



**Mini Project Report:**

# **Predict Customer Clicked Ads**

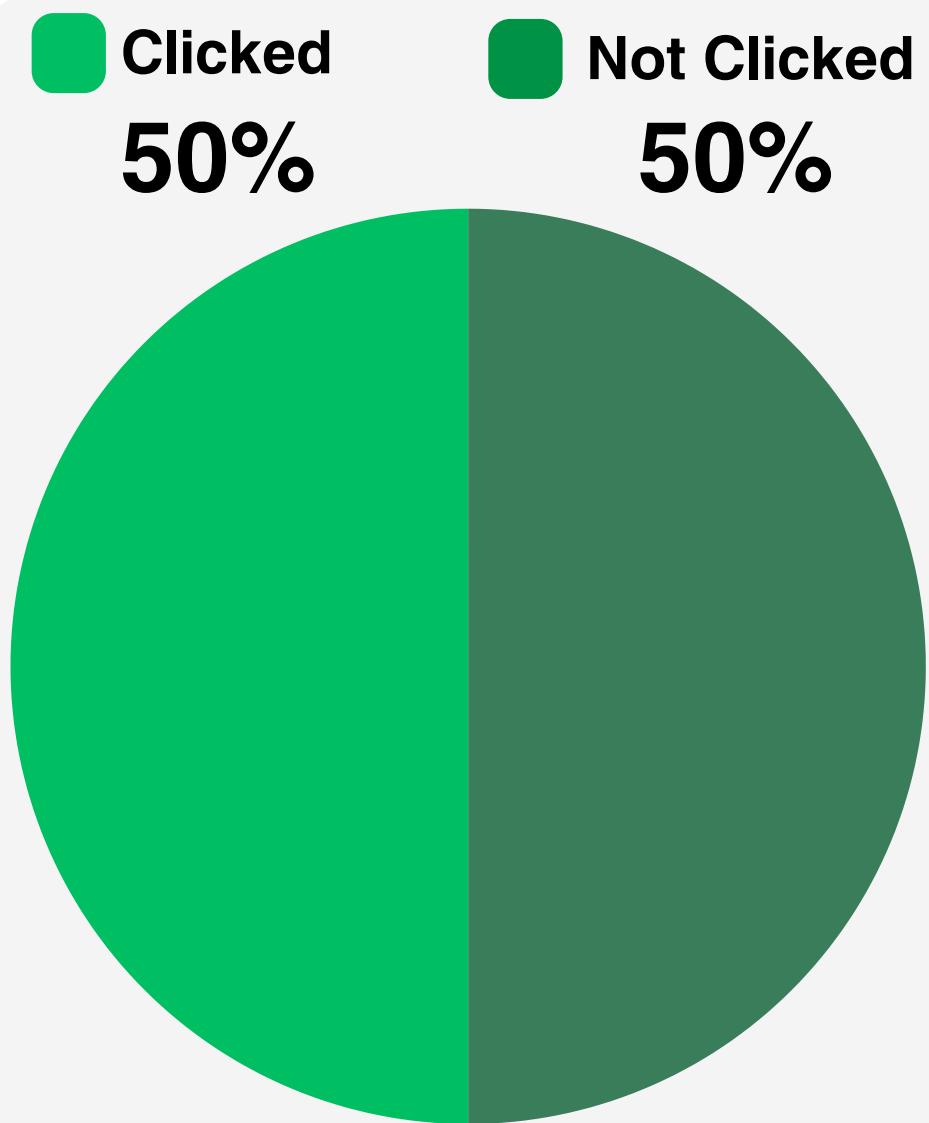
## **Classification by Using Machine**

### **Learning**

by Bhima Fairul Rifqi



# Business Overview



ABC Corp's ad campaign is underperforming, with only 50% of users clicking on promoted content. To improve results, they use machine learning to predict high-conversion users and optimize ad targeting.

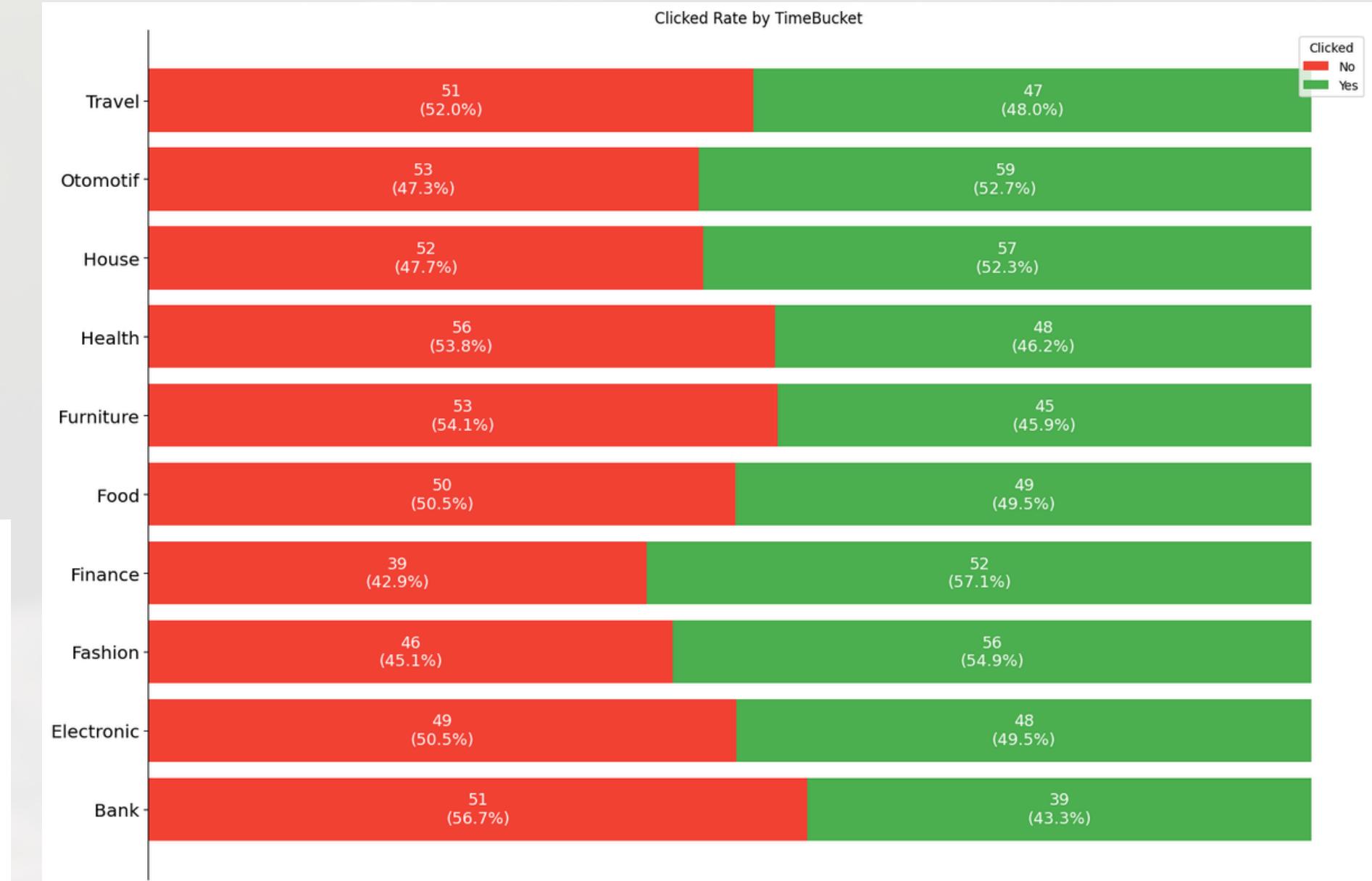


## Business Metric

ROI on Ads  
Promotion

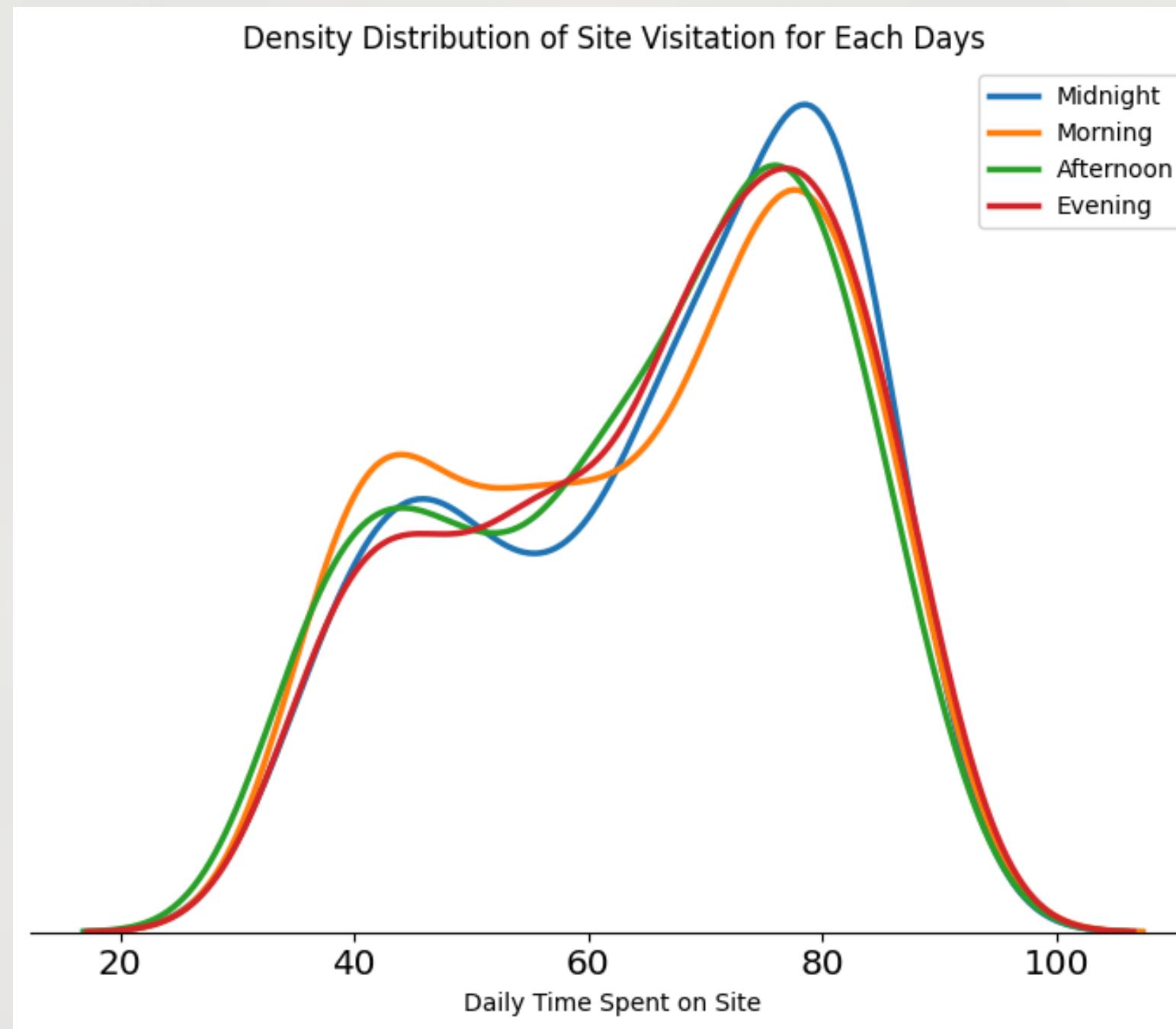
\*the focus target are labeled as 'Clicked on Ad' in the dataset.

# Exploratory Data Analysis



# Insights

1. User engagement is highest on Thursdays, especially during Evening and Midnight, where users spend the most time on site. This extended activity window boosts the likelihood of ad interaction.
2. Categories like Finance (57.1%) and Fashion (54.9%) outperform others, especially when aligned with high-engagement times and days. In contrast, Bank and Furniture struggle regardless of timing.



## Recommendation

1. Schedule key ad campaigns (especially for high-performing categories like Finance and Fashion) on Thursday evenings and midnights to maximize exposure and clicks.
2. Focus ad spend on top-performing categories during peak times (e.g., Finance ads on Sunday midnight). Reevaluate or redesign underperforming category ads, or run them in niche time slots for testing.

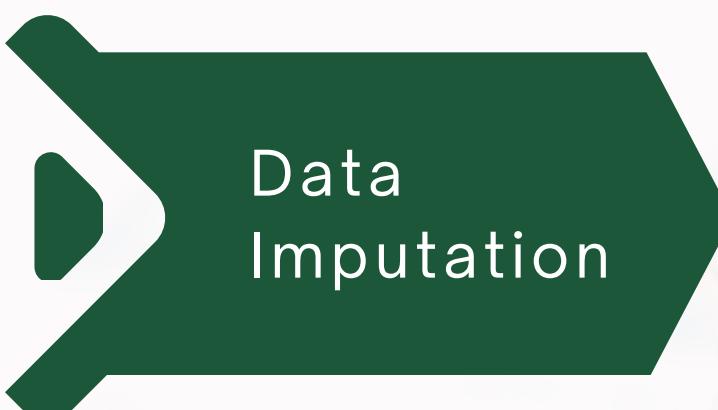
# About The Data: Complete Dataset

Feature Name	Feature Group	Type	To Do
Daily Time Spent on Site	Numeric	Numerical	None
Age	Numeric	Numerical	None
Area Income	Numeric	Numerical	None
Daily Internet Usage	Numeric	Numerical	None
Male	Binary	Categorical	Encode
Timestamp	Timestamp	Timestamp	Feature Engineering & Drop
Clicked on Ad (Target)	Binary	Categorical	Encode
city	Categoric	Categorical	Drop
Province	Categoric	Categorical	Drop
category	Categoric	Categorical	Drop
Weekday	Categoric	Categorical	None
TimeBucket	Categoric	Categorical	None

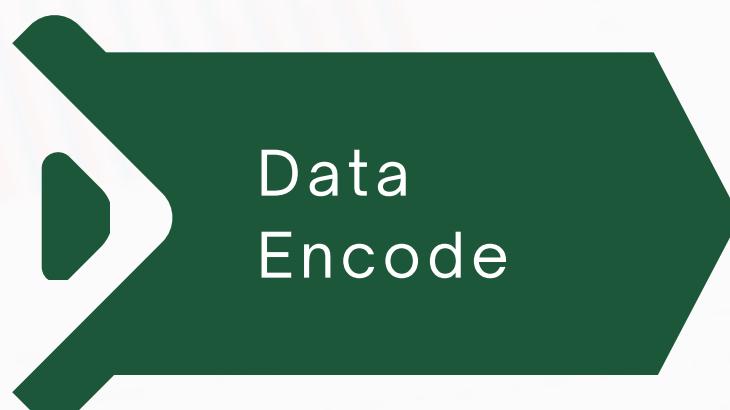
**Number of Observations: 1000**

# Data Preprocessing

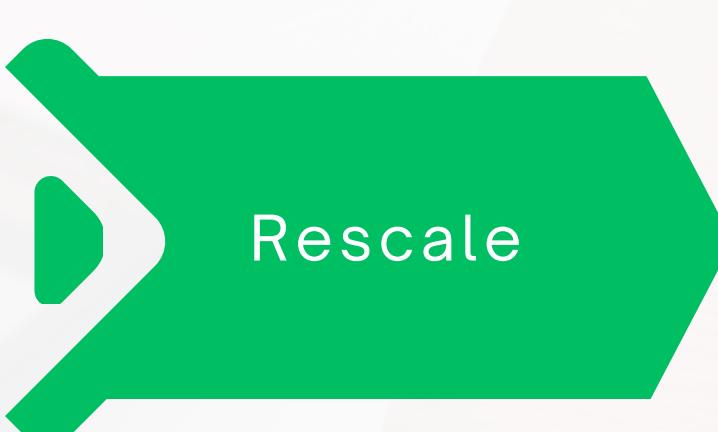
The data are splitted into train and test set on 80:20 ratio respectively.



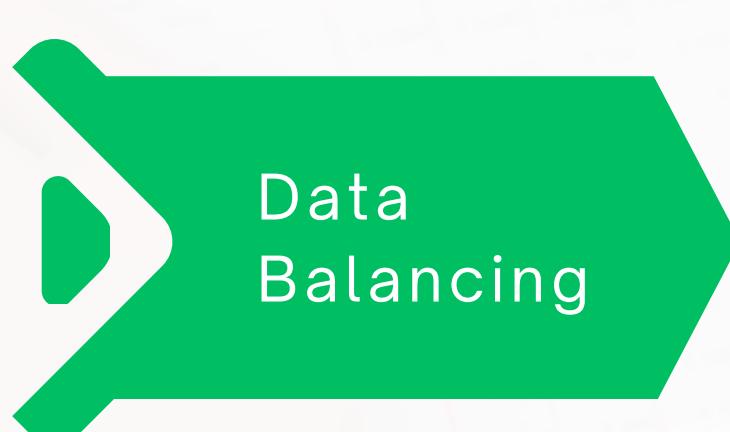
Mode imputation for categoric features  
Median imputation for numeric features



Binary Encoding for binary features  
Ordinal Encoding for ordinal categoric features



Rescaling the numeric features using robust scaler



Applied SMOTE on the target of the data train

# Applied Models and Its Evaluation

Model	Accuracy	Precision	Recall	F1 Score	ROC AUC	Best Threshold
logreg	0,96	0,9725	0,955	0,9636	0,9806	0,53
nb	0,955	0,9474	0,973	0,96	0,9816	0,39
xgb	0,945	0,9464	0,955	0,9507	0,9782	0,34
rf	0,945	0,963	0,9369	0,9498	0,9822	0,6
cat	0,94	0,9304	0,964	0,9469	0,979	0,1
gb	0,94	0,9381	0,955	0,9464	0,9791	0,1
knn	0,935	0,9298	0,955	0,9422	0,9744	0,33
ada	0,93	0,9217	0,955	0,9381	0,9787	0,47
dt	0,92	0,9279	0,9279	0,9279	0,9168	0,34

## Selected Model:

logreg (Logistic Regression based model), since it has the highest precision.

# Model Impact Analysis

	Predicted to Click	Predicted to Not Click
Actually Clicked	107 (TP)	4 (FN)
Actually Not Clicked	5 (FP)	84 (TN)
Total Employees	220	

Category	Count	Unit Value (Rp)	Total (Rp)	Type
Predicted Clicks (TP + FP)	112	2.000	224.000	Cost
Actual Clicks (TP)	107	20.000	2.140.000	Revenue
Opportunity Loss (FN)	4	20.000	80.000	Lost Revenue
False Positive (FP)	5	2.000	10.000	Wasted Cost

The campaign spends Rp 224,000 to serve ads but generates Rp 2,140,000 in revenue.



# Model Explainability Analysis

Feature	Coefficient
Age	1,673613
Weekday	0,588148
TimeBucket	-0,039832
Male	-0,528487
Area Income	-2,029972
Daily Time Spent on Site	-4,072866
Daily Internet Usage	-4,551744

## Insight

Older users and those browsing closer to the weekend are more likely to click ads, while males, high-income users, and heavy internet users are less likely to engage.

# Conclusions



**Best Time to do Promotion**  
Midnight.

**Impact of Model Implementation**  
Developed model successfully increased ROI.

**Main Factors of Clicking Tendency**  
Age and Weekdays.

# THANK YOU

**Project Resource:**

 [GitHub / Repo](#)

