

ANALISIS PREDIKTIF UNTUK MENGOPTIMALKAN ALOKASI BIAYA IKLAN

Rizky Apriansyah Eska Pratama

Data Science

Batch 23B



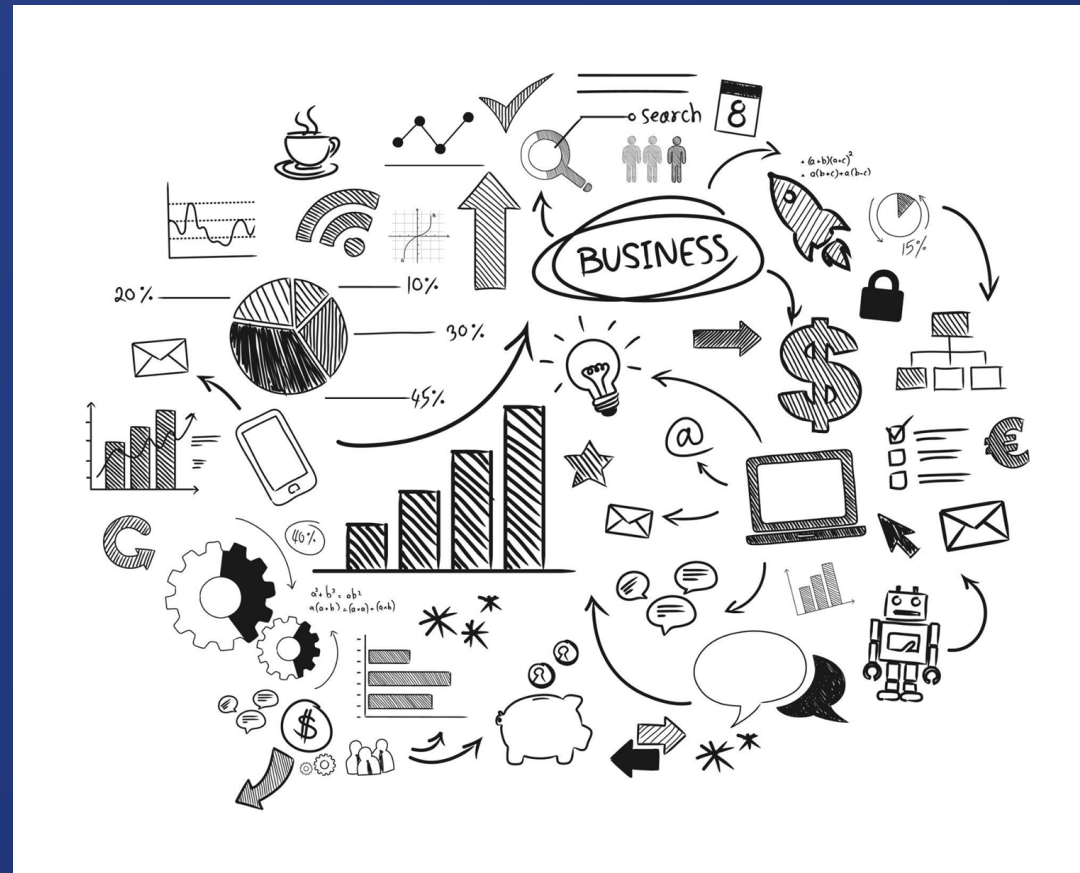
Outline :

- Background
- Business Question
- Preliminary look & Data Cleansing
- Data Understanding
- Exploratory Data Analysis
- Machine Learning Model
- EDA Questions
- Conclusion

Background

Dalam lingkungan bisnis yang semakin kompetitif, sebuah perusahaan akan menghadapi tekanan untuk memaksimalkan efisiensi sumber daya mereka, termasuk alokasi biaya pemasaran. Penentuan alokasi biaya yang efektif menjadi kunci dalam mencapai hasil pemasaran yang optimal.

Di tengah tantangan bisnis dan preferensi pelanggan yang terus berubah, banyak perusahaan mulai mengadopsi analisis prediktif sebagai alat untuk membantu dalam pengambilan keputusan pemasaran.



Business Question

Apakah alokasi biaya iklan sudah efektif?

Bagaimana cara mengoptimalkan alokasi biaya iklan agar mendapatkan penjualan yang maksimal?

Preliminary Look & Data Cleansing

	TV	Billboards	Google_Ads	Social_Media	Influencer_Marketing	Affiliate_Marketing	Product_Sold
0	281.42	538.80	123.94	349.30	242.77	910.10	7164.00
1	702.97	296.53	558.13	180.55	781.06	132.43	5055.00
2	313.14	295.94	642.96	505.71	438.91	464.23	6154.00
3	898.52	61.27	548.73	240.93	278.96	432.27	5480.00
4	766.52	550.72	651.91	666.33	396.33	841.93	9669.00

Keterangan:

- Dataset : Advertising Product Data
- terdapat 7 kolom : 6 kolom fitur iklan dan 1 kolom penjualan
- Semua kolom berisi numeric

tidak ada missing value & duplicate data !



Data Understanding

```
df.describe()
```

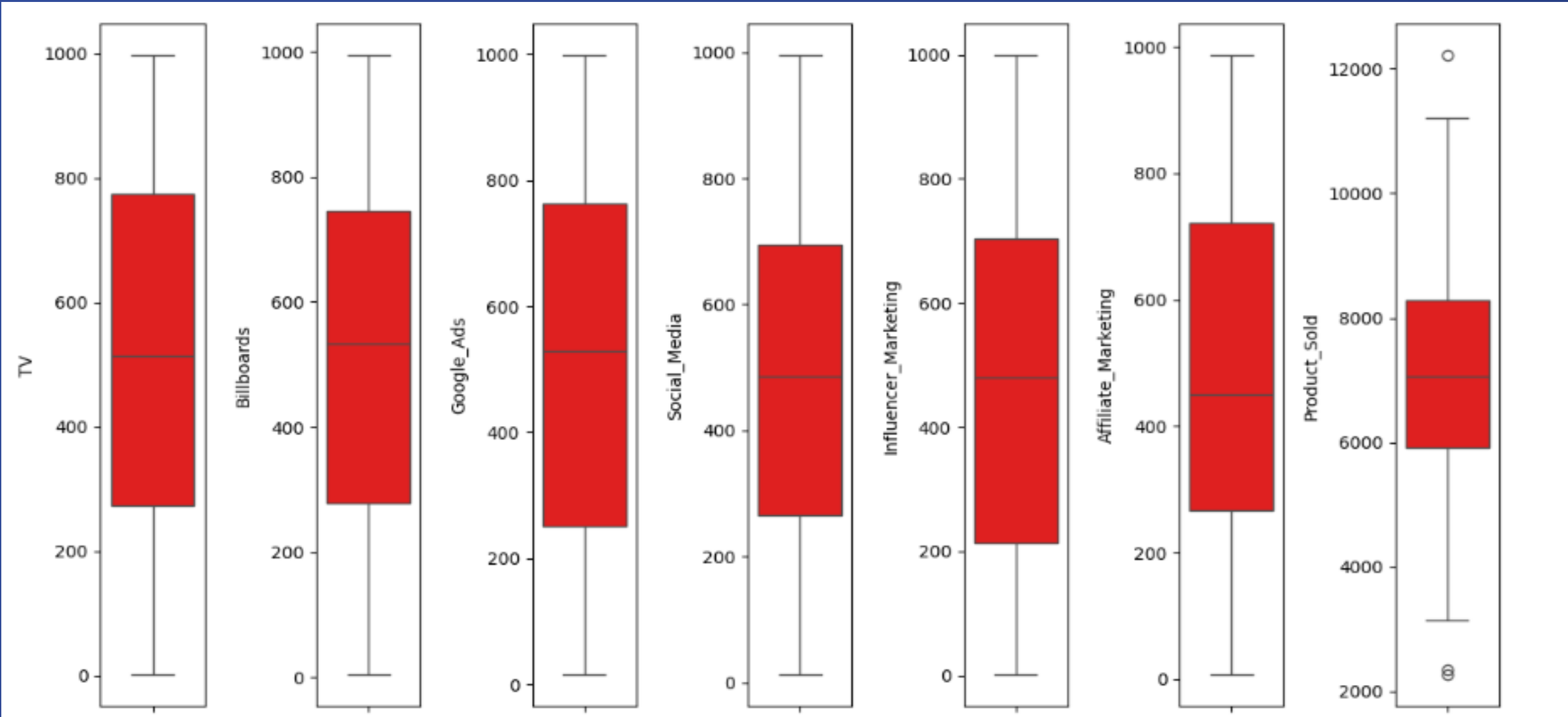
	TV	Billboards	Google_Ads	Social_Media	Influencer_Marketing	Affiliate_Marketing	Product_Sold
count	300.000000	300.000000	300.000000	300.000000	300.000000	300.000000	300.000000
mean	517.431000	502.644933	512.444133	489.800100	465.732567	484.430633	7031.523333
std	288.114792	275.842369	285.422376	273.883915	288.314080	277.955458	1703.614951
min	1.040000	3.630000	14.860000	11.690000	0.770000	6.740000	2259.000000
25%	273.090000	277.912500	250.530000	265.612500	214.480000	267.677500	5922.500000
50%	513.970000	533.020000	528.965000	486.385000	480.355000	451.315000	7051.000000
75%	774.275000	745.005000	763.345000	695.117500	703.092500	721.065000	8278.000000
max	998.100000	995.320000	999.230000	996.160000	999.830000	987.580000	12227.000000

Keterangan:

- Minimum dan maksimum values terlihat normal
- Product_sold berisi penjualan dan akan dijadikan target untuk modeling
- Tv, Billboards, Google_Ads, Social_media, Influencer_marketing, Affiliate_Marketing berisi biaya iklan yang akan menjadi features di modeling



Exploratory Data Analyst



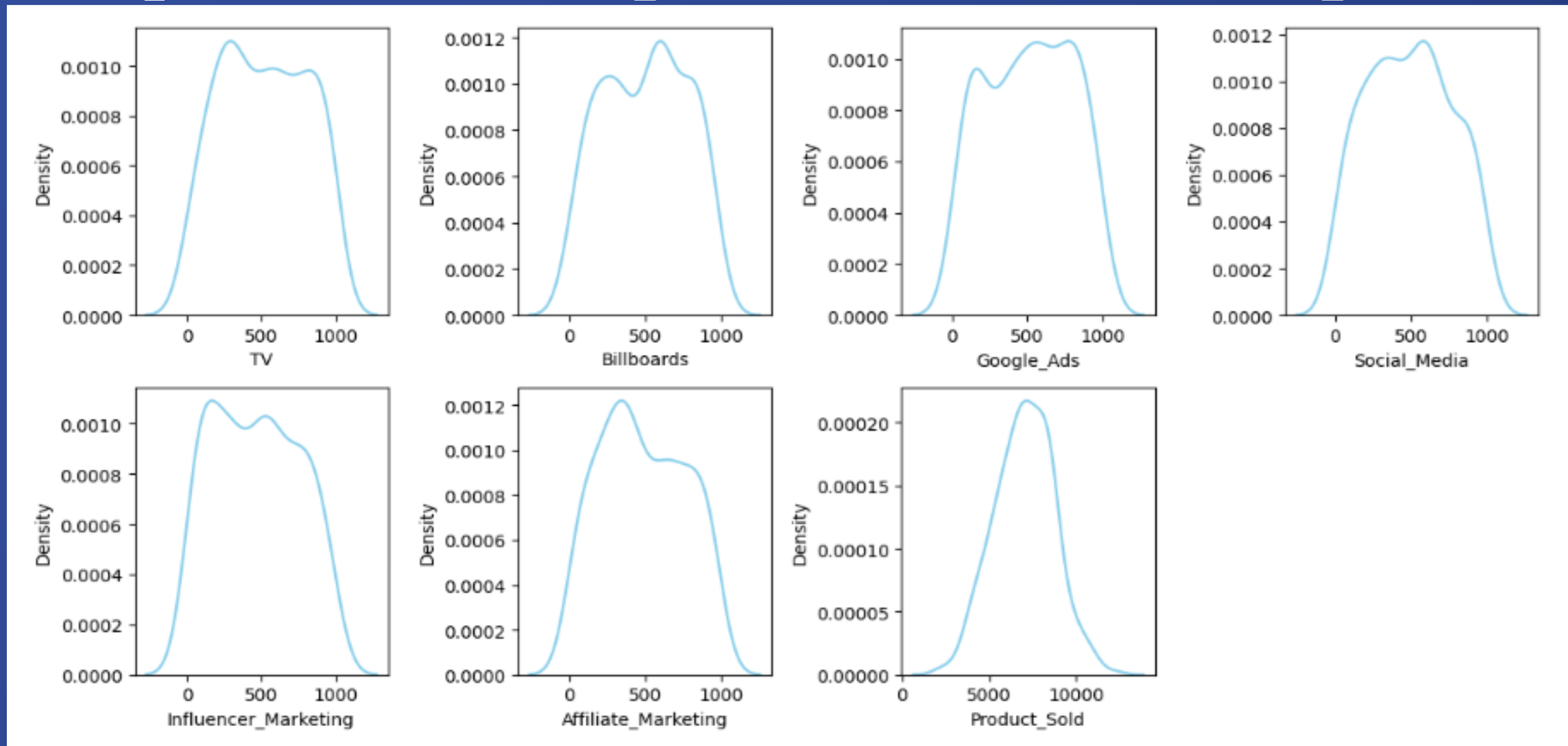
	TV	Billboards	Google_Ads	Social_Media	Influencer_Marketing	Affiliate_Marketing	Product_Sold
11	787.33	947.73	884.15	992.30	871.55	742.36	12227.0
145	1.04	19.68	900.71	55.67	379.03	63.10	2259.0
281	480.67	135.75	82.66	217.72	217.90	15.41	2353.0

Keterangan:

- terdeteksi ada outlier di product_sold, namun tidak perlu di drop



Exploratory Data Analyst

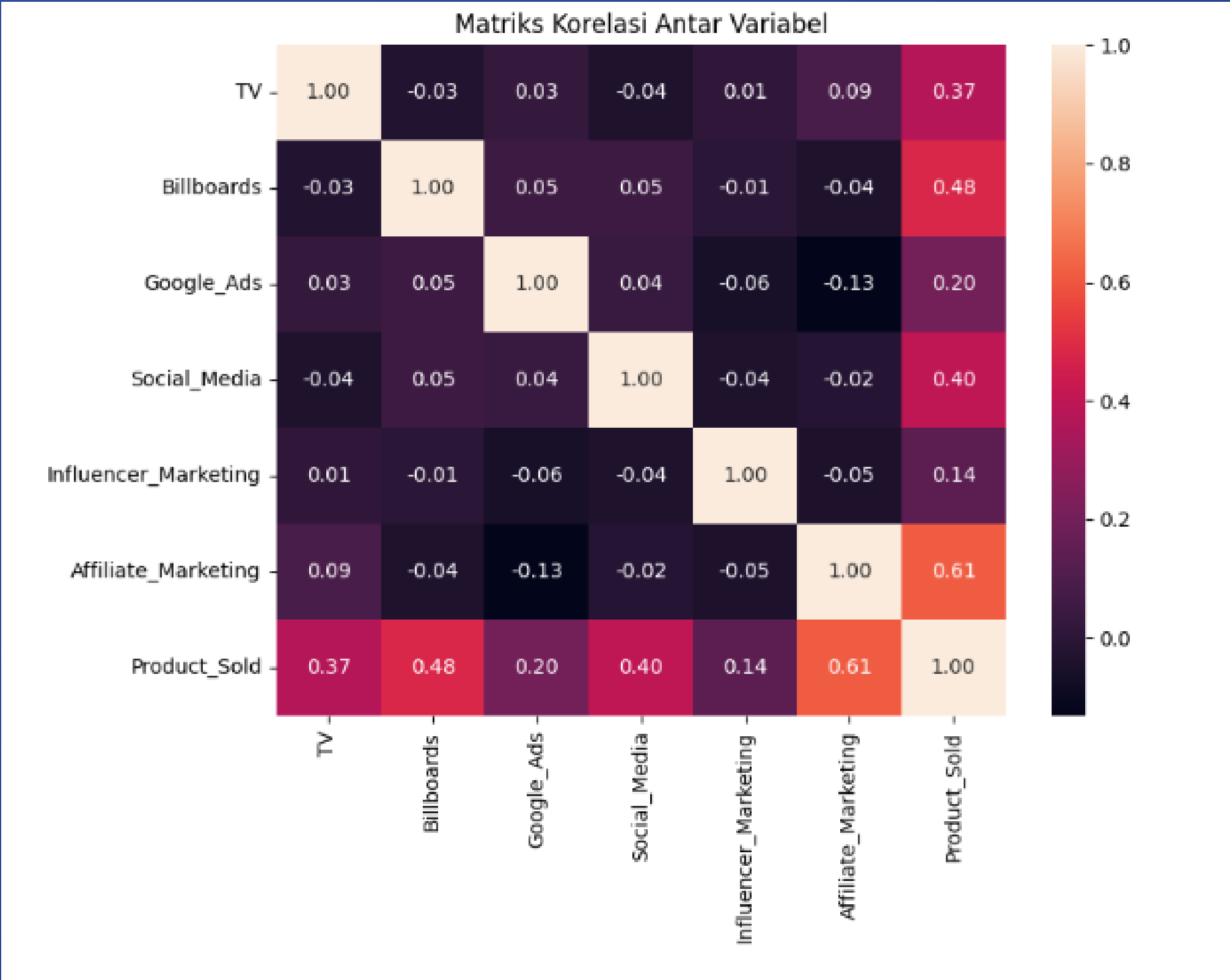


Keterangan:

- Tv, Influencer_marketing, Affiliate_marketing condong kekiri, mengartikan biaya yang dikeluarkan relatif lebih rendah
- Google_Ads condong ke kanan, mengartikan biaya yang dikeluarkan relatif lebih tinggi
- Billboards, Social_media, Product_sold terlihat normal



Exploratory Data Analyst



Keterangan:

- Product_sold memiliki korelasi paling tinggi dengan Affiliate_Marketing
- Product_sold memiliki korelasi paling rendah dengan influencer_Marketing



Machine Learning Model

SPLIT DATA

	TV	Billboards	Google_Ads	Social_Media	Influencer_Marketing	Affiliate_Marketing	Product_Sold
0	281.42	538.80	123.94	349.30	242.77	910.10	7164.00
1	702.97	296.53	558.13	180.55	781.06	132.43	5055.00
2	313.14	295.94	642.96	505.71	438.91	464.23	6154.00
3	898.52	61.27	548.73	240.93	278.96	432.27	5480.00
4	766.52	550.72	651.91	666.33	396.33	841.93	9669.00

Features

Target

data dibagi menjadi 2:

training data n = 240 -> X_train, y_train

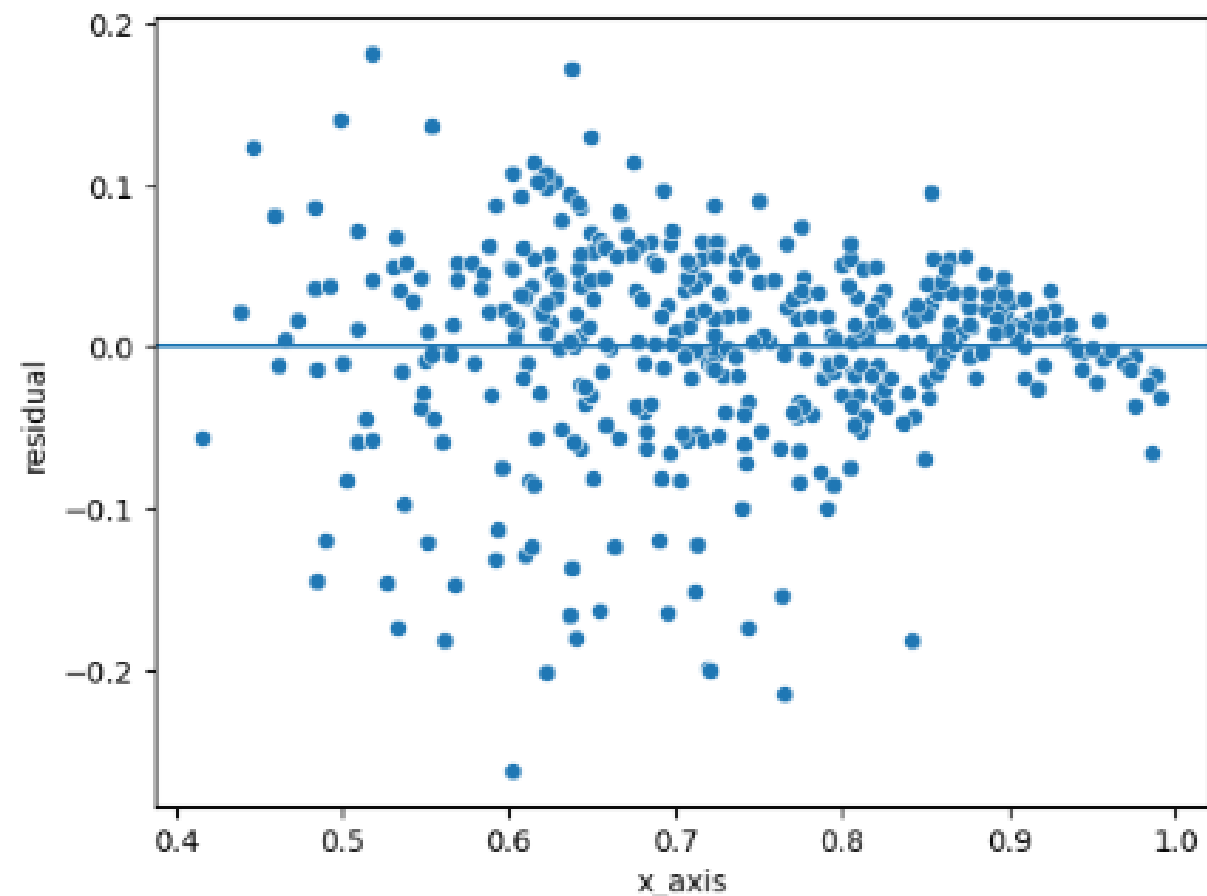
test data n = 60 -> X_test, y_test

Machine Learning Model

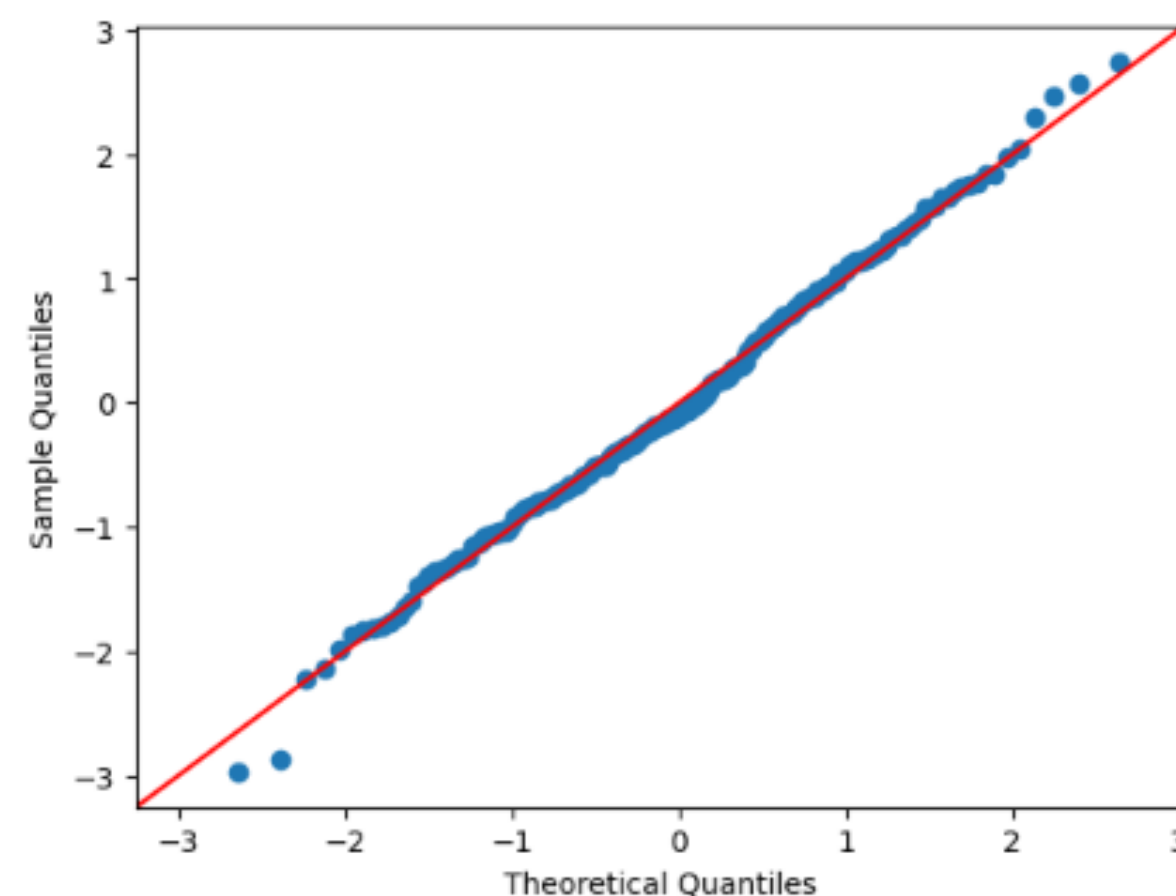
MODEL TRAINING

LinearRegression() --> **model yang digunakan LinearRegression**

MODEL DIAGNOSTIC



residul memiliki variance yang konstan



residual terdistribusi secara normal

Machine Learning Model

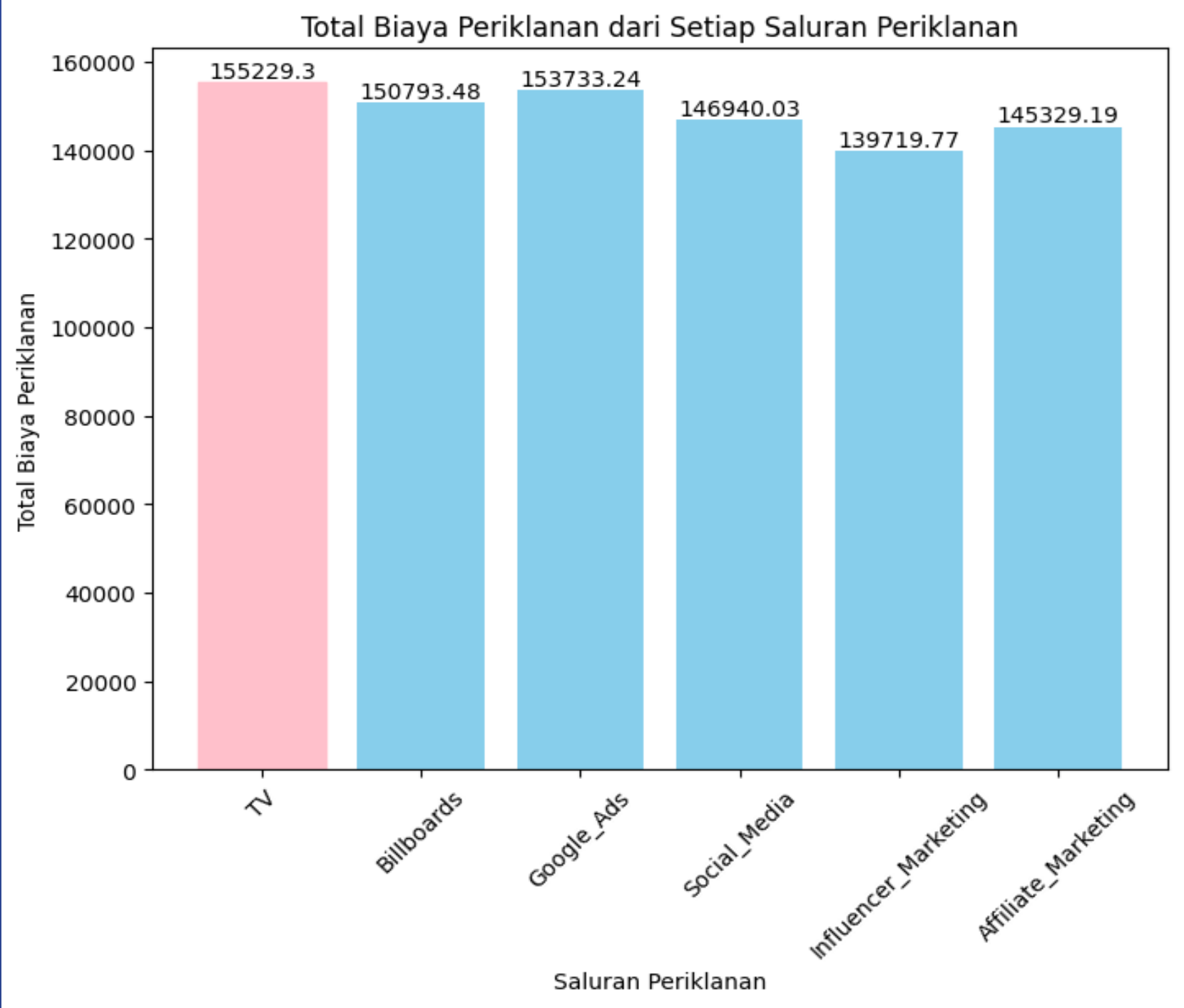
EVALUASI

	Metrics	Values
0	Mean Squared Error	75.37
1	Mean Absolute Error	7.09
2	Root Mean Squared Error	8.68
3	Relative Mean Error	0.10

**evaluasi menunjukkan nilai cukup rendah,
mengartikan model berkerja dengan baik**

Exploratory Data Analyst Questions

Fitur apa dengan jumlah biaya paling besar?

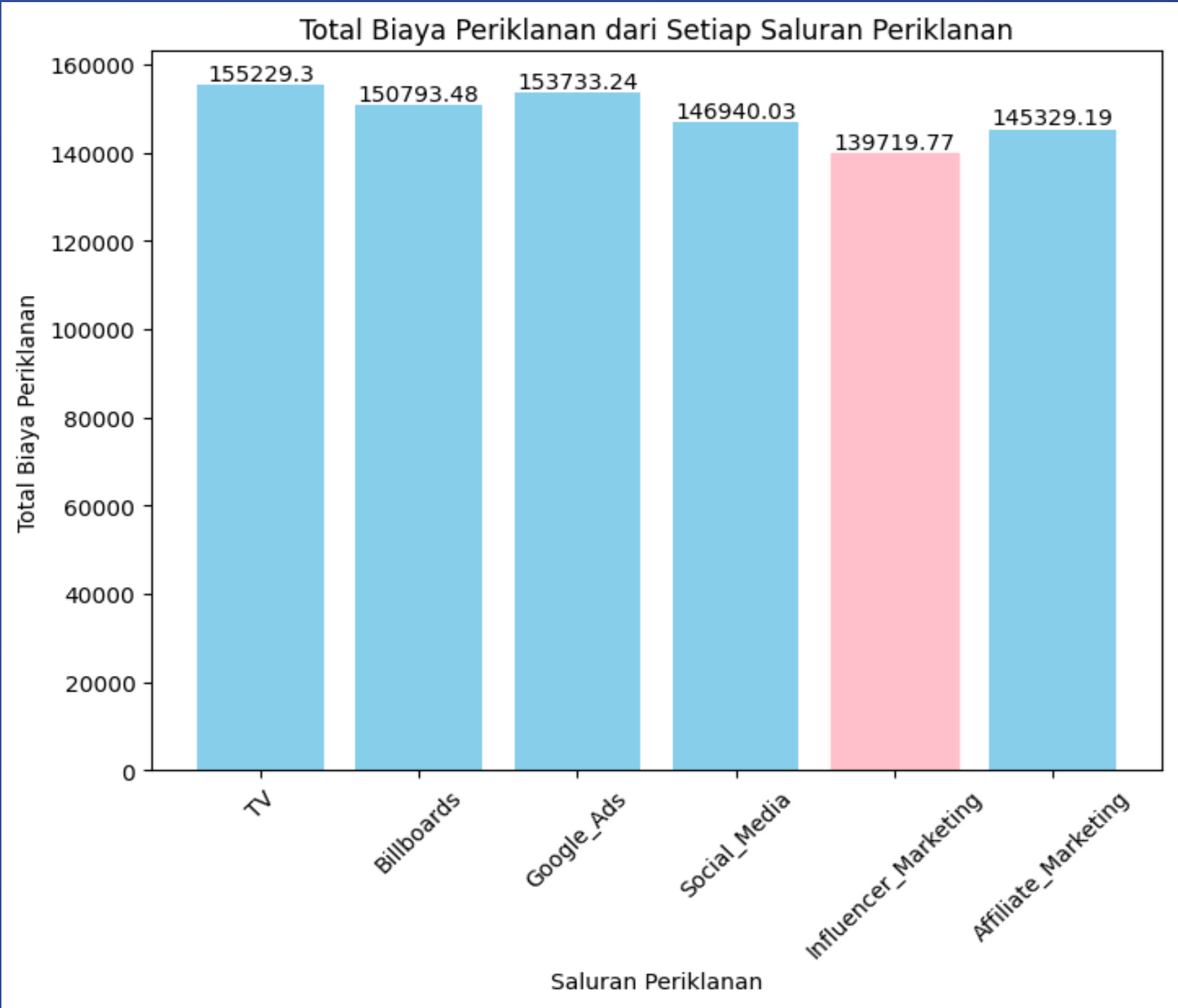


Fitur dengan total biaya paling besar adalah TV sebesar 155229.3



Exploratory Data Analyst Questions

Fitur apa dengan jumlah biaya paling kecil?

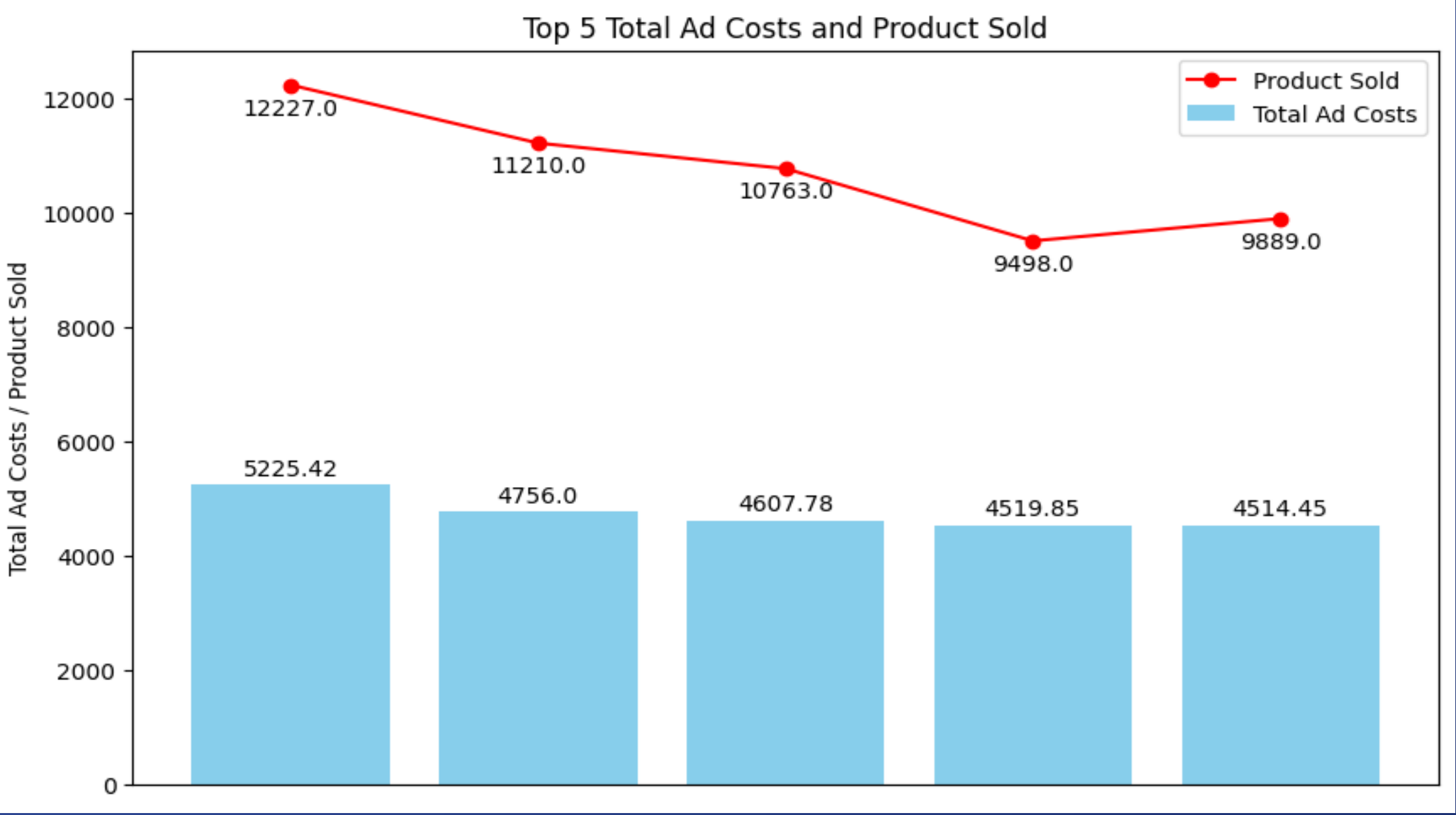
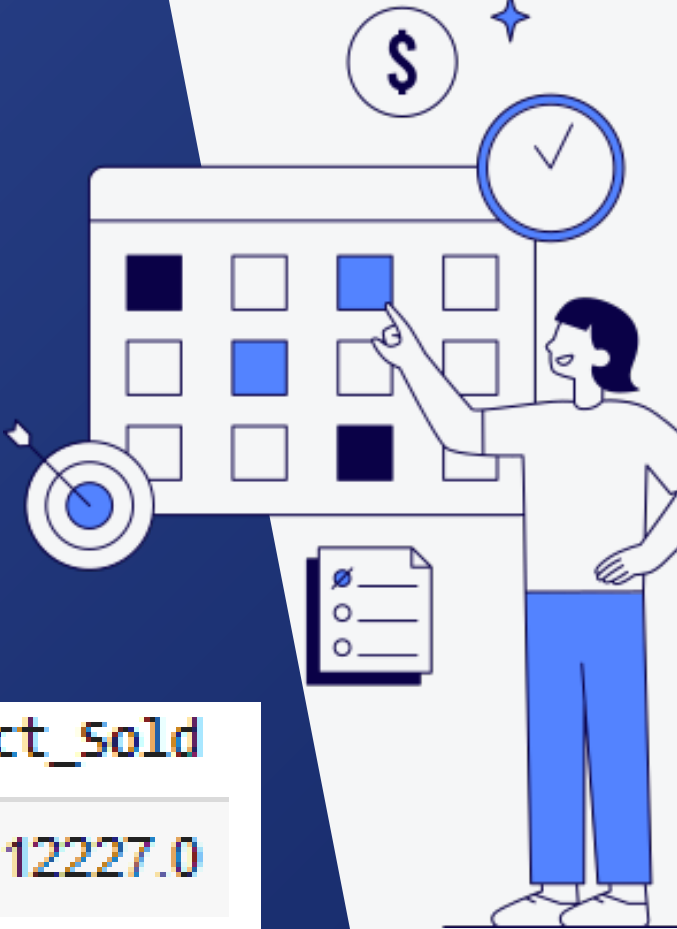


Fitur dengan total biaya paling kecil adalah **Influencer_marketing** sebesar **139719.77**



Exploratory Data Analyst Questions

Top 5 total biaya ads dan berapa penjualan product nya?

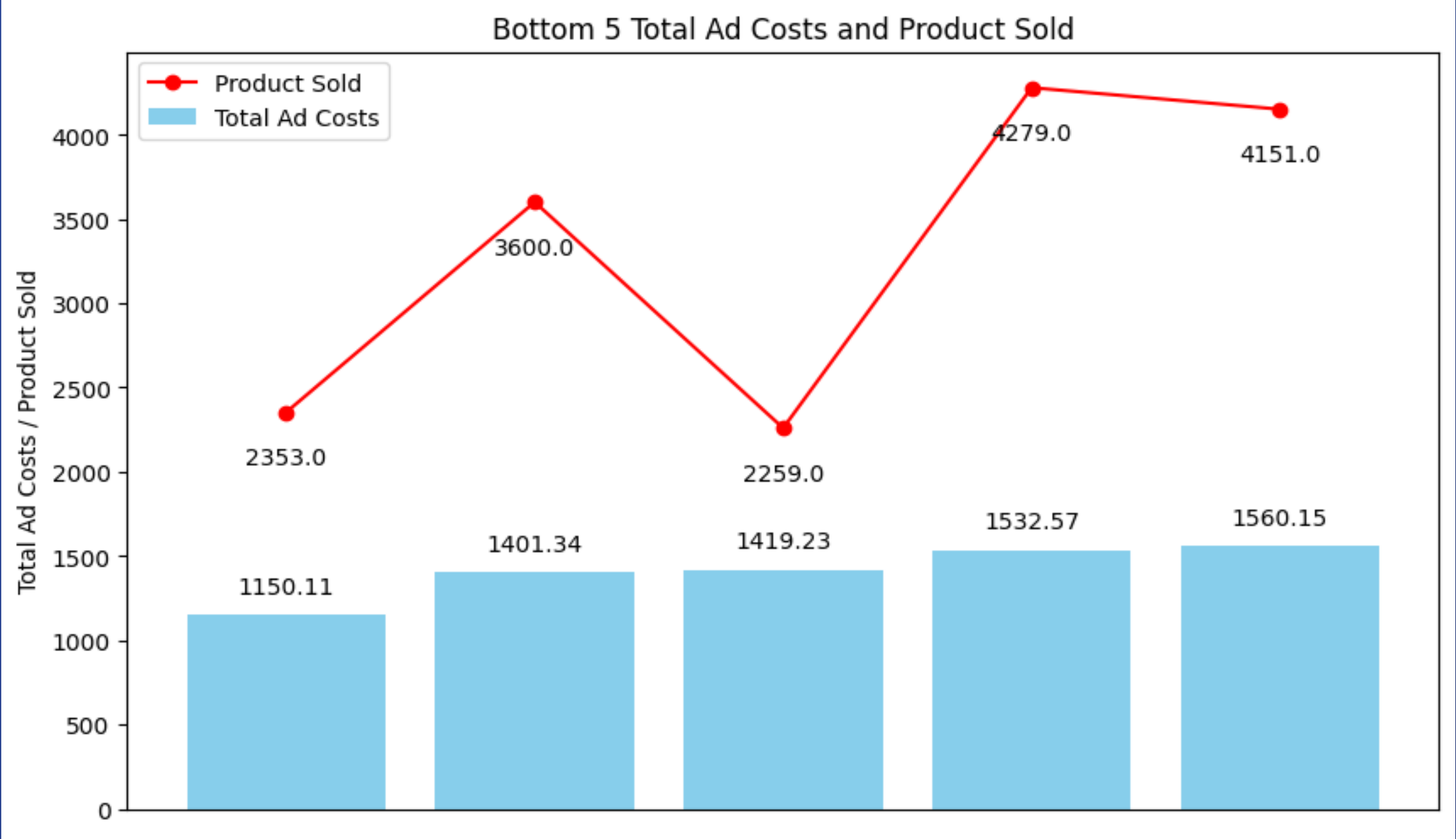
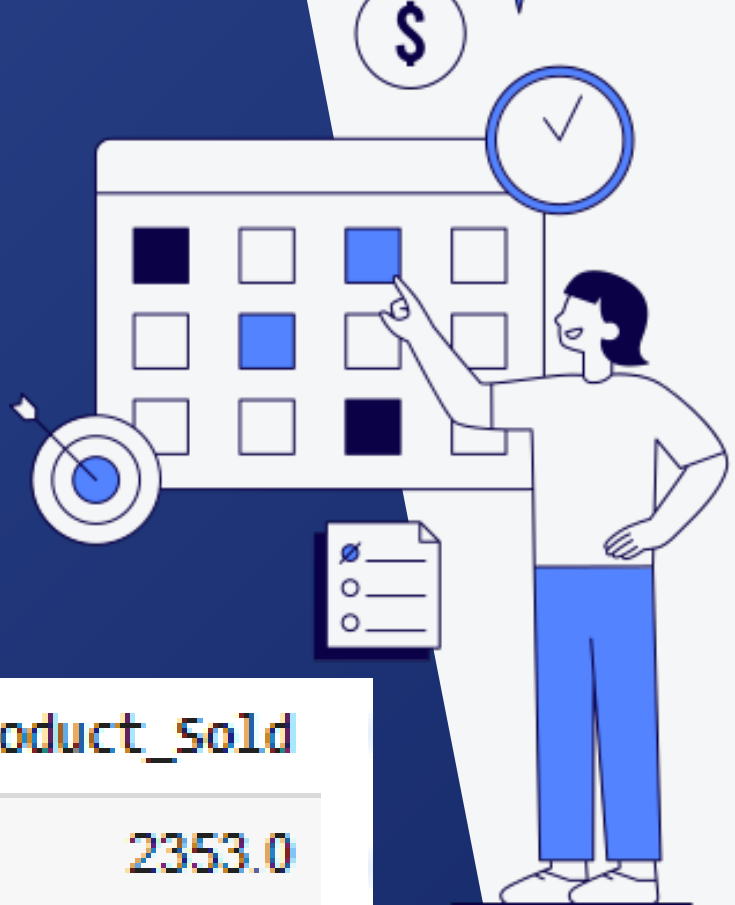


	Total_Ad_Costs	Product_Sold
11	5225.42	12227.0
10	4756.00	11210.0
96	4607.78	10763.0
165	4519.85	9498.0
147	4514.45	9889.0

Total_Ads terbesar 5225.42
dengan penjualan produk
12227.0

Exploratory Data Analyst Questions

Bottom 5 total biaya ads dan berapa penjualan product nya?

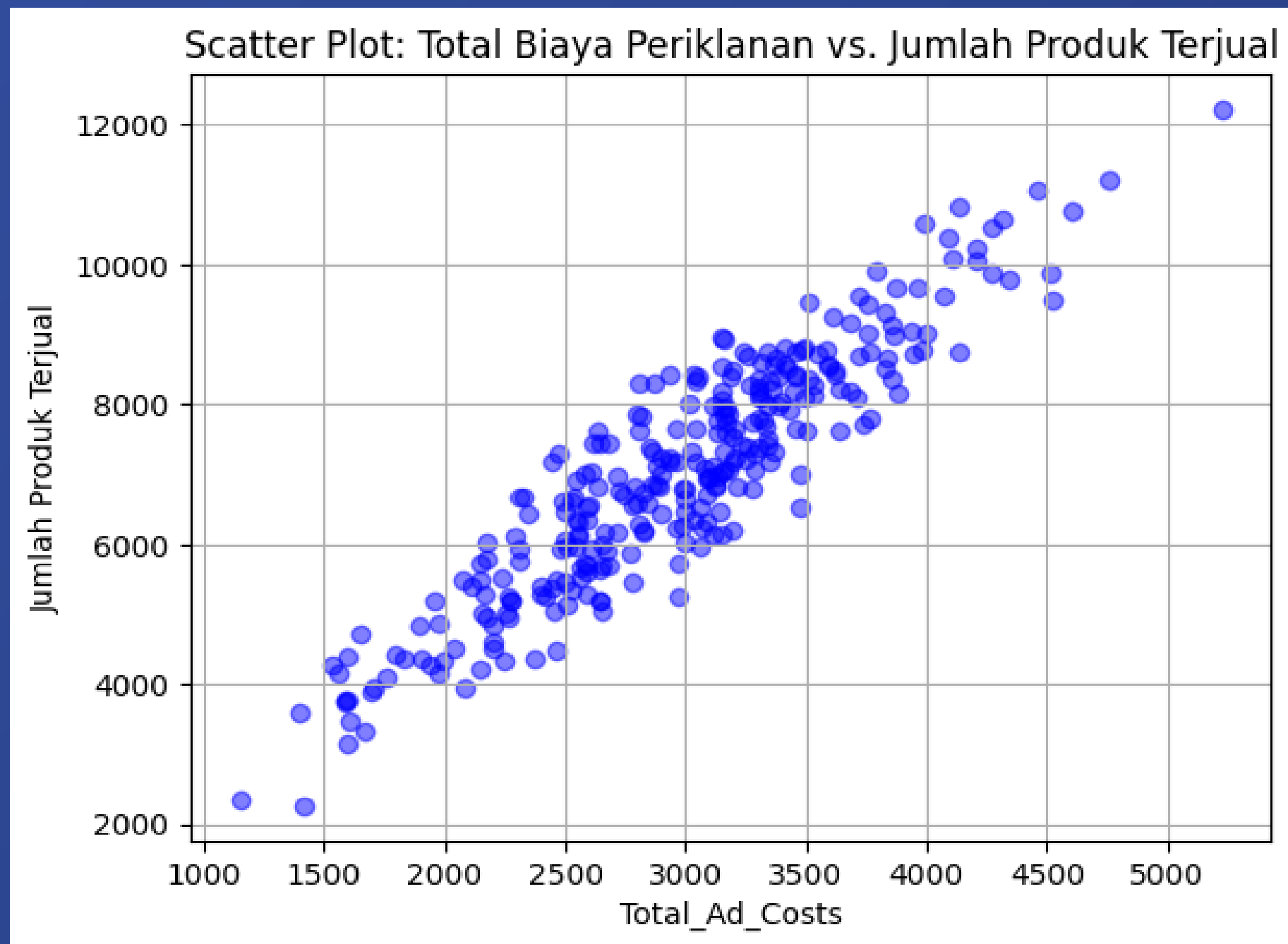


	Total_Ad_Costs	Product_Sold
281	1150.11	2353.0
36	1401.34	3600.0
145	1419.23	2259.0
171	1532.57	4279.0
78	1560.15	4151.0

Total_Ads terkecil 1150.11
dengan penjualan produk
2353.0

Exploratory Data Analyst Questions

Bagaimana hubungan biaya saluran periklanan mempengaruhi penjualan produk?



**semakin tinggi biaya ads, maka
semakin besar penjualannya**



Exploratory Data Analyst Questions

Saluran periklanan mana yang memberikan kontribusi terbesar terhadap penjualan produk?

	Fitur	Koefisien
0	TV	575.792808
1	Billboards	825.742566
2	Google_Ads	426.869826
3	Social_Media	683.512159
4	Influencer_Marketing	345.301761
5	Affiliate_Marketing	1109.030676

berdasarkan coefisiennya, **Affiliate_marketing** memberikan kontribusi terbesar terhadap penjualan produk



Exploratory Data Analyst

Questions

saluran periklanan mana yang memberikan kontribusi terkecil terhadap penjualan produk?

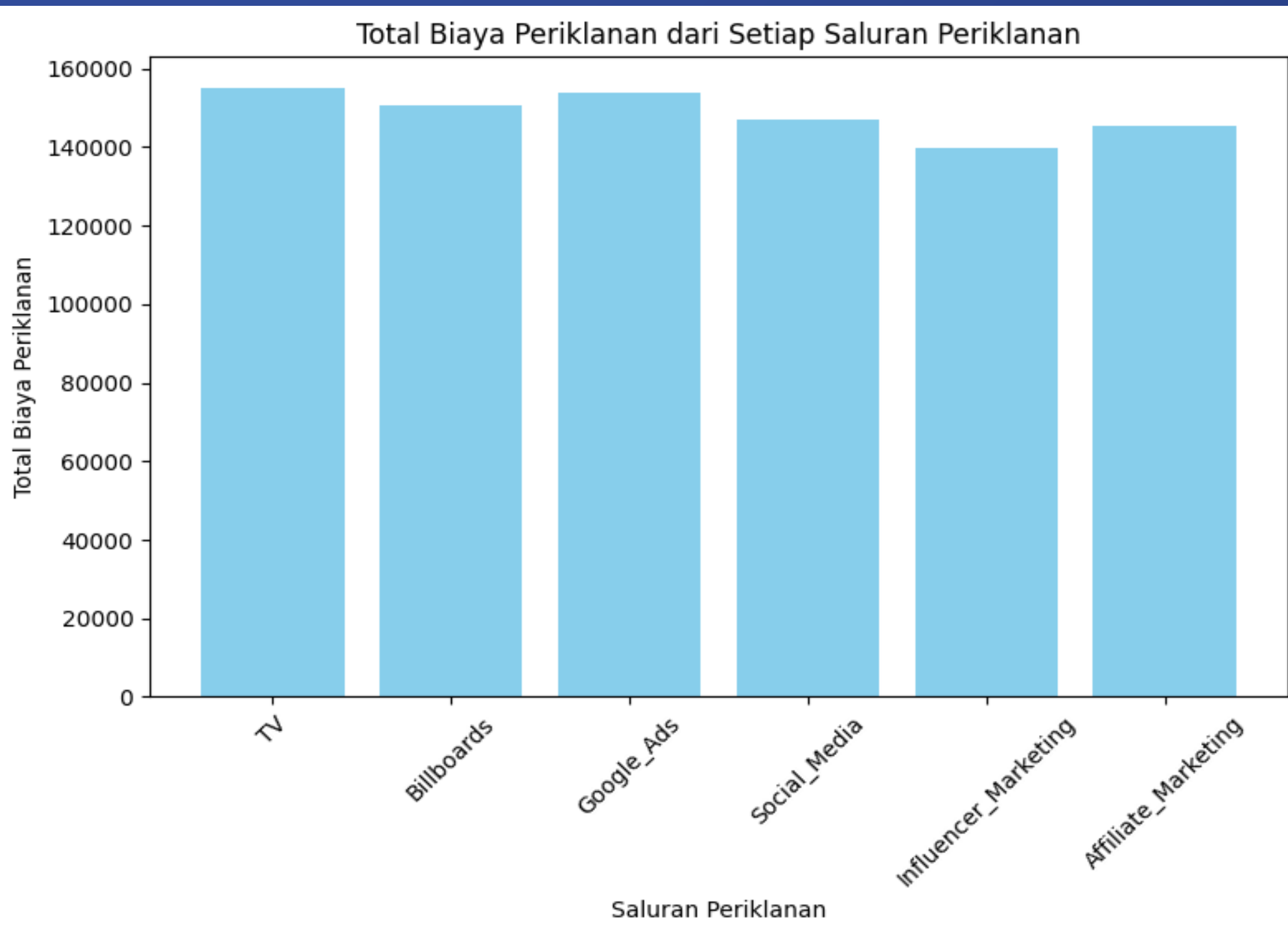
	Fitur	Koefisien
0	TV	575.792808
1	Billboards	825.742566
2	Google_Ads	426.869826
3	Social_Media	683.512159
4	Influencer_Marketing	345.301761
5	Affiliate_Marketing	1109.030676

berdasarkan coefisiennya, Influencer_marketing memberikan kontribusi terkecil terhadap penjualan produk



Exploratory Data Analyst Questions

Apakah alokasi biaya iklan sudah efektif?

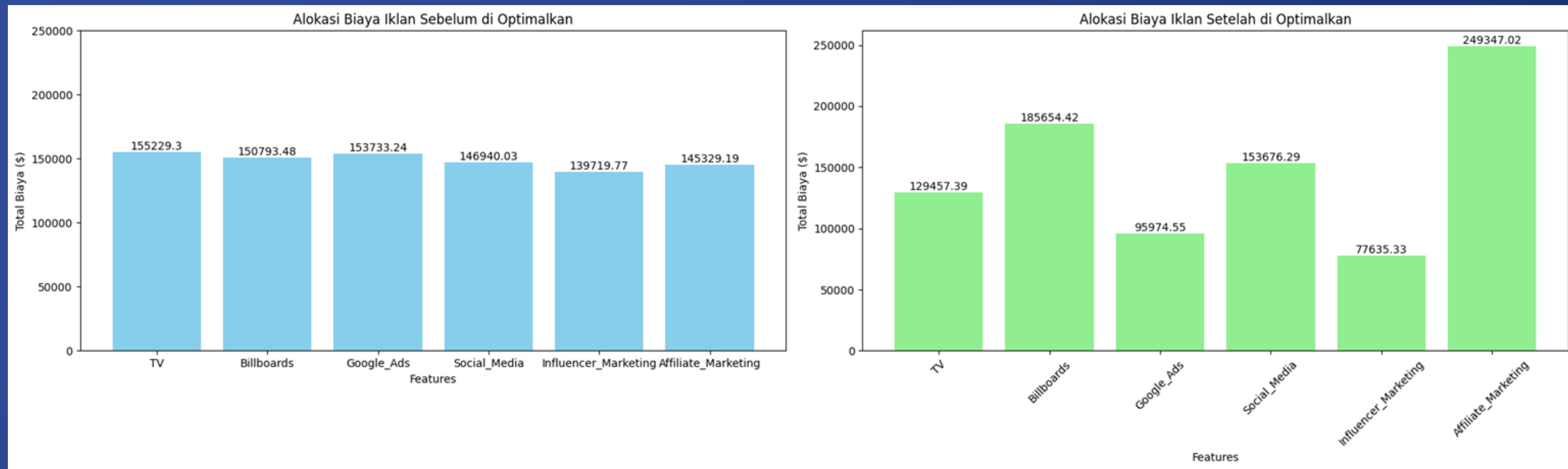


dari data yang ada, alokasi biaya yang dikeluarkan relatif hampir merata, namun secara coefision setiap fitur berbeda. jawabannya tidak!, alokasi biaya tidak efektif !



Exploratory Data Analyst Questions

Bagaimana mengoptimalkan alokasi biaya iklan agar lebih efektif?



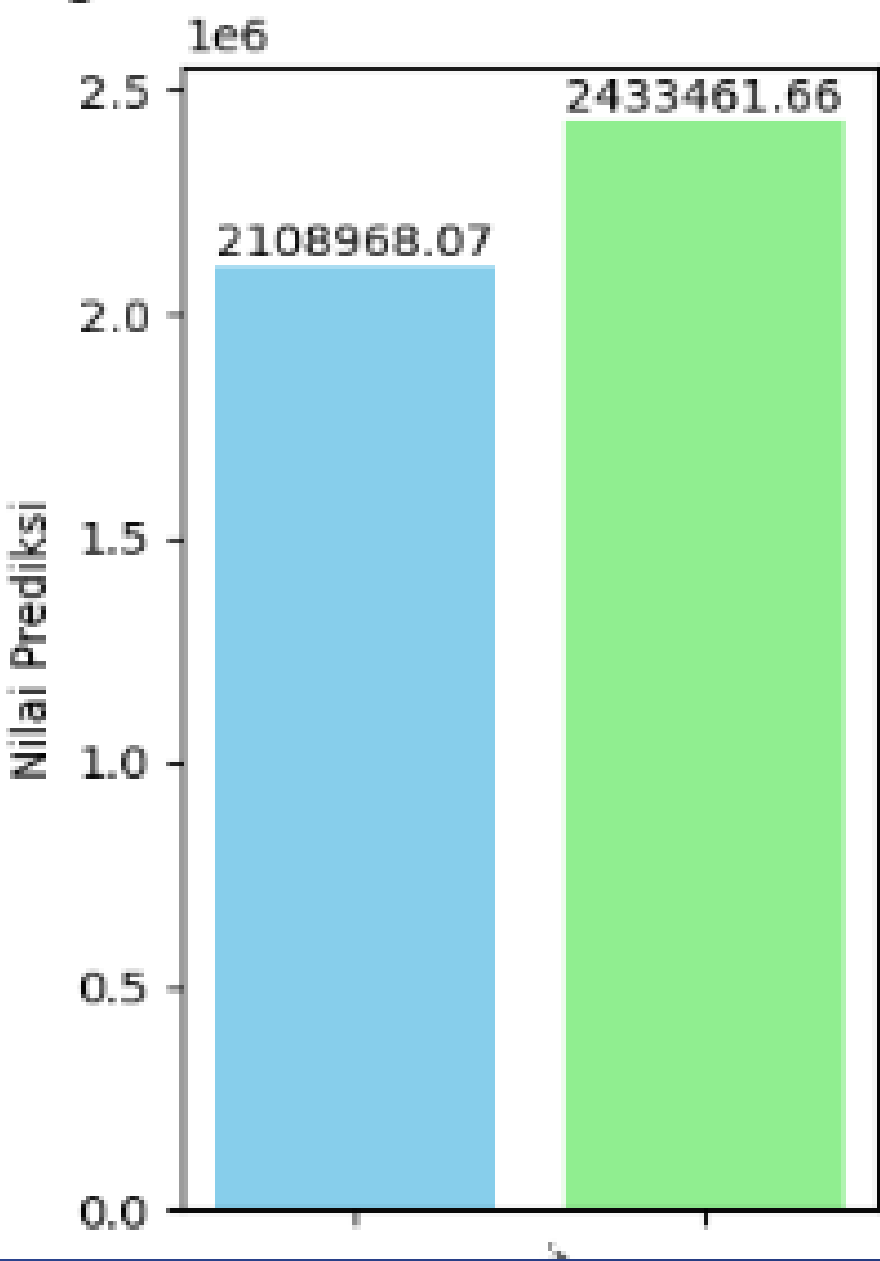
terlihat alokasi biaya menjadi lebih efektif sesuai dengan pengaruh fitur terhadap penjualan dengan biaya yang sama

Exploratory Data Analyst Questions

Berapa persen kenaikan / Penurunan penjualan setelah mengoptimalkan alokasi biaya?



Perbandingan Prediksi Sebelum dan Setelah Optimasi



	Keterangan	Nilai
0	Prediksi Sebelum Optimasi	2108968.07
1	Prediksi Setelah Optimasi	2433461.66
2	Selisih	324493.58
3	Persentase Kenaikan/Penurunan	15.39

terjadi kenaikan penjualan produk sebesar 324493.58 dengan presentase 15,39% setelah mengoptimalkan alokasi biaya iklan.


Note:
prediksi penjualan produk secara total

Conclusion

INSIGHT

- Saluran iklan memiliki pengaruh yang berbeda-beda terhadap penjualan
- berdasarkan koefisien nya, alokasi biaya yang ada tidak efektif

Recommendation:

- Alokasi kan biaya sesuai dengan pengaruh saluran iklan nya
 - Evaluasi kualitas setiap saluran iklan, contoh: apakah Influencer_marketing yang dipakai memiliki engagement tinggi?
 - Lakukan uji coba terhadap alokasi biaya baru dan terus pantau kinerjanya.
 - tambahkan fitur-fitur baru
- 

**THANK
YOU**