



[MENU](#)

[ANALYSIS](#)

[CONTACT](#)

DATA SCIENCE

DATA SCIENCE PORTOFOLIO

Assa Trissia Rizal | 2022





assatrissiarzl@gmail.com

+62895377154926

Assa Trissia Rizal

TENTANG PENULIS

Hai. Nama saya Assa Trisia Rizal. Kebanyakan orang memanggilku Assa. Saya lulusan S1 dari Universitas Tanjungpura dengan konsentrasi jurusan Statistika. Melalui portofolio ini, saya ingin menunjukkan apa yang telah saya lakukan sejauh ini dalam proyek berbasis Data Science. Semoga pada akhir portofolio ini, Anda akan mendapatkan wawasan berharga dan kita dapat mendiskusikan kemungkinan peluang di luar sana. Selamat membaca!

Pengalaman

Project Machine Learning for Retail with R: Product Packaging | DQLAB

April 2022

Project: Data Scientist Assessment Using R

April 2022

Project Analisa Klasifikasi Pinjaman untuk Sektor UMKM

April 2022

Pendidikan

Universitas Tanjungpura

Sep 2017 - Nov 2021

Tingkat Sarjana – Statistika

Pengalaman Organisasi

Anggota |

Himpunan Mahasiswa Statistika

Mar 2018 - Apr 2019

Skill

Data Analyst

Data Scientist

Tools

SPSS

Minitab

KNIME

Language

Indonesia (native)

Inggris (menengah)

1

Project Machine Learning for Retail with R: Product Packaging DQLAB

DQLAB



RINGKASAN

DQLab.id Fashion adalah sebuah toko fashion yang menjual berbagai produk seperti jeans, kemeja, kosmetik, dan lain-lain. Walaupun cukup berkembang, namun dengan semakin banyaknya kompetitor dan banyak produk yang stoknya masih banyak tentunya membuat kuatir Pak Agus, manajer DQLab.id Fashion. Pada proyek portofolio individu kali ini, DQLAB berkolaborasi bersama Xeratic memberikan sebuah dataset transaksi penjualan produk yang memiliki 33,699 baris data dan terdapat 3,450 kode transaksi.



PROJECT

Import Library dan Membaca Data

```
#import library
library(arules)
#import data
data_transaksi <- read.transactions(file="../input/transaksi-
retail/transaksi_dqlab_retail.tsv", format="single", sep="\t", cols=c(1,2), skip=1)
data_transaksi
```

```
transactions in sparse format with
 3450 transactions (rows) and
 69 items (columns)
```

Dari output di atas, didapat informasi bahwa ada 3450 transaksi dan ada 69 item yang terdapat pada objek transactions bernama transaksi_tabular tersebut.

MENAMPILKAN DATA TRANSAKSI TOP 10

Untuk melihat statistik top 10 (item dengan frequency paling banyak) value data akan diurutkan secara decreasing (decreasing = TRUE). Lalu untuk menampilkan data top 10, maka dipilih 10 data teratas. Terakhir, untuk menampilkan data, format data diubah menjadi data frame melalui base function data.frame.

Menampilkan Data Transaksi Top 10

```
data_item_top <- itemFrequency(data_transaksi, type="absolute")
data_item_top <- sort(data_item_top, decreasing = TRUE)
data_item_top <- data_item_top[1:10]
data_item_top <- data.frame("Nama Produk"=names(data_item_top),
"Jumlah"=data_item_top, row.names=NULL)
data_item_top
```

Hasil

Nama.Produk	Jumlah
Shampo Biasa	2075
Serum Vitamin	1685
Baju Batik Wanita	1312
Baju Kemeja Putih	1255
Celana Jogger Casual	1136
Cover Koper	1086
Sepatu Sandal Anak	1062
Tali Pinggang Gesper Pria	1003
Sepatu Sport merk Z	888
Wedges Hitam	849

MENAMPILKAN DATA TRANSAKSI BOTTOM 10

Tahap untuk mendapatkan statistik bottom 10 mirip seperti tahap sebelumnya, hanya saja berbeda di value dari parameter sort, untuk mendapatkan item dengan frequency paling sedikit, value dari paramter decreasing diubah menjadi decreasing = FALSE.

Menampilkan Data Transaksi Bottom 10

```
data_item_top <- itemFrequency(data_transaksi, type="absolute")
data_item_top <- sort(data_item_top, decreasing = TRUE)
data_item_top <- data_item_top[1:10]
data_item_top <- data.frame("Nama Produk"=names(data_item_top),
"Jumlah"=data_item_top, row.names=NULL)
data_item_top
```

Hasil

Nama.Produk	Jumlah
Celana Jeans Sobek Pria	9
Tas Kosmetik	11
Stripe Pants	19
Pelembab	24
Tali Ban Ikat Pinggang	27
Baju Renang Pria Anak-anak	32
Hair Dye	46
Atasan Baju Belang	56
Tas Sekolah Anak Perempuan	71
Dompot Unisex	75

MENDAPATKAN KOMBINASI PRODUK YANG MENARIK

Selanjutnya, Manager DQLab.id Fashion meminta daftar 10 paket kombinasi produk yang menarik. Kombinasi menarik yang dimaksud manager disini adalah kombinasi yang memiliki asosiasi atau hubungan erat. Untuk mendapatkan kombinasi produk tersebut, kita dapat menggunakan algoritma apriori melalui function apriori pada library arules.

MENAMPILKAN KOMBINASI PRODUK YANG MENARIK

```
mba <- apriori(data_transaksi, parameter = list(supp = 10/length(data_transaksi), confidence = 0.5, minlen = 2, maxlen = 3))
mba <- head(sort(mba, by="lift", decreasing=TRUE), n=10) #mba
inspect(mba)
```

Hasil

	lhs	rhs	support	confidence	lift	count
[1]	{Tas Makeup, Tas Pinggang Wanita}	=> {Baju Renang Anak Perempuan}	0.010434783	0.8780488	24.42958	36
[2]	{Tas Makeup, Tas Travel}	=> {Baju Renang Anak Perempuan}	0.010144928	0.8139535	22.64629	35
[3]	{Tas Makeup, Tas Ransel Mini}	=> {Baju Renang Anak Perempuan}	0.011304348	0.7358491	20.47322	39
[4]	{Sunblock Cream, Tas Pinggang Wanita}	=> {Kuas Makeup }	0.016231884	0.6913580	20.21343	56
[5]	{Baju Renang Anak Perempuan, Tas Pinggang Wanita}	=> {Tas Makeup}	0.010434783	0.8000000	19.57447	36
[6]	{Baju Renang Anak Perempuan, Tas Ransel Mini}	=> {Tas Makeup}	0.011304348	0.7959184	19.47460	39
[7]	{Baju Renang Anak Perempuan, Celana Pendek Green/Hijau}	=> {Tas Makeup}	0.010144928	0.7777778	19.03073	35
[8]	{Tas Makeup, Tas Waist Bag}	=> {Baju Renang Anak Perempuan}	0.004347826	0.6818182	18.96994	15
[9]	{Celana Pendek Green/Hijau, Tas Makeup}	=> {Baju Renang Anak Perempuan}	0.010144928	0.6730769	18.72674	35
[10]	{Dompot Flip Cover, Sunblock Cream}	=> {Kuas Makeup }	0.016231884	0.6292135	18.39650	56

Jika diperhatikan, value dari *lift* 10 kombinasi tersebut semuanya > 1 , artinya ada hubungan positif antara kombinasi antar item di atas.

MENCARI PAKET PRODUK YANG BISA DIPASANGKAN DENGAN ITEM SLOW-MOVING

Slow-moving item adalah produk yang pergerakan penjualannya lambat atau kurang cepat. Ini akan bermasalah apabila item produk tersebut masih menumpuk. Jika tidak dijual satuan maka perlu mencari asosiasi kuat dari item produk ini dengan produk lain sehingga jika dipaketkan akan menjadi lebih menarik.

Manager DQLab.id Fashion juga meyakini hal ini, setelah di analisis, diidentifikasi bahwa ada dua item produk yang menurut dia stoknya masih banyak dan perlu dicari pasangan item untuk pemaketannya. Dua item produk tersebut adalah “Tas Makeup” dan “Baju Renang Pria Anak-anak”. Manager ingin meminta kombinasi yang bisa dipaketkan dengan kedua produk tersebut.

MENDAPATKAN PAKET PRODUK YANG BISA DIPASANGKAN DENGAN ITEM SLOW-MOVING

```
mba2 <- apriori(data_transaksi, parameter = list(supp = 10/length(data_transaksi), confidence = 0.1, minlen = 2, maxlen = 3))
#subset tas makeup
mba_1 <- subset(mba2, rhs %in% "Tas Makeup")
mba_1 <- head(sort(mba_1, by="lift", decreasing=TRUE), n=3) #subset baju renang pria anak-anak\mba_2 <- subset(mba2, rhs %in%
"Baju Renang Pria Anak-anak")
mba_2 <- head(sort(mba_2, by="lift", decreasing=TRUE), n=3)
mba2 <- c(mba_1, mba_2) #mba2 inspect(mba2)
```

Hasil

	lhs	rhs	support	confidence	lift	count
[1]	{Baju Renang Anak Perempuan, Tas Pinggang Wanita}	=> {Tas Makeup}	0.010434783	0.8000000	19.57447	36
[2]	{Baju Renang Anak Perempuan, Tas Ransel Mini}	=> {Tas Makeup}	0.011304348	0.7959184	19.47460	39
[3]	{Baju Renang Anak Perempuan, Celana Pendek Green/Hijau}	=> {Tas Makeup}	0.010144928	0.7777778	19.03073	35
[4]	{Gembok Koper, Tas Waist Bag}	=> {Baju Renang Pria Anak-anak}	0.004057971	0.2745098	29.59559	14
[5]	{Flat Shoes Ballerina, Gembok Koper}	=> {Baju Renang Pria Anak-anak}	0.004057971	0.1866667	20.12500	14
[6]	{Celana Jeans Sobek Wanita, Jeans Jumbo}	=> {Baju Renang Pria Anak-anak}	0.005507246	0.1210191	13.04737	19

HASIL

Dari hasil di atas, didapat hasil 3 kombinasi masing-masing untuk 2 item *slow-moving* item yang memiliki asosiasi paling kuat yaitu:

Tas Makeup:

- {Baju Renang Anak Perempuan, Tas Pinggang Wanita} => {Tas Makeup}
 - {Baju Renang Anak Perempuan, Tas Ransel Mini} => {Tas Makeup}
 - {Baju Renang Anak Perempuan, Celana Pendek Green/Hijau} => {Tas Makeup}
 - Baju Renang Pria Anak-anak:
 - {Gembok Koper, Tas Waist Bag} => {Baju Renang Pria Anak-anak}
 - {Flat Shoes Ballerina, Gembok Koper} => {Baju Renang Pria Anak-anak}
 - {Celana Jeans Sobek Wanita, Jeans Jumbo} => {Baju Renang Pria Anak-anak}
-



KESIMPULAN



1. Tas Makeup (item slow-moving) dapat dikombinasikan/dipaketkan dengan Baju Renang Anak Perempuan dan Tas Pinggang Wanita.
2. Baju Renang Pria Anak-anak (item slow-moving) dapat dikombinasikan/dipaketkan dengan Gembok Koper dan Tas Waist Bag.
3. Dari rekomendasi paket produk tersebut, dapat disimpulkan bahwa target konsumen adalah Ibu Rumah Tangga.



REKOMENDASI

1. Beberapa rekomendasi New Bundling Package telah berhasil didapatkan untuk mengurangi stok persediaan barang sehingga dapat meningkatkan penjualan.
 2. Peluncuran berbagai diskon menarik atau harga khusus dengan paket penjualan produk diatas. Sehingga produk yang memiliki status “item slow moving” stok produknya bisa habis terjual menggunakan rekomendasi paket produk yang baru.
 3. Rak display dengan lebih didekatkan misalnya rak display “flat shoes ballerina, gembok koper bisa didekatkan dengan baju renang pria anak-anak” begitupun dengan rekomendasi paket produk lainnya berdasarkan hasil implementasi algoritma apriori,
 4. Jika sudah memiliki web commerce maka display penawaran produk dapat didesain ulang menyesuaikan rekomendasi hasil analisis.
-

2

Eksplorasi dan Analisis Data COVID-19 Indonesia menggunakan R

DQLAB



RINGKASAN

Proyek ini disediakan oleh DQLab. COVID-19 adalah penyakit yang kini telah menjadi pandemi global. Pemerintah di berbagai negara umumnya sigap membentuk satuan tugas untuk menangani penyebaran COVID19 di masyarakat, termasuk pemerintah di Indonesia. Salah satu bentuk tindakan yang dilakukan pemerintah adalah dengan mengumpulkan dan memberikan data perkembangan kasus COVID-19 kepada masyarakat. Data perkembangan kasus seringkali juga dilengkapi dengan dashboard dan grafik visualisasi pendukung dengan harapan masyarakat dapat lebih mudah memahami informasi tersebut. Misalnya, portal covid19.go.id yang dibuat oleh Gugus Tugas Percepatan Penanganan Covid-19 Nasional dan portal PIKOBAR milik Pemprov Jabar. Serta banyak portal data COVID-19 lainnya yang disediakan oleh masing-masing pemerintah daerah.

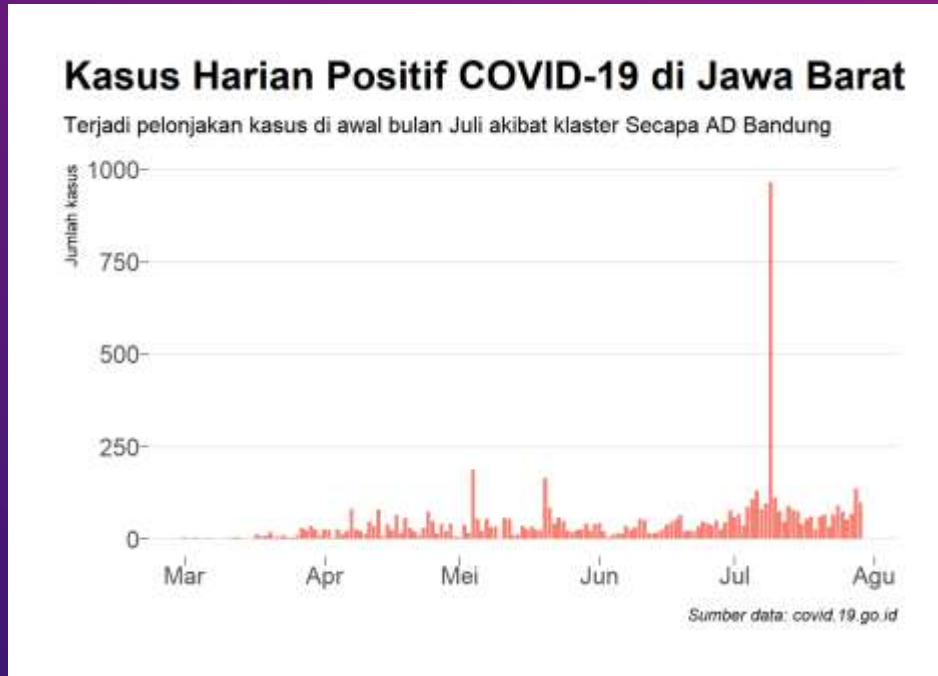


Kasus Harian di Jawa Barat

Buat query dengan menggunakan R

```
library(ggplot2)
library(hrbrthemes)
ggplot(new_cov_jabar, aes(tanggal, kasus_baru)) +
  geom_col(fill = "salmon") +
  labs(
    x=NULL,
    y="Jumlah kasus",
    title="Kasus Harian Positif COVID-19 di Jawa Barat",
    subtitle="Terjadi pelonjakan kasus di awal bulan Juli akibat klaster Secapa AD
    bandung",
    caption = "Sumber data: covid.19.go.id"
  )+
  theme_ipsum(
    base_size = 13,
    plot_title_size = 21,
    grid="Y",
    ticks = TRUE
  )+
  theme(plot.title.position = "plot")
```

HASIL



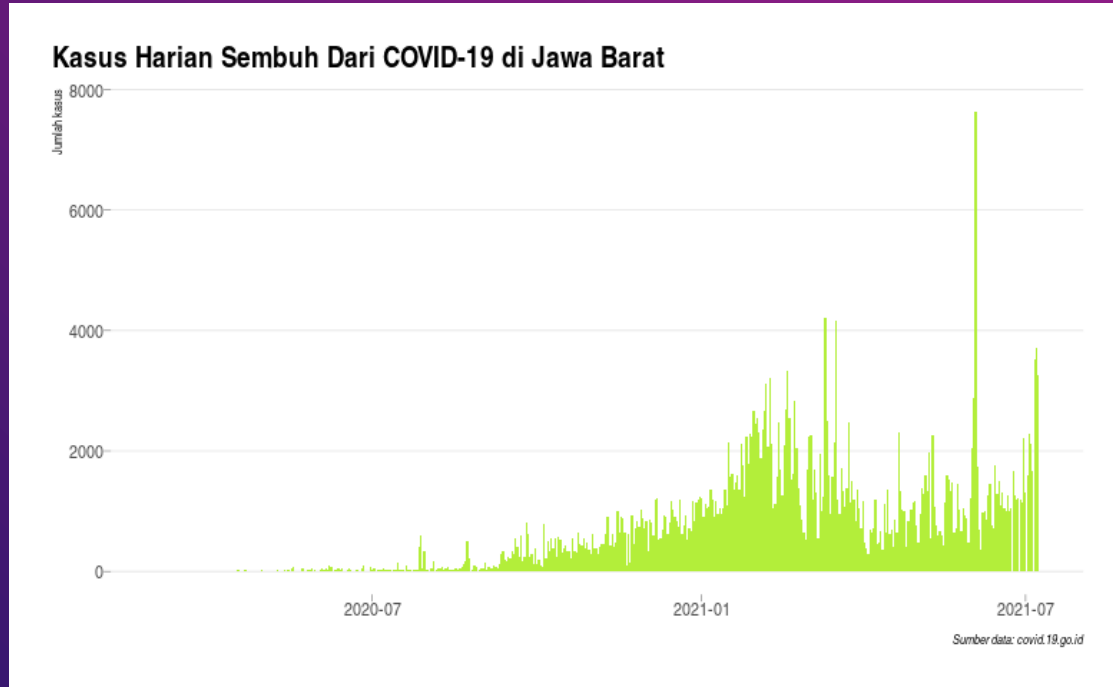
Terlihat pada grafik harian kasus Jawa Barat jumlah kasus tertinggi pada bulan Juli dengan total kasus tertinggi lebih dari 1000 kasus dan kasus terendah pada bulan Maret dengan rata-rata jumlah kasus kurang dari 250.

Grafik untuk Kasus Sembuh

Buat query dengan menggunakan R

```
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_ipsum(
    base_size = 13,
    plot_title_size = 21,
    grid = "Y",
    ticks = TRUE
  ) +
  theme(plot.title.position = "plot")
```

HASIL



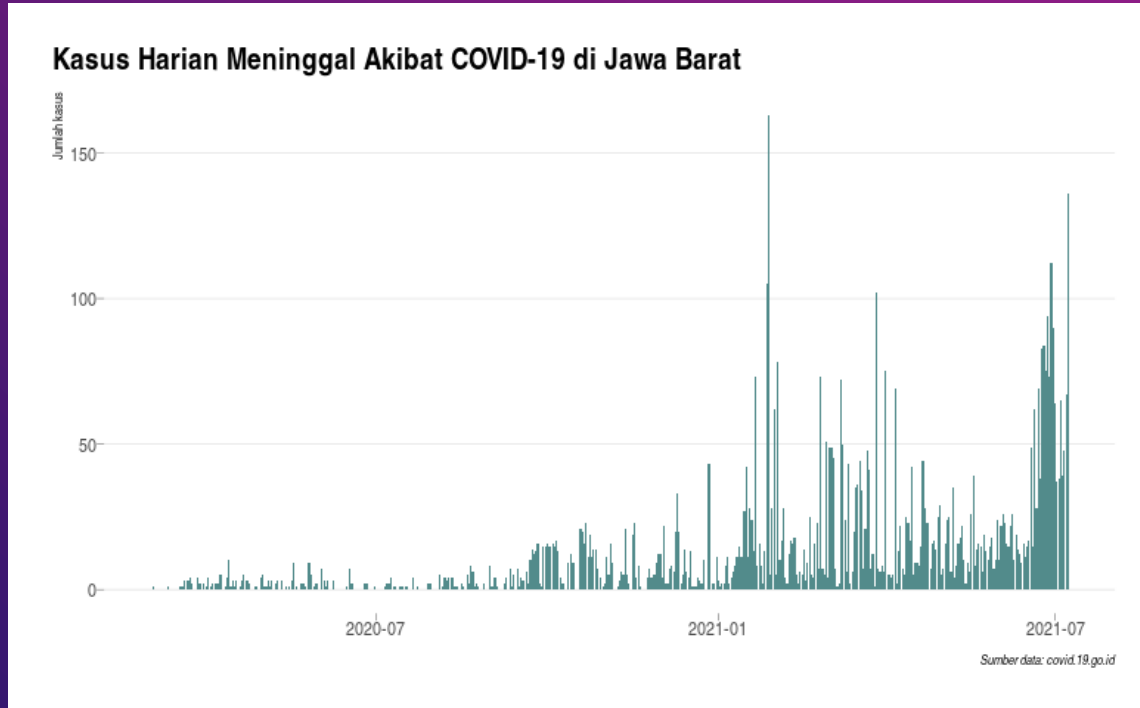
Pada Grafik Kasus Sembuh di Jawa Barat terjadi peningkatan kesembuhan dari bulan Januari hingga Juli dengan jumlah kasus tertinggi lebih dari 6000 kasus .

Grafik untuk Kasus Kematian

Buat query dengan menggunakan R

```
library(ggplot2)
library(hrbrthemes)
ggplot(new_cov_jabar, aes(tanggal, meninggal)) +
  geom_col(fill = "darkslategray4") +
  labs(
    x = NULL,
    y = "Jumlah kasus",
    title = "Kasus Harian Meninggal Akibat 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.position = "plot")
```

HASIL



Grafik Kasus Kematian mulai meningkat pada bulan Januari hingga Juli dengan nilai kasus tertinggi menjadi lebih dari 150 pada bulan Januari.

Kasus Positif Mingguan COVID-19 di Jawa Barat

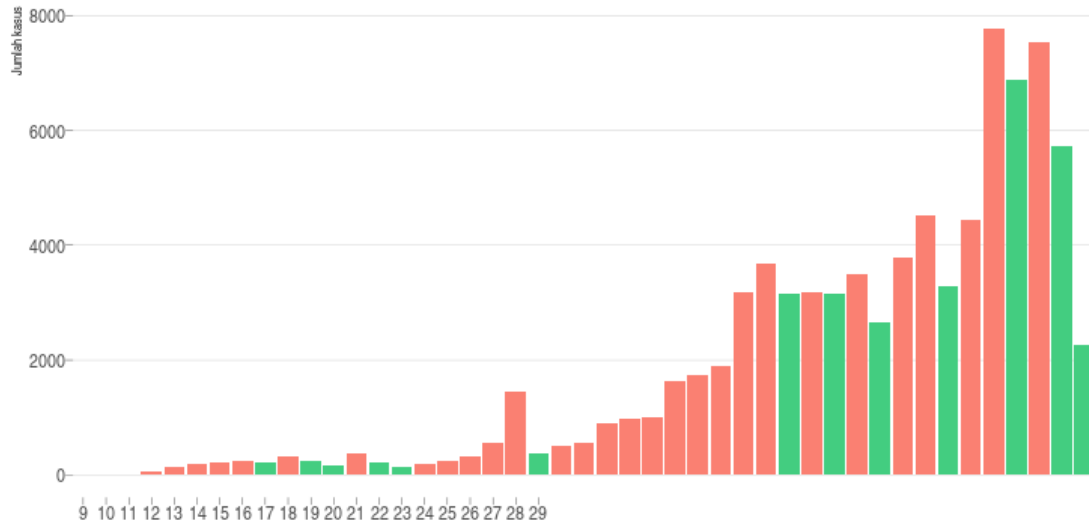
Buat query dengan menggunakan R

```
library(ggplot2)
library(hrbrthemes)
ggplot(cov_jabar_pekanan[cov_jabar_pekanan$tahun==2020,], aes(pekan_ke, jumlah, fill
= lebih_baik)) +
geom_col(show.legend = FALSE) +
scale_x_continuous(breaks = 9:29, 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 menunjukkan penambahan kasus baru lebih sedikit dibandingkan
satu pekan sebelumnya",
caption = "Sumber data: covid.19.go.id"
)+
theme_ipsum(
base_size = 13,
plot_title_size = 21,
grid = "Y",
ticks = TRUE
) +
theme(plot.title.position = "plot")
```


HASIL

Kasus Pekanan Positif COVID-19 di Jawa Barat

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



Sumber data: covid.19.go.id

Kasus positif COVID-19 mingguan di Jawa Barat pada kolom hijau menunjukkan kasus baru lebih sedikit dibandingkan minggu sebelumnya. Pada minggu ke 30 terjadi penurunan kasus

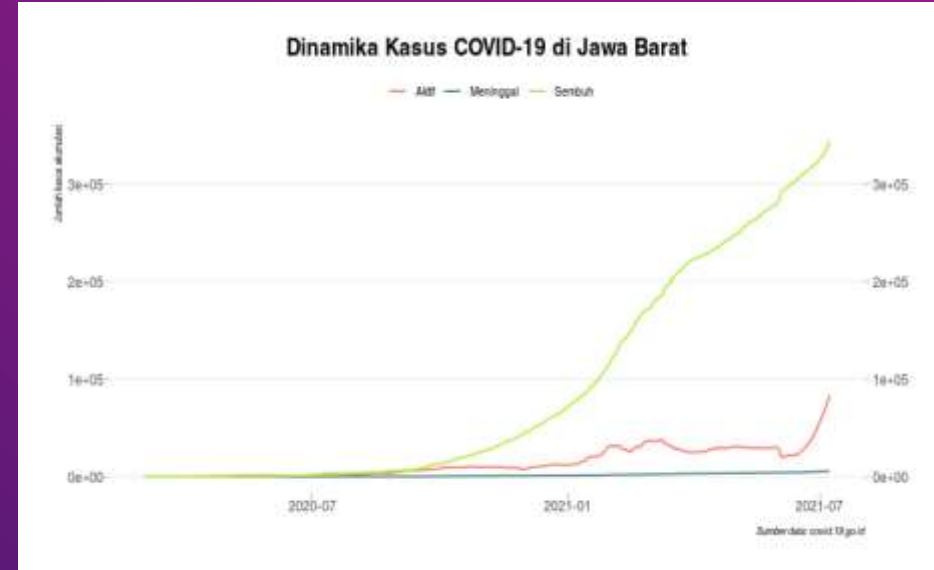
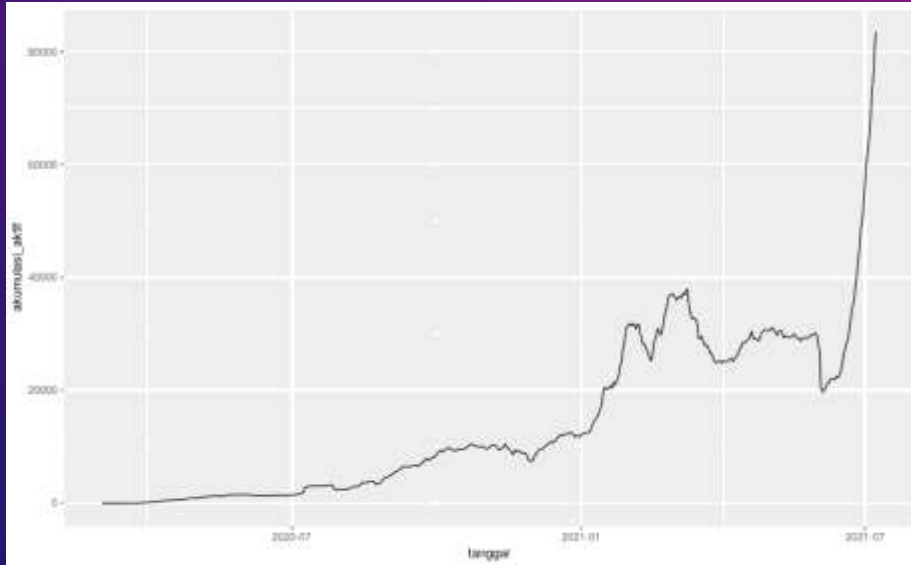
Pola dan Dinamika

Buat query dengan menggunakan R

```
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" = "salmon",
      "meninggal" = "darkslategray4",
      "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
  ) +
```

DINAMIKA KASUS COVID-19 DI JABAR

Hasil



Kita bisa melihat akumulasi aktif COVID-19 di Jawa Barat pada grafik pada bulan Juli terjadi peningkatan yang signifikan.

Dinamika Kasus COVID-19 di Jawa Barat. dapat dilihat pada grafik untuk kasus akumulasi pulih yang meningkat tinggi. untuk kasus aktif dan mati terlihat sangat miring dan tidak ada peningkatan yang sangat signifikan

3

Mini Project: Analisis Penjualan dengan SQL

DQLAB

Project 1. Fundamental SQL Using SELECT Statement

Question:

1. Menyiapkan data transaksi penjualan dengan total revenue \geq IDR 100.000
2. Format datanya yang akan ditampilkan adalah: kode_pelanggan, nama_produk, qty, harga, dan total.
3. Diurutkan mulai dari total revenue terbesar. (NB : kolom total diperoleh dari perkalian antara kolom qty dan kolom harga).

tabel tr_penjualan

kode_transaksi	kode_pelanggan	no_urut	kode_produk	nama_produk	qty	harga
tr-001	dqlabcust07	1	prod-01	Kotak Pensil DQLab	5	62500
tr-001	dqlabcust07	2	prod-03	Flash disk DQLab 32 GB	1	100000
tr-001	dqlabcust07	3	prod-09	Buku Planner Agenda DQLab	3	92000
tr-001	dqlabcust07	4	prod-04	Flashdisk DQLab 32 GB	3	40000
tr-002	dqlabcust01	1	prod-03	Gift Voucher DQLab 100rb	2	100000
tr-002	dqlabcust01	2	prod-10	Sticky Notes DQLab 500 sheets	4	55000
tr-002	dqlabcust01	3	prod-07	Tas Travel Organizer DQLab	1	48000
tr-003	dqlabcust03	1	prod-02	Flashdisk DQLab 64 GB	2	55000
tr-004	dqlabcust03	1	prod-10	Sticky Notes DQLab 500 sheets	5	55000
tr-004	dqlabcust03	2	prod-04	Flashdisk DQLab 32 GB	4	40000
tr-005	dqlabcust05	1	prod-09	Buku Planner Agenda DQLab	3	92000
tr-005	dqlabcust05	2	prod-01	Kotak Pensil DQLab	1	62500
tr-005	dqlabcust05	3	prod-04	Flashdisk DQLab 32 GB	2	40000
tr-006	dqlabcust02	1	prod-05	Gift Voucher DQLab 250rb	4	250000
tr-006	dqlabcust02	2	prod-08	Gantungan Kunci DQLab	2	15800

Untuk menjawab soal pada project 1, kita akan meng-query data dari tabel tr_penjualan. Dan membuat kolom baru dengan nama total yang akan menyimpan value dari perkalian antara kolom qty dan kolom harga. Berikut simple query untuk menampilkan informasi yang dibutuhkan.

```
SELECT kode_pelanggan, nama_produk, qty, harga,  
qty*harga AS total FROM tr_penjualan WHERE  
qty*harga >= 100000 ORDER BY total DESC;
```

kode_pelanggan	nama_produk	qty	harga	total
dqlabcust02	Gift Voucher DQLab 250rb	4	250000	1000000
dqlabcust07	Kotak Pensil DQLab	5	62500	312500
dqlabcust07	Buku Planner Agenda DQLab	3	92000	276000
dqlabcust05	Buku Planner Agenda DQLab	3	92000	276000
dqlabcust03	Sticky Notes DQLab 500 sheets	5	55000	275000
dqlabcust01	Sticky Notes DQLab 500 sheets	4	55000	220000
dqlabcust01	Gift Voucher DQLab 100rb	2	100000	200000
dqlabcust03	Flashdisk DQLab 32 GB	4	40000	160000
dqlabcust07	Flashdisk DQLab 32 GB	3	40000	120000
dqlabcust03	Flashdisk DQLab 64 GB	2	55000	110000
dqlabcust07	Flash disk DQLab 32 GB	1	100000	100000

1. Project 2. Sql Fundamental Using Function and Group By

Pertanyaan:

1. Melakukan analisis penjualan di suatu store. Adapun laporan yang diminta sebagai berikut:
2. Total jumlah seluruh penjualan (total/revenue).
3. Total quantity seluruh produk yang terjual.
4. Total quantity dan total revenue untuk setiap kode produk.
5. Rata - Rata total belanja per kode pelanggan.
6. Selain itu, jangan lupa untuk menambahkan kolom baru dengan nama 'kategori' yang mengkategorikan total/revenue ke dalam 3 kategori: High: > 300K; Medium: 100K - 300K; Low: <100K.
7. NB: Informasi tabel menggunakan tabel tr_penjualan pada soal project 1.

Sama seperti project 1, untuk menjawab soal diatas kita akan menggunakan informasi pada tabel tr_penjualan. Namun pada project ke-2 ini kita akan menggunakan query yang sedikit advance, dimana kita akan menggunakan function aggregate dan grouping serta function case..when. Berikut query untuk menampilkan informasi yang dibutuhkan

```
## 1. Total jumlah seluruh penjualan (total/revenue).  
SELECT SUM(total) as total  
FROM tr_penjualan;  
## 2. Total quantity seluruh produk yang terjual.  
SELECT SUM(qty) as qty  
FROM tr_penjualan;  
## 3. Total quantity dan total revenue untuk setiap kode produk.  
SELECT kode_produk, SUM(qty) as qty, SUM(total) as total  
FROM tr_penjualan  
GROUP BY kode_produk;
```

```
+-----+  
| total |  
+-----+  
| 3271600 |  
+-----+  
  
+-----+  
| qty |  
+-----+  
| 42 |  
+-----+  
  
+-----+-----+-----+  
| kode_produk | qty | total |  
+-----+-----+-----+  
| prod-01 | 6 | 375000 |  
| prod-02 | 2 | 110000 |  
| prod-03 | 3 | 300000 |  
| prod-04 | 9 | 360000 |  
| prod-05 | 4 | 1000000 |  
| prod-07 | 1 | 48000 |  
| prod-08 | 2 | 31600 |  
| prod-09 | 6 | 552000 |  
| prod-10 | 9 | 495000 |  
+-----+-----+-----+
```



```

## 4. Rata - Rata total belanja per kode pelanggan.
SELECT kode_pelanggan, AVG(total) as avg_total
FROM tr_penjualan
GROUP BY kode_pelanggan;
## 5. Selain itu, jangan lupa untuk menambahkan kolom baru dengan nama 'kategori' yang
mengkategorikan total/revenue ke dalam 3 kategori: High: > 300K; Medium: 100K - 300K; Low:
<100K.
SELECT kode_transaksi,kode_pelanggan,no_urut,kode_produk, nama_produk, qty, total,
CASE
    WHEN total > 300000 THEN 'high'
    WHEN total < 100000 THEN 'low'
    ELSE 'Medium'
END as kategori
FROM tr_penjualan;

```

kode_pelanggan	avg_total
dqlabcust01	156000.0000
dqlabcust02	515800.0000
dqlabcust03	181666.6667
dqlabcust05	139500.0000
dqlabcust07	202125.0000

kode_transaksi	kode_pelanggan	no_urut	kode_produk	nama_produk	qty	total	kategori
tr-001	dqlabcust07	1	prod-01	Kotak Pensil DQLab	5	312500	high
tr-001	dqlabcust07	2	prod-03	Flash disk DQLab 32 GB	1	100000	Medium
tr-001	dqlabcust07	3	prod-09	Buku Planner Agenda DQLab	3	276000	Medium
tr-001	dqlabcust07	4	prod-04	Flashdisk DQLab 32 GB	3	120000	Medium
tr-002	dqlabcust01	1	prod-03	Gift Voucher DQLab 100rb	2	200000	Medium
tr-002	dqlabcust01	2	prod-10	Sticky Notes DQLab 500 sheets	4	220000	Medium
tr-002	dqlabcust01	3	prod-07	Tas Travel Organizer DQLab	1	48000	low
tr-003	dqlabcust03	1	prod-02	Flashdisk DQLab 64 GB	2	110000	Medium
tr-004	dqlabcust03	1	prod-10	Sticky Notes DQLab 500 sheets	5	275000	Medium
tr-004	dqlabcust03	2	prod-04	Flashdisk DQLab 32 GB	4	160000	Medium
tr-005	dqlabcust05	1	prod-09	Buku Planner Agenda DQLab	3	276000	Medium
tr-005	dqlabcust05	2	prod-01	Kotak Pensil DQLab	1	62500	low
tr-005	dqlabcust05	3	prod-04	Flashdisk DQLab 32 GB	2	80000	low
tr-006	dqlabcust02	1	prod-05	Gift Voucher DQLab 250rb	4	1000000	high
tr-006	dqlabcust02	2	prod-08	Gantungan Kunci DQLab	2	31600	low

Project 3. SQL Fundamental Using INNER JOIN and UNION Part 2

Pertanyaan:

Persiapkanlah data katalog mengenai mengenai nama-nama produk yang akan dijual di suatu store. Data tersebut akan digunakan dalam meeting untuk mereview produk mana saja yang akan dilanjutkan penjualannya dan mana yang tidak akan dilanjutkan.

Siapkan hanya data produk dengan harga di bawah 100K untuk kode produk prod-1 sampai prod-5; dan dibawah 50K untuk kode produk prod-6 sampai prod-10, tanpa mencantumkan kolom no_urut.

```
SELECT nama_produk, kode_produk, harga from ms_produk_1 where harga < 100000
UNION
SELECT nama_produk, kode_produk, harga from ms_produk_2 where harga < 50000;
```

Tabel ms_produk_1

no_urut	kode_produk	nama_produk	harga
1	prod-01	Kotak Pensil DQLab	62500
2	prod-02	Flashdisk DQLab 64 GB	55000
3	prod-03	Gift Voucher DQLab 100rb	100000
4	prod-04	Flashdisk DQLab 32 GB	40000
5	prod-05	Gift Voucher DQLab 250rb	250000

Tabel ms_produk_2

no_urut	kode_produk	nama_produk	harga
6	prod-06	Pulpen Multifunction + Laser DQLab	92500
7	prod-07	Tas Travel Organizer DQLab	48000
8	prod-08	Gantungan Kunci DQLab	15800
9	prod-09	Buku Planner Agenda DQLab	92000
10	prod-10	Sticky Notes DQLab 500 sheets	55000

Untuk menjawab pertanyaan pada *project* union ini, kita akan menggabungkan kedua buah tabel, sama seperti kasus inner join sebelumnya. Namun perbedaannya adalah jika pada kasus inner join kita menggabungkan tabel secara horizontal dan menentukan kolom kunci pada kedua buah tabel, maka pada kasus union ini, kita akan menggabungkan tabel secara vertikal dengan syarat jumlah kolom pada kedua tabel harus sama dan memiliki tipe data sama pada masing-masing kolom dengan urutan atau posisi yang sama pula . Berikut query untuk menggabungkan tabel menggunakan perintah union.

```
SELECT nama_produk, kode_produk, harga from ms_produk_1 where harga < 100000
UNION
SELECT nama_produk, kode_produk, harga from ms_produk_2 where harga < 50000;
```

nama_produk	kode_produk	harga
Kotak Pensil DQLab	prod-01	62500
Flashdisk DQLab 64 GB	prod-02	55000
Flashdisk DQLab 32 GB	prod-04	40000
Tas Travel Organizer DQLab	prod-07	48000
Gantungan Kunci DQLab	prod-08	15800