

Analisis Hubungan Kepadatan Penduduk dengan Pola Penyebaran COVID-19 Provinsi DKI Jakarta menggunakan Regresi Robust

Tiara Shofi Edriani ^{a,b}, Anisa Rahmadani^{a,b}, D. Michiko M. Noor ^{a,b}

^a Program Studi Matematika, Institut Teknologi Sumatera

^b Riset Prediksi dan Pemodelan Risiko Institut Teknologi Sumatera

* Corresponding E-mail: tiara.edriani@ma.itera.ac.id

Abstract: COVID-19 pandemic have been spread around the world since the first outbreak on Desember 2019 in Wuhan, China. DKI Jakarta as one of the highest population density among 34 provinces in Indonesia, has become an endemic area of COVID-19 with the rate of new cases show some fluctuation for each month along 2020. This is a secondary data research which drawn from Health Ministry of Indonesia as well as Center of Statistics for DKI Jakarta. Focus and the scope of this paper is on analyzing the relation between new cases of COVID-19 with population density of Jakarta's districts. Descriptive and inferential analysis that combined with Robust Regression Test are conducted due to some outliers data. This unbiased method shows a good regression model of spreading new positive cases. M-Estimator Robust Regression with Tukey Bisquare function, $Y = 1.8203 + 0.3882X$, shows the best result with the least Residual Standar Error (RSE), that is 0.411. Analysis on statistical test for the chosen model shows that population density has significant impacts on *outbreak* pattern of COVID-19 in Jakarta. But mobilities and interactions between citizens has also give a great impact.

Keywords: linear regression analysis, robust regression, Tukey Bisquare function, COVID-19, population density

Abstrak: Pandemi COVID-19 bermula sejak Bulan Desember 2019 di Wuhan, China. Di Indonesia, DKI Jakarta merupakan wilayah yang sangat padat dan endemik di masa ini dengan angka penambahan kasus positif tiap bulannya pada tahun 2020 mengalami fluktuasi. Penelitian ini merupakan penelitian observasional dengan menggunakan data sekunder dari situs Kementerian Kesehatan Republik Indonesia dan Badan Pusat Statistik Provinsi DKI Jakarta. Fokus penelitian ini adalah menganalisis hubungan kasus COVID-19 dengan kepadatan penduduk di wilayah DKI Jakarta. Metode penelitian ini menggunakan analisis deskriptif dan inferensial dengan regresi *robust* karena terdapat beberapa data pencilan. Penggunaan regresi *robust* ini memberikan hasil model regresi yang baik dengan nilai yang tidak bias. Regresi *robust* menggunakan metode Estimasi-M dengan fungsi *Bisquare* Tukey, $Y = 1.8203 + 0.3882X$, menunjukkan hasil terbaik dengan *Residual Standard Error* (RSE) terkecil, yakni sebesar 0.411. Pada analisis hasil uji statistik atas model ini, diketahui bahwa kepadatan penduduk memiliki pengaruh yang signifikan terhadap kasus terkonfirmasi COVID-19. Namun banyak faktor lain yang juga mempengaruhi, diantaranya yang jelas terlihat adalah intensitas interaksi antarwilayah serta mobilitas penduduknya.

Kata Kunci: analisis regresi linear sederhana, regresi robust, fungsi Bisquare Tukey, COVID-19, kepadatan penduduk

Pendahuluan

Sejak diumumkan sebagai pandemi pada Maret 2020 oleh WHO, penyebaran *Coronavirus Disease* (COVID-19) telah menelan tak kurang dari 50 juta korban jiwa di seluruh dunia [1]. Penyakit ini disebabkan oleh adanya infeksi *Severe Acute Respiratory Syndrome Coronavirus 2* (SARS-CoV-2)

[2]. Saat penelitian ini dilakukan, Indonesia berada pada urutan ke-24 total kasus positif terbanyak, dari 215 negara yang terkonfirmasi positif kasus COVID-19, dengan angka kematian mencapai 4.320 jiwa [1]. Sementara itu, DKI Jakarta berada di urutan ke-2 terbanyak kasus positif COVID-19 di antara 34 provinsi di Indonesia. Angka

Original Article

terkonfirmasi COVID-19 di Ibu Kota Negara RI ini tercatat sebanyak 17.279 kasus, dengan 748 diantaranya adalah kasus kematian.

Menurut Carl J. Neiderud (2015), kota-kota modern yang merupakan destinasi dari arus urbanisasi dapat menjadi katalis untuk penyebaran penyakit menular yang cepat [3]. DKI Jakarta merupakan kawasan perkotaan modern yang padat penduduk, sekaligus kota megapolitan terbesar di Indonesia dan aglomerasi urban terbesar kedua di dunia [4]. Karakteristik DKI Jakarta memiliki kepadatan penduduk sebesar 16.704 jiwa/km². Menurut Emili A., dkk (2011), kepadatan populasi mempengaruhi laju penularan penyakit, khususnya yang ditularkan melalui jalur pernapasan dan tranmisi fekal-oral (diare, polis, hepatitis, dan sebagainya)[5].

Tingkat kepadatan penduduk yang tinggi akan menimbulkan tingginya frekuensi interaksi antar individu pada wilayah tersebut. Tercatat bahwa kasus positif dan kematian yang tinggi pada pandemi Flu Spanyol di tahun 1918 terjadi pada negara dengan densitas populasi yang tinggi^[5]. Chandra menemukan bahwa terdapat suatu ambang batas terhadap suatu nilai dari kepadatan penduduk tertentu di suatu wilayah yang menjadi tolak ukur utama dalam menentukan efektifitas diberlakukannya karantina atau *physical distancing* pada kasus Flu Spanyol satu dekade yang lalu. Area dengan tingkat kepadatan melebihi ambang batas tersebut mengalami penurunan penyebaran kasus Flu Spanyol saat diberlakukan karantina, jauh lebih efektif dibandingkan dengan area yang tingkat kepadatannya di bawah ambang batas [6].

Sejalan dengan hal tersebut, saat ini seluruh dunia memberlakukan PSBB (Pembatasan Sosial Berskala Besar) sebagai langkah dalam mencegah penyebaran COVID-19. Tercatat bahwa jumlah kasus baru pada berbagai negara terkofimasi positif dengan jumlah besar (seperti China, Iran, UK, Spanyol) mengalami penurunan yang signifikan saat diberlakukan PSBB pada distrik yang paling padat penduduknya [7,8]. Kendati demikian, hubungan langsung antara kepadatan penduduk dengan tingkat mortalitas pada kasus COVID-19 tidak menunjukkan korelasi yang kuat pada beberapa studi berbasis metode statistik [9]. Hamidi (2020) membandingkan dua variabel, yaitu *activity density* dan luas area sebuah kota metropolitan, pada kaitannya dengan laju penyebaran COVID-19

serta laju mortalitasnya. Hasilnya menunjukkan bahwa luas area kota metropolitan berkorelasi positif dengan laju penyebaran maupun mortalitas pada *outbreak* ini. Sedangkan *activity density* berkorelasi negatif yang dipengaruhi oleh baiknya sistem layanan kesehatan pada kota padat penduduk [9].

Variabel paling jelas yang berkorelasi dengan laju penyebaran maupun mortalitas COVID-19 adalah tingkat mobilitas penduduk suatu daerah [10]. Hal ini mendukung penelitian Hamidi (2020) mengenai adanya korelasi positif cukup kuat antara luasnya area kota metropolitan dengan laju penyebaran maupun mortalitas. Sebab semakin luas area metropolitan tersebut, semakin tinggi mobilitas yang dilakukan oleh penduduk setempat dan semakin banyak moda transportasi umum yang digunakan sehingga penularan mudah terjadi [9].

DKI Jakarta bukan hanya area dengan kepadatan penduduk tinggi, namun juga kota megapolitan dengan total area yang luas. Penelitian yang dilakukan Ruiqi Li dkk. (2018) mengenai efek kepadatan penduduk terhadap fenomena epidemi menjelaskan bahwa terdapat batasan dan asumsi yang perlu diperhatikan untuk melihat korelasi antara kepadatan penduduk dengan penyebaran suatu penyakit epidemik. Jika *background noise* tersebut dihilangkan atau dikontrol, maka akan terlihat adanya korelasi yang erat antara kepadatan penduduk dengan tingkat penyebaran tersebut [11].

Salah satu *background noise* yang disebutkan oleh Ruiqi Li (2018) adalah adanya pencilan pada data [11]. Oleh sebab itu, penelitian ini ditujukan untuk menganalisis hubungan antara kepadatan penduduk dengan pola penyebaran COVID-19 di Jakarta dengan kombinasi beberapa metode statistik, termasuk metode pada regresi *robust* yang dapat menangkap fenomena pencilan dengan baik tanpa menghapus data pencilan tersebut. Selanjutnya *Residual Standar Error (RSE)* pada model regresi *robust* Estimasi-M menggunakan fungsi Huber akan dibandingkan dengan Estimasi-M menggunakan fungsi *Bisquare* Tukey untuk

memilih model regresi terbaik. Analisis uji asumsi klasik untuk Metode Kuadrat Terkecil atau *Ordinary Least Square* (OLS) juga dilakukan pada penelitian ini sehingga memberi argumen yang tepat dalam menerapkan regresi *robust*. Koefisien korelasi dan determinasi ditentukan untuk pengujian ketetapan model yang telah diperoleh, serta dipaparkan hasil analisis model tersebut dalam menjelaskan fenomena hubungan kepadatan penduduk dengan jumlah kasus terkonfirmasi positif COVID-19 di DKI Jakarta.

Metode

Data yang digunakan pada penelitian ini berupa data sekunder, yaitu data jumlah penduduk (jiwa) dan luas wilayah (ha) dari lima wilayah kotamadya dan satu kabupaten administratif DKI Jakarta yang diperoleh dari situs Badan Pusat Statistik (BPS). Selain itu, digunakan pula data sebaran kasus COVID-19 DKI Jakarta dari situs Badan Nasional Penanggulangan Bencana (BNPB) dan corona.jakarta.go.id.

Variabel penelitian yang akan dianalisis adalah variabel terikat (Y), yaitu jumlah kasus COVID-19 per kecamatan di wilayah DKI Jakarta periode Maret 2020 hingga Juli 2020 dan variabel bebas (X), yaitu kepadatan penduduk dari 44 wilayah kecamatan yang tersebar di lima wilayah kotamadya dan satu kabupaten administratif DKI Jakarta tahun 2019.

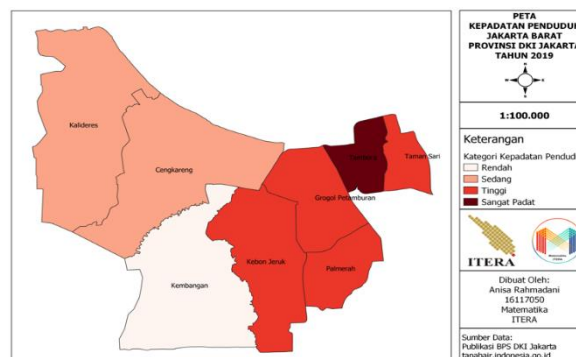
Berdasarkan peraturan SNI 03-1733-2004 Tata Cara Perencanaan Lingkungan Perumahan di Perkotaan, klasifikasi kepadatan penduduk dibagi dalam empat kelompok berikut:

Tabel 1. Klasifikasi Kepadatan Penduduk

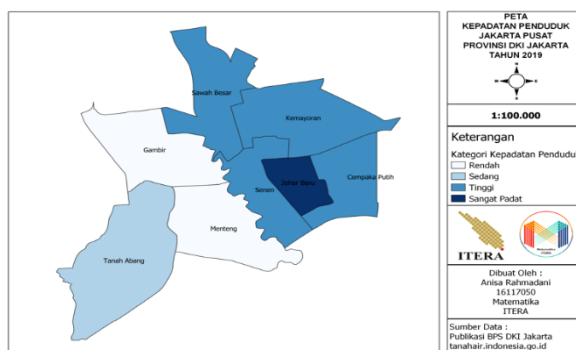
No	Klasifikasi	Kepadatan (Jiwa/ha)
1.	Rendah	< 150
2.	Sedang	151 - 200
3.	Tinggi	200 - 400
4.	Sangat Padat	> 400

Peta kepadatan penduduk masing-masing kecamatan di wilayah DKI Jakarta tahun 2019 diperoleh dengan membandingkan jumlah penduduk (jiwa) per kecamatan dengan luas wilayah (ha), lalu diolah menggunakan *software* ArcGIS. Hasil pemetaan 44 wilayah kecamatan DKI

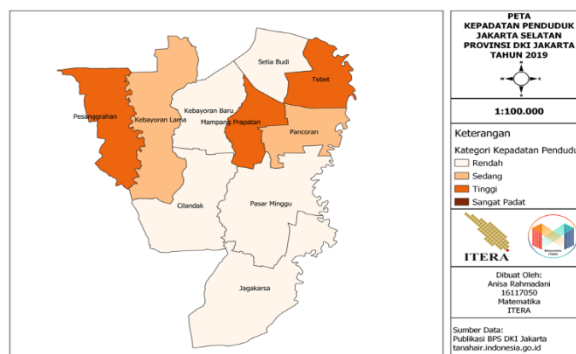
Jakarta ditampilkan pada Gambar 1 hingga Gambar 5.



Gambar 1. Peta Kepadatan Penduduk Jakarta Barat tahun 2019.

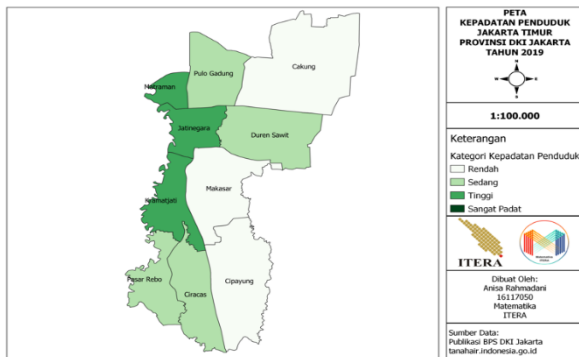


Gambar 2. Peta Kepadatan Penduduk Jakarta Pusat tahun 2019

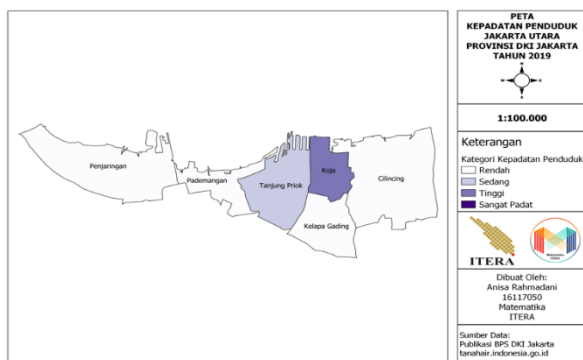


Gambar 3. Peta Kepadatan Penduduk Jakarta Selatan tahun 2019

Original Article

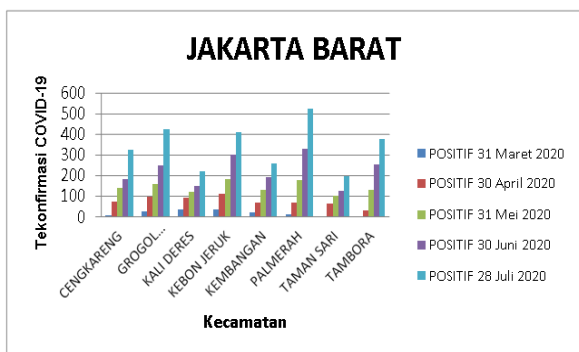


Gambar 4. Peta Kepadatan Penduduk Jakarta Timur tahun 2019

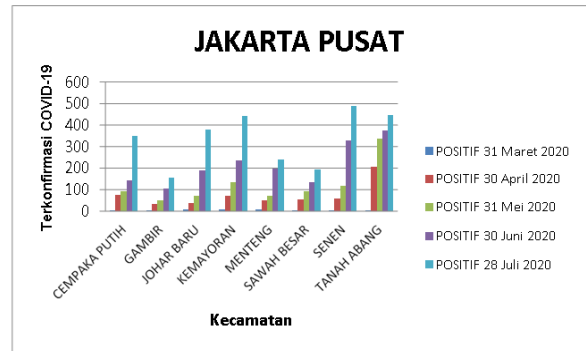


Gambar 5. Peta Kepadatan Penduduk Jakarta Utara tahun 2019

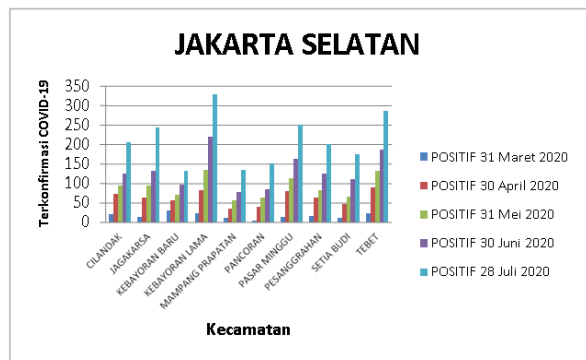
Penyebaran COVID-19 di wilayah DKI Jakarta sangat bervariasi dengan klasifikasi kepadatan penduduk yang berbeda per kecamatan. Berdasarkan data yang dibagikan Satuan Tugas Penanganan COVID-19, DKI Jakarta tercatat memiliki 12.812 kasus terkonfirmasi positif COVID-19 dari Maret hingga Juli 2020. Informasi tersebut kemudian diolah dan ditampilkan dalam bentuk grafik pada Gambar 6 hingga Gambar 10.



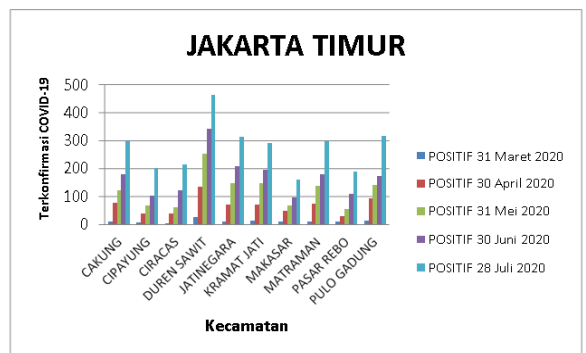
Gambar 6. Grafik Kasus Terkonfirmasi Positif COVID-19 Jakarta Barat pada Maret 2020 hingga Juli 2020



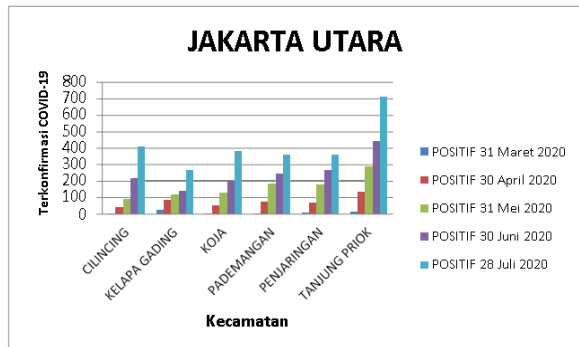
Gambar 7. Grafik Kasus Terkonfirmasi Positif COVID-19 Jakarta Pusat pada Maret 2020 hingga Juli 2020



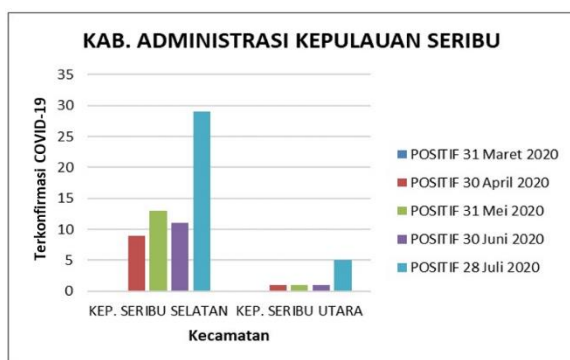
Gambar 8. Grafik Kasus Terkonfirmasi Positif COVID-19 Jakarta Selatan pada Maret 2020 hingga Juli 2020



Gambar 9. Grafik Kasus Terkonfirmasi Positif COVID-19 Jakarta Timur pada Maret 2020 hingga Juli 2020



Gambar 10. Grafik Kasus Terkonfirmasi Positif COVID-19 Jakarta Utara pada Maret 2020 hingga Juli 2020



Gambar 11. Grafik Kasus Terkonfirmasi Positif COVID-19 Kepulauan Seribu pada Maret 2020 hingga Juli 2020

Analisis data yang digunakan pada penelitian ini adalah analisis regresi linear sederhana dan regresi *robust*. Analisis regresi linear sederhana dilakukan dengan bantuan *software* SPSS bertujuan untuk mengetahui besarnya pengaruh variabel X terhadap variabel Y melalui suatu model regresi linear terbaik. Dalam pembentukan model, salah satu metode yang digunakan untuk menentukan koefisien-koefisien regresi linear yang merupakan nilai duga dari parameter model regresi linear adalah Metode Kuadrat Terkecil atau *Ordinary Least Square* (OLS). Metode ini dapat meminimumkan kesalahan pendugaan dengan syarat data residual memenuhi beberapa asumsi klasik regresi linear. Selanjutnya, dapat dilakukan tahap uji parsial dan simultan dari koefisien regresi linear untuk mengetahui hubungan linear antara variabel-variabel bebas terhadap variabel terikat.

Dalam pemodelan regresi linear menggunakan metode OLS, adakalanya asumsi klasik tidak terpenuhi, salah satu penyebabnya adalah terdapat pencilan (*outlier*) pada data pengamatan. Hal ini akan mempengaruhi model regresi linear dengan terbentuknya koefisien regresi yang bersifat bias dan hasil interpretasi yang tidak valid. Oleh sebab itu, diperlukan metode regresi lain yang dapat mengatasi kondisi ini.

Regresi *robust* merupakan alternatif dari metode OLS yang digunakan ketika distribusi dari residual tidak normal dan/atau terdapat beberapa *outlier* yang mempengaruhi model [12]. Regresi *Robust* dapat digunakan untuk mendeteksi pencilan (*outlier*) dan memberikan hasil yang resisten (stabil) dengan adanya pencilan (*outlier*) [13]. Salah satu metode estimasi parameter dari regresi *robust* yang paling sering digunakan adalah Metode Estimasi-M. Metode ini membutuhkan suatu fungsi pembobot yang dapat meminimumkan pengaruh pencilan terhadap model sehingga akan diperoleh model regresi linear yang terbaik. Pada penelitian ini, akan digunakan *software* RStudio untuk mengkaji regresi *robust* menggunakan metode estimasi-M dengan fungsi pembobot Huber dan *Bisquare* Tukey.

Hasil dan Pembahasan

Hubungan kepadatan penduduk dan jumlah total kasus COVID-19 per kecamatan di wilayah DKI Jakarta periode Maret 2020 hingga Juli 2020 akan dianalisis menggunakan regresi linear sederhana. Mula-mula dilakukan pembentukan model menggunakan metode OLS dan diperoleh $Y = -3.675 + 0.9723X$. Metode ini akan menghasilkan pendekatan terbaik jika memenuhi uji asumsi klasik. Hal ini diperlukan untuk menjamin perolehan koefisien regresi yang tidak menimbulkan nilai bias dan hasil interpretasi yang valid.

Uji Asumsi Klasik

Penerapan uji asumsi klasik regresi linear sederhana dilakukan pada data residual, yaitu selisih data pengamatan dengan data hasil estimasi pada model regresi linier. Pada penelitian ini, uji

Original Article

asumsi klasik regresi linear sederhana yang dibahas terdiri dari Uji Normalitas, Uji Heteroskedastisitas, serta Uji Autokorelasi. Seluruh pengujian dilakukan pada data residual dari data jumlah kasus COVID-19 dan kepadatan penduduk di wilayah DKI Jakarta.

A. Uji Normalitas Residual

Uji normalitas residual dilakukan untuk menguji apakah data residual yang diolah berdistribusi normal atau tidak. Teknik pengujian normalitas residual yang digunakan pada penelitian ini adalah *Kolmogorov-Smirnov Test* dan *Shapiro-Wilk Test*. Konsep dasar dari *Kolmogorov-Smirnov Test* adalah membandingkan distribusi data yang akan diuji normalitasnya dengan distribusi normal baku. Jika distribusi data yang akan diuji menunjukkan kemiripan (tidak terdapat perbedaan yang signifikan) dengan distribusi normal baku, maka dapat disimpulkan bahwa data yang diuji berdistribusi normal [14].

Hasil uji normalitas menggunakan uji statistik *Kolmogorov-Smirnov Test* dan *Shapiro-Wilk Test* dengan taraf signifikansi $\alpha = 0.05$ dinyatakan sebagai berikut

Tabel 2. Hasil Uji Normalitas
Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	.105	44	.200*	.945	44	.036

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Dari **Tabel 2** dapat dilihat bahwa nilai signifikansi (p) pada *Kolmogorov-Smirnov Test* adalah 0.2 ($p > 0.05$), artinya nilai residual telah normal karena memenuhi syarat pengambilan keputusan residual berdistribusi normal, yaitu nilai signifikansi pada uji *Kolmogorov-Smirnov* yang bernilai lebih dari taraf signifikansi $\alpha = 0.05$. Hal yang berbeda terjadi pada hasil *Shapiro-Wilk Test*. Nilai signifikansi (p) yang diperoleh adalah 0,036 ($p < 0.05$) sehingga tidak memenuhi syarat pengambilan keputusan residual berdistribusi normal, yaitu nilai signifikansi pada uji *Shapiro-Wilk* yang seharusnya bernilai lebih dari taraf signifikansi $\alpha = 0.05$. Secara umum, *Shapiro-Wilk Test* dikatakan sebagai uji normalitas yang lebih *powerful* dibandingkan dengan *Kolmogorov-Smirnov Test* [14].

B. Uji Heteroskedastisitas

Uji heteroskedastisitas dilakukan untuk menguji ketergantungan variansi residual terhadap variabel bebas. Deteksi heteroskedastisitas dalam penelitian ini menggunakan Uji *Breusch-Pagan*, yaitu statistik uji yang dilakukan dengan cara meregresikan residu kuadrat (sebagai variabel terikat) dengan variabel bebas [15].

Hasil uji heteroskedastisitas menggunakan uji *Breusch-Pagan* dengan taraf signifikansi $\alpha = 0.05$ dapat dilihat pada **Gambar 12**. Berdasarkan hasil uji tersebut diperoleh informasi bahwa variabel bebas cukup memberikan pengaruh terhadap variansi residual. Hal ini terlihat dari nilai probabilitas Chi-Square sebesar $1.014368e^{-07}$ (prob. Chi-Squ. $e < 0.05$). Artinya terjadi gejala heteroskedastisitas pada metode, yaitu keadaan dengan variansi residual yang tidak konstan atau terdapat ketidaksamaan variansi dari suatu residual dengan residual lainnya.

Gambar 12. Hasil Uji Heteroskedastisitas

```

Breusch Pagan Test for Heteroskedasticity
-----
Ho: the variance is constant
Ha: the variance is not constant

-----
Data
-----
Response : log(positif_juli)
variables: fitted values of log(positif_juli)

-----
Test Summary
-----
DF                = 1
Chi2              = 28.34637
Prob > Chi2      = 1.014368e-07

```

C. Uji Autokorelasi

Uji autokorelasi dilakukan untuk mengetahui adanya korelasi antar residual pada model regresi dari waktu ke waktu. Autokorelasi seringkali terjadi pada data yang mengandung unsur waktu (*time series*). Pada penelitian ini, autokorelasi akan dideteksi menggunakan uji *Durbin-Watson*, yaitu statistik uji yang akan membandingkan nilai statistik *Durbin-Watson* (DW) dengan nilai-nilai pada Tabel *Durbin Watson* [16].

Hasil uji autokorelasi menggunakan statistik uji *Durbin-Watson* disajikan pada **Gambar 13** berikut.

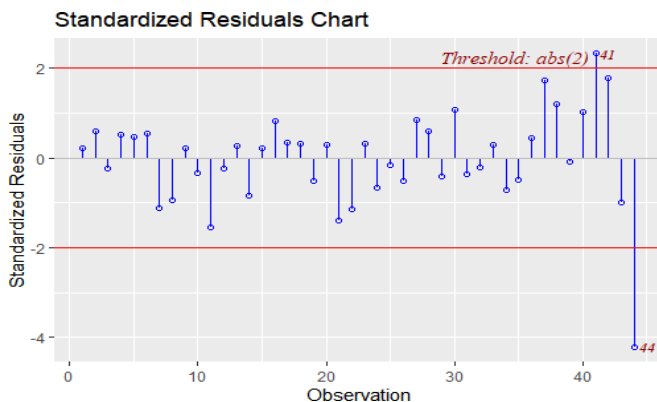
```
> #///non-autokorelasi
> durbinwatsonTest(ols_md11, max.lag = 1)
lag Autocorrelation D-w Statistic p-value
1      0.2653825      1.161515      0
Alternative hypothesis: rho != 0
```

Gambar 13. Hasil Uji Autokorelasi

Output Rstudio di atas menunjukkan bahwa nilai statistik *Durbin-Watson* (*DW*) adalah 1.161515. Kemudian, nilai tersebut dibandingkan dengan nilai *dL* dan *dU* pada Tabel *Durbin Watson*. Berdasarkan Tabel *Durbin Watson* dengan $k = 2, n = 44$ serta $\alpha = 0.05$ diperoleh nilai $dL = 1.4692$ dan $dU = 1.5619$. Dengan demikian nilai $DW (1.1615) < dL (1.4692)$ sehingga dapat disimpulkan bahwa pada residual terdapat fenomena autokorelasi.

Berdasarkan uji asumsi klasik yang telah dilakukan, terdapat beberapa asumsi yang tidak terpenuhi oleh data residual sehingga metode OLS akan memberikan nilai pada koefisien regresi model yang bersifat bias dan hasil interpretasi yang tidak valid. Salah satu penyebabnya adalah adanya pencilan (*outlier*) pada data pengamatan.

Berikut hasil deteksi *outlier* berdasarkan nilai *standardized residual* pada data kasus terkonfirmasi positif COVID-19 serta kepadatan penduduk di wilayah DKI Jakarta.



Gambar 14. Hasil Deteksi *Outlier*

Gambar 14 menunjukkan bahwa terdapat dua subjek pada residual yang memiliki nilai ekstrim

dan terdeteksi mengalami *outlier*, yaitu data ke-41 (Kec. Penjarangan) dan data ke-44 (Kec. Kep. Seribu Utara). *Outlier* berpengaruh besar terhadap koefisien regresi sehingga dapat mengubah kesimpulan penelitian yang diambil dari hasil analisis statistik.

Regresi Robust

Regresi *Robust* merupakan alat penting dalam menganalisis data yang terkontaminasi pencilan (*outlier*) [13]. Pada penelitian ini, data pengamatan yang terkontaminasi *outlier* akan dikenakan regresi *robust* Estimasi-M dengan fungsi Huber dan Tukey *Bisquare* untuk menentukan model regresi yang paling baik dalam mengestimasi parameter atau koefisien regresi bagi model data pengamatan.

A. Metode Estimasi-M dengan fungsi Huber

Hasil estimasi koefisien model regresi *robust* dengan menggunakan Estimasi-M fungsi Huber terlihat pada **Gambar 15**, sehingga model regresi *robust* yang dapat dibentuk adalah

$$Y = -1.0598 + 0.6678X \quad (1)$$

Konstanta sebesar -1.0598 menunjukkan bahwa jika tidak memperhatikan kepadatan penduduk (X) maka nilai konsisten kasus COVID-19 adalah -1.0598. Sedangkan koefisien regresi sebesar 0.6678 berarti bahwa setiap penambahan 1% tingkat kepadatan penduduk maka kasus COVID-19 meningkat sebesar 0.67%.

```
> #//huber-weight
> mhud_md11 <- regboot(form = formula(log(positif_juli) ~ log(kepadatan_perkm2)),
+                       data = df, model = "M", psi = MASS::psi.huber,
+                       maxit = 200, nboot = 1000) #-estimasi
> mhud_md11$bootreg$to #-est
      (Intercept) log(kepadatan_perkm2)
-1.0598087      0.6777846
> cibootreg(mhud_md11, conf = .95) #-selang kepercayaan 95%
      (Intercept)      0.025      0.975
-7.3715589      3.552787
log(kepadatan_perkm2) 0.2133175 1.320162
> gof(mhud_md11) #-Rsq dan MSE
      R-sq      MSE      MAE
0.5077668 0.3326646 0.3885546
```

Gambar 15. Hasil Estimasi-M dengan fungsi Huber

Original Article

B. Metode Estimasi-M dengan fungsi *Bisquare* Tukey

Hasil estimasi koefisien model regresi *robust* dengan menggunakan Estimasi-M fungsi *Bisquare* Tukey tertera pada **Gambar 16**

```
> #//bisquare-weight
> mbis_md11 <- regboot(log(positif_juli) ~ log(kepadatan_perkm2),
+ data = df, model = "M", psi = MASS::psi.bisquare,
+ maxit = (700), nboot = 1000) #-estimasi
> mbis_md11$bootreg$to #-est
      (Intercept) log(kepadatan_perkm2)
      1.8203070      0.3881738
> cibootreg(mbis_md11, conf = .95) #-selang kepercayaan 95%
      (Intercept)      0.025      0.975
      -7.8241566  4.232903
log(kepadatan_perkm2) 0.1436602  1.351371
> gof(mbis_md11) #-Rsq dan MSE
      R-sq      MSE      MAE
0.5077668  0.4252525  0.3925393
```

Gambar 16. Hasil Estimasi-M dengan fungsi *Bisquare* Tukey

Berdasarkan **Gambar 16**, model regresi *robust* yang sesuai adalah

$$Y = 1.8203 + 0.3882X \quad (2)$$

Pada metode yang menggunakan fungsi *Bisquare* Tukey ini menunjukkan bahwa tanpa memperhatikan kepadatan penduduk (X), nilai konsisten kasus COVID-19 adalah 1.8203. Selain itu, koefisien regresi sebesar 0.3882 berarti bahwa setiap penambahan 1% tingkat kepadatan penduduk menyebabkan kasus COVID-19 meningkat sebesar 0.39%.

Penentuan Metode Regresi Terbaik

Setelah mencari estimasi parameter model regresi, selanjutnya dilakukan penghitungan nilai *Residual Standard Error* (RSE) dari masing-masing model. Keseluruhan model beserta nilai RSE yang sudah didapat terlihat pada **Tabel 3**:

Tabel 3. Perbandingan Nilai RSE Metode Regresi

Metode Regresi	Model yang dihasilkan	<i>Residual Standard Error</i> (RSE)
Metode Kuadrat Terkecil	$Y = -3.675 + 0.9723X$	0.5629
Estimasi-M fungsi Huber	$Y = -1.0598 + 0.6678X$	0.4258
Estimasi-M fungsi <i>Bisquare</i> Tukey	$Y = 1.8203 + 0.3882X$	0.411

Berdasarkan **Tabel 3**, model regresi linear terbaik bagi data kepadatan penduduk dan jumlah total kasus COVID-19 per kecamatan di wilayah DKI

Jakarta periode Maret 2020 hingga Juli 2020 dengan kriteria nilai *Residual Standard Error* (RSE) terkecil adalah metode regresi *robust* Estimasi-M menggunakan fungsi *Bisquare* Tukey.

Selanjutnya, pengujian parsial dilakukan untuk parameter model regresi *robust* Estimasi-M menggunakan fungsi *Bisquare* Tukey. Uji ini dilakukan untuk mengetahui ada tidaknya pengaruh variabel X terhadap variabel Y . Dengan β_1 menunjukkan *slope*, koefisien regresi untuk variabel X dari model, hipotesis yang digunakan adalah^[17]

$$H_0: \beta_1 = 0$$

$$H_1: \beta_1 \neq 0$$

H_0 menyatakan bahwa tidak ada pengaruh kepadatan penduduk terhadap total kasus COVID-19 di Jakarta. Dari hasil statistik uji diperoleh nilai $t_{hitung} = 6.582$ dan $t_{tabel} = 2,0181$ sehingga $t_{hitung} > t_{tabel}$. Dengan demikian H_0 ditolak, artinya kepadatan penduduk berpengaruh signifikan terhadap jumlah total kasus COVID-19.

Uji Ketetapan Model

Untuk mengetahui eratnya hubungan antara dua variabel penelitian, dinyatakan dengan suatu bilangan, yaitu koefisien korelasi. Selain itu, untuk mengetahui besarnya pengaruh dan kemampuan variabel bebas dalam menerangkan variabilitas variabel terikat, dinyatakan dengan koefisien determinasi (*R Square*).

Hasil koefisien korelasi dan determinasi disajikan pada **Tabel 4**. Hasil yang diperoleh menunjukkan bahwa nilai korelasi sebesar 0.713. Nilai ini dapat diinterpretasikan bahwa hubungan antara variabel jumlah kasus terkonfirmasi positif COVID-19 dengan kepadatan penduduk ada di kategori kuat. Kemudian koefisien dari *R Square* adalah 0.508 atau 50.8% sehingga dapat disimpulkan bahwa pengaruh kepadatan penduduk terhadap jumlah kasus terkonfirmasi COVID-19 sebesar 50.8%, sedangkan sebesar 49.2% sisanya dipengaruhi oleh variabel lain yang tidak disertakan dalam model penelitian ini.

Tabel 4. Hasil Koefisien Korelasi dan Determinasi

Model Summary ^a										
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.713 ^a	.508	.496	.56286	.508	43.326	1	42	.000	1.162

a. Predictors: (Constant), Transformasi_Keypad

b. Dependent Variable: Transformasi_Covid19

Pola penyebaran COVID-19 sangat bervariasi. Hasil statistik yang telah diperoleh sebelumnya menunjukkan bahwa kepadatan penduduk memiliki pengaruh yang signifikan terhadap kasus terkonfirmasi COVID-19. Namun jika diperhatikan lebih dalam, masih banyak faktor lain yang mempengaruhi nilai terkonfirmasi COVID-19 diantaranya intensitas interaksi antarwilayah serta mobilitas penduduk.

Pada peta pemetaan wilayah berdasarkan kepadatan penduduk dengan grafik kasus terkonfirmasi positif COVID-19 yang disajikan bahwa tidak selalu pada wilayah dengan kepadatan tinggi memiliki angka kasus terkonfirmasi COVID-19 yang tinggi pula. Kasus COVID-19 tertinggi ada pada Kecamatan Duren Sawit (Jakarta Timur), Kecamatan Tanjung Priok (Jakarta Utara), dan Kecamatan Kebayoran Lama (Jakarta Selatan) dengan klasifikasi kepadatan penduduk tingkat rendah. Kecamatan Duren Sawit berbatasan dengan wilayah Bekasi dan pada bagian selatan berbatasan dengan Kota Depok dimana wilayah tersebut merupakan wilayah yang mempunyai kasus COVID-19 tertinggi di Jawa Barat. Pada kasus ini, faktor lokasi dan interaksi antar wilayah menjadi faktor penting dalam penyebaran virus. Perdagangan dan perjalanan internasional berpotensi dalam kontribusi terhadap terjadinya pandemi di seluruh dunia [3]. Pernyataan tersebut sesuai dengan penyebaran kasus COVID-19 wilayah Kecamatan Tanjung Priok yang terkenal dengan pelabuhan pengangkutan barang maupun penumpang. Kegiatan keagamaan yang dilakukan Warga Negara Asing (WNA) pada kelurahan Sunter Agung juga menjadi pemicu penyebaran.

Kesimpulan

Dari penelitian ini dapat disimpulkan bahwa model regresi yang paling baik untuk menganalisis

hubungan antara kepadatan penduduk dengan jumlah kasus terkonfirmasi positif COVID-19 di DKI Jakarta periode Maret 2020 hingga Juli 2020 adalah regresi *robust* menggunakan metode M-estimator fungsi *Bisquare* Tukey dengan persamaan $Y = 1.8203 + 0.3882X$. Kriteria penentuan model regresi terbaik ini berdasarkan nilai *Residual Standard Error* (RSE) yang paling minimum. Hubungan antara dua variabel penelitian dikategorikan kuat dengan nilai koefisien korelasi sebesar 0.713. Selain itu, nilai koefisien determinasi (*R Square*) menunjukkan pengaruh kepadatan penduduk terhadap jumlah kasus terkonfirmasi positif COVID-19 sebesar 50.8%. Sisanya sebesar 49.2% dipengaruhi oleh variabel-variabel lain, misalkan intensitas interaksi antar wilayah serta mobilitas penduduk. Hal ini tampak dari peta kepadatan penduduk pada **Gambar 1-5** dan grafik kasus terkonfirmasi positif COVID-19 pada **Gambar 6-11**. Terdapat beberapa kecamatan yang memiliki kepadatan penduduk tingkat rendah, namun masuk pada kategori kecamatan dengan kasus terkonfirmasi positif COVID-19 yang tinggi.

Konflik Kepentingan

Tidak ada konflik kepentingan pada penulisan jurnal ini.

Ucapan Terima Kasih

Terimakasih penulis ucapkan atas partisipasi Pusat Penelitian Kependudukan LIPI dalam memberikan masukan yang berharga pada penelitian yang menyertai penyusunan jurnal ini.

Referensi

- [1] Worldometer 2020, Coronavirus Cases, <https://worldometer.info/coronavirus/> (diakses 19 November 2020).
- [2] WHO.int (2020), Timeline: WHO's COVID-19 response, www.who.int/emergencies/diseases/novel-coronavirus-2019 (diakses 19 November 2020).
- [3] Neiderud, C. J., dan MD, "How urbanization affects the epidemiology of emerging infectious disease," *Infection Ecology and Epidemiology*, 2015.

Original Article

- [4] Martinez, R., dan Irna N. M. "Jakarta: A city of cities", Elsevier, 2020.
- [5] Emilie A, dkk. "Urbanisation and infectious diseases in a globalised world", Elsevier vol. 11, 2011.
- [6] Chandra S., dkk. "A geographic analysis of population density thresholds in influenza pandemic of 1918-19," International Journal of Helath Geographics.
- [7] Tammes, P. "Social distancing, population density, and spread of COVID-19 in England."
- [8] Jawad, A. J., "Effectiveness of population density as natural social distancing in COVID-19 spreading", Elsevier Masson SAS.
- [9] Hamidi, S., "Does Density Aggravate the COVID-19 Pandemic," Journal of American Planning Association, vol. 86 Number 4, pp. 495-508, June 2020.
- [10] Ghiffari, R. A., "Dampak populasi dan mobilitas perkotaan terhadap penyebaran pandemi COVID-19 di Jakarta", Jurnal Tunas Geografi Vol. 09 No. 01, pp.81-88, Juli 2020.
- [11] Li, R. dkk, "Effect of population density on epidemics," Physica A: Statistical Mechanics and Its Application.
- [12] Ryan, T. P., "Modern Regression Methods", Canada: John Wiley & Son, Inc., 1997.
- [13] Chen, Colin., "Robust Regression and Outlier Detection with the ROBUSTREG Procedure," SUGI Paper 265-27, 2002.
- [14] Razali, N. M., dkk. "Power Comparison of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Test," Journal of Statistical Modeling and Analytics Vol. 1, No. 1, pp. 21-33, 2011.
- [15] Breusch, T. S., dan A. R. Pagan, "A Simple Test for Heteroscedasticity and Random Coefficient Variation," Journal of The Economics Society Vol. 47 No. 5, pp. 1287-1294, 1979.
- [16] Durbin, J., dan G. S. Watson, "Testing for Serial Correlation in Least Squares Regression II," Biometrika Vol. 38 No.1/2, pp. 159-177, 1951.
- [17] Walpole, R. E, dan R. H. Myers, "Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuan Edisi ke-4", Bandung: ITB Press, 1995.