

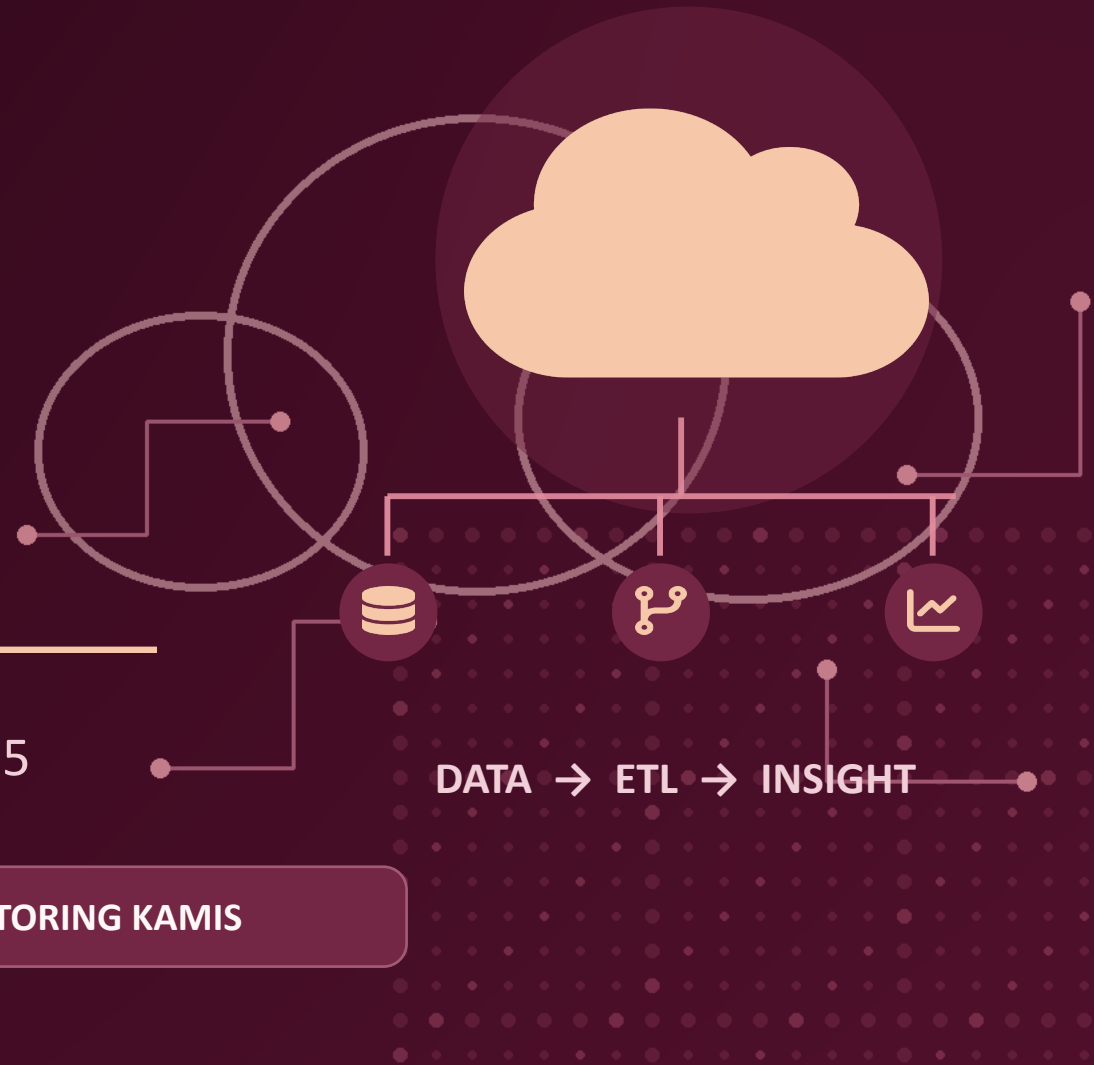
BIG DATA & CLOUD COMPUTING

Panduan proyek terstruktur untuk mahasiswa semester 5

5 KELOMPOK

6 MINGGU

MONITORING KAMIS



Bagian A–K: dari ide sampai demonstrasi



A Filosofi & capaian

B Aturan & alat

C 6 pilihan sektor proyek

D Clue menemukan ide

E Rencana 6 minggu

F Monitoring Rabu

G Rubrik penilaian

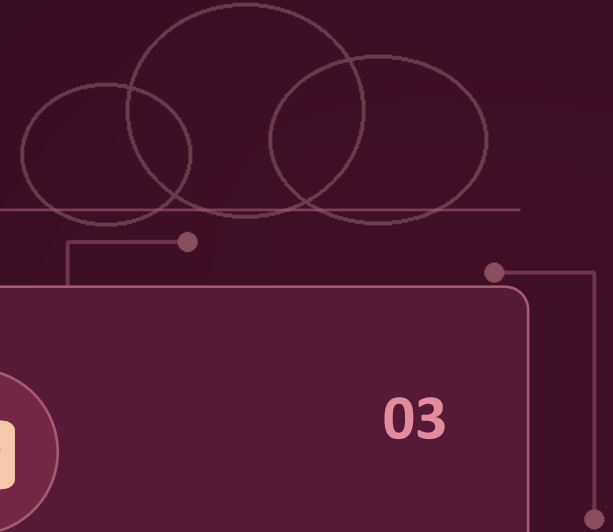
H Laporan akhir

I Presentasi akhir

J Checklist dosen

K Instruksi mahasiswa

Mengapa project-based learning?



01

MENEMUKAN

Berangkat dari masalah yang nyata dan kebutuhan pengguna yang terukur.



02

MEMBANGUN

Rakit pipeline data dan hubungkan hasilnya dengan layanan cloud.



03

MEMBUKTIKAN

Uji mutu data, fungsi dashboard, serta manfaat keputusan.

Pertanyaan utama proyek



Data apa yang sulit diolah atau dipantau, dan bagaimana cloud membuat hasilnya mudah diakses?



BERMANFAAT

Membantu keputusan



TERUKUR

Bukti analitik jelas



REALISTIS

Rampung 6 minggu

Capaian pembelajaran (1/2)



01

MASALAH

Definisikan pengguna, kebutuhan, dan keputusan yang ingin dibantu.



02

DATA

Nilai sumber, izin penggunaan, format, dan kualitas data.



03

ARSITEKTUR

Rancang aliran data dari sumber hingga aplikasi cloud.

Capaian pembelajaran (2/2)



ETL & ANALITIK

Bersihkan, gabungkan, ringkas, dan analisis data untuk menjawab masalah.



CLOUD

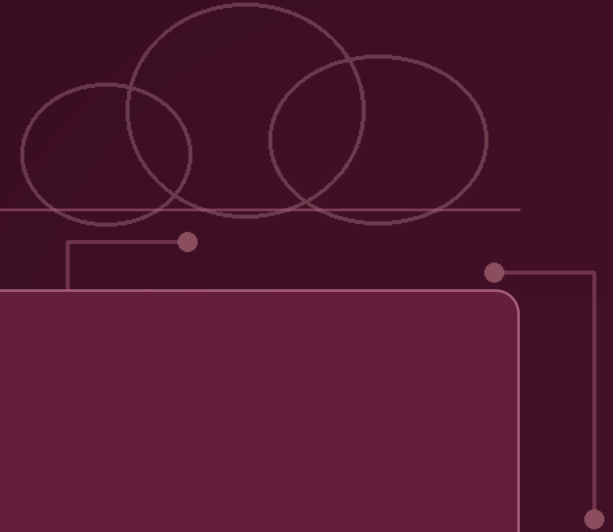
Simpan output dan sajikan dashboard yang dapat diakses secara daring.



VALIDASI

Periksa kebenaran angka, fungsi sistem, serta batas penggunaan hasil.

Big Data berbeda dari Cloud



BIG DATA

Tantangan volume, variasi, atau kecepatan data menuntut metode pengolahan yang sesuai.

Volume • Variety • Velocity

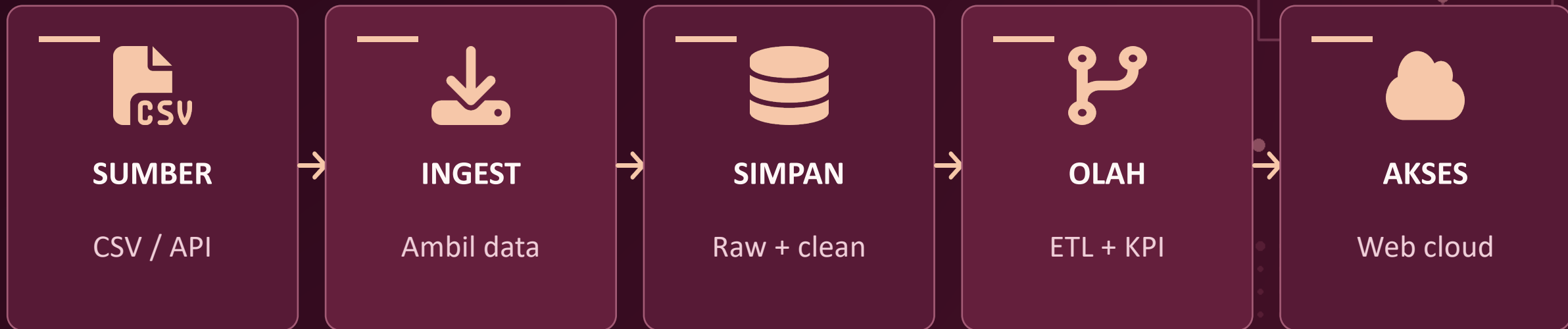


CLOUD COMPUTING

Komputasi, penyimpanan, dan layanan tersedia melalui jaringan tanpa membangun server sendiri.

Storage • Compute • Access

Alur kerja proyek end-to-end



BUKTI: dataset → pipeline → hasil analisis → tautan cloud app

Kontrak proyek kelas



KELOMPOK



MINGGU



MONITORING



PIPELINE



CLOUD APP



DEMO + LAPORAN

Empat bukti wajib setiap kelompok



DATA

Sumber, lisensi, raw dan clean.



PIPELINE

Kode ingest, ETL, dan analitik.



CLOUD

Aplikasi daring & hasil tersimpan.



DOKUMEN

Laporan, hasil uji, repo, dan video.

Apa yang membuatnya proyek Big Data?



VOLUME

Olah data dalam jumlah memadai untuk menunjukkan tantangan pemrosesan.



VARIETY

Gabungkan format, sumber, kategori, atau karakteristik data yang beragam.



VELOCITY

Peragakan pembaruan data secara berkala atau bertahap.

Minimum pipeline yang harus bekerja



01 / INGEST

Ambil CSV, API, atau log yang dapat diulang.



02 / TRANSFORM

Bersihkan, ubah, dan hitung indikator.



03 / STORAGE

Simpan raw dan cleaned secara terpisah.



04 / KONSUMSI

Tampilkan sekurang-kurangnya 3 KPI.

Alat yang realistis untuk 6 minggu



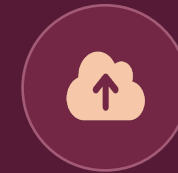
PEMROSESAN

Python + Colab; DuckDB, Polars, atau PySpark sesuai kebutuhan.



KOLABORASI

GitHub, notebook, README dan dokumentasi skema data.



DEPLOY CLOUD

Streamlit Community Cloud atau layanan hosting setara.

Privasi, etika, dan biaya



DATA AMAN

Gunakan data legal, anonim, atau sintetis. Tanpa data pasien, nasabah, dan militer operasional.



IZIN & LISENSI

Dokumentasikan hak akses, izin penggunaan, dan atribusi sumber.



BIAYA CLOUD

Utamakan free tier dan cegah biaya layanan tak terduga.

6 sektor pilihan untuk 5 kelompok



P1 / PENDIDIKAN

Dashboard Kehadiran & Aktivitas Belajar



P2 / KESEHATAN

Pemetaan Waktu Tunggu Puskesmas



P3 / MILITER

Dashboard Logistik Peralatan Latihan



P4 / PERBANKAN

Analitik Transaksi Bank Sintetis



P5 / UMKM

Dashboard Penjualan & Stok UMKM



P6 / TRANSPORTASI

Dashboard Ketepatan Waktu Transportasi

SETIAP KELOMPOK MEMILIH 1 TEMA • 1 SEKTOR BOLEH TIDAK TERPILIH

Dashboard Kehadiran & Aktivitas Belajar

PENDIDIKAN

Pengguna: Dosen / pengelola kelas

**01 / SUMBER DATA**

Log kehadiran & tugas anonim
atau sintetis

**02 / ETL & ANALITIK**

ETL + agregasi kelas dan periode

**03 / CLOUD APP**

Dasbor kehadiran & ketuntasan
tugas

MUDAH**MINIMUM / 3 KPI**

Kehadiran, ketepatan tugas, tren kelas

**BATASAN ETIKA & KLAIM**

Jangan menampilkan identitas dan nilai pribadi
mahasiswa.

Apa yang harus dibuktikan?



1

Dataset aktivitas belajar

≥20.000 rekaman sintetis/anonim; jelaskan skema.

2

Indikator yang terukur

Persentase hadir, tugas tepat waktu, tren mingguan.

3

Validasi analitik

Cocokkan total sesi, jumlah tugas, dan agregasi.



Jangan unggah NIM, nama, atau data akademik pribadi.

Pemetaan Waktu Tunggu Puskesmas

KESEHATAN



Pengguna: Petugas administrasi



01 / SUMBER DATA

Log antrean Puskesmas sintetis



02 / ETL & ANALITIK

Analitik durasi layanan & jam sibuk



03 / CLOUD APP

Dashboard antrean & beban loket

MUDAH



MINIMUM / 3 KPI

Tunggu rata-rata, P90, jam sibuk



BATASAN ETIKA & KLAIM

Tanpa NIK, nama pasien, diagnosis, atau rekam medis.

Apa yang harus dibuktikan?



1

Log layanan anonim

≥20.000 antrean sintetis dengan stempel waktu.

2

Analisis waktu tunggu

Rerata, persentil 90, dan intensitas tiap loket.

3

Pengujian logika

Uji duplikat, waktu negatif, dan antrean kosong.



Data pasien sesungguhnya tidak digunakan dalam proyek.

Dashboard Logistik Peralatan Latihan

MILITER

Pengguna: Pengelola perlengkapan (simulasi)

**01 / SUMBER DATA**

Inventaris & peminjaman sintetis

**02 / ETL & ANALITIK**

ETL stok, permintaan, dan pengembalian

**03 / CLOUD APP**

Pantauan stok perlengkapan latihan

SEDANG**MINIMUM / 3 KPI**

Stok tersedia, frekuensi pakai, pengembalian

**BATASAN ETIKA & KLAIM**

Tanpa lokasi satuan, personel, senjata, atau data operasi.

Apa yang harus dibuktikan?

1

Data non-operasional

≥30.000 catatan perlengkapan latihan sintetis.

2

KPI logistik dasar

Stok tersedia, tingkat penggunaan, dan retur.

3

Rekonsiliasi stok

Jumlah awal + masuk – keluar = saldo akhir.



Hanya simulasi inventaris umum; tidak memakai data pertahanan rahasia.

Analitik Transaksi Bank Sintetis

PERBANKAN



Pengguna: Analis layanan (simulasi)



01 / SUMBER DATA

Data transaksi fiktif & kanal layanan



02 / ETL & ANALITIK

ETL pola waktu, kanal, dan transaksi



03 / CLOUD APP

Dasbor transaksi & anomali sederhana



MINIMUM / 3 KPI

Volume, nilai, pola anomali berbasis aturan



BATASAN ETIKA & KLAIM

Bukan sistem fraud/AML nyata; tanpa rekening nasabah.

SEDANG

Apa yang harus dibuktikan?



1

Dataset transaksi aman

≥100.000 transaksi sintetis; tak ada rekening asli.

2

Analitik deskriptif

Volume transaksi, nilai agregat, pola per kanal.

3

Audit kualitas hasil

Uji duplikat, aturan anomali, dan jumlah total.



Jangan gunakan data nasabah atau mengklaim deteksi penipuan resmi.

Dashboard Penjualan & Stok UMKM

UMKM



Pengguna: Pemilik toko / pelaku UMKM



01 / SUMBER DATA

Transaksi sintetis atau berizin



02 / ETL & ANALITIK

ETL penjualan, produk, dan stok



03 / CLOUD APP

Dashboard omzet & persediaan

MUDAH



MINIMUM / 3 KPI

Omzet, barang laris, tren penjualan



BATASAN ETIKA & KLAIM

Data sintetis ditandai, bukan kondisi riil usaha.

Apa yang harus dibuktikan?



1

Data untuk uji skala

100.000–300.000 transaksi sintetis/berizin.

2

Indikator bisnis

Omzet, produk laris, dan frekuensi transaksi.

3

Cek konsistensi

Jumlah barang, omzet, dan saldo stok sesuai.



Jangan tampilkan data pelanggan riil tanpa izin dan anonimisasi.

Dashboard Ketepatan Waktu Transportasi

TRANSPORTASI



Pengguna: Pengelola rute / penumpang



01 / SUMBER DATA

Jadwal & kedatangan sintetis / terbuka



02 / ETL & ANALITIK

Normalisasi rute, waktu, dan durasi



03 / CLOUD APP

Peta rute & grafik keterlambatan



MINIMUM / 3 KPI

Tepat waktu, rerata terlambat, volume rute



BATASAN ETIKA & KLAIM

Bukan pelacakan langsung atau kendali lalu lintas.

SEDANG

Apa yang harus dibuktikan?

1

Data perjalanan terstruktur

≥40.000 log sintetis atau terbuka berlisensi.

2

KPI layanan transportasi

Persen tepat waktu, keterlambatan, volume rute.

3

Validasi waktu & lokasi

Uji zona waktu, durasi negatif, dan duplikasi.



Jangan mengklaim data simulasi sebagai jadwal real-time resmi.

Boleh memodifikasi ide proyek



SEKTOR

Tetap berada dalam Pendidikan, Kesehatan, Militer, Perbankan, UMKM, atau Transportasi.



DATA

Dataset legal, berizin, anonim, atau sepenuhnya sintetis.



CLOUD

Ada pipeline dan dashboard daring yang dapat diperiksa.

Jangan mulai dari memilih platform



01

MASALAH

Keputusan apa yang saat ini lambat, keliru, atau sulit dipantau?



02

DATA

Data aman dan legal apa yang dapat menjawab keputusan itu?



03

AKSI

Apa manfaat nyata dari hasil analisis dan dashboard?

Lima pertanyaan untuk calon pengguna



01

SIAPA?

Siapa pengguna dashboard?

02

MASALAH?

Apa keputusan utamanya?

03

KAPAN?

Kapan informasi dibutuhkan?

04

DATA?

Data apa tersedia dan dapat dipakai secara legal?

05

AKSI?

Apa tindakan yang akan diambil setelah melihat KPI?

Loloskan ide dengan empat filter



AKSES DATA

Sumber data, izin, dan lisensi jelas.



TEKNIS

ETL dapat dibangun dan diuji oleh tim.



CLOUD

Mungkin di-deploy dengan biaya nol/rendah.



WAKTU

Lingkup layak selesai dalam 6 minggu.

Ubah keluhan menjadi proyek data



NILAI PROYEK = MANFAAT PENGGUNA + PROSES YANG TERUJI

Enam minggu menuju demonstrasi



SETIAP RABU: tunjukkan artefak dan hasil yang dapat diperiksa

Masalah, data, dan proposal



MINGGU

01

PILIH IDE

✓ Pilih pengguna dan pertanyaan analitik.

✓ Pastikan dataset dapat diakses secara sah.

✓ Tulis proposal 2–3 halaman + 3 KPI.

✓ Tetapkan peran anggota dan risiko.

GATE: Ide & akses data disetujui dosen.

Eksplorasi data dan rancangan



MINGGU

02

RANCANG

✓ Bangun data dictionary dan skema.

✓ Lakukan profiling: null dan duplikat.

✓ Gambar arsitektur cloud end-to-end.

✓ Siapkan notebook dan repositori.

GATE: Data terbaca dan rancangan nyata.

Pipeline minimum harus hidup



MINGGU

03

BANGUN



Ingest dataset berjalan tanpa manual.



ETL membersihkan dan mengubah data.



Simpan output clean ke cloud storage.



Hitung minimal satu KPI dengan benar.

GATE: MVP wajib dapat didemokan.

Integrasi cloud dan dashboard



MINGGU

04

INTEGRASI

✓ Tampilkan minimal tiga indikator (KPI).

✓ Deploy dashboard di hosting cloud.

✓ Tambahkan filter waktu / kategori / rute.

✓ Uji lima skenario fungsi dashboard.

GATE: Dosen menerima tautan aplikasi.

Uji mutu, keamanan, dan perbaiki



MINGGU

05

VALIDASI

- ✓ Jalankan minimal 10 kasus uji akhir.
- ✓ Cocokkan sampel KPI dengan perhitungan.
- ✓ Pastikan tidak ada data sensitif / secret.
- ✓ Selesaikan 80% laporan dan video.

GATE: Sistem stabil untuk presentasi.

Presentasi dan demonstrasi proyek



MINGGU

06

DEMO AKHIR

- ✓ Jelaskan masalah, data, dan arsitektur.
- ✓ Jalankan demo dari sumber ke dashboard.
- ✓ Tunjukkan uji, keterbatasan, biaya.
- ✓ Serahkan laporan, kode, izin data, video.

GATE: Proyek dinilai final.

Minggu 3: jalur data wajib berhasil



LANJUT

Sumber → ETL → penyimpanan → satu KPI berhasil dijalankan ulang.



SEDERHANAKAN

Kurangi fitur; prioritaskan data yang tersedia, pemrosesan batch, dan dashboard.

Agenda kontrol 100 menit

00–10

TARGET

Ulas milestone dan kendala umum.

10–70

LIMA KELOMPOK

12 menit per kelompok: 3 + 5 + 4.

70–90

KLINIK

Bantu kendala ETL, cloud, dan data.

90–100

KOMITMEN

Catat target dan penanggung jawab.

Lima pertanyaan wajib tiap Rabu



- 01 Apa target yang disepakati minggu lalu?
- 02 Bukti apa yang dapat diperlihatkan sekarang?
- 03 Apa masalah data / kode / cloud yang terjadi?
- 04 Siapa mengerjakan bagian apa?
- 05 Apa target terukur untuk Rabu berikutnya?

Formulir ringkas monitoring mingguan



KELOMPOK / MINGGU

Nama proyek dan Rabu ke-...



TARGET & REALISASI

Target minggu ini dan bukti hasil.



KENDALA / SOLUSI

Masalah yang ditemukan dan tindakan perbaikan.



SKOR & ARAHAN

Skor 0–100, target, dan PIC berikutnya.

Gunakan traffic light progres



HIJAU / 80–100

Lanjut ke milestone berikutnya dengan tetap melaporkan bukti.



KUNING / 60–79

Tetapkan perbaikan yang spesifik, terukur, dan memiliki penanggung jawab.



MERAH / 0–59

Perlu pendampingan serta penyederhanaan lingkup atau data.

Rubrik penilaian akhir

**MASALAH & NILAI****15%****TEKNIK DATA / ETL****20%****CLOUD & FUNGSI****25%****UJI & ETIKA DATA****15%****PROGRES & TIM****15%****DEMO & LAPORAN****10%****TOTAL BOBOT PENILAIAN: 100%**

Nilai tim bukan otomatis nilai individu

NILAI AKHIR = 75% KELOMPOK + 25% INDIVIDU



KELOMPOK / 75%

Pipeline, cloud, mutu analisis, pengujian,
dan laporan.



INDIVIDU / 25%

Kontribusi, logbook, demo kode, dan
penilaian teman.

Kinerja diuji, bukan hanya dipamerkan



01

BENAR

Nilai KPI cocok dengan sampel hasil perhitungan manual.



02

STABIL

Input tidak valid ditangani tanpa membuat aplikasi berhenti.



03

DAPAT DIAKSES

Dashboard cloud dapat dibuka langsung oleh dosen.

Laporan ringkas 8–12 halaman



BAB 1 / PENDAHULUAN

Masalah, sasaran, pengguna, dan manfaat.



BAB 2 / DATA

Sumber, lisensi, skema, dan data dictionary.



BAB 3 / METODE

Arsitektur, ETL, dan layanan cloud.



BAB 4 / HASIL

Pengujian, temuan, dan dashboard.

Berkas yang wajib diserahkan



REPRODUKSI

Kode, README, requirements, dan diagram arsitektur.



DATA & UJI

Sampel dataset, lisensi, kasus uji, dan hasil KPI.



DEMONSTRASI

URL cloud app, slide presentasi, dan video singkat.

Alokasi demo lima kelompok



05 MENIT

Masalah, sumber data, dan rancangan arsitektur.



07 MENIT

Jalankan pipeline, dashboard, dan filter pada cloud app.



05 MENIT

Pengujian, keterbatasan, serta tanya jawab.

5 KELOMPOK × 17 MENIT = 85 MENIT

Tiga pertanyaan saat demonstrasi



01

ASAL DATA

Angka pada dashboard berasal dari kolom dan transformasi mana?



02

CLOUD

Apakah tautan dashboard dapat dibuka pada perangkat dosen?



03

KETERBATASAN

Dalam keadaan apa hasil tersebut tidak dapat dipercaya?

Pemeriksaan minggu 1 sampai 3

MASALAH & DATA

M1

Pengguna, masalah, izin akses, dan KPI disetujui.

PROFILING & SKEMA

M2

Data dictionary, eksplorasi, rancangan, dan repo tersedia.

PIPELINE MINIMUM

M3

Raw → clean → satu KPI berjalan dan dapat diuji.

Pemeriksaan minggu 4 sampai 6

CLOUD & 3 KPI

M4

Tautan dashboard cloud dan tiga KPI dapat diuji.

UJI AKHIR

M5

Sepuluh uji, audit hasil, laporan 80% selesai.

DEMO & SERAH TERIMA

M6

Demo, laporan, repositori, dan video diserahkan.

Ketentuan tugas kelompok



01 / PILIH MASALAH

Satu dari enam sektor dan dataset legal.



02 / BANGUN PIPELINE

Alur ETL, tiga KPI, dan dashboard cloud.



03 / LAPOR RABU

Tunjukkan bukti progres setiap Rabu.



04 / DEMO MINGGU 6

Uji sistem, jelaskan batasan, dan presentasi.

TANTANGAN 6 MINGGU

SATU MASALAH
SATU PIPELINE
SATU CLOUD APP



Buktikan manfaat dengan data, bukan hanya tampilan.