



Analisis Epidemiologi Karakteristik Outcome Pasien COVID-19 di Seluruh Dunia dari tahun 2020-2023

Tema SDG's : Kehidupan Sehat dan Sejahtera

Latar Belakang Masalah

Penelitian ini juga bertujuan untuk menganalisis pola perubahan epidemiologi selama periode waktu yang diteliti, memahami apakah terjadi evolusi dalam keparahan penyakit, apakah terdapat peningkatan dalam pengetahuan dan perawatan medis terkait COVID-19, serta sejauh mana efektivitas vaksinasi dan kepatuhan terhadap protokol kesehatan publik dalam mengendalikan penyebaran virus.

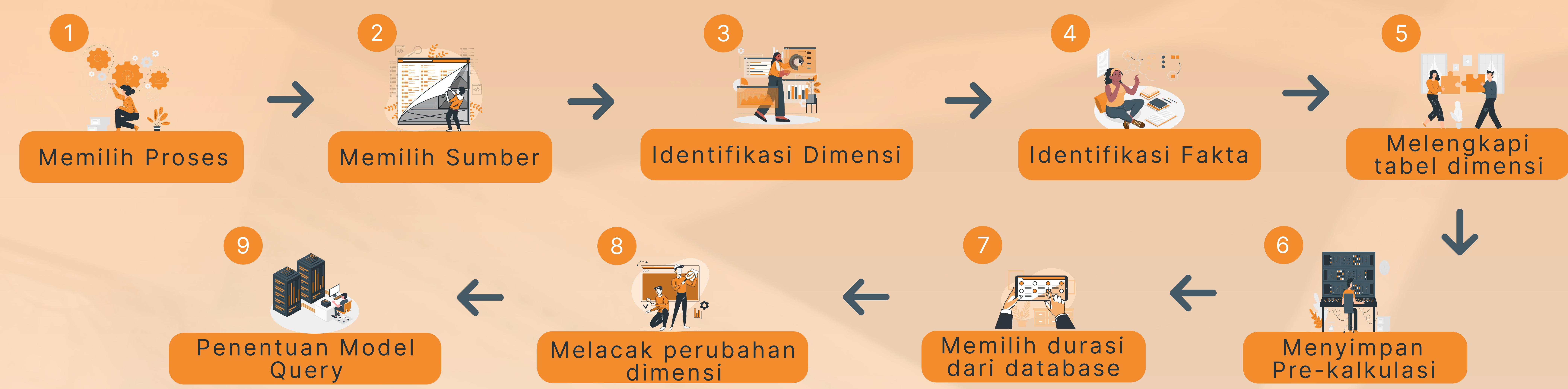
Tujuan

Memahami profil karakteristik pasien COVID-19, termasuk faktor-faktor yang memengaruhi tingkat keparahan penyakit dan hasil akhir pasien.

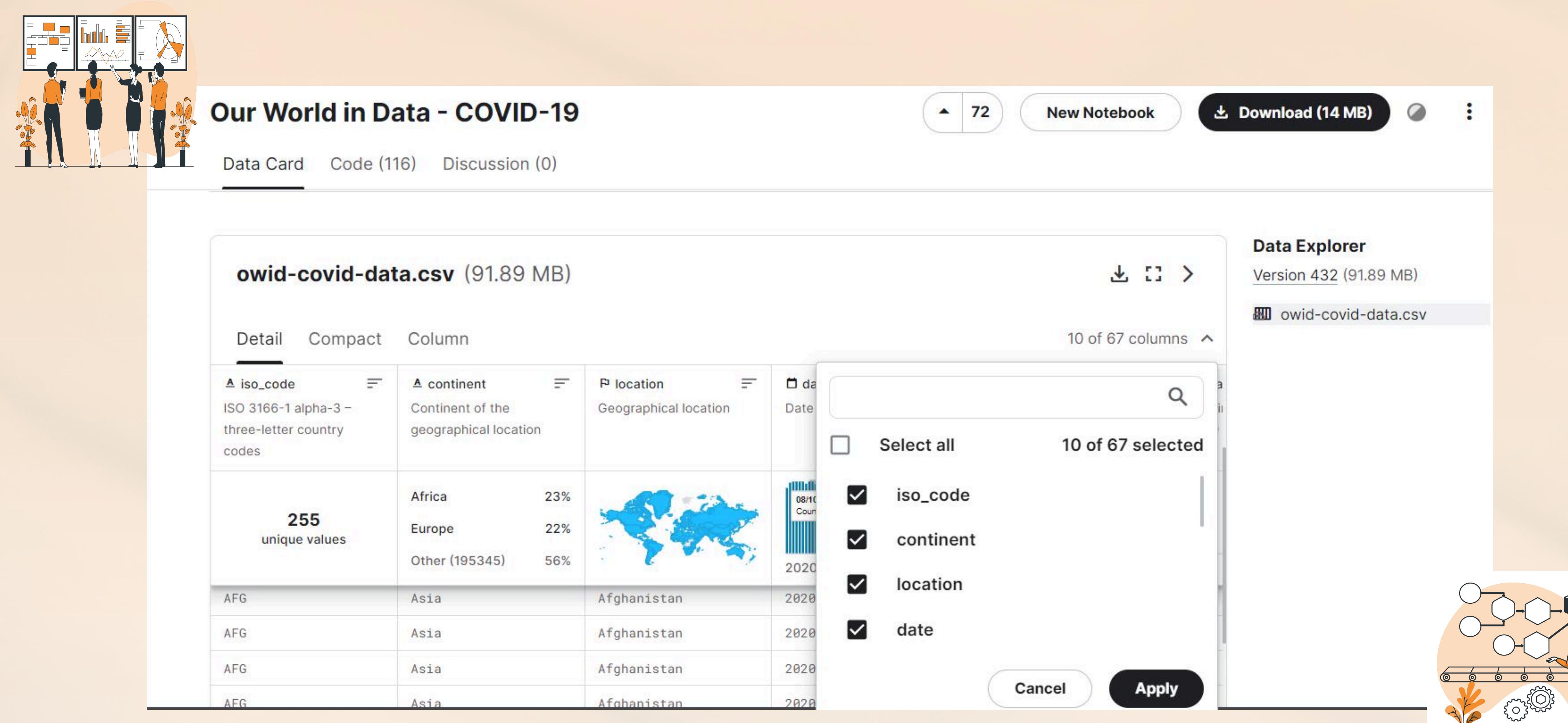


Metodologi Nine-Step Kimball

Merancang dan mengembangkan Data Warehouse dalam mengelola data pasien COVID-19 dari berbagai sumber dan menyajikannya secara terstruktur untuk analisis Epidemiologi yang lebih mendalam dari tahun 2020 hingga 2023.



Sumber Data : Kaggle (<https://www.kaggle.com/datasets/caesarmario/our-world-in-data-covid19-dataset/>) (01 Januari 2020-24 Oktober 2023)



Menerapkan Penggunaan Matriks untuk Identifikasi Fakta Berdasarkan Measurement dan Dimensi

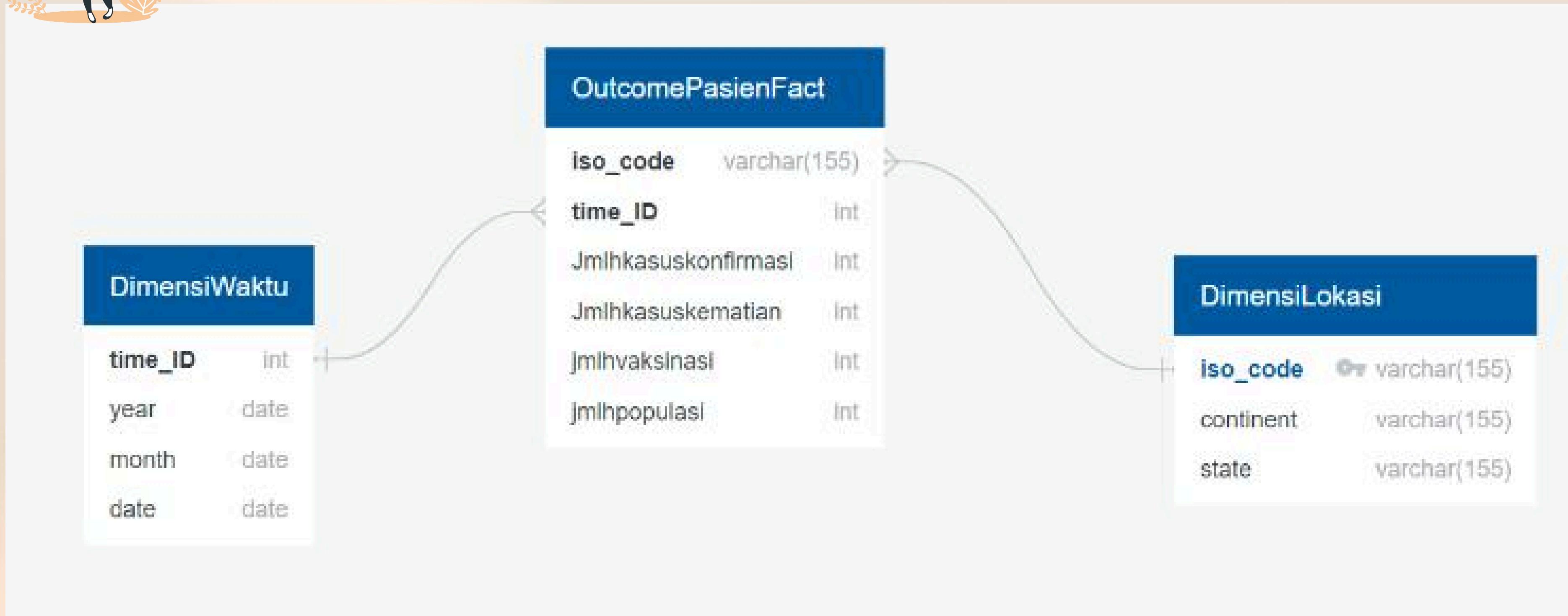
Measurement/Dimensi	Lokasi	Waktu(Bulan)
Jumlah kasus terkonfirmasi COVID 19	✓	✓
Jumlah kasus kematian COVID 19	✓	✓
Jumlah Vaksinasi COVID 19	✓	✓
Jumlah Populasi	✓	✓

Melengkapi Tabel Dimensi

Dimensi	Field	Deskripsi
Lokasi	continent	Data dapat dilihat berdasarkan benua
	state	Data dapat dilihat berdasarkan negara
Waktu (tahun, bulan, dan tanggal)	Date	Data dapat dilihat berdasarkan tanggal, bulan dan tahun

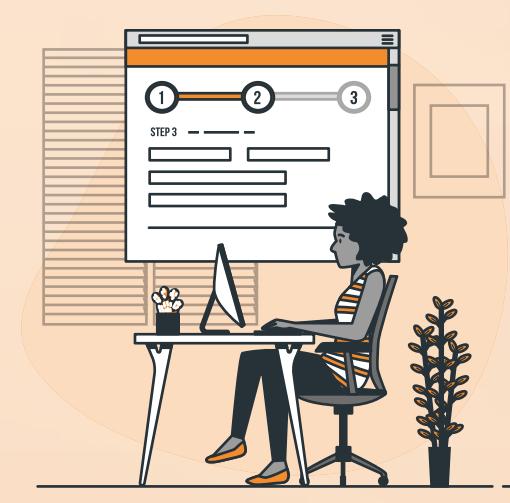


Rancangan Desain Model Star Schema



Keterangan

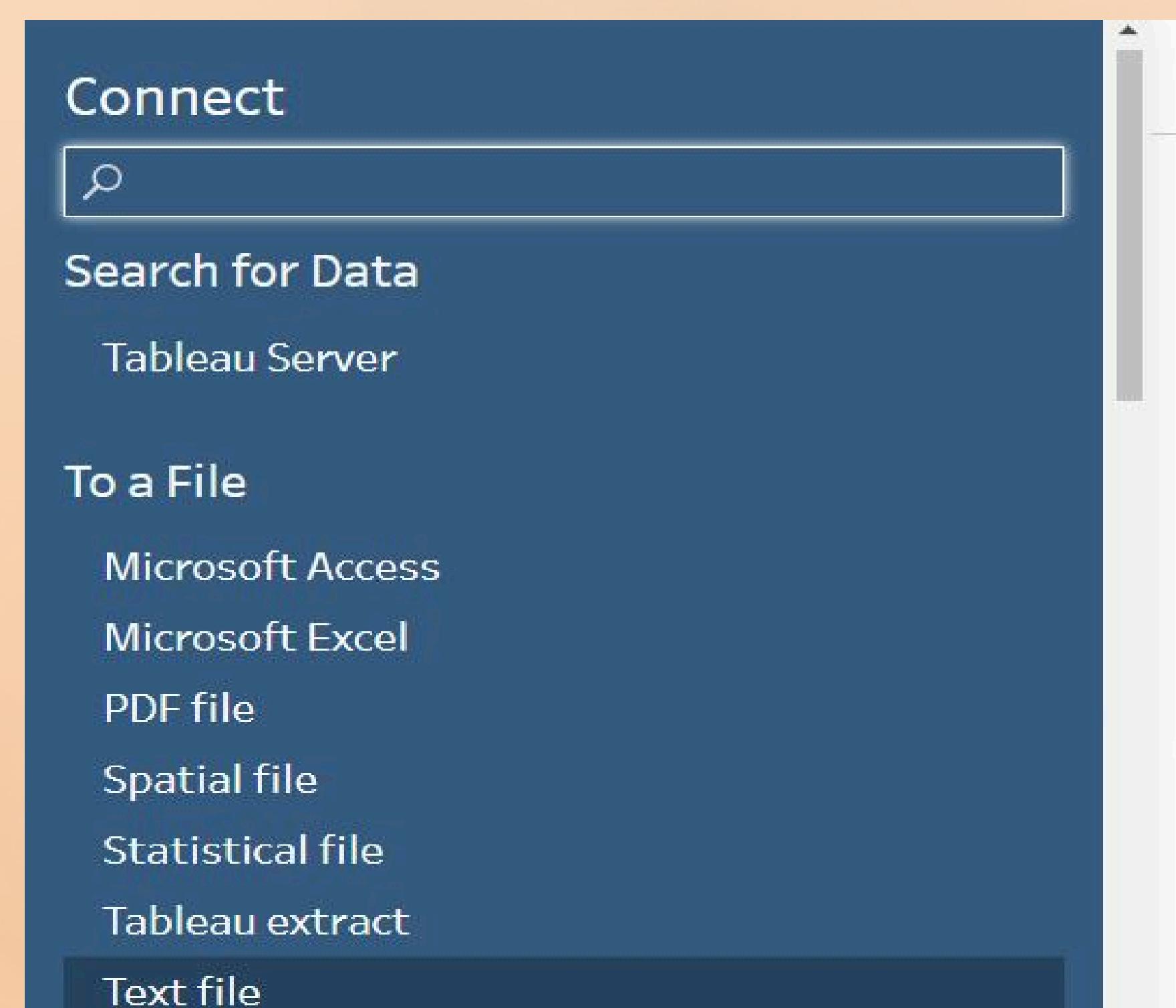
1. **OutcomePasienFact:** berisi berisi data numerik yang dapat diukur atau dihitung dari informasi terkait dengan hasil pasien. Tabel ini memiliki foreign key (FK) yang terhubung ke dua tabel dimensi.
2. **DimensiLokasi:** berisi informasi terkait lokasi. Ada kolom iso_code yang merupakan primary key dan terhubung sebagai foreign key pada kolom iso_code di OutcomePasienFact.
3. **DimensiWaktu:** berisi informasi terkait waktu. Tabel ini memiliki kolom time_ID yang menjadi foreign key pada OutcomePasienFact.



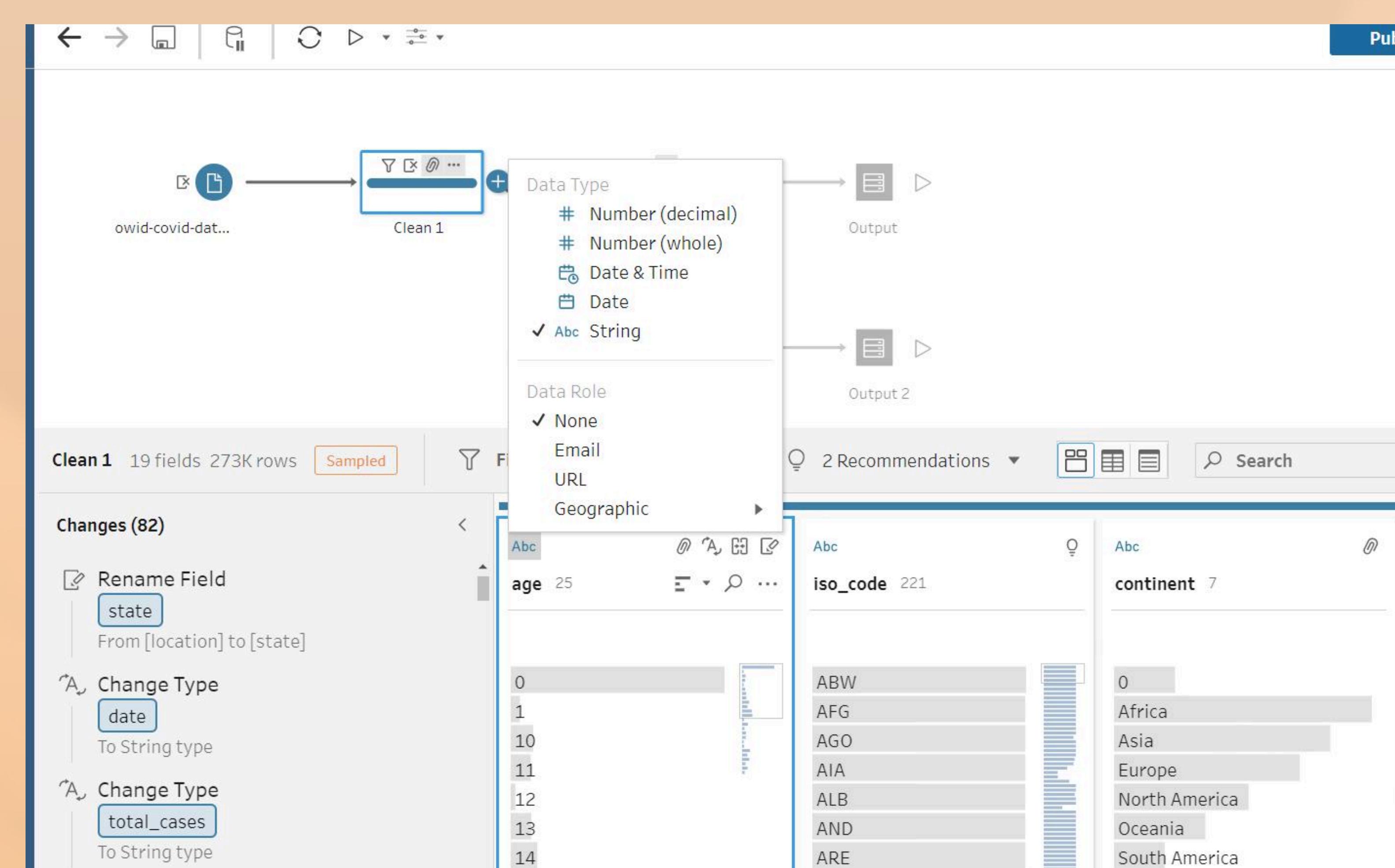
Pembuatan ETL

Proses menggabungkan data dari berbagai sumber ke dalam repositori pusat yang besar disebut gudang data guna untuk membuat profil, mengaudit, dan membersihkan data, memastikan bahwa data tersebut dapat dipercaya.

Masukkan sumber data ke tableau prep untuk melakukan pembersihan data



Membersihkan data di dalam data warehouse untuk memastikan kualitas data yang tepat dan terpercaya, menghilangkan inkonsistensi serta anomali untuk mendukung pengambilan keputusan yang akurat dan efektif.



Hasil Analisis (Visualisasi Data)

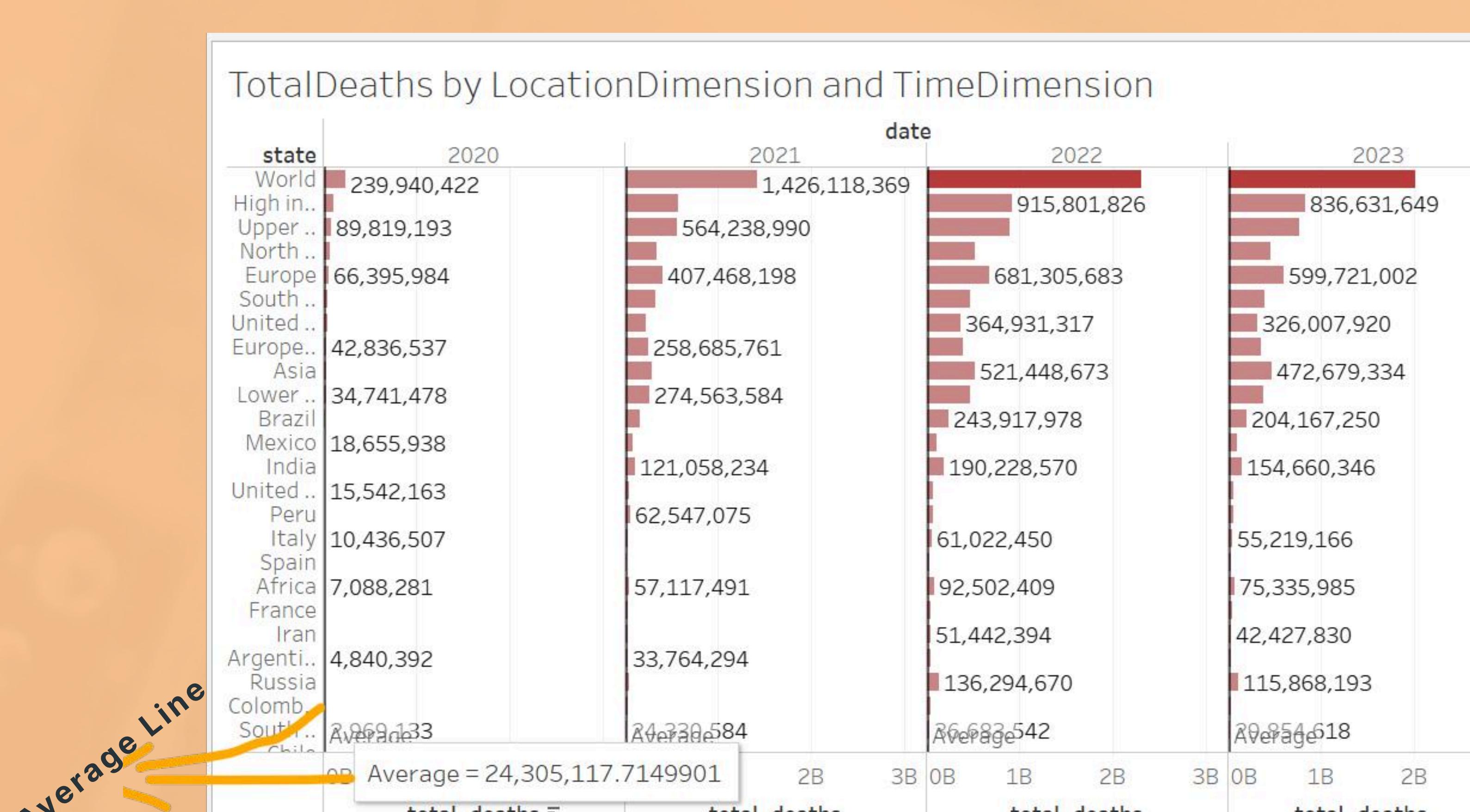


Menunjukkan variasi yang signifikan dalam jumlah kasus terkonfirmasi dan kematian di berbagai negara dari tahun 2020 hingga 2023. Ada perbedaan yang mencolok dalam tren kasus terkonfirmasi dan kematian, serta faktor-faktor yang mempengaruhinya antara negara-negara yang diselidiki. Temuan utama meliputi jumlah kasus kematian dan terkonfirmasi tertinggi di Amerika Serikat (AS) dengan kecenderungan penurunan pada tahun 2023, jumlah populasi yang padat, serta jumlah vaksinasi yang tinggi di China yang dapat berkontribusi pada pengurangan jumlah kasus.

Analisis Deskriptif (2020-2023)

Proses menganalisis data untuk mendeskripsikan, merangkum, dan meringkas informasi yang ada dalam data. Ini melibatkan teknik statistik dan visualisasi data untuk memahami pola, tren, dan karakteristik dari data yang telah terjadi.

TotalDeaths by LocationDimension and TimeDimension



continent: North America
state: United States

2020

continent: North America
state: United States
Year of date: 2020
total_deaths: 47,048,156

2021

continent: North America
state: United States
Year of date: 2021
total_deaths: 223,942,585

2022

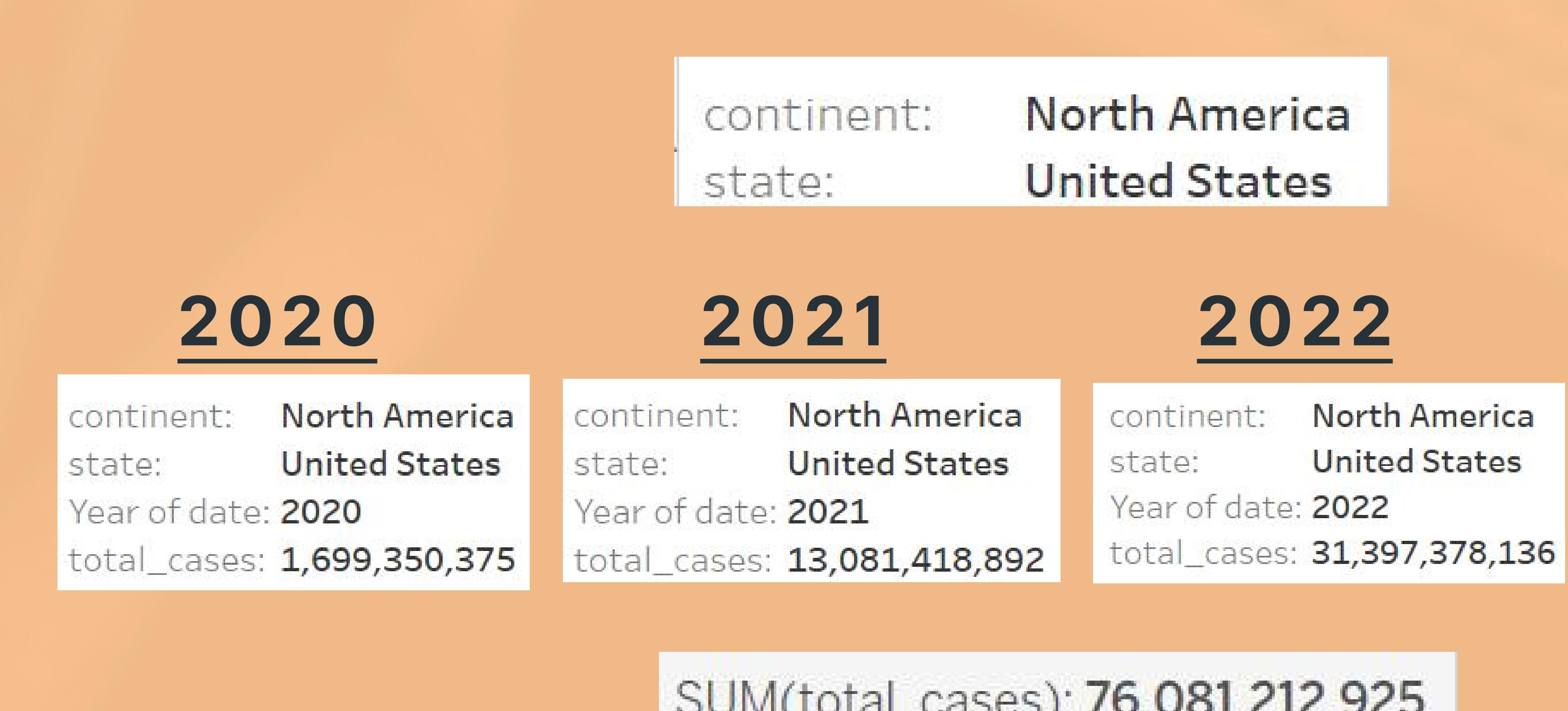
continent: North America
state: United States
Year of date: 2022
total_deaths: 364,931,317

2023

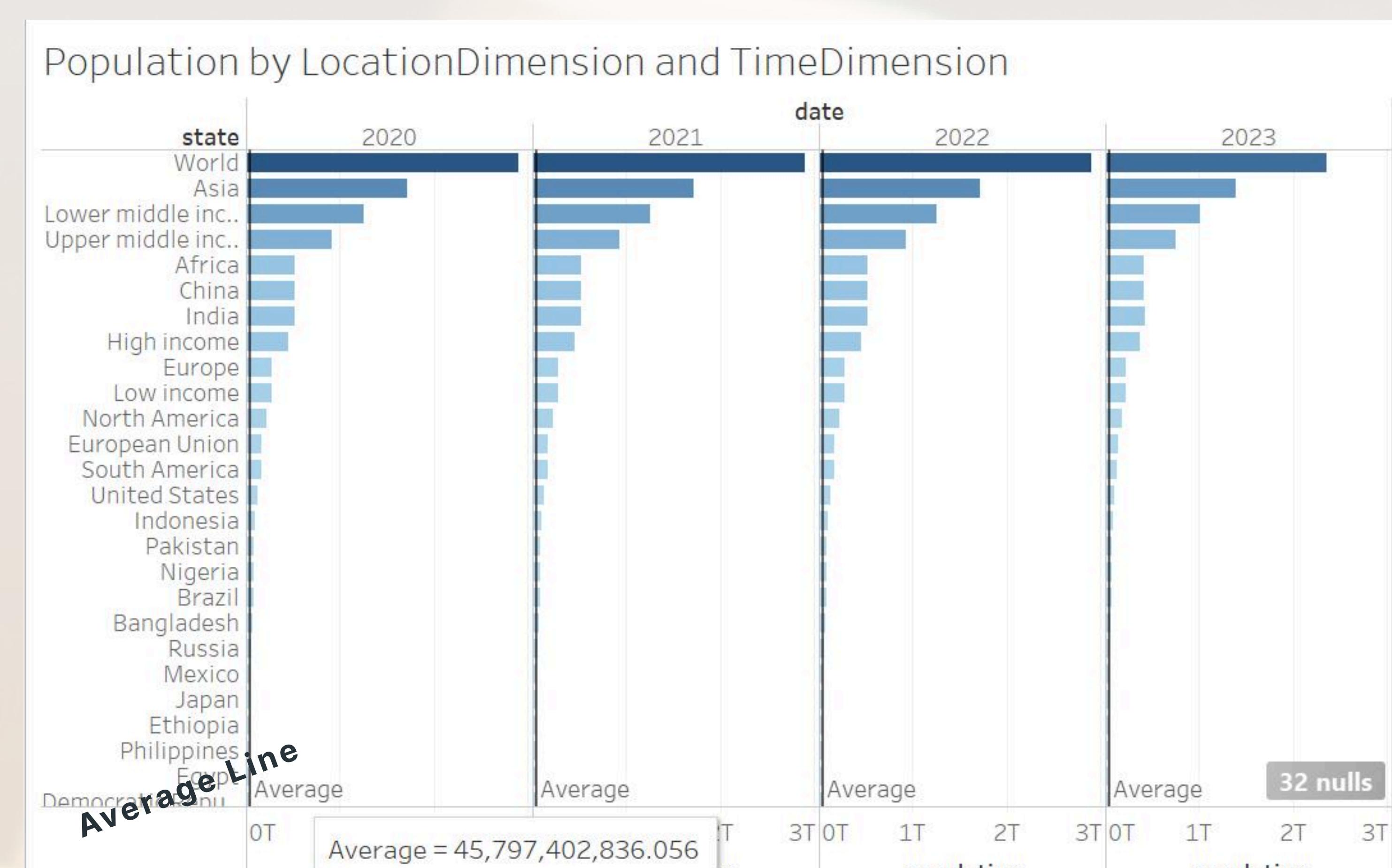
continent: North America
state: United States
Year of date: 2023
total_deaths: 326,007,920

SUM(total_deaths): 961,929,978

TotalCases by LocationDimension and TimeDimension



Populations by LocationDimension and TimeDimension



continent: Asia
state: China

2020

continent: Asia
state: China
Year of date: 2020
population: 519,022,999,040

2021

continent: Asia
state: China
Year of date: 2021
population: 520,448,886,400

2022

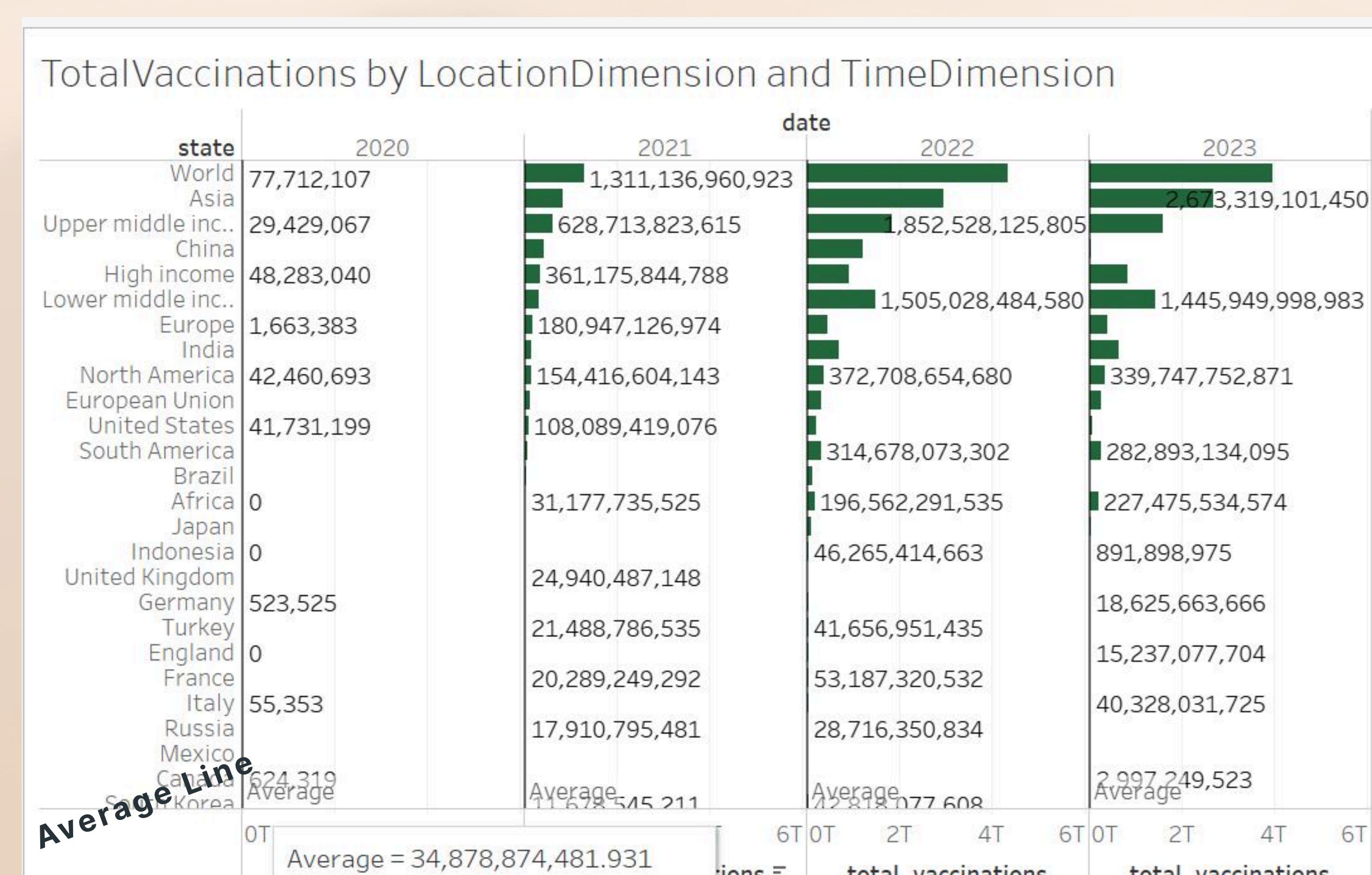
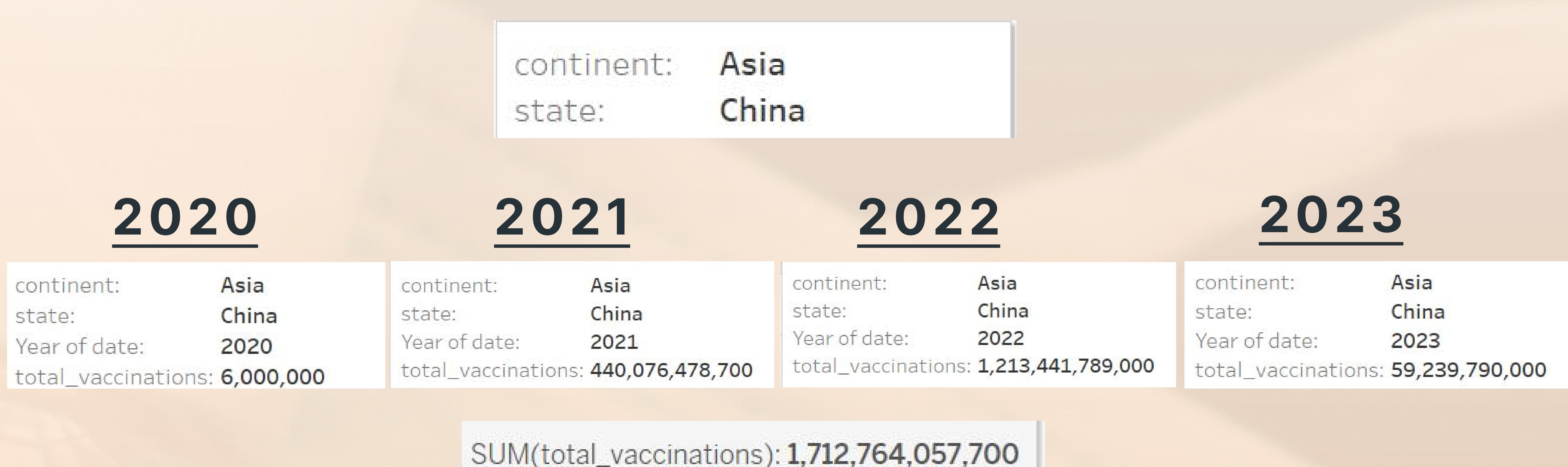
continent: Asia
state: China
Year of date: 2023
population: 414,933,221,760

2023

continent: Asia
state: China
Year of date: 2023
population: 414,933,221,760

SUM(population): 1,974,853,993,600

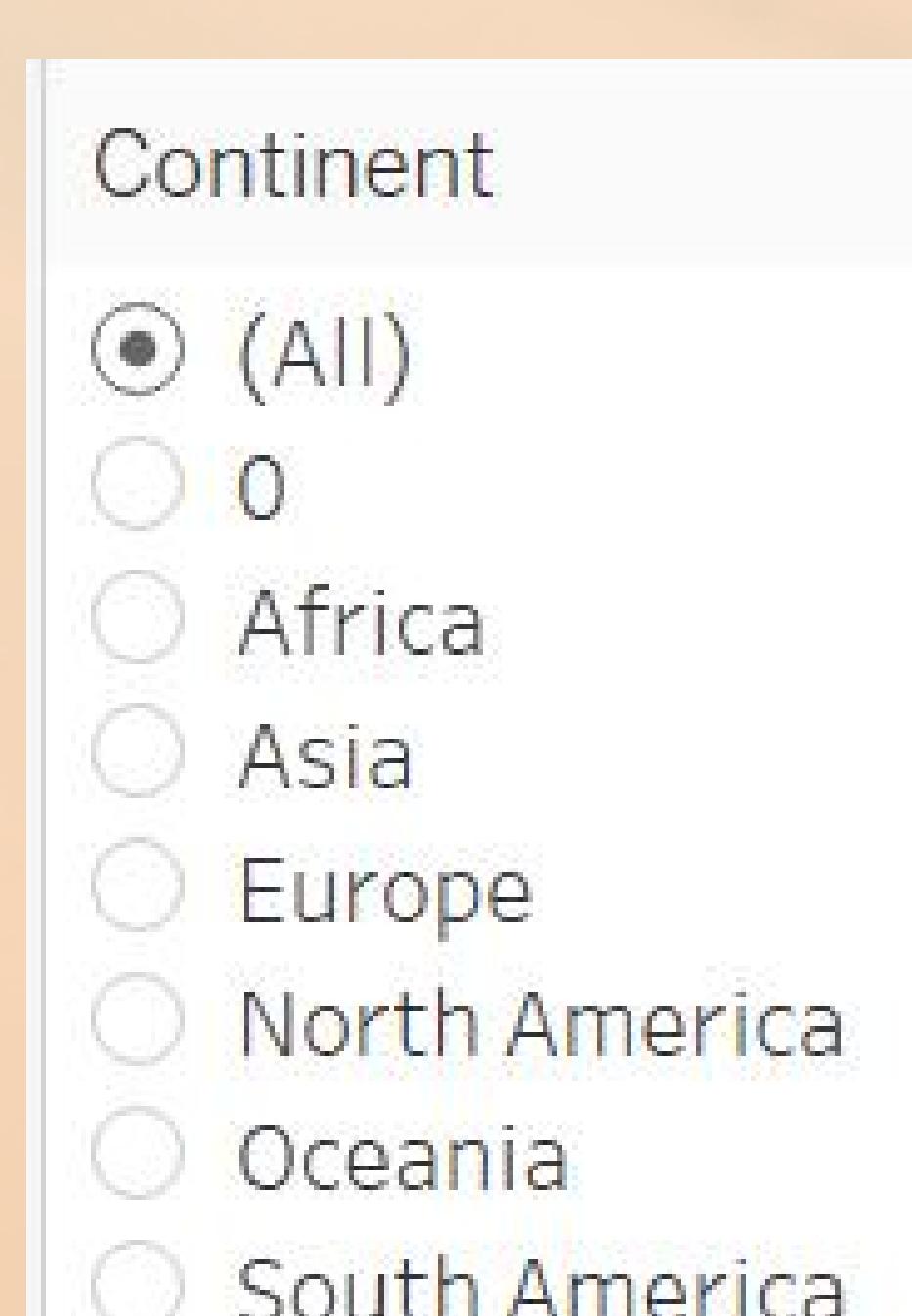
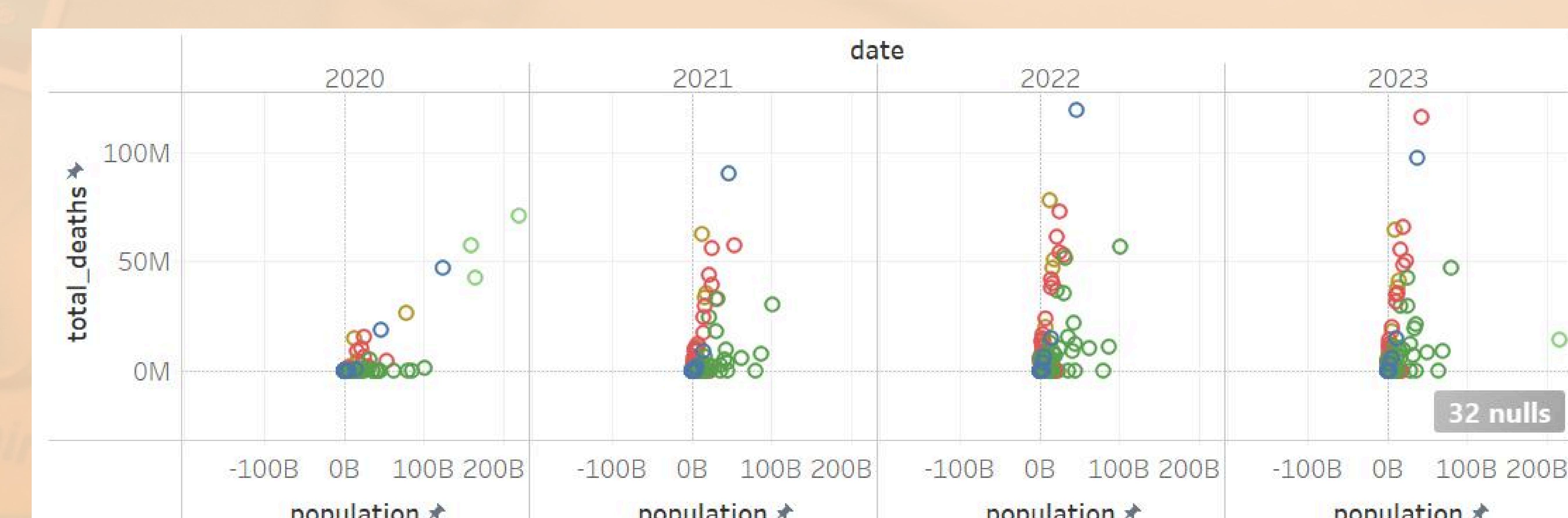
Total Vaccinations by LocationDimension and TimeDimension



Analisis Prediktif (Clustering)

Langkah dalam analisis data yang menggunakan metode statistik dan machine learning untuk membuat prediksi tentang peristiwa yang akan datang atau hasil yang mungkin terjadi.

Date dan Populations Terhadap TotalDeaths



Summary Diagnostics

Number of Clusters: 7
Number of Points: 982
Between-group Sum of Squares: 805.34
Within-group Sum of Squares: 5.9052
Total Sum of Squares: 811.25

Clusters	Number of Items	Centers	Sum of population	Sum of total_deaths	Most Common continent
Cluster 1	156	5.3313e+09	8.6524e+06	North America	
Cluster 2	201	3.2562e+10	6.1291e+06	Asia	
Cluster 3	216	5.1988e+09	7.6884e+06	Europe	
Cluster 4	84	7.4285e+08	1.0937e+05	Oceania	
Cluster 5	225	8.7788e+09	6.7491e+05	Africa	
Cluster 6	48	7.048e+11	3.9589e+08	0	
Cluster 7	52	1.1639e+10	2.3832e+07	South America	
Not Clustered	32				



- Cluster 1 - North America**
Populasi Rendah & Kematian Rendah
- Cluster 2 - Asia**
Populasi Tinggi & Kematian Tinggi
- Cluster 3 - Europe**
Populasi Rendah & Kematian Rendah
- Cluster 4 - Oceania**
Populasi Rendah & Kematian Rendah
- Cluster 5 - Africa**
Populasi Tinggi & Kematian Rendah
- Cluster 6-0**
Populasi dan Kematian yang Sangat Tinggi
- Cluster 7 - South America**
Populasi dan Kematian Sedang

Kategori Rendah (Low): Wilayah dengan jumlah populasi dan kematian yang rendah.

Contohnya: Cluster 1 (North America) dan Cluster 4 (Oceania).

Kategori Menengah (Medium): Wilayah dengan populasi sedang dan kematian bervariasi.

Contohnya: Cluster 7 (South America).

Kategori Tinggi (High): Wilayah dengan jumlah populasi yang besar dan tingkat kematian yang cenderung tinggi.

Contohnya: Cluster 2 (Asia) dan Cluster 5 (Africa).

Kategori Sangat Tinggi (Very High): Wilayah dengan populasi dan kematian sangat tinggi.

Contohnya: Cluster 6.

Kategori Rendah Kematian (Low Death): Wilayah dengan populasi besar namun angka kematian rendah.

Contohnya: Cluster 3 (Europe).

Kesimpulan

Dari hasil analisis di atas di harapkan untuk tindak lanjut pemerintah agar seluruh negara bisa dalam **prinsip wilayah dengan jumlah populasi stabil dan tingkat kematian yang rendah**. Diperlukan tindakan pencegahan yang kuat, perencanaan respons krisis yang efektif, serta penguatan infrastruktur kesehatan untuk mengurangi angka kematian yang tinggi, pengurangan pertumbuhan penduduk, dan meningkatkan pemerataan pembangunan untuk menanggulangi dampak pandemi ini.

Rekomendasi

- Peningkatan strategi vaksinasi
- Implementasi kebijakan kesehatan yang ketat
- Penguatan sistem pelacakan dan pengujian
- Dukungan yang lebih besar terhadap upaya masyarakat untuk mematuhi protokol kesehatan yang telah ditetapkan

