

**ANALISIS KORELASI ANTARA KONDISI SOSIAL
EKONOMI MAHASISWA DAN PRESTASI AKADEMIK:
PENDEKATAN BIG DATA DI PERGURUAN TINGGI**



LAPORAN MBKM PENELITIAN

Ferdyan Hidayat

00000072575

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

**ANALISIS KORELASI ANTARA KONDISI SOSIAL
EKONOMI MAHASISWA DAN PRESTASI AKADEMIK:
PENDEKATAN BIG DATA DI PERGURUAN TINGGI**



LAPORAN MBKM PENELITIAN

Diajukan Sebagai Salah Satu Syarat untuk Memperoleh

Gelar Sarjana Komputer (S.Kom)

Ferdyan Hidayat

00000072575

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2025

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Ferdyan Hidayat

Nomor Induk Mahasiswa : 00000072575

Program studi : Sistem Informasi

Laporan MBKM Penelitian dengan judul:

ANALISIS KORELASI ANTARA KONDISI SOSIAL EKONOMI MAHASISWA DAN PRESTASI AKADEMIK: PENDEKATAN BIG DATA DI PERGURUAN TINGGI

merupakan hasil karya saya sendiri bukan plagiat dari karya ilmiah yang ditulis oleh orang lain, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/ penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan MBKM, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk laporan MBKM yang telah saya tempuh.

Tangerang, 17 July 2025



Ferdyan Hidayat

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Multimedia Nusantara, saya yang bertandatangan di bawah ini:

Nama : Ferdyan Hidayat
NIM : 00000072575
Program Studi : Sistem Informasi
Fakultas : Teknik dan Informasi
JenisKarya : Laporan MBKM Penelitian

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Multimedia Nusantara Hak Bebas Royalti Noneksklusif (Non-exclusive Royalty-Free Right) atas karya ilmiah saya yang berjudul:

ANALISIS KORELASI ANTARA KONDISI SOSIAL EKONOMI MAHASISWA DAN PRESTASI AKADEMIK: PENDEKATAN BIG DATA DI PERGURUAN TINGGI

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non eksklusif ini Universitas Multimedia Nusantara berhak menyimpan, mengalih media / format-kan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis / pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Tangerang, 17 July 2025

Yang menyatakan,



Ferdyan Hidayat

KATA PENGANTAR

Puji Syukur atas selesainya penulisan penelitian ini dengan judul: “Studi Korelasi Antara Kondisi Sosial Ekonomi Mahasiswa dan Prestasi Akademik: Pendekatan Big Data di Perguruan Tinggi” dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Jurusan Sistem Informasi Pada Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan tugas akhir ini, sangatlah sulit bagi saya untuk menyelesaikan tugas akhir ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Dr. Ir. Andrey Andoko, M.Sc., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Ibu Ririn Ikana Desanti, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi Universitas Multimedia Nusantara.
4. Ibu Ririn Ikana Desanti, S.Kom., M.Kom., sebagai Pembimbing yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya tesis ini.
5. Bapak Iwan Prasetiawan, S.Kom., M.M., sebagai Pembimbing Lapangan yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya laporan MBKM Penelitian.
6. Keluarga yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan laporan MBKM ini.

Karya ini diharapkan dapat berkontribusi dalam pengembangan studi pendidikan berbasis data, serta menjadi referensi untuk penelitian dan kebijakan pendidikan di masa mendatang.

Tangerang, 17 July 2025



Ferdyan Hidayat

ANALISIS KORELASI ANTARA KONDISI SOSIAL EKONOMI MAHASISWA DAN PRESTASI AKADEMIK: PENDEKATAN BIG DATA DI PERGURUAN TINGGI

Ferdyan Hidayat

ABSTRAK

Penelitian ini dilatarbelakangi oleh pentingnya memahami pengaruh kondisi sosial ekonomi terhadap prestasi akademik mahasiswa aktif Universitas Multimedia Nusantara guna mendukung kebijakan pendidikan yang lebih inklusif dan adil. Tujuan penelitian ini adalah untuk menganalisis hubungan antara latar belakang sosial ekonomi mahasiswa seperti pendidikan orang tua, pekerjaan, dan pendapatan keluarga dengan capaian akademik yang diukur melalui Indeks Prestasi Kumulatif (IPK). Penelitian ini menggunakan pendekatan kuantitatif dengan metode korelasi Spearman dan kerangka kerja CRISP-DM. Data diperoleh dari survei sosial ekonomi terhadap 65 mahasiswa aktif dan transkrip akademik 961 mahasiswa angkatan 2010–2023 Universitas Multimedia Nusantara. Proses analisis dilakukan menggunakan bahasa pemrograman Python melalui tahapan data preparation, modeling, dan evaluasi. Hasil menunjukkan tidak terdapat korelasi signifikan antara faktor sosial ekonomi dan IPK mahasiswa. Variabel penanggung biaya memiliki korelasi tertinggi ($p = 0.211$), namun tetap tidak signifikan secara statistik ($p > 0.05$). Temuan ini menekankan bahwa faktor internal seperti motivasi dan strategi belajar kemungkinan lebih berperan dalam menentukan prestasi akademik mahasiswa. Hasil analisis diterapkan dalam bentuk aplikasi web interaktif menggunakan Streamlit agar dapat diakses dan dimanfaatkan secara lebih luas oleh pemangku kepentingan.

Kata kunci: Big Data, CRISP-DM, Kondisi Sosial Ekonomi, Korelasi Spearman, Prestasi Akademik.

***CORRELATION ANALYSIS BETWEEN STUDENTS
SOCIOECONOMIC CONDITIONS AND ACADEMIC
PERFORMANCE: A BIG DATA APPROACH IN HIGHER
EDUCATION***

Ferdyan Hidayat

ABSTRACT (English)

This study is motivated by the importance of understanding the influence of socioeconomic conditions on the academic performance of active students at Universitas Multimedia Nusantara, in order to support the development of more inclusive and equitable educational policies. The primary objective of this research is to analyze the relationship between students' socioeconomic backgrounds such as parental education, occupation, and household income and academic achievement, measured by Grade Point Average (GPA). A quantitative approach was employed using the Spearman correlation method and the CRISP-DM framework. Data were obtained from a socioeconomic survey involving 65 active students and academic transcripts of 961 students from the 2010–2023 cohorts at Universitas Multimedia Nusantara. The analysis process was conducted using the Python programming language, covering the stages of data preparation, modeling, and evaluation. The results indicate no significant correlation between socioeconomic factors and students' GPA. The tuition payer variable showed the highest correlation ($\rho = 0.211$), yet remained statistically insignificant ($p > 0.05$). These findings highlight that internal factors such as motivation and learning strategies are likely to play a more substantial role in determining students' academic performance. The analysis results are implemented in an interactive web application using Streamlit to broaden accessibility for stakeholders.

Keywords: *Academic Achievement, Big Data, CRISP-DM, Socioeconomic Condition, Spearman correlation*

DAFTAR ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT	ii
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS.....	iii
KATA PENGANTAR.....	iv
ABSTRAK	v
<i>ABSTRACT (English)</i>	vi
DAFTAR ISI.....	vii
DAFTAR TABEL	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Tujuan Penelitian	4
1.4 Urgensi Penelitian	4
1.5 Luaran Penelitian.....	5
1.6 Manfaat Penelitian	5
BAB II TINJAUAN PUSTAKA.....	6
2.1 Teori Tentang Topik.....	6
2.1.1 Pendidikan	6
2.1.2 Sosial Ekonomi	7
2.1.3 Korelasi	7
2.1.4 Sumber Data	8
2.2 Teori Tentang Framework dan Metode.....	9
2.2.1 CRISP-DM	9
2.2.2 Spearman Corellation	11
2.3 Teori Tentang Tools dan Tenik	12
2.3.1 Apps Script.....	12
2.3.2 Python	12
2.3.3 Jupyter	13

2.3.4 Github	14
2.3.4 Streamlit	14
2.4 Penelitian Terdahulu	15
BAB III METODE PENELITIAN	18
3.1 Objek Penelitian	18
3.2 Metode Penelitian	18
3.2.1 Business Understanding	21
3.2.2 Data Understanding	21
3.2.3 Data Preparation	22
3.2.4 Modeling	23
3.2.5 Evaluation	23
3.2.6 Deployment	24
3.3 Teknik Pengumpulan Data	25
3.4 Variable Penelitian	25
3.4.1 Variable Dependent	25
3.4.2 Variable Independent	26
3.5 Teknik Analisis Data	26
BAB IV HASIL DAN PEMBAHASAN	28
4.1 Business Understanding	28
4.2 Data Understanding	29
4.2.1 Data Survey Sosial Ekonomi Mahasiswa	29
4.2.2 Data Transkrip Nilai Akademik Mahasiswa	31
4.2.3 Visualisasi Data	31
4.3 Data Preparation	36
4.3.1 Data Preparation Data Survey	36
4.3.1.1 Handling Missing Value Data Survey	36
4.3.1.2 Encoding Data Survey	37
4.3.1.3 Handling Outlier Data Survey	38
4.3.2 Data Preparation Data Transkrip Nilai	39
4.3.2.1 Handling Struktur Data Transkrip Nilai	39
4.3.2.2 Menambahkan Kolom Baru IPK Data Transkrip Nilai	40
4.3.2.3 Handling Missing Data Transkrip Nilai	45

4.3.2.5	Encoding Data Transkrip Nilai	46
4.3.2.5	Handling Outlier Data Transkrip Nilai	47
4.4	Modeling	48
4.4.1	Uji Korelasi Data Survey	48
4.4.2	Uji Korelasi Data Transkrip Nilai Akademik Mahasiswa	50
4.5	Evaluation	51
4.6	Deployment	54
BAB V	SIMPULAN SARAN	59
5.1	Simpulan	59
5.2	Saran	60
DAFTAR PUSTAKA		61
LAMPIRAN		66

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu	15
Tabel 3. 1 Perbandingan Framework Data Mining.....	19
Tabel 3. 2 Perbandingan Bahasa Pemograman	27
Tabel 4. 1 Pengertian Data Survey Sosial Ekonomi	30
Tabel 4. 2 Pengertian Data Transkrip Nilai Mahasiswa.....	31
Tabel 4. 3 Hasil Uji Korelasi Data Survey	49
Tabel 4. 4 Hasil Uji Korelasi Data Transkrip Nilai.....	50

DAFTAR GAMBAR

Gambar 1. 1 Sebaran Peserta Didik Berprestasi berdasarkan Tahun 2021-2024 ...	2
Gambar 2. 1 Alur CRISP-DM.....	9
Gambar 3. 1 Alur Penelitian.....	20
Gambar 3. 2 Diagram Alur Data Preparation.....	22
Gambar 3. 3 Diagram Alur Modeling	23
Gambar 4. 1 Visualisasi Distribusi IPK Data Survey	32
Gambar 4. 2 Visualisasi Distribusi IPK Data Transkrip Nilai	33
Gambar 4. 3 Persebaran Data Pendidikan Orang Tua Data Survey.....	34
Gambar 4. 4 Persebaran Data Pendidikan Orang Tua Data Transkrip Nilai.....	35
Gambar 4. 15 Cek Nilai Null Data Survey	36
Gambar 4. 18 Encoding Data Survey	37
Gambar 4. 20 Cek Oulier Data Survey	38
Gambar 4. 5 Code Apps Script Handling Struktur Data	39
Gambar 4. 6 Hasil Handling Struktur Data	40
Gambar 4. 7 Penentuan Nama Kolom Nilai, Kehadiran, dan Mata Kuliah	41
Gambar 4. 8 Hasil Kolom Nilai, Kehadiran, dan Mata Kuliah.....	41
Gambar 4. 9 Proses Konversi Kolom Kehadiran ke Format Numerik	42
Gambar 4. 10 Tipe Data Kolom Kehadiran Setelah Konversi.....	42
Gambar 4. 11 Proses Konversi Nilai Angka ke Nilai Huruf Mahasiswa.....	43
Gambar 4. 12 Hasil Distribusi Frekuensi Nilai Huruf Mahasiswa	44
Gambar 4. 13 Penambahan Kolom SKS dan Perhitungan IPK	44
Gambar 4. 14 Hasil IPK Mahasiswa Berdasarkan NIM	45
Gambar 4. 16 Cek Nilai Null Data Transkrip Nilai.....	45
Gambar 4. 17 Code Hapus Nilai Null Data Transkrip Nilai	46
Gambar 4. 19 Encoding Data Transkrip Nilai.....	46

Gambar 4. 21 Hapus Outlier Data Transkrip Nilai	47
Gambar 4. 22 Visualisasi Sebelum dan Sesudah Handling Outlier	48
Gambar 4. 23 Code Uji Korelasi Spearman Data Survey	49
Gambar 4. 24 Code Uji Korelasi Spearman Data Transkrip Nilai	50
Gambar 4. 25 Visualisasi Heatmap Uji Korelasi Spearman Data Survey	52
Gambar 4. 26 Visualisasi Heatmap Uji Korelasi Spearman Data Transkrip Nilai	53
Gambar 4. 27 Kode Untuk Kesimpulan Hasil Korelasi Spearman	54
Gambar 4. 28 Kode uji korelasi Spearman untuk data survei	55
Gambar 4. 29 Kode pemilihan jenis data pada antarmuka aplikasi Streamlit	56
Gambar 4. 30 Tampilan hasil korelasi data survei pada aplikasi Streamlit	57
Gambar 4. 31 Tampilan form upload file transkrip nilai mahasiswa	57
Gambar 4. 32 Tampilan hasil korelasi dan heatmap dari data transkrip nilai	58

DAFTAR LAMPIRAN

Lampiran A Surat Pengantar MBKM - MBKM 01	66
Lampiran B Kartu MBKM - MBKM 02.....	67
Lampiran C Daily Task MBKM - MBKM 03	125
Lampiran D Lembar Verifikasi Laporan MBKM - MBKM 04.....	126
Lampiran E Surat Penerimaan MBKM (LoA).....	127
Lampiran F Lampiran Pengecekan Hasil Turnitin.....	128
Lampiran G Semua hasil karya tugas yang dilakukan selama MBKM	129