



## Visualisasi Data Program Vaksinasi Covid-19 di Kota Depok dengan Big Data Analytics

Rizki Elisa Nalawati\*, Dewi Yanti Liliana

Jurusan Teknik Informatika dan Komputer, Teknik Informatika, Politeknik Negeri Jakarta, Depok, Indonesia

Email: <sup>1,\*</sup>rizkielisa@tik.pnj.ac.id, <sup>2</sup>dewiyanti.liliana@tik.pnj.ac.id

Email Penulis Korespondensi: rizkielisa@tik.pnj.ac.id

**Abstrak**—Indonesia dan negara-negara di berbagai belahan dunia sedang menghadapi masalah pandemi Covid-19. Di Indonesia suspect Covid-19 ditemukan pada bulan Maret 2020 di Kota Depok, Jawa Barat. Sampai bulan februari 2021 jumlah pasien positif covid berjumlah 1.527.524. Kebutuhan pengawasan terhadap pemberian vaksin dilakukan oleh Dinas Kesehatan Kota Depok terhadap sejumlah fasilitas kesehatan yang dipercaya melakukan pemberian vaksin kepada masyarakat. Pengawasan ini dilakukan dengan menggunakan surveilans melalui visualisasi data pemberian vaksin dari masyarakat. Sejauh ini jumlah pemberian vaksin pada seluruh fasilitas kesehatan di Kota Depok dari bulan Januari sampai Agustus ada sekitar 613276 kali yang mencakup pemberian vaksin dosis 1, dosis 2 dan dosis 3. Besarnya data yang ada dapat dikelola dan divisualisasi dengan baik menggunakan *big data analytics*. Untuk mendapatkan bentuk dan visualisasi yang baik dalam pengambilan keputusan dilakukan beberapa proses cleansing data sampai pada tahap visualisasi. Penggunaan *big data analytics* dapat digunakan untuk visualisasi data deskriptif yang mampu menggambarkan ritme pemberian vaksinasi di Kota Depok, pengkategorian penerima vaksin, jenis vaksin yang diberikan sampai jumlah dosis yang diberikan. Sehingga dapat diperkirakan bahwa setiap bulan, penerima vaksin akan terus mengalami kenaikan, baik yang menerima dosis 1, dosis 2 dan dosis 3. Hal ini selaras dengan target pemerintah Kota Depok yang akan menuntaskan pemberian vaksin pada masyarakat Depok sampai dengan akhir tahun 2021.

**Kata Kunci:** Visualisasi Data; *Big Data Analytics*; Vaksinasi; Kota Depok; Covid-19

**Abstract**—Indonesia and various country in the world are facing the problem of the Covid-19 pandemic. In Indonesia, the suspect Covid-19 was found in March 2020 in Depok City, West Java. Until February 2021, the number of positive COVID patients was 1,527,524. The need for supervision of the administration of vaccines is carried out by the Depok City Health Office on a number of health facilities that are trusted to administer vaccines to the public. This supervision is carried out using surveillance through visualization of vaccine administration data from the community. So far, the number of vaccines given to all health facilities in Depok City from January to August is around 613276 times which includes the administration of dose 1, dose 2 and dose 3. The amount of existing data can be managed and visualized properly using big data analytics. To get a good shape and visualization in decision making, several data cleansing processes are carried out up to the visualization stage. The use of big data analytics can be used to visualize descriptive data that is able to describe the rhythm of vaccination in Depok City, categorization of vaccine recipients, the type of vaccine given to the number of doses given. So it can be estimated that every month, vaccine recipients will continue to increase, both receiving dose 1, dose 2 and dose 3. This is in line with the Depok government's target which will complete the provision of vaccines to the people of Depok by the end of 2021

**Keywords:** Visualization; Big Data Analytics; Vaccination; Covid-19

### 1. PENDAHULUAN

Program vaksinasi di Indonesia sudah dimulai dan menjangkau seluruh lapisan masyarakat mulai dari pejabat, tokoh agama, organisasi, birokrat pemerintahan sampai dengan lansia. Tujuan program vaksinasi ini adalah tercapainya Herd Immunity. Herd Immunity merupakan kondisi ketika sebagian besar orang dalam suatu kelompok telah memiliki kekebalan terhadap penyakit infeksi tertentu. Semakin banyak orang yang kebal terhadap suatu penyakit, semakin sulit bagi penyakit tersebut untuk menyebar karena tidak banyak orang yang dapat terinfeksi [1]. Pelaksanaan vaksinasi COVID-19 sudah dimulai pada 13 Januari 2021, adapun vaksin pertama diberikan pada Presiden RI Joko Widodo yang juga diterima oleh tokoh-tokoh pejabat lainnya. Dalam pelaksanaannya pemerintah Indonesia akan terus memastikan bahwa program vaksinasi berjalan dengan baik dan juga terus mengikuti hasil uji klinis vaksin Covid-19 diberbagai belahan dunia dengan beragam vaksin

WHO telah menyampaikan enam prinsip dasar agar proses imunisasi/ vaksinasi dapat berjalan dengan baik. Hal ini sudah menjadi dasar dan dicantumkan pada *Guidance on Developing a National Deployment and Vaccination Plan for Covid-19 vaccines* [2]. Prinsip-prinsip yang ada akan menjadi acuan bagi negara-negara yang akan memulai vaksinasi. Prinsip pertama adalah terjaminnya proses perencanaan vaksinasi nasional, berdasarkan kepemimpinan program imunisasi yang kuat, keputusan yang selalu berbasis ilmu pengetahuan, akuntabel, dan mampu bekerja sama dengan komponen kesehatan lainnya. Indonesia sudah sejak lama mempunyai kebijakan vaksin nasional dengan menerapkan vaksin-vaksin dini untuk mencegah terjadinya penyakit yang bisa dicegah melalui vaksin. Program vaksinasi Covid-19 telah dijadikan program vaksinasi nasional guna mencegah virus Covid-19 semakin meluas di seluruh penjuru negeri.

Prinsip kedua mencerminkan program imunisasi nasional yang berfungsi baik dan reponsif terhadap kemungkinan perkembangan di lapangan. Seperti yang kita ketahui antusiasme masyarakat terhadap vaksinasi Covid-19 begitu besar. Fasilitas- Terakhir adalah prinsip keenam dimana pilihan vaksin yang tepat untuk suatu negara. Tentunya vaksin yang dipilih harus aman dan tinggi efektifitasnya. Selanjutnya efek samping yang diberikan juga menjadi salah satu kriteria dalam memilih vaksin ini. Di Indonesia ada beberapa pilihan vaksin



yang siap untuk diberikan kepada masyarakat diantaranya Sinovac, Oxford-AstraZeneca, Sinopharm, Moderna, Novavax dan lain sebagainya.

Pada kondisi yang sama, visualisasi data dilakukan untuk melakukan pengawasan (surveilans) dan analisis impact dari Covid-19 di China pada banyak industri. Fasilitas pelayanan vaksinasi selalu dipenuhi oleh masyarakat yang ingin menerima vaksin. Hal ini karena 80 % masyarakat Indonesia antusias dengan vaksin Covid-19 [3].

Prinsip ketiga merupakan jaminan berjalannya lima kegiatan vaksinasi dasar mencakup tersedianya SDM yang terlatih dan termotivasi, komunikasi publik yang tepat serta sistem distribusi lainnya, yang keempat adalah jaminan cara pemberian vaksinasi yang aman dan pengelolaan efek samping serta monitoring dan evaluasi bermutu tinggi. Hal ini terlihat dari cara pemerintah menyiapkan SDM yang handal dalam melayani vaksin terhadap masyarakat. Selain itu rangkaian prosedur untuk vaksinasi juga dijalankan dengan baik.

Selanjutnya adalah prinsip kelima yang merupakan jaminan ketersediaan SDM dan anggaran yang memadai. Hal ini dapat terlihat dari besarnya anggaran pemerintah untuk mendukung program vaksinasi Covid-19 ini.

Pemodelan terhadap data yang ada dapat menggambarkan kondisi yang ada pada saat itu sehingga keterkaitan antara atribut bisa diketahui [4]. Menurut (James D Miller, 2017) Data visualization merupakan era baru yang merupakan sumber yang muncul dari kecerdasan, perkembangan teoretis dan kemajuan dalam pencitraan multidimensi dengan membentuk kembali nilai potensial yang analitik dan wawasan dapat memberikan peran visualisasi [5]. Big data merupakan sederat isu yang telah berkembang di kalangan data scientist. Istilah ini dipakai secara khusus untuk digunakan pada data yang melebihi kapasitas pemrosesan database konvensional. Dalam pengolahan data ini diperlukan alat yang handal dalam melakukan pengumpulan, pengelolaan dan analisis data. Hal ini biasa disebut dengan *big data analytics*.

Data-data tersebut dapat diolah dan divisualisasikan sesuai dengan kebutuhan. Hal ini sejalan dengan penelitian yang dilakukan [6], yang melakukan visualisasi data dari GIS (Geographics Information System) untuk open big data pada environment science. Dimana dalam penelitiannya berfokus pada masalah lingkungan yang memiliki suhu lebih tinggi harus dikurangi dengan angin laut yang mengair ke daerah pedalaman. Aliran angin terhalang oleh bangunan bertingkat tinggi di daerah pesisir, yang selanjutnya dengan model simulasi ini dapat diketahui secara kuantitatif pengaruh hambatan angin laut oleh cluster bangunan tinggi. Simulasi ini menggunakan WRF (Weather Research and Forecasting) dimana diperlukan suatu kelompok open big data seperti data tutupan lahan yang dibangun dari tata guna lahan dan data cuaca seperti data atmosfer dan suhu permukaan laut. Informasi Geospasial Jepang digunakan dalam analisis ini.

Penelitian ini juga didasarkan pada sistem prediksi untuk melihat tingkat penyebaran Covid-19 yang relatif tinggi. Virus corona merebak karena ada kontak antara penderita positif dengan masyarakat. Data statistik digunakan untuk menganalisis potensi penyebab sosial yang memengaruhi distribusi spasial jumlah terkonfirmasi kumulatif provinsi COVID-19 di Cina [7]. Untuk mendukung pemerintah dalam upaya mensukseskan program vaksinasi Covid-19 khususnya Kota Depok, diperlukan upaya dalam memantau tingkat keberhasilan masyarakat dalam menerima vaksin Covid-19. Hal ini sesuai dengan prinsip pertama dari *Guidance on Developing a National Deployment and Vaccination Plan for Covid-19 vaccines*, dimana peran pemerintah dalam mengambil keputusan berbasis ilmu pengetahuan. Pemerintah khususnya kota Depok perlu memantau progress pemberian vaksin Covid-19 pada masyarakat Depok. Dimana progress pemberian vaksinasi ini dapat divisualisasikan kedalam bentuk grafik setiap harinya.

Untuk melakukan visualisasi data yang jumlahnya besar, diperlukan alat yang mampu menyelesaikan berbagai permasalahan data dan membuat data-data tersebut dapat divisualisasi dengan baik guna mendukung dalam pemberian kebijakan [8]. Oleh sebab itu diperlukan tools yang mumpuni seperti *Big Data Analytic*.

Penggunaan perangkat sejenis juga dilakukan oleh Fernando, dimana dalam melakukan visualisasi data menggunakan Google Data sebagai alat pendukung untuk pembuatan laporan penjualan. Dimana data yang divisualisasi merupakan data penjualan tanpa harus memiliki kemampuan programming [9]. Bao, Ngan, Vung, Tommy menggunakan data Hybrid Accelerated Cosmology Code (HACC) sebagai objek visualisasi. Dimana bentuk visualisasi yang dilakukan adalah dengan melakukan rendering titik serta hidrodinamika partikel yang dihaluskan. Teknik dalam penyajian data menggunakan visualisasi data dalam 3D. Dalam proses visualisasi beberapa pendekatan dipakai dalam menemukan bentuk visualisasi terbaik [10]. Beberapa penelitian sebelumnya telah berhasil memproyeksikan data baku ke bentuk visualisasi data yang mudah dipahami oleh pembaca. Namun hal yang paling penting dalam visualisasi data adalah prediksi serta pengaruh dari masing-masing variabel data terhadap variabel lain. Dalam penelitian ini akan diperoleh visualisasi data yang bersifat prediktif terhadap laju pemberian vaksin Covid-19 di Depok. Selain itu dalam penelitian ini akan diperoleh pengaruh atas satu variabel dengan variabel lainnya. Sebagai contoh pengaruh fasilitas kesehatan terhadap jumlah penerima vaksin akan bisa dianalisis seberapa besar pengaruhnya mencakup pengaruh sedang, kuat ataupun lemah.

Seiring dengan pertambahan jumlah data di berbagai kasus di seluruh dunia. Bentuk *big data* sangat beranekaragam. Beberapa teknik dilakukan untuk melakukan dekomposisi dari jumlah dimensi *big data*. Masalah visualisasi data berdimensi tinggi dalam sebuah data mengakibatkan antarmuka perangkat lunak pemrosesan tidak dapat sepenuhnya ditampilkan dengan baik, hal ini disebabkan karena ukuran data melebihi dua dimensi, sehingga tidak dapat diproyeksikan ke dalam antarmuka dua dimensi. Mimoun dan Mohamed mencoba mengatasi masalah



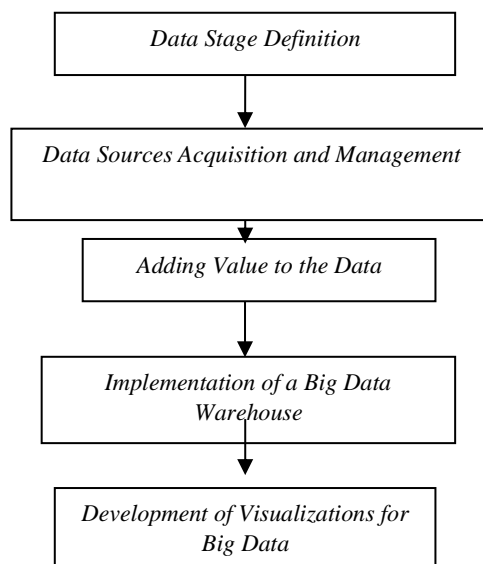
ini dengan melakukan integrasi dari Principal Component Analysis (PCA) yang dikombinasikan dengan Matrix by Block Decomposition (MBD) method (A.K.A block segmentation). Sehingga dapat mereduksi segmentasi dimensi data yang besar [11]. Salah satu perangkat yang dapat digunakan untuk memodelkan data dengan banyak dimensi dan mereduksi menjadi 2 dimensi adalah *SAP Analytics Cloud*. Solusi SAP Analytics Cloud menggabungkan BI, analitik augmented dan prediktif, dan kemampuan perencanaan ke dalam satu lingkungan cloud. Sebagai lapisan analitik Platform Teknologi Bisnis SAP, ini mendukung analitik tingkat lanjut di seluruh perusahaan maupun organisasi. SAP Analytics Cloud telah mengubah cara dalam pengambilan keputusan yang dibuat oleh ribuan pelanggannya yang mencakup lebih dari 75 negara. Dengan bergabung dengan mereka dan menggunakan SAP Analytics Cloud, sehingga dapat memberi setiap orang di organisasi terkait wawasan yang dibutuhkan untuk membuat keputusan yang lebih baik dan bertindak dengan percaya diri [12].

Dalam penelitian ini dilakukan visualisasi data program vaksinasi Covid-19 dengan menggunakan platform SAP Analytics Clouds yang akan membantu dalam menangani data yang berdimensi besar dan memiliki kapasitas pengolahan data yang besar menjadi bentuk 2 dimensi yang mudah dipahami. Kemudian hasil visualisasi menjadi mudah dipahami, diakses dan diprediksi.

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Tahapan penelitian yang dilakukan mengacu pada tahapan proses manajemen, analisis dan visualisasi Big Data. Tahapan tersebut mencakup, Data Stage Definition, Data Sources Acquisition and Management, Adding Value to the Data, Implementation of a Big Data Warehouse, Development of Visualizations for Big Data yang ditampilkan pada gambar 1 [8].



**Gambar 1.** Tahapan Proses Visualisasi Big Data

### 2.2 Data Stage Definition

Pada tahap pertama ini didefinisikan beberapa tahapan data sesuai untuk menganalisis persyaratan berdasarkan alur data. Disini ditambahkan detail tentang cara menentukan persyaratan analisis. Untuk tujuan ini, baik tipe repositori maupun data yang disimpan harus memenuhi persyaratan yang diidentifikasi selama fase ini sehingga harus diidentifikasi informasi (fungsional) dan persyaratan non-fungsional. Disini juga dianalisis, peluang analisis yang menentukan batasan waktu dari pengumpulan data hingga analisis, kualitas data persyaratan, dan query latency [13].

Dalam penelitian ini menggunakan 7 kolom yang terdiri dari tanggal, fasilitas\_kesehatan, Jenis\_kelamin, Range\_Umur, Kategori, Vaksin dan Jenis\_Vaksin dengan 613275 baris data, contoh data bisa dilihat pada table 1. Dari data ini dapat didefinisikan kebutuhan visualisasi mencakup bentuk visualisasi grafik yang berupa jumlah vaksin perhari yang dapat menyimpulkan trendnya akan naik atau akan turun. Selain itu bentuk kategorisasi dari penerima vaksin yang mencakup Gotong-royong, lansia, masyarakat umum, petugas publik, remaja serta SDM Kesehatan. Selain itu, kebutan jumlah vaksin yang diberikan serta jenisnya.



**Tabel 1.** Contoh Data Program Vaksin Covid-19

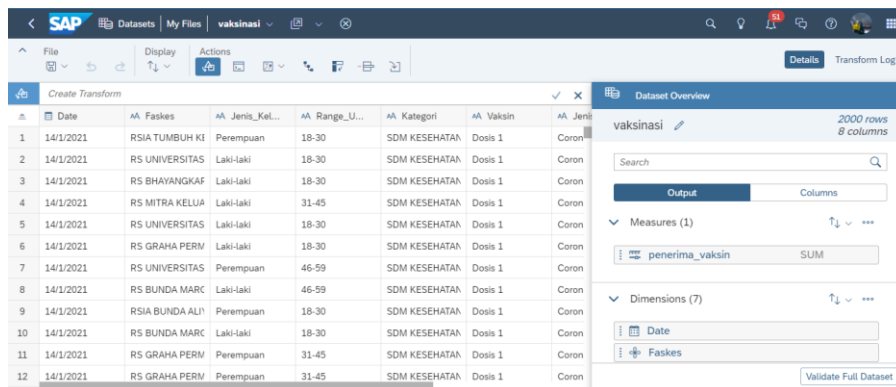
Date	Faskes	Gender	Range Age	Category	Dosis	Vaccine
1/14/2021	RSIA TUMBUH KEMBANG	Perempuan	18-30	SDM KESEHATAN	Dosis 1	CoronaVac
1/14/2021	RS UNIVERSITAS INDONESIA	Laki-laki	18-30	SDM KESEHATAN	Dosis 1	CoronaVac
1/14/2021	RS BHAYANGKARA BRIMOB	Laki-laki	18-30	SDM KESEHATAN	Dosis 1	CoronaVac
1/14/2021	RS MITRA KELUARGA DEPOK	Laki-laki	31-45	SDM KESEHATAN	Dosis 1	CoronaVac
1/14/2021	RS UNIVERSITAS INDONESIA	Laki-laki	18-30	SDM KESEHATAN	Dosis 1	CoronaVac

### 2.3 Data Sources Acquisition and Management

Tujuan dari tahap kedua ini adalah untuk mengumpulkan data mentah dan memuatnya ke dalam sistem file Big Data. Hal ini bergantung pada persyaratan yang diidentifikasi dalam Fase 1 dan produksi sumber data. Dua teknik berikut untuk mengumpulkan data:

- Batch Load: ketika requirement jauh dari data real time maka dibutuhkan metode pengambilan data menggunakan HDFS
- Streaming Load : proses dan analisis data secara realtime menggunakan HDFS

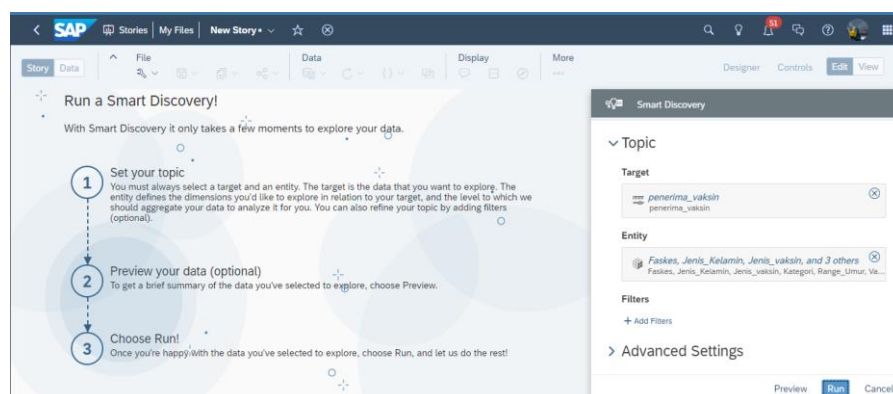
Data mentah dari tahap 1 akan diunggah kedalam sistem file *Big Data* milik SAP Analytics Cloud. Sehingga akan disesuaikan proses penggolongan *measure* dan *dimensions*. Pada gambar 2, kita bisa melihat bahwa measure merupakan variable yang digunakan untuk melakukan perhitungan jumlah data yang diinginkan sedangkan *dimesions* merupakan data yang akan dihitung.



**Gambar 2.** Load Data ke SAP Analytics Cloud

### 2.4 Adding Value to the Data

Langkah selanjutnya adalah menjelajahi data mentah yang disimpan di fase2 untuk mendefinisikan model data multidimensi (MD). Data mentah dikumpulkan di Tahap 2 berisi beberapa entitas dan hubungan tersembunyi yang tidak bisa diidentifikasi Dengan mendefinisikan model data MD, entitas data dan hubungan ini dijelaskan, dapat digunakan untuk analitik atas penambahan nilai pada data. Contoh pendefinisian model pada gambar 3.



**Gambar 3.** Pengaturan Entitas untuk Measure dan Dimensions



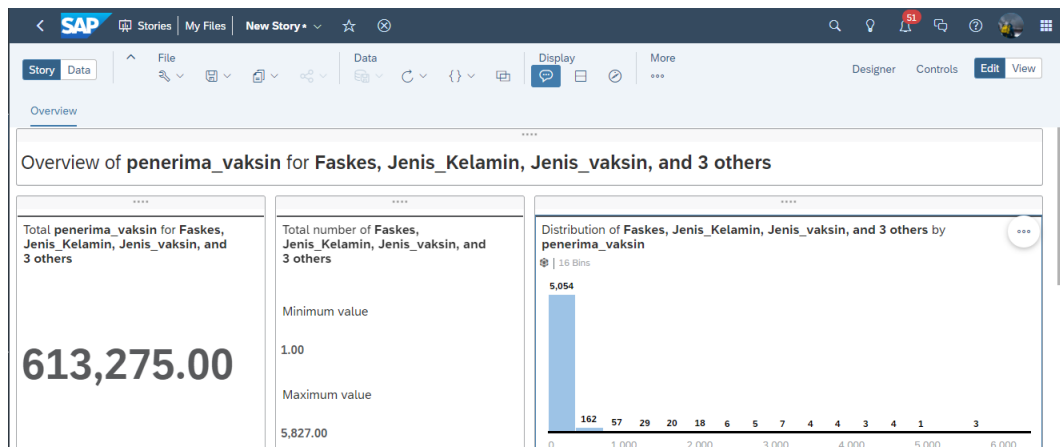


## 2.4 Implementing of Big Data Warehouse

Tujuan dari fase ini adalah untuk mengimplementasikan Big Data Warehouse yang menyediakan dukungan untuk sebagian besar alat Business Intelligent dan sesuai dengan persyaratan query latency yang diidentifikasi dalam fase ini diketahui alat mana yang akan digunakan berdasarkan persyaratan query latency, membantu Arsitek Data Warehouse serta memastikan bahwa fitur Volume Big Data dapat berjalan dengan baik.

## 2.5 Development of Visualizations for Big Data

Setelah data dimuat ke dalam Big Data Warehouse, penggunaan dapat dilakukan sebagai file sumber alat BI. Kemudian, dengan alat BI bisa didesain beberapa jenis visualisasi seperti tabel atau dasbor. Sehingga dapat dilakukan penafsiran data dengan lebih mudah berkat visualisasi. Hal ini dilakukan dengan menambahkan value lebih lanjut ke data. Bentuk visualisasi data bisa dilihat pada gambar 4.



Gambar 4. Contoh Visualisasi Data Vaksin

## 3. HASIL DAN PEMBAHASAN

Bertumbuhnya data secara digital mempengaruhi berbagai aspek bisnis. Nilai bisnis bisa bertambah dengan pemanfaatan data. Salah satu proses dalam pencarian nilai bisnis adalah dengan melakukan visualisasi data. Metode ini sudah dipakai pada sekitar abad 17, dimana visualisasi ini dapat menampilkan penanda lahan, kota, jalan, wilayah dan lain sebagainya. Trend perkembangan visualisasi data ini mengakibatkan banyak perangkat berlomba-lomba dalam mengembangkan kemampuan untuk melakukan visualisasi data yang fleksibel dan mampu mengelola jutaan data. Salah satu perangkat yang dapat mengelola puluhan juta data ini ada SAP Analytics Clouds. Dalam proses visualisasi ini dilakukan dengan mengikuti metode penelitian di Bab 2.

### 3.1 Dataset Program Vaksinasi Covid-19

Dalam proses pelaksanaan visualisasi data. Tahap terpenting adalah penyiapan dataset. Dalam penelitian ini menggunakan dataset Vaksinasi Covid-19 di Kota Depok. Dataset mempunyai timeframe yang di tandai dengan tanggal. Data yang dipakai didapatkan dari proses pertama pelaksanaan vaksinasi Covid-19 di tanggal 14 Januari 2021 sampai dengan tanggal 21 Agustus 2021. Selain itu ada entitas kesehatan yang dijadikan layanan pemberian vaksin di Kota Depok yang terdiri dari Abadi Jaya, Bakti Jaya, Beji, Bojongsari, Cilangkap, Cilodong, Cimanggis, Cimpalen, Cinangka, Cinere, Cipayung, CisalakPasar, DepokJaya, Depok Utara, Duren Seribu, Harjamukti, Jatijajar, Kalimulya, Kedaung, Kemiri Muka, Klinik Jaya GR, Klinik Kimia Farma Adelina Sawangan, Klinik Kimia Farma Margonda 352, Klinik Polres Metro Depok, Klimik Satelit Makara UI, Klinik Sikesjas Korbrimob, Limo, Mampang, Mekarsari, Pancoran Mas, Pasir Gunung Selatan, Pasir Putih, PD Sukmajaya, Pengasinan, Polkes 00.09.07 Depok, Pos Vaksin Dinkes Depok, Rangkapan Jaya, Ratu Jaya, RS Bhayangkara Brimob, RS Bunda Margonda, RS Brawijaya Depok, RS Citra MEDika Depok, RS Graha Permata Ibu, RS Hermina Depok, RS Hermina Depok GR, RS jantung Diagram, RS Meilia, RS Mitra KELuarga Depok, RS Permata Depok, RS Puri Cinere, RS Sentra Medika Cisalak Depok, RS Simpangan Depok, RS Sentra Medika Cisalak GR, RS Tugu Ibu, RS Universitas Indonesia, RSIA Asyifa Depok, RSIA Bunda Aliyahm RSIA Setya Bhakti, RSIA Tumbuh Kembang, RSU Bhakti Yudha, RSU Citra Arafik, RSU Hasanah Graha Afifah, RSUD Kota Depok, Sawangan, Sukmajaya Baru, Sukatani, Sukmajaya, Tanah Baru, Tapos, Tugu, Vila Pertiwi. Attribute gender juga digunakan untuk menandai berapa banyak laki-laki dan perempuan yang sudah di vaksin di Kota Depok. Attribute range\_umur dikategorikan menjadi 5 yaitu umur <18 tahun, > 60 tahun, 18-30 tahun, 31-45 tahun, 46-59 tahun. Kategori penerima vaksin terdiri dari gotong royong, lansia, masyarakat umum, petugas public, remaja, dan SDM Kesehatan.



### 3.1.1 Noise Treatment

Pada proses *data stage definition* dilakukan *big data preprocessing* [14]. Secara umum algoritma yang dipakai dalam data mining cenderung berasumsi bahwa setiap dataset merupakan sampel dari distribusi yang mendasari tanpa gangguan. Hal ini dalam artian bahwa walaupun ada anomali data tetap digolongkan sebagai data yang bisa dikelola. Namun hal ini akan berdampak pada capaian hasil yang diinginkan. Pada kenyataannya data jarang sekali sempurna, sehingga *noise* sering muncul. Kualitas hasil yang diperoleh dengan teknik data mining terutama pada *big data* bergantung pada kualitas data yang ingin dikelola. Penanganan khusus harus dilakukan untuk masalah *noise* pada data [15]. Untuk menangani *noise* dalam data mining, dua pendekatan utama biasanya digunakan dalam literatur *preprocessing data*. Yang pertama adalah memperbaiki *noise* dengan menggunakan metode perbaikan data, khususnya jika itu mempengaruhi pelabelan sebuah instance. Bahkan koreksi *noise* parsial diklaim bermanfaat, tetapi ini adalah tugas yang sulit dan biasanya terbatas pada sejumlah kecil *noise*. Yang kedua adalah menggunakan filter *noise*, yang mengidentifikasi dan menghapus *instance noise* dalam data pelatihan dan tidak memerlukan modifikasi teknik data mining. Dalam kasus ini implementasi *noise treatment* sangat penting. Karena ketersediaan jumlah data akan merefleksikan jumlah penerima vaksin di Kota Depok. Filter *noise* tipe 2 tidak dilakukan karena akan menghapus *instance noise* dalam data yang akan divisualisasikan. Sedangkan jumlah *noise* yang besar akan berpengaruh pada data deskriptif visualisasi. Oleh karena ini, dilakukan *treatment noise* jenis pertama yaitu melakukan perbaikan kepeada *noise* dengan melakukan perbaikan data. Hal ini dikarenakan akan mempengaruhi pelabelan data vaksinasi.

### 3.1.2 Pengurangan Dimensi

Ketika kumpulan data menjadi besar dalam jumlah variabel prediktor atau jumlah instance, algoritma data mining menghadapi masalah dimensionalitas [16]. Ini adalah masalah serius karena akan menghambat pengoperasian sebagian besar algoritma data mining seiring dengan kenaikan cost komputasi. Bagian ini akan menggaris bawahi algoritma reduksi dimensionalitas yang paling berpengaruh [17], [18]. Dalam kasus ini jumlah dimensi akan mempengaruhi proses visualisasi. Semakin banyak jumlah dimensi akan berbanding lurus dengan kenaikan cost komputasi. Dalam kasus ini terdapat pengurangan dimensi pada attribute sub kategori penerima vaksin. Hal ini dikarenakan sebagian besar *instance* tidak memiliki attribute jenis ini ini. Sehingga dalam penelitian ini menghapuskan attribute sub kategori penerima vaksin.

### 3.2 Pengelolaan Dataset

Sebagai bagian dari visualisasidata. Proses dalam *Data Sources Acquisition and Management* akan berpengaruh pada penyiapan data. Dalam perangkat SAP Analytics Cloud terdapat pengkategorian *instance* yaitu measure dan dimension. Measure disini merupakan instance yang ditambahkan sebagai wadah perhitungan nilai dari dimensi yang dipilih.

**Tabel 2.** Penambahan Instance Measure

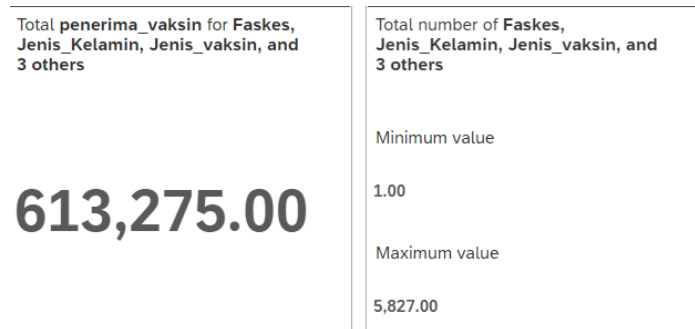
Date	Faskes	JenisKelamin	Umur	Kategori	Vaksin	Jenis_vaksin	Pnr_vaksin
14/1/2021	RSIA TUMBUH KEMBANG	Perempuan	18-30	SDM KESEHATAN	Dosis 1	CoronaVac	1
14/1/2021	RS UNIVERSITAS INDONESIA	Laki-laki	18-30	SDM KESEHATAN	Dosis 1	CoronaVac	1
14/1/2021	RS BHAYANGKARA BRIMOB	Laki-laki	18-30	SDM KESEHATAN	Dosis 1	CoronaVac	1
14/1/2021	RS MITRA KELUARGA DEPOK	Laki-laki	31-45	SDM KESEHATAN	Dosis 1	CoronaVac	1
14/1/2021	RS UNIVERSITAS INDONESIA	Laki-laki	18-30	SDM KESEHATAN	Dosis 1	CoronaVac	1

Pada tabel 2 terlihat bahwa Pnr\_vaksin merupakan *instance* baru yang ditambahkan dan akan dikategorisasikan sebagai measure. Measure memiliki nilai 1 yang dibakukan untuk semua baris data. Hal ini dikarenakan jumlah per baris mewakili satu orang yang menerima vaksin. Keseluruhan data akan di unggah pada SAP Analytics Cloud. Hal ini akan merepresentasikan *Implementing of Big Data Warehouse*. SAP sebagai wadah big data warehouse mampu mengelola data sebesar 2.000.000.000 baris data dengan 100 kolom. Setiap data dapat dikelola menjadi beberapa model data dan dapat diproyeksikan sebagai data deskriptif maupun data prediktif. 3.2 Pengembangan Visualisasi Big Data Vaksinasi Covid-19

Dalam proses pelaksanaan visualisasi data terkait data vaksinasi Covid-19 di Kota Depok akan diketahui bahwa jumlah penerima vaksin secara keseluruhan dari tanggal 14 januari 2021 sampai dengan 21 Agustus 2021 adalah sebesar 613.275. Dimana dalam 1 hari minimum terdapt 1 orang yang telah mendapatkan vaksin dan

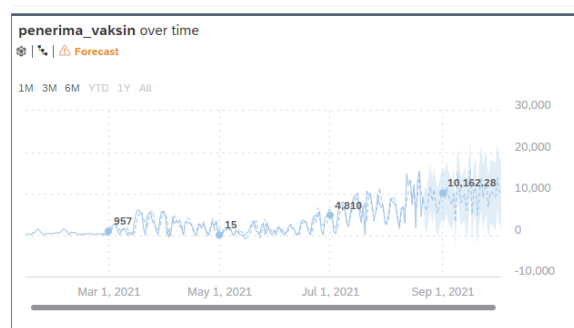


maksimum orang menerima vaksin adalah sebesar 5.827 dalam satu fasilitas kesehatan yang ada di Kota Depok. Detail dashboard terlihat pada gambar 5.



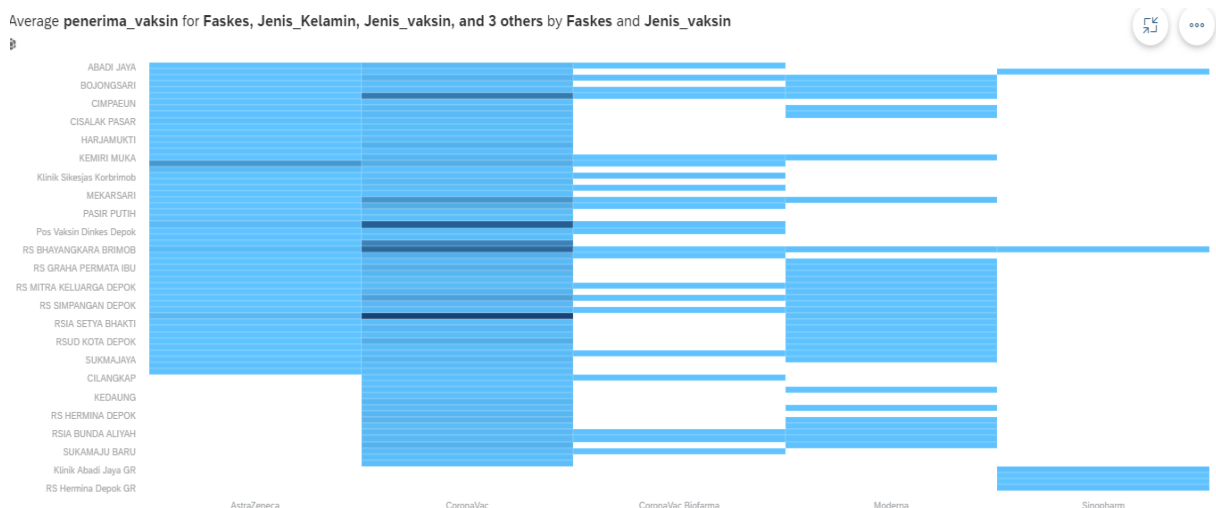
**Gambar 5.** Total Penerima Vaksin

Trend perkembangan pemberian vaksin pada masyarakat Kota Depok terlihat pada gambar 6. Dimana dalam setiap hari terjadi peningkatan masyarakat yang menerima vaksin. Vaksin yang diberikan mencakup dosis 1, dosis 2 dan dosis 3. Disini terlihat trend mendatang bahwa jumlah penerima vaksin sampai bulan oktober 2021 akan terus mengalami kenaikan. Hal ini juga didukung dengan kebijakan pemerintah yang mewajibkan seluruh elemen masyarakat untuk mendapatkan vaksin. Hal ini ditandai dengan pemanfaatan fasilitas-fasilitas umum harus menggunakan bukti telah menerima vaksin. Diperkiraan sampai akhir tahun 2021 jumlah penerima vaksin di Kota Depok akan terus mengalami kenaikan. Hal ini akan sesuai dengan target pemerintah Kota Depok di akhir 2021 untuk menuntaskan vaksin bagi masyarakat Depok.



**Gambar 6.** Trend Peningkatan Penerima Vaksin

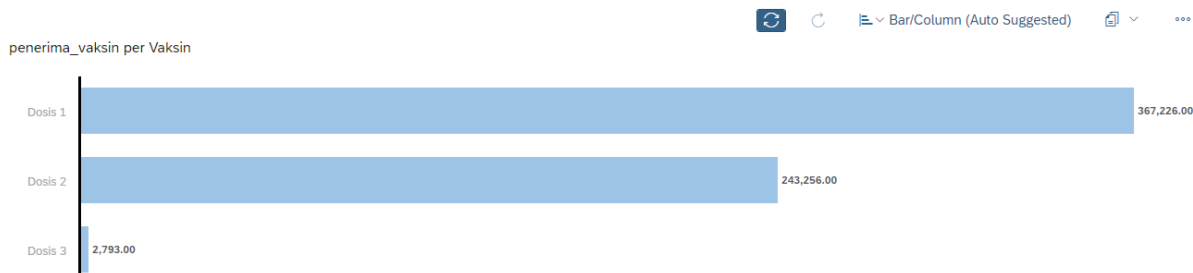
SAP Analytics Cloud juga dilengkapi dengan fitur Smart Discovery. Fitur ini menganalisis penerima\_vaksin untuk Faskes, Jenis\_Kelamin, Jenis\_vaksin, range\_usia, kategori, dan vaksin. Ini mengidentifikasi 6 kolom sebagai pemberi pengaruh utama. *Influencer* utama ini didasarkan pada keseluruhan dataset yang berjumlah 613.275. Disini digabungkan semua ukuran dan dimensi yang relevan ke tingkat entitas. Model pembelajaran mesin yang mendasari analisis dataset ini menunjukkan bahwa kualitas analisisnya bagus. Gambar 7. Menginterpretasikan entitas yang akan mempengaruhi entitas lain.



**Gambar 7.** Rata-rata penerima vaksin dari fasilitas kesehatan

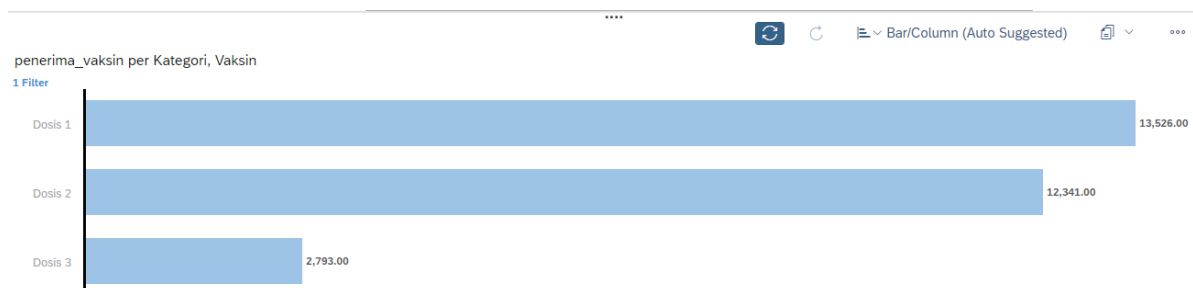


Dari gambar 7 terlihat bahwa setiap fasilitas kesehatan memberikan beberapa jenis vaksin tertentu. Memungkinkan bahwa dari 5 vaksin yang ada hanya beberapa vaksin yang dimiliki oleh fasilitas kesehatan tertentu. Dari keseluruhan vaksin yang ada jenis vaksin CoronaVac paling banyak dimiliki oleh semua fasilitas kesehatan yang ada di Kota Depok. Jumlah penerima vaksin AstraZeneca ada 30344, CoronaVac berjumlah 573703, CoronaVac Biofarma berjumlah 4634, Moderna berjumlah 2770 dan Sinopharm berjumlah 1824 untuk keseluruhan penerima vaksin.

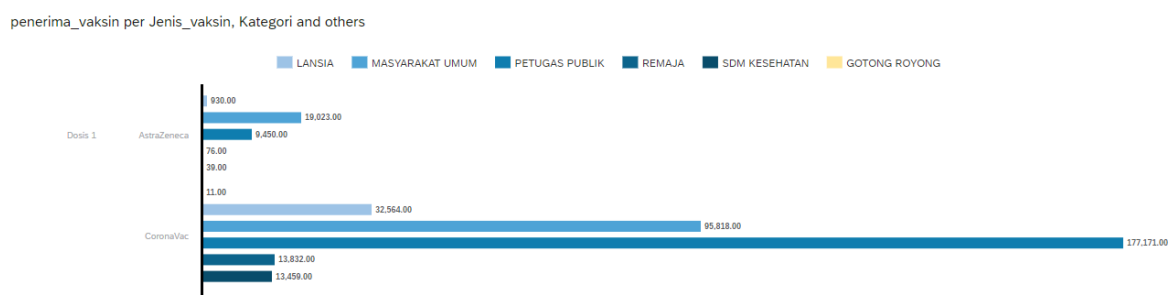


**Gambar 8.** Jumlah Penerima Vaksin Per Dosis

Gambar 8 mendeskripsikan bahwa jumlah penerima vaksin dosis satu berjumlah 367.226, dosis 2 sejumlah 243.256 dan dosis tiga sebesar 2.793. Se jauh ini penerima dosis satu cukup besar sedangkan penerima dosis 2 juga memiliki nilai yang sama. Namun hal ini terkait dengan interval penerimaan dosis ke 2. Sehingga sangat memungkinkan penerima dosis kedua lebih sedikit dari dosis pertama. Sedangkan pemberian dosis 3 hanya dilakukan untuk SDM Kesehatan hal ini bisa dilihat pada gambar 9 bahwa untuk SDM Kesehatan yang sudah menerima dosis 1 ada sebanyak 13526 sedangkan penerima dosis 2 ada sebanyak 12.341 dan penerima dosis 3 ada sebanyak 2793.



**Gambar 9.** Penerima Dosis 3 untuk SDM Kesehatan



**Gambar 10.** Cuplikasin kategori penerima vaksin berdasarkan jenis vaksin yang diberikan

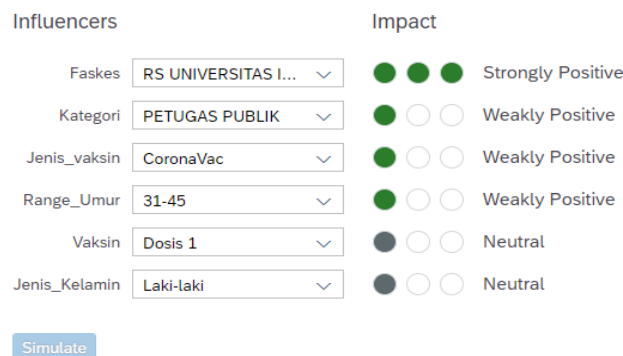
Berdasarkan gambar 10 bisa diketahui bahwa untuk SDM Kesehatan yang menerima dosis 3 hanya diberikan kepada jenis vaksin CoronaVac dan Moderna. Untuk penerima vaksin kategori lanjut usia pada dosis 1 telah menerima sebesar 930 untuk jenis vaksin AstraZeneca, sebanyak 32564 untuk CoronaVac, 129 untuk CoronaVac Biofarma, dan 2 untuk Sinopharm. Sedangkan pemberian dosis dua untuk lansia ada sebesar 78 untuk AstraZeneca, 29474 untuk CoronaVac, 33 untuk CoronaVac Biofarma dan belum ada sama sekali untuk Moderna dan Sinopharm.

Dalam SAP Analytics Cloud, dari data vaksinasi covid-19 dapat disimulasikan bahwa influencer ataupun dimensi pengukuran dalam kasus ini masing-masing mempunyai dampak terhadap jumlah penerima vaksin. Seperti yang terlihat pada gambar 11. Bahwa RS Universitas Indonesia mempunyai impact yang memiliki nilai strongly Positive. Hal ini dikarenakan jumlah terbanyak penerima vaksin di lingkungan Kota Depok didapat dari RS Universitas Indonesia yaitu berjumlah 87140 atau 14,20% dari total keseluruhan penerima vaksin. Jika diambil contoh lainnya seperti fasilitas kesehatan di sawangan mempunyai dampak negative lemah. Hal ini dikarenakan





dampak terhadap keberhasilan penerima vaksin tidak sebanyak pada RS Universitas Indonesia. Kemudian untuk kategori petugas public mempunyai dampak *weakly positive* pada tingkat pemberian vaksin hal ini karena petugas publik memiliki jumlah terbanyak pada kategori penerima vaksin. Untuk kategori gotong royong jumlah penerima vaksin sebesar 63210, kategori lansia sebesar 63210, untuk masyarakat umum sebesar 172958, untuk petugas publik sebesar 326894, untuk remaja ada sebesar 19765 dan untuk SDM Kesehatan ada 28660. Sehingga dapat dikatakan bahwa 53,30% keseluruhan penerima vaksin didapat oleh petugas publik.



**Gambar 11.** Simulasi Influencer terhadap Impact Pelaksanaan Vaksinasi Covid-19

Selain itu gambar 11 juga mendeskripsikan bahwa jenis vaksin CoronaVac mempunyai impact yang terkategori sebagai positif lemah. Hal ini dikarenakan jumlah vaksin jenis ini paling banyak dipakai. 93,54% penerima vaksin menggunakan vaksin ini. Kemudian untuk range umur 31-45 tahun berdampak positif lemah terhadap jumlah penerima vaksin keseluruhan. Jumlah penerima vaksin untuk umur <18 tahun sebesar 47292, untuk umur >60 berjumlah 67024, untuk umur 18-30 tahun berjumlah 186864, untuk umur 31-45 tahun berjumlah 185767 dan untuk umur 16-59 berjumlah 126328. Sehingga kategori umur 31-45 menyumbang 30,29% dari keseluruhan penerima vaksin.

Kemudian dosis vaksin 1 memiliki pengaruh netral terhadap keseluruhan penerima vaksin. Nilainya akan berimbang dengan pemberian vaksin tahap 2. Sehingga baik dosis 1 maupun dosis 2 akan memiliki dampak netral terhadap penerima vaksin dan terakhir adalah jenis kelamin yang juga memiliki nilai netral untuk kedua jenis tipe yaitu laki-laki dan perempuan. Jumlah penerima avaksin untuk laki-laki sebanyak 288373 dan untuk perempuan sebanyak 324902.

#### 4. KESIMPULAN

Visualisasi data program vaksinasi Covid-19 dapat dilakukan dengan menggunakan SAP Analytics Cloud sebagai solusi alternatif untuk pemegang kebijakan dalam melakukan visualisasi data. SAP Analytics Cloud memiliki dukungan berbagai sumber data yang fleksibel dan *powerfull* dalam mendukung proses cleansing data sampai dengan visualisasi dan sampai proses prediktif. Dalam kasus vaksinasi Covid-19, perangkat ini mampu menggambarkan kondisi pelaksanaan vaksinasi dari bulan Januari sampai dengan Agustus 2021. Dimana terlihat bahwa proses peningkatan penerima vaksin dari hari ke hari semakin bertambah. Proses visualisasi vaksinasi Covid-19 dilakukan dengan melakukan proses *cleansing dataset* dimana seringkali ditemukan anomali terhadap data. Perangkat ini mampu mendeteksi dan memvalidasi anomali terhadap data. Dalam proses visualisasi didapatkan bahwa total penerima vaksin dari tanggal 14 Januari sampai dengan 21 Agustus 2021 sebanyak 613276. Dimana Fasilitas kesehatan paling banyak memberikan vaksin adalah RS Universitas Indonesia. Dalam visualisasi data ini terlihat bahwa trend penerima vaksin di Kota Depok akan terus meningkat sampai dengan akhir tahun. Ada 70 fasilitas kesehatan yang setiap harinya akan berusaha meningkatkan layanan pemberian vaksin terhadap masyarakat. Untuk jenis vaksin yang digunakan, paling banyak menggunakan vaksin CoronaVac. Sedangkan penerima vaksin paling besar diperoleh petugas publik. Setiap *instance* akan berkorelasi mempengaruhi jumlah dari penerima vaksin seperti fasilitas kesehatan dan jenis vaksin yang akan diberikan. Pengaruh dari masing-masing *instance* akan terukur dari seberapa besar kontribusi dalam menyumbang jumlah penerima vaksin di Kota Depok.

#### UCAPAN TERIMA KASIH

Puji syukur kami panjatkan selama proses penelitian yang berjudul Visualisasi program vaksinasi Covid-19 di Kota Depok dengan Big Data Analytics. Pelaksanaan penelitian ini tidak terlepas dari dukungan Politeknik Negeri Jakarta yang telah memberikan wadah untuk mengembangkan minat dalam penelitian. Selain itu, tidak lupa ungkapan terima kasih juga kami sampaikan kepada Dinas Kesehatan Kota Depok yang telah membantu dalam



proses penelitian dengan meluangkan waktu untuk bersama-sama dapat memvisualisasikan keberhasilan Kota Depok dalam pemberian vaksin Covid-19 di Lingkungan Kota Depok.

## REFERENCES

- [1] WHO, "Coronavirus disease (COVID-19): Herd immunity, lockdowns and COVID-19," *WHO*, 2020. [https://www.who.int/news-room/q-a-detail/herd-immunity-lockdowns-and-covid-19?gclid=Cj0KCQjw5JSLBhCxARIsAHgO2Sd\\_T-5gy-CJfNASZngU0HDoELY6qSycDX5bVIYzjW7-Hql1HHASIM8aAi-OEALw\\_wcB#](https://www.who.int/news-room/q-a-detail/herd-immunity-lockdowns-and-covid-19?gclid=Cj0KCQjw5JSLBhCxARIsAHgO2Sd_T-5gy-CJfNASZngU0HDoELY6qSycDX5bVIYzjW7-Hql1HHASIM8aAi-OEALw_wcB#) (accessed Mar. 13, 2021).
- [2] WHO, "Coronavirus disease (COVID-19): VaccinesNo Title," *WHO*, 2020. [https://www.who.int/news-room/q-a-detail/coronavirus-disease-\(covid-19\)-vaccines?](https://www.who.int/news-room/q-a-detail/coronavirus-disease-(covid-19)-vaccines?) (accessed Apr. 01, 2021).
- [3] WHO, "Coronavirus disease (COVID-19) pandemic," *WHO*, 2020. [https://www.who.int/emergencies/diseases/novel-coronavirus-2019?adgroupsurvey=%7Badgroupsurvey%7D&gclid=Cj0KCQjw5JSLBhCxARIsAHgO2SdwrlmC1VlPO7wROFFb0YxQu0Q90l-bs2kFJaLU-3vYxe8sE-jBLU8aAuhaEALw\\_wcB](https://www.who.int/emergencies/diseases/novel-coronavirus-2019?adgroupsurvey=%7Badgroupsurvey%7D&gclid=Cj0KCQjw5JSLBhCxARIsAHgO2SdwrlmC1VlPO7wROFFb0YxQu0Q90l-bs2kFJaLU-3vYxe8sE-jBLU8aAuhaEALw_wcB) (accessed Apr. 01, 2021).
- [4] J. N. Charles Chen, Ling Chen, Mingjun Xiao, "The Impact Analysis of COVID-19 on China Various Industries Using Crawler Technology and Data Visualization Technology," 2021.
- [5] J. D. Miller, *Big Data Visualization*. UK: Packt Publishing Ltd, 2017.
- [6] Y. Hong, M. Zhang, and W. Q. Meeker, "Big data and reliability applications: The complexity dimension," *J. Qual. Technol.*, vol. 50, no. 2, pp. 135–149, 2018, doi: 10.1080/00224065.2018.1438007.
- [7] H. Y. Liyun SU, "Analysis of Factor Influencing the Spatial Distribution of Provincial Cumulative Confirmed Count of Novel Coronavirus Pneumonia (covid-19) in China," *2020 Int. Conf. Public Heal. Data Sci.*, pp. 5–9, 2020.
- [8] R. Tardio, A. Mate, and J. Trujillo, "An iterative methodology for big data management, analysis and visualization," *Proc. - 2015 IEEE Int. Conf. Big Data, IEEE Big Data 2015*, pp. 545–550, 2015, doi: 10.1109/BigData.2015.7363798.
- [9] D. Fernando, "Data Visualization Using Google Data Studio," *Natl. Semin. Inf. Technol. Eng.*, vol. 1, no. November, pp. 71–77, 2018, [Online]. Available: <https://ejurnal.lppmunsera.org/index.php/snartisi/issue/view/100>.
- [10] B. D. Nguyen, N. V. T. Nguyen, V. Pham, and T. Dang, "Visualization of Data from HACC Simulations by Paraview," *2019 IEEE Sci. Vis. Conf. SciVis 2019*, vol. 3, no. 2, pp. 31–32, 2019, doi: 10.1109/SciVis47405.2019.8968854.
- [11] M. Lamrini and M. Y. Chkouri, "Decomposition and Visualization of High-Dimensional Data in a Two Dimensional Interface," *ICSSD 2019 - Int. Conf. Smart Syst. Data Sci.*, 2019, doi: 10.1109/ICSSD47982.2019.9002846.
- [12] SAP, "SAP Analytics Cloud," *SAP*, 2020. <https://www.sap.com/sea/products/cloud-analytics.html?> (accessed Apr. 02, 2021).
- [13] Z. M. Khalid and S. R. M. Zebaree, "Big data analysis for data visualization: A review," *Int. J. Sci. Bus.*, vol. 5, no. 2, pp. 64–75, 2021, doi: 10.5281/zenodo.4462042.
- [14] T. Ahmad and M. N. Aziz, "Data preprocessing and feature selection for machine learning intrusion detection systems," *ICIC Express Lett.*, vol. 13, no. 2, pp. 93–101, 2019, doi: 10.24507/icicel.13.02.93.
- [15] A. Juneja and N. N. Das, "Big Data Quality Framework: Pre-Processing Data in Weather Monitoring Application," *Proc. Int. Conf. Mach. Learn. Big Data, Cloud Parallel Comput. Trends, Perspectives Prospect. Com. 2019*, pp. 559–563, 2019, doi: 10.1109/COMITCon.2019.8862267.
- [16] D. Fu, "The Density-Based Descending Dimension Algorithm LLE Publisher: IEEE Cite This PDF," 2019.
- [17] Y. Li and S. Li, "A real-time location privacy protection method based on space transformation," *Proc. - 14th Int. Conf. Comput. Intell. Secur. CIS 2018*, pp. 291–295, 2018, doi: 10.1109/CIS2018.2018.00071.
- [18] Z. Duygu, A. K. Ş. Eh, A. El, İ. B. O. L. Sedat, and E. Kiliç, "Veri Önişleme ve İstatistiksel Tekniklerin Kullanılmasıyla İş Kazası Verilerinin Analizi On the Analysis of Work Accidents Data by Using Data Preprocessing and Statistical Techniques," *2018 2nd Int. Symp. Multidiscip. Stud. Innov. Technol.*, 2018, doi: 10.1109/ISMSIT.2018.8567309.