

**LAPORAN PRAKTIKUM  
UTS DALIS**



**LAPORAN PRAKTIKUM UTS DALIS**

**Data Analysis COVID-19 Russia Regions Cases**

**Disusun Oleh:**

**11423004 : Marselino Tambunan  
11423018 : Eduward Simanjuntak  
11423019 : Hans Manalu**

**Teknologi Rekayasa Perangkat Lunak**

**FAKULTAS VOKASI**

**INSTITUT TEKNOLOGI DEL**

**2025**

# DAFTAR ISI

<b>BAB I.....</b>	<b>3</b>
<b>Pendahuluan.....</b>	<b>3</b>
<b>1.1 Latar Belakang.....</b>	<b>3</b>
<b>1.2 Tujuan.....</b>	<b>3</b>
<b>1.3 Rumusan Masalah .....</b>	<b>4</b>
<b>BAB II .....</b>	<b>5</b>
<b>Metodologi Penelitian .....</b>	<b>5</b>
<b>2.1 Tahapan CRISP-DM .....</b>	<b>5</b>
2.1.1 Business Understanding .....	5
2.1.2 Data Understanding.....	6
2.1.3 Data Preparation .....	9
2.1.4 Modelling.....	13
2.1.5 Evaluasi .....	13
2.1.6 Deployment.....	14
<b>BAB III.....</b>	<b>21</b>
<b>Hasil Eksplorasi dan Analisis.....</b>	<b>21</b>
<b>3.1 Hasil Eksplorasi Data .....</b>	<b>21</b>
<b>3.2 Analisis .....</b>	<b>22</b>
<b>3.3 Rekomendasi Bisnis .....</b>	<b>24</b>
1. Optimalisasi Alokasi Sumber Daya Kesehatan.....	24
2. Penguatan Kebijakan Pencegahan di Wilayah Padat Penduduk .....	24
3. Perencanaan Ekonomi dan Bisnis .....	24
4. Pengembangan Sistem Monitoring Berkelanjutan .....	25
5. Strategi Komunikasi Publik dan Edukasi .....	25
<b>BAB IV .....</b>	<b>26</b>
<b>Kesimpulan.....</b>	<b>26</b>

# **BAB I**

## **Pendahuluan**

### **1.1 Latar Belakang**

Project ini merupakan salah satu bentuk penerapan pembelajaran dalam bidang *Data Analytics* yang dilakukan secara tim oleh mahasiswa. Melalui project ini, tim berupaya memahami dan mengimplementasikan tahapan analisis data mulai dari pengumpulan data (*data collection*), pembersihan data (*data cleaning*), hingga analisis dan visualisasi hasil untuk mendapatkan insight yang bermakna.

Topik yang diangkat dalam project ini adalah data terkait penyebaran COVID-19. Pemilihan tema ini didasari oleh tingginya relevansi isu pandemi terhadap berbagai aspek kehidupan, terutama dalam bidang kesehatan dan kebijakan publik. Dengan menganalisis dataset COVID-19, tim berharap dapat memahami pola penyebaran, tingkat kasus, serta dampak yang muncul di berbagai wilayah.

### **1.2 Tujuan**

Tujuan utama dari proyek analisis ini adalah untuk memahami secara komprehensif kondisi dan dinamika penyebaran pandemi COVID-19 di Rusia melalui pendekatan berbasis data. Secara khusus, analisis ini bertujuan untuk:

1. Mengetahui tren perkembangan kasus COVID-19 di Rusia berdasarkan periode waktu tertentu, guna melihat pola kenaikan dan penurunan kasus.
2. Menganalisis distribusi kasus di berbagai wilayah Rusia, termasuk identifikasi wilayah dengan tingkat kasus aktif, kematian, dan kesembuhan tertinggi.
3. Mengevaluasi hubungan antara kasus aktif, kasus sembuh, dan kematian untuk memahami tingkat keparahan pandemi di masing-masing wilayah.
4. Memberikan insight dan rekomendasi yang dapat dijadikan dasar dalam pengambilan keputusan, kebijakan publik, maupun strategi kesehatan yang lebih efektif.
5. Mengembangkan dashboard visual interaktif yang mempermudah interpretasi data dan pemantauan kondisi pandemi secara real-time.

### **1.3 Rumusan Masalah**

Berdasarkan latar belakang dan tujuan di atas, maka rumusan masalah dalam proyek ini dapat dijabarkan sebagai berikut:

1. Bagaimana tren perkembangan kasus COVID-19 di Rusia dari waktu ke waktu?
2. Wilayah mana yang memiliki tingkat penyebaran kasus tertinggi dan terendah?
3. Bagaimana perbandingan antara jumlah kasus terkonfirmasi, kematian, dan kesembuhan di tiap wilayah?
4. Faktor apa saja yang berpotensi memengaruhi peningkatan atau penurunan kasus di wilayah tertentu?
5. Bagaimana visualisasi data melalui dashboard dapat membantu pemahaman dan pengambilan keputusan terkait pandemi COVID-19 di Rusia

## **BAB II**

### **Metodologi Penelitian**

#### **2.1 Tahapan CRISP-DM**

##### **2.1.1 Business Understanding**

Pandemi COVID-19 yang melanda sejak awal tahun 2020 telah memberikan dampak besar terhadap berbagai negara di dunia, termasuk Rusia. Dengan luas wilayah dan jumlah penduduk yang besar, Rusia menghadapi tantangan signifikan dalam memantau serta mengendalikan penyebaran virus ini. Oleh karena itu, analisis data COVID-19 di Rusia menjadi penting untuk memahami bagaimana pola persebaran, tren kenaikan maupun penurunan kasus, serta efektivitas penanganan pandemi di tiap wilayah.

Proyek analisis ini bertujuan untuk memahami kondisi dan dinamika penyebaran COVID-19 di Rusia melalui pendekatan berbasis data. Dengan menggunakan dataset yang berisi informasi mengenai jumlah kasus terkonfirmasi, kematian, dan kesembuhan, analisis ini membantu memberikan gambaran yang lebih komprehensif terhadap situasi pandemi di negara tersebut.

Secara umum, tujuan utama dari analisis ini adalah:

1. Mengetahui tren perkembangan kasus COVID-19 di Rusia berdasarkan periode waktu tertentu untuk melihat pola peningkatan atau penurunan kasus dari waktu ke waktu.
2. Menganalisis distribusi kasus di berbagai wilayah Rusia, termasuk identifikasi wilayah dengan kasus aktif tertinggi, tingkat kematian tertinggi, dan tingkat kesembuhan tertinggi.
3. Mengevaluasi hubungan antara kasus aktif, kasus sembuh, dan kematian, guna memperoleh pemahaman mengenai tingkat keparahan pandemi di masing-masing wilayah.
4. Memberikan insight dan rekomendasi yang dapat digunakan sebagai bahan pertimbangan dalam pengambilan keputusan atau kebijakan yang lebih efektif dalam penanggulangan COVID-19.

Analisis ini juga bertujuan untuk menjawab beberapa pertanyaan utama yaitu:

- Bagaimana tren perkembangan kasus COVID-19 di Rusia dari waktu ke waktu?
- Wilayah mana yang mengalami tingkat penyebaran kasus tertinggi dan terendah?

- Bagaimana perbandingan antara jumlah kasus terkonfirmasi, kasus sembuh, dan kematian di tiap wilayah?
- Apa saja faktor yang berpotensi mempengaruhi peningkatan atau penurunan kasus di wilayah tertentu?

Dari hasil analisis yang disajikan melalui dashboard visualisasi, diharapkan dapat ditemukan pola yang jelas mengenai situasi pandemi di Rusia

## 2.1.2 Data Understanding

### 2.1.2.1 Sumber Data

Dataset yang digunakan berjudul “COVID-19 Russia Regions Cases”, diambil dari platform Kaggle. Dataset ini berisi data perkembangan kasus COVID-19 di berbagai wilayah Rusia, yang dikumpulkan dari sumber resmi dan media publik.

Sumber utama: <https://www.kaggle.com/datasets/kapral42/covid19-russia-regions-cases?select=covid19-russia-cases.csv>

Sumber tambahan:

1. <https://meduza.io>
2. <https://coronavirus-monitor.ru>
3. <https://yandex.ru/company/researches/2020/podomam>
4. <https://xn--80aefpebagmfblc0a.xn--p1ai/information/>
5. Platform publikasi dataset: Kaggle - COVID-19 Russia Regions Cases

### 2.1.2.2 Deskripsi Data

Dataset yang digunakan dalam proyek ini berjudul “COVID-19 Russia Regions Cases”, bersumber dari platform Kaggle. Dataset ini berisi data perkembangan kasus COVID-19 di berbagai wilayah (region) di Rusia yang dikumpulkan dari sumber resmi pemerintah Rusia, yaitu Rospotrebnadzor, serta beberapa sumber publik seperti Meduza.io, Coronavirus-monitor.ru, Yandex Research, dan situs informasi resmi COVID-19 Rusia. Periode data mencakup Januari 2020 hingga saat ini, dan diperbarui secara harian. Dataset ini mencatat jumlah kasus terkonfirmasi (confirmed), sembuh (recovered), dan meninggal (deaths), baik dalam bentuk harian maupun kumulatif di setiap wilayah Rusia.

Dataset ini dikumpulkan dengan tujuan untuk menganalisis penyebaran COVID-19 di tingkat regional, bukan hanya pada level nasional. Dengan demikian, peneliti dan analis dapat memahami perbedaan tingkat penyebaran antarwilayah, tren waktu perkembangan kasus, serta efektivitas kebijakan lokal dalam menanggulangi pandemi.

### 2.1.2.3 Relevansi Dataset

Dataset ini relevan dengan tujuan proyek karena:

- Menyediakan informasi terperinci per wilayah, memungkinkan analisis spasial dan temporal.
- Dapat digunakan untuk mendeteksi wilayah risiko tinggi, memantau efektivitas kebijakan, serta menyusun rekomendasi mitigasi berbasis data.

### 2.1.2.4 Struktur Dataset

Jumlah kolom : 10    Jumlah Baris data : 5.843

Dataset utama yang digunakan adalah file COVID-19.csv , yang memiliki beberapa kolom utama sebagai berikut:

No	Nama Kolom	Deskripsi	Tipe Data
1	date	Tanggal pencatatan kasus	Date / Object
2	region	Nama wilayah atau provinsi di Rusia	Object (String)
3	confirmed	Jumlah kasus positif yang terkonfirmasi	Integer / Float
4	deaths	Jumlah kematian akibat COVID-19	Integer / Float
5	recovered	Jumlah pasien yang dinyatakan sembuh	Integer / Float
6	active ( <i>jika ada</i> )	Jumlah kasus aktif (confirmed – recovered – deaths)	Integer / Float

### 1.Data Numerik

- Region\_ID → kode numerik tiap region
- Day-Confirmed → jumlah kasus terkonfirmasi per hari
- Day-Deaths → jumlah kematian per hari
- Day-Recovered → jumlah pasien sembuh per hari
- Confirmed → total kasus terkonfirmasi kumulatif
- Deaths → total kematian kumulatif
- Recovered → total pasien sembuh kumulatif

## 2. Data Kategorikal

- Date → tanggal pelaporan kasus
- Region/City → nama wilayah atau kota
- Region/City\_Eng → nama wilayah dalam bahasa Inggris

### 2.1.2.5 Missing Value dan Outlier

Geographic roles

Country/Region:

State/Province:

City:  ⚠ 82 issues

Match values to locations

⚠ City

Your Data	Matching Location
Altai region	Unrecognized
Altai Republic	Unrecognized
Amurskaya Oblast	Unrecognized
Arkhangelsk region	Unrecognized
Astrakhan region	Unrecognized
Belgorod region	Unrecognized
Bryansk region	Unrecognized
Chechen Republic	Unrecognized
Chelyabinsk region	Unrecognized

☐ Show only unmatched locations in drop down list

Banyak data pada visualisasi peta tidak dapat terbaca dengan benar karena terdapat ketidaksesuaian penamaan wilayah atau kota pada dataset. Hal ini menyebabkan Tableau gagal mengenali lokasi geografis dari sebagian data yang seharusnya ditampilkan. Permasalahan ini umumnya terjadi karena format penulisan nama daerah yang tidak standar, seperti adanya singkatan, penulisan yang tidak lengkap, atau penggunaan karakter yang tidak dikenali akibat perbedaan encoding.



	A	B	C	D	E	F	G	H	I	J	K
1	Date	Region/City	Region/Ci	Region_IC	Day-Confi	Day-Death	Day-Reco	Confirmed	Deaths	Recovered	
2	#####	Diamond Princess	Diamond Princess		3.0	0.0	0.0	3.0	0.0	0.0	
3	#####	????????????????????	Tyumen r	72.0	1.0	0.0	0.0	1.0	0.0	0.0	
4	#####	????????????????????	Zabaykals	75.0	1.0	0.0	0.0	1.0	0.0	0.0	
5	#####	????????????????????	Tyumen r	72.0	0.0	0.0	1.0	1.0	0.0	1.0	
6	#####	????????????????????	Zabaykals	75.0	0.0	0.0	1.0	1.0	0.0	1.0	
7	#####	????????????????????	Moscow r	50.0	1.0	0.0	0.0	1.0	0.0	0.0	
8	#####	????????????-????	St. Peters	78.0	1.0	0.0	0.0	1.0	0.0	0.0	
9	#####	????????????	Moscow	77.0	5.0	0.0	0.0	5.0	0.0	0.0	
10	#####	????????????????????	Nizhny Nc	52.0	1.0	0.0	0.0	1.0	0.0	0.0	
11	#####	Diamond Princess	Diamond Princess		0.0	0.0	3.0	3.0	0.0	3.0	
12	#####	????????????	Moscow	77.0	0.0	0.0	1.0	5.0	0.0	1.0	
13	#####	????????????-????	St. Peters	78.0	1.0	0.0	0.0	2.0	0.0	0.0	
14	#####	????????????????????	Lipetsk re	48.0	3.0	0.0	0.0	3.0	0.0	0.0	
15	#####	????????????????????	Moscow r	50.0	1.0	0.0	0.0	2.0	0.0	0.0	

Beberapa nama wilayah pada dataset, terutama dari Rusia, awalnya tidak terbaca dengan benar dan berubah menjadi simbol aneh. Hal ini disebabkan oleh perbedaan encoding teks antara sumber data yang menggunakan format UTF-8 dan pengaturan regional Excel yang tidak mendukung karakter non-Latin secara default. Untuk mengatasinya, file CSV diimpor ulang menggunakan pengaturan Unicode (UTF-8) agar seluruh teks tampil dengan benar sebelum dianalisis di Tableau.

## 2.1.3 Data Preparation

### 2.1.3.1 Masalah Tampilan Data dan Perbedaan Encoding (Unreadable Characters)

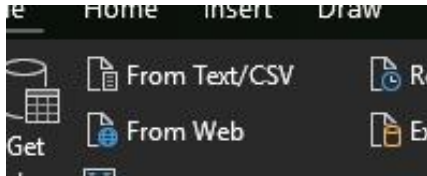
	A	B	C	D	E	F	G	H	I	J	K
1	Date	Region/City	Region/Ci	Region_IC	Day-Confi	Day-Death	Day-Reco	Confirmed	Deaths	Recovered	
2	#####	Diamond Princess	Diamond Princess		3.0	0.0	0.0	3.0	0.0	0.0	
3	#####	????????????????????	Tyumen r	72.0	1.0	0.0	0.0	1.0	0.0	0.0	
4	#####	????????????????????	Zabaykals	75.0	1.0	0.0	0.0	1.0	0.0	0.0	
5	#####	????????????????????	Tyumen r	72.0	0.0	0.0	1.0	1.0	0.0	1.0	
6	#####	????????????????????	Zabaykals	75.0	0.0	0.0	1.0	1.0	0.0	1.0	
7	#####	????????????????????	Moscow r	50.0	1.0	0.0	0.0	1.0	0.0	0.0	
8	#####	????????????-????	St. Peters	78.0	1.0	0.0	0.0	1.0	0.0	0.0	
9	#####	????????????	Moscow	77.0	5.0	0.0	0.0	5.0	0.0	0.0	
10	#####	????????????????????	Nizhny Nc	52.0	1.0	0.0	0.0	1.0	0.0	0.0	
11	#####	Diamond Princess	Diamond Princess		0.0	0.0	3.0	3.0	0.0	3.0	
12	#####	????????????	Moscow	77.0	0.0	0.0	1.0	5.0	0.0	1.0	
13	#####	????????????-????	St. Peters	78.0	1.0	0.0	0.0	2.0	0.0	0.0	
14	#####	????????????????????	Lipetsk re	48.0	3.0	0.0	0.0	3.0	0.0	0.0	
15	#####	????????????????????	Moscow r	50.0	1.0	0.0	0.0	2.0	0.0	0.0	

Beberapa nama wilayah dalam dataset, khususnya yang berasal dari Rusia, awalnya tidak dapat terbaca dengan benar dan berubah menjadi deretan simbol atau karakter aneh. Permasalahan ini muncul akibat perbedaan sistem encoding teks, di mana sumber data menggunakan format UTF-

8, sedangkan pengaturan regional Excel secara default tidak mendukung karakter non-Latin. Akibatnya, huruf-huruf berbahasa Rusia tidak terinterpretasi dengan benar saat file dibuka di Excel.

Solusi :

### 1. Mengubah format data menggunakan fitur From Text/CSV



### 2. Lalu Lakukan Transformasi Data

covid19-russia-cases.csv

File Origin: 65001: Unicode (UTF-8) | Delimiter: Comma | Data Type Detection: Based on first 200 rows

Date	Region/City	Region/City-Eng	Region_ID	Day-Confirmed	Day-Deaths	Day-Recovered	Confirmed	Deaths
26/01/2020	Diamond Princess	Diamond Princess	null	30	0	0	30	0
31/01/2020	Тюменская область	Tyumen region	720	10	0	0	10	0
31/01/2020	Забайкальский край	Zabaykalsky Krai	750	10	0	0	10	0
12/02/2020	Тюменская область	Tyumen region	720	0	0	10	10	0
12/02/2020	Забайкальский край	Zabaykalsky Krai	750	0	0	10	10	0
03/02/2020	Московская область	Moscow region	500	10	0	0	10	0
05/03/2020	Санкт-Петербург	St. Petersburg	780	10	0	0	10	0
06/03/2020	Москва	Moscow	770	50	0	0	50	0
06/03/2020	Нижегородская область	Nizhny Novgorod Region	520	10	0	0	10	0
06/03/2020	Diamond Princess	Diamond Princess	null	0	0	30	30	0
07/03/2020	Москва	Moscow	770	0	0	10	10	0
07/03/2020	Санкт-Петербург	St. Petersburg	780	10	0	0	10	0
07/03/2020	Липецкая область	Lipetsk region	480	30	0	0	30	0
08/03/2020	Московская область	Moscow region	500	10	0	0	10	0
08/03/2020	Белгородская область	Belgorod region	310	10	0	0	10	0
08/03/2020	Калининградская область	Kaliningrad region	390	10	0	0	10	0
09/03/2020	Москва	Moscow	770	30	0	0	30	0
11/03/2020	Москва	Moscow	770	60	0	0	60	0
11/03/2020	Московская область	Moscow region	500	20	0	0	20	0
12/03/2020	Москва	Moscow	770	40	0	0	40	0

Buttons: Load, Transform Data, Cancel

### 3. Setelah Data sudah benar klik close & load

Microsoft Excel interface showing the 'Data' tab ribbon and the 'Queries' pane.

Queries list: covid19-russia-cases

Formula bar: = Table.TransformColumnTypes(#"Promoted Headers",{"Date", type date}, {"Region/City-Eng", type text}, {"Region/City", type text}, {"Region\_ID", type text})

Date	Region/City	Region/City-Eng	Region_ID
26/01/2020	Diamond Princess	Diamond Princess	null
31/01/2020	Тюменская область	Tyumen region	720
31/01/2020	Забайкальский край	Zabaykalsky Krai	750
12/02/2020	Тюменская область	Tyumen region	720
12/02/2020	Забайкальский край	Zabaykalsky Krai	750
03/02/2020	Московская область	Moscow region	500
05/03/2020	Санкт-Петербург	St. Petersburg	780
06/03/2020	Москва	Moscow	770
06/03/2020	Нижегородская область	Nizhny Novgorod Region	520
06/03/2020	Diamond Princess	Diamond Princess	null
07/03/2020	Москва	Moscow	770
07/03/2020	Санкт-Петербург	St. Petersburg	780

### 5. Setelah semua rapi simpan ulang dengan ekstensi utf 8

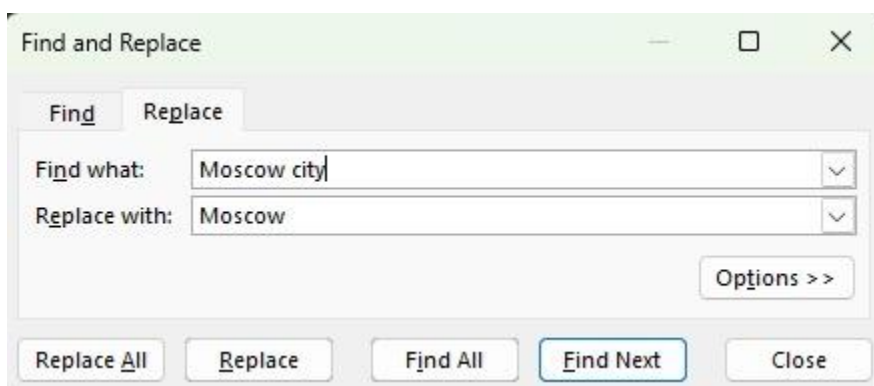
Date	Region/City	Region/City-Eng	Region_ID	Day-Confirmed	Day-Deaths	Day-Recovered	Confirmed	Deaths	Recovered
26/01/2020	Diamond Princess	Diamond Princess		30	0	0	30	0	0
31/01/2020	Тюменская область	Tyumen region	720	10	0	0	10	0	0
31/01/2020	Забайкальский край	Zabaykalsky Krai	750	10	0	0	10	0	0
12/02/2020	Тюменская область	Tyumen region	720	0	0	10	10	0	10
12/02/2020	Забайкальский край	Zabaykalsky Krai	750	0	0	10	10	0	10
03/02/2020	Московская область	Moscow region	500	10	0	0	10	0	0
05/03/2020	Санкт-Петербург	St. Petersburg	780	10	0	0	10	0	0
06/03/2020	Москва	Moscow	770	50	0	0	50	0	0
06/03/2020	Нижегородская область	Nizhny Novgorod Region	520	10	0	0	10	0	0
06/03/2020	Diamond Princess	Diamond Princess		0	0	30	30	0	30
07/03/2020	Москва	Moscow	770	0	0	10	50	0	10
07/03/2020	Санкт-Петербург	St. Petersburg	780	10	0	0	20	0	0
07/03/2020	Липецкая область	Lipetsk region	480	30	0	0	30	0	0
08/03/2020	Московская область	Moscow region	500	10	0	0	20	0	0
08/03/2020	Белгородская область	Belgorod region	310	10	0	0	10	0	0
08/03/2020	Калининградская область	Kaliningrad region	390	10	0	0	10	0	0
09/03/2020	Москва	Moscow	770	30	0	0	80	0	10
11/03/2020	Москва	Moscow	770	60	0	0	140	0	10

### 2.1.3.2 Ketidaksesuaian Data Lokasi (Unrecognized Geographic Locations)

Banyak data pada visualisasi peta yang tidak terbaca dengan benar karena adanya ketidaksesuaian penamaan wilayah atau kota pada dataset. Tableau tidak dapat mengenali beberapa nama lokasi tersebut akibat perbedaan format penulisan, seperti penggunaan singkatan, ejaan yang tidak standar, atau karakter khusus yang tidak dikenali sistem. Akibatnya, sebagian data tidak dapat ditampilkan di peta sesuai posisi geografisnya. Untuk mengatasinya, nama-nama wilayah perlu disesuaikan terlebih dahulu, baik dengan melakukan perbaikan langsung di Excel maupun melalui proses pembersihan data (data cleaning) di luar Tableau. Penyesuaian ini bertujuan agar seluruh nama wilayah sesuai dengan format geografis yang dikenali oleh Tableau, sehingga data dapat divisualisasikan secara lengkap dan akurat.

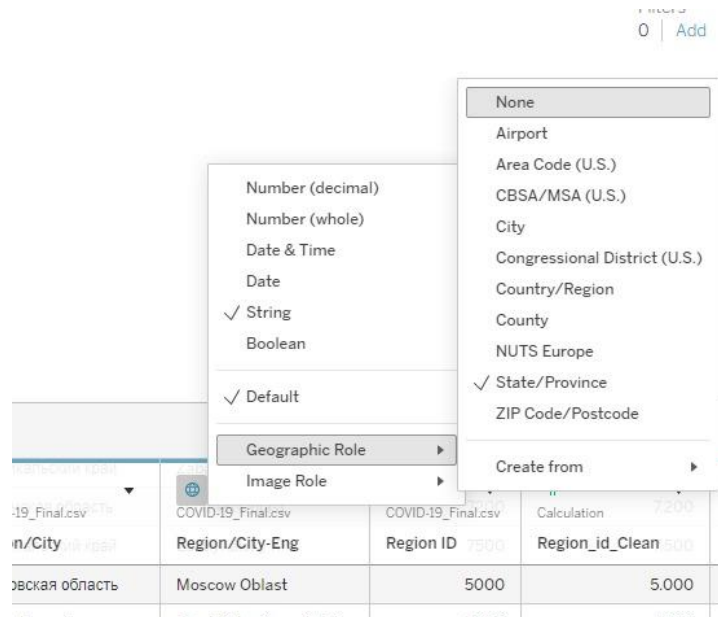
Solusi :

1. Mengubah Secara Manual di Excel lalu upload ulang



### 2.1.3.3 Mengatasi Missing Value dan Outlier

1. Menyiapkan field geografis Geographical Role → State/Province supaya dapat digunakan pada visualisasi map



2. Membuat rumus pada calculated field Region\_id\_Clean



Pada dataset COVID-19 Russia Regions Cases, ditemukan beberapa missing values pada kolom *Region ID*. Untuk mengatasinya, dibuat calculated field bernama *Region\_id\_Clean* dengan rumus `IFNULL([Region ID], 0)` yang mengganti nilai NULL menjadi 0. Langkah ini mencegah kesalahan saat agregasi dan pemetaan di Tableau, dengan 0 berfungsi sebagai placeholder bagi data tanpa ID valid sehingga seluruh data tetap dapat diproses dan divisualisasikan secara utuh.

#### 2.1.4 Modelling

##### Descriptive Analytics

Median confirmed	5,280
Recovered	47,777,100
STDEV	133,026
Variance	17,695,977,546

Analisis deskriptif menunjukkan bahwa jumlah kasus terkonfirmasi COVID-19 memiliki tingkat variasi yang sangat tinggi antar wilayah, dengan standar deviasi sebesar 133.026 dan variance sekitar  $1,77 \times 10^{10}$ . Nilai variasi yang besar ini mengindikasikan adanya ketimpangan distribusi kasus di antara berbagai region, di mana beberapa wilayah mencatat jumlah kasus yang sangat tinggi sementara wilayah lain relatif rendah. **Nilai** median sebesar 5.280 menunjukkan bahwa lebih dari setengah wilayah memiliki jumlah kasus di bawah angka tersebut, sehingga persebaran kasus tidak merata dan cenderung terpusat pada beberapa daerah dengan populasi besar atau aktivitas mobilitas tinggi.

Selain itu, total pasien sembuh mencapai sekitar 47,7 juta, yang menunjukkan tingkat pemulihan yang tinggi secara nasional. Angka ini mencerminkan efektivitas sistem penanganan kesehatan dan kapasitas medis di banyak wilayah Rusia dalam mengendalikan penyebaran serta meningkatkan kesembuhan pasien. Secara keseluruhan, hasil analisis ini memberikan gambaran bahwa pandemi berdampak tidak merata di berbagai wilayah, namun tingkat pemulihan yang tinggi menjadi indikator positif terhadap upaya mitigasi yang dilakukan.

#### 2.1.5 Evaluasi

Evaluasi dilakukan dengan meninjau kembali hasil analisis data COVID-19 di Rusia dan melihat sejauh mana temuan tersebut mendukung tujuan utama proyek, yaitu memahami pola penyebaran, tingkat keparahan, serta efektivitas penanganan pandemi di tiap wilayah. Berdasarkan hasil analisis deskriptif, terlihat bahwa jumlah kasus terkonfirmasi menunjukkan variasi yang sangat besar antar wilayah, dengan standar deviasi mencapai 133.026 dan varians sekitar  $1,77 \times 10^{10}$ . Nilai median kasus sebesar 5.280 menunjukkan bahwa sebagian besar wilayah memiliki jumlah kasus relatif rendah, meskipun ada beberapa wilayah yang mencatat lonjakan cukup tinggi.



Selain itu, total pasien sembuh yang mencapai sekitar 47,7 juta orang mencerminkan tingkat pemulihan yang tinggi secara nasional. Temuan ini memberikan gambaran bahwa penanganan pandemi di beberapa wilayah cukup efektif, meskipun masih terdapat ketimpangan dalam penyebaran dan pengendalian kasus di wilayah lainnya. Melalui visualisasi yang dibuat di Tableau, pola persebaran, tren kasus aktif, serta perbandingan tingkat kematian dan kesembuhan dapat diamati dengan lebih jelas dan informatif.

Secara keseluruhan, hasil analisis ini dianggap relevan dengan tujuan proyek karena mampu memberikan pemahaman yang lebih mendalam mengenai situasi pandemi di Rusia. Insight yang diperoleh dapat dimanfaatkan untuk mendukung pengambilan keputusan, seperti penentuan prioritas wilayah dengan kasus tinggi, evaluasi kebijakan kesehatan, serta perencanaan strategi yang lebih tepat dalam menghadapi potensi gelombang pandemi berikutnya.

## 2.1.6 Deployment

### 1. Persebaran Jumlah kasus tiap kota di Rusia



Ketiga peta ini menggambarkan persebaran geografis kasus COVID-19 di berbagai wilayah Rusia dengan tiga indikator utama, yaitu:

1. Persebaran Kasus Terkonfirmasi (peta kuning)
  - Menunjukkan jumlah total kasus positif COVID-19 di setiap wilayah.
  - Warna kuning semakin pekat pada wilayah dengan kasus terkonfirmasi yang tinggi.
  - Terlihat bahwa wilayah Moscow dan sekitarnya merupakan daerah dengan konsentrasi kasus tertinggi.
2. Persebaran Kasus Kematian (peta merah muda)
  - Memvisualisasikan jumlah kematian akibat COVID-19 di seluruh wilayah Rusia.

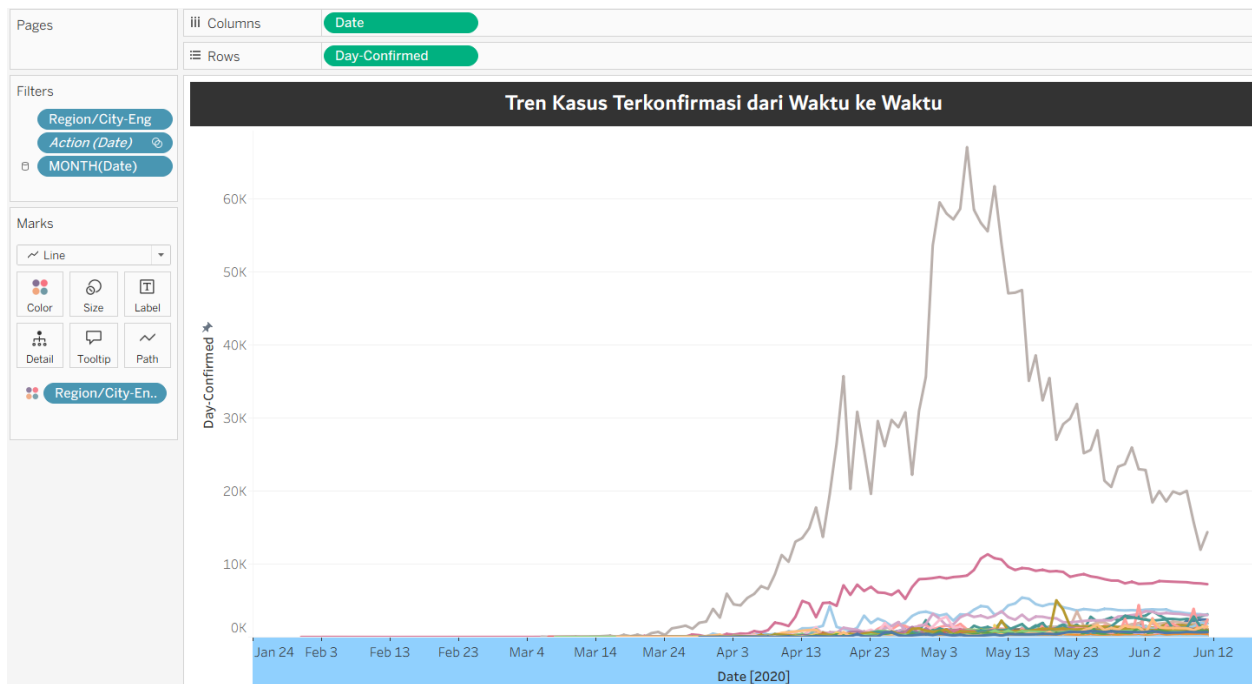
- Warna merah muda pekat menunjukkan wilayah dengan tingkat kematian lebih tinggi.
- Moscow kembali menjadi pusat dengan angka kematian tertinggi, sementara sebagian besar wilayah lainnya menunjukkan intensitas warna yang lebih ringan.

### 3. Persebaran Kasus Kesembuhan (peta hijau)

- Menggambarkan jumlah pasien yang telah dinyatakan sembuh di setiap wilayah.
- Warna hijau lebih pekat menandakan tingkat kesembuhan yang tinggi.
- Pola persebaran menunjukkan bahwa wilayah dengan kasus tertinggi seperti Moscow juga mencatat jumlah kesembuhan yang besar, menandakan efektivitas sistem kesehatan di wilayah tersebut.

Visualisasi ini mendukung tujuan analisis dengan memberikan gambaran yang jelas mengenai distribusi dan dampak pandemi di berbagai wilayah Rusia. Melalui ketiga peta tersebut, dapat dilihat daerah-daerah yang paling terdampak serta tingkat keberhasilan penanganan kasus di tiap wilayah. Informasi ini penting untuk menjawab pertanyaan penelitian mengenai wilayah dengan tingkat penyebaran tertinggi dan efektivitas penanganan pandemi, sekaligus menjadi dasar bagi pembuat kebijakan dalam menentukan prioritas penanganan serta pemerataan sumber daya kesehatan di seluruh wilayah Rusia.

### 2. Tren Kasus Terkonfirmasi dari Waktu ke Waktu Tiap Kota Dirusia



Grafik ini menampilkan perkembangan jumlah kasus COVID-19 yang terkonfirmasi setiap harinya di Rusia sepanjang periode awal pandemi tahun 2020. Beberapa hal penting yang terlihat dalam grafik ini antara lain:

1. Sumbu X (horizontal) menunjukkan rentang waktu, dimulai dari akhir Januari hingga pertengahan Juni 2020.
2. Sumbu Y (vertikal) menampilkan jumlah kasus harian yang terkonfirmasi (Day Confirmed) di seluruh wilayah Rusia.
3. Garis utama dengan nilai tertinggi memperlihatkan tren nasional atau wilayah dengan lonjakan kasus terbesar — yaitu Moscow (city).
4. Beberapa garis berwarna lain menggambarkan tren kasus di wilayah lain, dengan pola kenaikan dan penurunan yang lebih rendah.
5. Dari grafik ini tampak adanya lonjakan tajam pada awal Mei 2020, yang menjadi puncak penyebaran COVID-19 di Rusia, sebelum akhirnya menunjukkan penurunan yang bertahap hingga pertengahan Juni.

Visualisasi ini berperan penting dalam memahami tujuan pertama analisis, yaitu untuk melihat pola peningkatan dan penurunan kasus COVID-19 dari waktu ke waktu. Dengan grafik tren ini, dapat diidentifikasi periode kritis penyebaran virus, waktu terjadinya lonjakan kasus, serta fase di mana penyebaran mulai terkendali. Informasi tersebut sangat relevan untuk menilai efektivitas kebijakan pemerintah, kesiapan sistem kesehatan, serta respons publik terhadap upaya pencegahan selama pandemi berlangsung.

### 3. Perbandingan Kasus COVID-19: Confirmed vs Recovered vs Deaths



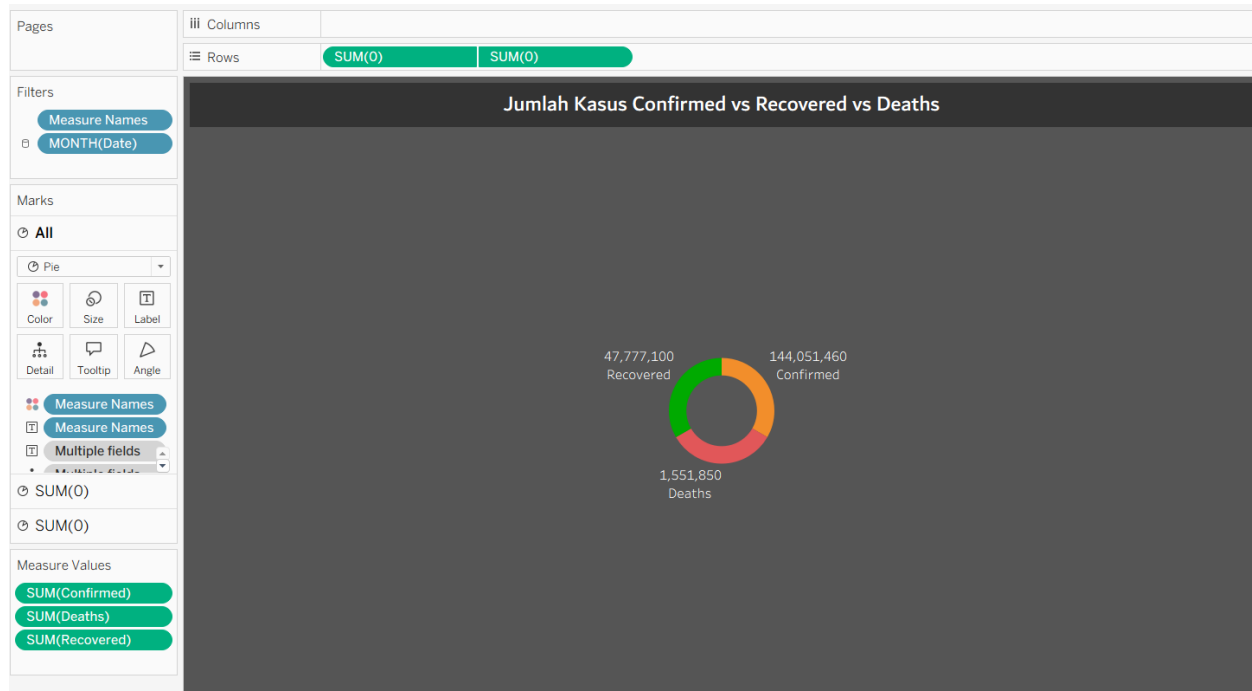


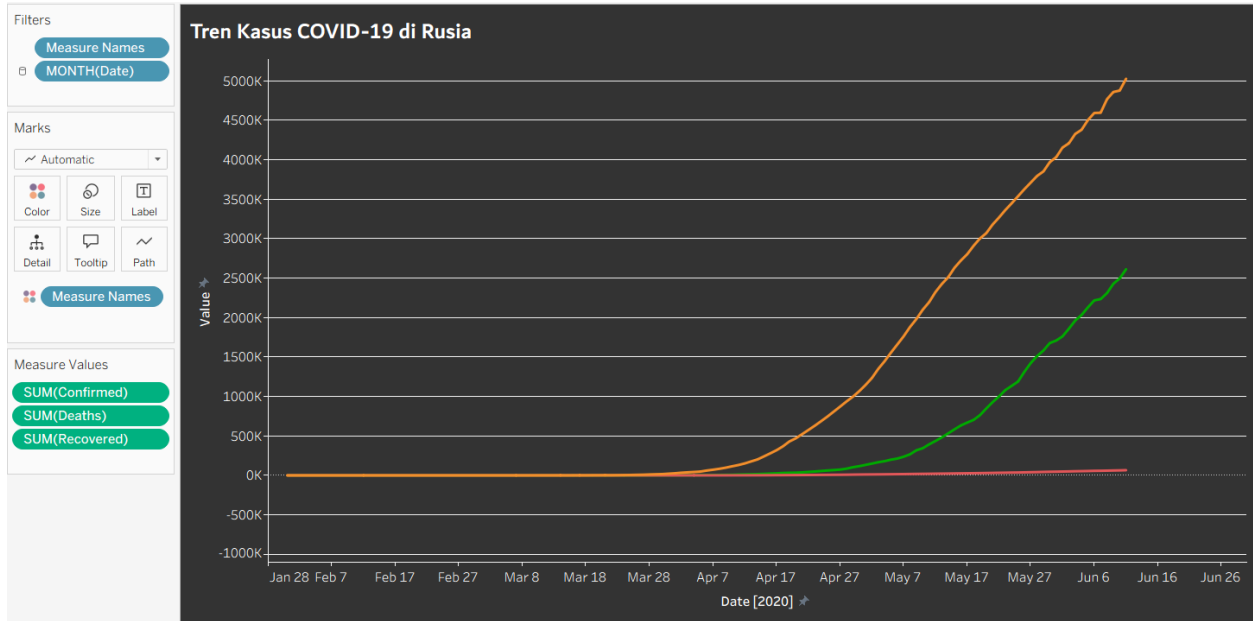
Diagram ini menggambarkan proporsi total dari tiga kategori utama:

- 144.051.460 kasus terkonfirmasi
- 47.777.100 pasien sembuh
- 1.551.850 kematian

Melalui tampilan visual berbentuk cincin, pengguna dapat dengan mudah memahami perbandingan skala antara jumlah kasus aktif, sembuh, dan meninggal. Nilai ini menunjukkan bahwa tingkat pemulihan di Rusia cukup tinggi, sementara angka kematian tergolong rendah dibandingkan total kasus keseluruhan.

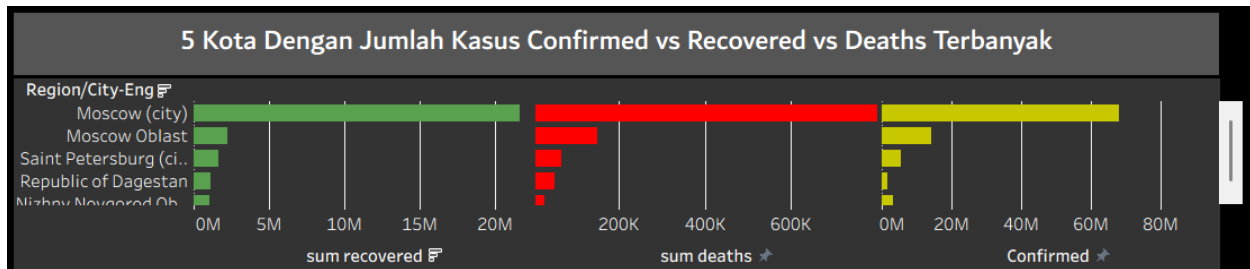
Visualisasi ini menggambarkan proporsi dan keseimbangan antara tiga indikator utama COVID-19. Data ini relevan untuk memahami tingkat keparahan pandemi dan efektivitas sistem kesehatan Rusia dalam menangani pasien. Perbandingan persentase juga membantu pembuat kebijakan untuk menilai rasio recovery rate dan fatality rate, yang menjadi tolok ukur penting dalam evaluasi penanganan pandemi.

#### 4. Tren Kasus COVID-19 di Rusia



Visualisasi ini menampilkan perkembangan jumlah kasus terkonfirmasi (confirmed), kematian (deaths), dan kesembuhan (recovered) dari waktu ke waktu. Terlihat jelas lonjakan kasus signifikan mulai April hingga Mei 2020, diikuti peningkatan cepat pada jumlah pasien sembuh. Sementara itu, grafik kematian relatif landai, menunjukkan bahwa sebagian besar kasus dapat ditangani dengan baik. Grafik ini berguna untuk melihat dinamika pandemi secara kronologis dan mengidentifikasi periode kritis penyebaran virus. Visualisasi ini mendukung tujuan pertama business understanding, yaitu menganalisis pola peningkatan dan penurunan kasus berdasarkan periode waktu tertentu. Dengan melihat tren kasus terkonfirmasi, sembuh, dan meninggal secara kronologis, pemerintah atau analis data dapat mengidentifikasi fase kritis pandemi, periode lonjakan, serta efektivitas kebijakan yang diterapkan pada waktu tertentu

## 5. 5 Kota Dengan Jumlah Kasus Confirmed vs Recovered vs Deaths Terbanyak



Visualisasi ini menampilkan perbandingan antara jumlah kasus terkonfirmasi (confirmed), sembuh (recovered), dan meninggal (deaths) akibat COVID-19 pada lima kota/wilayah dengan jumlah kasus terbanyak di Rusia. Grafik disajikan dalam bentuk batang horizontal dengan tiga warna berbeda, di mana:

- Kuning merepresentasikan jumlah kasus terkonfirmasi,
- Hijau menunjukkan jumlah pasien yang sembuh, dan
- Merah menggambarkan jumlah kematian.

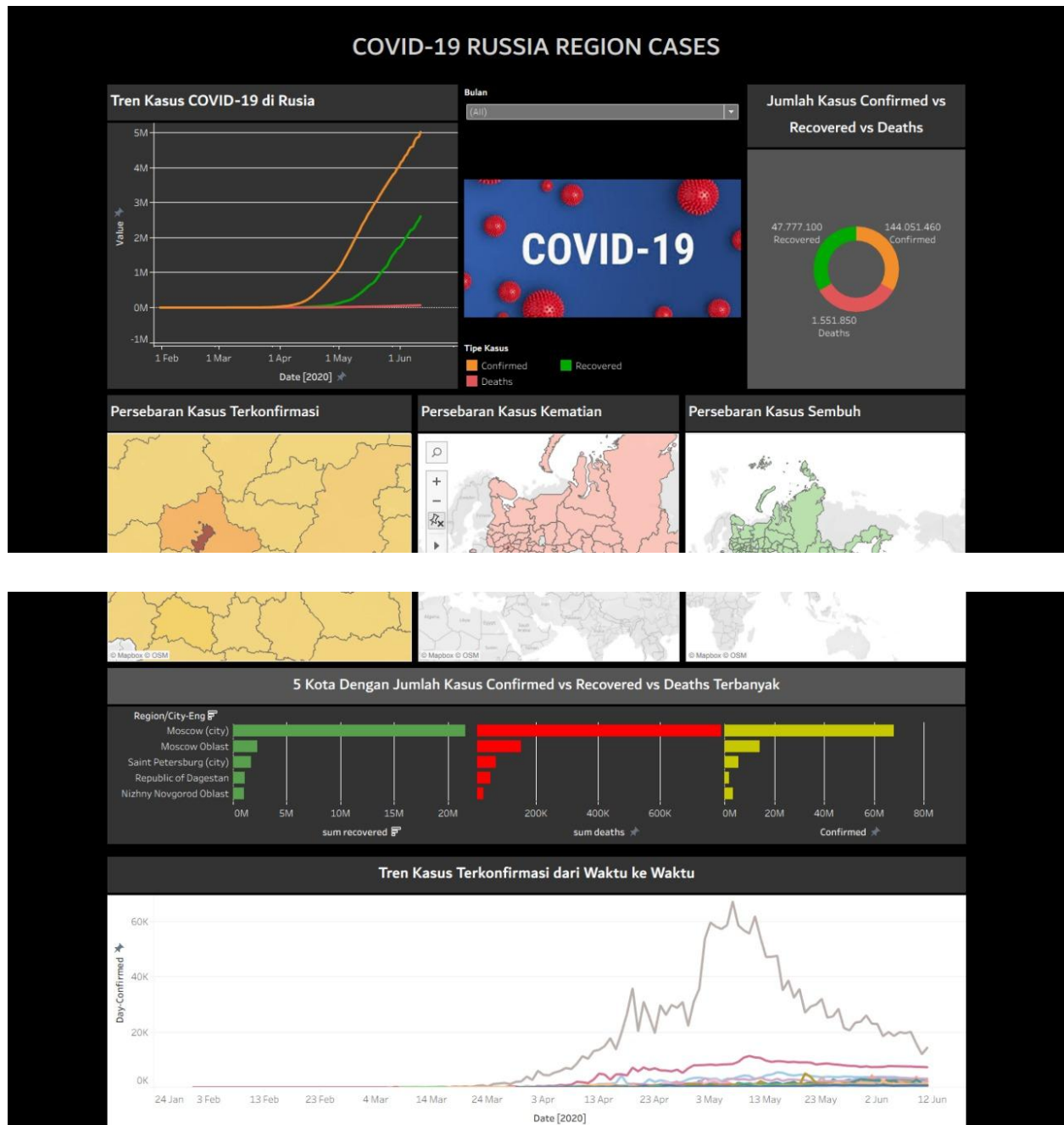
Dari grafik terlihat bahwa:

- Moscow (city) menempati posisi tertinggi di semua kategori, dengan lebih dari 20 juta kasus sembuh, sekitar 80 juta kasus terkonfirmasi, dan jumlah kematian tertinggi dibanding wilayah lain. Hal ini menunjukkan bahwa Moscow merupakan pusat utama penyebaran sekaligus penanganan pandemi di Rusia.
- Moscow Oblast berada di posisi kedua dengan jumlah kasus dan kesembuhan yang juga tinggi, diikuti oleh Saint Petersburg, Republic of Dagestan, dan Nizhny Novgorod Oblast.
- Meskipun jumlah kematian di wilayah-wilayah tersebut meningkat seiring tingginya jumlah kasus, tingkat kesembuhan secara umum juga menunjukkan hasil yang signifikan.
- Pola ini memperlihatkan bahwa wilayah dengan populasi besar dan aktivitas ekonomi tinggi juga memiliki angka kasus serta kesembuhan yang tinggi.

Visualisasi ini menegaskan bahwa konsentrasi penyebaran COVID-19 tertinggi berada di pusat-pusat urban, terutama di Moscow dan sekitarnya. Namun, tingginya angka kesembuhan di kota-kota besar tersebut juga menunjukkan efektivitas sistem layanan kesehatan dan penanganan pandemi yang relatif lebih baik dibanding wilayah lain.

Visualisasi ini relevan dengan tujuan analisis untuk mengidentifikasi wilayah dengan kasus COVID-19 tertinggi serta memahami hubungan antara jumlah kasus, tingkat kesembuhan, dan kematian. Temuan ini mendukung pemahaman dalam *business understanding* bahwa kota besar seperti Moscow berperan penting dalam dinamika penyebaran COVID-19 di Rusia. Selain itu, perbedaan signifikan antara wilayah utama dan daerah lain menunjukkan adanya kesenjangan fasilitas medis dan efektivitas penanganan, yang dapat menjadi dasar dalam penyusunan kebijakan pemerataan sumber daya kesehatan di masa mendatang.

## 6. Dashboard COVID-19



Dashboards ini terdiri dari lima visualisasi utama yang secara keseluruhan memberikan gambaran menyeluruh mengenai perkembangan dan distribusi kasus COVID-19 di Rusia. Setiap visualisasi memiliki fungsi analitis tersendiri, namun saling melengkapi untuk membantu memahami pola penyebaran, tingkat kematian, dan tingkat kesembuhan di berbagai wilayah.

Secara keseluruhan, dashboard ini berfungsi sebagai alat pemantauan, analisis, dan pengambilan keputusan berbasis data.

Beberapa manfaat utamanya:

- Memantau tren pandemi dan efektivitas penanganan dari waktu ke waktu.
- Mengenali wilayah prioritas dengan kasus tinggi atau tingkat kematian besar.
- Mendukung perencanaan kebijakan kesehatan publik, termasuk distribusi sumber daya medis.
- Menyediakan insight visual yang mudah dipahami, baik untuk analis data, pemerintah, maupun masyarakat umum.

Lima visualisasi dalam dashboard ini saling melengkapi untuk menggambarkan gambaran makro dan mikro pandemi COVID-19 di Rusia.

Tren kenaikan dan penurunan kasus, persebaran geografis, serta 5 Kota Dengan Jumlah Kasus Terbanyak memberikan pemahaman komprehensif mengenai situasi pandemi. Dashboard ini tidak hanya bernilai informatif, tetapi juga strategis sebagai dasar evaluasi kebijakan dan perencanaan penanggulangan pandemi ke depan.

## **BAB III**

### **Hasil Eksplorasi dan Analisis**

#### **3.1 Hasil Eksplorasi Data**

##### **3.1.1 Struktur dan Kualitas Data**

Dataset terdiri dari beberapa kolom utama seperti:

- Date: menunjukkan tanggal pencatatan kasus.
- Region/City dan Region/City-Eng: nama wilayah dalam bahasa Rusia dan terjemahannya dalam bahasa Inggris.
- Region\_ID: kode identitas tiap wilayah.
- Day-Confirmed, Day-Deaths, Day-Recovered: jumlah kasus harian yang terkonfirmasi, meninggal, dan sembuh.
- Confirmed, Deaths, Recovered: jumlah kumulatif kasus di tiap wilayah.

Selama proses eksplorasi ditemukan beberapa nilai kosong (missing values) pada kolom *Region\_ID* serta ketidaksesuaian nama wilayah akibat perbedaan *encoding* teks (UTF-8). Masalah ini menyebabkan sebagian data tidak terbaca saat divisualisasikan di Tableau. Untuk mengatasinya, data wilayah yang bermasalah diperbaiki secara manual di Excel sebelum diunggah kembali. Selain itu, nilai kosong pada *Region\_ID* diganti dengan 0 menggunakan *calculated field* IFNULL([Region ID], 0) agar seluruh data tetap dapat diolah tanpa error.

### 3.1.2 Statistik Dasar

Dari hasil analisis deskriptif diketahui bahwa:

- Jumlah kasus terkonfirmasi memiliki variasi sangat tinggi antar wilayah (STDEV = 133.026, Variance  $\approx 1,77 \times 10^{10}$ ).
- Median kasus sebesar 5.280, menunjukkan sebagian besar wilayah memiliki kasus relatif rendah dibandingkan beberapa wilayah dengan lonjakan tinggi seperti *Moscow City* dan *Moscow Oblast*.
- Total pasien sembuh mencapai 47,7 juta jiwa, menggambarkan tingkat pemulihan yang cukup tinggi secara keseluruhan.

## 3.2 Analisis

Hasil analisis data COVID-19 di wilayah Rusia divisualisasikan dalam bentuk dashboard interaktif yang terdiri dari lima visualisasi utama. Setiap visualisasi menggambarkan aspek berbeda dari persebaran dan perkembangan kasus, sehingga memberikan pemahaman yang komprehensif terhadap kondisi pandemi di berbagai wilayah.

### 1. Peta Persebaran Kasus Terkonfirmasi, Kematian, dan Sembuh

Tiga peta tematik menampilkan persebaran kasus berdasarkan kategori.

- Warna kuning menunjukkan tingkat kasus terkonfirmasi, dengan konsentrasi tertinggi di wilayah *Moscow City* dan sekitarnya.
- Warna merah memperlihatkan persebaran kasus kematian yang juga berpusat di area dengan jumlah kasus tertinggi.

- Warna hijau menunjukkan distribusi pasien sembuh, dengan beberapa wilayah menonjol dalam tingkat pemulihan.
  - Ketiga peta ini membantu mengidentifikasi pola geografis penyebaran COVID-19 serta tingkat keparahan di masing-masing daerah.
- 2. Tren Kasus Terkonfirmasi dari Waktu ke Waktu
 

Grafik garis ini menggambarkan perubahan jumlah kasus harian sepanjang tahun 2020. Terlihat adanya peningkatan signifikan mulai akhir Maret, mencapai puncaknya pada awal Mei, kemudian menurun secara bertahap hingga pertengahan Juni. Pola ini menggambarkan fase awal lonjakan kasus di Rusia, sejalan dengan periode awal pandemi global.
- 3. Perbandingan Kasus Kumulatif antar Wilayah (Bar Chart)
 

Diagram batang menampilkan wilayah dengan jumlah kasus tertinggi. *Moscow City* mendominasi secara signifikan dibandingkan wilayah lain seperti *Saint Petersburg* dan *Moscow Oblast*. Hal ini menegaskan bahwa pusat aktivitas ekonomi dan populasi padat menjadi faktor utama tingginya kasus.
- 4. Proporsi Kasus (Confirmed, Deaths, Recovered)
 

Visualisasi berbentuk *pie chart* menampilkan perbandingan total keseluruhan kasus terkonfirmasi, kematian, dan sembuh. Dari hasilnya terlihat bahwa sebagian besar kasus berakhir dengan kesembuhan, sementara proporsi kematian relatif kecil dibandingkan total kasus.
- 5. 5 Kota Dengan Jumlah Kasus Confirmed vs Recovered vs Deaths Terbanyak
 

Grafik batang ini memperlihatkan perbandingan jumlah kasus terkonfirmasi, sembuh, dan meninggal akibat COVID-19 di lima kota dengan kasus terbanyak di Rusia. Terlihat bahwa *Moscow (city)* mendominasi ketiga kategori, diikuti oleh *Moscow Oblast* dan *Saint Petersburg*. Pola ini menunjukkan bahwa wilayah perkotaan besar menjadi pusat utama penyebaran sekaligus penanganan pandemi di Rusia, dengan tingkat kesembuhan yang relatif tinggi seiring dengan besarnya jumlah kasus yang tercatat.

### 3.3 Rekomendasi Bisnis

Berdasarkan hasil eksplorasi dan analisis data COVID-19 di berbagai wilayah Rusia, terdapat beberapa langkah strategis yang dapat diambil baik oleh pemerintah, pelaku bisnis, maupun lembaga kesehatan untuk mendukung pengambilan keputusan berbasis data.

#### 1. Optimalisasi Alokasi Sumber Daya Kesehatan

- Fokuskan sumber daya medis di wilayah dengan kasus tinggi, terutama *Moscow City* dan sekitarnya, karena wilayah ini menunjukkan angka kasus terkonfirmasi dan kematian tertinggi.
- Tingkatkan kapasitas rumah sakit dan tenaga kesehatan di daerah berisiko tinggi untuk mencegah overload sistem kesehatan pada saat lonjakan kasus.
- Perluas fasilitas pemulihan dan karantina di wilayah dengan tingkat kesembuhan tinggi agar proses perawatan pasien dapat dilakukan lebih cepat dan efisien.

#### 2. Penguatan Kebijakan Pencegahan di Wilayah Padat Penduduk

- Berdasarkan pola persebaran, daerah metropolitan menjadi pusat penyebaran utama. Oleh karena itu, implementasi kebijakan pembatasan mobilitas dan kampanye kesadaran publik perlu difokuskan pada area ini.
- Digital tracking dan analitik mobilitas masyarakat dapat digunakan untuk memantau dan mengendalikan penyebaran secara lebih efektif.

#### 3. Perencanaan Ekonomi dan Bisnis

- Sektor bisnis dapat menggunakan informasi tren kasus untuk menentukan waktu dan lokasi optimal dalam melanjutkan operasi bisnis atau membuka cabang baru.
- Perusahaan ritel dan logistik dapat menyesuaikan strategi distribusi berdasarkan wilayah yang mulai pulih dan memiliki tingkat kesembuhan tinggi.
- Sektor pariwisata dan transportasi dapat merancang strategi pemulihan bertahap dengan fokus pada daerah yang menunjukkan stabilitas kasus rendah dan tingkat pemulihan tinggi.



#### 4. Pengembangan Sistem Monitoring Berkelanjutan

- Pemerintah dan lembaga analitik dapat mengembangkan dashboard interaktif berbasis waktu nyata (real-time) untuk terus memantau dinamika penyebaran kasus.
- Integrasi data multi-sumber (medis, ekonomi, mobilitas) akan meningkatkan akurasi prediksi dan membantu mitigasi risiko di masa depan.

#### 5. Strategi Komunikasi Publik dan Edukasi

- Berdasarkan tren kasus, masyarakat cenderung lebih patuh terhadap kebijakan kesehatan setelah puncak kasus terjadi. Oleh karena itu, komunikasi publik yang berkelanjutan dan berbasis data visual sangat penting untuk menjaga kesadaran dan kepatuhan.
- Edukasi terkait vaksinasi, kebersihan, dan tanggung jawab sosial harus disesuaikan dengan karakteristik tiap wilayah.

## **BAB IV**

### **Kesimpulan**

Analisis data COVID-19 di Rusia menunjukkan bahwa penyebaran kasus sangat bervariasi antar wilayah, dengan Moscow City menjadi pusat utama kasus tertinggi sekaligus memiliki tingkat kesembuhan yang besar. Tren kasus memperlihatkan lonjakan pada April–Mei 2020 sebelum menurun secara bertahap, menandakan efektivitas kebijakan pengendalian yang diterapkan.

Visualisasi dalam dashboard membantu menggambarkan pola persebaran, tren waktu, serta perbandingan antarwilayah, sehingga mempermudah pengambil keputusan dalam menentukan area prioritas dan strategi penanganan yang tepat.

Secara keseluruhan, proyek ini menegaskan bahwa pendekatan berbasis data sangat penting dalam memahami dinamika pandemi dan mendukung kebijakan yang lebih cepat, akurat, dan efektif di masa depan.

### **Lampiran**

(Tantowi & Siregar, 2025)Alwizain, T., Ectefania, I. M., & Hafifudin, M. (2023). *Visualisasi Data Menggunakan Tableau*. PT. Penerbit buku Pedia.  
<https://joiv.org/index.php/joiv/article/view/952/577>

Okwara Jerry Chizoba and Buba Abba Kyari. (2020). Global Journal of Engineering and Technology Advances. *Global Journal of Engineering and Technology Advances*, 05(02), 047–056. <https://doi.org/10.30574/gjeta>

Parhusip, H. A., Trihandaru, S., Heriadi, A. H., Santosa, P. P., & Puspasari, M. D. (2022). Data Exploration Using Tableau and Principal Component Analysis. *International Journal on Informatics Visualization*, 6(4), 911–920. <https://doi.org/10.30630/joiv.6.4.952>

Tantowi, R. P., & Siregar, S. D. (2025). *Covid-19 data visualization using tableau*. 13(1), 98–104.

(Okwara Jerry Chizoba and Buba Abba Kyari, 2020)

(Parhusip et al., 2022)(Alwizain et al., 2023)