

## Soal Final Project Data Science Academy COMPFEST 13

Sebagai bentuk pengaplikasian materi yang telah didapatkan selama Camp-Camp yang telah diikuti selama penyelenggaraan Data Science Academy, Final Project hadir sebagai penutup dari rangkaian acara Data Science Academy COMPFEST 13. Final project sendiri menghasilkan *deliverables* berupa presentasi final project.

Hasil pengerjaan dari Final Project ini akan dipresentasikan pada Camp 3 Day 3:

Hari : Sabtu

Tanggal : 04 September 2021

Pukul : 09.00 - 15.20 WIB

Dengan teknis presentasi sebagai berikut:

1. Setiap tim akan mempresentasikan hasil Final Project dalam bentuk Powerpoint dengan durasi 30 menit, dengan 15 menit pertama adalah presentasi dan 15 menit berikutnya dilanjutkan dengan sesi tanya jawab dengan panelis.
2. Terdapat 3 juri yang akan bertanggung jawab untuk melakukan penilaian terhadap Final Project setiap tim.
3. Urutan presentasi tim akan ditentukan secara acak, di mana hasil pengurutan tersebut akan diberitahukan satu hari sebelum hari presentasi oleh panitia DSA melalui channel Slack.
4. Seluruh peserta wajib mengikuti rangkaian acara secara penuh dari pembukaan hingga penutupan.

### **Dataset:**

Data COVID-19 dunia/global

<https://www.worldometers.info/coronavirus>

<https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset>

<https://www.kaggle.com/allen-institute-for-ai/CORD-19-research-challenge>

Data COVID-19 Indonesia & Jakarta

<https://covid19.go.id/peta-risiko>

<https://corona.jakarta.go.id/id>

<https://tiny.cc/Datacovidjakarta>

<https://riwayat-file-covid-19-dki-jakarta-jakartagis.hub.arcgis.com>

Data mobilitas global, Indonesia & Jakarta

<https://www.google.com/covid19/mobility>

<https://covid19.apple.com/mobility>

<https://dataforgood.fb.com/docs/covid19>

### **Problem Description:**

Pandemi COVID-19 telah berlangsung selama 1 tahun lebih dan memiliki dampak sangat signifikan secara global, termasuk di Indonesia. Sangatlah dibutuhkan pengambilan keputusan (strategis & operasional) berbasis data, sehingga kebijakan dan keputusan yang dipilih adalah yang paling baik dan relevan, baik itu oleh pemerintah, industri, dan masyarakat.

Melalui **data COVID-19 global, Indonesia, dan mobilitas**, silakan kaji mengenai masalah utama yang anda rasa paling penting untuk ditelusuri lebih jauh dan dibuatkan model prediksi. Cakupan masalah terkait COVID-19 dibebaskan sesuai dengan kreativitas, keahlian, dan pengalaman anda (misal: kondisi umum kesehatan, kondisi penularan, penanganan angka kematian, efektivitas pembatasan sosial, dll).

Pastikan anda dapat menjelaskan signifikansi/dampak dari cakupan masalah yang anda pilih, dan bagaimana analisis/rekomendasi/model prediksi anda mampu menjawab cakupan masalah tersebut.

### **Komponen hasil akhir yang diharapkan:**

- Paparan lengkap : Slides & MS Word (opsional)

- Code : Python
- File/link dashboard live : Tableau/Google Data Studio
- Link repository Github

### Anjuran/bonus:

Silakan eksplorasi dataset (publik) lain yang berguna untuk memperkaya dan mempertajam analisis dan hasil akhir rekomendasi kalian. Sangat dianjurkan untuk menggunakan/menambahkan dataset lain selain yang diberikan di sini.

### Instruksi:

- Pastikan proses dan paparan mencakup:
  - perumusan masalah,
  - hipotesis,
  - penentuan metodologi dan variabel,
  - analisis awal,
  - analisis mendalam,
  - kesimpulan dan rekomendasi.
- Lakukan *initial data assessment & pre-processing*.
- Lakukan exploratory data analysis (EDA) untuk mengetahui konteks data secara umum dan menyeluruh. Dari langkah ini, akan didapatkan aspek yang bisa dianalisis lebih dalam.
- Berikan pertimbangan dan evaluasi beberapa opsi data model dan usaha prediksi. Jelaskan secara detail dari aspek pemilihan variabel, *feature engineering*, dan *hyper-parameter tuning*.
- Jelaskan kaitannya dengan konsep “[wicked problem](#)” untuk melakukan usaha terkait prediksi pola COVID-19.