

## Exploratory Data Analysis (EDA)

Nama : Al Ghifary Akmal Nasheeri

NIM : 1103201242

Kelas : TK-44-06

### Analisis EDA untuk Dataset House Pricing:

1. **Distribusi Harga Rumah:** Pertama-tama, dalam melakukan EDA terhadap dataset House Pricing, penting untuk memeriksa distribusi harga rumah. Distribusi ini membantu untuk memahami apakah distribusi tersebut simetris atau miring. Dari hasil EDA, ditemukan bahwa distribusi harga rumah memiliki kemiringan positif, yang berarti sebagian besar harga rumah berada di kisaran harga yang lebih rendah, sementara ada beberapa rumah dengan harga yang jauh lebih tinggi (pencilan). Hal ini bisa menjadi petunjuk bahwa ada faktor-faktor tertentu yang menyebabkan harga rumah melonjak secara signifikan.
2. **Korelasi Luas Tanah dengan Harga Rumah:** Penting untuk memeriksa korelasi antara luas tanah dan harga rumah. Ini karena biasanya luas tanah memengaruhi harga rumah. Dari analisis, ditemukan bahwa ada korelasi positif yang cukup kuat antara luas tanah dan harga rumah. Artinya, semakin besar luas tanah, semakin tinggi harga rumahnya. Informasi ini bisa berguna bagi pembeli rumah atau investor properti.
3. **Pengaruh Jumlah Kamar Tidur:** Analisis juga dilakukan untuk melihat bagaimana jumlah kamar tidur mempengaruhi harga rumah. Hasilnya menunjukkan bahwa ada variasi yang signifikan dalam harga rumah tergantung pada jumlah kamar tidur. Ini memberikan wawasan bahwa jumlah kamar tidur adalah salah satu faktor penting dalam menentukan harga rumah.
4. **Karakteristik Pemilik Rumah:** Selanjutnya, dataset dianalisis untuk mengidentifikasi karakteristik pemilik rumah. Informasi ini dapat membantu memahami lebih banyak tentang profil pemilik rumah dalam dataset, seperti jenis kelamin atau pekerjaan mereka.

### Analisis EDA untuk Dataset Iris:

1. **Tiga Spesies Bunga Iris:** Dalam dataset Iris, terdapat tiga spesies bunga yang berbeda, yaitu "setosa," "versicolor," dan "virginica." Hal ini adalah hasil utama dari EDA awal. Mengetahui bahwa ada tiga spesies berbeda sangat penting karena masing-masing spesies memiliki karakteristik yang berbeda.
2. **Distribusi Sepal dan Petal:** Analisis dilanjutkan dengan memeriksa distribusi panjang dan lebar sepal serta panjang dan lebar petal untuk masing-masing spesies bunga Iris. Melalui visualisasi, dapat dibandingkan karakteristik morfologis bunga dari tiga spesies

ini. Hasilnya adalah bahwa beberapa spesies memiliki sepal atau petal yang lebih besar dibandingkan dengan spesies lainnya.

3. **Perbedaan Karakteristik Spesies:** Identifikasi perbedaan karakteristik antara spesies bunga Iris dilakukan. Misalnya, ditemukan bahwa spesies "setosa" cenderung memiliki petal yang lebih pendek dibandingkan dengan spesies "versicolor" atau "virginica." Ini adalah wawasan penting bagi para ahli botani atau peneliti yang ingin mengidentifikasi spesies bunga berdasarkan morfologi.

Kesimpulannya, analisis EDA dari kedua dataset tersebut memberikan pemahaman yang lebih dalam tentang data yang digunakan. EDA membantu mengidentifikasi wawasan, pola, dan perbedaan penting dalam data, yang pada gilirannya dapat digunakan untuk pengambilan keputusan atau penelitian lebih lanjut.