



eISSN 3090-6946 & pISSN 3090-6938

JURNAL TEOLOGI ISLAM

Vol. 1, No. 2, Tahun 2025

doi.org/10.63822/4kgpwy25

Hal. 258-274

Available online at <https://ojs.indopublishing.or.id/index.php/jti>

Analisis Korelasi, Regresi, dan Anova : Serta Contoh Penggunaannya dalam Penelitian Pendidikan Bahasa Arab

Ummu Salamah^{1*}, Fatmah Hambali², Ahmad Royani³

Magister Pendidikan Bahasa Arab, UIN Syarif Hidayatullah, Tangerang Selatan, Indonesia^{1,2,3}

*Email Korespondensi: ummhusal@gmail.com

Diterima: 07-07-2025 | Disetujui: 13-07-2025 | Diterbitkan: 15-07-2025

ABSTRACT

This article discusses three essential statistical methods in quantitative research, namely correlation analysis, regression, and ANOVA (Analysis of Variance), along with examples of their application in Arabic Language Education research. Correlation analysis is used to measure the relationship between two variables, regression aims to build predictive models between independent and dependent variables, while ANOVA is used to compare the means of three or more groups. This article systematically outlines the basic concepts of each technique, their differences, and statistical testing procedures. Equipped with real examples from research in the field of Arabic Language Education, this paper is expected to serve as a practical guide for students and researchers in accurately and effectively applying statistical analysis.

Keywords: Correlation, Regression, ANOVA, Statistics, Arabic Language Education

ABSTRAK

Artikel ini membahas tiga metode statistik penting dalam penelitian kuantitatif, yaitu analisis korelasi, regresi, dan ANOVA (Analysis of Variance), serta contoh penerapannya dalam penelitian Pendidikan Bahasa Arab. Analisis korelasi digunakan untuk mengukur hubungan antara dua variabel, regresi untuk membangun model prediktif antara variabel bebas dan terikat, sedangkan ANOVA digunakan untuk membandingkan rata-rata dari tiga kelompok atau lebih. Artikel ini menguraikan konsep dasar masing-masing teknik, perbedaannya, serta prosedur pengujian statistik secara sistematis. Dilengkapi dengan contoh-contoh nyata dari penelitian di bidang Pendidikan Bahasa Arab, tulisan ini diharapkan dapat menjadi panduan praktis bagi mahasiswa dan peneliti dalam menerapkan analisis statistik secara tepat dan valid.

Kata kunci: Korelasi, Regresi, ANOVA, Statistik, Pendidikan Bahasa Arab

Bagaimana Cara Sitasi Artikel ini:

Ummu Salamah, Fatmah Hambali, & Ahmad Royani. (2025). Analisis Korelasi, Regresi, dan Anova : Serta Contoh Penggunaannya dalam Penelitian Pendidikan Bahasa Arab. Jurnal Teologi Islam, 1(2), 258-274. <https://doi.org/10.63822/4kgpwy25>

PENDAHULUAN

Dalam penelitian kuantitatif, pemahaman terhadap hubungan antar variabel menjadi aspek penting dalam menggali, menjelaskan, dan memprediksi fenomena yang terjadi. Tiga teknik statistik yang umum digunakan untuk tujuan tersebut adalah analisis korelasi, regresi, dan ANOVA (Analysis of Variance). Ketiganya memiliki peran yang saling melengkapi dalam mengolah dan menafsirkan data.

Analisis korelasi digunakan untuk mengetahui sejauh mana hubungan antara dua variabel, apakah hubungan tersebut bersifat positif, negatif, atau tidak ada hubungan sama sekali. Sementara itu, analisis regresi tidak hanya mengukur hubungan, tetapi juga membangun model prediksi antara variabel bebas (independen) dan variabel terikat (dependen). Di sisi lain, ANOVA digunakan untuk membandingkan rata-rata lebih dari dua kelompok dan menentukan apakah terdapat perbedaan yang signifikan antar kelompok tersebut.

Makalah ini bertujuan untuk menguraikan konsep dasar, perbedaan, serta contoh penerapan dari analisis korelasi, regresi, dan ANOVA secara sistematis. Diharapkan pembahasan ini dapat memberikan gambaran yang jelas dan menjadi referensi praktis dalam melakukan analisis statistik dasar.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif deskriptif dengan metode studi pustaka dan dokumentasi, yang bertujuan menjelaskan secara sistematis konsep, perbedaan, dan prosedur penggunaan analisis korelasi, regresi, dan ANOVA. Data dan contoh yang dikaji bersumber dari hasil-hasil penelitian sebelumnya dalam bidang Pendidikan Bahasa Arab, seperti tesis dan studi akademik terkait.

PEMBAHASAN

A. Pengertian Pengujian Korelasi

Untuk mengukur hubungan antara satu variabel dengan variabel lainnya dilakukan analisa korelasi. Analisa korelasi adalah istilah analisis statistik yang digunakan untuk mengukur kekuatan hubungan antara satu variabel dengan variabel lainnya.

Penjelasan Supardi mengatakan bahwa analisa korelasi adalah bagian dari analisis inferensial, yang bertujuan untuk mendapati kekuatan suatu hubungan, bentuk atau hubungan kausal dan hubungan timbal-balik antara variabel-variabel yang sedang diteliti.

Analisa korelasi dibagi menjadi dua, *bivariate correlation* yakni hubungan antara dua variabel, dan *multivariate correlation* yakni mencari hubungan antara tiga variabel atau lebih. Tujuan korelasi adalah untuk mengetahui adanya hubungan antar variabel, namun menurut pendapat Walter R. Bog dan Meredith D. Gall korelasi tidak dapat digunakan untuk hubungan sebab akibat antara satu variabel dengan variabel yang lain.

Dalam korelasi ada dua arah, pertama adalah korelasi searah atau korelasi positif, dan kedua adalah korelasi berlawanan arah atau korelasi negative. Korelasi searah adalah jika satu variabel naik, maka variabel lainnya naik, dan jika satu variabel turun, maka variabel lainnya juga turun. Sedangkan korelasi berlawanan adalah ketika satu variabel naik maka variabel yang lain akan turun. Contoh dari korelasi searah, “Hubungan antara Penguasaan Kosakata Bahasa Arab dengan Kemampuan Menulis Bahasa Arab”,

semakin tinggi penguasaan kosakata, semakin tinggi pula kemampuan menulis bahasa Arabnya, maka kedua variabel bergerak searah atau positif. Sedangkan contoh dari korelasi berlawanan adalah, “Hubungan antara Tingkat Kecemasan Berbicara dengan Kemampuan Berbicara Bahasa Arab”, semakin tinggi Tingkat kecemasan, maka biasanya akan semakin rendah kemampuan berbicara bahasa Arabnya. Jadi hubungan antara kedua variabel bersifat berlawanan arah atau negatif.

Selain arah hubungan, dalam korelasi ada juga angka korelasi. Angka korelasi atau koefisien korelasi menunjukkan seberapa kuat dan seberapa erat hubungan antara dua variabel, dengan nilai berkisar antara 0 hingga $\pm 1,00$. Nilai ini tidak boleh melebihi $+1,00$ —jika terjadi, berarti ada kesalahan perhitungan. Tanda positif (+) menunjukkan hubungan searah atau korelasi positif, artinya jika satu variabel meningkat, variabel lainnya juga cenderung meningkat. Sebaliknya, tanda negatif (–) menunjukkan hubungan berlawanan arah atau korelasi negatif, yaitu jika satu variabel naik, yang lainnya cenderung turun.

1. **Korelasi Product Moment**

Korelasi product moment yang dikembangkan oleh Karl Pearson, adalah teknik statistik untuk mengetahui hubungan antara dua variabel interval atau rasio dengan asumsi hubungan linier. Teknik ini digunakan jika sampel diambil secara acak dan kedua variabel memiliki hubungan linier. Koefisien korelasi ini menunjukkan sejauh mana hubungan antara dua variabel, dan hasilnya berkisar antara -1 hingga 1. Jika koefisien mendekati 1 atau -1, hubungan antar variabel semakin kuat, sedangkan jika mendekati 0, maka hubungannya lemah atau tidak ada.

Perhitungan koefisien korelasi dapat dilakukan dengan beberapa cara. Cara pertama menggunakan deviasi skor dan standar deviasi masing-masing variabel. Cara kedua, tanpa menghitung standar deviasi terlebih dahulu, cukup dengan memasukkan data ke dalam rumus sederhana. Untuk sampel besar ($N \geq 30$), disarankan menggunakan peta korelasi (scatter diagram) karena perhitungan manual menjadi kurang praktis dan diagram ini juga membantu dalam melihat pola hubungan (linier atau tidak) antar variabel.

Rumus:

a. Rumus 1

Rumus ini digunakan ketika nilai yang ingin dihitung adalah nilai mentah.

$$r_{xy} = \frac{N \sum xy - (\sum x)(\sum y)}{\sqrt{[N \sum x^2 - (\sum x)^2][N \sum y^2 - (\sum y)^2]}}$$

- r_{xy} : Koefisien korelasi antara variabel X dan Y
- N : Jumlah pasangan data
- $\sum xy$: Jumlah hasil perkalian nilai X dan Y
- $\sum x$: Jumlah seluruh nilai X
- $\sum y$: Jumlah seluruh nilai Y
- $\sum x^2$: Jumlah kuadrat dari seluruh nilai X
- $\sum y^2$: Jumlah kuadrat dari seluruh nilai Y

b. Rumus 2

Rumus ini digunakan ketika sudah menghitung deviasi dan standar deviasi dari masing-masing variabel.

$$r_{xy} = \frac{\sum(x \cdot y)}{N \cdot S_{dx} \cdot S_{dy}}$$

Keterangan:

- r_{xy} = koefisien korelasi antara X dan Y
- x, y = deviasi dari skor masing-masing variabel
- S_{dx}, S_{dy} = standar deviasi variabel X dan Y
- N = jumlah sampel

2. Korelasi Rank/ Spearman

Teknik korelasi Rank atau Spearman digunakan ketika data yang dianalisis berbentuk ordinal. Berbeda dengan korelasi Pearson (product moment) yang mengukur hubungan linier antara dua variabel, korelasi Spearman tidak menuntut adanya hubungan linier dalam data yang dikorelasikan.

Rumus:

$$\rho = 1 - \frac{6 \sum D^2}{n(n^2 - 1)}$$

Keterangan:

D = difference (selisih antara X dan Y)

n = jumlah responden

Contoh data:

a. Membuat Tabel

X	Y	D	D ²
5	2	3	9
4	3	1	1
5	4	1	1
3	2	1	1
3	3	0	0
2	6	-4	16
4	3	1	1
3	5	-2	4
			$\sum D^2 = 33$

b. Mencari Harga Nilai Koefisien Spearman

$$\begin{aligned}\rho &= 1 - \frac{6 \sum D^2}{n(n^2 - 1)} \\ &= 1 - \frac{6 \times 33}{8(64 - 1)} \\ &= 1 - \frac{198}{504} \\ &= 1 - 0,39 \\ &= 0,61\end{aligned}$$

c. Menguji Signifikansi Korelasi Spearman

Rumus:

$$t = \rho \sqrt{\frac{n-2}{1-\rho^2}}$$

Masukkan perhitungan ke rumus:

$$\begin{aligned}&= 0,61 \sqrt{\frac{8-2}{1-0,61^2}} \\ &= 0,61 \sqrt{\frac{6}{1-0,37}} \\ &= 0,61 \sqrt{\frac{6}{0,63}} \\ &= 0,61 \times 3,08 \\ &= 1,88\end{aligned}$$

d. Membandingkan t Hitung dengan t Tabel

Membandingkan harga hitung (1,88) dengan harga pada tabel t dengan db = 8 – 2 = 6 dan taraf signifikansi 5% diperoleh harga t tabel = 2,45. Oleh karena harga t hitung 1,88 lebih kecil dari harga t tabel 2,45 maka hipotesis alternatif ditolak.

3. Korelasi Kendal Tau (τ)

Korelasi Kendall Tau digunakan untuk menganalisis hubungan antara dua data yang berskala ordinal. Hipotesis yang diuji adalah:

H₀: $\tau = 0$ (tidak ada hubungan)

H₁: $\tau \neq 0$ (ada hubungan)

Pengujian dilakukan dengan menghitung nilai z, yaitu dengan membandingkan z hitung dan z tabel.

Jika $z \text{ hitung} > z \text{ tabel}$, maka H_0 ditolak

Jika $z \text{ hitung} < z \text{ tabel}$, maka H_0 diterima

Untuk menghitung $z \text{ hitung}$, nilai τ dikonversi terlebih dahulu menggunakan rumus. Contoh bisa digunakan untuk meneliti hubungan antara kecerdasan intelektual (X) dengan prestasi kerja (Y). Langkah pertama adalah menyusun data X dari nilai tertinggi ke terendah, lalu menghitung nilai Kendall Tau dan mengubahnya menjadi $z \text{ hitung}$ untuk pengujian hipotesis

Rumus Kendal Tau (τ):

$$\tau = \frac{\sum X - \sum Y}{\frac{n(n-1)}{2}}$$

Di mana:

τ = koefisien korelasi Kendal Tau

$\sum X = \sum R_x$ = jumlah rangking atas

$\sum Y = \sum R_y$ = jumlah rangking bawah

n = banyaknya sampel (pasangan data)

Rumus mengkonversi nilai τ menjadi $z \text{ hitung}$:

$$z = \frac{\tau}{\sqrt{\frac{2(2n+5)}{9n(n-1)}}}$$

4. Korelasi Phi

Teknik korelasi Phi digunakan untuk menganalisis hubungan antara dua variabel diskrit yang masing-masing hanya memiliki dua kategori, seperti laki-laki dan perempuan, hidup dan mati, atau lulus dan tidak lulus. Lambang korelasi Phi adalah ϕ , dan untuk menghitung nilainya, terlebih dahulu harus dihitung nilai Kai Kuadrat (χ^2).

Rumus Korelasi Phi:

$$\phi = \sqrt{\frac{\chi^2}{N}}$$

Rumus Kai Kuadrat:

$$\chi^2 = \frac{\sum (f_o - f_t)^2}{f_t}$$

f_o = Frekuensi dari hasil observasi

f_t = Frekuensi teoritik

Sebagai contoh, dalam sebuah penelitian yang ingin mengetahui apakah ada hubungan yang signifikan antara asal sekolah dan hasil tes penerimaan mahasiswa baru, maka langkah awal yang dilakukan adalah menghitung nilai Kai Kuadrat. Berbeda dengan teknik korelasi lainnya, pada korelasi Phi, uji signifikansi dilakukan terhadap nilai Kai Kuadrat, bukan langsung terhadap nilai koefisien korelasinya.

Selain menggunakan rumus dasar yang melibatkan nilai χ^2 , korelasi Phi juga dapat dihitung dengan rumus alternatif berdasarkan nilai-nilai frekuensi dalam tabel kontingensi.

Rumus:

$$\phi = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$$

Contoh Data:

Hasil Test Sipenmaru	Asal Sekolah		Jumlah
	SMA Negeri	SMA Swasta	
Lulus	37	57	94
Tidak Lulus	28	94	122
Jumlah	65	151	N = 216

5. Korelasi Kontingensi (KK)

Teknik Korelasi Kontingensi (KK) digunakan untuk menyelidiki hubungan antara dua variabel yang masing-masing dinyatakan dalam bentuk kualitatif (kategori).

Dalam analisis ini, diperlukan perhitungan frekuensi yang diharapkan, yang selanjutnya digunakan untuk mencari nilai korelasi.

Langkah selanjutnya adalah membandingkan nilai Kai Kuadrat observasi (χ^2_o) dengan nilai kritik Kai Kuadrat (χ^2_t) yang tercantum dalam tabel statistik. Interpretasinya adalah sebagai berikut:

- Jika nilai χ^2_o lebih besar atau sama dengan χ^2_t , maka terdapat perbedaan yang signifikan antara kedua variabel. Artinya, hipotesis nol (H_o) yang menyatakan tidak ada hubungan signifikan ditolak.
- Sebaliknya, jika nilai χ^2_o lebih kecil dari χ^2_t , maka tidak ditemukan perbedaan yang signifikan. Ini menunjukkan bahwa tidak terdapat korelasi yang signifikan antara kedua variabel, sehingga hipotesis nol diterima.

Rumus Koefisien Kontingensi:

$$KK = \sqrt{\frac{\chi^2}{\chi^2 + N}}$$

Keterangan:

KK = Koefisien Kontingensi

χ^2 = Chi kuadrat.

Rumus Frekuensi Harapan:

$$f_t = \frac{\text{Total Baris}}{N} \times (\text{Total Kolom})$$

Jika salah satu variabel memiliki skala ordinal dan variabel lainnya berskala interval, maka teknik yang digunakan untuk menghitung hubungan keduanya adalah korelasi serial. Apabila variabel ordinal

tersebut hanya memiliki dua tingkatan (kategori), maka nilai korelasinya dihitung menggunakan rumus *serial dichotomized with standard scores* (rdiws).

Rumus:

$$r_{serial} = \frac{M_1 - M_0}{s} \cdot \sqrt{\frac{pq}{n}}$$

Keterangan:

- M_1 = rata-rata skor interval untuk kelompok pertama (misalnya kategori "tinggi")
- M_0 = rata-rata skor interval untuk kelompok kedua (misalnya kategori "rendah")
- s = simpangan baku (standar deviasi) seluruh data interval
- p = proporsi subjek dalam kelompok pertama
- q = proporsi subjek dalam kelompok kedua
- n = jumlah total subjek

6. Korelasi Ganda

Korelasi ganda digunakan ketika penelitian ingin mengetahui hubungan antara dua atau lebih variabel bebas terhadap satu variabel terikat. Dengan kata lain, variabel bebas yang dianalisis berjumlah dua atau lebih. Korelasi ganda menunjukkan seberapa kuat dan ke mana arah hubungan antara beberapa variabel bebas secara bersama-sama dengan satu variabel terikat. Nilai hubungan ini dikenal sebagai koefisien korelasi ganda dan biasanya dilambangkan dengan huruf R.

Rumus:

$$R_{x_1x_2y} = \sqrt{\frac{r^2 x_1 y + r^2 x_2 y - 2 r x_1 y r x_2 y r x_1 x_2}{1 - r^2 x_1 x_2}}$$

Di mana :

$R_{x_1x_2y}$ = koefisien korelasi ganda X_1 dan X_2 bersama-sama dengan Y

r_{x_1y} = koefisien korelasi antara X_1 dengan Y

r_{x_2y} = koefisien korelasi antara X_2 dengan Y

$r_{x_1x_2}$ = koefisien korelasi antara X_1 dengan X_2

7. Korelasi Parsial

Korelasi parsial digunakan ketika peneliti ingin mengetahui seberapa besar pengaruh murni dari variabel bebas terhadap variabel terikat. Biasanya, analisis ini digunakan bersamaan dengan korelasi ganda. Jika korelasi ganda digunakan untuk melihat hubungan secara kolektif antara dua atau lebih variabel bebas dengan variabel terikat, maka korelasi parsial bertujuan menelusuri hubungan masing-masing variabel bebas secara terpisah terhadap variabel terikat. Nilai dari korelasi parsial disebut koefisien korelasi parsial, yaitu angka yang menggambarkan arah dan kekuatan hubungan antara dua atau lebih variabel setelah mengontrol atau menetralkan pengaruh dari variabel lain yang diduga ikut memengaruhi hubungan tersebut.

Rumus:

$$r_{y12} = \frac{(ry_1 - ry_2 r_{12})}{\sqrt{(1 - r^2 y_2)(1 - r^2 y_{12})}}$$

Uji Keberartian Koefisien Korelasi Parsial:

$$t = \frac{ry_{12} \sqrt{n-3}}{\sqrt{1 - r^2 y_{12}}}$$

B. Pengertian Analisis Varians (ANOVA)

Analisis Varians (ANAVA) atau disebut juga dalam bahasa Inggris *Analysis of Variance* (ANOVA) adalah metode statistik yang digunakan untuk membandingkan tiga kelompok sampel atau lebih, dan termasuk dalam jenis penelitian komparatif. Penelitian komparatif sebaiknya tidak hanya menunjukkan ada atau tidaknya perbedaan, tetapi juga menjelaskan penyebab perbedaan tersebut. Misalnya, jika ditemukan perbedaan frekuensi membaca antara siswa pria dan wanita, maka penelitian perlu mengungkap alasan di balik perbedaan itu agar dapat menunjukkan hubungan sebab-akibat.

Sebelum melakukan Analisis Varians, terdapat beberapa syarat yang harus dipenuhi, yaitu:

- Terdapat perbedaan antar kelompok yang ingin dibandingkan.
- Data yang digunakan berada dalam skala interval atau rasio untuk kedua variabel yang dianalisis.
- Variansi antar kelompok harus homogen atau setidaknya mendekati homogen.
- Data yang dianalisis berdistribusi normal, meskipun normalitas ini tidak menjadi syarat mutlak

Langkah Uji Hipotesis dalam ANOVA:

- Hipotesis

$$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = \dots = \sigma_a^2$$

$$H_a: \exists \sigma_i^2 \neq \sigma_j^2, i \neq j, i = 1, 2, 3, 4, \dots, a$$

- Taraf nyata (signifikansi) : α dan tingkat kepercayaan = $1 - \alpha$
- Menentukan Uji statistik F (dalam artian Fhit)
- Uji signifikansi

Fhit > Ftab: : H0 ditolak artinya ada yang berbeda dari beberapa variansi.

Fhit < Ftab: : H0 diterima artinya semua variansinya sama

Sig < 0.05 : H0 ditolak

Sig > 0.05 : H0 diterima

1. ANOVA Satu Arah (One Way)

ANOVA satu arah, atau yang sering disebut ANOVA satu jalur (one-way ANOVA), digunakan untuk menganalisis masalah yang melibatkan dua variabel: satu variabel independen dan satu variabel dependen. Variabel independen minimal harus memiliki tiga kategori, sementara variabel dependen harus berskala interval atau rasio dan diasumsikan berdistribusi normal.

ANOVA Satu Arah

- Membuat tabel kerja dengan mencari K1, K2, K3. Dengan K adalah nilai dari kategori.
- Mencari jumlah kuadrat total dengan rumus : $DK_{total} = \{EX^2\} - \{EX^2\}/N$

3. Mencari jumlah kuadrat antara dengan rumus $DK_{\text{ant}} = (\sum X_1^2)/n_1 + (\sum X_2^2)/n_2 + (\sum X_3^2)/n_3 - (\sum X^2)/N$
4. Mencari jumlah kuadrat dalam dengan rumus : $DK_{\text{dal}} = DK_{\text{total}} - DK_{\text{ant}}$
5. Mencari mean kuadrat antara kelompok dengan rumus: $MK_{\text{ant}} = DK_{\text{ant}}/db_{\text{ant}}$
6. Mencari mean kuadrat dalam kelompok dengan rumus : $MK_{\text{dal}} = DK_{\text{dal}}/db_{\text{dal}}$
7. Mencari Fhitung dengan rumus: $MK_{\text{ant}}/MK_{\text{dal}}$
8. Mencari F table
9. Membandingkan Fhitung dengan Ftabel. Jika Fhitung lebih besar maka hasilnya signifikan

2. ANOVA Dua Arah

Analisis varians dua arah, yang juga dikenal sebagai anava dua jalur (two-way ANOVA), digunakan untuk meneliti permasalahan yang melibatkan dua variabel bebas. Masing-masing variabel ini terdiri atas beberapa kelompok. Menurut Irianto (2004), terdapat beberapa asumsi yang perlu dipenuhi dalam penerapan analisis varians dua arah, yaitu:

- a. Distribusi normal diharapkan pada setiap skor dalam tiap sel.
- b. Skor pada setiap kelompok sel sebaiknya memiliki variansi yang seragam atau homogen.
- c. Skor yang digunakan seharusnya tidak dipengaruhi oleh variabel luar yang tidak diteliti.

Anova Dua Arah:

1. Mencari Jumlah Kuadrat (JK) dengan rumus: $JKT = \sum X^2. (\sum X)^2/N$
2. Mencari derajat kebebasan dengan rumus: $dk_{\text{antar kelompok}} = \text{banyak kelompok} - 1$
3. Menghitung rata-rata jumlah kuadrat dengan rumus: $RJK = JK_{\text{antar Kelompok}}/dk_{\text{Antar Kelompok}}$
4. Menentukan Fhitung dengan Ftabel

C. Pengertian Pengujian Regresi

Istilah regresi pertama kali diperkenalkan oleh Galton. Menurut Sudjana (1983), konsep regresi dapat dijelaskan melalui ilustrasi yang menunjukkan bahwa seorang peneliti seringkali menghadapi situasi yang melibatkan dua atau lebih variabel yang saling berkaitan atau diduga memiliki hubungan tertentu. Dalam dunia pendidikan, misalnya, peneliti mungkin ingin memprediksi nilai akhir siswa berdasarkan hasil tes masuk, menilai peningkatan daya ingat siswa berdasarkan usaha belajar yang ditingkatkan, atau memperkirakan kemampuan mengajar guru berdasarkan tingkat motivasinya.

Berdasarkan uraian tersebut, regresi dapat dipahami sebagai bentuk hubungan antara variabel prediktor dan variabel respons. Hubungan ini biasanya diwujudkan dalam bentuk persamaan matematis yang dapat bersifat linier atau non-linier. Persamaan ini berfungsi sebagai alat prediksi yang menggambarkan keterkaitan antara variabel prediktor dengan variabel respons.

Namun, seperti halnya makna kata prediksi, hasil dari persamaan regresi bukanlah sesuatu yang pasti, melainkan hanya mendekati kenyataan atau bersifat estimatif (Irianto, 2004).

Regresi memiliki kaitan yang erat dengan korelasi. Setiap analisis regresi selalu mengandung unsur korelasi, tetapi tidak semua korelasi diikuti oleh analisis regresi. Korelasi yang tidak dilanjutkan dengan regresi biasanya terjadi pada variabel-variabel yang tidak memiliki hubungan sebab-akibat atau hubungan

fungsional. Oleh karena itu, analisis regresi hanya tepat dilakukan ketika terdapat hubungan kausal atau fungsional antara variabel-variabel yang dikaji.

1. Regresi Sederhana/Tunggal

Regresi sederhana, atau disebut juga regresi tunggal, digunakan ketika peneliti ingin mengetahui apakah terdapat hubungan yang bersifat linier antara satu variabel bebas (X) dengan satu variabel terikat (Y). Selain itu, regresi ini juga berguna untuk memperkirakan seberapa besar perubahan atau peningkatan pada variabel terikat (Y) jika nilai variabel bebas (X) diketahui.

Rumus Regresi Sederhana:

$$Y = a + bX$$

Keterangan:

Y = variabel terikat (dependent variable)

X = variabel bebas (independent variable)

a = konstanta (intersep), yaitu nilai Y saat $X = 0$

b = koefisien regresi, menunjukkan besar perubahan Y untuk setiap kenaikan satu satuan pada X

Rumus mencari nilai a :

$$a = \frac{(\sum Y)(\sum X^2) - (\sum X)(\sum XY)}{n(\sum X^2) - (\sum X)^2}$$

Rumus mencari nilai b :

$$b = \frac{n(\sum XY) - (\sum X)(\sum Y)}{n(\sum X^2) - (\sum X)^2}$$

2. Regresi Ganda

Dalam ilmu sosial, terutama di bidang pendidikan, hubungan yang terjadi jarang hanya melibatkan dua variabel. Umumnya, satu variabel terikat berkaitan dengan beberapa variabel bebas sekaligus. Oleh karena itu, dalam analisis statistik sebaiknya digunakan metode yang mampu menangani hubungan antar banyak variabel, yaitu analisis regresi ganda. Regresi ganda digunakan untuk mengetahui apakah terdapat hubungan yang bersifat linier antara dua atau lebih variabel bebas dengan satu variabel terikat. Selain itu, regresi ini juga berguna untuk memprediksi nilai variabel terikat berdasarkan nilai-nilai dari variabel-variabel bebas yang diketahui.

Rumus:

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_nX_n$$

Keterangan:

Y = variabel terikat (dependent variable)

a = konstanta (intersep), nilai Y saat semua $X = 0$

$b_1, b_2, \dots, b_n$ = koefisien regresi masing-masing variabel independen

$X_1, X_2, \dots, X_n$ = variabel-variabel bebas (independent variables)

Sebelum menerapkan analisis regresi ganda, ada beberapa syarat yang perlu dipenuhi agar hasil analisis valid, sebagaimana dijelaskan oleh Irianto (2004):

- a) Sampel harus diambil secara acak dari populasi yang berdistribusi normal.
- b) Sampel juga harus berdistribusi normal. Normalitas bisa dibantu dengan jumlah sampel yang besar atau diuji menggunakan uji Liliefors.

- c) Variabel dependen harus berskala interval atau rasio, sementara variabel independen bisa berskala lebih rendah, asalkan telah dikodekan (coding).
- d) Harus ada hubungan teoritis antara variabel bebas dan terikat, dan hubungan ini signifikan secara statistik berdasarkan uji korelasi sederhana.
- e) Hubungan dalam model regresi harus bersifat linier.

D. Contoh Penggunaan Pengujian Korelasi, Analisis Regresi, dan ANOVA dalam Penelitian PBA

Contoh Penggunaan Analisis Regresi dan ANOVA	
	Idhokus Samsiah. Program Magister Prodi PBA Pascasarjana IAIN Ponorogo, Tahun 2021. Pengaruh Penguasaan Kosakata Bahasa Arab dan <i>Tahfiz</i> Al-Qur'an Terhadap Motivasi Belajar Siswa Masa <i>Home Visit</i> Era Pandemi Covid-19 Tujuan Penelitian: 1. Pengaruh Penguasaan Kosakata Bahasa Arab Terhadap Motivasi Belajar Siswa. 2. Pengaruh <i>Tahfiz</i> Al-Qur'an Terhadap Motivasi Belajar Siswa. 3. Pengaruh Penguasaan Kosakata Bahasa Arab Dan <i>Tahfiz</i> Al-Qur'an Terhadap Motivasi Belajar Siswa Masa <i>Home Visit</i> Era Pandemi Covid 19 Di MI MiftaHul Ulum II Kedungpanji, Lembeyan, Magetan Tahun Pelajaran 2020/2021. Instrumen Penelitian: 1. Penguasaan Kosakata Bahasa Arab 2. Penguasaan <i>Tahfiz</i> Al-Qur'an (Nilai <i>Tahfiz</i>) 3. Motivasi Siswa (Angket) Olah dan Analisis Data : 1. Tes Validitas (Menggunakan Korelasi <i>Product Moment</i>) $R_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{(N \sum X^2 - (\sum X)^2)(N \sum Y^2 - (\sum Y)^2)}}$ 2. Tes Reliabilitas a. Masing-masing Item (Rumus Varians) $\sigma_i^2 = \frac{\sum x_i^2}{N} - \left(\frac{\sum x_i}{N}\right)^2$ b. Keseluruhan Item (<i>Alpha Cronbach</i>) $r_{11} = \left[\frac{k}{k-1} \right] \left[1 - \frac{\sum \sigma_i^2}{\sigma_t^2} \right]$ Keterangan:

r_{11}	= koefisien reliabilitas tes
k	= banyaknya butir item
$\sum \sigma_i^2$	= total jumlah varian
σ_i^2	= jumlah varian skor dari tiap-tiap butir item
1	= bilangan konstanta

3. Menghitung Nilai Rata-rata dan Standar Deviasi

a. Rumus menghitung rata-rata

$M_{x_1} = \frac{\sum f_{x_1}}{n}$	$M_{x_2} = \frac{\sum f_{x_2}}{n}$	$M_{x_3} = \frac{\sum f_{x_3}}{n}$
Keterangan:		
M_{x_1} , M_{x_2} , dan M_{x_3} : rata-rata variabel yang dicari		
$\sum f_{x_1}$, $\sum f_{x_2}$ dan $\sum f_{x_3}$: jumlah perkalian dari masing-masing point interval dengan frekuensi		
n : jumlah data		

b. Rumus menghitung standar deviasi

$SD_{x_1} = i \sqrt{\frac{\sum f_{x_1}^2}{n} + \left(\frac{\sum f_{x_1}}{n}\right)^2}$	$SD_{x_2} = i \sqrt{\frac{\sum f_{x_2}^2}{n} + \left(\frac{\sum f_{x_2}}{n}\right)^2}$
Keterangan:	
SD_{x_1} , SD_{x_2} , SD_{x_3} : standar deviasi	
$\sum f_{x_1}^2$, $\sum f_{x_2}^2$, $\sum f_{x_3}^2$: jumlah dari masing-masing interval dengan X^2 atau Y^2	
n : jumlah data	
i : interval kelas	

4. Uji Asumsi

1) Uji Normalitas

Untuk mengetahui data normal atau tidak normal. Jika data normal maka dapat dilakukan uji parametrik, dan jika tidak normal maka akan menggunakan uji statistik nonparametrik.

2) Uji Linieritas

Uji linieritas digunakan untuk mengetahui apakah hubungan variabel bebas dan terikat bersifat linear. Untuk menguji penelitian ini, peneliti menggunakan SPSS versi 25 dengan Tingkat signifikansi 5% atau 0,05.

3) Uji Multikolinieritas

Uji ini dilakukan untuk mengetahui antara variabel X1 dan variabel X2 ada hubungan yang kuat atau tidak. (Penguji menggunakan SPSS)

4) Uji Heteroskedastisitas

Uji ini dilakukan untuk mengamati apakah dalam regresi terdapat ketidaksamaan varians dari analisis residual (kesalahan) ke analisis residual analisis lainnya.

5. Analisis Regresi (Linear Berganda)

Uji regresi linear berganda digunakan untuk penentuan hipotesis dengan kondisi jumlah variabel bebas lebih dari satu mempengaruhi satu variabel terikat. Dalam penelitian ini X1 adalah penguasaan kosakata bahasa Arab, X2 adalah *tahfiz* al-

Qur'an, dan Y adalah motivasi belajar siswa. Regresi linear dilakukan untuk menguji masing-masing pengaruh variabel X terhadap variabel Y.

Rumus regresi linear:

$$\hat{y} = b_0 + b_1x_1 + b_2x_2$$

$$b_1 = \frac{(\sum_{i=1}^n x_2^2)(\sum_{i=1}^n x_1y) - (\sum_{i=1}^n x_2y)(\sum_{i=1}^n x_1x_2)}{(\sum_{i=1}^n x_1^2)(\sum_{i=1}^n x_2^2) - (\sum_{i=1}^n x_1x_2)^2}$$

$$b_2 = \frac{(\sum_{i=1}^n x_1^2)(\sum_{i=1}^n x_2y) - (\sum_{i=1}^n x_1y)(\sum_{i=1}^n x_1x_2)}{(\sum_{i=1}^n x_1^2)(\sum_{i=1}^n x_2^2) - (\sum_{i=1}^n x_1x_2)^2}$$

$$b_0 = \frac{\sum_{i=1}^n y - b_1 \sum_{i=1}^n x_1 - b_2 \sum_{i=1}^n x_2}{n}$$

6. Analisis Variansi (ANOVA)

Setelah rumus regresi terbentuk, ANOVA digunakan untuk menguji signifikansi model secara keseluruhan. Apakah semua variabel X secara bersama-sama berpengaruh signifikan terhadap Y.

Ho : Penguasaan kosakata bahasa Arab dan *tahfiz* Qur'an tidak berpengaruh secara signifikan terhadap motivasi belajar siswa di MI Miftahul Ulum II Kedungpanji tahun pelajaran 2020/2021.

Ha : Penguasaan kosakata bahasa Arab dan *tahfiz* Qur'an berpengaruh secara signifikan terhadap motivasi belajar siswa di MI Miftahul Ulum II Kedungpanji tahun pelajaran 2020/2021.

Tabel ANOVA

Analysis of Varians			
Sumber Variasi	Degree of Freedom (df)	Sum of Square (SS)	Mean Square (MS)
Regresi	P	$SSR = (b_0 \sum y + b_1 \sum x_1 y + b_2 \sum x_2 y) - \frac{(\sum y)^2}{n}$	$MSR = \frac{SSR}{df}$
Error	n-P-1	$SSE = \sum y^2 - (b_0 \sum y + b_1 \sum x_1 y + b_2 \sum x_2 y)$	$MSR = \frac{SSE}{n-2}$
Total	n-1	$SST = \sum y^2 - \frac{(\sum y)^2}{n}$	

7. Uji t dan Uji F:

1) Uji t:

a. Jika nilai sig < 0,05, atau t hitung > t tabel maka terdapat pengaruh variabel X terhadap variabel Y.

b. Jika nilai sig > 0,05, atau t hitung < t tabel maka tidak terdapat pengaruh variabel X terhadap variabel Y.

2) Uji F:

	a. Jika nilai sig < 0,05, atau F hitung > F tabel maka terdapat pengaruh variabel X secara signifikan terhadap variabel Y. b. Jika nilai sig > 0,05, atau F hitung < F tabel maka tidak terdapat pengaruh variabel X secara signifikan terhadap variabel Y. Kesimpulan dari penelitian di atas didapati: Penguasaan kosakata bahasa Arab dan tahfiz al-Qur'an (X1,X2) berpengaruh terhadap motivasi belajar siswa (Y) sebesar 50,5%.
	Contoh Analisis Korelasi العلاقة بين خلفية التربية لدى طلاب قسم تعليم اللغة العربية وقدرتهم الإنشائية بجامعة شريف هداية الله الإسلامية الحكومية جاكارتا Ummu Salamah. PBA UIN Jakarta, 2023. Hubungan antara Latar Belakang Pendidikan Mahasiswa PBA dan Kemampuan Insya di UIN Syarif Hidayatullah Jakarta. Tujuan Penelitian: 1. Mengukur Latar Belakang Mahasiswa 2. Mengukur Perbedaan Hasil Insya Mahasiswa yang Berasal dari Pesantren dan dari Non Pesantren 3. Mengukur Hubungan antara Latar Belakang Pendidikan Mahasiswa dengan Kemampuan Insya. Instrumen Penelitian 1. Angket 2. Wawancara 3. Dokumen Olah dan Analisis Data 1. Uji Validitas Uji ini dilakukan untuk memeriksa keabsahan masing-masing butir soal. Peneliti menggunakan <i>pearson product moment</i>. Rumus: $r = \frac{n(\sum XY) - (\sum X)(\sum Y)}{\sqrt{[n \sum X^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}}$ 2. Uji Reliabilitas 3. Analisis Korelasi Setelah mendapat nilai X (didapatkan dari nilai angket berkaitan dengan latar belakang pendidikan) dan nilai Y (Nilai insya). Peneliti membuat tabel kerja sebagai berikut:

Y ²	X ²	XY	Y	X	الرقم
3600	2401	2940	60	49	1
5625	2209	3020	75	47	2
4225	1764	2730	65	42	3
3600	1444	2280	60	38	4
5281	2000	3900	73	50	5

Penghitungan dapat dilakukan dengan manual menggunakan rumus atau memanfaatkan aplikasi SPSS.

Rumus Korelasi Product Moment:

$$r = \frac{n(\sum XY) - (\sum X)(\sum Y)}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}}$$

Keterangan:

- r = koefisien korelasi
- n = jumlah pasangan data
- X = skor variabel X
- Y = skor variabel Y
- $\sum XY$ = jumlah hasil perkalian X dan Y
- $\sum X$ = jumlah total nilai X
- $\sum Y$ = jumlah total nilai Y
- $\sum X^2$ = jumlah kuadrat nilai X
- $\sum Y^2$ = jumlah kuadrat nilai Y

Setelah itu dilakukan uji interpretasi:

- Jika r hitung $\geq r$ tabel, maka hubungan antar variabel signifikan
- Jika r hitung $< r$ tabel, maka tidak signifikan

Kesimpulan dari penelitian di atas:

Dari hasil perhitungan didapati terdapat hubungan yang **sedang** antara latar belakang pendidikan dengan kemampuan insya.

KESIMPULAN

Berdasarkan pembahasan yang telah diuraikan, dapat disimpulkan bahwa analisis korelasi, regresi, dan ANOVA merupakan tiga metode statistik yang memiliki fungsi penting dalam penelitian kuantitatif. Analisis korelasi digunakan untuk mengetahui hubungan antar dua variabel, analisis regresi bertujuan untuk memprediksi nilai variabel dependen berdasarkan variabel independen, sedangkan ANOVA digunakan untuk membandingkan rata-rata dari tiga kelompok atau lebih guna melihat adanya perbedaan yang signifikan.

Untuk peneliti, mahasiswa, atau siapa pun yang menggunakan metode kuantitatif, disarankan untuk mempelajari secara mendalam teori serta aplikasi praktis dari analisis korelasi, regresi, dan ANOVA. Penggunaan software statistik seperti SPSS, R, atau Python juga sangat membantu dalam mempercepat proses analisis dan meminimalisasi kesalahan perhitungan. Selain itu, penting untuk memahami asumsi-asumsi dasar dari setiap metode agar hasil analisis lebih valid dan dapat dipertanggungjawabkan secara ilmiah.

DAFTAR PUSTAKA

- Ananda, R., & Fadhli, M. (2018). *Statistik pendidikan: Teori dan praktik dalam penelitian*. Medan: CV Widya Puspita.
- Nuryadi, S. P., Astuti, T. D., Sri Utami, E., & Budiantara, M. (2017). *Dasar-dasar statistik penelitian*. Yogyakarta: Sibuku Media.
- Samsiah, I. (2021). *Pengaruh penguasaan kosakata bahasa Arab dan tahfiz Al-Qur'an terhadap motivasi belajar siswa masa home visit era pandemi Covid-19* (Tesis Magister, Program Studi Pendidikan Bahasa Arab, Pascasarjana IAIN Ponorogo).
- سلمة، أم. العلاقة بين خلفية التربية لدى طلاب قسم تعليم اللغة العربية وقدرتهم الإنشائية بجامعة شريف هداية الله الإسلامية الحكومية جاكارتا (2023). بحث في طلاب قسم تعليم اللغة العربية جامعة شريف هداية الله الإسلامية الحكومية جاكارتا.