

## **MEMPREDIKSI EFEK VAKSINASI TERHADAP ANGKA KASUS COVID-19 DENGAN REGRESI LINIER**

**Rikardo H. Siahaan<sup>1)</sup> dan Swingly Purba<sup>2)</sup>**

Dosen Fakultas Teknologi Industri  
Institut Sains Dan Teknologi TD. Pardede Medan

<sup>1)</sup>rikardosiahaan@istp.ac.id, <sup>2)</sup>swinglypurba@istp.ac.id

### **Abstrak**

Pandemi covid-19 di dunia telah berlangsung selama dua tahun lebih. bahkan hingga saat ini (29 mei 2022) WHO masih belum menurunkan status wabah pandemi ini. Selama pandemi ini berlangsung telah banyak penelitian dilakukan terhadap wabah covid-19 ini.

Salah satunya adalah upaya untuk memprediksi angka kasus covid-19 dengan menggunakan *data time series*. Namun berbeda dengan penelitian yang disebutkan sebelumnya, pada penelitian ini akan dilakukan prediksi ataupun pemodelan angka kasus covid berdasarkan faktor yang langsung dapat mempengaruhi angka kasus covid-19 yaitu jumlah dosis vaksinasi yang telah diterima masyarakat dengan menggunakan metode regresi linier.

**Kata kunci:** Covid-19, Regresi Linier

### **Abstract**

*The COVID-19 pandemic in the world has lasted for more than two years. even today (29 may 2022) WHO still has not lowered the status of this pandemic outbreak. During this pandemic, many studies have been carried out on this COVID-19 outbreak.*

*One of them is an effort to predict the number of COVID-19 cases using time series data. However, in contrast to the research previously mentioned, this study will predict or model the number of Covid cases based on factors that can directly affect the number of Covid-19 cases, namely the number of vaccination doses that have been received by the community using the linear regression method.*

**Keywords:** Covid-19, Linier Regression

### **I. Pendahuluan**

Salah satu cara untuk mencegah seseorang terpaparnya covid-19 selain daripada menggunakan masker dan melakukan *social distancing* adalah dengan diberikannya vaksin covid-19 dosis lengkap yaitu vaksinasi dosis 1 dan 2 dengan jenis vaksin yang sama serta vaksin tambahan booster untuk kekebalan tubuh lebih.

Dengan semakin tersedianya fasilitas vaksinasi bagi masyarakat umum, data dapat menjadi sebuah alat handall yang bila dimanfaatkan dengan baik akan membantu upaya penanganan pandemi ini secara signifikan. Berbeda dengan situasi dunia pada saat pandemi Flu Spanyol tahun 1918, saat ini kita memiliki data dan informasi yang sangat melimpah. Mulai dari pengumpulan, pengolahan, hingga pendistribusian data dapat dilakukan melalui penggunaan teknologi informasi dan komunikasi

yang semakin canggih.

Adapun studi kasus ini menggunakan metode regresi linier yang merupakan pemodelan statistik untuk menyelidiki hubungan antara variabel dependen, y, dengan satu atau lebih variabel bebas, X. Penerapan regresi linier sangat luas dan dapat diaplikasikan di hampir semua bidang ilmu, termasuk bidang komputer, teknik, biologi, dan bidang lainnya.

Melalui pengolahan data vaksinasi dan angka kasus covid-19, penulis mencoba mencari korelasi antara efektivitas vaksin terhadap angka kasus covid-19 yang terjadi di Indonesia. Seiring dengan semakin banyaknya jumlah vaksinasi yang telah diterima masyarakat, angka vaksinasi ini diasumsikan akan semakin memiliki dampak terhadap angka kasus covid-19.

Adapun hasil penelitian ini dapat digunakan sebagai acuan bagi masyarakat untuk mengetahui hal seputar vaksinasi pada tingkat pemahaman rudimental (*rudimental knowledge*) dan diperlukan kajian lebih dalam dengan metode ilmiah lainnya jika informasi ini ingin digunakan oleh pemerintah untuk pengambilan keputusan lebih lanjut seperti kebijakan seputar PPKM. Hal ini dikarenakan peristiwa pandemik covid-19 ini merupakan hal yang tergolong cukup baru bahkan untuk bidang epidemiologi. Besar sekali kemungkinan ada faktor lain yang dapat mempengaruhi angka kasus covid-19 secara signifikan yang masih belum dipelajari lebih dalam oleh para peneliti.

## II. Landasan Teori

### 1. Regresi Linier

Pada regresi linier, akan ada dua variabel yang digunakan pada diagram analisis: pertama adalah variabel dependen (pada sumbu-y) yang menjadi dasar prediksi, dan yang lainnya adalah variabel independen (pada sumbu x) yang digunakan untuk membuat prediksi. Prediksi berbasis variabel ini dapat didasari oleh satu variabel maupun beberapa variabel (*Moore et al., 2013*). Garis regresi adalah garis lurus yang menjelaskan bagaimana variabel dependen berubah dengan perubahan variabel bebas. Untuk membedakan prediksi model dengan beberapa set data lapangan, penulis menggunakan data dosis vaksin untuk menghitung jumlah kasus COVID-19 dan jumlah orang yang meninggal dengan mempertimbangkan faktor-faktor di atas. Adapun persamaannya :

$$Y = a + bX$$

Keterangan:

Y: Variable dependen

X: Variabel Bebas

a: Intercap (nilai Y pada saat X = 0)

b: Slope (perubahan rata-rata Y terhadap X)

### 2. RMSE dan nilai $r^2$

RMSE atau yang disebut juga Root Mean Squared Error adalah perhitungan kuadrat kesalahan maupun perbedaan antara nilai hasil prediksi dengan nilai sebenarnya. Nilai kesalahan ini memberikan gambaran tentang kecocokan model, yaitu bagaimana nilai menyimpang dari nilai sebenarnya. RMSE menggunakan kecocokan mutlak, sedangkan R-kuadrat menggunakan kecocokan relatif. RMSE dapat diartikan sebagai deviasi standar dari varians yang tidak dapat dijelaskan karena merupakan akar kuadrat dari varians. Ini memiliki keuntungan berada di unit yang sama dengan variabel jawaban. Semakin rendah nilai RMSE semakin baik prediksinya. Jika tujuan utama model

adalah prediksi, RMSE adalah kriteria penting untuk kecocokan karena merupakan standar yang valid tentang seberapa baik model memprediksi respons.

### 3. Python

merupakan sebuah bahasa pemrograman tingkat tinggi, yang merupakan bahasa pemrograman yang dirancang untuk membuat aplikasi umum. Bahasa ini berfokus pada kemudahan membaca kode program dengan penggunaan indentasi yang cukup unik. Python termasuk bahasa pemrograman yang *Dynamically Typed* dan *Garbage collected*. Peneliti juga menggunakan library *Panda* sebagai alat bantu dalam menganalisis data.

### III. Metode penelitian

Penelitian ini menggunakan metode regresi linier yang diterapkan pada data angka kasus covid-19 yang ada di Indonesia, adapun sumber data berasal dari website resmi Satuan Tugas Penanganan COVID-19 <https://covid19.go.id/>. Data yang dapat diambil dari website tersebut tersimpan dalam suatu file berformat .json dan masih perlu dilakukan *Pre-Processing* dan juga normalisasi sehingga siap untuk dianalisis lebih lanjut.

Pada pemilihan metode analisis data, dikarenakan penelitian ini lebih difokuskan sebagai media untuk memberikan edukasi seputar vaksinasi dan covid-19 kepada masyarakat dan bukan untuk langsung dijadikan sebagai sumber informasi pengambilan keputusan bagi pemerintah, maka penulis memilih menggunakan metode regresi linier sederhana sehingga dapat lebih mudah dimengerti oleh masyarakat umum.

### Dataset Preprocessing

Terdapat dua file json yang dapat diambil dari website covid19.go.id yaitu file pertama yang berisi tentang angka testing/pemeriksaan terhadap covid dan vaksinasi, dan yang kedua berisi jumlah angka kasus covid di Indonesia. Namun perlu diperhatikan bahwa angka jumlah vaksinasi yang disediakan pemerintah hanya berupa vaksinasi pertama dan kedua, untuk vaksinasi booster tidak disediakan pada file *dataset* yang tersedia.

Oleh karena data disajikan dalam bentuk json, dan belum dapat dibaca dengan baik, dataset ini perlu dipetakan ke dalam bentuk tabel lalu kemudian dilanjutkan dengan menghapus *data point* yang tidak terpakai pada penelitian ini sehingga hanya menyisakan variabel yang diperlukan saja seperti jumlah vaksinasi, angka kasus positif dan angka kasus meninggal.

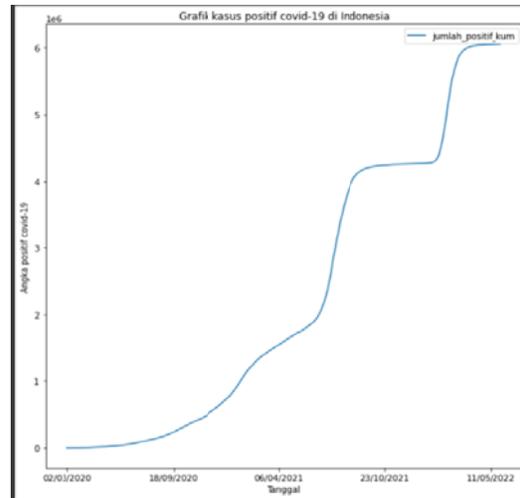
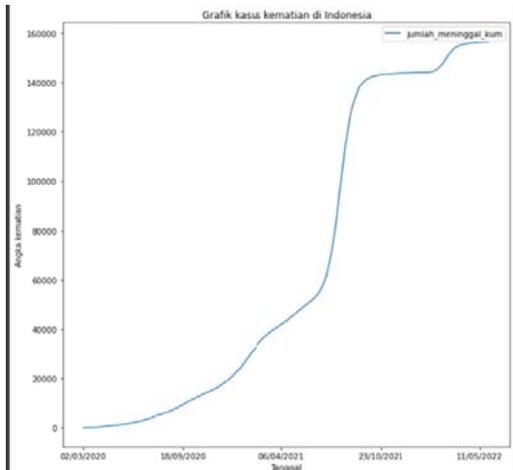
```
, "jumlah_orang_antigen_kum": {"value": 40795},
"jumlah_vaksinasi": {"penambahan": {"jumlah_vaksinasi":
600000, "doc_count": 1, "jumlah_vaksinasi_2":
{"key_as_string": "2021-02-02", "key": 16122},
"jumlah_jumlah_vaksinasi_2_kum": {"value": 0},
"jumlah_vaksinasi_1_kum": {"value": 0}, "jumlah_jumlah_vaksi
sinasi_1": {"value": 0}, "jumlah_jumlah_vaksinasi_2": {"value": 0}, "jumlah_vaksinasi_1": {"v
00000, "doc_count": 0, "jumlah_vaksinasi_2":
{"key_as_string": "2021-02-07", "key": 16126},
"jumlah_jumlah_vaksinasi_2_kum": {"value": 0},
"jumlah_vaksinasi_1_kum": {"value": 0}, "jumlah_jumlah_vaksi
sinasi_1": {"value": 0}, "jumlah_jumlah_vaksinasi_2": {"v
00000, "doc_count": 0, "jumlah_vaksinasi_2":
{"key_as_string": "2021-02-12", "key": 16130},
"jumlah_jumlah_vaksinasi_2_kum": {"value": 0},
"jumlah_vaksinasi_1_kum": {"value": 0}, "jumlah_jumlah_vaksi
sinasi_1": {"value": 0}, "jumlah_jumlah_vaksinasi_2": {"v
```

```
'jumlah_vaksinasi_2',
'jumlah_vaksinasi_1',
'jumlah_jumlah_vaksinasi_1_kum',
'jumlah_jumlah_vaksinasi_2_kum']
```

```
'jumlah_meninggal',
'jumlah_sembuh',
'jumlah_positif',
'jumlah_dirawat',
'jumlah_positif_kum',
'jumlah_sembuh_kum',
'jumlah_meninggal_kum',
'jumlah_dirawat_kum']
```

key_as_string	key	doc_count	jumlah_vaksinasi_2	jumlah_vaksinasi_1	
465	2022-05-12	1652313600000	1	129279	54811
466	2022-05-13	1652400000000	1	111396	62878
467	2022-05-14	1652486400000	1	79967	42913
468	2022-05-15	1652572800000	1	91850	48038
469	2022-05-16	1652659200000	1	20831	20913
470	2022-05-17	1652745600000	1	17579	19068
471	2022-05-18	1652832000000	1	116223	70974
472	2022-05-19	1652918400000	1	85293	26357
473	2022-05-20	1653004800000	1	140428	56908
474	2022-05-21	1653091200000	1	132018	45566

Adapun beberapa poin data yang menarik pada dataset ini seperti :



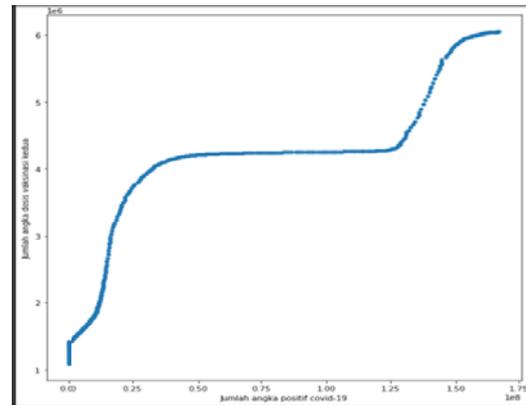
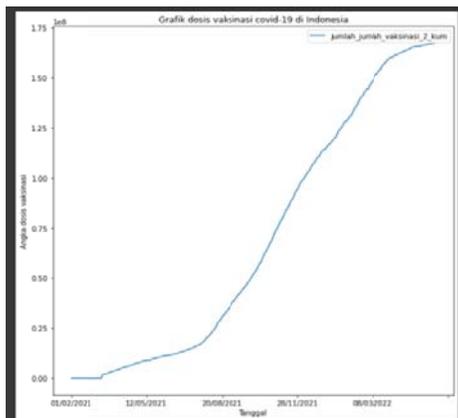
#### IV. Pembahasan

Ada beberapa hal yang perlu diperhatikan terlebih dahulu sebelum melanjutkan ke bagian ini. Hasil pemodelan dari prediksi ini sangat terbatas karena penelitian ini hanya menggunakan data rangkuman dari website satgas covid-19 yang tidak memiliki data unik untuk setiap kasus pemeriksaan tetapi merupakan laporan angka kasus covid-19 yang

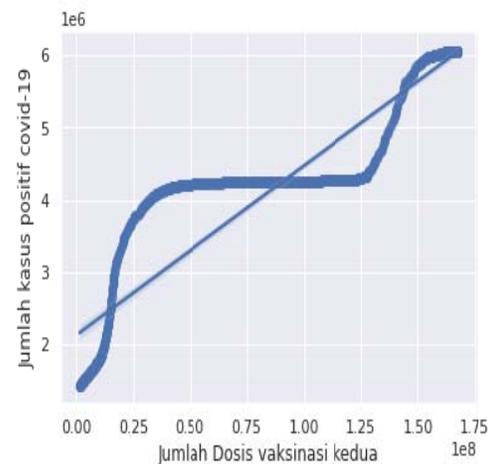
sudah digeneralisir. Penelitian ini masih dapat dikembangkan lebih lanjut untuk akurasi prediksi yang lebih tepat jika kami memiliki data yang lebih kompleks. Pada data yang digunakan terdapat pula beberapa angka pada data bernilai kosong dan juga adanya pencatatan data yang kurang konsisten sehingga membuat penafsiran data menjadi salah.

Pemodelan data ini didasarkan oleh orang yang telah menerima vaksin pertama atau kedua, namun lebih difokuskan pada vaksinasi kedua. Sesuai dengan arahan WHO, masyarakat perlu diberikan vaksinasi secara massal karena vaksinasi berguna untuk menghasilkan respons imun tubuh yang baik. Dosis vaksinasi kedua juga diperlukan karena akan menghasilkan tingkat respons imun tubuh yang jauh lebih baik dan memungkinkan masyarakat untuk melawan infeksi virus dengan lebih efektif daripada orang yang tidak divaksinasi.

Sebelum dilakukan metode regresi linier, data yang akan dianalisis perlu dipetakan terlebih dahulu :



kemudian dilakukan pemanggilan library Seaborn untuk pembuatan model linear regresi



## V. Kesimpulan

Sangat banyak faktor yang dapat mempengaruhi angka kasus covid-19. Hal ini juga yang menyebabkan sulitnya untuk dapat memprediksi kasus covid-19 di masa mendatang. Salah satu faktor utama tersebut adalah pola perilaku manusia, rendahnya kesadaran masyarakat akan pentingnya protokol PPKM terutama pada

## Referensi

- Algifari. **Analisis Regresi : teori, kasus, dan solusi**. Ed.2, Cet.1, BPFE UGM (2000).
- A. S. Albana and S. Azhari, **“Prediksi Penyebaran COVID-19 Kota Surabaya dengan Simulasi,”** J. Adv. Inf. Ind. Technol., vol. 2, no. 1, pp. 36–42, 2020.

waktu liburan atau perayaan dapat mengakibatkan lonjakan kasus covid-19.

Menurut penelitian *Jama network*, kasus covid pada tahun 2022 yang sudah didominasi oleh omicron memiliki kasus positif covid 19 tanpa gejala yang jauh lebih besar dari sebelumnya, hal ini dapat mengakibatkan pencatatan angka kasus yang tidak tepat karena kurangnya testing massal pada masyarakat sehingga masyarakat yang terkena covid tanpa gejala tidak tercatat pada sistem.

Buana, D. R. **Analisis Perilaku Masyarakat Indonesia dalam Menghadapi Pandemi Virus Corona (Covid-19) dan Kiat Menjaga Kesejahteraan Jiwa**. National Research Tomsk State University, Universitas Mercu Buana.(2020)

**Data COVID-19 di Indonesia** [Online]. Sumber: <https://covid19.go.id/>

Ndii, M.Z., **Pemodelan Matematika Dinamika Teori, Aplikasi, Dan Numerik**, Deepublish, (2018).

**Populasi Dan Penyebaran Penyakit**

N. Kusumawati, F. Marisa, and I. D. Wijaya, **“Prediksi Kurs Rupiah Terhadap Dollar Amerika Dengan Menggunakan Metode Regresi Linear,”** J I M P -J. Inform. Merdeka Pasuruan, vol. 2, no. 3, pp. 45–56, 2017

Valerisha, Anggia, and Marshell Adi Putra. **“Pandemi Global COVID-19 Dan Problematika Negara-Bangsa: Transparansi Data Sebagai Vaksin Socio-Digital?”** *Jurnal Ilmiah Hubungan Internasional, Special Edition* (2020).

Woodie, A. **Datanami. Retrieved from Data Transparency: Lessons from COVID-19:**  
<https://www.datanami.com/2020/04/09/data-transparency-lessons-from-covid-19/>