**Analisis Eksplorasi Data (EDA)**

**1. Ringkasan Dataset**

Dataset ini berisi 62.547 entri terkait pencatatan berat sampah yang ditimbang setiap hari selama periode 2019–2023. Setiap entri mencakup informasi tanggal, waktu, identitas kendaraan, jenis barang (dalam hal ini selalu Sampah), nama suplier, berat bersih (Netto\_kg), serta informasi petugas yang terlibat (sopir dan admin).

**2. Tinjauan Data**

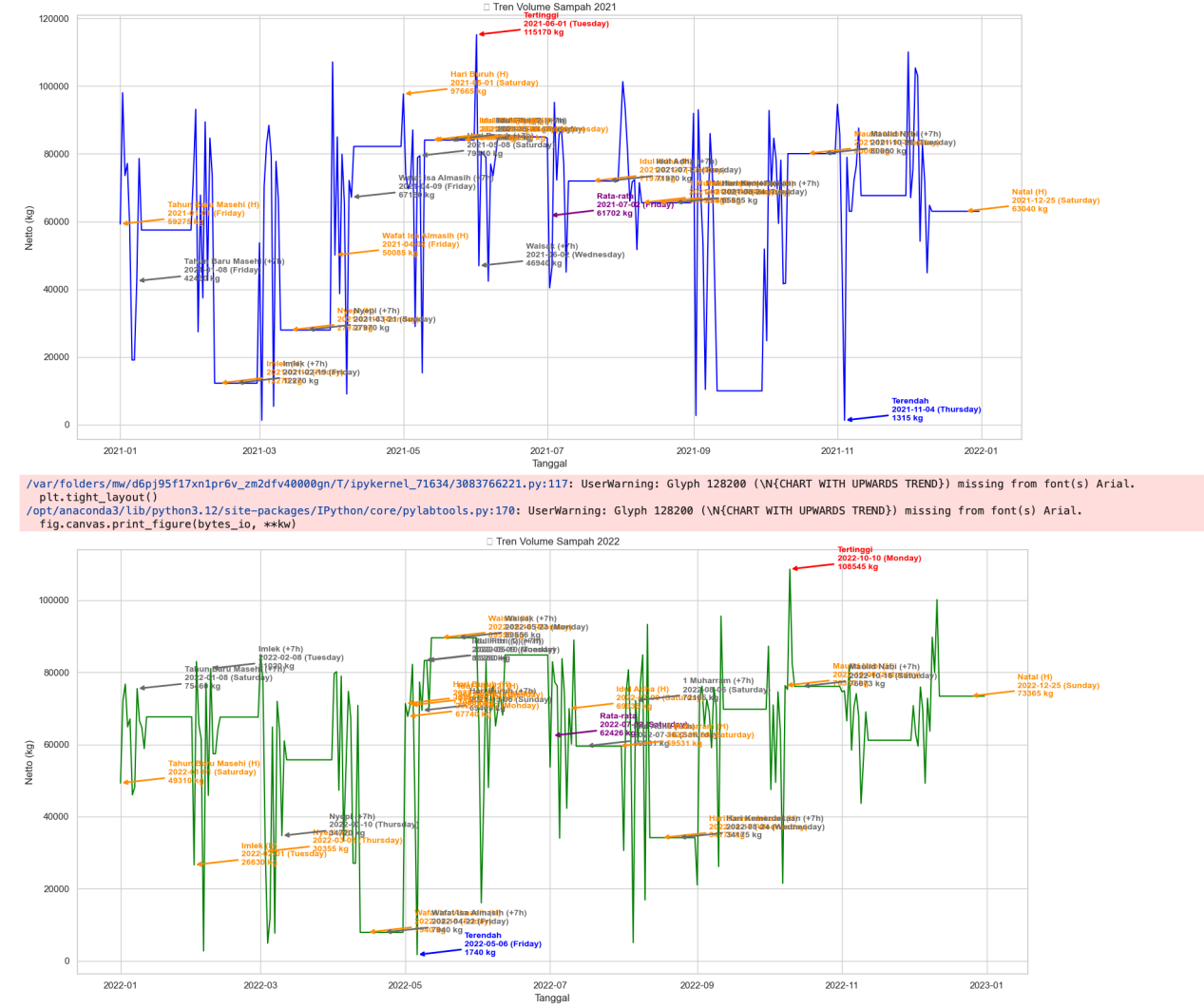
* Kolom Tanggal awalnya bertipe object dan telah dikonversi menjadi datetime untuk analisis waktu.
* Kolom Netto\_kg (target utama untuk prediksi) juga dikonversi ke tipe numerik. Tidak terdapat nilai kosong pada kolom ini setelah cleaning data.
* Kolom Sopir memiliki 325 nilai hilang dari total 62.547 entri (sekitar 0,5%)
* Statistik deskriptif menunjukkan:
  + Rata-rata berat harian (Netto\_kg) adalah 1.540 kg, dengan rentang dari 2 kg hingga 12.670 kg.
  + Penyebaran nilai Netto\_kg cukup besar (standar deviasi ~1.331 kg), mengindikasikan adanya fluktuasi yang cukup signifikan antar hari.

**3. Pembersihan dan Agregasi Data**

Data dibersihkan dengan menghapus entri yang memiliki nilai kosong di kolom Tanggal dan Netto\_kg. Kemudian dilakukan agregasi total berat harian per tanggal, dan difokuskan hanya pada tahun 2021 dan 2022 untuk dianalisis secara lebih mendalam. Hilangnya data pada tanggal-tanggal tertentu diatasi dengan metode forward fill, yang mengasumsikan nilai sebelumnya masih berlaku bila tidak ada entri baru.

1. **Visualisasi Tren Waktu**

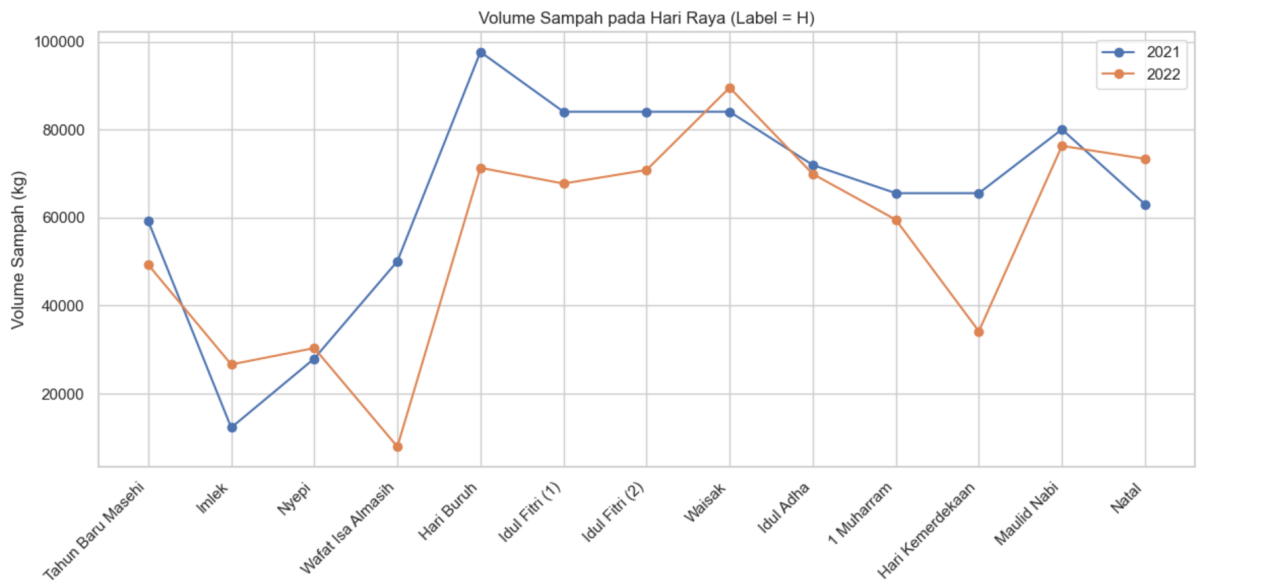
Berdasarkan data yang telah diexplorasi untuk periode tahun 2021 dan 2022, terdapat beberapa temuan yang cukup menarik terkait dengan hari besar dan volume sampah. Berdasarkan grafik yang telah saya buat, ada beberapa hari besar dengan volume sampah yang lebih tinggi dibandingkan dengan hari biasa.



Pada hari-hari besar ini, terjadi lonjakan dan penurunan yang cukup signifikan dalam volume sampah yang tercatat. Di sisi lain, hari-hari biasa cenderung menunjukkan volume sampah yang cukup konsisten, yang hampir mirip dengan rata-rata harian. Pola ini menunjukkan bahwa faktor-faktor tertentu, seperti liburan dan hari besar, berkontribusi terhadap peningkatan produksi sampah yang signifikan.

Data berat sampah saat hari Besar

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Tahun | Nama Hari Besar | Label | Tanggal | Hari | Netto\_kg |
| 2021 | Tahun Baru Masehi | (H) | 2021-01-01 | Friday | 59275.0 |
| 2021 | Tahun Baru Masehi | (+7h) | 2021-01-08 | Friday | 42460.0 |
| 2021 | Imlek | (H) | 2021-02-12 | Friday | 12270.0 |
| 2021 | Imlek | (+7h) | 2021-02-19 | Friday | 12270.0 |
| 2021 | Nyepi | (H) | 2021-03-14 | Sunday | 27970.0 |
| 2021 | Nyepi | (+7h) | 2021-03-21 | Sunday | 27970.0 |
| 2021 | Wafat Isa Almasih | (H) | 2021-04-02 | Friday | 50085.0 |
| 2021 | Wafat Isa Almasih | (+7h) | 2021-04-09 | Friday | 67160.0 |
| 2021 | Hari Buruh | (H) | 2021-05-01 | Saturday | 97665.0 |
| 2021 | Hari Buruh | (+7h) | 2021-05-08 | Saturday | 79440.0 |
| 2021 | Idul Fitri (1) | (H) | 2021-05-13 | Thursday | 84070.0 |
| 2021 | Idul Fitri (1) | (+7h) | 2021-05-20 | Thursday | 84070.0 |
| 2021 | Idul Fitri (2) | (H) | 2021-05-14 | Friday | 84070.0 |
| 2021 | Idul Fitri (2) | (+7h) | 2021-05-21 | Friday | 84070.0 |
| 2021 | Waisak | (H) | 2021-05-26 | Wednesday | 84070.0 |
| 2021 | Waisak | (+7h) | 2021-06-02 | Wednesday | 46940.0 |
| 2021 | Idul Adha | (H) | 2021-07-20 | Tuesday | 71970.0 |
| 2021 | Idul Adha | (+7h) | 2021-07-27 | Tuesday | 71970.0 |
| 2021 | 1 Muharram | (H) | 2021-08-10 | Tuesday | 65555.0 |
| 2021 | 1 Muharram | (+7h) | 2021-08-17 | Tuesday | 65555.0 |
| 2021 | Hari Kemerdekaan | (H) | 2021-08-17 | Tuesday | 65555.0 |
| 2021 | Hari Kemerdekaan | (+7h) | 2021-08-24 | Tuesday | 65555.0 |
| 2021 | Maulid Nabi | (H) | 2021-10-19 | Tuesday | 80060.0 |
| 2021 | Maulid Nabi | (+7h) | 2021-10-26 | Tuesday | 80060.0 |
| 2021 | Natal | (H) | 2021-12-25 | Saturday | 63040.0 |
| 2022 | Tahun Baru Masehi | (H) | 2022-01-01 | Saturday | 49310.0 |
| 2022 | Tahun Baru Masehi | (+7h) | 2022-01-08 | Saturday | 75460.0 |
| 2022 | Imlek | (H) | 2022-02-01 | Tuesday | 26630.0 |
| 2022 | Imlek | (+7h) | 2022-02-08 | Tuesday | 81020.0 |
| 2022 | Nyepi | (H) | 2022-03-03 | Thursday | 30355.0 |
| 2022 | Nyepi | (+7h) | 2022-03-10 | Thursday | 34720.0 |
| 2022 | Wafat Isa Almasih | (H) | 2022-04-15 | Friday | 7940.0 |
| 2022 | Wafat Isa Almasih | (+7h) | 2022-04-22 | Friday | 7940.0 |
| 2022 | Hari Buruh | (H) | 2022-05-01 | Sunday | 71325.0 |
| 2022 | Hari Buruh | (+7h) | 2022-05-08 | Sunday | 69405.0 |
| 2022 | Idul Fitri (1) | (H) | 2022-05-02 | Monday | 67740.0 |
| 2022 | Idul Fitri (1) | (+7h) | 2022-05-09 | Monday | 83260.0 |
| 2022 | Idul Fitri (2) | (H) | 2022-05-03 | Tuesday | 70830.0 |
| 2022 | Idul Fitri (2) | (+7h) | 2022-05-10 | Tuesday | 83260.0 |
| 2022 | Waisak | (H) | 2022-05-16 | Monday | 89556.0 |
| 2022 | Waisak | (+7h) | 2022-05-23 | Monday | 89556.0 |
| 2022 | Idul Adha | (H) | 2022-07-09 | Saturday | 69935.0 |
| 2022 | Idul Adha | (+7h) | 2022-07-16 | Saturday | 59531.0 |
| 2022 | 1 Muharram | (H) | 2022-07-30 | Saturday | 59531.0 |
| 2022 | 1 Muharram | (+7h) | 2022-08-06 | Saturday | 72165.0 |
| 2022 | Hari Kemerdekaan | (H) | 2022-08-17 | Wednesday | 34175.0 |
| 2022 | Hari Kemerdekaan | (+7h) | 2022-08-24 | Wednesday | 34175.0 |
| 2022 | Maulid Nabi | (H) | 2022-10-08 | Saturday | 76310.0 |
| 2022 | Maulid Nabi | (+7h) | 2022-10-15 | Saturday | 76073.0 |
| 2022 | Natal | (H) | 2022-12-25 | Sunday | 73365.0 |



**5. Hari besar dengan Sampah Tertinggi**

Berikut adalah hari-hari besar dengan volume sampah tertinggi berdasarkan kolom Netto\_kg:

* Hari Buruh 2021 (1 Mei 2021) dengan volume sampah sebesar 97,665 kg.
* Hari Buruh 2022 (1 Mei 2022) dengan volume sampah sebesar 71,325 kg.
* Idul Fitri 2021 (1) & (2) (13-14 Mei 2021) dengan volume sampah masing-masing 84,070 kg.
* Idul Fitri 2022 (1) & (2) (2-3 Mei 2022) dengan volume sampah masing-masing 67,740 kg dan 70,830 kg.
* Waisak 2022 (16 Mei 2022) dengan volume sampah sebesar 89,556 kg.

Dari sini, dapat dilihat bahwa Hari Buruh (baik pada 2021 dan 2022) dan Idul Fitri memiliki volume sampah yang sangat tinggi. Hari-hari ini kemungkinan besar diikuti dengan banyaknya aktivitas atau pebesaran yang menyebabkan peningkatan konsumsi dan produksi sampah.

**6. Hari besar dengan Sampah Terendah**

Sedangkan, berikut adalah hari besar dengan volume sampah terendah:

* Waisak 2022 (+7h) (23 Mei 2022) dengan volume sampah sebesar 46,940 kg.
* Waisak 2021 (+7h) (2 Juni 2021) dengan volume sampah sebesar 46,940 kg.
* Waisak 2022 (16 Mei 2022) dengan volume sampah sebesar 89,556 kg.
* Waisak 2021 (26 Mei 2021) dengan volume sampah sebesar 84,070 kg.

Terlihat bahwa selama pebesaran besar seperti Waisak dan Natal, volume sampah yang tercatat bisa sedikit lebih rendah dibandingkan dengan beberapa hari besar lain.

**7. Rata-rata Volume Sampah Pada Hari besar**

Berikut adalah perhitungan rata-rata volume sampah pada hari besar per tahun:

Tahun 2021:

* Tahun Baru Masehi: 59,275 kg
* Imlek: 12,270 kg
* Nyepi: 27,970 kg
* Waisak: 84,070 kg
* Idul Fitri (1 & 2): 84,070 kg
* Hari Buruh: 97,665 kg
* Idul Adha: 71,970 kg
* Maulid Nabi: 80,060 kg

Rata-rata 2021: 52,880.44 kg

Tahun 2022:

* Tahun Baru Masehi: 49,310 kg
* Imlek: 26,630 kg
* Nyepi: 30,355 kg
* Waisak: 89,556 kg
* Idul Fitri (1 & 2): 67,740 kg & 70,830 kg
* Hari Buruh: 71,325 kg
* Idul Adha: 69,935 kg
* Maulid Nabi: 76,310 kg

Rata-rata 2022: 59,035.63 kg

**8. Kesimpulan**

* Hari besar dengan volume sampah tertinggi adalah Hari Buruh di tahun 2021, dengan volume sampah sebesar 97,665 kg, diikuti oleh Idul Fitri dan Waisak.
* Hari besar dengan volume sampah terendah adalah Waisak pada tahun 2021 dan 2022, dengan volume sampah 46,940 kg, yang bisa dipengaruhi oleh faktor pebesaran yang tidak sebesar hari besar lainnya.

Rata-rata volume sampah pada tahun 2021 dan 2022:

* 2021: Volume sampah rata-rata pada hari besar sebesar 52,880.44 kg.
* 2022: Volume sampah rata-rata pada hari besar sebesar 59,035.63 kg.