



**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN
RISET, DAN TEKNOLOGI
UNIVERSITAS DIPONEGORO
FAKULTAS SAINS DAN MATEMATIKA**

Jalan Prof. Jacob Rais
Kampus Universitas Diponegoro
Tembalang, Semarang, Kode Pos 50275
Telp (024) 7474754 Fax (024) 70460690
Laman: <https://fsm.undip.ac.id>
Pos-el: fsm@atundip.ac.id

UJIAN TENGAH SEMESTER GASAL 2023/2024

Mata Kuliah	:	Penambangan Data
Kelas	:	A, B, C, D
Pengampu	:	Nurdin Bahtiar, S.Si., M.T Dr. Indra Waspada, S.T., M.TI.
Departemen/Program Studi	:	Informatika
Hari/Tanggal	:	Senin, 16 Oktober 2023
Jam/Ruang	:	10:00:00-11:40:00 (100 menit) / E101, E102, E103, A303
Sifat Ujian	:	Buku Tertutup

Capaian Pembelajaran Lulusan (CPL)	CPL-12: Mampu menghasilkan rancangan solusi manajemen informasi dan mengimplementasikan solusi menggunakan pendekatan data analytics.
Capaian Pembelajaran Mata Kuliah (CPMK) dan Sub-CPMK	CPMK-12.2: Mampu mengimplementasikan solusi menggunakan pendekatan data analytics. Sub CPMK-12.2: 1. Mampu mengimplementasikan (C3) framework data mining untuk solusi data analytics 2. Mampu mengidentifikasi (C4) tipe atribut serta target untuk berbagai masalah data dan solusinya 3. Mampu mempersiapkan dan memproses (C3) data untuk model solusi pengelompokan, regresi, dan klasifikasi 4. Mampu mengimplementasi (C3) berbagai jenis visualisasi untuk data analytics dalam membantu memperoleh pengetahuan / insight dari data 5. Mampu mengimplementasi (C3) model untuk data analytics

Petunjuk Pengerjaan:

- Tuliskan identitas **Nama**, **NIM**, dan **Tanda Tangan** Anda pada lembar jawab!
- Jawablah soal-soal berikut pada lembar jawab dan bila perlu disertai asumsi/gambar!
- Jawablah **SOAL A** dan **SOAL B** pada **lembar jawab yang terpisah**.

SOAL A [Sub CPMK-12.2 – 1, 3 Bobot 30%]

- {10} Jelaskan gambaran karakteristik data sehingga menyebabkan diperlukannya *data preprocessing* menurut Jiawei Han!
- {10} Jelaskan macam-macam data berdasarkan derajat pengukurannya disertai contohnya!
- {10} Jelaskan disertai gambar tahapan yang ada pada CRISP-DM!



SOAL B [Sub CPMK-12.2 – 2.3 Bobot 70%]

Perhatikan tabel data pelanggan toko online ABC berikut:

ID Pelanggan	Jenis Kelamin	Asal Negara	Tahun Lahir	Usia	Penghasilan Tahunan (juta)	Nilai Belanja (1-100)
1	Pria	Indonesia	1999	24	100	39
2	Pria	Indonesia	1990	33	150	81
3	Wanita	Indonesia	2000	23		6
4		Indonesia	1999	24	80	77
5	Wanita	Indonesia	1998	25	50	30
6	Wanita	Indonesia	1993	30	120	76
7	Wanita	Indonesia	1985	38	300	6
8	Perempuan	Indonesia	1980	43	200	94
9	Pria	Indonesia	1995	28	100	85
10	Wanita	Indonesia	2012	11	150	72

Budi sebagai manajer toko online ABC ingin memahami perilaku pelanggannya untuk mendapatkan pola-pola menarik yang dapat mendukung strategi pemasaran. Sebagai langkah awal penambangan data, bantulah Budi mengolah data tabel tersebut agar menjadi lebih baik (terutama untuk menyiapkan tahap pemodelan) mengacu pada metodologi CRISP-DM:

- {25} Terapkan cara memahami data (*data understanding*) langkah demi langkah menurut panduan CRISP-DM. Pastikan pula dalam langkah tersebut saudara mengidentifikasi dan menjelaskan kemungkinan adanya permasalahan pada data, misalnya: *missing value*, data *outlier*, kebutuhan normalisasi, data tidak konsisten, potensi reduksi atribut, dan sebagainya.
- {25} Lanjutkan dengan menguraikan tahap persiapan data (*data preparation*) terutama dalam memberi solusi terhadap hasil identifikasi permasalahan pada bagian (a).
- {10} Rangkum uraian permasalahan beserta solusi pada bagian (a) dan (b) ke dalam tabel dengan format berikut:

No	Jenis Permasalahan	Lokasi Terjadinya Permasalahan	Alternatif solusi	Solusi dipilih	Alasan pemilihan solusi
1	Permasalahan x	atribut xyz, id 2	rumus p, rumus q, Teknik m	rumus p	rumus p sesuai dengan....
2	Permasalahan y	atribut abc dan atribut xyz	cara 1, cara 2	cara 2	cara 2 lebih ...
...

- {10} Sajikan tabel data pelanggan online versi akhir setelah melalui dua tahapan (a) dan (b).