

# Data Visualisasi

Week 9 - Advanced Chart : Correlation

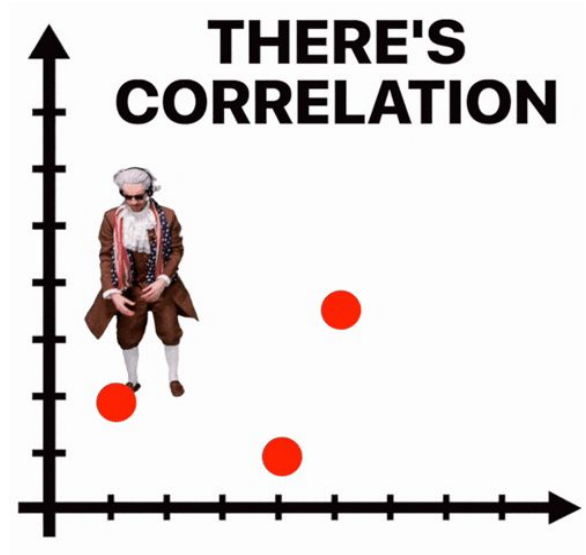
# What is Correlation?

Korelasi adalah teknik statistik yang digunakan untuk mengukur dan mendeskripsikan hubungan (seberapa kuat dan ke arah mana) antara dua variabel kuantitatif (angka).

*Jika nilai variabel A berubah (naik/turun), apakah nilai variabel B juga ikut berubah?*

Digunakan untuk :

- Analisis pola
- Prediksi awal
- Insight bisnis



# Why Do We Look for Correlations?

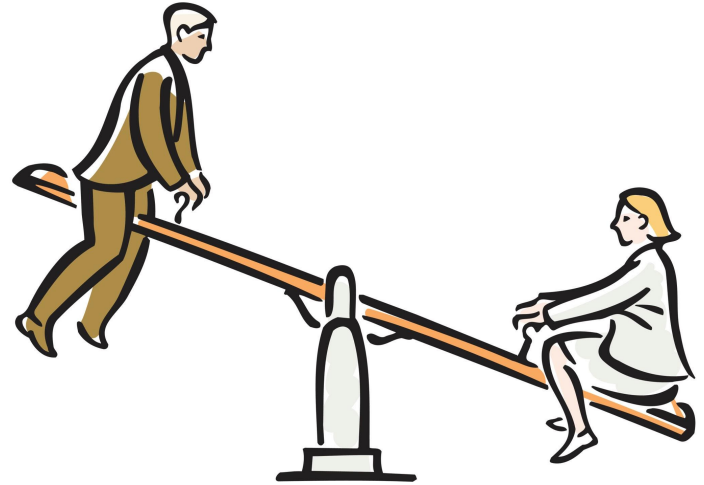
Analisis korelasi bukan sekadar melihat angka, tetapi:

- ❑ Mencari pola tersembunyi dalam data
- ❑ Membantu pengambilan keputusan berbasis data
- ❑ Menjadi langkah awal sebelum:
  - Prediksi
  - Modeling
  - Eksperimen



*“Data tidak berbicara langsung, kita yang harus membaca polanya.”*

Apakah 'A' berhubungan dengan 'B'?



# Correlation Does not Imply Causation

Hanya karena dua variabel terlihat berhubungan, bukan berarti yang satu menyebabkan yang lain.

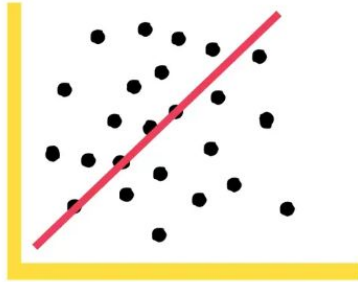


## Poin Utama

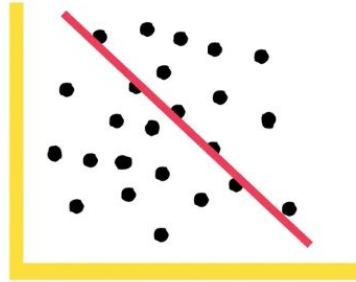
- Korelasi Semu
- Ada variabel tersembunyi (confounding variable)
- Data tidak menangkap seluruh konteks

Perlu analisis lebih lanjut, penjelasan logis atau konteks bisnis

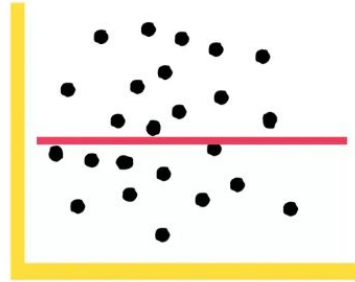
# Where are These Dots Moving?



Nilai positif



Nilai negatif



Nilai nol

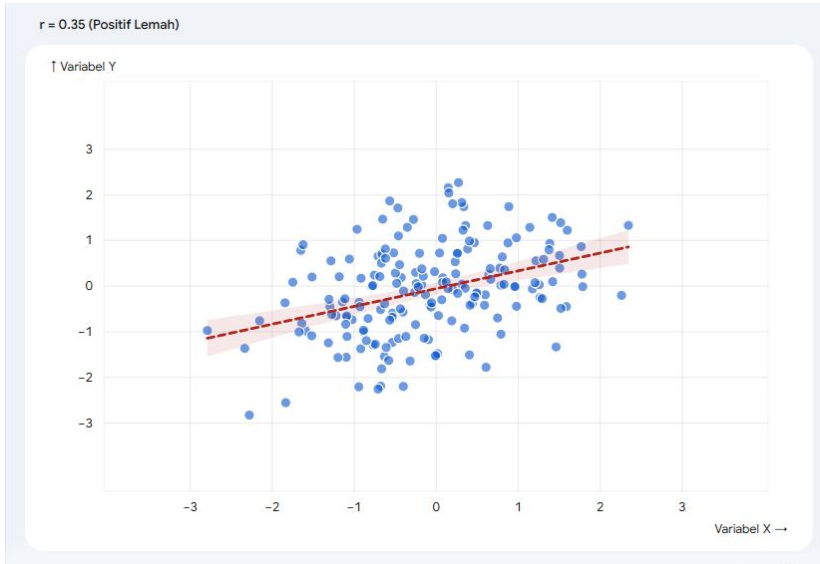
Sumber :  
<https://www.revou.co/kosakata/korelasi>

3 Nilai korelasi :

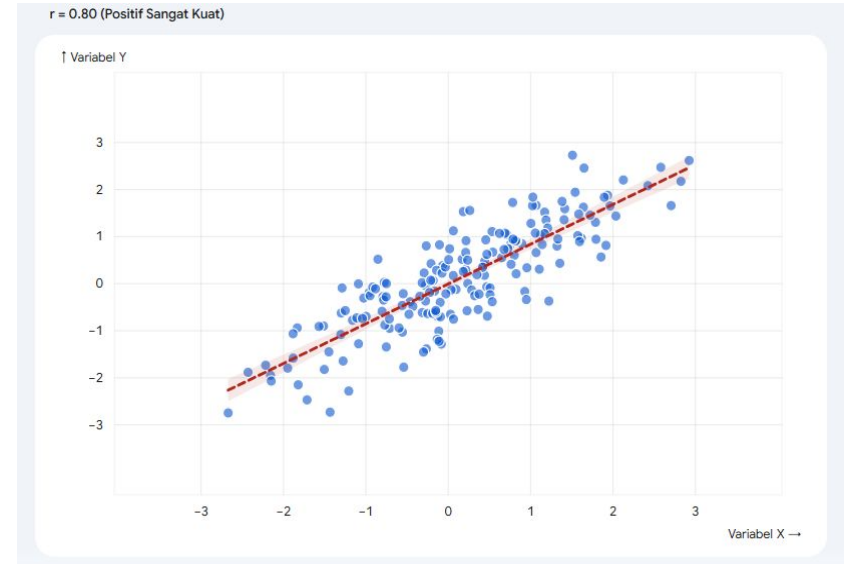
- ❑ Korelasi Positif: Saat X naik, Y ikut naik.
- ❑ Korelasi Negatif: Saat X naik, Y malah turun.
- ❑ Tidak Ada Korelasi : Titik menyebar tanpa pola yang jelas.

# How Close the Dots are?

Cek kerapatan titik



Korelasi Lemah

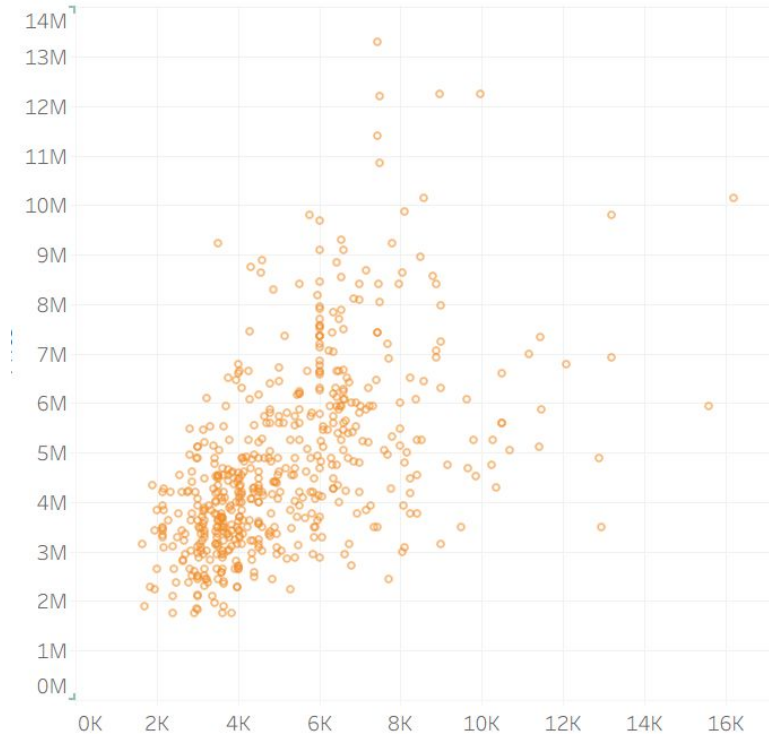


Korelasi Kuat

# Scatter Plot

Scatter Plot adalah visualisasi wajib untuk analisis korelasi

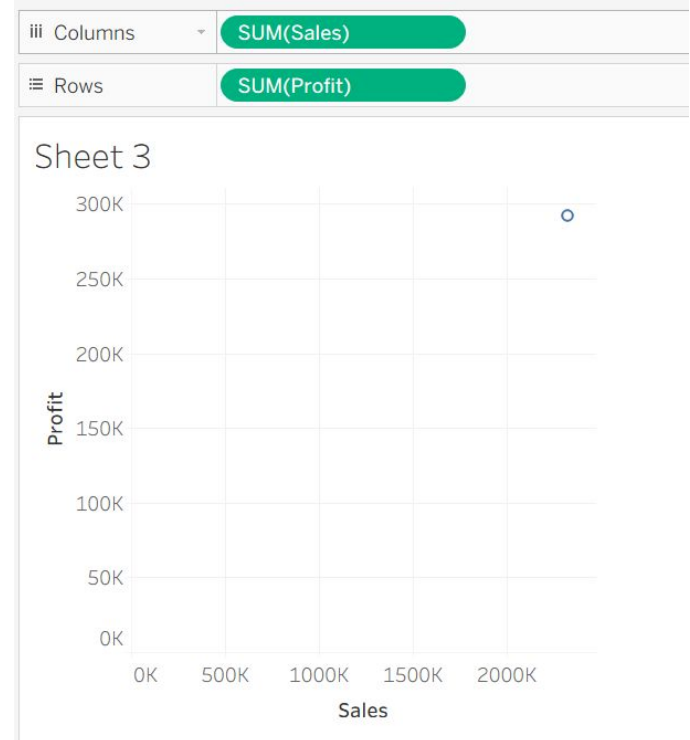
Memerlukan 2 variabel numerik untuk sumbu X dan sumbu Y



# Scatter Plot in Tableau

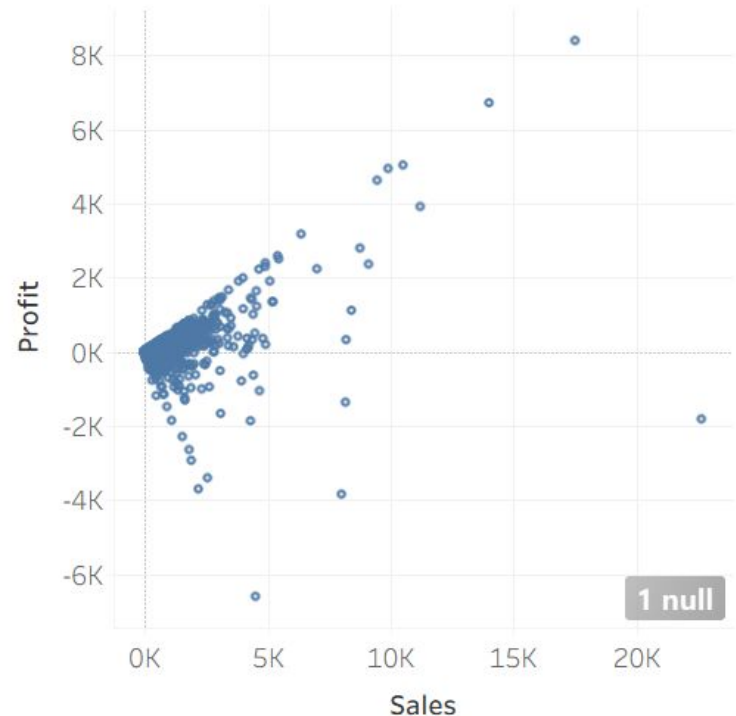
- ❑ Di tableau, untuk membuat scatter plot memerlukan 2 measure di columns dan rows.
- ❑ Menggunakan dataset Sample - Superstore.xls, letakkan sales dan profit pada masing-masing columns dan rows

Mengapa titiknya hanya satu yang muncul?



# Memecah Agregasi

- ❑ Di tableau, satu titik yang muncul karena agregasi data (diringkas). Dimana yang muncul adalah kalkulasi atau total dari seluruh data
- ❑ Melihat seluruh data -> menambahkan dimension ke Detail atau
- ❑ Uncheck Aggreate Measures



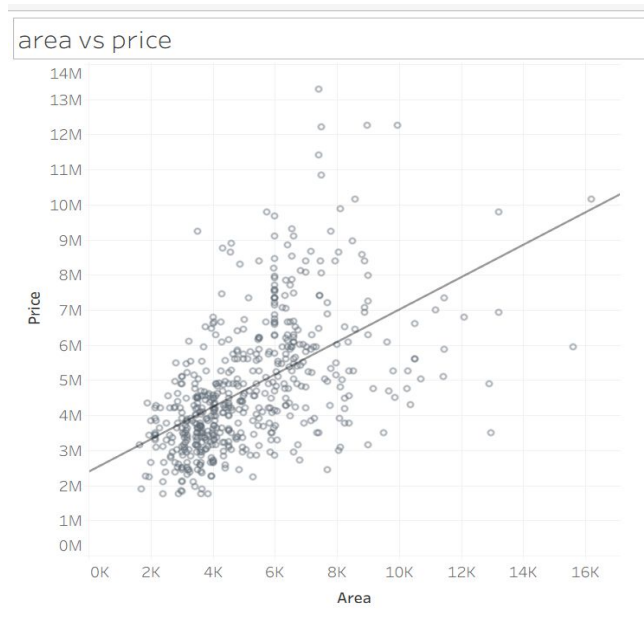
# Instruksi

## Uncovering the Secret of Home Prices

- 1) Gunakan dataset Sample dari <https://www.kaggle.com/datasets/yasserh/housing-prices-dataset>
- 2) Buat worksheet baru bernama: "Luas vs Harga"
- 3) Buat perbandingan antara column Area dengan Price
- 4) Pada menu Analysis, hilangkan tanda centang (uncheck) opsi Aggregate Measures. Amati hasilnya.
- 5) Buka tab Analytics di panel kiri, tarik Trend Line ke kanvas (kemudian pilih Linear). Amati hasilnya
- 6) Kemudian coba tambahkan column lain (Furnishingstatus atau Airconditioning) ke kotak Color di *Marks Card*. Amati hasilnya

# Instruksi

Contoh tampilan



# Instruksi

## **Pertanyaan**

- 1) Berdasarkan kemiringan garis tren (sebelum dipisah warnanya), apa jenis korelasi antara Area dan Price? Apakah semakin luas rumah, harganya cenderung naik, turun, atau tidak jelas?
- 2) Arahkan cursor (hover) ke garis tren Anda dan lihat nilai R-Squared. Apakah titik-titiknya menempel erat pada garis, atau sangat menyebar? Apakah aman jika agen sales kita memprediksi harga rumah hanya bermodalkan luas tanahnya saja?
- 3) Coba masukkan variabel Airconditioning ke dalam warna, maka akan ada dua garis tren (Rumah dengan AC vs Tanpa AC). Tarik kesimpulan dari dua garis tren tersebut! Mana yang kemiringannya lebih tajam, dan apa arti bisnis dari perbedaan tersebut jika agen kita ingin merenovasi rumah sebelum dijual?

# Tugas

- Buat laporan yang berisi link tableau public, screenshot hasil dan jawaban hasil investigasi
- Kemudian laporan di upload di ILMU2