

Regresi - Open Dataset

Instruktur: Muhammad Hilman, Ph.D

Penyusun: Fariz Darari, Ph.D

Prasyarat

Pastikan aplikasi **Orange** sudah terpasang dengan baik di komputer anda.

1 Pendahuluan

Lakukan prediksi terhadap dua soal berikut ini.

1. Pilih 1 diantara tiga dataset yang telah disediakan (bills-and-tipping.zip, car-factory-electricity, dan toy-advertising-sales.zip).
2. Pilih 1 diantara sekian banyak dataset APBN yang disediakan pada laman Open Data Kementerian Keuangan.

2 Tugas

Hal-hal yang perlu dilakukan dan dijawab untuk setiap soal adalah sebagai berikut:

1. Lakukan eksplorasi data, perhatikan atribut yang ada beserta jumlah data yang tersedia. Perlukah dilakukan prapemrosesan data? Mengapa?
2. Apa saja observasi EDA yang dapat dilakukan? Mengapa?
3. Lakukan pemodelan regresi. Bagaimana nilai koefisiennya (intercept dan slope)?
4. Gunakan model tersebut untuk melakukan prediksi. Apakah prediksi yang dibuat cukup tepat?
5. Buat pemodelan regresi dengan cara membalik sumbu x dan sumbu y / dependent dan independent variable. Buatlah sendiri data untuk proses prediksi ini.