

# REGRESI LINEAR UNTUK DATA SALES & MARKETING

Link notebook [disini](#).

## Latar Belakang

Perusahaan ingin mengetahui promosi apa yang paling harus difokuskan dan Influencer mana yang paling menghasilkan sales paling banyak. Maka dari itu kita akan mencoba melakukan uji hipotesis dan juga membuat model regresi linear untuk mengetahui bagaimana hubungan promosi TV, Radio, Media Sosial dan Influencer terhadap Sales.

## Objektif

1. Bagaimana analisis eksplorasi datanya?
  - Apa yang ditemukan dari hasil eksplorasi dan bagaimana uji hipotesisnya?
2. Bagaimana hasil dari pembuatan model regresi-nya?

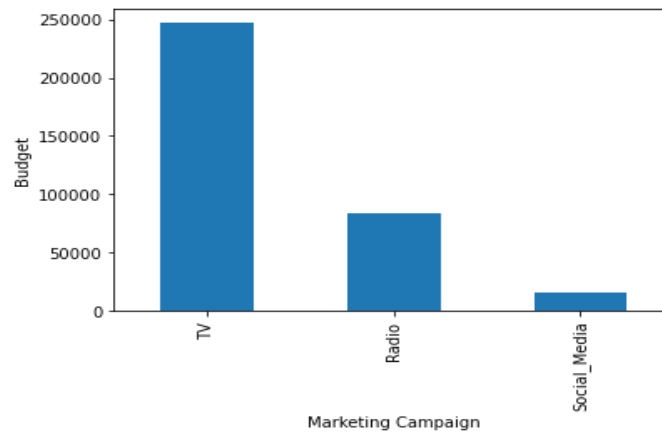
## Goal

1. Mengetahui promotion channel apa yang paling mempengaruhi sales dan bagaimana pengaruhnya

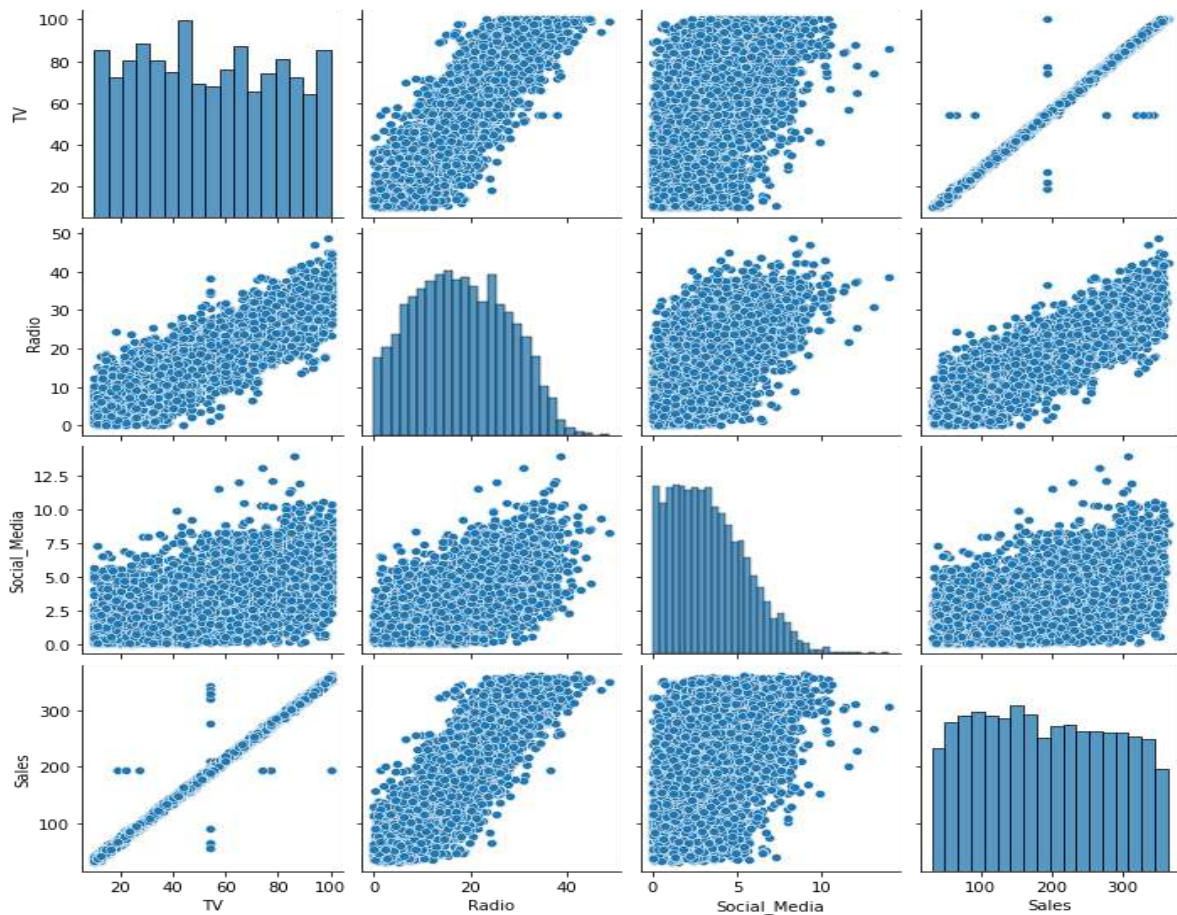
## Data Contains

- TV promotion budget (in million)
- Social Media promotion budget (in million)
- Radio promotion budget (in million)
- Influencer: Whether the promotion collaborate with Mega, Macro, Nano, Micro influencer
- Sales (in million)

## Data Eksplorasi

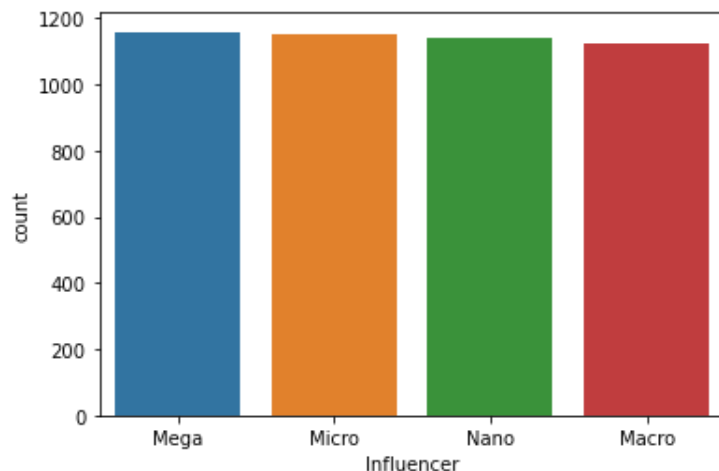


Dapat dilihat bahwa budget yang dikeluarkan oleh perusahaan untuk promosi TV itu sangat besar melebihi dari promosi TV dan Sosial Media. Untuk mengetahui bagaimana hubungan antara setiap channel promosi dengan Sales kita perlu melihat grafik korelasi.

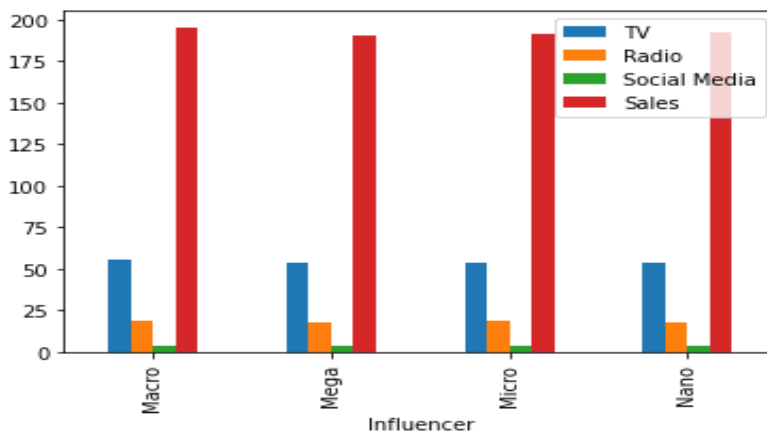


Berdasarkan plot TV dan Sales dapat ditarik kesimpulan bahwa Sales akan terus meningkat seiring dengan meningkatnya anggaran promosi TV. Hal ini menunjukkan bahwa terdapat korelasi yang positif antara Sales dan marketing campaign di TV. Radio terlihat memiliki korelasi yang positif juga namun tidak sebaik dengan TV. Campaign marketing Social Media membutuhkan improvement. Hal ini bisa dilihat dari plotnya terlalu menyebar yang menunjukkan rendahnya korelasi dengan Sales.

Perusahaan memiliki 4 Influencer untuk menjalankan promosinya yaitu Micro, Macro, Nano dan Mega.



Dapat dilihat dari grafik diatas bahwa setiap influencer telah melakukan project marketing yang hampir sama jumlah nya. Yang paling banyak adalah influencer Mega yaitu sebanyak 1158 kali dan yang paling sedikit adalah influencer Macro yaitu sebanyak 1122 kali.



Setiap Influencer terlihat memiliki nilai yang mirip, tetapi Influencer Macro menunjukkan rata-rata sales tertinggi dari influencer lain yaitu sebesar 195.610796. Maka akan dilakukan uji hipotesis untuk membuktikan pernyataan tersebut.

## Uji Hipotesis

Uji hipotesis ini dilakukan untuk menguji apakah inflencer Macro memiliki rata-rata sales yang lebih besar dari inflencer Micro, Nano dan Mega. Metode yang digunakan adalah *Two Sample t-test for Mean*. Berikut hipotesisnya :

$H_0$ : Inflencer Macro dan Micro/Nano/Mega sama-sama menghasilkan rata-rata sales yang sama

$$H_0 : \mu_A = \mu_B$$

$H_1$ : Inflencer Macro memiliki rata-rata sales yang lebih besar dari Inflencer Micro/Nano/Mega

$$H_1 : \mu_A > \mu_B$$

Significance level = 0.05

1. Perbandingan Macro dan Micro

t stats : 0.9706 dan p-value : 0.1659

Nilai p-value > significance level. hal ini menunjukkan bahwa kita gagal menolak  $H_0$ .

2. Perbandingan Macro dan Nano

t stats : 0.9405 dan p-value : 0.1735

Nilai p-value > significance level. hal ini menunjukkan bahwa kita gagal menolak  $H_0$ .

3. Perbandingan Macro dan Mega

t stats : 1.2956 dan p-value : 0.0976

Nilai p-value > significance level. hal ini menunjukkan bahwa kita gagal menolak  $H_0$ .

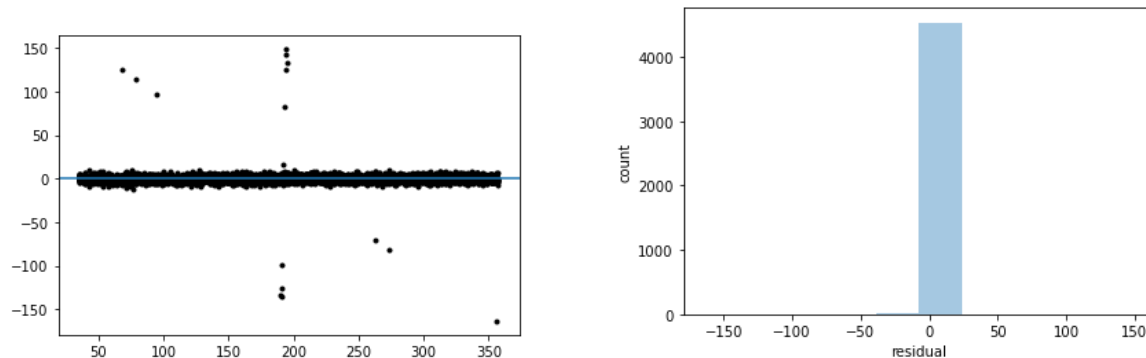
Berdasarkan Uji Hipotesis diatas, Kita gagal menolak hipotesis nol sebanyak 3 kali. Hal ini menunjukkan bahwa tidak cukup evidence untuk membuktikan bahwa inflencer Macro memiliki rata-rata sales yang lebih besar dari inflencer Micro, Nano dan Mega. Dapat disimpulkan bahwa setiap Inflencer menghasilkan sales yang sama.

## Pemodelan

Setelah melakukan pemodelan di Jupyter Notebook. Kita mendapatkan persamaan linear sebagai berikut:

$$\text{Sales} = 0.1816 + 3.5124 \text{ TV} + 0.1311 \text{ Radio}$$

Dengan nilai R-squared sebesar 99.3% yang mana nilai ini adalah nilai yang baik untuk model regresi tersebut. Lalu, sebelumnya kita sudah menambahkan variabel Social Media, Influencer Micro, Influencer Macro, Influencer Nano dan Influencer Mega sebelumnya. namun setelah dilakukan F-test dan t-test kami terpaksa harus mengeluarkannya dari model dikarenakan tidak signifikan secara statistik. Pada uji Asumsi, model ini telah lolos asumsi *independence of error* dan *distribution of the error*. Berikut grafiknya:



## Rekomendasi

1. Perusahaan bisa fokus meningkatkan sales dengan melakukan promosi pada campaign TV saja. Karena semakin besar budget spent terhadap promosi TV maka semakin besar juga tingkat penjualannya. Radio menjadi medium promosi yang baik juga setelah TV.
2. Socia Media menjadi medium yang paling banyak digunakan di era digital sekarang. Maka dari itu perusahaan harus mengevaluasi lebih dinamika sosmed kita.
3. Influencer tidak memiliki pengaruh yang signifikan terhadap sales. Perusahaan bisa lebih selektif dalam pemilihan influencer agar tetap relevan.