

# **“Implementasi Fungsi Menghitung Koefisien Korelasi Spearman antara Dua Set Data”**

<sup>1st</sup> Elia Meylani Simanjuntak  
122450026

<sup>2nd</sup> Priska Silvia Ferantiana  
122450053

<sup>3rd</sup> Residen Nusantara  
122450080

<sup>4th</sup> Anwar Muslim  
122450117

<sup>5th</sup> Muhammad Zaky Zaiddan  
122450119

*Jurusan Sains Data, Fakultas Sains, Institut Teknologi Sumatera  
Jl. Terusan Ryacudu, Way Huwi, Kec. Jati Agung, Kabupaten Lampung Selatan, Lampung 35365  
Penulis Korespondensi : [elia.122450026@student.itera.ac.id](mailto:elia.122450026@student.itera.ac.id),  
[priska.122450053@student.itera.ac.id](mailto:priska.122450053@student.itera.ac.id), [residen.122450080@student.itera.ac.id](mailto:residen.122450080@student.itera.ac.id),  
[anwar.122450117@student.itera.ac.id](mailto:anwar.122450117@student.itera.ac.id), [muhammad.122450119@student.itera.ac.id](mailto:muhammad.122450119@student.itera.ac.id)*

## **Abstrak**

Analisis data menjadi salah langkah yang harus dilakukan sebelum mengambil keputusan di berbagai bidang. Untuk menganalisis data yang semakin banyak tiap harinya, diperlukan adanya teknologi yang semakin canggih dan metode yang efektif. Salah satu metode yang dapat dilakukan untuk menganalisis data yaitu menggunakan Koefisien Korelasi Spearman. Koefisien korelasi Spearman merupakan ukuran kuatnya erat kaitan antara dua variabel ordinal atau ukuran derajat hubungan antar data yang di rangking menurut rangkingnya. Dataset yang kita pakai berisi informasi tentang berat badan dan tinggi badan sejumlah orang atau subjek yang akan digunakan untuk menghitung koefisien korelasi antara berat badan dan tinggi badan.

**Kata Kunci:** Analisis Data, Dataset, Korelasi Spearman.

## **1. Pendahuluan**

### **1.1 Latar Belakang**

Dalam era teknologi yang semakin canggih, analisis data menjadi salah langkah yang harus dilakukan sebelum mengambil keputusan di berbagai bidang. Untuk menganalisis data yang semakin banyak tiap harinya, diperlukan adanya teknologi yang semakin canggih dan metode yang efektif. Salah satu metode yang dapat dilakukan untuk menganalisis data yaitu menggunakan Koefisien Korelasi Spearman.

Koefisien Korelasi Spearman dapat digunakan ketika data tidak memenuhi asumsi distribusi normal. Pada umumnya, metode ini adalah salah satu metode statistika non-parametrik yang digunakan untuk mengukur derajat hubungan antara dua variabel (Daniel, 1989).

Dalam analisis kali ini, kita menggunakan dataset. pengertian dari dataset adalah kumpulan data yang terstruktur yang dapat digunakan untuk analisis, penelitian, atau pengembangan model. Dataset berisi informasi tentang berbagai entitas, seperti individu, percobaan, atau kejadian, yang diwakili dalam baris dan kolom untuk atribut atau variabel yang diamati. Dataset dapat berbentuk tabel, array, atau format lainnya.

#### 1.2 Rumusan Masalah

- a) Bagaimana penerapan metode koefisien korelasi Spearman dalam analisis data?
- b) Apa keunggulan metode koefisien korelasi Spearman dibandingkan metode lainnya dalam analisis data?

#### 1.3 Tujuan

- a) Menerapkan metode koefisien korelasi Spearman pada dua data set.
- b) Menganalisis kinerja metode koefisien korelasi Spearman untuk memudahkan analisis data.

## 2. Metode

Untuk mengukur korelasi pada skala ordinal atau statistik nonparametrik, digunakan korelasi Spearman (Bin, 2019). Korelasi spearman ini adalah ukuran korelasi yang menghubungkan kedua variabel, yang diukur sekurang-kurangnya dalam skala ordinal, sehingga subjek penelitian dapat diurutkan dalam dua rangkaian berurutan (Andreas, 2020). Koefisien korelasi Spearman merupakan ukuran kuatnya erat kaitan antara dua variabel ordinal atau ukuran derajat hubungan antar data yang di rangking menurut rangkingnya. Koefisien korelasi digunakan untuk mengukur seberapa erat keterkaitan suatu variabel dengan variabel lainnya, dimana pengamatan terhadap setiap variabel didasarkan pada pemberian rangking tertentu berdasarkan pengamatan dan pembandingnya.

## 3. Pembahasan

### 3.1 Koefisien Korelasi Spearman

Koefisien korelasi peringkat Spearman dikembangkan oleh Charles Spearman (1908). Koefisien korelasi peringkat Spearman-rho digunakan untuk mengukur korelasi pada statistik nonparametrik (skala ordinal), seperti yang dilakukan dengan koefisien korelasi Pearson (r). Ini adalah ukuran korelasi yang menuntut bahwa kedua variabel diukur dalam skala ordinal setidaknya. Dengan demikian, subjek penelitian dapat diurutkan dalam dua rangkaian berurutan (Daniel, 1989). Koefisien korelasi Spearman dapat dihitung dengan menggunakan rumus:

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

dimana:

$r_s$  = koefisien korelasi spearman

$\sum d^2$  = jumlah dari selisih peringkat kuadrat antara kedua variabel

n = jumlah data

Peringkat atau masing-masing variabel akan dihitung menggunakan dataset yang disediakan. Selanjutnya, persamaan di atas akan digunakan untuk menghitung koefisien korelasi spearman. Ini adalah metode analisis statistik yang umum untuk mengidentifikasi hubungan antara dua variabel dalam penelitian. Penelitian tersebut dapat menghasilkan informasi yang berguna tentang korelasi antara variabel-variabel yang dipelajari dengan menggunakan dataset dan persamaan matematika yang sesuai. Landasan teoritis dan konteks penelitian yang lebih mendalam terkait dengan penerapan fungsi dan analisis data akan diberikan melalui referensi artikel ilmiah tertentu (Djarwanto, 1997).

Untuk menghitung koefisien korelasi Spearman-rho ( $\rho$ ), lakukan hal-hal berikut:

- Beri peringkat untuk masing-masing pengamatan X dari 1 hingga n, dan untuk pengamatan Y dari 1 hingga n.
- Tentukan harga  $\sum_{i=1}^n d_i^2$ , yaitu jumlah kuadrat dari selisih-selisih antara rank-rank  $X_i$  dan  $Y_i$  untuk masing-masing pengamatan.
- Gunakan persamaan [6] untuk menghitung  $\rho$

### 3.2 Dataset

Dataset yang kita pakai berisi informasi tentang berat badan dan tinggi badan sejumlah orang atau subjek yang akan digunakan untuk menghitung koefisien korelasi antara berat badan dan tinggi badan. Data ini biasanya digunakan dalam penelitian di berbagai bidang seperti kedokteran, antropometri, nutrisi, olahraga, dan psikologi, di mana hubungan antara berat badan dan tinggi badan sering menjadi subjek. Kita mengambil dataset ini karena dataset ini yang paling mudah ditemukan dan kita mengambil dataset ini dari Kaggle. Ada 2 variabel yang kita pakai pada analisis kali ini, yaitu:

- a) Berat Badan: Variabel ini mengacu pada massa tubuh seseorang, yang diukur dalam satuan metrik seperti kilogram (kg) atau pound (lbs). Berat badan adalah ukuran objektif dari jumlah materi yang ada dalam tubuh seseorang. Ini bisa dipengaruhi oleh banyak hal, seperti genetika, pola makan, tingkat aktivitas fisik, dan kondisi tempat tinggal.
- b) Tinggi Badan: Variabel ini mengacu pada jarak vertikal yang diukur dalam sentimeter (cm) atau inci (in). Tinggi badan seseorang adalah ukuran umum dari dimensi fisiknya. Faktor genetik dan faktor lingkungan seperti gizi, kesehatan, dan kondisi pertumbuhan memengaruhi hal ini.

Koefisien korelasi Spearman untuk berat badan dan tinggi badan menunjukkan seberapa erat atau kuat hubungan monotonik antara keduanya dalam populasi. Dengan kata lain, jika koefisien korelasi Spearman positif dan signifikan, itu menunjukkan bahwa dalam populasi tersebut, kecenderungan untuk meningkatnya berat badan juga diikuti oleh kecenderungan untuk meningkatnya tinggi badan, dan sebaliknya. Ini menunjukkan bahwa ada hubungan sistematis antara berat badan dan tinggi badan, tetapi tidak harus linier. Koefisien korelasi Spearman dapat menunjukkan arah dan kekuatan hubungan antara kedua variabel ini tanpa mengganggu hubungan tersebut berbentuk linier.

### 3.3 Penguraian Kode

```
[1] def spearman_correlation(x, y):  
    n = len(x)  
    rank_x = {val: rank for rank, val in enumerate(sorted(set(x)), 1)}  
    rank_y = {val: rank for rank, val in enumerate(sorted(set(y)), 1)}  
    d = sum((rank_x[xi]-rank_y[yi]) ** 2 for xi, yi in zip(x, y))  
    return 1- (6*d) / (n*(n**2-1))
```

**Gambar 1.** Pendefinisian Fungsi

```
berat_badan = [87, 60, 49, 60, 55, 59, 58, 62, 67, 69]  
tinggi_badan = [168, 173, 150, 168, 157, 167, 159, 167, 166, 176]  
  
spearman_coefficient = spearman_correlation(berat_badan, tinggi_badan)  
print("Koefisien Korelasi Spearman:", spearman_coefficient)  
  
Koefisien Korelasi Spearman: 0.8484848484848485
```

**Gambar 2.** Inisiasi dua *variable*

Penjelasan tentang kode ini adalah implementasi dari perhitungan koefisien korelasi Spearman antara dua set data:

- Fungsi `spearman_correlation(x, y)` mengambil dua set data, `x` dan `y`.
- Peringkat Variabel: Dalam set data, setiap nilai diurutkan dan diberi peringkat relative berdasarkan urutan nilai tersebut.
- Perhitungan Selisih Peringkat: Ini adalah cara untuk menghitung perbedaan peringkat antara nilai kedua variabel yang sesuai.
- Perhitungan Koefisien Korelasi: Selisih peringkat dihitung dengan menggunakan rumus koefisien korelasi Spearman.
- Output: Nilai koefisien korelasi Spearman antara dua set data dikembalikan oleh fungsi.

Secara keseluruhan, fungsi tersebut memungkinkan penggunaan koefisien korelasi Spearman untuk mengetahui seberapa erat hubungan monotonik antara dua set data.

## 4. Kesimpulan

Berdasarkan hasil pembahasan di atas dapat ditarik beberapa kesimpulan, yaitu:

- Penerapan metode koefisien korelasi Spearman pada dua data set membantu menentukan keterkaitan suatu variabel dengan variabel lainnya.
- Koefisien korelasi Spearman membantu mengidentifikasi adanya hubungan monotonik antara dua variabel atau data set.

- c) Nilai koefisien korelasi pada kasus ini adalah sebesar 0.8484848484848485, dimana nilai tersebut diketahui mendekati 1. Hal ini menunjukkan bahwa terdapat hubungan monotonik positif antara kedua variabel dari data set tersebut.
- d) Besarnya nilai koefisien korelasi Spearman ini menunjukkan kekuatan hubungan antara kedua data set.
- e) Nilai koefisiensi korelasi Spearman berdasarkan variabel kedua data set ini sebesar 0.8484848484848485. Nilai tersebut dikatakan mendekati 1 yang berarti hubungan antara kedua variabel pada data set tersebut menunjukkan hubungan yang kuat.
- f) Koefisien korelasi Spearman ini membantu dalam menentukan jenis hubungan antara kedua variabel atau data set.
- g) Hubungan antara kedua data set yang telah dianalisis menunjukkan hubungan yang positif, artinya jika nilai suatu variabel meningkat, maka nilai variabel lainnya juga akan cenderung meningkat.

Link github : <https://github.com/sains-data/Praktikum-SD2201-Pemrograman-Berbasis-Fungsi>

Link Collab : [https://colab.research.google.com/drive/1tunUuBEMt\\_mF2o-cm81Qk9xWA8-dcWhj?usp=sharing](https://colab.research.google.com/drive/1tunUuBEMt_mF2o-cm81Qk9xWA8-dcWhj?usp=sharing)

## ***References***

- Andreas, H. &. (2020). Spearman rank correlation of the bivariate Student and scale mixtures of normal distributions. *Journal of Multivariate Analysis*, 1-11.
- Bin, W. R. (2019). Compatible matrices of spearman's rank correlation. *Statistics and Probability Letters*, 67-72.
- Daniel, W. (1989). *Statistika Nonparametrik Terapan*. Jakarta: PT. Gramedia.
- Djarwanto. (1997). *Statistika Nonparametrik*. Yogyakarta: BPFE.