

Modeling The Gender Development Index of West Java Province Using A Fourier Series Nonparametric Regression Approach

Dhea Putri Rizkia, Fadhilah Fitri*, Dony Permana, Dina Fitria

Departemen Statistika, Universitas Negeri Padang, Padang, Indonesia

*Corresponding author: fadhilahfitri@fmipa.unp.ac.id

Submitted : 24 Mei 2024

Revised : 30 Mei 2024

Accepted : 31 Mei 2024

ABSTRACT

Gender equality is a development target in many countries. The ideal condition in human development that is expected is that male and female population groups have equal access to play a role in development, control over existing development resources, and receive benefits from development equally and fairly. The gender gap still occurs today in all aspects. The condition of the gender gap can be known by looking at the Gender Development Index. In observing the data curve, do the Gender Development Index and each independent variable not form a certain pattern. In addition, the data patterns that are formed tend to repeat. Nonparametric regression analysis is the solution. Fourier series is a nonparametric analysis used for repetitive data. Modeling was performed using 1, 2, and 3 oscillation parameters. Of the three parameters, the best model resulted from the $K=3$ oscillation with a GCV value of 2.8084 and a coefficient of determination of 42.39%.

Keywords: *Fourier Series, GCV, Gender, Nonparametric Regression, Oscillation Parameters.*



This is an open access article under the Creative Commons 4.0 Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2022 by author and Universitas Negeri Padang.

I. PENDAHULUAN

Pembangunan adalah proses perubahan pada aspek sosial, seperti ekonomi, pendidikan dan teknologi, infrastruktur, kelembagaan, dan budaya tanpa mengabaikan keragaman kebutuhan dan dasar masing-masing aspek. Tujuan pembangunan yakni mewujudkan masyarakat adil, makmur, dan sejahtera (Samad, dkk, 2019). Keberhasilan pembangunan diukur melalui beberapa indikator salah satunya adalah kesetaraan gender. Gender diartikan sebagai kondisi yang setara antara perempuan maupun laki-laki dalam memperoleh hak serta kesempatan dalam berbagai bidang.

Kesetaraan gender merupakan salah satu target pembangunan di berbagai negara. Pembangunan tidak dapat mencapai potensinya secara maksimal apabila terjadi ketimpangan pembangunan berdasarkan gender (Hadiyanti, 2021). Kondisi ideal yang diharapkan dalam pembangunan manusia adalah setiap individu, baik laki-laki maupun perempuan, memiliki akses yang sama untuk berpartisipasi dalam pembangunan. Mereka harus memiliki kendali atas sumber daya yang ada serta menerima hasil pembangunan secara setara dan adil. (Badan Pusat Statistik, 2022).

Kesenjangan gender masih terjadi sampai saat ini di segala aspek. Masih adanya ketidakadilan bagi perempuan di Indonesia adalah bukti kesenjangan gender (Zuhri dan Amalia, 2022). Salah satu bentuk ketidakadilan bagi perempuan adalah terjadinya kasus kekerasan terhadap perempuan di berbagai daerah di Indonesia. Menurut Komisi Nasional Anti-Kekerasan terhadap Perempuan (Komnas Perempuan) pada tahun 2022 terdapat 457.895 jumlah pengaduan kasus kekerasan dengan rata-rata 17 pengaduan kasus per hari. Informasi terkait kondisi kesenjangan gender ini dipublikasikan setiap tahunnya melalui Indeks Pembangunan Gender (IPG).

Berdasarkan data dari Badan Pusat Statistik (BPS) mengenai jumlah penduduk tahun 2022, Provinsi Jawa Barat memiliki jumlah penduduk terbanyak di Indonesia, dengan total sebanyak 49,4 juta jiwa. Meskipun memiliki jumlah penduduk terbesar, Indeks Pembangunan Gender (IPG) Provinsi Jawa Barat berada pada posisi ke-13 secara nasional, menunjukkan bahwa provinsi ini masih menghadapi tantangan dalam mencapai kesetaraan gender. Nilai IPG sebesar 89.8% menjadikan Provinsi Jawa Barat berada di posisi terendah di Pulau Jawa. IPG yang rendah juga akan mempengaruhi pertumbuhan ekonomi suatu daerah. Sehingga diperlukan kebijakan dari pemerintahan daerah untuk

mendorong Pertumbuhan Gender kabupaten/kota di provinsi Jawa Barat agar kesetaraan gender tercapai dan dapat memperbaiki pertumbuhan ekonomi di provinsi Jawa Barat.

Masalah kesenjangan IPG harus di atasi, maka perlu diketahui model yang terbentuk dari faktor yang mempengaruhi kesenjangan tersebut. Faktor tersebut perlu diperhatikan agar ketimpangan IPG dapat diatasi. Teknik statistika yang dapat digunakan untuk mengetahui, menjelaskan dan memodelkan pengaruh adalah analisis regresi (Masruroh dan Subekti, 2016). Dalam analisis regresi terdapat tiga jenis pendekatan regresi untuk mengestimasi kurva regresi, yaitu regresi parametrik, nonparametrik dan semiparametrik (Rencher, 2002). Regresi parametrik dalam analisisnya mempunyai asumsi yang ketat yaitu bentuk kurva harus membentuk suatu pola tertentu. Berbeda dari regresi parametrik, regresi nonparametrik mengasumsikan bentuk kurva yang tidak membentuk suatu pola tertentu. Sementara regresi semiparametrik adalah gabungan dari regresi parametrik dan regresi nonparametrik.

Pada penelitian ini, kurva yang terbentuk dari data tidak membentuk pola tertentu serta cenderung berulang. maka, untuk analisisnya digunakan regresi nonparametrik. Adapun metode regresi nonparametrik yang biasa digunakan pada data yang berulang adalah Deret Fourier (Prahutama, 2013). Deret Fourier dipopulerkan oleh Bilodeau pada Tahun 1992. Deret Fourier menggunakan konsep polinomial trigonometri yang mempunyai fleksibilitas tinggi, sehingga mampu menyesuaikan secara efektif terhadap sifat lokal data.

Penelitian ini bertujuan untuk memodelkan dan menggambarkan grafik dari IPG di Provinsi Jawa Barat menggunakan regresi nonparametrik estimator Deret Fourier. Penelitian sebelumnya yang menggunakan regresi nonparametrik pendekatan Deret Fourier antara lain, Prahutama (2013) melakukan pemodelan regresi nonparametrik dengan pendekatan Deret Fourier untuk kasus tingkat pengangguran terbuka di Jawa Timur dan diperoleh akurasi sebesar 96,67%. Dani dkk (2022) melakukan penelitian dengan regresi nonparametrik dengan pendekatan Deret Fourier untuk memodelkan kasus Rata-rata Lama Sekolah di Provinsi Nusa Tenggara Timur dan diperoleh akurasi sebesar 89,18%. Ketrin (2023) melakukan penelitian regresi nonparametrik dengan pendekatan Deret Fourier untuk memodelkan tingkat kemiskinan di Provinsi Sumatera Barat dengan akurasi sebesar 92,44%.

II. METODE PENELITIAN

A. Jenis Penelitian dan Sumber Data

Penelitian ini merupakan penelitian terapan yang menerapkan analisis regresi nonparametrik dengan menggunakan estimator Deret Fourier. Penelitian ini menggunakan data IPG beserta faktor-faktor yang diduga mempengaruhinya. Data ini diperoleh dari publikasi Badan Pusat Statistik Provinsi Jawa Barat Tahun 2022, yang diakses melalui *website* resmi BPS. Variabel yang digunakan adalah IPG (Y) dengan satuan persen, Umur Harapan Hidup (X_1) dengan satuan tahun, Angka Melek Huruf (X_2) dan Tingkat Pengangguran Terbuka (X_3) dengan satuan persen.

B. Teknik Analisis Data

Teknik analisis yang dilakukan dalam pemodelan Provinsi Jawa Barat dengan metode regresi nonparametrik estimator Deret Fourier adalah.

1. Mengumpulkan data IPG beserta faktor-faktor yang diduga mempengaruhinya dari situs resmi BPS
2. Melakukan analisis deskripsi dari variabel yang digunakan. Adapun analisis yang dilakukan yaitu, minimal, maksimal, rata-rata dan Standar Deviasi. Untuk mean menggunakan persamaan.

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} \text{ atau } \bar{X} = \frac{\sum_{i=1}^n X_i}{n} \quad (1)$$

Diman nilai \bar{X} rata-rata, nilai $\sum_{i=1}^n X_i$ jumlah seluruh nilai data dan n jumlah seluruh frekuensi dengan Standar Deviasi menggunakan rumus

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (2)$$

3. Membentuk *scatterplot* variabel dependen (Y) dengan setiap variabel independen (X) untuk mengetahui pola data yang akan dianalisis.
4. Melakukan standardisasi data. Standardisasi sangat disarankan Ketika nilai dari variabel yang akan dianalisis memiliki skala yang berbeda jauh maupun beda satuan dengan tujuan untuk memperkecil perbedaan nilai antar variabel (Kassambara, 2017). Untuk melakukan standardisasi dapat digunakan persamaan.

$$z_{ik} = \frac{x_{ik} - \bar{x}_k}{s_k} \quad (3)$$

dimana nilai z baris ke- i variabel ke- k (z_{ik}), nilai x baris ke- i variabel ke- k (x_{ik}), rata-rata variabel ke- k (\bar{x}_k), dan simpangan baku variabel ke- k (s_k).

- Melakukan pemodelan dengan regresi nonparametrik Deret Fourier dengan tiga parameter osilasi yaitu 1, 2 dan 3. Fungsi dari Deret Fourier dapat dilihat pada Persamaan (4).

$$f(x) = \frac{1}{2}\alpha_0 + bx + \sum_{k=1}^K \alpha_k \cos_k x \tag{4}$$

(Bilodeau, 1992)

Dimana $\alpha_0, b, \alpha_k, k = 1, 2, \dots, K$ merupakan parameter-parameter model. Adapun estimasi parameter model yang digunakan adalah *Least Square*. Metode *Least Square* bekerja dengan meminimkan jumlah kuadrat residual yang didapat dari model sehingga model dapat menjelaskan data dengan baik. Berdasarkan Persamaan (4) dalam bentuk matriks pada Persamaan (5).

Dengan:

$$\underline{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \cos x_{11} & \cdots & \cos_k x_{11} & \vdots & \cdots & \vdots & x_{j1} & \cos x_{j1} & \cdots & \cos_k x_{j1} \\ 1 & x_{12} & \cos x_{12} & \cdots & \cos_k x_{12} & \vdots & \cdots & \vdots & x_{j2} & \cos x_{j2} & \cdots & \cos_k x_{j2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & \cos x_{1n} & \cdots & \cos_k x_{1n} & \vdots & \cdots & \vdots & x_{jn} & \cos x_{jn} & \cdots & \cos_k x_{jn} \end{bmatrix} \begin{bmatrix} \frac{1}{2}\alpha_0 \\ b_j \\ \vdots \\ \alpha_{kj} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \tag{5}$$

Dari Persamaan (5) maka,

$$\underline{y} = \underline{X}\underline{\beta} + \underline{\varepsilon} \tag{6}$$

Sehingga diperoleh penduga $\hat{\underline{\beta}}$ dengan Persamaan (7)

$$\hat{\underline{\beta}} = (\underline{X}'_{(k)}\underline{X}_{(k)})^{-1}\underline{X}'_{(k)}\underline{y} \tag{7}$$

- Menghitung nilai GCV, MSE dan Koefisien Determinasi dari masing-masing parameter osilasi yang dicobakan. *Generalized Cross Validation (GCV)* adalah metode yang digunakan untuk menentukan jumlah optimal dari osilasi dalam pendekatan regresi nonparametrik, seperti pendekatan Deret Fourier. (Dani dan Adrianingsih, 2021). Rumus GCV ditunjukkan pada persamaan:

$$GCV = \frac{MSE(\lambda)}{(n^{-1}\text{trace}(I-H(x)))^2} \tag{8}$$

$$H(x) = X(X'X)^{-1}X'$$

(Eubank, 1999:84)

Dimana matriks identitas (I), matriks Hessian ($H(x)$) dan jumlah pengamatan (n), dengan

$$MSE(\lambda) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \tag{9}$$

y_i , adalah variabel dependen ke- i dan prediksi variabel dependen ke- i (\hat{y}_i). Nilai K yang optimal dihasilkan berdasarkan nilai GCV terkecil (Prahutama, 2013). Adapun rumus dari Koefisien Determinasi (R^2) dapat dilihat pada Persamaan (10).

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \tag{10}$$

(Montgomery, 2012:332)

Dengan variabel dependen ke- i (y_i), prediksi variabel dependen ke- i (\hat{y}_i) dan rata-rata variabel dependen (\bar{y}). Dalam penjelasannya, semakin tinggi nilai R^2 yang diperoleh dari suatu model atau semakin mendekati 100%, maka model regresi yang diperoleh semakin akurat pula.

- Memperoleh model terbaik berdasarkan parameter osilasi dengan nilai GCV, MSE terendah dan Koefisien Determinasi tertinggi.
- Membentuk grafik dari model terbaik yang telah diperoleh.

III. HASIL DAN PEMBAHASAN

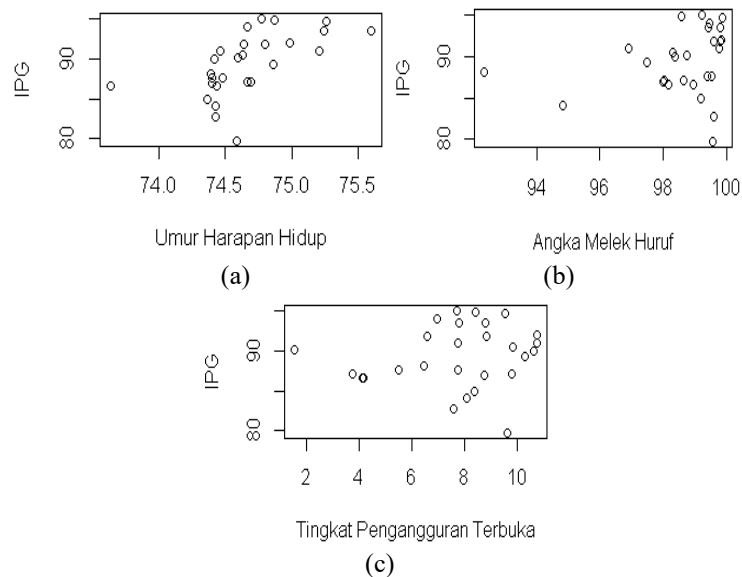
A. Statistik Deskriptif

Gambaran umum dari variabel yang digunakan pada penelitian ini dapat dilihat pada Tabel 1.

Tabel 1. Statistik Deskriptif

Variabel	Maksimal	Minimal	Rata-Rata	Standar Deviasi
Y	95.09	79.69	89.3589	3.91416
X ₁	75.60	73.64	74.6648	.37898
X ₂	99.88	92.34	98.5656	1.67336
X ₃	10.78	1.56	7.8011	2.33196

Berdasarkan Tabel 1, dapat dilihat variabel X₂ memiliki nilai maksimal 99,88 persen dibandingkan variabel X lainnya di Kota Bandung dengan nilai minimal 92,34 persen di Kabupaten Indramayu. Sedangkan variabel Y memiliki nilai maksimal dengan nilai 95,09 persen pada Kabupaten Sumedang dan nilai minimal pada Kabupaten Bandung Barat 79.69 persen. Sementara X₃ memiliki nilai terendah dibandingkan variabel X lainnya, yaitu pada Kabupaten Pangandaran dengan nilai 1,56 persen. Untuk mengetahui pola data antara IPG dengan setiap variabel independennya, digunakan *scatterplot* seperti yang terlihat pada Gambar 1.



Gambar 1. *Scatterplot* IPG dan umur harapan hidup (a), *Scatterplot* IPG dan angka melek huruf (b), dan *Scatterplot* IPG dan tingkat pengangguran terbuka di Provinsi Jawa Barat.

Gambar 1 menunjukkan bahwa data antara IPG dengan masing-masing variabel independennya tidak menunjukkan adanya pola tertentu. Selain itu, pola data yang terbentuk cenderung berulang. Maka untuk penelitian ini, dengan pola data tidak diketahui dan data yang cenderung berulang analisis yang digunakan adalah regresi nonparametrik Deret Fourier. Selain itu, variabel yang diamati memiliki satuan yang berbeda dan perlu dilakukan standarisasi. Perhitungan standarisasi dilakukan dengan Persamaan 11, hasil standarisasi ditunjukkan pada Tabel 2.

Tabel 2. Standarisasi Data

Provinsi	Y	X ₁	X ₂	X ₃
Bogor	0.1510	-0.6459	-0.1109	1.2174
Sukabumi	-0.4391	-0.6987	0.5763	-0.0133
Cianjur	-1.1392	-0.7779	0.3672	0.2611
Bandung	1.1781	0.0137	0.5405	-0.3521
Garut	-1.6808	-0.6196	0.6182	-0.0862
Tasikmalaya	-0.6946	-2.7041	0.2357	-1.5571
Ciamis	-0.5924	0.0664	0.0385	-1.7372
Kuningan	-0.5592	0.0137	-0.3200	0.8615

Provinsi	Y	X ₁	X ₂	X ₃
Cirebon	-1.3665	-0.6196	-2.2443	0.1325
⋮	⋮	⋮	⋮	⋮
Kota Banjar	-0.4315	-0.4877	0.4987	-0.9739

B. Deret Fourier

Langkah pertama melakukan pemodelan regresi pendekatan deret fourier adalah menentukan parameter osilasi (K) dengan nilai terbaik menggunakan nilai GCV, MSE terkecil dan Koefisien Determinasi tertinggi. Dimana nilai K yang digunakan ialah 1, 2 dan 3. Tabel 3 menunjukkan nilai GCV, MSE dan Koefisien Determinasi pada masing-masing nilai K yang diperoleh.

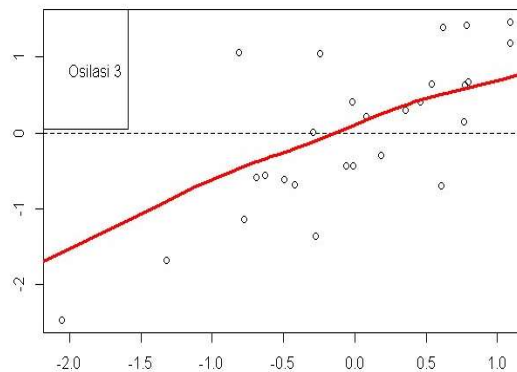
Tabel 3. Standardisasi Data

Nilai K	GCV	MSE	Koefisien Determinasi
K = 1	11.3874	0.5623	41.60%
K = 2	5.0230	0.5581	42.04%
K = 3	2.8084	0.5547	42.39%

Berdasarkan Tabel 3, dapat diketahui bahwa nilai K = 3 memiliki nilai GCV, MSE terendah dan koefisien determinasi tertinggi dibandingkan nilai K lainnya. Maka nilai K yang akan digunakan untuk melakukan pemodelan IPG di Provinsi Jawa Barat dengan masing-masing variabel independennya adalah K = 3. Adapun model dari parameter menggunakan tiga parameter osilasi yang dihasilkan adalah sebagai berikut.

$$\hat{y} = -0.740 + 0.650x_{1i} - 0.287\cos x_{1i} + 0.024\cos 2x_{1i} + 0.031\cos 3x_{1i} - 0.361x_{2i} + 0.885\cos x_{2i} + 0.334\cos 2x_{2i} - 0.235\cos 3x_{2i} - 0.121x_{3i} + 0.152\cos x_{3i} + 0.155\cos 2x_{3i} + 0.033\cos 3x_{3i}$$

Setelah diperoleh model, scatterplot yang memuat kurva regresi dengan nilai K dari model terbaik yang terpilih dapat disajikan. Adapun *scatterplot* model terbaik disajikan pada Gambar 2.



Gambar 2. Scatterplot dari model terbaik dengan K=3

Gambar 2 menunjukkan grafik yang diperoleh dari model terbaik dengan nilai K=3. Nilai GCV terendah dan nilai R² tertinggi diperoleh ketika K=3. Hal ini dikarenakan kurva regresi dengan parameter osilasi K=3 mampu mengikuti pola data asli dengan baik. Model terbaik yang dipilih adalah model dengan tiga parameter osilasi dengan R² 42.39%.

IV. KESIMPULAN

Model terbaik digunakan adalah model dengan tiga parameter osilasi. Berdasarkan grafik tersebut dapat diprediksi penurunan atau peningkatan IPG ditahun-tahun berikutnya berdasarkan persebaran data. Ketika persebaran data berada dibawah grafik, dapat diidentifikasi bahwa IPG mengalami penurunan. Sebaliknya, ketika data berada diatas grafik dapat diidentifikasi bahwa IPG mengalami peningkatan. Informasi ini berguna untuk mengetahui kondisi IPG jika terjadi penurunan atau peningkatan, berdasarkan tren yang diperlihatkan oleh model regresi nonparametrik Deret Fourier dengan tiga parameter osilasi. Sehingga, pemerintah dapat mengambil keputusan agar penurunan IPG yang akan terjadi ditahun berikutnya dapat dicegah. Hasil pemodelan menunjukkan bahwa analisis ini tidak efektif dalam memodelkan IPG di Provinsi Jawa Barat. Hal ini dapat diketahui berdasarkan nilai koefisien determinasi yang kecil, yaitu 42.39%. Nilai ini mengindikasikan bahwa variabel umur harapan hidup, angka melek huruf, dan tingkat pengangguran terbuka hanya mampu menjelaskan variasi atau persentase kemiskinan sebesar 42.39%.

DAFTAR PUSTAKA

- Badan Pusat Statistik. (2022). *Kajian Penghitungan Indeks Ketimpangan Gender 2022*. (diunduh tanggal 1 Februari 2024).
- Bilodeau, M. (1992). Fourier Smoother and Additive Models, *The Canadian Journal of Statistics*, 20(3), 257-269.
- Dani, A. T. R., Dewi, A. F., dan Ni'matuzzahroh, L. (2022). "Estimator Deret Fourier pada Pemodelan Regresi Nonparametrik", *Prosiding Seminar Nasional Matematika, Statistika dan Aplikasinya*, 2, 279-288.
- Dani, A. T. R., dan Adrianingsih, N. Y. (2021). Pemodelan Regresi Nonparametrik Dengan Estimator Spline Truncated dan Deret Fourier. *Jambura Journal Of Mathematics*, 3(1): 26-36.
- Eubank, R. L. (1999). *Nonparametric Regression and Spline Smoothing*. 2th edition, New York: Marcel Dekker.
- Hadiyanti, R. 2021. Korelasi Antara Ketimpangan Gender Bidang Pendidikan dan Ketenagakerjaan Terhadap Laju Pertumbuhan Ekonomi di Kabupaten Belitang Timur. *Equity: Jurnal Ekonomi*, 9 (2): 1-8.
- Kassambara, A. (2017). *Practical Guide To Cluster Analysis in R Unsupervised Machine Learning*. USA: STHDA.
- Ketrin, M. W., Fitri, F., Putra, A. A., Zilrahmi. (2023). Nonparametric Regression Modeling with Fourier Series Approach on Poverty Cases in West Sumatra Province. *UNP Journal Of Statistics and Data Science*, 1(2): 53-58.
- Masruroh, M., dan Subekti, R. (2016). Aplikasi Regresi Partial Least Square untuk Analisis Hubungan Faktor-Faktor yang Mempengaruhi Indeks Pembangunan Manusia Di Kota Yogyakarta. *Media Statistika* 9(2): 75-84.
- Montgomery, D. C., Peck, E. A., dan Vining, G. G. (2012). *Introduction to Linear Regression Analysis*, Canada: John Wiley & Sons, Inc.
- Prahutama, A., (2013). "Model Regresi Nonparametrik dengan Pendekatan Deret Fourier pada Kasus Tingkat Pengangguran Terbuka di Jawa Timur". *Prosiding Seminar Nasional Statistika. Undip*, 10: 69-76.
- Rencher, A. C. (2002). *Methods of Multivariate Analysis*, Second Edition. New York: John Wiley & Sons, Inc.
- Samad, Z., Mustanir, A., Pratama, M. Y. P. (2019). Partisipasi Masyarakat dalam Musyawarah Rencana Pembangunan untuk Mewujudkan *Good Governance Kabupaten Enrekang*. *Jurnal MODERAT*, 5(4):379-395.
- Zuhri, S., dan Amalia. D. 2022. Ketidakadilan Gender Dan Budaya Patriarki Di Kehidupan Masyarakat Indonesia. *Murabbi : Jurnal Ilmiah dalam Bidang Pendidikan* 5(1): 17-41.