

CAPSTONE PROJECT MODULE 2

Pada Modul 2: Data Analysis, siswa telah mempelajari materi SQL, *data manipulation*, *data visualization*, dan *statistics*. *Capstone Project* modul 2 ini bertujuan untuk mengimplementasikan materi yang telah dipelajari oleh siswa dalam sebuah proyek.

Pengerjaan *Capstone Project* modul 2 ini diharapkan mampu melatih siswa dalam membuat sebuah rangkaian analisis. Dalam proyek ini, siswa akan memosisikan diri sebagai data analyst dalam sebuah perusahaan. Siswa akan membuat serangkaian pertanyaan sesuai dengan fokus analisis, pertanyaan dibuat berdasarkan keperluan bisnis dan masalah yang mungkin dihadapi oleh perusahaan terkait sesuai dengan fokus analisisnya.

Siswa akan menjawab pertanyaan yang telah dibuat melalui analisis data, membuat cerita dan visualisasi, serta mempresentasikan insight dari analisis yang kemudian bisa digunakan *stakeholders* sebagai dasar pengambilan keputusan bisnis.

Penilaian

- **Video & Dokumentasi: 25**

1. Penjelasan tentang data dan konteks permasalahan yang hendak dianalisis.
2. Penjelasan tentang bagaimana pertanyaan dijawab, dan apakah jawaban tersebut sudah mampu menjawab keperluan/permasalahan bisnis.
3. Dokumentasi *code* dibuat secara lengkap dalam format Jupyter Notebook (.ipynb).

- **SQL: 20**

1. Koneksi SQL ke Python.
2. Penggunaan fungsi CTE dan/atau *Window Function* (Jika memang dibutuhkan).
3. Pemanfaatan *built in function* sesuai kebutuhan analisis.

- **Python for Data Manipulation: 15**

1. Pencarian anomali: *missing values*, *outlier*, *data formatting*, dan/atau data duplikat.
2. Cara dan justifikasi penanganan *data anomalies* (jika ada).
3. Penggunaan *built-in function* di Pandas yang sesuai dengan kebutuhan analisis.

- **Data Visualization: 25**

1. Penggunaan *tools* dalam visualisasi tidak dibatasi. Sertakan alasan singkat mengapa *tools* tersebut digunakan (baik Python atau Tableau).
2. Penggunaan grafik yang sesuai untuk menjawab pertanyaan bisnis yang telah dibuat.
3. Penjelasan *insight*/cerita/informasi dari grafik yang digunakan berdasarkan keperluan/masalah bisnis.

- **Statistics: 15**

1. Uji perbandingan nilai/proporsi dengan metode yang sesuai kebutuhan analisis.
2. Analisis hubungan antara fokus analisis dengan variabel lain.
3. Penjelasan *insight* dari hasil uji statistika deskriptif dan inferensial berdasarkan keperluan/masalah bisnis.

Dataset

1. [Northwind](#)

- 'customers' sebagai fokus analisis
- 'suppliers' sebagai fokus analisis
- 'products' sebagai fokus analisis
- 'orders' sebagai fokus analisis
- 'employees' sebagai fokus analisis

2. [Human Resource](#)

- Analisis bebas

3. **Sakila (SUDAH TERSEDIA DI DATABASE MASING-MASING)**

- 'Film' sebagai fokus analisis
- 'Customers' sebagai fokus analisis
- 'Store & staff' sebagai fokus analisis

Lama Pengerjaan & Pengumpulan

Capstone Project modul 2 ini memiliki batas waktu pengumpulan **paling lambat 10 hari kerja** setelah ujian modul 2. Pengumpulan dilakukan dengan cara:

1. Unggah video penjelasan proyek ke dalam *cloud storage* (Youtube, Google Drive, Dropbox) masing-masing siswa (buka hak akses untuk publik).
2. Jika siswa menggunakan Tableau untuk visualisasi, siswa dapat:
 - a. Memasukan grafik dari Tableau sebagai *image* di Jupyter Notebook.
 - b. Memakai `iframe` untuk memasukkan *image* dari Tableau ke jupyter Notebook.
 - c. Memasukkan *file* tableau ke dalam *cloud storage* masing-masing. *File* dapat diunduh dari Tableau Public.
3. Video penjelasan berdurasi 10-15 menit.
4. Membuat *file readme* (dalam bentuk *markdown*) yang menjelaskan tentang proyek secara garis besar.
5. Upload *file* jupyter Notebook proyek ke Github masing-masing siswa.
6. Menginformasikan *link* data tersebut ke form yang telah disediakan oleh operasional.