**Bagian 1:** Analisis Dataset

**Laporan Analisis Covid-19 di Jakarta**

**Latar Belakang**

Ini belum...

**Jawaban Soal**

1. Dari dataset yang disediakan, temukan nilai mean, median, dan modus dari positif COVID-19 harian Jakarta.

![Graphical user interface, text, application, email

Description automatically generated]()

**Gambar 1. 1** Mencari Agg pada kolom positif\_harian

Dari *syntax* yang ditampilkan pada Gambar 1.1 di atas. Dapat diketahui bahwa nilai *mean ­­*atau rata-rata kasus positif harian di Jakarta adalah 1356 kasus, nilai *median* untuk kolom positif\_harian adalah *884*, dan *modus* (data yang paling sering muncul) pada kolom positif\_harian adalah 0.

1. Dari dataset yang disediakan, temukan nilai minimal dan maksimal dari positif COVID-19 harian Jakarta.

![Graphical user interface, text, application

Description automatically generated]()

**Gambar 1. 2** Nilai minimal dan maksimal kasus positif harian

Dari syntax yang ditampilkan pada Gambar 1.2 di atas. Nilai maksimal untuk kasus positif harian di Jakarta adalah 14619 kasus, dan nilai minimal atau kasus terendah positif harian adalah 0 (*tidak ada kasus positif harian*).

1. Dari dataset yang disediakan, temukan nilai-nilai outlier yang ada (menggunakan variabel yang kalian tentukan).

Graphical user interface, application

Description automatically generated

**Gambar 1. 3** BoxPlot data numerik

Berdasarkan Gambar 1.3 di atas, dari beberapa kolom yang ditampilkan dalam bentuk *boxplot*, ternyata terdapat beberapa kolom yang memiliki *outlier* diantaranya adalah kolom self\_isolation, masih\_perawatan, positif\_harian, positif\_aktif, dan kolom sembuh\_harian. Oleh karena itu, kami memilih menganalisis *oulier* yang ada di kolom self\_isolation. Grafik distribusi dari kolom self\_isolation ditampilkan pada Gambar 1.4 di bawah ini.

Chart, histogram

Description automatically generated

**Gambar 1. 4** DistPlot kolom self\_isolation

Untuk melihar batas yang dirasa merupakan oulier, maka kami menggunakan IQR untuk menghitungnya batas atas dan batas bawahnya. Rumus mencari nilai IQR dan batas-batasnya adalah sebagai berikut.

(1.1)

(1.2)

(1.3)

Keterangan:

kuartil 1 (25% data)

kuartil 3 (75% data)

Sesuai dengan rumus di atas, kami melakukan perhitungan terhadap kolom self\_isolation.

![Text

Description automatically generated]()

**Gambar 1. 5** Perhitungan IQR dan Batas

Hasil dari perhitungan yang ditampilkan pada Gambar 1.5 di atas adalalah nilai IQR adalah 6870, batas atas adalah 17673, dan batas bawah adalah -5808 atau 0 karena tidak ada data yang negatif. Sedangkan data pada kolom self\_isolation yang sudah dibersikan *outlier* mempunyai rentang data dari 0 sampai 17524. Grafik distibusi kolom self\_isolation setelah dihapus *outlier* dapat dilihat pada Gambar 1.6 di bawah ini.

Chart, histogram

Description automatically generated

**Gambar 1. 6** DistPlot self\_isolation setelah dihapus outlier

1. Dari dataset yang disediakan, usulkan dua buah variabel dan berikan analisis korelasi antara kedua variabel tersebut. Jelaskan apa kesimpulan yang dapat diambil berdasarkan analisis kalian.

Chart, treemap chart

Description automatically generated

**Gambar 1. 7** Heatmap data covid-19 Jakarta

Berdasarkan grafik heatmap yang dapat dilihat pada Gambar 1.7 di atas, diketahui bahwa kolom yang memiliki korelasi paling erat (*berdasarkan data*) adalah kolom *positif\_harian* dan *positif\_aktif*. Oleh karena itu,kami memilih kolom tersebut untuk kami analisis. Grafik untuk kedua kolom dapat dilihat pada gambar di bawah ini.

Chart, scatter chart

Description automatically generated

**Gambar 1. 8** JointPlot kolom positif\_harian dan positif\_aktif

Dari Gambar 1.8 di atas dapat dilihat bahwa pertambahan jumlah positif aktif selaras dengan bertambahnya jumlah positif harian yang ada. Dari grafik dapat diketahui bahwa korelasi dari kedua variabel tersebut mencapai yang berarti memiliki korelasi yang kuat. Kesimpulan yang kami berikan dari analisis grafik di atas adalah kasus positif harian dan positif aktif yang meningkat secara bersamaan. Selama jumlah kasus positif harian yang ada naik, maka kasus positif aktif juga akan naik. Dan juga menandakan pasien yang positif aktif belum banyak yang sembuh.

**Hasil Analisis Tambahan**

1. Problem Statement

Berdasarkan datasets yang ada dan analisis dari data tersebut, permasalaha yang ada adalah:

* Kasus positif aktif yang meningkat lagi dari pada tahun 2021 terutama pada bulan Juni dan Juli.

1. Hypothesis

Ini belum...

1. Exploratory Data Analisis

Untuk melakukan analisis kami melakukan beberapa plotting terhadap data. Gambar di bawah ini adalah gambar grafik perkembangan covid-19 di Jakarta menurut bulan, dan dibagi menurut tahun.

Chart, line chart

Description automatically generated

**Gambar 1. 9** Perkembangan Covid-19 di Jakarta

Berdasarkan grafik pada gambar di atas dapat dikatakan bahwa kasus covid-19 di jakarta mengalami kenaikan jumlah kasus yang sangat tinggi.

1. Initial Findings
2. Deep Dive Analysis
3. Conclusion and Reccomendation

**Kesimpulan**

**Daftar Pustaka**