

# Predict Clicked Ads Customer Classification by using Machine Learning

Supported by:  
**Rakamin Academy**  
Career Acceleration School  
[www.rakamin.com](http://www.rakamin.com)



**Created by:**

**RAMADHANI NURUL FAZRI**

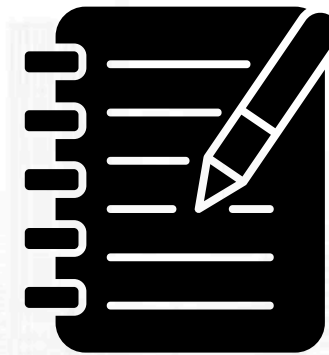
[rnurul08@gmail.com](mailto:rnurul08@gmail.com)

[www.linkedin.com/in/ramadhani-nurul-fazri](https://www.linkedin.com/in/ramadhani-nurul-fazri)

“Saya individu yang berkomitmen dan berdedikasi, memiliki latar belakang pendidikan di bidang Teknik Geologi. Kemampuan analitis yang kuat dan keterampilan komunikatif yang baik, saya mampu bekerja secara efektif dalam tim maupun secara mandiri. Saya memiliki minat yang besar dalam analisis data dan selalu berusaha untuk belajar dan berkembang, serta berkontribusi secara positif terhadap tujuan organisasi. Selain itu, saya memiliki ketertarikan dalam penerapan teknologi dan inovasi untuk meningkatkan efisiensi serta hasil kerja.”

“Sebuah perusahaan di Indonesia ingin mengetahui efektifitas sebuah iklan yang mereka tayangkan, hal ini penting bagi perusahaan agar dapat mengetahui seberapa besar ketercapainnya iklan yang dipasarkan sehingga dapat menarik customers untuk melihat iklan. Dengan mengolah data historical advertisement serta menemukan insight serta pola yang terjadi, maka dapat membantu perusahaan dalam menentukan target marketing, fokus case ini adalah membuat model machine learning classification yang berfungsi menentukan target customers yang tepat ”





## 1003 Data Observasi 11 Variabel

1. **Variabel Dependen** dari data tersebut adalah **Click on Ad** dimana luaran dari kolom tersebut adalah 1 untuk mengklik iklan dan 0 untuk tidak mengklik.
2. **Variabel Independen** dari data tersebut adalah Daily Time Spent on Site, Age, Area Income, Daily Internet Usage, Male, Timestamp, city, province, category.
3. Terdapat tipe kolom yang tidak sesuai dan perlu di ubah.
  - a. Tipe kolom 'Male' seharusnya bertipe interger.
  - b. Tipe kolom 'Clicked on Ad' seharusnya bertipe interger.
4. Kolom Unnamed: 0 tidak termasuk variabel independen karena hanya melabelkan identitas pelanggan.



# **CUSTOMER TYPE AND BEHAVIOUR ANALYSIS ON ADVERTISEMENT**

## STATISTIKA ANALISIS

### Data Numerik

	Unnamed: 0	Daily Time Spent on Site	Age	Area Income	Daily Internet Usage	Male	Clicked on Ad
count	963.000000	963.000000	963.000000	9.630000e+02	963.000000	963.000000	963.000000
mean	499.855659	64.829200	36.049844	3.855628e+08	179.716106	0.478712	0.503634
std	286.827132	15.892075	8.764154	9.380390e+07	43.867289	0.499806	0.500247
min	0.000000	32.600000	19.000000	9.797550e+07	104.780000	0.000000	0.000000
25%	253.500000	50.600000	29.000000	3.296658e+08	138.615000	0.000000	0.000000
50%	499.000000	68.010000	35.000000	3.991039e+08	182.200000	0.000000	1.000000
75%	749.500000	78.365000	42.000000	4.591870e+08	218.550000	1.000000	1.000000
max	999.000000	91.430000	61.000000	5.563936e+08	267.010000	1.000000	1.000000

#### 1. Kebiasaan Online:

- Waktu harian di situs berkisar antara 50–78 menit, sedangkan penggunaan internet harian sekitar 3 jam.
- Pengguna dengan waktu internet tinggi (>218 menit) mungkin lebih mudah dijangkau melalui iklan online.

2. Efektivitas Iklan: Tingkat klik yang cukup seimbang (50%) menunjukkan potensi yang baik untuk meningkatkan keterlibatan melalui strategi iklan yang lebih terarah.

### Data Katagori

	Timestamp	city	province	category
count	963	963	963	963
unique	960	30	16	10
top	5/26/2016 15:40	Bandung	Daerah Khusus Ibukota Jakarta	Otomotif
freq	2	64	244	108

1. Timestamp Unik: Mayoritas entri memiliki waktu yang unik, menunjukkan bahwa data mencakup berbagai momen aktivitas.
2. Dominasi Kota dan Provinsi: Kota Bandung dan Provinsi DKI Jakarta memiliki jumlah data terbanyak.
3. Kategori: Kategori Otomotif sebagai kategori teratas, ini mengindikasikan bahwa sektor ini mungkin sedang berkembang atau memiliki perhatian besar di pasar.

Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)



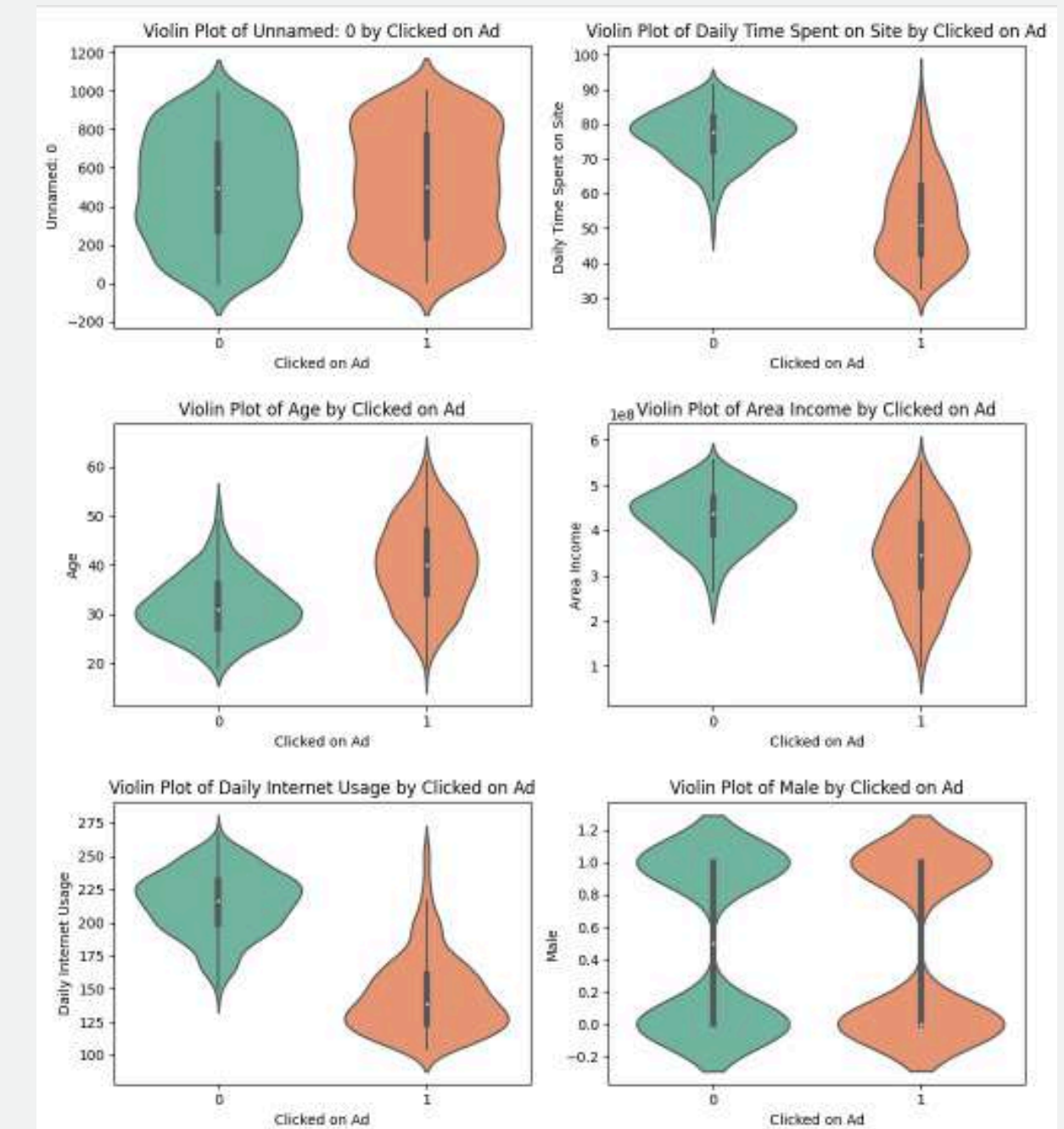


## **VISUALISASI DATA**

## UNIVARIATE ANALYSIS

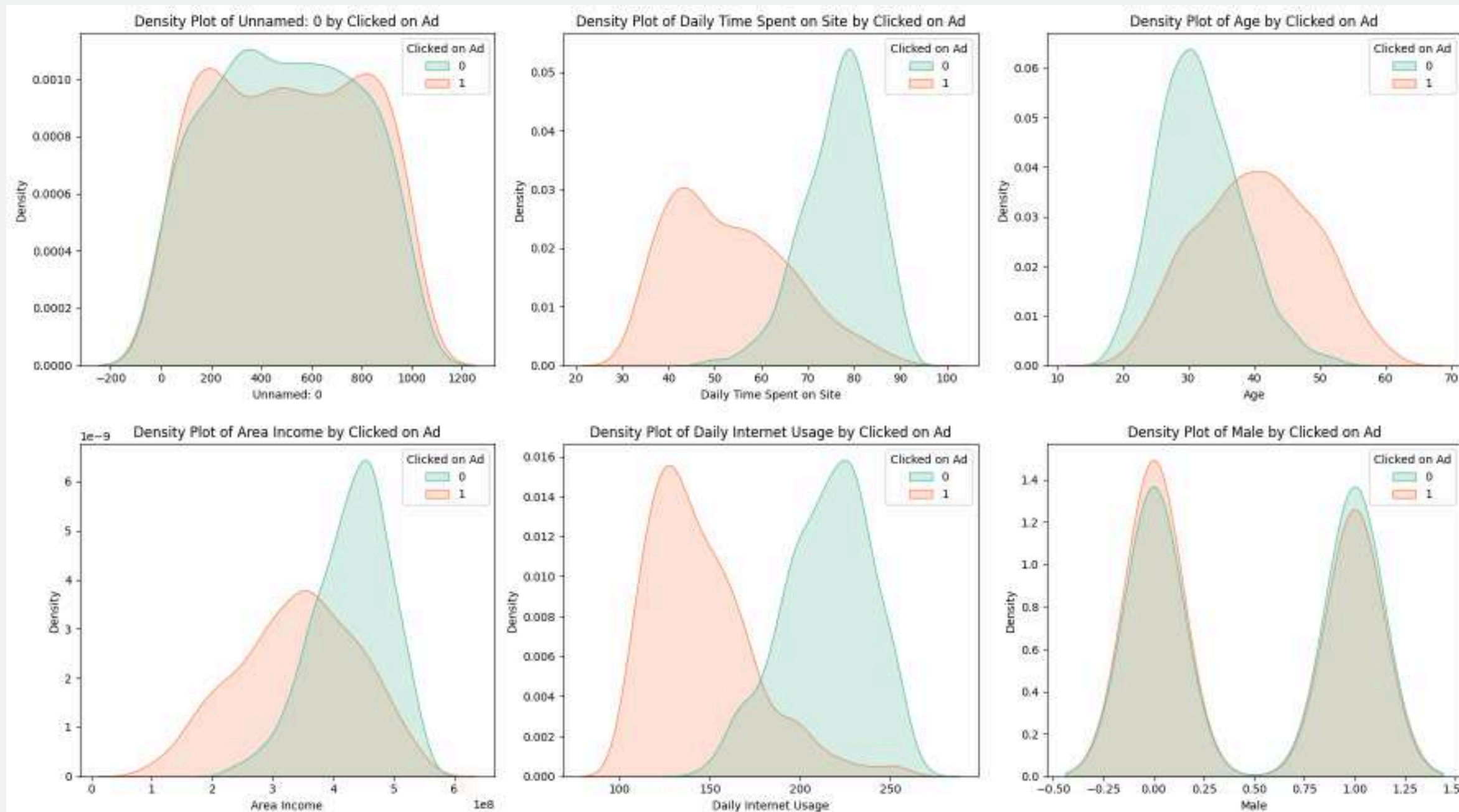
### Violin Plot

1. Plot Daily Time Spent terhadap Clicked on Ad menunjukkan yang meng-klik memiliki lebar atau puncaknya bergeser ke kanan dibandingkan dengan yang tidak mengklik, badan yang lebih lebar menunjukkan area dengan kepadatan yang lebih tinggi dalam data.
2. Plot Age terhadap Clicked on Ad menunjukkan kelompok usia yang lebih cenderung meng-klik iklan dibandingkan yang tidak.



Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)

## UNIVARIATE ANALYSIS



### Density Plot

Analisis density plot pada Daily Time Spent on Site terhadap Clicked on Ad menunjukkan kurva yang bergeser ke arah kanan hal ini menunjukkan bahwa pengguna cenderung menghabiskan lebih banyak waktu di situs dibandingkan dengan pengguna yang tidak mengklik iklan.

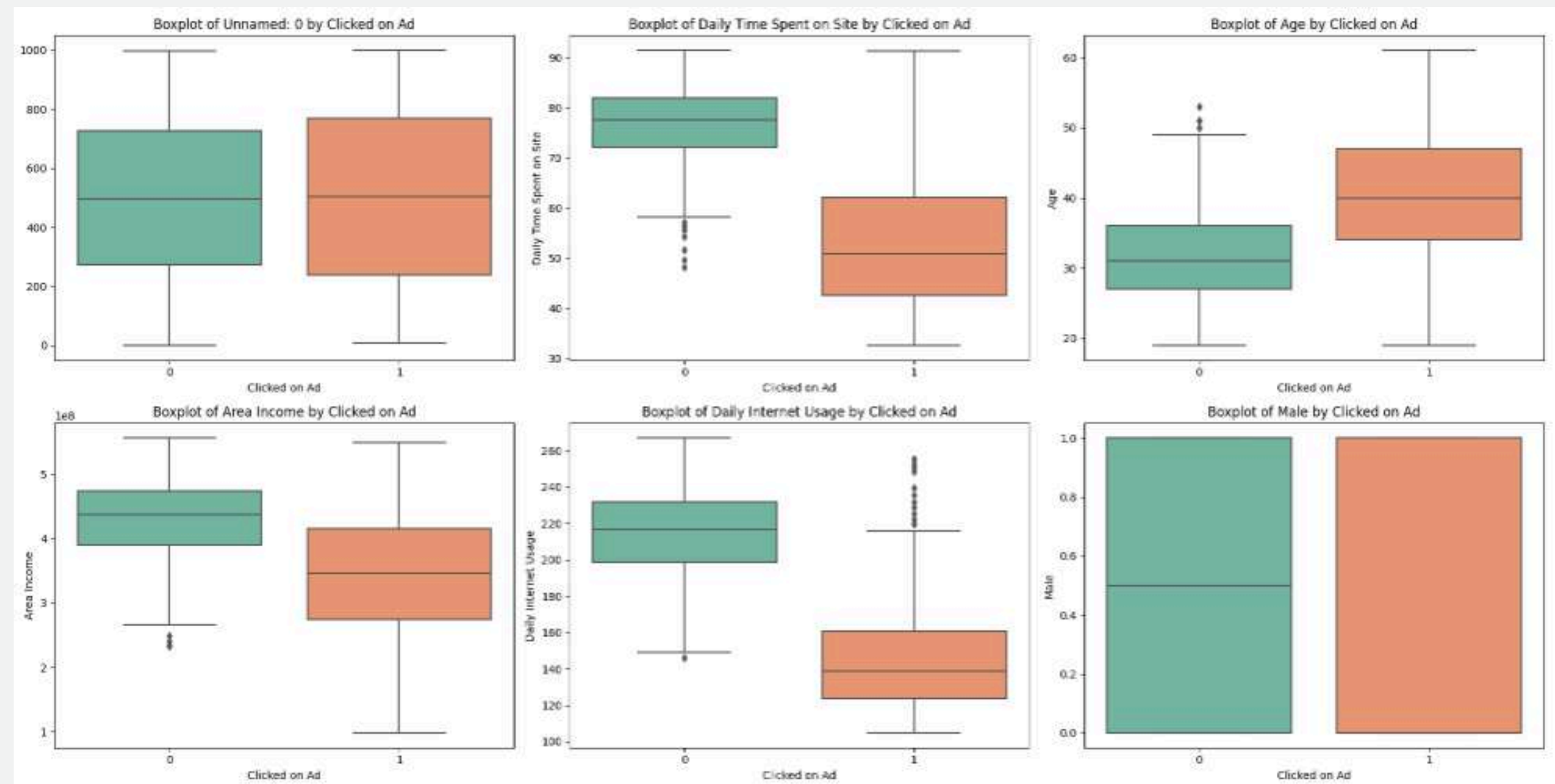
Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)



## BIVARIATE ANALYSIS

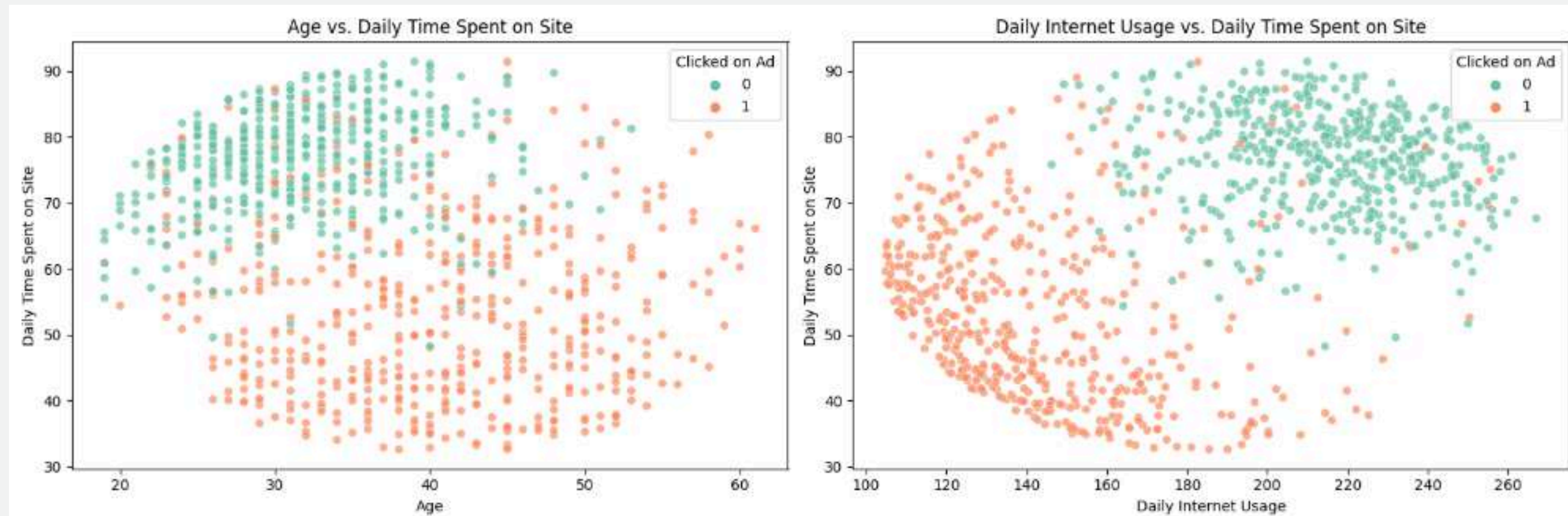
### Box Plot

1. Kolom Daily Time Spent on Site terhadap Clicked on Ad memiliki outliers, mengindikasikan pengguna yang menghabiskan lebih banyak waktu di site secara signifikan daripada yang lain.
2. Kolom Area Income terhadap Clicked on Ad memiliki kotak yang sempit mengindikasikan lebih sedikit variabilitas, sementara kotak yang lebar menunjukkan rentang pendapatan yang lebih luas.
3. Kolom Daily Internet Usage terhadap Clicked on Ad menunjukkan adanya outliers.



## BIVARIATE ANALYSIS

### Scatter plot



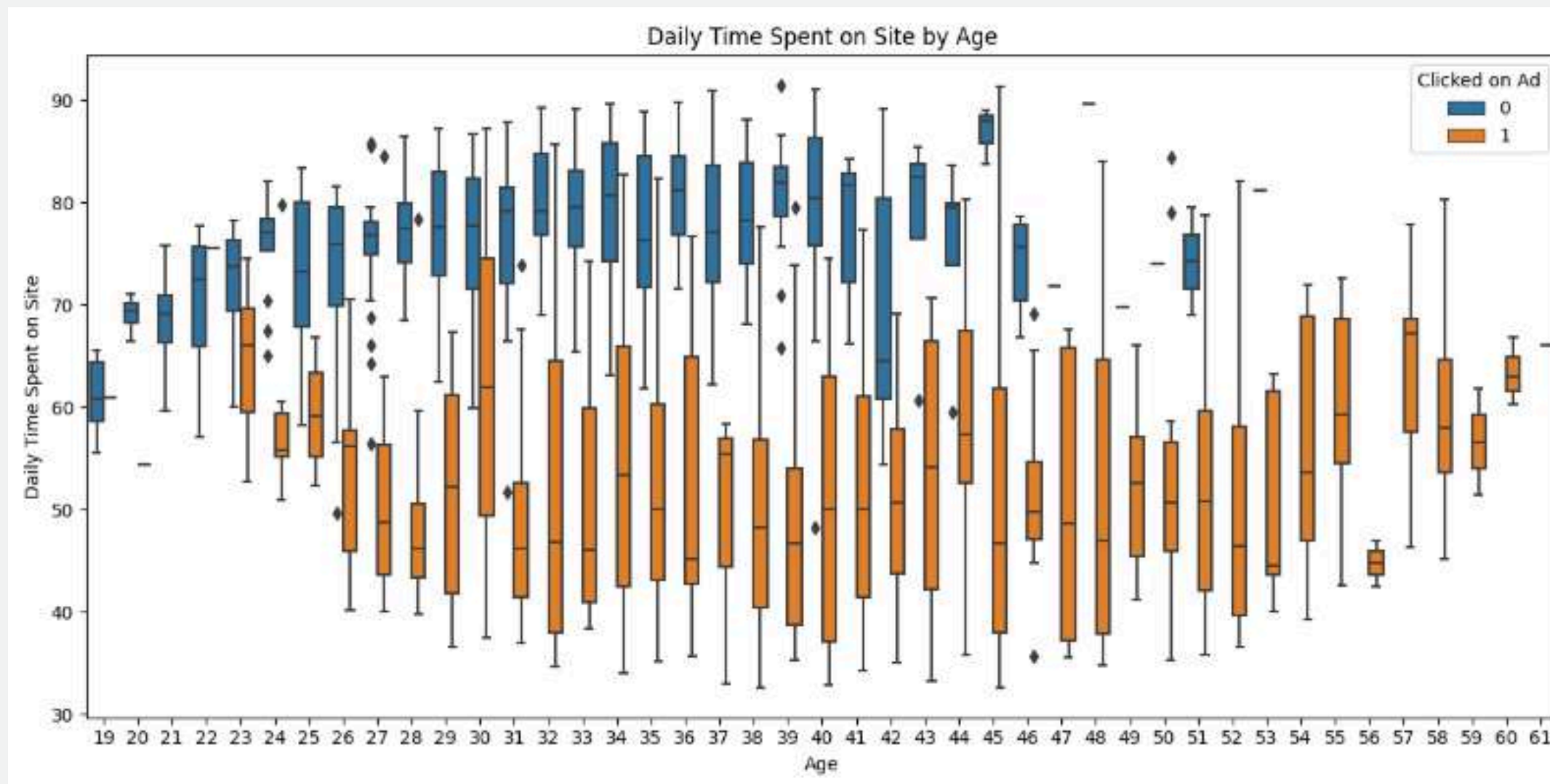
1. Age vs. Daily Time Spent on Site: Tidak ada tren linier yang jelas terlihat, titik-titik data agak tersebar menunjukkan korelasi linier lemah atau tidak ada sama sekali.
2. Daily Internet Usage vs. Daily Time Spent on Site: Plot menunjukkan korelasi yang agak positif, hubungan tidak linear sempurna terdapat beberapa sebaran tetapi tren kenaikan secara umum terlihat.

Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)



## BIVARIATE ANALYSIS

### Scatter plot

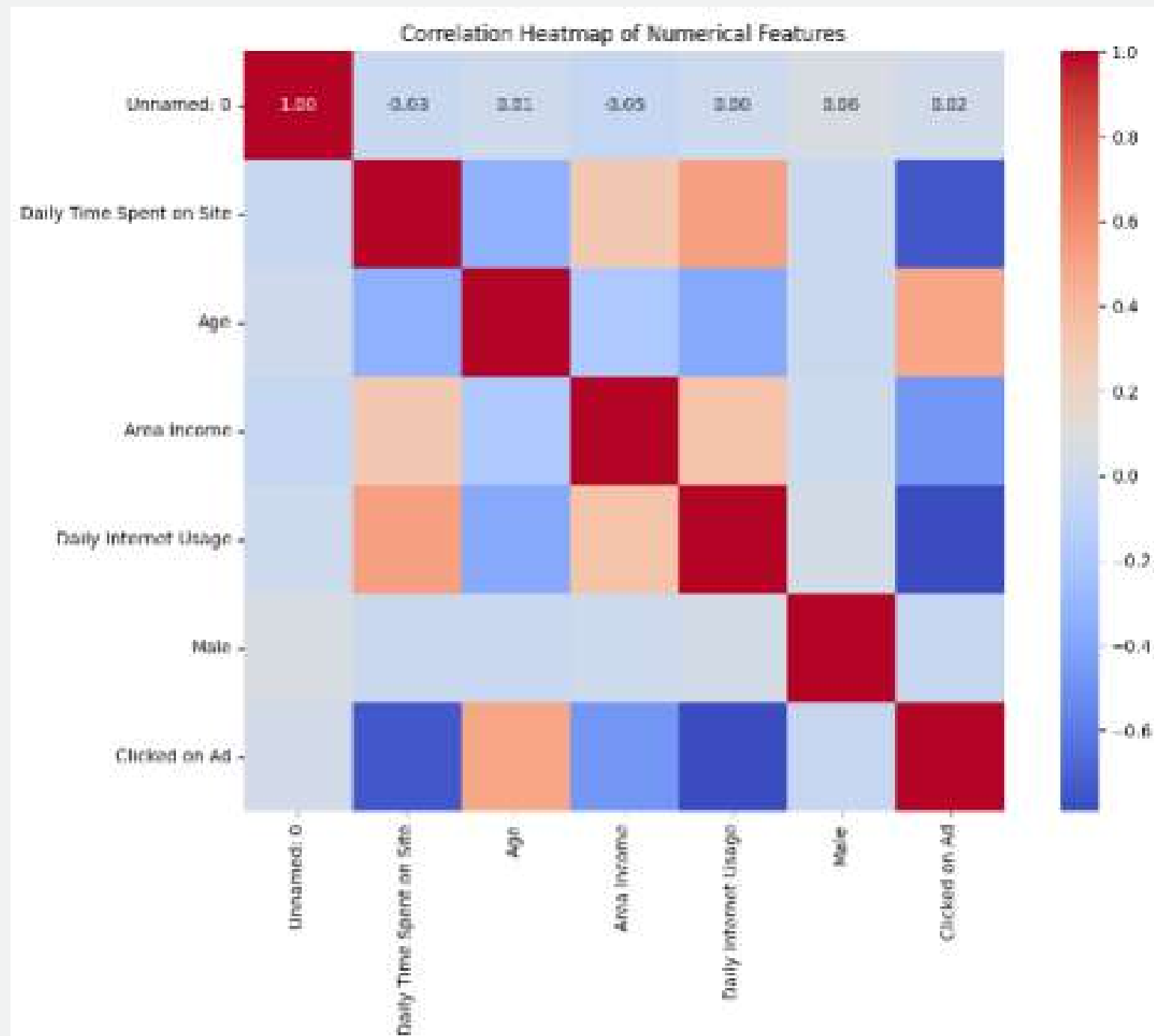


1. No Clear Age Trend: Tidak ada tren yang kuat secara keseluruhan yang menunjukkan hubungan antar Age dan Daily Time Spent on The Site. Median (garis dalam kotak) tidak secara konsisten naik atau turun seiring bertambahnya usia. Distribusi waktu yang dihabiskan tampak relatif konsisten di seluruh rentang usia.
2. Variability: Ada variasi yang cukup besar dalam waktu yang di habiskan pengguna dengan usia yang sama.
3. Outliers: Terlihat beberapa outliers ini mewakili pengguna yang menghabiskan lebih banyak atau lebih sedikit waktu di situs dibandingkan pengguna lain dalam kelompok usia mereka.
4. Ad Click Impact (Minimal): Waktu yang dihabiskan pengguna di situs tidak secara kuat memprediksi apakah mereka akan meng-klik iklan.

Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)



## MULTIVARIATE ANALYSIS



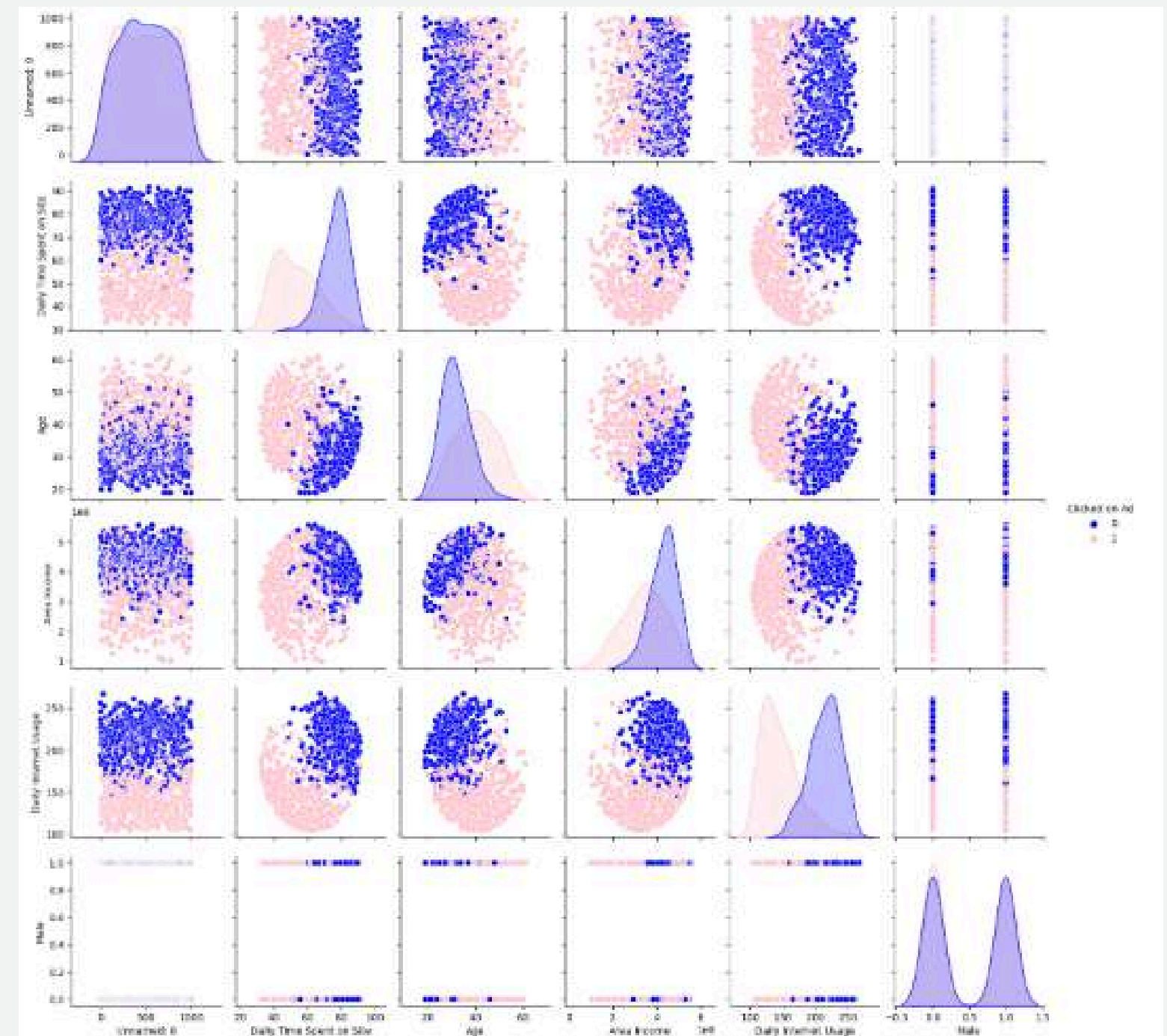
### Heatmap

1. Korelasi antara Daily Time Spent on Site dengan Clicked on Ad bernilai positif, bahwa pengguna menghabiskan lebih banyak waktu di website dan mereka cenderung mengklik iklan.
2. Korelasi antara Age dengan Clicked on Ad memiliki nilai rendah, usia mungkin tidak secara signifikan mempengaruhi klik iklan.
3. Korelasi Area Income dengan Clicked on Ad memiliki korelasi lemah, tingkat pendapatan mungkin tidak secara dramatis mempengaruhi perilaku mengklik.
4. Korelasi Daily Internet Usage dengan Clicked on Ad memiliki nilai positif, hal ini menunjukkan bahwa pengguna yang menghabiskan lebih banyak waktu online mungkin juga lebih sering mengklik iklan.
5. Korelasi antara Male dengan Clicked on Ad memiliki nilai korelasi lemah, hal ini mungkin menyiratkan target yang lebih universal.

## MULTIVARIATE ANALYSIS

### Pair Plot

1. Kolom Daily Time Spent on Site dengan variabel lain menunjukkan tren yang jelas dimana lebih banyak waktu yang dihabiskan di situs berkorelasi dengan kemungkinan lebih tinggi untuk meng-klik iklan. Plot kepadatan untuk fitur ini menunjukkan puncak yang berbeda untuk mereka yang meng-klik, keterlibatan dipengaruhi oleh waktu yang dihabiskan.
2. Area Income bukanlah faktor kuat yang menentukan pengunjung meng-klik iklan atau tidak.
3. Daily Internet Usage mungkin berinteraksi dengan iklan.



Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)



# **DATA CLEANING & PREPROCESSING**



## Missing Value

```
Unnamed: 0      0
Daily Time Spent on Site  13
Age                0
Area Income        13
Daily Internet Usage  11
Male              3
Timestamp         0
Clicked on Ad      0
city              0
province          0
category          0
```

Ditemukan baris bernilai kosong pada kolom Daily Time Spent on Site, Area Income, Daily Internet Usage, dan Male. Jumlah baris kosong kurang dari 5% jadi akan di dilakukan drop baris.

## Duplicated Value

```
#Mendeteksi duplikasi baris
df.duplicated().sum()
```

```
0
```

Tidak ditemukan data duplikat.

## EKSTRAKSI KOLOM

	Unnamed: 0	Daily Time Spent on Site	Age	Area Income	Daily Internet Usage	Male	Timestamp	Clicked on Ad	city	province	category	Year	Month	Week	Day
0	0	68.95	35	432837300.0	256.09	0	2016-03-27 00:53:00	0	Jakarta Timur	Daerah Khusus Ibukota Jakarta	Furniture	2016	3	12	27
1	1	80.23	31	479092950.0	193.77	1	2016-04-04 01:39:00	0	Denpasar	Bali	Food	2016	4	14	4
2	2	69.47	26	418501580.0	236.50	0	2016-03-13 20:35:00	0	Surabaya	Jawa Timur	Electronic	2016	3	10	13
3	3	74.15	29	383643260.0	245.89	1	2016-01-10 02:31:00	0	Batam	Kepulauan Riau	House	2016	1	1	10
4	4	68.37	35	517229930.0	225.58	0	2016-06-03 03:36:00	0	Medan	Sumatra Utara	Finance	2016	6	22	3

kolom yang berhubungan dengan waktu dalam dataset tersebut adalah:

**Timestamp:** Waktu saat konsumen mengklik iklan atau menutup iklan

Untuk melakukan analisis lebih lanjut atau mengekstrak fitur dari kolom waktu, dilakukan proses mengonversi kolom dan mengekstrak elemen seperti tahun, bulan, minggu, dan hari

Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)

## FEATURE ENCODING

	Daily Time Spent on Site	Age	Area Income	Daily Internet Usage	Male	Clicked on Ad	Year	Month	Week	Day	...	category_Bank	category_Electronic	category_Fashion	category_Finance	category_Food	category_Furniture	category_Health
0	68.95	35	432837300.0	256.09	0	0	2016	3	12	27	...	False	False	False	False	False	True	False
1	80.23	31	479092950.0	193.77	1	0	2016	4	14	4	...	False	False	False	False	True	False	False
2	69.47	26	418501580.0	236.50	0	0	2016	3	10	13	...	False	True	False	False	False	False	False
3	74.15	29	383643260.0	245.89	1	0	2016	1	1	10	...	False	False	False	False	False	False	False
4	68.37	35	517229930.0	225.58	0	0	2016	6	22	3	...	False	False	False	True	False	False	False

5 rows × 66 columns

Melakukan feature encoding dengan menggunakan **get\_dummy** untuk **kolom kategori**



## FEATURE & TARGET

### FEATURE

Variabel yang digunakan untuk memprediksi target. Pada analisis kali ini feature adalah semua kolom kecuali kolom Clicked on Ad.

### TARGET

Variabel yang ingin di prediksi, pada penelitian kali ini kolom **Clicked on Ad** adalah targetnya.

# DATA MODELING



## SPLIT DATA

### DATA TRAIN

Data train sebanyak  
**80%**

Jumlah data train: **770**

### DATA TEST

Data test sebanyak  
**20%**

Jumlah data test: **193**



## KONTEKS PENGUKURAN KINERJA MODEL

### TP (TRUE POSITIVE)

Model prediksi pelanggan meng-klik iklan, nyatanya meng-klik iklan.

### TN (TRUE NEGATIVE)

Model prediksi pelanggan tidak meng-klik iklan, nyatanya tidak meng-klik iklan.

### FP (FALSE POSITIVE)

Model prediksi pelanggan meng-klik iklan, nyatanya tidak meng-klik iklan.

### FN (FALSE NEGATIVE)

Model prediksi pelanggan tidak meng-klik iklan, nyatanya meng-klik iklan.

## MACHINE LEARNING

Penelitian ini memiliki hubunga **Presisi (Precision)**

$$\text{Precision} = (TP)/(TP + FP)$$

Dari semua prediksi positif, berapa banyak yang benar-benar positif. Ini mengukur akurasi dari prediksi positif. Semakin tinggi precision, semakin sedikit false positives, yang berarti model lebih akurat dalam memprediksi kasus positif.

## HASIL EXPERIMENT 1

Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)



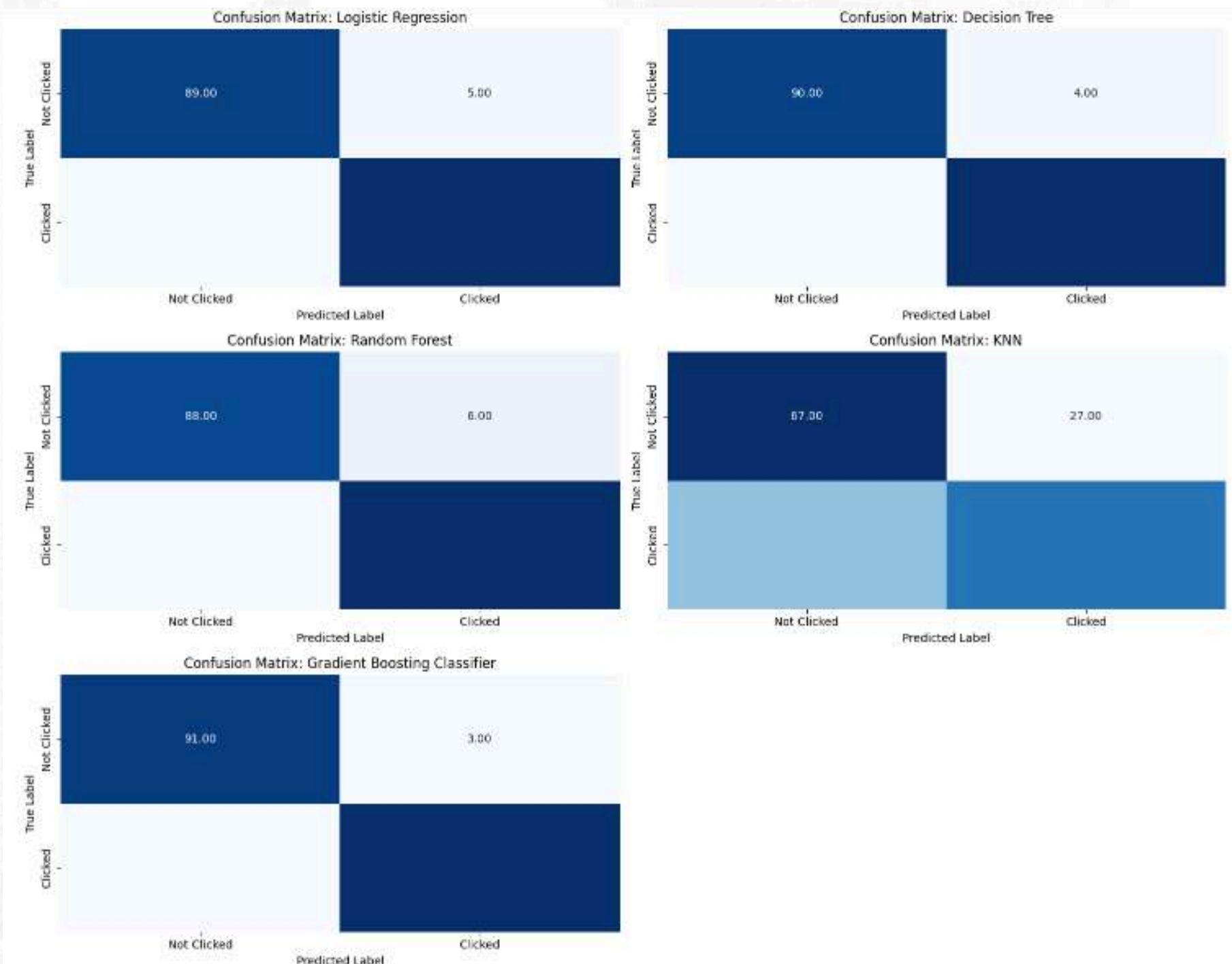
## MACHINE LEARNING ANALYSIS

	Model	Akurasi	Precision	Recall	F1 Score
0	Logistic Regression	0.9534	0.9500	0.9596	0.9548
1	Decision Tree	0.9482	0.9406	0.9596	0.9500
2	Random Forest	0.9689	0.9604	0.9798	0.9700
3	K-Nearest Neighbor	0.6373	0.6747	0.5657	0.6154
4	Gradient Boosting Classifier	0.9689	0.9697	0.9697	0.9697

1. **Gradient Boosting Classifier:** Precision dan recall berada pada tingkat yang sama.
2. **Decision Tree dan Random Forest:** Random Forest memiliki precision yang lebih rendah (0.9406) meskipun recall-nya sangat tinggi (0.9798).
3. **Logistic Regression:** Precision dan recall yang tinggi menunjukkan bahwa model ini cukup baik dalam mengklasifikasikan kelas positif.
4. **K-Nearest Neighbor (KNN):** Nilai precision dan recall yang tidak seimbang (precision tinggi tetapi recall rendah) menunjukkan masalah dalam menangkap kelas positif.

## MACHINE LEARNING ANALYSIS

1. **Random Forest** menunjukkan performa terbaik di antara semua model dengan TP tertinggi dan hanya 1 FN.
2. **Gradient Boosting Classifier** dan **Decision Tree** juga menunjukkan hasil yang sangat baik, dengan keseimbangan antara TP dan FP/FN yang rendah.
3. **Logistic Regression** juga baik, tetapi tidak seoptimal model-model di atas.
4. **KNN**, bagaimanapun, dengan kesalahan yang tinggi dalam klasifikasi, memerlukan perhatian lebih untuk dioptimalkan. Model lain jauh lebih efektif untuk dataset ini.



Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)

## HASIL EXPERIMENT 2

Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)



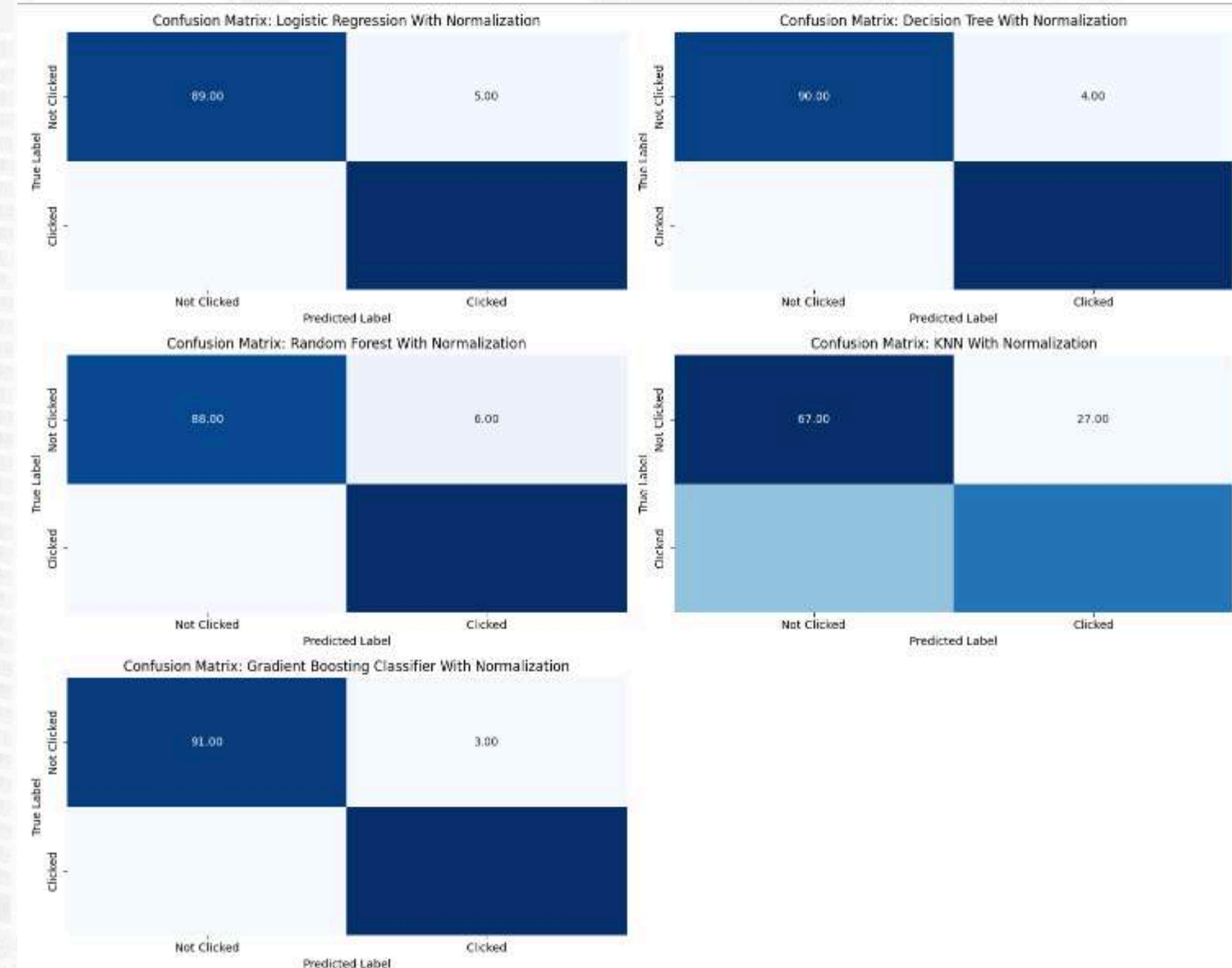
## MACHINE LEARNING ANALYSIS

	Model With Normalization	Akurasi	Precision	Recall	F1 Score
0	Logistic Regression With Normalization	0.9637	0.9694	0.9596	0.9645
1	Decision Tree With Normalization	0.9482	0.9406	0.9596	0.9500
2	Random Forest With Normalization	0.9482	0.9423	0.9899	0.9655
3	K-Nearest Neighbor With Normalization	0.9482	0.7789	0.7475	0.7629
4	Gradient Boosting Classifier With Normalization	0.9482	0.9697	0.9697	0.9697

1. **Logistic Regression With Normalization:** Model ini memiliki akurasi tertinggi (0.9637), dengan precision dan recall yang juga baik, serta F1 score yang sangat baik.
2. **Decision Tree With Normalization:** Meskipun memiliki akurasi (0.9482) yang sama dengan beberapa model lain, precision-nya sedikit lebih rendah dibandingkan dengan recall.
3. **Random Forest With Normalization:** Model ini juga memiliki akurasi 0.9482, recall yang sangat tinggi.
4. **K-Nearest Neighbor With Normalization:** Precision dan recall rendah menunjukkan bahwa tidak efektif dalam dataset ini meskipun telah dinormalisasi.
5. **Gradient Boosting Classifier With Normalization:** Model ini menunjukkan akurasi yang sama dengan precision dan recall yang seimbang.

## MACHINE LEARNING ANALYSIS

1. **Random Forest** adalah model yang paling efektif dengan angka TP tertinggi dan FN yang sangat rendah.
2. **Logistic Regression dan Gradient Boosting Classifier** juga menunjukkan kinerja yang sangat baik, dengan kesalahan minimal dalam klasifikasi.
3. **Decision Tree** memiliki hasil yang baik, tetapi sedikit lebih banyak kesalahan dibandingkan model lain.
4. **KNN** menunjukkan kinerja yang poor, dengan banyak kesalahan klasifikasi, dan perlu penyesuaian untuk meningkatkan hasilnya. Penerapan normalisasi tidak memperbaiki kinerja KNN secara signifikan.



## MACHINE LEARNING

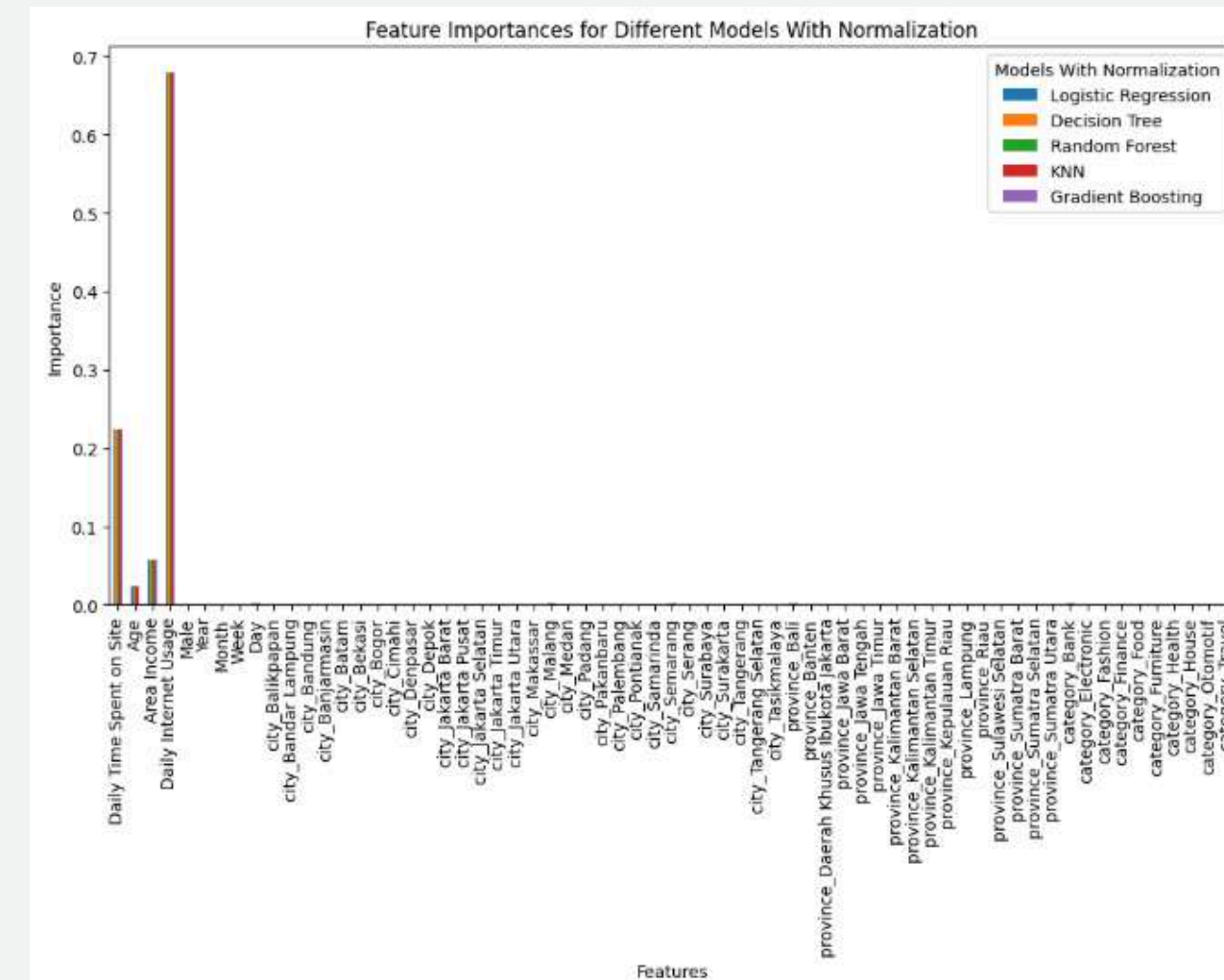
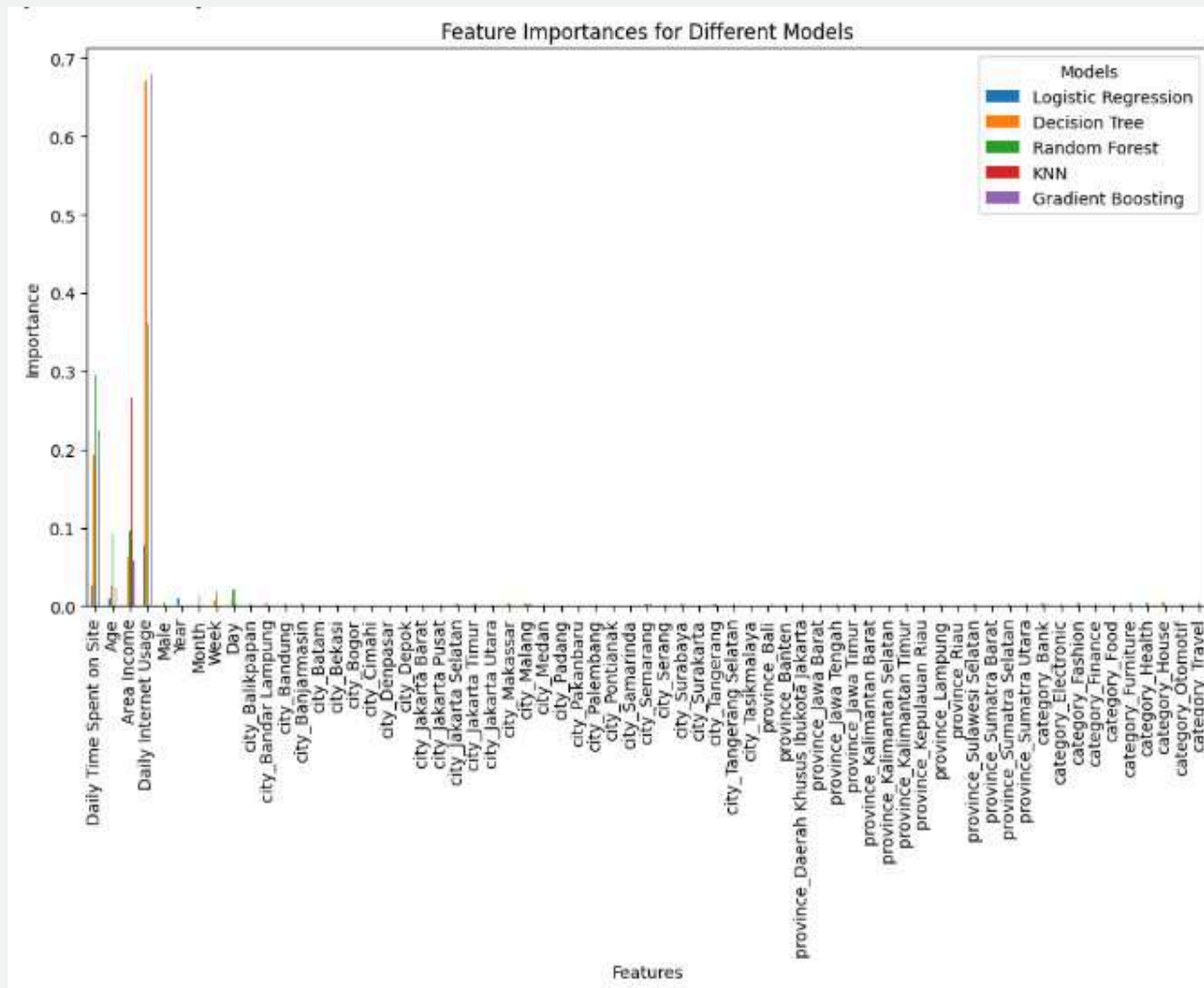
Berdasarkan hasil eksperimen pertama dan kedua, model machine learning menggunakan algoritma **Gradient Boosting Classifier** menunjukkan kinerja yang sangat baik dengan **precision mencapai 0.9697**.

Model ini mencatat jumlah **true positives (TP) yang tinggi**, serta **false positives (FP) dan false negatives (FN) yang rendah**, yang secara keseluruhan mencerminkan **akurasi tinggi dalam mengklasifikasi data**.

Hasil ini membuktikan efektivitas Gradient Boosting Classifier dalam mencapai keseimbangan optimal antara kemampuan mendeteksi kelas positif dan meminimalkan kesalahan klasifikasi.



## FEATURE IMPORTANCE



Pada eksperimen pertama dan kedua, feature importance ditunjukkan untuk data berikut:

1. Daily Time Spent on Site
2. Age
3. Area Income

Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)



# **BUSINESS RECOMMENDATION & SIMULATION**

The background features a faded city skyline. On the left, there is a circular graphic with diagonal stripes. On the right, there are four grey chevron arrows pointing to the right. A large, light-grey rectangular box serves as a backdrop for the main title.

## **BUSINESS RECOMMENDATION**

Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)



## DAILY TIME SPENT ON SITE

### INSIGHT

Pengguna yang menghabiskan lebih banyak waktu di situs lebih cenderung mengklik iklan.

### STRATEGI

- Tingkatkan engagement situs dengan konten yang relevan (artikel, video, atau fitur interaktif).
- Optimalkan pengalaman pengguna di halaman utama dan area dengan aktivitas tinggi.

### TINDAKAN

Tambahkan iklan di halaman yang memiliki waktu interaksi tertinggi.

## AGE

### INSIGHT

Usia tertentu mungkin memiliki konversi lebih tinggi

### STRATEGI

- Sesuaikan konten iklan berdasarkan usia. Misalnya, untuk kelompok usia muda, gunakan visual dinamis; untuk kelompok usia tua, fokus pada nilai produk.
- Gunakan segmentasi usia untuk kampanye pemasaran yang lebih bertarget.

### TINDAKAN

Alokasikan anggaran iklan lebih banyak pada kelompok usia dengan klik tertinggi.

## AREA INCOME

### INSIGHT

Pengguna dari area dengan pendapatan lebih tinggi cenderung mengklik iklan lebih banyak.

### STRATEGI

- Fokus pada wilayah dengan pendapatan tinggi untuk kampanye iklan premium.
- Untuk wilayah berpendapatan rendah, gunakan iklan promosi atau diskon.

### TINDAKAN

Lakukan geo-targeting pada platform iklan untuk wilayah strategis.



The background features a faded city skyline. On the left, there is a circular graphic with diagonal stripes. On the right, there are four grey chevron arrows pointing to the right. A large, light-grey rectangular box is centered on the slide, containing the main title.

## **BUSINESS SIMULATION**

Untuk selengkapnya, dapat melihat jupyter notebook [disini](#)

## BUSINESS SIMULATION

Pendekatan	Conversion Rate	Total Conversions	Cost	Revenue	Profit
Tanpa Model ML	50.36%	485	Rp14,445,000	Rp36,375,000	Rp21,930,000
Dengan Model ML	94.95%	914	Rp14,445,000	Rp68,550,000	Rp54,105,000

### Conversion Rate (Tingkat Konversi)

- **Tanpa Model ML:** Conversion rate sebesar 50.36%.
  - Ini berarti dari seluruh pengguna yang terlibat dalam kampanye pemasaran, sekitar 50.36% melakukan konversi, yaitu mengklik iklan.
- **Dengan Model ML:** Conversion rate meningkat menjadi 94.95%.
  - Model machine learning **berhasil meningkatkan conversion rate secara signifikan**, yaitu lebih dari 90% dari pengguna yang terlibat akhirnya melakukan konversi.

## BUSINESS SIMULATION

Pendekatan	Conversion Rate	Total Conversions	Cost	Revenue	Profit
Tanpa Model ML	50.36%	485	Rp14,445,000	Rp36,375,000	Rp21,930,000
Dengan Model ML	94.95%	914	Rp14,445,000	Rp68,550,000	Rp54,105,000

### Total Conversions (Total Konversi)

- **Tanpa Model ML:** Dengan 50.36% konversi, total konversi yang tercatat adalah 485 konversi.
- **Dengan Model ML:** Dengan peningkatan conversion rate menjadi 94.95%, total konversi meningkat drastis menjadi 914 konversi.
  - Peningkatan konversi ini menunjukkan bahwa model ML dapat mengidentifikasi dan menargetkan pengguna yang lebih cenderung mengklik iklan, meningkatkan hasil kampanye pemasaran.



## BUSINESS SIMULATION

Pendekatan	Conversion Rate	Total Conversions	Cost	Revenue	Profit
Tanpa Model ML	50.36%	485	Rp14,445,000	Rp36,375,000	Rp21,930,000
Dengan Model ML	94.95%	914	Rp14,445,000	Rp68,550,000	Rp54,105,000

### Cost (Biaya Pemasaran)

- **Tanpa Model ML dan Dengan Model ML:** Biaya pemasaran yang dikeluarkan adalah sama, yaitu Rp14,445,000 di kedua pendekatan.
  - Ini menunjukkan bahwa meskipun model ML meningkatkan conversion rate, **biaya pemasaran tetap konstan dan tidak dipengaruhi oleh penggunaan model.**

## BUSINESS SIMULATION

Pendekatan	Conversion Rate	Total Conversions	Cost	Revenue	Profit
Tanpa Model ML	50.36%	485	Rp14,445,000	Rp36,375,000	Rp21,930,000
Dengan Model ML	94.95%	914	Rp14,445,000	Rp68,550,000	Rp54,105,000

### Revenue (Pendapatan)

- **Tanpa Model ML:** Pendapatan yang dihasilkan dari 485 konversi adalah Rp36,375,000.
- **Dengan Model ML:** Pendapatan meningkat tajam menjadi Rp68,550,000 karena total konversi yang lebih tinggi (914 konversi).
  - Peningkatan pendapatan ini mencerminkan dampak positif dari peningkatan conversion rate yang dihasilkan oleh model machine learning.

## BUSINESS SIMULATION

Pendekatan	Conversion Rate	Total Conversions	Cost	Revenue	Profit
Tanpa Model ML	50.36%	485	Rp14,445,000	Rp36,375,000	Rp21,930,000
Dengan Model ML	94.95%	914	Rp14,445,000	Rp68,550,000	Rp54,105,000

### Profit (Laba)

- **Tanpa Model ML:** Laba yang diperoleh adalah Rp21,930,000.
- **Dengan Model ML:** Laba meningkat signifikan menjadi Rp54,105,000.
  - Peningkatan laba ini sangat signifikan, menunjukkan bahwa meskipun biaya pemasaran tetap sama, peningkatan konversi dan pendapatan yang dihasilkan dengan **menggunakan model ML menghasilkan keuntungan yang jauh lebih tinggi.**



## KESIMPULAN

- **Peningkatan Conversion Rate:** Penggunaan model machine learning memberikan **peningkatan signifikan dalam conversion rate dari 50.36% menjadi 94.95%**. Hal ini sangat berdampak pada total konversi, yang **meningkat dari 485 konversi menjadi 914 konversi**.
- **Peningkatan Pendapatan dan Laba:** Meskipun biaya pemasaran tetap sama, hasil yang lebih baik dari model ML menghasilkan pendapatan dan laba yang lebih tinggi. **Pendapatan meningkat lebih dari dua kali lipat, dan laba meningkat hampir dua kali lipat**.
- **Keunggulan Penggunaan Model ML:** Menerapkan machine learning dalam kampanye pemasaran dapat memberikan keuntungan yang sangat besar, terutama dalam hal mengoptimalkan target pengguna dan meningkatkan hasil konversi. Hal ini dapat menjadi strategi bisnis yang sangat efektif dalam **meningkatkan efektivitas pemasaran tanpa perlu meningkatkan biaya**.

Penerapan model ML dalam kampanye pemasaran berpotensi besar dalam **meningkatkan efisiensi dan memberikan ROI (Return on Investment) yang lebih tinggi**.



# THANK YOU