %>%
  mutate(
    kategori = sub(pattern = "akumulasi_", replacement = "", kategori)
  )
```


159 · 4

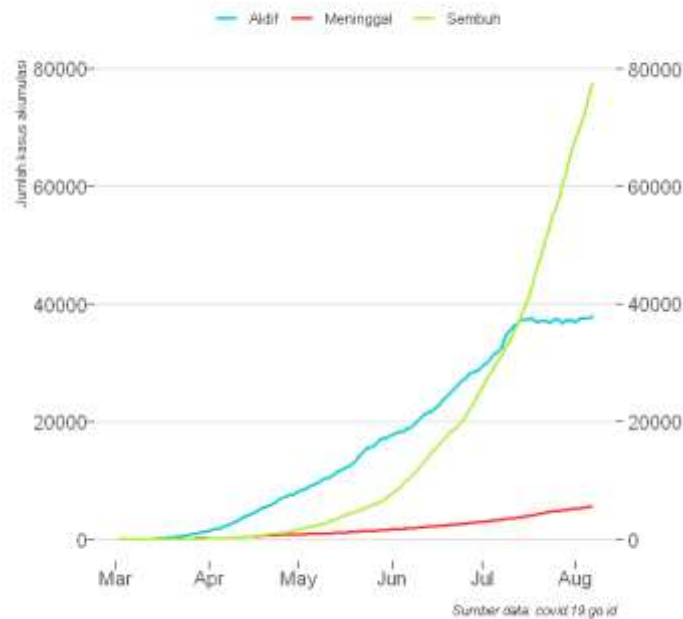
477 · 3

```
Rows: 477
Columns: 3
$ Tanggal <date> 2020-03-02, 2020-03-03, 2020-03-04, 2020-03-05, 2020-03-0...
$ kategori <chr> "aktif", "aktif", "aktif", "aktif", "aktif", "aktif", "akt...
$ Jumlah <int> 2, 2, 2, 2, 4, 4, 6, 19, 25, 31, 31, 63, 83, 104, 121, 158...
```

Setelah melakukan *pivot* atau transformasi data, maka selanjutnya kita akan membuat grafik komparasi.

```
#Membuat komparasi antara kasus aktif, sembuh, dan meninggal di Indonesia
library(ggplot2)
library(hrbrthemes)
ggplot(cov_id_akumulasi_pivot, aes(Tanggal, Jumlah, colour = (kategori))) +
  geom_line(size = 0.9) +
  scale_y_continuous(sec.axis = dup_axis(name = NULL)) +
  scale_colour_manual(
    values = c(
      "aktif" = "darkturquoise",
      "meninggal" = "firebrick1",
      "sembuh" = "olivedrab2"
    ),
    labels = c("Aktif", "Meninggal", "Sembuh")
  ) +
  labs(
    x = NULL,
    y = "Jumlah kasus akumulasi",
    colour = NULL,
    title = "Dinamika Kasus COVID-19 di Indonesia",
    caption = "Sumber data: covid.19.go.id"
  ) +
  theme_ipsum(
    base_size = 13,
    plot_title_size = 21,
    grid = "Y",
    ticks = TRUE
  ) +
  legend.position = "top"
)
```

Dinamika Kasus COVID-19 di Indonesia



Grafik Perkembangan Kasus Pekan

Setelah mengamati grafik perkembangan kasus kita menyadari bahwa terjadi fluktuasi pertambahan kasus harian. Dilandasi hal tersebut kita kemudian ingin mencoba mengamati bagaimana perkembangan kasus dalam rentang waktu pekan.

Kita dapat dengan mudah bekerja dengan data tanggal apabila menggunakan paket **lubridate**. Adapun yang akan digunakan untuk mengekstrak informasi pekan dalam satu tahun adalah fungsi **week()**.

Kita akan menyimpan hasilnya sebagai **cov_id_pekanan**. Kita juga akan menggunakan fungsi **glimpse()** dari **dplyr** untuk melakukan inspeksi data.

```
#Mengamati kasus pekanan dan melakukan inspeksi data
library(dplyr)
library(lubridate)

cov_id_pekanan <- cov_id_new %>%
  count(
    tahun = year(Tanggal),
    pekan_ke = week(Tanggal),
    wt = Jumlah_Positif,
    name = "Jumlah"
  )

glimpse(cov_id_pekanan)
```

```

Rows: 24
Columns: 3
$ tahun      <dbl> 2020, 2020, 2020, 2020, 2020, 2020, 2020, 2020, 2020, 2020...
$ pekan_ke   <dbl> 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23,...
$ Jumlah     <int> 2, 25, 145, 513, 843, 1210, 2101, 2296, 2376, 2560, 2678, ...

```

Pertanyaan baru muncul di benak kita setelah melakukan inspeksi terhadap data `cov_id_pekanan` tersebut: “Apakah pekan ini lebih baik dari pekan kemarin?”.

Demi menjawab hal tersebut kita melakukan kalkulasi sederhana dengan tahapan berikut:

- Membuat kolom baru yang berisi jumlah kasus baru dalam satu pekan sebelumnya. Kolom ini diberi nama “jumlah_pekanlalu”
- Mengganti nilai `NA` pada kolom “jumlah_pekanlalu” dengan nilai 0
- Melakukan komparasi antara kolom “jumlah” dengan kolom “jumlah_pekanlalu”. Hasil komparasi ini disimpan dalam kolom baru dengan nama “lebih_baik”, isinya adalah `TRUE` apabila jumlah kasus baru pekan ini lebih rendah dibandingkan jumlah kasus pekan lalu

Kita akan menggunakan fungsi `lag()` dari `dplyr` untuk membuat kolom “jumlah_pekanlalu”. Perhatikan bahwa disini fungsi tersebut dituliskan sebagai `dplyr::lag()` untuk menghindari konflik dengan fungsi `lag()` dari paket `stats`. Selanjutnya kita akan melakukan inspeksi hasil pekerjaan dengan menggunakan fungsi `glimpse()`.

```

#Apakah pekan ini Lebih baik dari pekan kemarin?
library(dplyr)
cov_id_pekanan <-
  cov_id_pekanan %>%
  mutate(
    jumlah_pekanlalu = dplyr::lag(Jumlah, 1),
    jumlah_pekanlalu = ifelse(is.na(jumlah_pekanlalu), 0, jumlah_pekanlalu),
    lebih_baik = Jumlah < jumlah_pekanlalu
  )
glimpse(cov_id_pekanan)

```

```

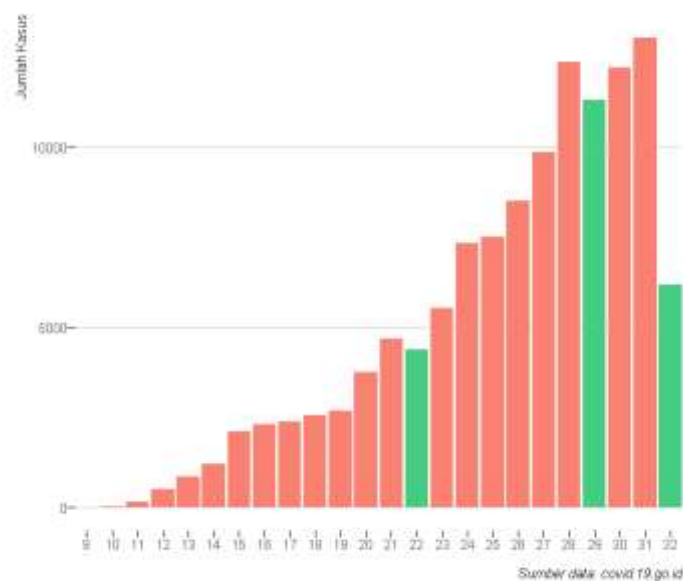
Rows: 24
Columns: 5
$ tahun      <dbl> 2020, 2020, 2020, 2020, 2020, 2020, 2020, 2020, 20...
$ pekan_ke    <dbl> 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21,...
$ Jumlah      <int> 2, 25, 145, 513, 843, 1210, 2101, 2296, 2376, 2560...
$ jumlah_pekanlalu <dbl> 0, 2, 25, 145, 513, 843, 1210, 2101, 2296, 2376, 2...
$ lebih_baik  <lgl> FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, F...

```

```
#Membuat bar chart yang menunjukkan kondisi kasus pekanan
library(ggplot2)
library(hrbrthemes)
ggplot(cov_id_pekanan, aes(pekan_ke, Jumlah, fill = lebih_baik)) +
  geom_col(show.legend = FALSE) +
  scale_x_continuous(breaks = 9:32, expand = c(0, 0)) +
  scale_fill_manual(values = c("TRUE" = "seagreen3", "FALSE" = "salmon")) +
  labs(
    x = NULL,
    y = "Jumlah Kasus",
    title = "Kasus Pekan Positif COVID-19 di Indonesia",
    subtitle = "Kolom hijau menunjukkan penambahan kasus baru lebih sedikit
dibandingkan satu pekan sebelumnya",
    caption = "Sumber data: covid.19.go.id"
  ) +
  theme_ipsum(
    base_size = 9,
    plot_title_size = 18,
    grid = "Y",
    ticks = TRUE
  ) +
  theme(plot.title.position = "plot")
```

Kasus Pekan Positif COVID-19 di Indonesia

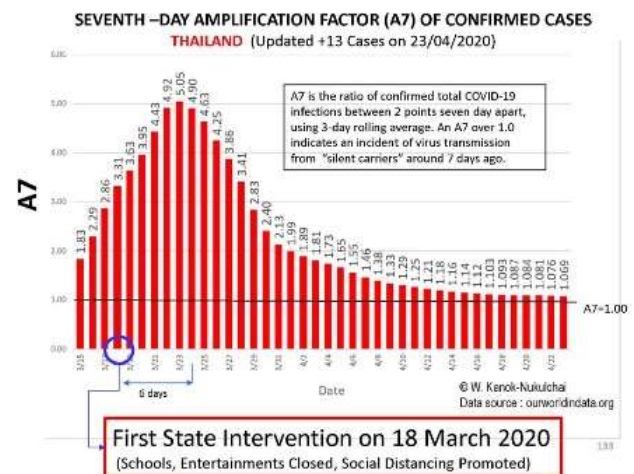
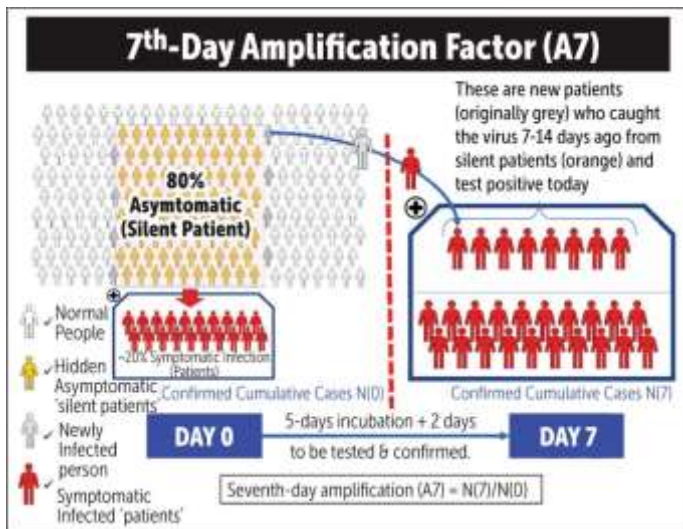
Kolom hijau menunjukkan penambahan kasus baru lebih sedikit dibandingkan satu pekan sebelumnya



Seventh-Day Amplification Factor

Seventh-Day Amplification Factor atau *A7 factor* adalah rasio hasil bagi dari jumlah akumulasi kasus positif di hari ke-7 dengan jumlah akumulasi kasus positif di hari ke-0 (7 hari lalu sebelum perhitungan A7).

Nilai lebih besar dari 1.0 menunjukkan bahwa banyak transmisi masih terjadi. Hal tersebut menyiratkan bahwa banyak Orang Tanpa Gejala (OTG) masih berkeliaran di ruang publik sekitar tujuh hari yang lalu dari dihitungnya *A7 factor*. Sebaliknya, rasio $A7 = 1.0$ yang sempurna menyiratkan tidak ada penularan selama seminggu, dan karenanya tidak ada lagi OTG yang aktif di masyarakat. Untaian panjang nilai 1.0 pada grafik *A7* masih diperlukan untuk memastikan kondisi stabil sebelum mendeklarasikan bahwa negara tersebut bebas dari corona (seperti dalam kasus Cina).



Gambar di atas merupakan gambaran dari *A7 factor* dan contoh aplikasinya di Thailand. Adapun juga asumsi yang diperoleh dari *Seventh-Day Amplification Factor*:

- Pasien yang dites positif hari ini telah terinfeksi virus setidaknya tujuh hari yang lalu (dengan asumsi lima hari inkubasi dan dua hari *testing*).
- Pasien-pasien ini (pasien hari ke-7) terinfeksi oleh OTG di hari ke-0, dan bukan oleh pasien positif di hari ke-0 (karena mereka sudah dikarantina).
- OTG mewakili 80% dari semua kasus yang terinfeksi, atau sekitar empat kali jumlah pasien positif di hari ke-0.

Penggunaan *A7* dapat bermanfaat untuk:

- Mengawasi trend pergerakan COVID-19
- Menganalisis keefektifan dari intervensi suatu negara
- Untuk melakukan *forecasting* kapan penyebaran COVID-19 akan berakhir
- Mendeteksi akan adanya Orang Tanpa Gejala (OTG)
- Mengontrol kebijakan (*social distancing*, PSBB, dll)
- Mengambil langkah di berbagai bidang (ekonomi, politik, sosial, budaya, dll) dengan tujuan untuk menstabilkan keamanan dan ketahanan negara

Sekarang kita akan menghitung A7 *factor* di Indonesia. Indonesia mendata kasus positif per 2 Maret 2020. Oleh karena itu kita baru dapat menghitung A7 di 9 Maret 2020 sampai 7 Agustus 2020. Pertama kita akan menghitung akumulasi kasus positif.

```
#Membuat akumulasi Jumlah_Positif di Indonesia
library(dplyr)
cov_id_akumulasi_positif <-
  cov_id_new %>%
  transmute(
    Tanggal, akumulasi_positif = cumsum(Jumlah_Positif)
  )

tail(cov_id_akumulasi_positif)
```

A data.frame: 6 × 2

	Tanggal	akumulasi_positif
	<date>	<int>
154	2020-08-02	111455
155	2020-08-03	113134
156	2020-08-04	115056
157	2020-08-05	116871
158	2020-08-06	118753
159	2020-08-07	121226

Setelah itu, kita akan mengambil data N(0) dan N(7). Kemudian akan dimasukkan ke dalam satu *data frame* yang berisi tanggal, N(0), N(7), A7 normal (yang seharusnya), dan A7 Indonesia (yang terjadi).

```
#Mengambil data N(0)
library(dplyr)

cov_id_akumulasi_positif$Tanggal <-
as.Date(cov_id_akumulasi_positif$Tanggal)
picked_date_N_0_id <- filter(cov_id_akumulasi_positif,
                             Tanggal >= "2020-03-02", Tanggal <= "2020-07-31")
N0_id <- picked_date_N_0_id$akumulasi_positif
```

```
#Mengambil data N(7)
library(dplyr)

cov_id_akumulasi_positif$Tanggal <-
as.Date(cov_id_akumulasi_positif$Tanggal)
picked_date_N_7_id <- filter(cov_id_akumulasi_positif,
                             Tanggal >= "2020-03-09", Tanggal <= "2020-08-07")
N7_id <- picked_date_N_7_id$akumulasi_positif
```


Setelah dirapikan datanya, maka kita akan membuat grafik A7. Dilakukan proses penggabungan antara grafik batang dengan garis. Garis biru menunjukkan A7 yang terjadi di Indonesia dan garis merah menunjukkan A7 yang seharusnya yaitu bernilai 1.0.

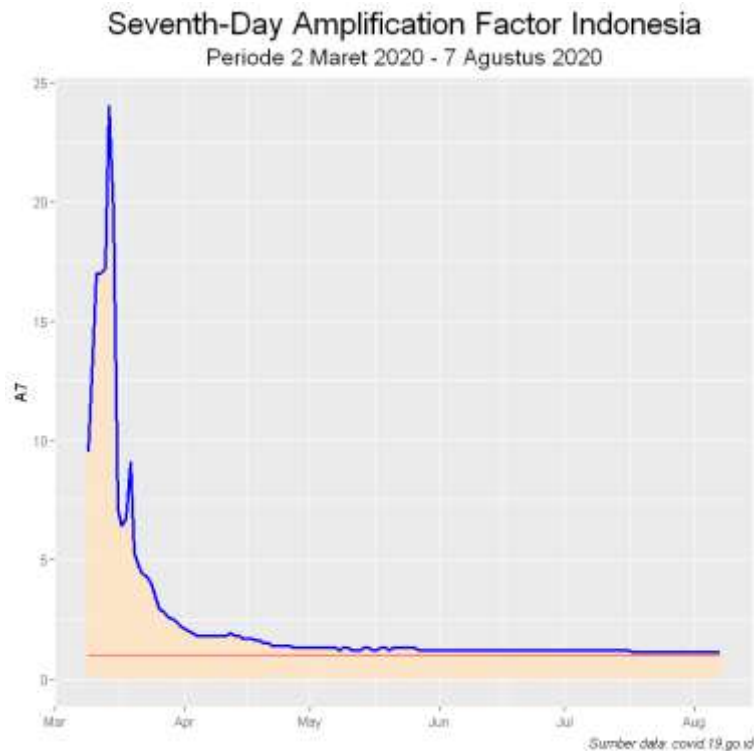
```
#Memasukkan data N0 ke sebuah kolom
picked_date_N_7_id$N0_id <- (N0_id)
#Mengganti kolom "akumulasi_positif" menjadi "N7"
names(picked_date_N_7_id)[names(picked_date_N_7_id)
                           == "akumulasi_positif"] <- "N7_indo"
#Menambahkan kolom A7 Factor yang seharusnya, yaitu A7 = 1.0
picked_date_N_7_id$A7_normal <- (1.0)
#Menghitung A7 Factor (N7/N0)
picked_date_N_7_id$A7_id <- (N7_id / N0_id)
#Proses pembulatan satu desimal
picked_date_N_7_id$A7_id <- round(picked_date_N_7_id$A7_id, digits=1)

head(picked_date_N_7_id)
```

A data.frame: 6 × 5

	Tanggal	N7_indo	N0_id	A7_normal	A7_id
	<date>	<int>	<int>	<dbl>	<dbl>
1	2020-03-09	19	2	1	9.5
2	2020-03-10	27	2	1	13.5
3	2020-03-11	34	2	1	17.0
4	2020-03-12	34	2	1	17.0
5	2020-03-13	69	4	1	17.2
6	2020-03-14	96	4	1	24.0

```
#Menampilkan bar chart A7 disertai dengan line
library(ggplot2)
library(hrbrthemes)
plot.N7_id <- ggplot(picked_date_N_7_id, aes(x = Tanggal, y = A7_id)) +
  geom_col(fill="bisque1") +
  labs(
    x = NULL,
    y = "A7",
    title = "Seventh-Day Amplification Factor Indonesia",
    subtitle = "Periode 2 Maret 2020 - 7 Agustus 2020",
    caption = "Sumber data: covid.19.go.id"
  ) + theme(
    plot.title = element_text(hjust = 0.5, size = 21),
    plot.subtitle = element_text(hjust = 0.5, size = 15),
    plot.caption = element_text(hjust = 1, face = "italic")
  )
plot.N7_id <- plot.N7_id +
  geom_line(aes(x=Tanggal, y=A7_normal, fill=A7_normal), lwd=0.5, colour="red")
plot.N7_id +
  geom_line(aes(x=Tanggal, y=A7_id, fill=A7_id), lwd=1, colour="blue")
```



Kesimpulan

Data COVID-19 di Indonesia telah direkam sejak 2 Maret 2020. Pada awalnya, kasus positif dan aktif tercatat sejumlah 2 orang Warga Negara Indonesia (WNI) yang tinggal di wilayah Depok, Jawa Barat. Salah satunya tertular dari warga negara Jepang di salah satu tempat dansa. Warga Jepang itu baru terdeteksi positif corona setelah meninggalkan Indonesia dan tiba di Malaysia. Dari analisis dan visualisasi data COVID-19 di Indonesia, kita dapat simpulkan:

- Data pertama kali di rekam 2 Maret 2020. Dengan jumlah aktif dan positif berjumlah 2 orang, jumlah meninggal dan sembuh berjumlah 0.
- Data terakhir *diupdate* tanggal 7 Agustus 2020. Penambahan kasus sembuh adalah 1,912 orang, penambahan kasus meninggal adalah 72 orang, total jumlah kasus positif adalah 121,226 orang, dan total jumlah kasus meninggal 5,593.
- Per tanggal 7 Agustus 2020, persentase kasus aktif adalah 31.4%, persentase kasus sembuh adalah 64%, dan persentase kasus meninggal adalah 4.6%.
- Jumlah penduduk di Indonesia sejumlah 267.7 juta jiwa (2018). Berarti persentase pasien positif COVID-19 di Indonesia tanggal 7 Agustus 2020 adalah 0.045% dari total jumlah penduduk Indonesia.
- Kenaikan total jumlah kasus positif dari tanggal 2 Maret 2020 sampai 7 Agustus 2020 sebesar 6,061,200 %.

- Dilihat dari grafik komparasi antara akumulasi kasus aktif, sembuh, dan meninggal. Kasus meninggal memang seharusnya berada di bawah garis kasus aktif dan sembuh. Garis kasus meninggal harus ditekan agar tidak terjadi kenaikan lagi. Jika garis akumulasi kasus meninggal sudah mendatar (stabil) dalam waktu yang cukup lama, maka sudah tidak ada penambahan kasus meninggal. Dan jika hal itu terjadi, maka wajib dipertahankan serta diambil beberapa keputusan guna menekan terjadinya kasus meninggal.
- Kasus aktif saat ini masih naik turun, tetapi berada di kondisi cukup mendatar. Kasus aktif lebih baik jika menurun dan berada pada titik terendah, yaitu nol. Hal itu menandakan tidak adanya pasien dirawat, positif corona, atau sedang dikarantina. Penekanan kasus aktif dan meninggal sebenarnya dapat ditekan dengan meningkatkan penanganan terhadap pasien positif. Adapun juga kebijakan lainnya guna menurunkan jumlah penambahan kasus.
- Kasus sembuh menandakan makin naik pada grafik. Hal tersebut sudah bagus dan perlu dipertahankan agar terus naik. Pemerintah seharusnya menyediakan fasilitas kesehatan yang lebih canggih dan lebih banyak dalam penanganan pasien positif, mengingat jumlah pasien positif semakin banyak dan tenaga kesehatan semakin kewalahan. Pemerintah Indonesia juga dapat bekerja sama dengan negara lain dalam penyediaan fasilitas kesehatan. Dengan begitu, maka grafik kasus sembuh akan semakin naik dan grafik kasus meninggal akan stabil (tidak ada penambahan).
- Grafik kasus pekanan COVID-19 di Indonesia diambil dari minggu ke-9 sampai ke-32 di tahun 2020. Terlihat hanya 3 pekan atau hanya 9.4% yang mengalami perbaikan dari pekan sebelumnya. Yang mana sebanyak 90.6% kasus makin parah di masing-masing pekan dari pekan sebelumnya.
- *Seventh-Day Amplification Factor* tertinggi di Indonesia sebesar 24.0 di bulan Maret. Sejak 17 Juli 2020 sampai 7 Agustus 2020, A7 di Indonesia menandakan kestabilan di nilai pembulatan 1.1. Jika kebijakan pemerintah terkait protokol kesehatan dilakukan dengan baik di seluruh lapisan masyarakat, berdasarkan data yang diperoleh maka A7 dapat menyentuh 1.0 di tahun ini.

III. Analisis dan Visualisasi Data di DKI Jakarta

Introduction

Proses analisis dan visualisasi data di DKI Jakarta memiliki alur pengerjaan yang sama dengan analisis dan visualisasi data di Indonesia. Hanya saja beberapa *variable*, situs yang diakses, struktur *data frame*, dan lain-lain terdapat sedikit perbedaan. Data yang kita akses sekarang (7 Agustus 2020) terakhir diperbarui tanggal 6 Agustus 2020. Oleh karena itu, selama analisis dan visualisasi data di DKI Jakarta saat ini, kita hanya dapat mengakses data dari tanggal 1 Maret 2020 sampai 6 Agustus 2020 saja.

Mengakses API covid.go.id

```
#Read API
library(httr)
resp_jkt <-
GET("https://data.covid19.go.id/public/api/prov_detail_DKI_JAKARTA.json")
```

Mengekstrak Isi Respon

```
#Ekstrak konten JSON
cov_jkt_raw <- content(resp_jkt, as = "parsed", simplifyVector = TRUE)

#Komponen cov_jkt_raw
names(cov_jkt_raw)

'last_date' · 'provinsi' · 'kasus_total' · 'kasus_tanpa_tgl' · 'kasus_dengan_tgl' ·
'meninggal_persen' · 'meninggal_tanpa_tgl' · 'meninggal_dengan_tgl' · 'sembuh_persen' ·
'sembuh_tanpa_tgl' · 'sembuh_dengan_tgl' · 'list_perkembangan' · 'data'
```

Analisis Data

Sekarang kita fokus pada obyek `cov_jkt_raw` untuk menjawab pertanyaan-pertanyaan berikut:

- Berapa jumlah total jumlah kasus positif?
- Berapa persentase pasien meninggal?
- Berapa persentase pasien sembuh?
- Berapa persentase pasien aktif?

```
#Analisa COVID-19 di DKI Jakarta
cov_jkt_raw$kasus_total
cov_jkt_raw$meninggal_persen
cov_jkt_raw$sembuh_persen
```

```
#Menghitung persentase pasien aktif
cov_jkt_aktif <-
(100 - cov_jkt_raw$sembuh_persen - cov_jkt_raw$meninggal_persen)
cov_jkt_aktif
```

23936

3.76420454545455

62.692179144385

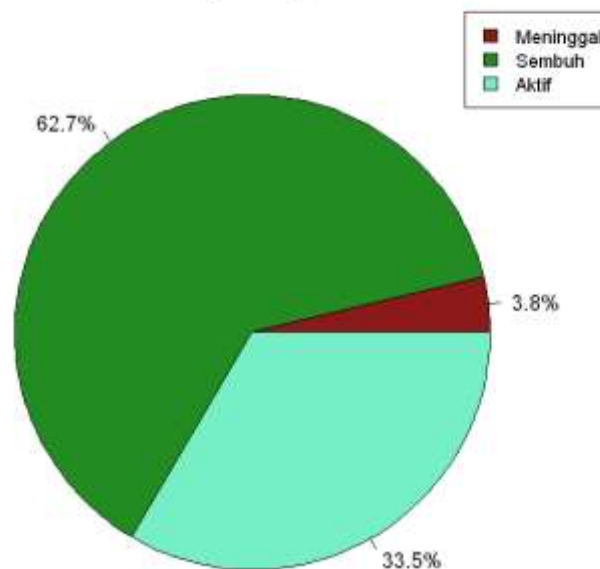
33.5436163101604

Lalu kita juga dapat membuat visualisasi melalui *pie chart* untuk melihat persentase jumlah total kasus meninggal, sembuh, dan aktif di DKI Jakarta.

```
#Membuat visualisasi data pasien aktif, sembuh, dan meninggal
x <- c(cov_jkt_raw$meninggal_persen, cov_jkt_raw$sembuh_persen, cov_jkt_aktif)
labels <- c("meninggal_persen", "sembuh_persen", "cov_jkt_aktif")
x <- round(x, digits=1)
labels <- paste(x, "%", sep="")

# Plot the chart.
pie(x, labels = labels,
    main = "Persentase Total Pasien Aktif, Sembuh, dan Meninggal di
DKI Jakarta per 6 Agustus 2020",
    col = c("firebrick4", "forestgreen", "aquamarine2"))
legend("topright", c("Meninggal", "Sembuh", "Aktif"),
    cex = 0.9,
    fill = c("firebrick4", "forestgreen", "aquamarine2"))
```

Persentase Total Pasien Aktif, Sembuh, dan Meninggal di
DKI Jakarta per 6 Agustus 2020



Memperoleh Informasi yang Lebih Lengkap

```
#Kita ambil list_perkembangan dan masukkan ke variable cov_jkt
cov_jkt <- cov_jkt_raw$list_perkembangan
str(cov_jkt)
```

```
#Mengambil bagian teratas dari list_perkembangan
head(cov_jkt)
```

```
'data.frame': 159 obs. of 9 variables:
 $ tanggal          : num  1.58e+12 1.58e+12 1.58e+12 1.58e+12 1.58e
+12 ...
 $ KASUS              : int   2 2 2 2 0 0 0 0 0 0 ...
 $ MENINGGAL          : int   0 0 0 0 1 0 2 0 1 0 ...
 $ SEMBUH             : int   0 0 0 0 0 0 0 0 0 0 ...
 $ DIRAWAT_OR_ISOLASI : int   2 2 2 2 -1 0 -2 0 -1 0 ...
 $ AKUMULASI_KASUS    : int   2 4 6 8 8 8 8 8 8 8 ...
 $ AKUMULASI_SEMBUH   : int   0 0 0 0 0 0 0 0 0 0 ...
 $ AKUMULASI_MENINGGAL : int   0 0 0 0 1 1 3 3 4 4 ...
 $ AKUMULASI_DIRAWAT_OR_ISOLASI: int  2 4 6 8 7 7 5 5 4 4 ...
```

A data.frame: 6 × 9

	tanggal	KASUS	MENINGGAL	SEMBUH	DIRAWAT_OR_ISOLASI	AKUMULASI_KASUS	AK
	<dbl>	<int>	<int>	<int>	<int>	<int>	<int>
1	1.583021e+12	2	0	0	2	2	2
2	1.583107e+12	2	0	0	2	4	4
3	1.583194e+12	2	0	0	2	6	6
4	1.583280e+12	2	0	0	2	8	8
5	1.583366e+12	0	1	0	-1	8	8
6	1.583453e+12	0	0	0	0	8	8

Menjinakkan Data

Setelah mengekstrak dan mengamati `cov_jkt`, kita menemukan beberapa kejanggalan pada data tersebut. Diantaranya adalah kejanggalan data pada kolom tanggal dan format penulisan kolom yang tidak konsisten. Sekarang kita akan mencoba melakukan beberapa tahapan untuk menjinakkan data tersebut sehingga dapat diolah dan dianalisis dengan lebih mudah.

Sebelum itu, kita aktifkan paket `dplyr` yang akan dipergunakan untuk melakukan pengolahan data.

Ada beberapa tahapan yang akan kita lakukan untuk menjinakkan data `cov_jkt`, yaitu:

1. Menghapus kolom “DIRAWAT_OR_ISOLASI” dan “AKUMULASI_DIRAWAT_OR_ISOLASI”

2. Menghapus semua kolom yang berisi nilai kumulatif
3. Mengganti nama kolom "KASUS" menjadi "kasus_baru"
4. Merubah format penulisan kolom berikut menjadi huruf kecil
 - kolom MENINGGAL
 - kolom SEMBUH
5. Memperbaiki data pada kolom tanggal

Kita akan menggunakan operator *pipe* (`%>%`) untuk merangkai fungsi menjadi sebuah *pipeline*. Lalu simpan hasil pengolahan dengan nama `new_cov_jkt`.

```
#Menjinakkan data dan merapikan kolom tanggal yang janggal
library(dplyr)
new_cov_jkt <-
  cov_jkt %>%
  select(-contains("DIRAWAT_OR_ISOLASI")) %>%
  rename(
    kasus_baru = KASUS,
    meninggal = MENINGGAL,
    sembuh = SEMBUH
  ) %>%
  mutate(
    tanggal = as.POSIXct(tanggal / 1000, origin = "1970-01-01"),
    tanggal = as.Date(tanggal)
  )
str(new_cov_jkt)
```

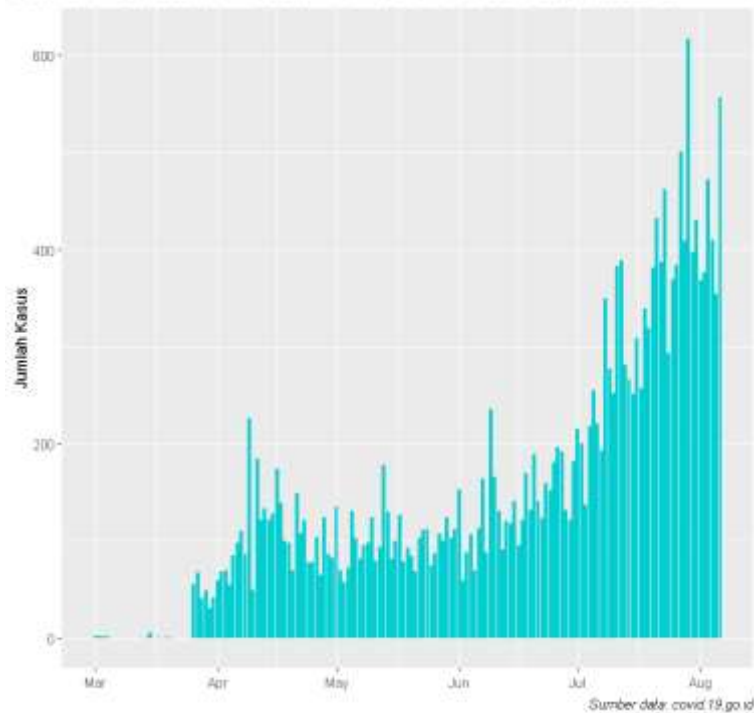
```
'data.frame': 159 obs. of 7 variables:
 $ tanggal      : Date, format: "2020-03-01" "2020-03-02" ...
 $ kasus_baru    : int  2 2 2 2 0 0 0 0 0 0 ...
 $ meninggal     : int  0 0 0 0 1 0 2 0 1 0 ...
 $ sembuh       : int  0 0 0 0 0 0 0 0 0 0 ...
 $ AKUMULASI_KASUS : int  2 4 6 8 8 8 8 8 8 8 ...
 $ AKUMULASI_SEMBUH : int  0 0 0 0 0 0 0 0 0 0 ...
 $ AKUMULASI_MENINGGAL: int  0 0 0 0 1 1 3 3 4 4 ...
```

Visualisasi Data Dengan Grafik

```
#Menunjukkan melalui gambar grafik positif beserta beberapa keterangan tambahan
library(ggplot2)
library(hrbrthemes)
ggplot(new_cov_jkt, aes(tanggal, kasus_baru)) +
  geom_col(fill = "darkturquoise") +
  labs(
    x = NULL,
    y = "Jumlah Kasus",
    title = "Kasus Harian Positif COVID-19 di DKI Jakarta",
    caption = "Sumber data: covid.19.go.id"
  ) + theme(
    plot.title = element_text(hjust = 0, size = 21, face = "bold"),
    plot.subtitle = element_text(hjust = 0, size = 12),
    plot.caption = element_text(hjust = 1, face = "italic"),
```

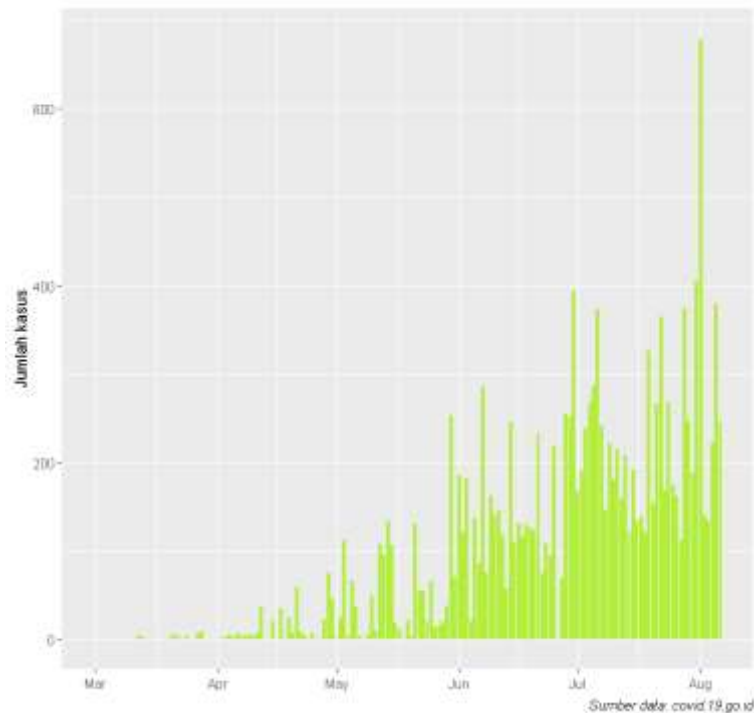
```
) +  
theme(plot.title.position = "plot")
```

Kasus Harian Positif COVID-19 di DKI Jakarta



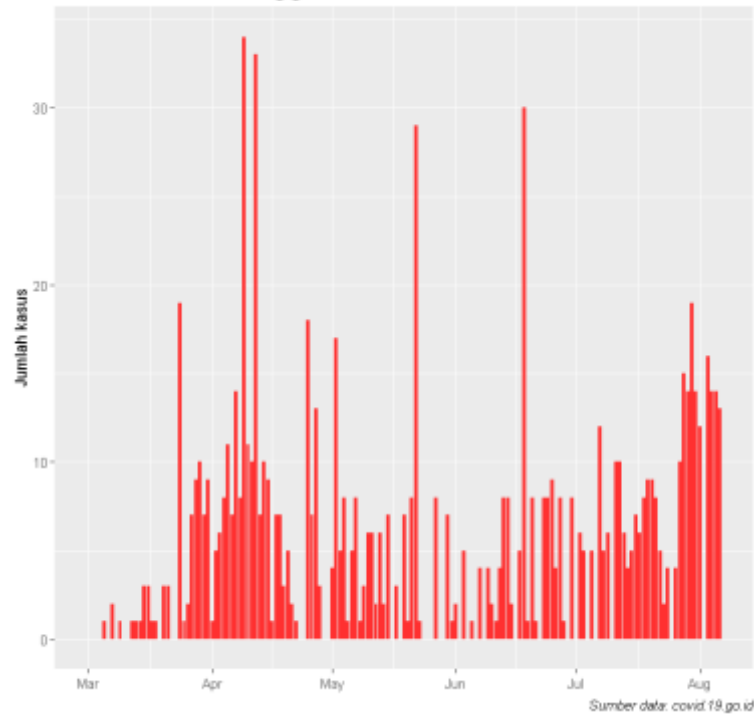
```
#Grafik untuk kasus sembuh  
library(ggplot2)  
library(hrbrthemes)  
ggplot(new_cov_jkt, aes(tanggal, sembuh)) +  
  geom_col(fill = "olivedrab2") +  
  labs(  
    x = NULL,  
    y = "Jumlah kasus",  
    title = "Kasus Harian Sembuh Dari COVID-19 di DKI Jakarta",  
    caption = "Sumber data: covid.19.go.id"  
  ) + theme(  
    plot.title = element_text(hjust = 0, size = 20, face = "bold"),  
    plot.subtitle = element_text(hjust = 0, size = 12),  
    plot.caption = element_text(hjust = 1, face = "italic"),  
  ) +  
  theme(plot.title.position = "plot")
```

Kasus Harian Sembuh Dari COVID-19 di DKI Jakarta



```
#Grafik untuk kasus meninggal
library(ggplot2)
library(hrbrthemes)
ggplot(new_cov_jkt, aes(tanggal, meninggal)) +
  geom_col(fill = "firebrick1") +
  labs(
    x = NULL,
    y = "Jumlah kasus",
    title = "Kasus Harian Meninggal Akibat COVID-19 di DKI Jakarta",
    caption = "Sumber data: covid.19.go.id"
  ) + theme(
    plot.title = element_text(hjust = 0, size = 18, face = "bold"),
    plot.subtitle = element_text(hjust = 0, size = 12),
    plot.caption = element_text(hjust = 1, face = "italic"),
  ) +
  theme(plot.title.position = "plot")
```

Kasus Harian Meninggal Akibat COVID-19 di DKI Jakarta



Pola dan Dinamika

```
#Menghitung nilai akumulasi dari suatu vektor numerik
library(dplyr)
cov_jkt_akumulasi <-
  new_cov_jkt %>%
  transmute(
    tanggal,
    akumulasi_aktif = cumsum(kasus_baru) - cumsum(sembuh) - cumsum(meninggal),
    akumulasi_sembuh = cumsum(sembuh),
    akumulasi_meninggal = cumsum(meninggal)
  )
tail(cov_jkt_akumulasi)
```

A data.frame: 6 × 4

	tanggal	akumulasi_aktif	akumulasi_sembuh	akumulasi_meninggal
	<date>	<int>	<int>	<int>
154	2020-08-01	6612	13844	824
155	2020-08-02	6848	13985	824
156	2020-08-03	7170	14119	840
157	2020-08-04	7344	14341	854
158	2020-08-05	7305	14720	868
159	2020-08-06	7601	14967	881

Grafik Komparasi Data

```
#Melakukan pivot, transformasi data dari wide ke Long
library(dplyr)
library(tidyr)

dim(cov_jkt_akumulasi)

cov_jkt_akumulasi_pivot <-
  cov_jkt_akumulasi %>%
  gather(
    key = "kategori",
    value = "jumlah",
    -tanggal
  ) %>%
  mutate(
    kategori = sub(pattern = "akumulasi_", replacement = "", kategori)
  )

dim(cov_jkt_akumulasi_pivot)

glimpse(cov_jkt_akumulasi_pivot)
```

```
#Membuat komparasi antara kasus aktif, sembuh, dan meninggal
library(ggplot2)
library(hrbrthemes)
ggplot(cov_jkt_akumulasi_pivot, aes(tanggal, jumlah, colour = (kategori))) +
  geom_line(size = 0.9) +
  scale_y_continuous(sec.axis = dup_axis(name = NULL)) +
  scale_colour_manual(
    values = c(
      "aktif" = "darkturquoise",
      "meninggal" = "firebrick1",
      "sembuh" = "olivedrab2"
    ),
    labels = c("Aktif", "Meninggal", "Sembuh")
  ) +
  labs(
    x = NULL,
    y = "Jumlah kasus akumulasi",
    colour = NULL,
    title = "Dinamika kasus COVID-19 di DKI Jakarta",
    caption = "Sumber data: covid.19.go.id"
  ) +
  theme_ipsum(
    base_size = 13,
    plot_title_size = 21,
    grid = "Y",
    ticks = TRUE
```

```
) +
theme(
  plot.title = element_text(hjust = 0.5),
  legend.position = "top"
)
```



Grafik Perkembangan Kasus Pekan

```
#Mengamati kasus pekan dan melakukan inspeksi data
library(dplyr)
library(lubridate)

cov_jkt_pekanan <- new_cov_jkt %>%
  count(
    tahun = year(tanggal),
    pekan_ke = week(tanggal),
    wt = kasus_baru,
    name = "jumlah"
  )

glimpse(cov_jkt_pekanan)
```

```
Rows: 24
Columns: 3
$ tahun    <dbl> 2020, 2020, 2020, 2020, 2020, 2020, 2020, 2020, 2020, 2020...
$ pekan_ke <dbl> 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23,...
$ jumlah   <int> 6, 2, 8, 2, 287, 546, 924, 858, 679, 635, 677, 789, 641, 7...
```

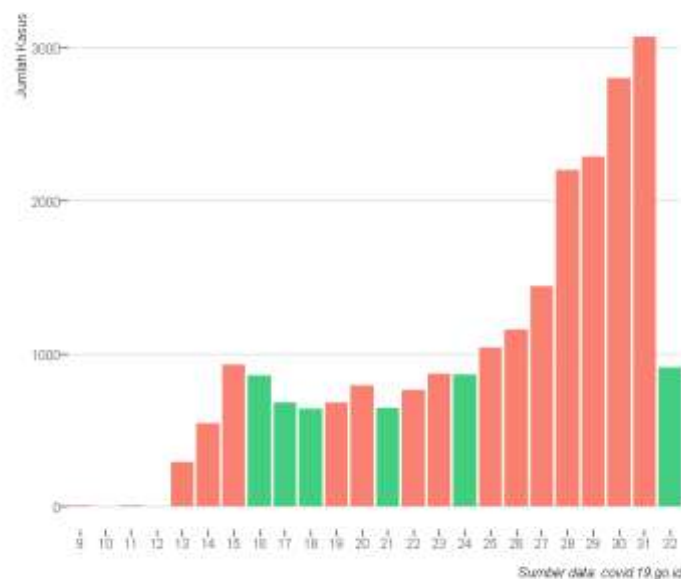
```
#Apakah pekan ini lebih baik dari pekan kemarin?
library(dplyr)
cov_jkt_pekanan <-
  cov_jkt_pekanan %>%
  mutate(
    jumlah_pekanlalu = dplyr::lag(jumlah, 1),
    jumlah_pekanlalu = ifelse(is.na(jumlah_pekanlalu), 0, jumlah_pekanlalu),
    lebih_baik = jumlah < jumlah_pekanlalu
  )
glimpse(cov_jkt_pekanan)

Rows: 24
Columns: 5
$ tahun      <dbl> 2020, 2020, 2020, 2020, 2020, 2020, 2020, 2020, 20...
$ pekan_ke   <dbl> 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21,...
$ jumlah     <int> 6, 2, 8, 2, 287, 546, 924, 858, 679, 635, 677, 789...
$ jumlah_pekanlalu <dbl> 0, 6, 2, 8, 2, 287, 546, 924, 858, 679, 635, 677, ...
$ lebih_baik <lgl> FALSE, TRUE, FALSE, TRUE, FALSE, FALSE, FALSE, TRU...
```

```
#Membuat bar chart yang menunjukkan kondisi kasus pekanan
library(ggplot2)
library(hrbrthemes)
ggplot(cov_jkt_pekanan, aes(pekan_ke, jumlah, fill = lebih_baik)) +
  geom_col(show.legend = FALSE) +
  scale_x_continuous(breaks = 9:32, expand = c(0, 0)) +
  scale_fill_manual(values = c("TRUE" = "seagreen3", "FALSE" = "salmon")) +
  labs(
    x = NULL,
    y = "Jumlah Kasus",
    title = "Kasus Pekan Positif COVID-19 di DKI Jakarta",
    subtitle = "Kolom hijau menunjukkan penambahan kasus baru lebih sedikit
dibandingkan satu pekan sebelumnya",
    caption = "Sumber data: covid.19.go.id"
  ) +
  theme_ipsum(
    base_size = 9,
    plot_title_size = 18,
    grid = "Y",
    ticks = TRUE
  ) +
  theme(plot.title.position = "plot")
```

Kasus Pekan Positif COVID-19 di DKI Jakarta

Kolom hijau menunjukan penambahan kasus baru lebih sedikit dibandingkan satu pekan sebelumnya



Seventh-Day Amplification Factor

```
#Mengambil data N(0)
library(dplyr)

new_cov_jkt$tanggal <- as.Date(new_cov_jkt$tanggal)
picked_date_N_0_jkt <- filter(new_cov_jkt, tanggal >= "2020-03-01",
                              tanggal <= "2020-07-30")
N0_jkt <- picked_date_N_0_jkt$AKUMULASI_KASUS
```

```
#Mengambil data N(7)
library(dplyr)

new_cov_jkt$tanggal <- as.Date(new_cov_jkt$tanggal)
picked_date_N_7_jkt <- filter(new_cov_jkt, tanggal >= "2020-03-08",
                              tanggal <= "2020-08-06")
N7_jkt <- picked_date_N_7_jkt$AKUMULASI_KASUS
```

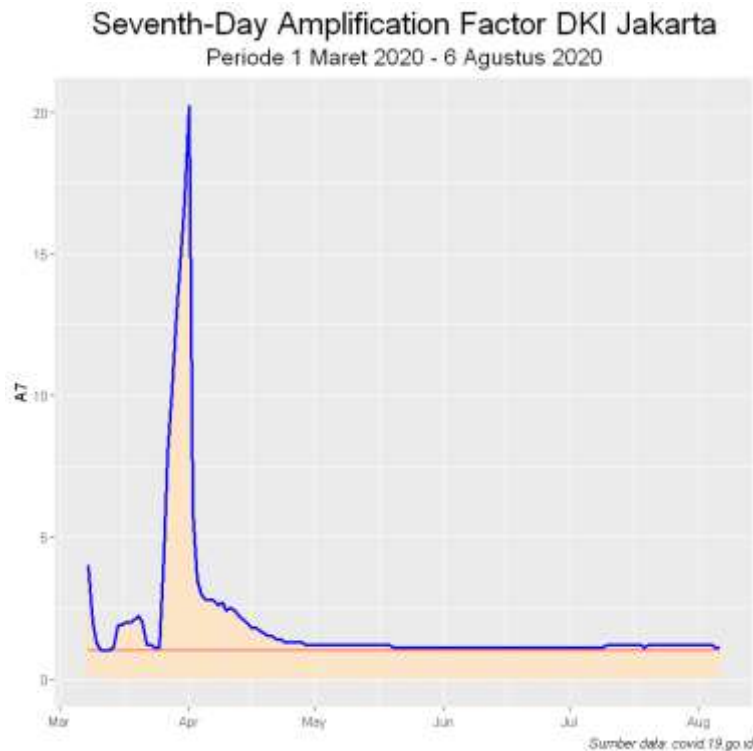
```
#Memasukkan data N0 ke sebuah kolom
picked_date_N_7_jkt$N0_jkt <- (N0_jkt)
#Mengganti kolom "AKUMULASI_KASUS" menjadi "N7"
names(picked_date_N_7_jkt)[names(picked_date_N_7_jkt)
                           == "AKUMULASI_KASUS"] <- "N7_jkt"
#Menghapus kolom yang tidak diperlukan dalam perhitungan A7
select (picked_date_N_7_jkt, -c(meninggal, sembuh, AKUMULASI_SEMBUH,
                              AKUMULASI_MENINGGAL, kasus_baru))
#Menambahkan kolom A7 Factor yang seharusnya, yaitu A7 = 1.0
picked_date_N_7_jkt$A7_normal <- (1.0)
```

```
#Menghitung A7 Factor (N7/N0)
picked_date_N_7_jkt$A7_jkt <- (N7_jkt / N0_jkt)
#Proses pembulatan satu desimal
picked_date_N_7_jkt$A7_jkt <- round(picked_date_N_7_jkt$A7_jkt, digits=1)
```

A data.frame: 152 × 5

tanggal	N7_jkt	N0_jkt	A7_normal	A7_jkt
<date>	<int>	<int>	<dbl>	<dbl>
2020-03-08	8	2	1	4.0
2020-03-09	8	4	1	2.0
2020-03-10	8	6	1	1.3
2020-03-11	8	8	1	1.0
2020-03-12	8	8	1	1.0
2020-03-13	8	8	1	1.0

```
#Menampilkan bar chart disertai dengan line
library(ggplot2)
library(hrbrthemes)
plot.N7_jkt <- ggplot(picked_date_N_7_jkt, aes(x = tanggal, y = A7_jkt)) +
  geom_col(fill="bisque1") +
  labs(
    x = NULL,
    y = "A7",
    title = "Seventh-Day Amplification Factor DKI Jakarta",
    subtitle = "Periode 1 Maret 2020 - 6 Agustus 2020",
    caption = "Sumber data: covid.19.go.id"
  ) + theme(
    plot.title = element_text(hjust = 0.5, size = 21),
    plot.subtitle = element_text(hjust = 0.5, size = 15),
    plot.caption = element_text(hjust = 1, face = "italic")
  )
plot.N7_jkt <- plot.N7_jkt +
  geom_line(aes(x=tanggal, y=A7_normal, fill=A7_normal), lwd=0.5, colour="red")
plot.N7_jkt + geom_line(aes(x=tanggal, y=A7_jkt, fill=A7_jkt), lwd=1,
  colour="blue")
```



Kesimpulan

Data yang kita lakukan analisis dan visualisasi adalah data dari 1 Maret 2020 sampai 6 Agustus 2020. Terdapat perbedaan pada data DKI Jakarta dan Indonesia, yaitu terletak pada tanggal kasus pertama. Kasus pertama di DKI Jakarta teridentifikasi pada tanggal 1 Maret 2020, sedangkan kasus pertama pada data Indonesia baru teridentifikasi pada tanggal 2 Maret 2020 sebanyak 2 pasien positif. Dapat kita simpulkan bahwa data nasional yang diperbarui setiap harinya cenderung terlambat. Hal tersebut karena membutuhkan proses panjang untuk melaporkan jumlah kasus pada tiap provinsi ke nasional. Dari analisis dan visualisasi data COVID-19 di DKI Jakarta, kita dapat simpulkan:

- Data pertama kali di rekam 1 Maret 2020. Dengan jumlah aktif dan positif berjumlah 2, jumlah meninggal dan sembuh berjumlah 0.
- Data terakhir diupdate tanggal 6 Agustus 2020. Penambahan kasus sembuh adalah 246 orang, penambahan kasus meninggal adalah 13 orang, total jumlah kasus positif adalah 23,936 orang, dan total jumlah kasus meninggal 901.
- Per tanggal 6 Agustus 2020, persentase kasus aktif adalah 33.5%, persentase kasus sembuh adalah 62.7%, dan persentase kasus meninggal adalah 3.8%.

- Jumlah penduduk DKI Jakarta mencapai sekitar 10,47 juta jiwa (2018). Berarti pasien positif COVID-19 di DKI Jakarta 6 Agustus 2020 adalah 0,229% dari total jumlah penduduk Provinsi DKI Jakarta.
- Kenaikan total jumlah kasus positif dari tanggal 1 Maret 2020 sampai 6 Agustus 2020 sebesar 1,196,700%
- Dilihat dari grafik komparasi antara akumulasi kasus aktif, sembuh, dan meninggal. Kasus meninggal memang seharusnya berada di bawah garis kasus aktif dan sembuh. Garis kasus meninggal harus ditekan agar tidak terjadi kenaikan lagi. Jika garis akumulasi kasus meninggal sudah mendatar (stabil) dalam waktu yang cukup lama, maka sudah tidak ada penambahan kasus meninggal. Dan jika hal itu terjadi, maka wajib dipertahankan serta diambil beberapa keputusan guna menekan terjadinya kasus meninggal.
- Kasus meninggal di DKI Jakarta juga perlu ditekan karena terbatasnya lahan pemakaman. Kemungkinan pemerintah setempat dapat membebaskan sebagian lahan untuk dijadikan lahan pemakaman apabila jumlah pasien meninggal sudah tidak dapat tertampung.
- Kasus aktif saat ini cenderung naik. Grafik kasus aktif yang semakin naik ini harus ditangani secepatnya. Kasus aktif lebih baik jika menurun dan berada pada titik terendah, yaitu nol. Hal itu menandakan tidak adanya pasien dirawat, positif corona, atau sedang dikarantina. Penekanan kasus aktif dan meninggal sebenarnya dapat ditekan dengan meningkatkan penanganan terhadap pasien positif. Adapun juga kebijakan lainnya guna menurunkan jumlah penambahan kasus.
- DKI Jakarta dapat disebut sebagai pusat ekonomi Indonesia. Jika pemerintah daerah di DKI Jakarta menetapkan PSBB atau *lockdown* yang terlalu ketat, maka ditakutkan pertumbuhan ekonomi Indonesia dapat terhambat. Akhir-akhir ini terdapat isu penyebaran COVID-19 pada kluster perkantoran. Pemerintah harus memikirkan cara terbaik agar dapat menekan kasus aktif COVID-19, tetapi perekonomian terus berkembang agar tidak mengalami resesi. Segala protokol kesehatan perlu diterapkan di seluruh bidang industri secara wajib dan harus dijadikan sebagai suatu kebiasaan yang baru.
- Kasus sembuh menandakan makin naik pada grafik. Hal tersebut sudah bagus dan perlu dipertahankan agar terus naik. Pemerintah seharusnya menyediakan fasilitas kesehatan yang lebih canggih dan lebih banyak dalam penanganan pasien positif, mengingat

jumlah pasien positif semakin banyak dan tenaga kesehatan semakin kewalahan. DKI Jakarta yang merupakan ibu kota, memperoleh bantuan fasilitas kesehatan yang lebih cepat dan lebih banyak dari daerah lain. Seharusnya kesempatan tersebut digunakan secara maksimal dalam penanganan kasus pasien positif COVID-19

- Grafik kasus pekanan COVID-19 di DKI Jakarta diambil dari minggu ke-9 sampai ke-32 di tahun 2020. Terlihat hanya 6 pekan atau hanya 18.75% yang mengalami perbaikan dari pekan sebelumnya. Yang mana sebanyak 81.25% kasus makin parah di masing-masing pekan dari pekan sebelumnya.
- *Seventh-Day Amplification Factor* tertinggi di DKI Jakarta sebesar 20.2 di 1 April 2020. Sejak 20 Juli 2020 sampai 4 Agustus 2020, A7 di DKI Jakarta menandakan kestabilan di nilai pembulatan 1.2. Lalu untuk 5 Agustus 2020 & 6 Agustus 2020, A7 di DKI Jakarta mengalami penurunan ke angka 1.1. Diharapkan A7 di DKI Jakarta dapat stabil atau bahkan mengalami penurunan. Seperti yang kita tahu, cukup sulit untuk mengontrol grafik persebaran COVID-19 di DKI Jakarta karena tingkat mobilitas masyarakat yang tinggi dan DKI Jakarta menduduki peringkat ke-2 dengan kasus positif terbanyak di Indonesia.
- Presiden Indonesia, Joko Widodo melakukan *rapid test* masal COVID-19 pertama kali di wilayah Jakarta Selatan pada 20 Maret 2020 karena banyak warga yang diketahui melakukan kontak dengan pasien positif COVID-19. Lalu mulai sekitar 24 Maret 2020, DKI Jakarta mulai mengadakan *rapid test* masal COVID-19. Jika kita bandingkan dengan grafik *Seventh-Day Amplification Factor* (A7) di DKI Jakarta, A7 mengalami kenaikan mulai dari tanggal 25 Maret 2020 dan mencapai puncak tertinggi di 1 April 2020. Sekarang kita anggap 25 Maret 2020 sebagai $N(0)$ dan berarti 1 April 2020 adalah $N(7)$, dengan hasil A7 di 1 April 2020 sebesar 20.2. Jika kita analisis berdasarkan teori A7, berarti sebelum 25 Maret 2020, dimana *rapid test* baru diadakan 20 Maret 2020, terdapat banyak Orang Tanpa Gejala atau pasien positif yang belum dites masih melakukan mobilisasi di publik. Sehingga data pasien positif yang masuk pada 1 April 2020 sangatlah banyak (karena faktor baru mengadakan *rapid test* masal).

IV. Analisis dan Visualisasi Data di Jawa Barat

Introduction

Proses analisis dan visualisasi data di Jawa Barat memiliki alur pengerjaan yang sama dengan analisis dan visualisasi data di Indonesia dan DKI Jakarta. Hanya saja beberapa *variable*, situs yang diakses, struktur *data frame*, dan lain-lain terdapat sedikit perbedaan. Data yang kita akses sekarang (7 Agustus 2020) terakhir diperbarui tanggal 6 Agustus 2020. Oleh karena itu, selama analisis dan visualisasi data di Jawa Barat saat ini, kita hanya dapat mengakses data dari tanggal 1 Maret 2020 sampai 6 Agustus 2020 saja.

Mengakses API covid.go.id

```
#Read API
library(httr)
resp_jabar <-
GET("https://data.covid19.go.id/public/api/prov_detail_JAWA_BARAT.json")
```

Mengekstrak Isi Respon

```
#Ekstrak konten JSON
cov_jabar_raw <- content(resp_jabar, as = "parsed", simplifyVector = TRUE)

#Komponen cov_jabar_raw
names(cov_jabar_raw)

'last_date' · 'provinsi' · 'kasus_total' · 'kasus_tanpa_tgl' · 'kasus_dengan_tgl' ·
'meninggal_persen' · 'meninggal_tanpa_tgl' · 'meninggal_dengan_tgl' · 'sembuh_persen' ·
'sembuh_tanpa_tgl' · 'sembuh_dengan_tgl' · 'list_perkembangan' · 'data'
```

Analisis Data

Sekarang kita fokus pada obyek `cov_jabar_raw` untuk menjawab pertanyaan-pertanyaan berikut:

- Berapa jumlah total jumlah kasus positif?
- Berapa persentase pasien meninggal?
- Berapa persentase pasien sembuh?
- Berapa persentase pasien aktif?

```
#Analisa COVID-19 di Jawa Barat
cov_jabar_raw$kasus_total
cov_jabar_raw$meninggal_persen
cov_jabar_raw$sembuh_persen
```

```
#Menghitung persentase pasien aktif
cov_jabar_aktif <-
(100 - cov_jabar_raw$sembuh_persen - cov_jabar_raw$meninggal_persen)
cov_jabar_aktif
```

6995

3.10221586847748

60.7862759113653

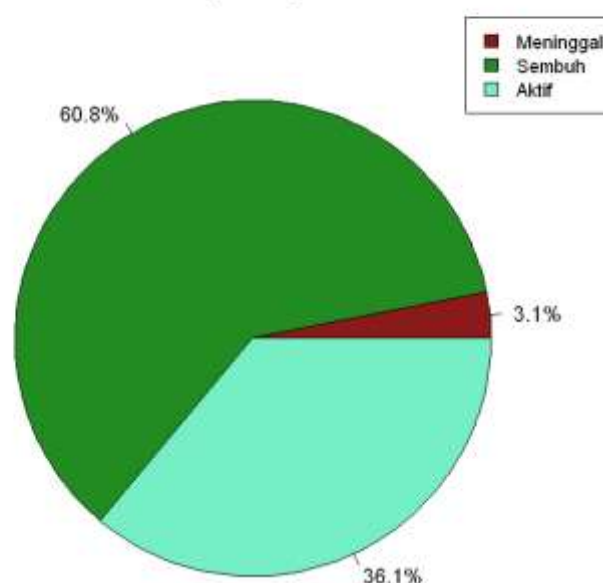
36.1115082201573

Lalu kita juga dapat membuat visualisasi melalui *pie chart* untuk melihat persentase jumlah total kasus meninggal, sembuh, dan aktif di Jawa Barat.

```
#Membuat visualisasi data pasien aktif, sembuh, dan meninggal
x <- c(cov_jabar_raw$meninggal_persen, cov_jabar_raw$sembuh_persen,
      cov_jabar_aktif)
labels <- c("meninggal_persen", "sembuh_persen", "cov_jabar_aktif")
x <- round(x, digits=1)
labels <- paste(x, "%", sep="")

# Plot the chart.
pie(x, labels = labels,
    main = "Persentase Total Pasien Aktif, Sembuh, dan Meninggal di
Jawa Barat per 6 Agustus 2020",
    col = c("firebrick4", "forestgreen", "aquamarine2"))
legend("topright", c("Meninggal", "Sembuh", "Aktif"),
    cex = 0.9,
    fill = c("firebrick4", "forestgreen", "aquamarine2"))
```

Persentase Total Pasien Aktif, Sembuh, dan Meninggal di Jawa Barat per 6 Agustus 2020



Memperoleh Informasi yang Lebih Lengkap

```
#Kita ambil list_perkembangan dan masukkan ke variable cov_jabar
cov_jabar <- cov_jabar_raw$list_perkembangan
str(cov_jabar)
```

```
#Mengambil bagian teratas dari list_perkembangan
head(cov_jabar)
```

```
'data.frame': 159 obs. of 9 variables:
 $ tanggal          : num  1.58e+12 1.58e+12 1.58e+12 1.58e+12 1.58e
+12 ...
 $ KASUS              : int   3 0 1 1 0 1 0 0 0 0 ...
 $ MENINGGAL          : int   0 0 1 0 0 0 0 0 0 0 ...
 $ SEMBUH             : int   0 0 0 0 0 0 0 0 0 0 ...
 $ DIRAWAT_OR_ISOLASI : int   3 0 0 1 0 1 0 0 0 0 ...
 $ AKUMULASI_KASUS    : int   3 3 4 5 5 6 6 6 6 6 ...
 $ AKUMULASI_SEMBUH   : int   0 0 0 0 0 0 0 0 0 0 ...
 $ AKUMULASI_MENINGGAL : int   0 0 1 1 1 1 1 1 1 1 ...
 $ AKUMULASI_DIRAWAT_OR_ISOLASI : int  3 3 3 4 4 5 5 5 5 ...
```

A data.frame: 6 × 9

	tanggal	KASUS	MENINGGAL	SEMBUH	DIRAWAT_OR_ISOLASI	AKUMULASI_KASUS	AKUMULASI_MENINGGAL	AKUMULASI_SEMBUH
	<dbl>	<int>	<int>	<int>	<int>	<int>	<int>	<int>
1	1.583021e+12	2	0	0	2	2	0	0
2	1.583107e+12	2	0	0	2	4	0	0
3	1.583194e+12	2	0	0	2	6	0	0
4	1.583280e+12	2	0	0	2	8	0	0
5	1.583366e+12	0	1	0	-1	8	1	0
6	1.583453e+12	0	0	0	0	8	1	0

Menjinakkan Data

Setelah mengekstrak dan mengamati `cov_jabar`, kita menemukan beberapa kejanggalan pada data tersebut. Diantaranya adalah kejanggalan data pada kolom tanggal dan format penulisan kolom yang tidak konsisten. Sekarang kita akan mencoba melakukan beberapa tahapan untuk menjinakkan data tersebut sehingga dapat diolah dan dianalisis dengan lebih mudah.

Sebelum itu, kita aktifkan paket `dplyr` yang akan dipergunakan untuk melakukan pengolahan data.

Ada beberapa tahapan yang akan kita lakukan untuk menjinakkan data `cov_jabar`, yaitu:

6. Menghapus kolom “DIRAWAT_OR_ISOLASI” dan “AKUMULASI_DIRAWAT_OR_ISOLASI”

7. Menghapus semua kolom yang berisi nilai kumulatif
8. Mengganti nama kolom "KASUS" menjadi "kasus_baru"
9. Merubah format penulisan kolom berikut menjadi huruf kecil
 - kolom MENINGGAL
 - kolom SEMBUH
10. Memperbaiki data pada kolom tanggal

Kita akan menggunakan operator *pipe* (`%>%`) untuk merangkai fungsi menjadi sebuah *pipeline*. Lalu simpan hasil pengolahan dengan nama `new_cov_jkt`.

```
#Menjinakkan data dan merapikan kolom tanggal yang janggal
library(dplyr)
new_cov_jabar <-
  cov_jabar %>%
  select(-contains("DIRAWAT_OR_ISOLASI")) %>%
  rename(
    kasus_baru = KASUS,
    meninggal = MENINGGAL,
    sembuh = SEMBUH
  ) %>%
  mutate(
    tanggal = as.POSIXct(tanggal / 1000, origin = "1970-01-01"),
    tanggal = as.Date(tanggal)
  )
str(new_cov_jabar)
```

```
'data.frame': 159 obs. of 7 variables:
 $ tanggal      : Date, format: "2020-03-01" "2020-03-02" ...
 $ kasus_baru    : int  3 0 1 1 0 1 0 0 0 0 ...
 $ meninggal     : int  0 0 1 0 0 0 0 0 0 0 ...
 $ sembuh       : int  0 0 0 0 0 0 0 0 0 0 ...
 $ AKUMULASI_KASUS : int  3 3 4 5 5 6 6 6 6 6 ...
 $ AKUMULASI_SEMBUH : int  0 0 0 0 0 0 0 0 0 0 ...
 $ AKUMULASI_MENINGGAL: int  0 0 1 1 1 1 1 1 1 1 ...
```

Visualisasi Data Dengan Grafik

```
#Menunjukkan melalui gambar grafik positif beserta beberapa keterangan tambahan
library(ggplot2)
library(hrbrthemes)
ggplot(new_cov_jabar, aes(tanggal, kasus_baru)) +
  geom_col(fill = "darkturquoise") +
  labs(
    x = NULL,
    y = "Jumlah Kasus",
    title = "Kasus Harian Positif COVID-19 di Jawa Barat",
    subtitle = "Terjadi pelonjakan kasus di awal bulan Juli akibat kluster Secapa AD Bandung",
    caption = "Sumber data: covid.19.go.id"
  ) + theme(
    plot.title = element_text(hjust = 0, size = 21, face = "bold"),
    plot.subtitle = element_text(hjust = 0, size = 12),
  )
```

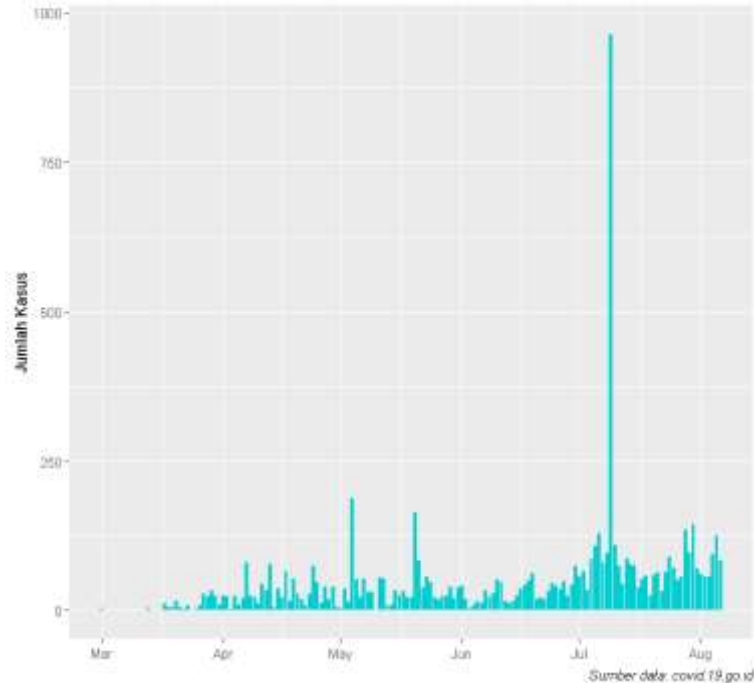
```

    plot.caption = element_text(hjust = 1, face = "italic"),
  ) +
  theme(plot.title.position = "plot")

```

Kasus Harian Positif COVID-19 di Jawa Barat

Terjadi pelonjakan kasus di awal bulan Juli akibat kluster Secapa AD Bandung

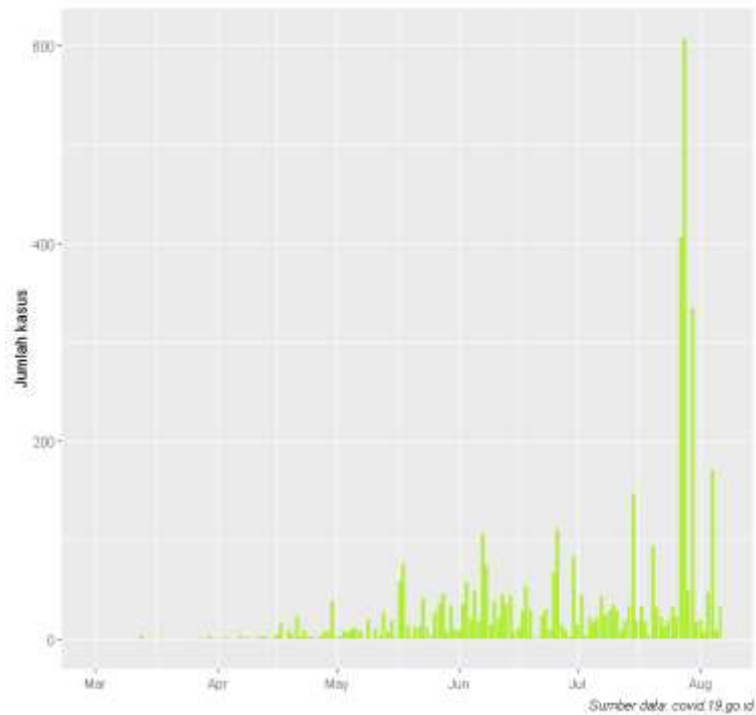


```

#Grafik untuk kasus sembuh
library(ggplot2)
library(hrbrthemes)
ggplot(new_cov_jabar, aes(tanggal, sembuh)) +
  geom_col(fill = "olivedrab2") +
  labs(
    x = NULL,
    y = "Jumlah kasus",
    title = "Kasus Harian Sembuh Dari COVID-19 di Jawa Barat",
    caption = "Sumber data: covid.19.go.id"
  ) + theme(
    plot.title = element_text(hjust = 0, size = 20, face = "bold"),
    plot.subtitle = element_text(hjust = 0, size = 12),
    plot.caption = element_text(hjust = 1, face = "italic"),
  ) +
  theme(plot.title.position = "plot")

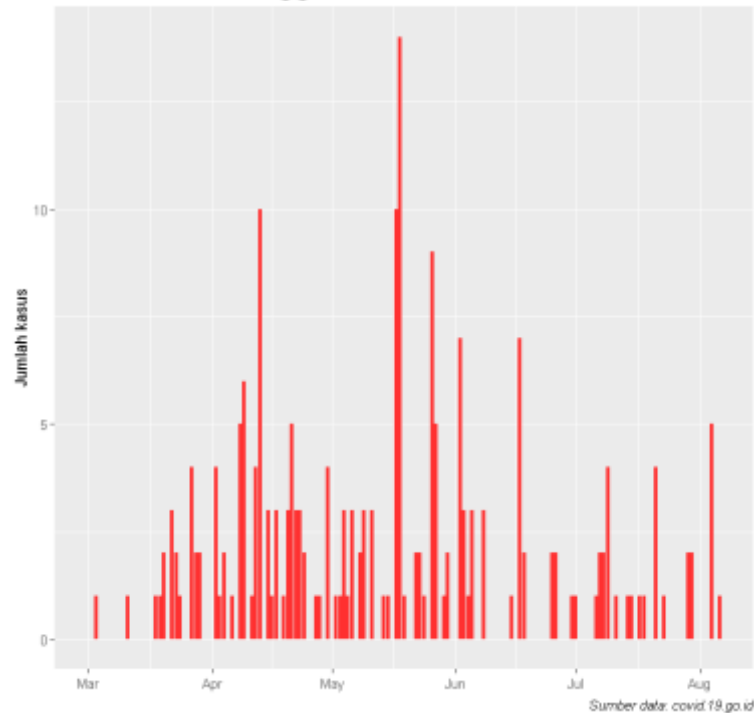
```

Kasus Harian Sembuh Dari COVID-19 di Jawa Barat



```
#Grafik untuk kasus meninggal
library(ggplot2)
library(hrbrthemes)
ggplot(new_cov_jabar, aes(tanggal, meninggal)) +
  geom_col(fill = "firebrick1") +
  labs(
    x = NULL,
    y = "Jumlah kasus",
    title = "Kasus Harian Meninggal Akibat COVID-19 di Jawa Barat",
    caption = "Sumber data: covid.19.go.id"
  ) + theme(
    plot.title = element_text(hjust = 0, size = 18.5, face = "bold"),
    plot.subtitle = element_text(hjust = 0, size = 12),
    plot.caption = element_text(hjust = 1, face = "italic"),
  ) +
  theme(plot.title.position = "plot")
```

Kasus Harian Meninggal Akibat COVID-19 di Jawa Barat



Pola dan Dinamika

```
#Menghitung nilai akumulasi dari suatu vektor numerik
library(dplyr)
cov_jabar_akumulasi <-
  new_cov_jabar %>%
  transmute(
    tanggal,
    akumulasi_aktif = cumsum(kasus_baru) - cumsum(sembuh) - cumsum(meninggal),
    akumulasi_sembuh = cumsum(sembuh),
    akumulasi_meninggal = cumsum(meninggal)
  )
tail(cov_jabar_akumulasi)
```

A data.frame: 6 × 4

	tanggal	akumulasi_aktif	akumulasi_sembuh	akumulasi_meninggal
	<date>	<int>	<int>	<int>
154	2020-08-01	2384	3981	206
155	2020-08-02	2434	3988	206
156	2020-08-03	2443	4035	206
157	2020-08-04	2361	4206	211
158	2020-08-05	2477	4215	211
159	2020-08-06	2527	4247	212

Grafik Komparasi Data

```
#Melakukan pivot, transformasi data dari wide ke Long
library(dplyr)
library(tidyr)

dim(cov_jabar_akumulasi)

cov_jabar_akumulasi_pivot <-
  cov_jabar_akumulasi %>%
  gather(
    key = "kategori",
    value = "jumlah",
    -tanggal
  ) %>%
  mutate(
    kategori = sub(pattern = "akumulasi_", replacement = "", kategori)
  )

dim(cov_jabar_akumulasi_pivot)

glimpse(cov_jabar_akumulasi_pivot)

159 · 4

477 · 3

Rows: 477
Columns: 3
$ tanggal <date> 2020-03-01, 2020-03-02, 2020-03-03, 2020-03-04, 2020-03-0...
$ kategori <chr> "aktif", "aktif", "aktif", "aktif", "aktif", "aktif", "akt...
$ jumlah <int> 3, 3, 3, 4, 4, 5, 5, 5, 5, 5, 5, 6, 6, 6, 6, 7, 18, 21, 27...
```

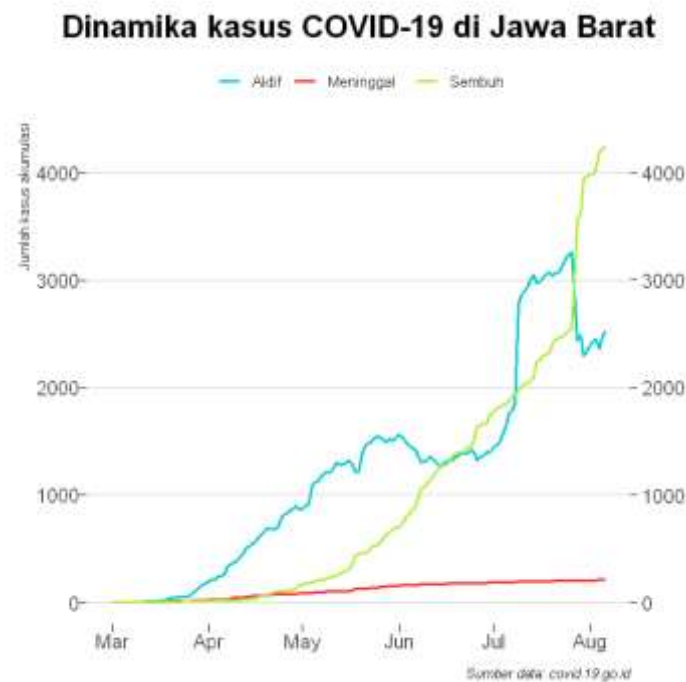
```
#Membuat komparasi antara kasus aktif, sembuh, dan meninggal
library(ggplot2)
library(hrbrthemes)
ggplot(cov_jabar_akumulasi_pivot, aes(tanggal, jumlah, colour = (kategori))) +
  geom_line(size = 0.9) +
  scale_y_continuous(sec.axis = dup_axis(name = NULL)) +
  scale_colour_manual(
    values = c(
      "aktif" = "darkturquoise",
      "meninggal" = "firebrick1",
      "sembuh" = "olivedrab2"
    ),
    labels = c("Aktif", "Meninggal", "Sembuh")
  ) +
  labs(
    x = NULL,
    y = "Jumlah kasus akumulasi",
    colour = NULL,
    title = "Dinamika kasus COVID-19 di Jawa Barat",
    caption = "Sumber data: covid.19.go.id"
```



```

) +
theme_ipsum(
  base_size = 13,
  plot_title_size = 21,
  grid = "Y",
  ticks = TRUE
) +
theme(
  plot.title = element_text(hjust = 0.5),
  legend.position = "top"
)

```



Grafik Perkembangan Kasus Pekan

#Mengamati kasus pekan dan melakukan inspeksi data

```

library(dplyr)
library(lubridate)

cov_jabar_pekanan <- new_cov_jabar %>%
  count(
    tahun = year(tanggal),
    pekan_ke = week(tanggal),
    wt = kasus_baru,
    name = "jumlah"
  )

glimpse(cov_jabar_pekanan)

```

```

Rows: 24
Columns: 3
$ tahun    <dbl> 2020, 2020, 2020, 2020, 2020, 2020, 2020, 2020, 2020, 2020...
$ pekan_ke <dbl> 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23,...
$ jumlah   <int> 4, 2, 18, 45, 129, 185, 217, 238, 227, 338, 243, 151, 427,...

```

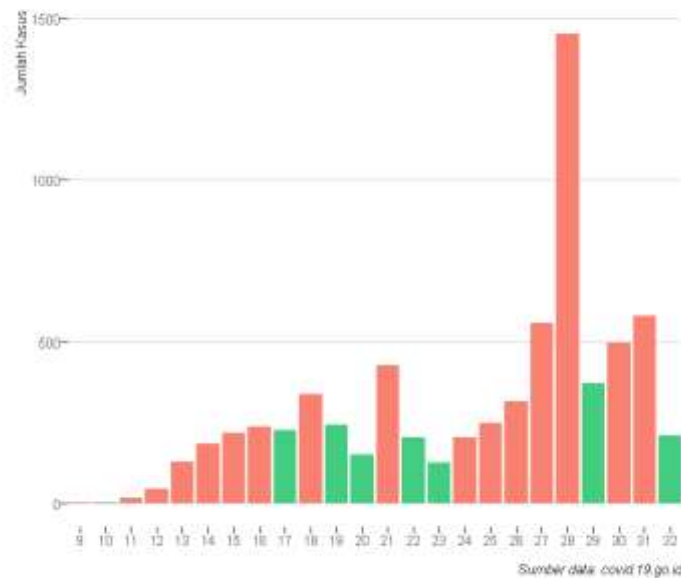
```
#Apakah pekan ini Lebih baik dari pekan kemarin?
library(dplyr)
cov_jabar_pekanan <-
  cov_jabar_pekanan %>%
  mutate(
    jumlah_pekanlalu = dplyr::lag(jumlah, 1),
    jumlah_pekanlalu = ifelse(is.na(jumlah_pekanlalu), 0, jumlah_pekanlalu),
    lebih_baik = jumlah < jumlah_pekanlalu
  )
glimpse(cov_jabar_pekanan)

Rows: 24
Columns: 5
$ tahun      <dbl> 2020, 2020, 2020, 2020, 2020, 2020, 2020, 2020, 20...
$ pekan_ke   <dbl> 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21,...
$ jumlah     <int> 4, 2, 18, 45, 129, 185, 217, 238, 227, 338, 243, 1...
$ jumlah_pekanlalu <dbl> 0, 4, 2, 18, 45, 129, 185, 217, 238, 227, 338, 243...
$ lebih_baik <lgl> FALSE, TRUE, FALSE, FALSE, FALSE, FALSE, FALSE, FA...
```

```
#Membuat bar chart yang menunjukkan kondisi kasus pekanan
library(ggplot2)
library(hrbrthemes)
ggplot(cov_jabar_pekanan, aes(pekan_ke, jumlah, fill = lebih_baik)) +
  geom_col(show.legend = FALSE) +
  scale_x_continuous(breaks = 9:32, expand = c(0, 0)) +
  scale_fill_manual(values = c("TRUE" = "seagreen3", "FALSE" = "salmon")) +
  labs(
    x = NULL,
    y = "Jumlah Kasus",
    title = "Kasus Pekan Positif COVID-19 di Jawa Barat",
    subtitle = "Kolom hijau menunjukan penambahan kasus baru lebih sedikit
dibandingkan satu pekan sebelumnya",
    caption = "Sumber data: covid.19.go.id"
  ) +
  theme_ipsum(
    base_size = 9,
    plot_title_size = 18,
    grid = "Y",
    ticks = TRUE
  ) +
  theme(plot.title.position = "plot")
```

Kasus Pekan Positif COVID-19 di Jawa Barat

Kolom hijau menunjukan penambahan kasus baru lebih sedikit dibandingkan satu pekan sebelumnya



Seventh-Day Amplification Factor

```
#Mengambil data N(0)
library(dplyr)

new_cov_jabar$tanggal <- as.Date(new_cov_jabar$tanggal)
picked_date_N_0_jabar <-
filter(new_cov_jabar, tanggal >= "2020-03-01", tanggal <= "2020-07-30")
N0_jabar <- picked_date_N_0_jabar$AKUMULASI_KASUS
```

```
#Mengambil data N(7)
library(dplyr)

new_cov_jabar$tanggal <- as.Date(new_cov_jabar$tanggal)
picked_date_N_7_jabar <-
filter(new_cov_jabar, tanggal >= "2020-03-08", tanggal <= "2020-08-06")
N7_jabar <- picked_date_N_7_jabar$AKUMULASI_KASUS
```

```
#Memasukkan data N0 ke sebuah kolom
picked_date_N_7_jabar$N0_jabar <- (N0_jabar)
#Mengganti kolom "AKUMULASI_KASUS" menjadi "N7"
names(picked_date_N_7_jabar)[names(picked_date_N_7_jabar)
                              == "AKUMULASI_KASUS"] <- "N7_jabar"
#Menghapus kolom yang tidak diperlukan dalam perhitungan A7
select (picked_date_N_7_jabar, -c(meninggal, sembuh, AKUMULASI_SEMBUH,
                                  AKUMULASI_MENINGGAL, kasus_baru))
#Menambahkan kolom A7 Factor yang seharusnya, yaitu A7 = 1.0
picked_date_N_7_jabar$A7_normal <- (1.0)
#Menghitung A7 Factor (N7/N0)
picked_date_N_7_jabar$A7_jabar <- (N7_jabar / N0_jabar)
```

```
#Proses pembulatan satu desimal
```

```
picked_date_N_7_jabar$A7_jabar <- round(picked_date_N_7_jabar$A7_jabar,  
                                          digits=1)
```

A data.frame: 152 × 5

tanggal	N7_jabar	N0_jabar	A7_normal	A7_jabar
<date>	<int>	<int>	<dbl>	<dbl>
2020-03-08	6	3	1	2.0
2020-03-09	6	3	1	2.0
2020-03-10	6	4	1	1.5
2020-03-11	7	5	1	1.4
2020-03-12	8	5	1	1.6
2020-03-13	12	6	1	2.0

```
#Menampilkan bar chart disertai dengan line
```

```
library(ggplot2)
```

```
library(hrbrthemes)
```

```
plot.N7_jabar <- ggplot(picked_date_N_7_jabar, aes(x = tanggal, y = A7_jabar)) +  
  geom_col(fill="bisque1") +
```

```
  labs(
```

```
    x = NULL,
```

```
    y = "A7",
```

```
    title = "Seventh-Day Amplification Factor Jawa Barat",
```

```
    subtitle = "Periode 1 Maret 2020 - 6 Agustus 2020",
```

```
    caption = "Sumber data: covid.19.go.id"
```

```
  ) + theme(
```

```
    plot.title = element_text(hjust = 0.5, size = 21),
```

```
    plot.subtitle = element_text(hjust = 0.5, size = 15),
```

```
    plot.caption = element_text(hjust = 1, face = "italic")
```

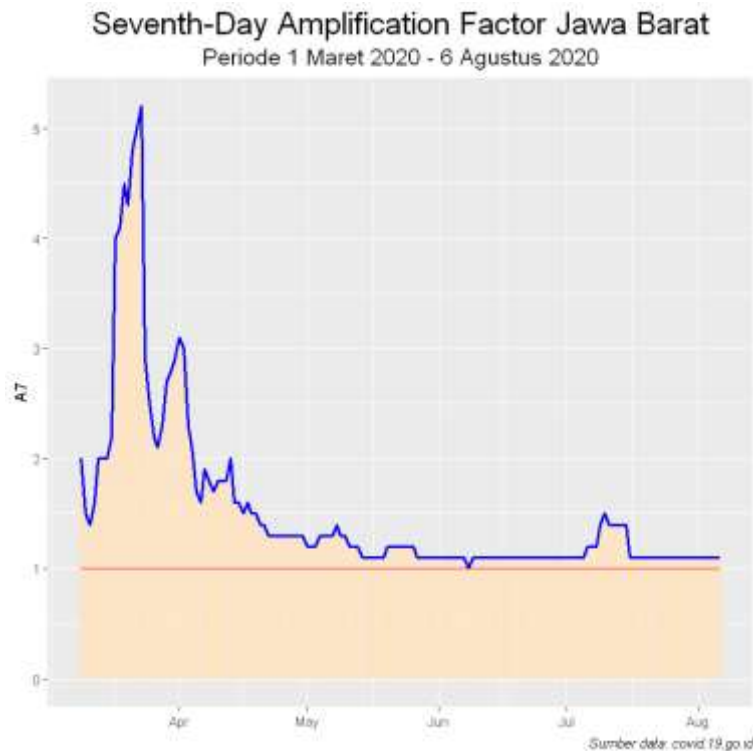
```
  )
```

```
plot.N7_jabar <- plot.N7_jabar +
```

```
geom_line(aes(x=tanggal, y=A7_normal, fill=A7_normal), lwd=0.5, colour="red")
```

```
plot.N7_jabar +
```

```
geom_line(aes(x=tanggal, y=A7_jabar, fill=A7_jabar), lwd=1, colour="blue")
```



Kesimpulan

Data yang kita lakukan analisis dan visualisasi adalah data dari 1 Maret 2020 sampai 6 Agustus 2020. Pada 1 Maret 2020, DKI Jakarta mengonfirmasi akan adanya kasus pertama sejumlah 2 orang positif. Sedangkan pada tanggal yang sama, Jawa Barat juga mengonfirmasi adanya kasus positif pertama sejumlah 3 orang. Namun, data kasus positif pertama yang diumumkan oleh nasional pada 2 Maret 2020 hanya sejumlah 2 orang positif. Hal ini menunjukkan adanya kemungkinan bahwa data dari Jawa Barat belum masuk atau diterima nasional pada 2 Maret 2020 (keterlambatan penerimaan atau *update* data). Dari analisis dan visualisasi data COVID-19 di Jawa Barat, kita dapat simpulkan:

- Data pertama kali di rekam 1 Maret 2020. Dengan jumlah aktif dan positif berjumlah 3 orang, jumlah meninggal dan sembuh berjumlah 0.
- Data terakhir diupdate tanggal 6 Agustus 2020. Penambahan kasus sembuh adalah 32 orang, penambahan kasus meninggal adalah 1 orang, total jumlah kasus positif adalah 6,995 orang, dan total jumlah kasus meninggal 217.
- Per tanggal 6 Agustus 2020, persentase kasus aktif adalah 36.1%, persentase kasus sembuh adalah 60.8%, dan persentase kasus meninggal adalah 3.1%.

- Jumlah penduduk di Jawa Barat sejumlah 48,68 juta jiwa (2018). Berarti pasien positif COVID-19 6 Agustus 2020 adalah 0.014% dari total jumlah penduduk di Provinsi Jawa Barat.
- Kenaikan total jumlah kasus positif dari tanggal 1 Maret 2020 sampai 6 Agustus 2020 sebesar 233,067%
- Dilihat dari grafik komparasi antara akumulasi kasus aktif, sembuh, dan meninggal. Kasus meninggal memang seharusnya berada di bawah garis kasus aktif dan sembuh. Garis kasus meninggal harus ditekan agar tidak terjadi kenaikan lagi. Jika garis akumulasi kasus meninggal sudah mendatar (stabil) dalam waktu yang cukup lama, maka sudah tidak ada penambahan kasus meninggal. Dan jika hal itu terjadi, maka wajib dipertahankan serta diambil beberapa keputusan guna menekan terjadinya kasus meninggal.
- Kasus meninggal di Jawa Barat lebih baik daripada yang ada di DKI Jakarta. Namun, tetap penanganan kasus pasien positif terus ditingkatkan guna menekan grafik kasus meninggal.
- Kasus aktif saat ini naik dan turun. Grafik kasus aktif sempat mengalami penurunan, namun sekarang cenderung naik kembali. Diharapkan grafik kasus aktif tidak mengalami kenaikan secara berkelanjutan Kasus aktif lebih baik jika menurun dan berada pada titik terendah, yaitu nol. Hal itu menandakan tidak adanya pasien dirawat, positif corona, atau sedang dikarantina. Penekanan kasus aktif dan meninggal sebenarnya dapat ditekan dengan meningkatkan penanganan terhadap pasien positif. Adapun juga kebijakan lainnya guna menurunkan jumlah penambahan kasus.
- Kasus positif mengalami pelonjakan pada awal bulan Juli akibat klaster Secapa AD Bandung. Per tanggal 16 Juli 2020, tercatat jumlah pasien positif di Secapa AD mencapai 1,307 orang. Hingga 7 Agustus 2020 92,9% pasien corona di fasilitas pendidikan di Kota Bandung, Jawa Barat tersebut dinyatakan pulih.
- Kasus sembuh menandakan makin naik pada grafik. Hal tersebut sudah bagus dan perlu dipertahankan agar terus naik. Pemerintah seharusnya menyediakan fasilitas kesehatan yang lebih canggih dan lebih banyak dalam penanganan pasien positif, mengingat jumlah pasien positif semakin banyak dan tenaga kesehatan semakin kewalahan. Pemerintah pusat juga dapat menyalurkan bantuan berupa fasilitas dan tenaga kesehatan dari pusat ke daerah. Jangan sampai terdapat ketidakmerataan akses fasilitas pendukung perawatan pasien positif COVID-19

- Grafik kasus pekanan COVID-19 di Jawa Barat diambil dari minggu ke-9 sampai ke-32 di tahun 2020. Terlihat hanya 7 pekan atau hanya 21,88% yang mengalami perbaikan dari pekan sebelumnya. Yang mana sebanyak 78.13% kasus makin parah di masing-masing pekan dari pekan sebelumnya.
- *Seventh-Day Amplification Factor* tertinggi di Jawa Barat sebesar 5.2 di 23 Maret 2020. Sejak 16 Juli 2020 sampai 6 Agustus 2020, A7 di Jawa Barat menandakan kestabilan di nilai pembulatan 1.1. Kini Jawa Barat menempati peringkat ke-5 dalam penyebaran kasus COVID-19 tertinggi di Indonesia.

V. Komparasi Provinsi DKI Jakarta dan Jawa Barat

Introduction

Setelah melakukan analisis dan visualisasi data di DKI Jakarta dan Jawa Barat, maka kita akan melakukan perbandingan terhadap akumulasi kasus pasien aktif, sembuh, dan meninggal dari 1 Maret 2020 sampai 6 Agustus 2020. Dan juga akan melakukan komparasi A7 yang dapat dihitung dari tanggal 8 Maret 2020 sampai 6 Agustus 2020. Manakah provinsi yang memiliki kasus lebih tinggi?

Pola dan Dinamika

Kita akan menghitung nilai akumulasi dari vektor numerik untuk kasus aktif, sembuh, dan meninggal. Lalu kita masukkan ke *variable* baru yaitu, `cov_perbandingan_aktif`, `cov_perbandingan_sembuh`, dan `cov_perbandingan_meninggal`.

```
#Menghitung nilai akumulasi dari suatu vektor numerik
library(dplyr)
cov_perbandingan_aktif <-
  new_cov_jkt %>%
  transmute(
    tanggal,
    akumulasi_aktif_jkt =
      cumsum(kasus_baru) - cumsum(sembuh) - cumsum(meninggal),
    akumulasi_aktif_jabar =
      cumsum(new_cov_jabar$kasus_baru) - cumsum(new_cov_jabar$sembuh) -
      cumsum(new_cov_jabar$meninggal)
  )

tail(cov_perbandingan_aktif)
```

A data.frame: 6 × 3

	tanggal	akumulasi_aktif_jkt	akumulasi_aktif_jabar
	<date>	<int>	<int>
154	2020-08-01	6612	2384
155	2020-08-02	6848	2434
156	2020-08-03	7170	2443
157	2020-08-04	7344	2361
158	2020-08-05	7305	2477
159	2020-08-06	7601	2527


```
#Menghitung nilai akumulasi dari suatu vektor numerik
library(dplyr)
cov_perbandingan_sembuh <-
  new_cov_jkt %>%
  transmute(
    tanggal,
    akumulasi_sembuh_jkt= cumsum(sembuh),
    akumulasi_sembuh_jabar = cumsum(new_cov_jabar$sembuh)
  )

tail(cov_perbandingan_sembuh)
```

A data.frame: 6 × 3

	tanggal	akumulasi_sembuh_jkt	akumulasi_sembuh_jabar
	<date>	<int>	<int>
154	2020-08-01	13844	3981
155	2020-08-02	13985	3988
156	2020-08-03	14119	4035
157	2020-08-04	14341	4206
158	2020-08-05	14720	4215
159	2020-08-06	14967	4247

```
#Menghitung nilai akumulasi dari suatu vektor numerik
library(dplyr)
cov_perbandingan_meninggal <-
  new_cov_jkt %>%
  transmute(
    tanggal,
    akumulasi_meninggal_jkt= cumsum(meninggal),
    akumulasi_meninggal_jabar = cumsum(new_cov_jabar$meninggal)
  )

tail(cov_perbandingan_meninggal)
```

A data.frame: 6 × 3

	tanggal	akumulasi_meninggal_jkt	akumulasi_meninggal_jabar
	<date>	<int>	<int>
154	2020-08-01	824	206
155	2020-08-02	824	206
156	2020-08-03	840	206
157	2020-08-04	854	211
158	2020-08-05	868	211
159	2020-08-06	881	212

Kita juga dapat mengambil nilai dari vektor numerik. Untuk kasus *Seventh-Day Amplification Factor*, kita akan mengambil `A7_jabar` dan digabungkan dengan `A7_jkt` dari `picked_date_N_7_jkt`. Selanjutnya akan disimpan dalam *variable* `cov_perbandingan_A7`.

```
#Mengambil nilai dari suatu vektor numerik
library(dplyr)

cov_perbandingan_A7 <-
  picked_date_N_7_jkt %>%
  transmute(
    tanggal,
    A7_jkt = A7_jkt,
    A7_jabar = picked_date_N_7_jabar$A7_jabar
  )

tail (cov_perbandingan_A7)
```

A data.frame: 6 × 3

	tanggal	A7_jkt	A7_jabar
	<date>	<dbl>	<dbl>
147	2020-08-01	1.2	1.1
148	2020-08-02	1.2	1.1
149	2020-08-03	1.2	1.1
150	2020-08-04	1.2	1.1
151	2020-08-05	1.1	1.1
152	2020-08-06	1.1	1.1

Grafik Komparasi Data

Seperti pada praktik sebelum-sebelumnya, kita akan melakukan *pivot*, transformasi data dari *wide* ke *long* terlebih dahulu. Setelah itu, akan dibuat grafik komparasi untuk kedua provinsi tersebut.

```
#Melakukan pivot, transformasi data dari wide ke Long
library(dplyr)
library(tidyr)

dim(cov_perbandingan_aktif)

cov_perbandingan_aktif_pivot <-
  cov_perbandingan_aktif %>%
  gather(
    key = "kategori",
    value = "jumlah",
    -tanggal
```

```

) %>%
mutate(
  kategori = sub(pattern = "akumulasi_", replacement = "", kategori)
)

dim(cov_perbandingan_aktif_pivot)

glimpse(cov_perbandingan_aktif_pivot)

```

159 · 3

318 · 3

```

Rows: 318
Columns: 3
$ tanggal <date> 2020-03-01, 2020-03-02, 2020-03-03, 2020-03-04, 2020-03-0...
$ kategori <chr> "aktif_jkt", "aktif_jkt", "aktif_jkt", "aktif_jkt", "aktif...
$ jumlah <int> 2, 4, 6, 8, 7, 7, 5, 5, 4, 4, 4, -1, -5, -6, -3, -6, -6, -...

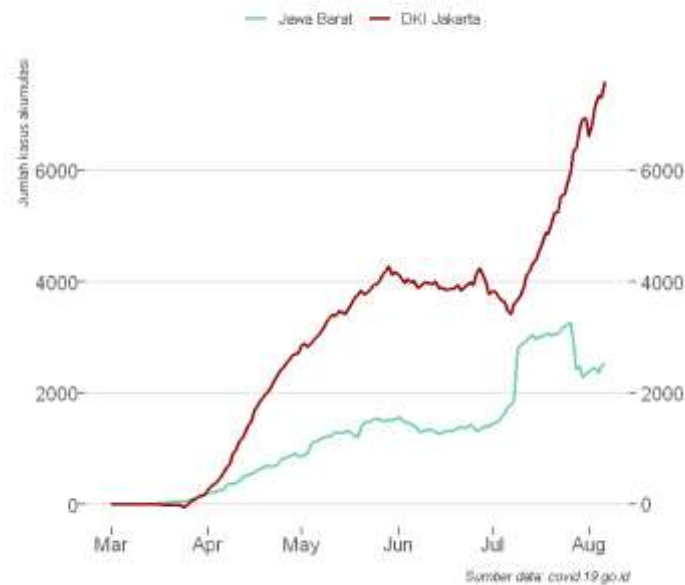
```

```

#Membuat komparasi antara kasus aktif di DKI Jakarta & Jawa Barat
library(ggplot2)
library(hrbrthemes)
ggplot(cov_perbandingan_aktif_pivot, aes(tanggal, jumlah, colour = (kategori)))
  geom_line(size = 0.9) +
  scale_y_continuous(sec.axis = dup_axis(name = NULL)) +
  scale_colour_manual(
    values = c(
      "aktif_jkt" = "darkred",
      "aktif_jabar" = "aquamarine3"
    ),
    labels = c("Jawa Barat", "DKI Jakarta")
  ) +
  labs(
    x = NULL,
    y = "Jumlah kasus akumulasi",
    colour = NULL,
    title = "Dinamika Perbandingan Kasus Aktif
di DKI Jakarta dan Jawa Barat",
    caption = "Sumber data: covid.19.go.id"
  ) +
  theme_ipsum(
    base_size = 13,
    plot_title_size = 21,
    grid = "Y",
    ticks = TRUE
  ) +
  theme(
    plot.title = element_text(hjust = 0.5),
    legend.position = "top"
  )

```

Dinamika Perbandingan Kasus Aktif di DKI Jakarta dan Jawa Barat



```
#Melakukan pivot, transformasi data dari wide ke Long
library(dplyr)
library(tidyr)

dim(cov_perbandingan_sembuh)

cov_perbandingan_sembuh_pivot <-
  cov_perbandingan_sembuh %>%
  gather(
    key = "kategori",
    value = "jumlah",
    -tanggal
  ) %>%
  mutate(
    kategori = sub(pattern = "akumulasi_", replacement = "", kategori)
  )

dim(cov_perbandingan_sembuh_pivot)

glimpse(cov_perbandingan_sembuh_pivot)
```

159 · 3

318 · 3

```
Rows: 318
Columns: 3
$ tanggal <date> 2020-03-01, 2020-03-02, 2020-03-03, 2020-03-04, 2020-03-0...
$ kategori <chr> "sembuh_jkt", "sembuh_jkt", "sembuh_jkt", "sembuh_jkt", "s...
$ jumlah <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 4, 7, 8, 8, 8, 8, 8, 9, 1...
```

```

#Membuat komparasi antara kasus sembuh di DKI Jakarta & Jawa Barat
library(ggplot2)
library(hrbrthemes)
ggplot(cov_perbandingan_sembuh_pivot, aes(tanggal, jumlah, colour = (kategori)))
  geom_line(size = 0.9) +
  scale_y_continuous(sec.axis = dup_axis(name = NULL)) +
  scale_colour_manual(
    values = c(
      "sembuh_jkt" = "darkred",
      "sembuh_jabar" = "aquamarine3"
    ),
    labels = c("Jawa Barat", "DKI Jakarta")
  ) +
  labs(
    x = NULL,
    y = "Jumlah kasus akumulasi",
    colour = NULL,
    title = "Dinamika Perbandingan Kasus Sembuh
di DKI Jakarta dan Jawa Barat",
    caption = "Sumber data: covid.19.go.id"
  ) +
  theme_ipsum(
    base_size = 13,
    plot_title_size = 21,
    grid = "Y",
    ticks = TRUE
  ) +
  theme(
    plot.title = element_text(hjust = 0.5),
    legend.position = "top"
  )

```



```
#Melakukan pivot, transformasi data dari wide ke Long
library(dplyr)
library(tidyr)

dim(cov_perbandingan_meninggal)

cov_perbandingan_meninggal_pivot <-
  cov_perbandingan_meninggal %>%
  gather(
    key = "kategori",
    value = "jumlah",
    -tanggal
  ) %>%
  mutate(
    kategori = sub(pattern = "akumulasi_", replacement = "", kategori)
  )

dim(cov_perbandingan_meninggal_pivot)

glimpse(cov_perbandingan_meninggal_pivot)
```

```
159 · 3
```

```
318 · 3
```

```
Rows: 318
```

```
Columns: 3
```

```
$ tanggal <date> 2020-03-01, 2020-03-02, 2020-03-03, 2020-03-04, 2020-03-0...
```

```
$ kategori <chr> "meninggal_jkt", "meninggal_jkt", "meninggal_jkt", "mening...
```

```
$ jumlah <int> 0, 0, 0, 0, 1, 1, 3, 3, 4, 4, 4, 5, 6, 7, 10, 13, 14, 15, ...
```

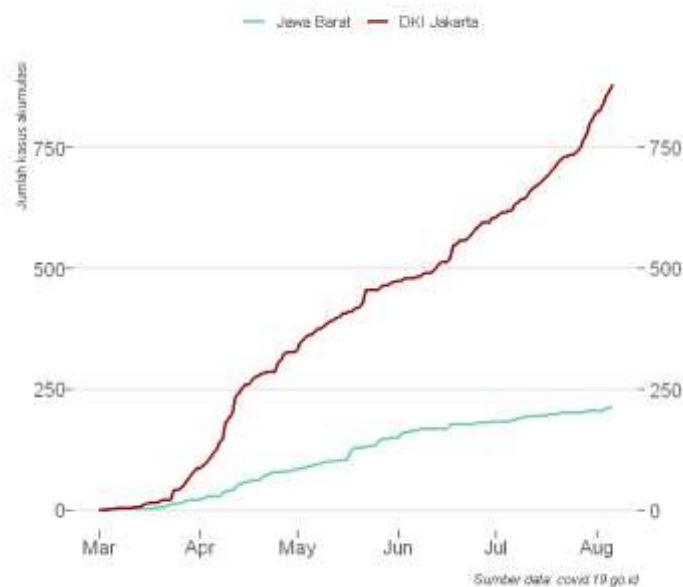
```
#Membuat komparasi antara kasus sembuh di DKI Jakarta & Jawa Barat
library(ggplot2)
library(hrbrthemes)
ggplot(cov_perbandingan_meninggal_pivot, aes(tanggal, jumlah, colour = (kategori
  geom_line(size = 0.9) +
  scale_y_continuous(sec.axis = dup_axis(name = NULL)) +
  scale_colour_manual(
    values = c(
      "meninggal_jkt" = "darkred",
      "meninggal_jabar" = "aquamarine3"
    ),
    labels = c("Jawa Barat", "DKI Jakarta")
  ) +
  labs(
    x = NULL,
    y = "Jumlah kasus akumulasi",
    colour = NULL,
    title = "Dinamika Perbandingan Kasus Meninggal
di DKI Jakarta dan Jawa Barat",
    caption = "Sumber data: covid.19.go.id"
```

```

) +
theme_ipsum(
  base_size = 13,
  plot_title_size = 21,
  grid = "Y",
  ticks = TRUE
) +
theme(
  plot.title = element_text(hjust = 0.5),
  legend.position = "top"
)

```

Dinamika Perbandingan Kasus Meninggal di DKI Jakarta dan Jawa Barat



```

#Melakukan pivot, transformasi data dari wide ke Long
library(dplyr)
library(tidyr)

dim(cov_perbandingan_A7)

cov_perbandingan_A7_pivot <-
  cov_perbandingan_A7 %>%
  gather(
    key = "kategori",
    value = "jumlah",
    -tanggal
  ) %>%
  mutate(
    kategori = sub(pattern = "A7_", replacement = "", kategori)
  )

dim(cov_perbandingan_A7_pivot)

glimpse(cov_perbandingan_A7_pivot)

```

152 · 3

304 · 3

Rows: 304

Columns: 3

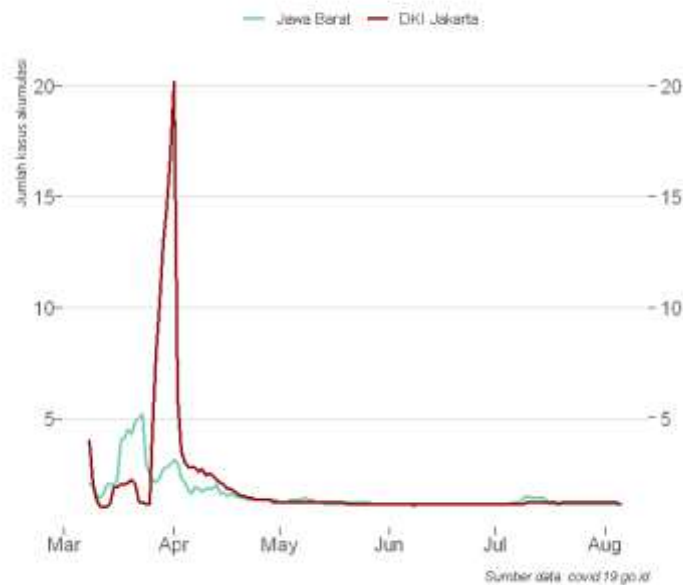
\$ tanggal <date> 2020-03-08, 2020-03-09, 2020-03-10, 2020-03-11, 2020-03-1...

\$ kategori <chr> "jkt", "jkt", "jkt", "jkt", "jkt", "jkt", "jkt", "jkt", "j...

\$ jumlah <dbl> 4.0, 2.0, 1.3, 1.0, 1.0, 1.0, 1.1, 1.9, 1.9, 2.0, 2.0, 2.1...

```
#Membuat komparasi antara kasus A7 di DKI Jakarta & Jawa Barat
library(ggplot2)
library(hrbrthemes)
ggplot(cov_perbandingan_A7_pivot, aes(tanggal, jumlah, colour = (kategori))) +
  geom_line(size = 0.9) +
  scale_y_continuous(sec.axis = dup_axis(name = NULL)) +
  scale_colour_manual(
    values = c(
      "jkt" = "darkred",
      "jabar" = "aquamarine3"
    ),
    labels = c("Jawa Barat", "DKI Jakarta")
  ) +
  labs(
    x = NULL,
    y = "Jumlah kasus akumulasi",
    colour = NULL,
    title = "Dinamika Perbandingan A7
di DKI Jakarta dan Jawa Barat",
    caption = "Sumber data: covid.19.go.id"
  ) +
  theme_ipsum(
    base_size = 13,
    plot_title_size = 21,
    grid = "Y",
    ticks = TRUE
  ) +
  theme(
    plot.title = element_text(hjust = 0.5),
    legend.position = "top"
  )
```


Dinamika Perbandingan A7 di DKI Jakarta dan Jawa Barat



Kesimpulan

Jumlah penduduk di Jawa Barat memang lebih besar dari Provinsi DKI Jakarta, namun hasil visualisasi data menunjukkan bahwa DKI Jakarta memiliki kasus lebih banyak dari Jawa Barat. Kita dapat mengambil beberapa asumsi dari kasus ini:

1. DKI Jakarta merupakan pusat pemerintahan, dan juga pusat ekonomi dan bisnis
2. Tingkat mobilitas di DKI Jakarta sangatlah tinggi dan padat
3. DKI Jakarta dapat disebut sebagai gerbang masuk Warga Negara Asing (WNA). Jika kita amati dari kasus pertama di Indonesia yang terjadi di wilayah Jakarta Selatan akibat berinteraksi dengan WNA asal Jepang
4. Tingkat identifikasi virus corona dan usaha pencegahan di Indonesia pada awal tahun 2020 masih rendah, akibatnya WNA yang datang ke Indonesia melalui DKI Jakarta tidak terdeteksi di awal
5. Jawa Barat memiliki mayoritas wilayah berupa pedesaan dan masyarakatnya agak tertutup. Sehingga warganya cenderung hanya berinteraksi dengan orang yang sama dan di lokasi yang sama (tidak banyak arus keluar masuk seperti DKI Jakarta)

VI. Kesimpulan

Keterbukaan data terkait penanganan COVID-19 memiliki manfaat bagi masyarakat. Keterbukaan data bisa meningkatkan kewaspadaan masyarakat dalam menghadapi wabah yang disebabkan virus corona. Keterbukaan data juga berpengaruh bagi pemerintah untuk mengambil kebijakan terkait penanganan COVID-19. Serta, bisa dijadikan bahan introspeksi terhadap kebijakan penanganan wabah yang selama ini sudah diterapkan. Gugus Tugas Percepatan Penanganan COVID-19 telah berinisiatif untuk membuat sistem integrasi data antara pemerintah pusat dan daerah. Data seputar COVID-19 dapat diakses melalui situs covid.go.id.

Dari analisis dan visualisasi data yang kita lakukan, data yang masuk memiliki starting point di tanggal yang berbeda. Data nasional dimulai tanggal 2 Maret 2020 serta DKI Jakarta dan Jawa Barat mulai tanggal 1 Maret 2020. Data yang dikirim dari provinsi ke nasional juga memiliki keterlambatan. Memiliki probabilitas yang tinggi jika data yang masuk di suatu provinsi pada tanggal A, baru masuk ke data nasional pada tanggal C. Jika hal tersebut terjadi pada arus provinsi ke nasional, maka dapat kita ambil kemungkinan bahwa keterlambatan masuknya data juga berlaku pada lingkup wilayah dibawahnya.

Data yang direkam oleh Gugus Tugas Percepatan Penanganan COVID-19 di Indonesia tidaklah 100% maksimal. Dikarenakan Indonesia masih belum mengadakan *test* COVID-19 secara masif. *World Health Organization* (WHO) menargetkan suatu negara untuk mampu melakukan *testing* 1,000 per 1 juta penduduk per minggu. Namun, jika kita menggunakan asumsi jumlah penduduk Indonesia 260 juta jiwa, maka dibutuhkan tes COVID-19 sebanyak 260 ribu per minggu. Dengan demikian, target per hari berada di kisaran angka 37 ribu. Tetapi *testing* di Indonesia saat ini hanya 13 ribu sampai 14 ribu per hari. Masih kurang tiga kali lipat dari target *testing* seharusnya. Selain itu, tes yang dilakukan juga belum sepenuhnya dilakukan untuk kasus baru melainkan juga pemeriksaan lanjutan pasien yang sudah positif. Akibatnya pelaksanaan tes menjadi tak maksimal.

Sehingga kita tidak dapat mempercayai 100% kepada hasil analisis dan visualisasi data yang kita lakukan pada praktik kali ini. Tetapi data yang diperoleh dari pemerintah tetap sangatlah berpengaruh dalam membuat kebijakan demi kestabilan pertahanan dan keamanan negara di segala bidang pada saat yang tidak tentu ini.

Pandemi ini menjadi penghambat pertumbuhan hampir di segala bidang. Kejadian ini merupakan tak terduga dan tak tentu, sehingga tidak ada upaya pencegahan yang dilakukan maksimal. Tidak hanya Indonesia, namun hampir seluruh negara di dunia mengalami hal yang

sama. Tetapi pemerintah terus meningkatkan kinerja dalam menyelesaikan masalah ini. Kali ini kita akan membandingkan visualisasi data yang kita peroleh dengan beberapa bidang yang terpengaruh pandemi:

- **Bidang Ekonomi**

Negara ingin menyelamatkan kesehatan warga negaranya dengan beberapa kebijakan seperti *lockdown* dan PSBB. Namun di lain sisi, beberapa bidang seperti ekonomi menjadi terhambat. Banyak pekerja yang di PHK, hampir semua industri mengalami penurunan kinerja, dan masih banyak lagi. Badan Pusat Statistik (BPS) Indonesia mencatat pertumbuhan ekonomi Indonesia pada kuartal I-2020 sebesar +2.97% dan kuartal II-2020 sebesar -5.32%. Diharapkan di kuartal III-2020, Indonesia mengalami pertumbuhan ekonomi ke arah positif agar tidak mengalami resesi. Hal tersebut dikarenakan bulan Juli – September, Indonesia telah melonggarkan PSBB sedikit demi sedikit dan telah memasuki era normal baru. Sebagian besar industri telah berjalan seperti biasa. Pemerintah memiliki tugas yang berat dalam memastikan semua aspek negara terseleamatkan.

Selain itu, kita dapat bandingkan dengan IHSG (Indeks Harga Saham Gabungan) di Indonesia. Tercatat bahwa data pertama kasus positif yang masuk berasal dari provinsi di Indonesia tanggal 1 Maret 2020. Lalu mulai 5 Maret 2020, IHSG mulai menurun. Dan menyentuh titik terendah pada tanggal 24 Maret 2020 (Rp 3,937.63). Per tanggal 7 Agustus 2020, IHSG menyentuh angka Rp 5,143.89 dan terus mengalami kenaikan sedikit demi sedikit seiring pelonggaran PSBB di Indonesia, khususnya DKI Jakarta.

- **Bidang Politik**

Dampak negatif ekonomi dan sosial yang terasa akibat pandemi COVID-19 berpotensi menimbulkan krisis politik, terutama soal kepercayaan publik kepada pemerintah. Solusinya adalah pemerintah harus sukses salurkan stimulus ekonomi di tengah pandemi COVID-19. Pemerintah juga harus sigap dalam mengatur beberapa peraturan baru yang sifatnya mendadak dan sementara untuk menanggulangi COVID-19 di Indonesia serta menjaga keamanan dan ketahanan negara.

Timbul juga banyak pertentangan argumen dan ketidaksetujuan pendapat antar tingkat pemerintahan. Hal tersebut dikarenakan kurangnya musyawarah dan semua pergerakan akibat COVID-19 harus dilakukan secepat mungkin, sehingga muncul miskomunikasi antar tingkat pemerintahan.

- **Bidang Budaya**

Memasuki era normal baru, maka masyarakat harus memiliki *culture* yang baru. Semua itu berhubungan dengan kesehatan masyarakat. Seperti tidak berjabat tangan, menggunakan masker, mencuci tangan, *physical distancing*, dan lain-lain. Hal-hal tersebut sangatlah tidak biasa bagi masyarakat Indonesia. Namun, kita harus menjadikan hal tersebut sebagai kebiasaan yang baru demi kesehatan kita bersama. Adapun juga *culture* yang baru dengan teknologi. Hampir semua hal kini dapat dikerjakan secara *remote*. Masyarakat Indonesia dituntut untuk *agile* dan *adaptive* terhadap perkembangan teknologi yang sangat pesat di era normal baru.

- **Bidang Pendidikan**

Semua beralih menggunakan teknologi berupa *conference call*. Tapi sayangnya, pandemi membuat ketidakmerataan akses pendidikan di seluruh wilayah Indonesia. Tidak semua masyarakat dan wilayah dapat mengakses internet yang baik. Pemerintah perlu memikirkan hal ini terlebih lanjut, karena masih banyak anak Indonesia yang masih belum memperoleh akses pendidikan yang maksimal selama pandemi.

Tenaga pengajar harus beradaptasi dengan teknologi yang ada. Hal ini menjadi sebuah kesulitan, terutama bagi tenaga pengajar yang sudah berusia lanjut. Mereka dituntut untuk mempelajari teknologi dengan cepat untuk menyalurkan pengetahuan mereka kepada anak didiknya.

Di lain sisi, bagi pelajar yang mampu, mereka diberi kesempatan untuk *explore* diri lebih lagi di rumah. Mereka tidak harus belajar dari institusi pendidikan, namun juga dari akses internet. Banyak hal yang dapat dipelajari selama pandemi ini dan memanfaatkan waktu sebijak mungkin.

Setelah melakukan analisis dan visualisasi data, kita disadarkan akan pentingnya mengikuti segala kebijakan yang diterapkan oleh pemerintah demi kepentingan kita bersama. Semua masyarakat ingin segala aktivitas berjalan seperti normal untuk kehidupan yang layak. Analisis dan visualisasi data COVID-19 juga dapat dijadikan sebagai alat introspeksi di berbagai bidang industri. Kita dapat menentukan tujuan, target, dan hal lainnya demi stabilitas suatu industri dengan melakukan analisis dan visualisasi data yang ada. Misalnya saja kita dapat membuat beberapa asumsi dan perkiraan kapan pandemi ini akan berakhir dari *Seventh-Day Amplification Factor*. Tetap jaga kebersihan, kesehatan, dan selalu *physical distancing*. *New normal, new habit*.

References

Badan Pusat Statistik Provinsi DKI Jakarta. 2019. *3.1.2. Jumlah Penduduk Provinsi DKI Jakarta Menurut Kelompok Umur dan Jenis Kelamin, 2018-2019*, diakses dari <https://jakarta.bps.go.id/dynamictable/2019/09/16/58/jumlah-penduduk-provinsi-dki-jakarta-menurut-kelompok-umur-dan-jenis-kelamin-2018-.html>, pada 4 Agustus 2020.

CNN Indonesia. 2020. *Target Spesimen Corona Pemerintah Belum Capai Standar WHO*, diakses dari <https://www.cnnindonesia.com/nasional/20200806073210-20-532671/target-tes-spesimen-corona-pemerintah-belum-capai-standar-who>, pada 6 Agustus 2020.

Correia S, Luck S, Verner E. *Pandemics depress the economy, public health interventions do not: evidence from the 1918 Flu*. Published online on March 30, 2020.

Darmajati, Danu. 2020. *Anies: Saat ini Alat Rapid Test Corona Bukan untuk Tes Random-Massal*, diakses dari <https://news.detik.com/berita/d-4950919/anies-saat-ini-alat-rapid-test-corona-bukan-untuk-tes-random-massal>, pada 5 Agustus 2020.

Gugus Tugas Percepatan Penanganan COVID-19. 2020. *Peta Sebaran*, diakses dari <https://covid19.go.id/peta-sebaran>, pada 7 Agustus 2020.

Indonesia Stock Exchange. 2020. *IDX Quarterly Statistics – 1st Quarter 2020*. Jakarta: Indonesia Stock Exchange.

Indonesia Stock Exchange. 2020. *IDX Quarterly Statistics – 2nd Quarter 2020*. Jakarta: Indonesia Stock Exchange.

Ulya, Fika Nurul. 2020. *Pertumbuhan Ekonomi RI Minus 5,32 Persen pada Kuartal II-2020*, diakses dari <https://money.kompas.com/read/2020/08/05/120854826/pertumbuhan-ekonomi-ri-minus-532-persen-pada-kuartal-ii-2020>, pada 4 Agustus 2020.

Wikipedia. 2020. *Jawa Barat*, diakses dari https://id.wikipedia.org/wiki/Jawa_Barat, pada 4 Agustus 2020.

Worldometer. 2020. *Indonesia Population*, diakses dari <https://www.worldometers.info/world-population/indonesia-population/>, pada 4 Agustus 2020.