

## Menganalisis Hubungan Antara Variabel X dan Variabel Y Menggunakan Simple Linear

Silvi Nurinsan

Sains Data Fakultas Sains & Teknologi Universitas Koperasi Indonesia

Email : [silvinurinsan7@gmail.com](mailto:silvinurinsan7@gmail.com)

### ABSTRAK

*Dataset ini berisi informasi mengenai data pelatihan mahasiswa di sebuah institusi pendidikan. Data mencakup berbagai atribut seperti ID mahasiswa, nama, umur, gender, nilai akhir, mata kuliah, tanggal pelaksanaan, serta nilai UTS dan UAS. Dataset ini dapat digunakan untuk menganalisis performa akademik berdasarkan berbagai faktor seperti gender, mata kuliah, dan nilai ujian. Informasi yang terkandung dalam dataset ini dapat membantu dalam pengambilan keputusan terkait peningkatan kurikulum, pembelajaran, dan evaluasi mahasiswa.*

**Kata kunci :** Data Mahasiswa, Performa Akademik, Gender, Nilai Akhir, Mata Kuliah, Ujian Tengah Semester (UTS), Ujian Akhir Semester (UAS), Analisis Pendidikan

### 1. PENDAHULUAN

Analisis data merupakan salah satu tahapan penting dalam proses pengambilan keputusan, terutama dalam konteks pendidikan. Dataset yang kita miliki berisi informasi terkait dengan performa mahasiswa dalam berbagai mata kuliah. Data ini mencakup beberapa atribut penting seperti ID mahasiswa, nama, umur, gender, nilai total, nama mata kuliah, tanggal ujian, serta nilai Ujian Tengah Semester (UTS) dan Ujian Akhir Semester (UAS).

Tujuan utama dari analisis ini adalah untuk memahami karakteristik dan tren performa akademik mahasiswa berdasarkan berbagai faktor seperti gender, umur, dan mata kuliah yang diambil. Dengan memahami pola-pola ini, institusi pendidikan dapat merancang strategi yang lebih efektif untuk meningkatkan hasil belajar dan memberikan dukungan yang lebih baik kepada mahasiswa.

## 2. TINJAUAN PUSTAKA

### 2.1. Pengantar

Analisis regresi linier sederhana digunakan untuk memprediksi nilai dari variabel dependen (Y) berdasarkan satu variabel independen (X). Dalam konteks dataset yang diberikan, tujuan kita adalah memprediksi nilai UAS (Ujian Akhir Semester) berdasarkan nilai UTS (Ujian Tengah Semester).

### 2.2. Deskripsi Data

Dataset yang dianalisis terdiri dari beberapa kolom yang mencakup informasi tentang mahasiswa, termasuk ID, Nama, Umur, Gender, Nilai, Mata Kuliah, Tanggal, UTS, dan UAS. Berikut adalah penjelasan singkat mengenai kolom-kolom tersebut:

- ID: Identifikasi unik untuk setiap mahasiswa.
- Nama: Nama mahasiswa.
- Umur: Usia mahasiswa.
- Gender: Jenis kelamin mahasiswa.
- Nilai: Nilai rata-rata dari mahasiswa.
- Matkul: Mata kuliah yang diambil.
- Tanggal: Tanggal ujian.
- UTS: Nilai Ujian Tengah Semester.
- UAS: Nilai Ujian Akhir Semester.

### 2.3. Analisis Regresi Linier Sederhana

Regresi linier sederhana digunakan untuk memahami hubungan antara dua variabel, dalam hal ini antara UTS (X) dan UAS (Y). Persamaan umum regresi linier sederhana adalah:

$$Y = \beta_0 + \beta_1 X + \epsilon$$

Di mana:

- Y : adalah variabel dependen (UAS).
- X : adalah variabel independen (UTS).
- $\beta_0$  : adalah intercept (konstanta).
- $\beta_1$  : adalah slope (kemiringan garis regresi).
- $\epsilon$  : adalah error term (galat).

### 3. HASIL PENELITIAN

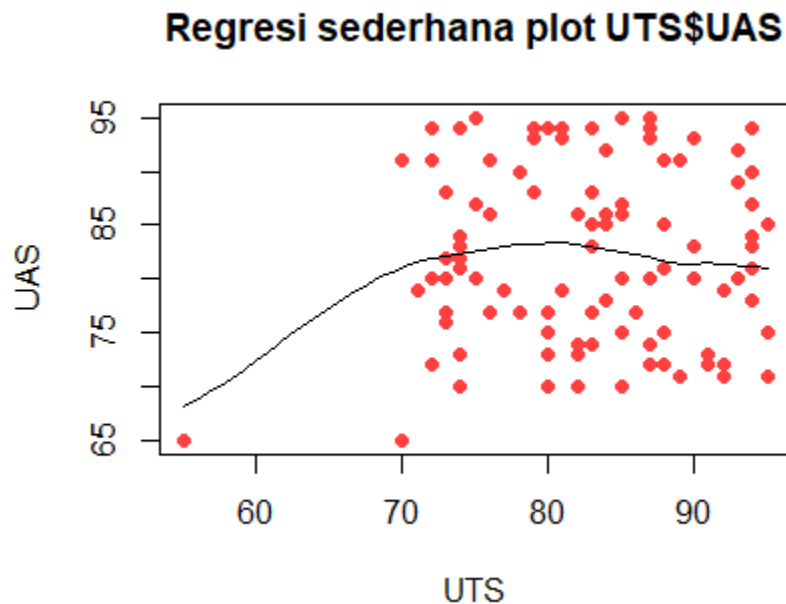
Persamaan regresi linier sederhana yang kita dapatkan untuk memprediksi nilai UAS berdasarkan nilai UTS dalam dataset Anda adalah:

$$\text{UAS} = 78.38 + 0.046 \cdot \text{UTS}$$

Hasil persamaan Regresi linier Sederhana :

```
Coefficients:  
(Intercept)          x  
78.38268          0.04623
```

Hasil Plot dari persamaan Regresi Sederhana UTS & UAS



Mari kita jelaskan komponen-komponen dari persamaan ini dan apa arti dari masing-masing koefisien.

### 3.1. Komponen-komponen Persamaan

Intercept ( $\beta_0$ ) = 78.38:

- Intercept adalah nilai konstanta dalam persamaan regresi. Ini adalah nilai prediksi dari UAS ketika nilai UTS adalah 0.
- Dalam konteks ini, jika seorang mahasiswa memiliki nilai UTS sebesar 0, maka nilai UAS mereka diprediksi akan berada di sekitar 78.38. Namun, nilai UTS sebesar 0 mungkin tidak realistis dalam skenario nyata, tetapi intercept tetap memberikan titik awal untuk garis regresi.

Slope ( $\beta_1$ ) = 0.046:

- Slope atau kemiringan adalah koefisien yang mengalikan variabel independen (UTS) dalam persamaan regresi. Ini menunjukkan perubahan rata-rata dalam nilai UAS untuk setiap satu unit perubahan dalam nilai UTS.
- Dalam hal ini, setiap peningkatan satu unit dalam nilai UTS diprediksi akan menghasilkan peningkatan sebesar 0.046 dalam nilai UAS.
- Misalnya, jika nilai UTS meningkat dari 70 menjadi 71, maka nilai UAS diprediksi akan meningkat sebesar 0.046.

### 3.1 Interpretasi Persamaan

Fungsi Prediktif:

- Dengan menggunakan persamaan ini, kita dapat memprediksi nilai UAS untuk setiap nilai UTS yang diberikan. Misalnya, jika seorang mahasiswa memiliki nilai UTS sebesar 80, maka nilai UAS yang diprediksi adalah:

$$\text{UAS} = 78.38 + 0.046 \times 80 = 78.38 + 3.68 = 82.06$$

Hubungan Linear:

- Persamaan ini mengasumsikan bahwa ada hubungan linear antara UTS dan UAS. Artinya, kenaikan dalam nilai UTS secara konsisten diikuti oleh kenaikan dalam nilai UAS sesuai dengan kemiringan (0.046) yang telah dihitung.

Model Sederhana:

- Model regresi linier sederhana hanya menggunakan satu variabel independen (UTS) untuk memprediksi variabel dependen (UAS). Ini memberikan gambaran

dasar tentang hubungan antara kedua variabel, tetapi mungkin tidak menangkap semua faktor yang mempengaruhi nilai UAS.

### 3.2 Validasi Model

R-squared ( $R^2$ ):

- R-squared adalah statistik yang memberikan informasi tentang seberapa baik variabel independen menjelaskan variabilitas dalam variabel dependen. Nilai R-squared berkisar antara 0 dan 1.
- Nilai R-squared yang lebih tinggi menunjukkan bahwa model lebih baik dalam menjelaskan variasi nilai UAS berdasarkan nilai UTS.

Signifikansi Statistik:

- Selain melihat nilai R-squared, penting juga untuk memeriksa signifikansi statistik dari koefisien regresi (baik intercept maupun slope) untuk memastikan bahwa hubungan yang diidentifikasi dalam model bukanlah hasil dari kebetulan semata.

## 4. KESIMPULAN

Model regresi linier sederhana yang digunakan untuk memprediksi nilai UAS berdasarkan nilai UTS tidak menunjukkan hubungan yang signifikan antara kedua variabel. Koefisien slope (0.046) tidak signifikan secara statistik, dan nilai R-squared yang sangat rendah menunjukkan bahwa variabel UTS hampir tidak memberikan informasi yang berguna untuk memprediksi nilai UAS. Untuk meningkatkan model prediksi nilai UAS, mungkin perlu mempertimbangkan variabel independen lainnya yang lebih relevan atau menggunakan model yang lebih kompleks. Analisis lebih lanjut juga dapat dilakukan untuk mengeksplorasi faktor-faktor lain yang mungkin mempengaruhi nilai UAS.

## 5. DAFTAR PUSTAKA

**Kutner, M. H., Nachtsheim, C. J., & Neter, J. (2004).** *Applied Linear Regression Models* (4th ed.). McGraw-Hill/Irwin.

**Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012).** *Introduction to Linear Regression Analysis* (5th ed.). Wiley.

**Frost, J. (2019).** *Regression Analysis: An Intuitive Guide for Using and Interpreting Linear Models*. Statistics by Jim Publishing.

**Field, A. (2013).** *Discovering Statistics Using IBM SPSS Statistics (4th ed.)*. Sage Publications Ltd.

**Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003).** *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences (3rd ed.)*. Routledge.

**Gujarati, D. N., & Porter, D. C. (2009).** *Basic Econometrics (5th ed.)*. McGraw-Hill/Irwin.

**Wooldridge, J. M. (2016).** *Introductory Econometrics: A Modern Approach (6th ed.)*. Cengage Learning.

**Chatterjee, S., & Hadi, A. S. (2012).** *Regression Analysis by Example (5th ed.)*. Wiley.