

Tugas 2

1. Terdapat suatu data “Volume Pengangkutan Sampah di Kali Sungai Situ Waduk Bulan Januari Tahun 2021”

Penjelasan variabel data ini berisi antara lain :

1. bulan : Bulan
2. titik_lokasi : Titik Lokasi

Penampungan Sampah

3. kecamatan : Kecamatan dimana titik lokasi penampungan sampah berada
4. wilayah : Wilayah dimana titik lokasi penampungan sampah berada
5. panjang/luas : Panjang atau luas area penanganan sampah
6. satuan : Satuan Panjang atau Luas
7. tanggal : Tanggal
8. volume_sampah_perhari (m3) : Volume Sampah per Hari (dengan satuan m3)

Data tersebut adalah data mentah yang masih kotor karena terdapat banyak noise. Lakukan pre-processing yaitu **mengatasi noise** (membuat data tidak konsisten menjadi konsisten) dan **mengisi missing value** terhadap dataset tersebut. (HINT: Anda dapat melakukan konsistensi data terlebih dahulu, misal mengatasi perbedaan pada kolom “satuan_panjang/luas”. Konsistensikan dengan nilai “panjang/luas”.) Berikan penjelasan langkah-langkah pengerjaan dan asumsi untuk penyelesaian soal tersebut. Penjelasan dapat disertai dengan diagram.

Anda dapat menggunakan Bahasa pemrograman Python dan beberapa libraries seperti pandas.

Jika Anda menggunakan alat bantu lainnya, misal Ms Excel, jelaskan urutan-urutan pengerjaannya.

2. Berikut ini adalah data resep dan nutrisi. Apakah ada outlier pada data tersebut? Jelaskan.