

Case Study Seleksi DSLS 2023



Anda adalah seorang *data consultant* yang menangani klien *New York City Taxi and Limousine Commission (TLC)* di New York. Anda diminta untuk menyelesaikan masalah dan menawarkan solusi pada masalah yang ada di bawah ini. Silakan untuk menggunakan kaidah-kaidah *data analytics* dan *problem solving* kalian untuk menyelesaikan masalah tersebut.

Dataset dan informasi pendukung lain bisa diakses pada link berikut : [Dataset DSLS 2023](#)

Latar Belakang

Mayoritas warga New York bergantung kepada layanan transportasi publik atau layanan taksi. Hanya 22% dari penduduk Manhattan memiliki mobil jika dibandingkan dengan rata-rata 91% rumah tangga di seluruh Amerika Serikat memiliki setidaknya satu mobil. Layanan taksi di New York merupakan jaringan transportasi terbesar keempat di Amerika Serikat. Sistem tersebut diatur oleh *New York City Taxi and Limousine Commission (TLC)*, yang mengawasi *yellow taxi*, taksi sewaan, mobil commuter, kendaraan transit, dan limosin tertentu. Terlepas dari jaringannya yang besar, sistem yang ada sekarang tidak melayani daerah (*boroughs*) secara merata. Karena hal tersebut, penduduk lebih memilih layanan taksi ilegal yang memiliki akses dan tarif yang tidak konsisten.

Detail Tambahan :

- **Borough** : Semacam daerah/regional tertentu di New York. *Borough* di New York ada lima, yaitu Bronx; Brooklyn; Manhattan; Queens; dan Staten Islands.

- Informasi terkait *borough* pada dataset dapat dilihat di kolom **'PULocationID'** dan **'DOLocationID'**. Detail *borough* tersebut dapat juga dilihat pada peta setiap *borough* yang telah disediakan.
1. Menurutmu, apa yang menyebabkan layanan *yellow taxi* hanya berpusat di *boroughs* tertentu sehingga menyebabkan ketidakseimbangan permintaan dan penawaran pada *boroughs* lain?
 2. TLC ingin agar layanan *yellow taxi* mereka tersebar secara merata di seluruh kota New York sehingga dapat menjawab *demand* yang ada. Apa yang harus TLC lakukan secara *short-term* (6 bulan) untuk mengatasi tantangan ini?
 3. Buatlah script berbentuk *python notebook* (.ipynb) yang berisi tentang insight-insight (dapat didukung oleh visualisasi) yang anda temukan untuk menunjang analisa anda dalam menyelesaikan permasalahan yang ada? (Tidak ada standar khusus dalam pembuatan *script* ini. Hasil analisa anda lah yang akan menjadi penilaian)
 4. Buatlah **salah satu** dari :
 - **Dashboard** (Google Data Studio, Tableau, dll)
 - **Model Machine Learning** (dengan algoritma dan metode bebas)Salah satu dari *output* diatas harus dijadikan dasar untuk rekomendasi pada soal selanjutnya
 5. Dari hasil analisamu dengan data yang tersedia, apa rekomendasimu untuk menyelesaikan permasalahan yang ada di TLC?

Dalam studi kasus ini, *deliverables* yang diminta :

1. *Project report* untuk menjawab nomor 1, 2, dan 5
2. *Script* dalam bentuk format .ipynb dalam menjawab nomor 3
3. Link untuk mengakses dashboard atau script machine learning yang sudah dibuat dalam menjawab nomor 4

Tata Cara Pengumpulan :

1. Upload file-file yang dibutuhkan ke sebuah folder google drive anda
2. Beri nama **“Study Case_Nama Lengkap”** untuk semua file yang dikumpulkan
3. Pastikan folder tersebut dapat diakses oleh panitia
4. Submit link folder tersebut pada form yang sudah disediakan