



# Mini Project DS

---

FARHAN SHIDIQ SUDAMA

1A

# Business Understanding

---

Di zaman modern saat ini, transportasi menjadi kebutuhan penting bagi masyarakat untuk memudahkan aktivitas sehari-hari. Perkembangan ilmu pengetahuan dan teknologi telah mendorong kemajuan alat transportasi secara pesat, sehingga terjadi peningkatan jumlah kendaraan. Kendaraan sebagai alat transportasi memberikan dampak positif dalam memenuhi kebutuhan manusia seperti mempercepat aktivitas bisnis. Namun, peningkatan jumlah kendaraan juga mengakibatkan dampak negatif seperti kemacetan lalu lintas.

Untuk mengetahui tingkat kemacetan di beberapa kota di Indonesia, termasuk Jakarta yang merupakan kota termacet ke-12 di dunia, dilakukan pengumpulan data oleh Inrix dari 1.360 kota di 38 negara dengan fokus pada kemacetan sepanjang hari dan minggu. Rata-rata masyarakat Indonesia menghabiskan waktu sampai 47 jam dalam setahun karena macet. Oleh karena itu, perlu dilakukan cluster analysis menggunakan machine learning untuk menetapkan tingkat potensi kemacetan.

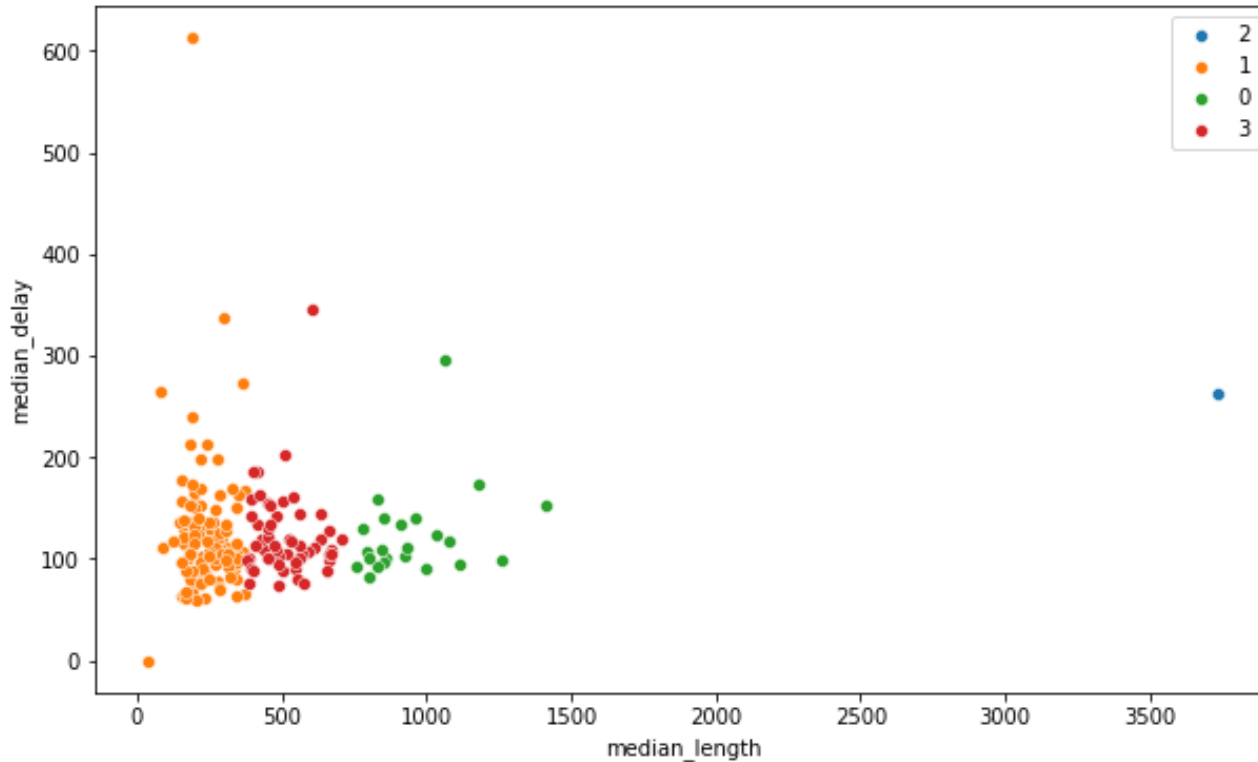
Proyek ini menggunakan dataset dari Waze, sebuah aplikasi navigasi, untuk mengetahui tingkat kemacetan di beberapa daerah di Indonesia seperti Cimahi, Bandung, Bekasi, Depok, Bogor, dan Tasikmalaya.

# K-Means Clustering

---

K-Means clustering adalah algoritma machine learning tanpa pengawasan (unsupervised) yang populer digunakan untuk mengelompokkan titik data menjadi kelompok. Tujuan dari K-Means clustering adalah mempartisi sejumlah titik data menjadi  $k$  kelompok berdasarkan jarak euclidean dari titik pusat kelompok yang disebut centroid. Algoritma ini biasanya digunakan untuk analisis data, segmentasi pelanggan, klasifikasi dokumen, dan pengelompokan gambar.

# Clustering



Cluster 2 merupakan cluster yang memiliki waktu dan panjang kemacetan pendek hingga menengah kebawah

Ada 1 jalan di Kota Cimahi masuk cluster 2

Cluster 3 merupakan cluster yang memiliki waktu dan panjang kemacetan yang paling lama

Ada 54 jalan di Kota Cimahi masuk cluster 3

Cluster 0 merupakan cluster yang memiliki waktu dan panjang kemacetan terpendek

Ada 23 jalan di Kota Cimahi masuk cluster 0

Cluster 1 merupakan cluster yang memiliki waktu dan panjang kemacetan menengah hingga menengah ke atas

Ada 116 jalan di Kota Cimahi masuk cluster 1

# Evaluasi

---

- K-Means clustering memiliki DB Score sebesar 0.44844077512821245
- Agglomerative clustering memiliki DB Score sebesar 0.4557898569131484

Dari nilai DB Score tersebut, dapat disimpulkan bahwa K-Means clustering memiliki kualitas clustering yang sedikit lebih baik dibandingkan Agglomerative clustering, karena nilai DB Score K-Means lebih rendah (lebih baik) daripada DB Score Agglomerative

*Your best quote that reflects your approach... “It’s one small step for man, one giant leap for mankind.”*

- NEIL ARMSTRONG