



INSTITUT TEKNOLOGI DEL

UJIAN TENGAH SEMESTER

Semester I 2025/2026

4143203 - Data Science

Date / Time	: 16 – 17 Oktober 2025
Course	: 4143203 - Data Science
Assignment Type	: Proyek (Presentasi)
Time	: 8 minutes / Team
Lecturer	Oppir Hutapea, S.Tr.Kom., M.Kom
Academic Assistant	- Revi Angeli Siahaan, S.Tr.Kom. - Tantri Harianti Silaen, S.Tr.Kom.
Submission	: Ecourse

PETUNJUK

Sebelum mengerjakan soal UTS, Anda diharapkan untuk membaca petunjuk pengerjaan berikut ini:

1. Total skor untuk pengerjaan UTS Teori ini adalah **100**
2. Ujian bersifat *team project* yang akan dipresentasikan sesuai dengan waktu yang akan ditentukan.
3. Project disubmit dalam bentuk :
 - a. Source Code (**UTS-Prak-NIM1_NIM2_NIM3**), NIM = 3 digit angka terakhir Nim Anda.
 - b. Slide Presentasi (.ppt, pdf, dll) dengan penamaan (**Presentasi_UTS_Prak_NoKelompok**)
 - c. Laporan Praktikum UTS dengan penamaan : **Lap_UTS_Prak_NoKelompok**
 - d. Ketiga file tersebut disubmit ke ecourse dalam satu bentuk file .zip dengan penamaan : UTS-PRAK-NoKelompok. ZIP
 - e. Semua folder project diunggah ke github dengan menambahkan account saya (**oppir07**)
4. Anda diharapkan memenuhi kode etik mahasiswa ketika ujian. Gagal dalam memenuhi hal tersebut, seperti mencontek atau bekerja sama dengan teman Anda, maka Anda dianggap melakukan pelanggaran dan dapat diberi sanksi sesuai kode etik mahasiswa IT Del.
5. Gagal dalam melakukan poin nomor 4 maka nilai secara otomatis adalah **0**

Syarat dan Ketentuan Proyek Data Science Mahasiswa

1. Ruang Lingkup Topik

Proyek hanya mencakup:

1. **Data Collection**
2. **Data Visualization**
3. **Data Processing and Techniques (Advance Preprocessing)**
4. **Statistical Analysis**

2. Dataset

1. Dataset **bebas** (boleh dari sumber terbuka, scraping, atau dikumpulkan sendiri).
2. Harus memiliki **minimal 20–30 fitur (features/kolom)**.
3. Harus memiliki **minimal 2000 baris data (rows/observations)**.
4. Dataset harus **asli** (tidak boleh hanya dummy/random tanpa konteks).

3. Data Collection

1. Mahasiswa wajib menjelaskan **sumber data**.
2. Jika scraping, sertakan **script/program** dan **referensi situs**.
3. Jika dataset publik, sertakan **tautan sumber resmi**.

4. Data Visualization

1. Visualisasi data **tidak hanya menampilkan grafik**, tetapi harus:
 - a) Menunjukkan **alasan pemilihan jenis visualisasi**.
 - b) Memberikan **interpretasi/penjelasan** terhadap pola atau insight yang muncul.
 - c) Menggunakan minimal **3 jenis visualisasi berbeda** (misalnya: scatter plot, bar chart, heatmap, histogram, boxplot, dll).

5. Data Processing and Techniques (Advance Preprocessing)

1. Harus mencakup minimal 2 teknik preprocessing lanjutan, misalnya:
 - a) Handling missing values (mean/mode imputation, interpolation, dll).
 - b) Handling outliers (IQR, Winsorization, z-score).
 - c) Feature scaling (Normalization/Standardization).
 - d) Encoding categorical variables (One-Hot, Target Encoding).
 - e) Feature selection/reduction (PCA, LDA, Feature Importance).

6. Statistical Analysis

1. Mahasiswa wajib melakukan **minimal 1 uji parametrik** (misalnya: t-test, ANOVA, Pearson correlation).
2. Mahasiswa wajib melakukan **minimal 1 uji non-parametrik** (misalnya: Mann-Whitney U, Kruskal-Wallis, Spearman correlation).
3. Hasil uji harus disertai **interpretasi nilai p-value, effect size, dan confidence interval** jika relevan.

7. Laporan

1. Laporan harus terdiri dari:
 - a) Pendahuluan (latar belakang, tujuan, dan rumusan masalah).

- b) Metode (penjelasan alur pengerjaan data science: collection, visualization, preprocessing, analysis).
- c) Hasil dan Pembahasan.
- d) Kesimpulan.

Matriks Penilaian

Aspek Penilaian	Kriteria	Bobot
Dataset	Dataset sesuai syarat (≥ 20 fitur, ≥ 2000 baris)	15
Data Collection	Penjelasan sumber data jelas, valid, dan dapat dipertanggungjawabkan	10
Data Processing (Advance Preprocessing)	<ul style="list-style-type: none"> Minimal 3 jenis visualisasi berbeda Alasan pemilihan jenis visualisasi jelas Interpretasi/insight dari visualisasi 	20
Data Visualization	<ul style="list-style-type: none"> Melakukan minimal 2 teknik preprocessing lanjutan Penjelasan metode yang digunakan & alasan memilihnya 	20
Statistical Analysis	<ul style="list-style-type: none"> Minimal 1 uji parametrik + 1 uji non-parametrik Interpretasi hasil (p-value, CI, effect size) Relevansi dengan dataset 	20
Laporan & Penyajian	<ul style="list-style-type: none"> Struktur laporan lengkap Penulisan rapi, sistematis, dan akademis Hasil analisis jelas dan mendalam 	20