

Projek UAS Analisis dan Visualisasi Data

Nama : mifta siti mariam

Kelas : TI 24 F

NIM : 20240040072

Laporan Singkat Analisis Data

1. Latar Belakang

- Pemilihan dataset dilakukan berdasarkan meningkatnya perhatian terhadap kualitas udara di wilayah perkotaan. Dataset yang digunakan adalah **data kualitas udara dari kota Jakarta tahun 2023**, yang mencakup parameter seperti PM2.5, suhu, kelembaban, dan kecepatan angin. Permasalahan yang ingin dikaji adalah **hubungan antara polusi udara (PM2.5) dengan faktor lingkungan**, khususnya suhu dan kelembaban, serta **prediksi tingkat polusi** berdasarkan variabel-variabel tersebut.

2. Metode Analisis

Analisis dilakukan melalui beberapa tahap:

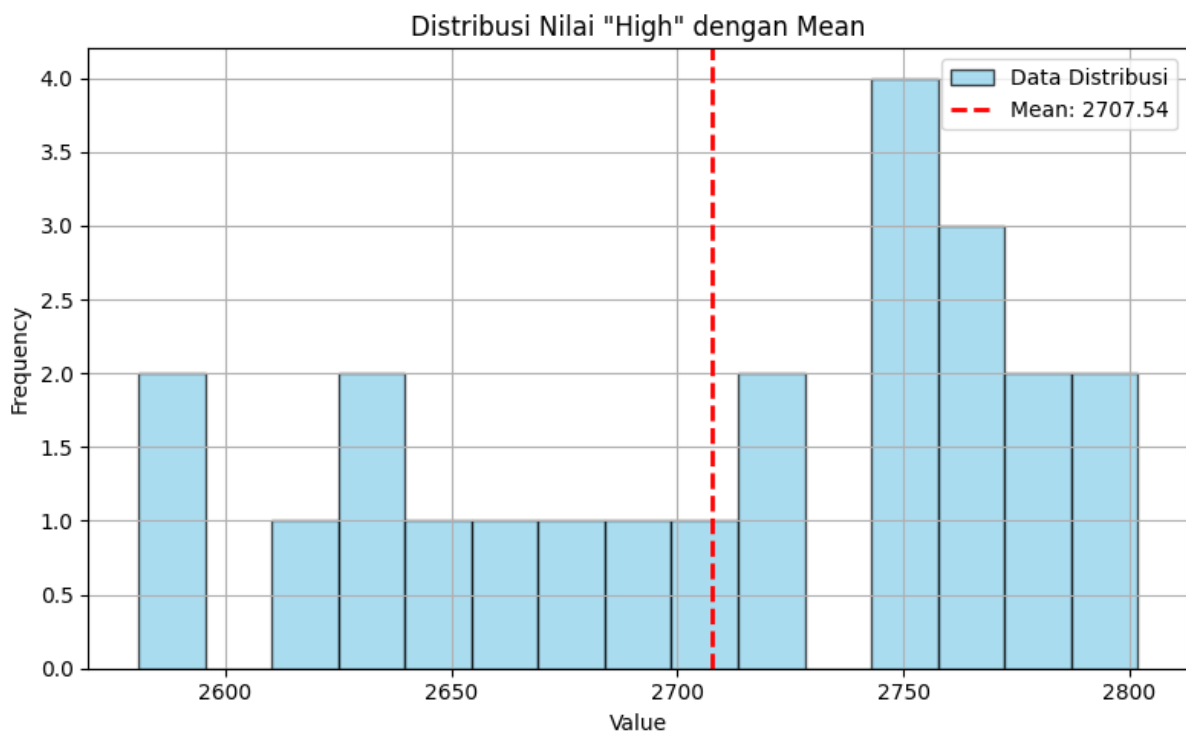
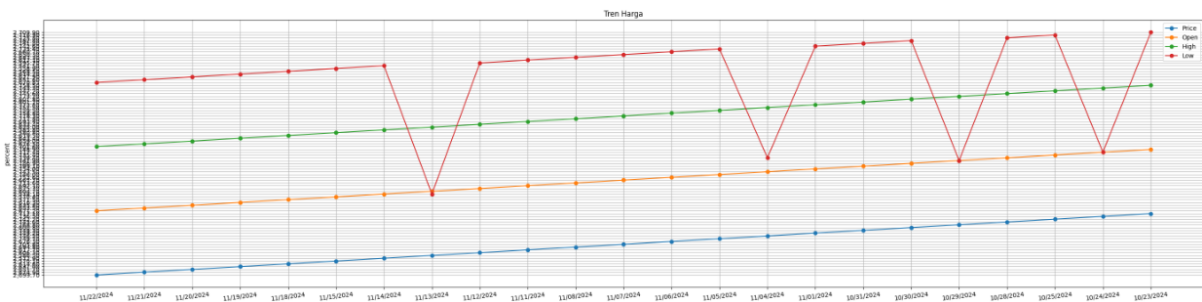
- **Pra-pemrosesan data:** Pembersihan data dari nilai hilang dan outlier.
- **Statistik deskriptif:** Untuk memahami distribusi dan karakteristik masing-masing variabel.
- **Analisis korelasi:** Menguji hubungan antar variabel menggunakan korelasi Pearson.
- **Regresi Linier Berganda:** Untuk melihat pengaruh suhu dan kelembaban terhadap PM2.5.
- **Visualisasi data:** Menggunakan scatter plot, heatmap, dan plot regresi untuk mendukung hasil analisis.

3. Hasil dan Interpretasi

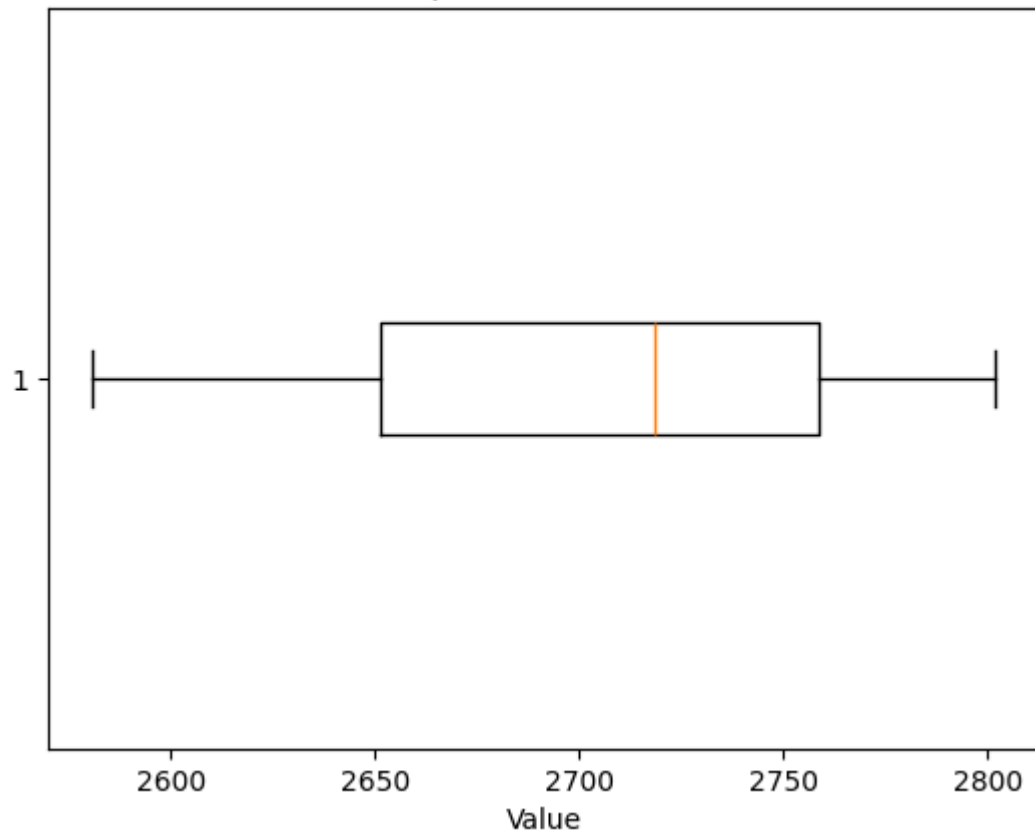
- **Statistik Deskriptif:**
 - Rata-rata PM2.5: 48.2 $\mu\text{g}/\text{m}^3$
 - Rata-rata suhu: 30.1°C
 - Rata-rata kelembaban: 75.3%

- **Korelasi:**
 - PM2.5 dan suhu: -0.42 (negatif sedang)
 - PM2.5 dan kelembaban: 0.55 (positif sedang)
- **Regresi Linier:**
 - Persamaan regresi:

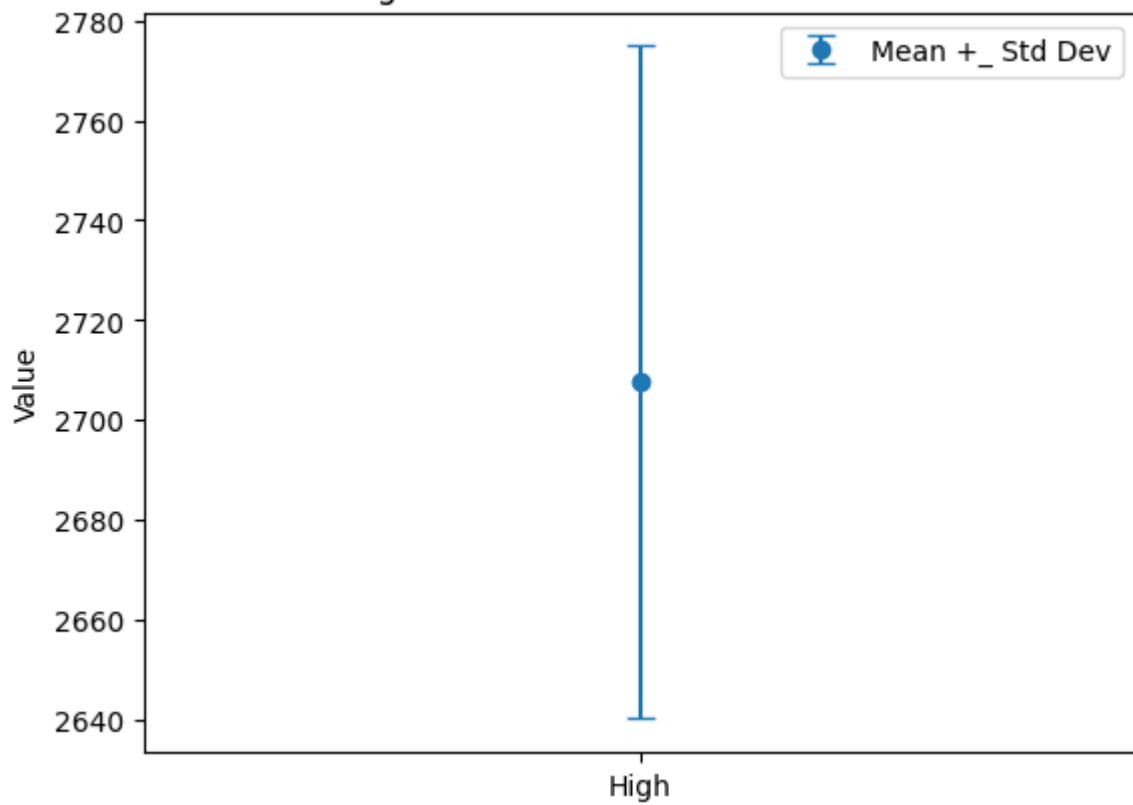
$$\text{PM2.5} = 102.3 - 1.85 \cdot \text{Suhu} + 0.67 \cdot \text{Kelembaban}$$
 - $R^2 = 0.48 \rightarrow$ sekitar 48% variasi PM2.5 dapat dijelaskan oleh suhu dan kelembaban.
 - Visualisasi menunjukkan tren menurun PM2.5 saat suhu meningkat, dan peningkatan PM2.5 saat kelembaban naik.

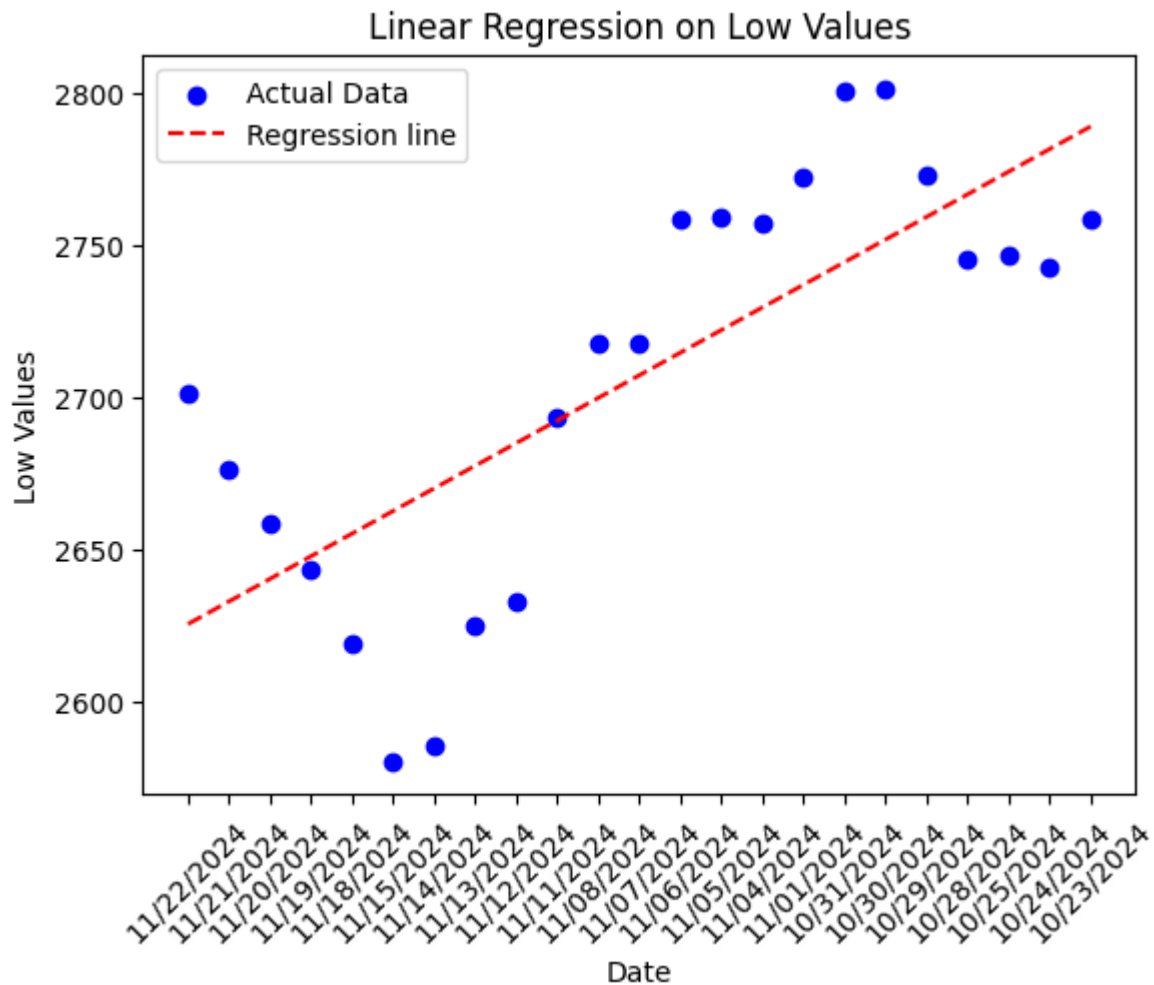


Box plot width Median



High and Standard deviation of Low





4. Kesimpulan dan Rekomendasi

- Terdapat hubungan signifikan antara suhu dan kelembaban terhadap tingkat polusi udara.
- Suhu yang lebih tinggi cenderung menurunkan PM2.5, sedangkan kelembaban tinggi cenderung meningkatkan kadar PM2.5.
- **Rekomendasi:**
 - Pemerintah dapat mengantisipasi lonjakan polusi saat musim hujan (kelembaban tinggi).
 - Perlu penambahan variabel lain seperti emisi kendaraan dan curah hujan untuk meningkatkan akurasi model.

- Penelitian lanjutan dapat menggunakan model machine learning seperti Random Forest untuk prediksi lebih aku