

# ONLINE SHOPPER PURCHASING INTENTION

ec-Team





# Our team



**Mira Amelia Rosvita**



**Dania Dwi Pani**



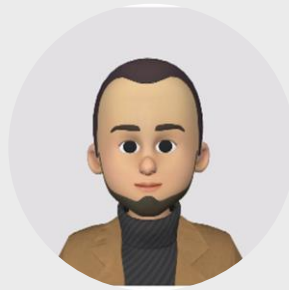
**Yanyan Gatot M.**



**Muthmainah**



**Haidar Aldy Eka N.**



**Yanuar Wachyudi**



**M. Raihan AKbar**

---



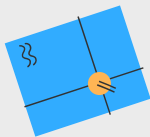
**ec-Team ?**





ec-Team sebagai tim data scientist berperan dalam memimpin dan mengelola proyek analisis data untuk memprediksi niat beli pelanggan dengan tujuan membantu meningkatkan transaksi dan revenue perusahaan e-commerce.





# Table of contents



**01**

**Problem  
Statement**

**02**

**EDA & Insight**

**03**

**Pre-Processing**

**04**

**Modelling**

**05**

**Recomendation**

---

01

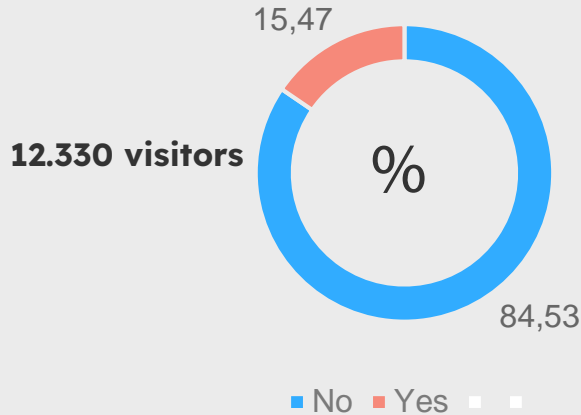
# Problem Statement





# Problem Statement

Revenue



**15,47%**

Customers yang mengunjungi Toko dan melanjutkan transaksi

**84,53%**

Customers yang mengunjungi Toko tetapi tidak melanjutkan transaksi

**Apa yang harus kami lakukan?**



# Goals, Objectives & Business Metrics



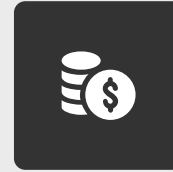
## Goals

Meningkatkan konversi penjualan dengan memahami pola dan niat beli pelanggan



## Objectives

- Mengembangkan model prediktif untuk memprediksi niat beli pelanggan secara akurat.
- Mengoptimalkan strategi penjualan untuk meningkatkan konversi penjualan dan kepuasan pelanggan.



## Business Metrics

Conversion Rate (%)

---



02

# EDA & Insight





# 12,330 row data dengan 18 feature

## 9 feature numerical

- Administrative
- Administrative\_Duration
- Informational
- Informational\_Duration
- ProductRelated
- ProductRelated\_Duration
- BounceRates
- ExitRates
- PageValues

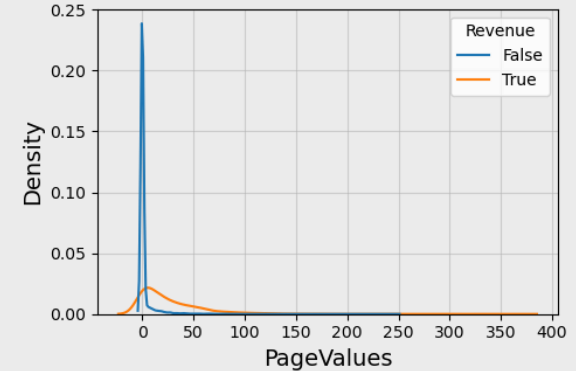
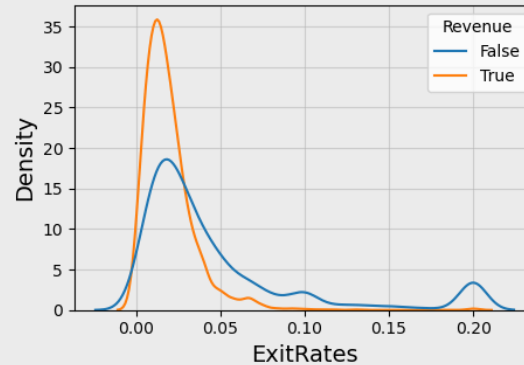
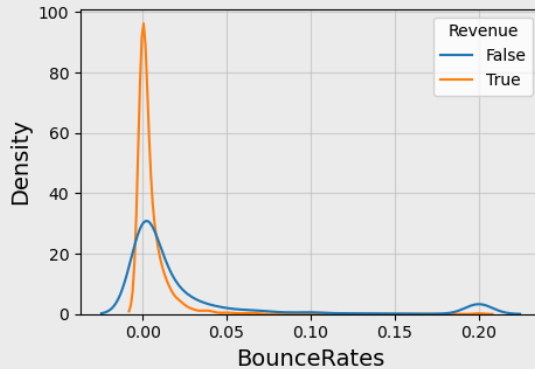
## 9 feature categorical

- SpecialDay
- Month
- OperatingSystems
- Browser
- Region,
- TrafficType
- VisitorType
- Weekend
- **Revenue → TARGET**





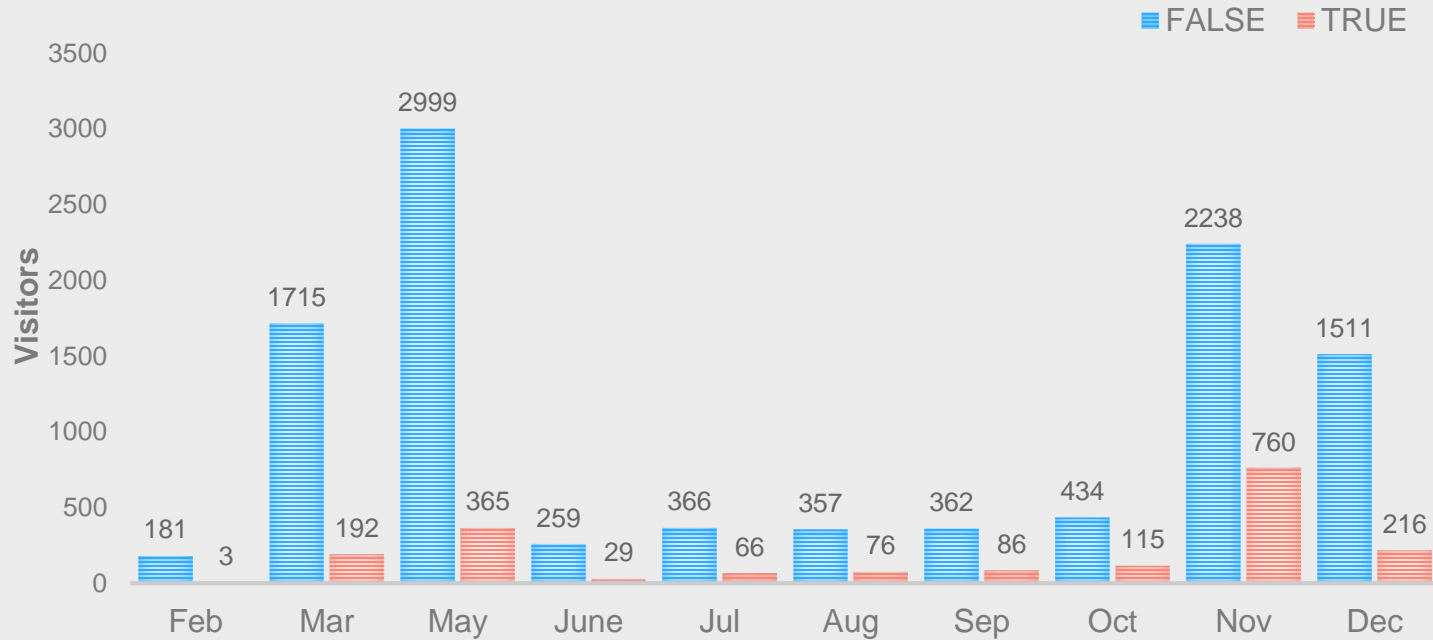
# EDA & Insight



Pengunjung dengan **Page Values** yang tinggi, **bounce rates** dan **exit rate** yang rendah cenderung melakukan **transaksi**

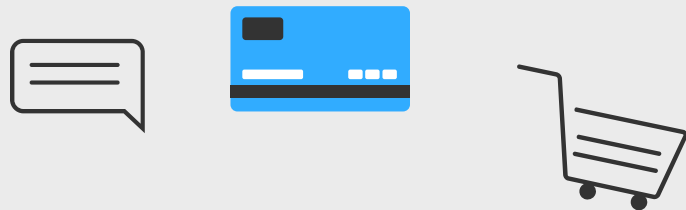
# EDA & Insight

MONTH



Pengunjung lebih cenderung melakukan transaksi pada bulan Mei dan November – Desember

\*Bulan Januari & April tidak ada data transaksi yang ditemukan

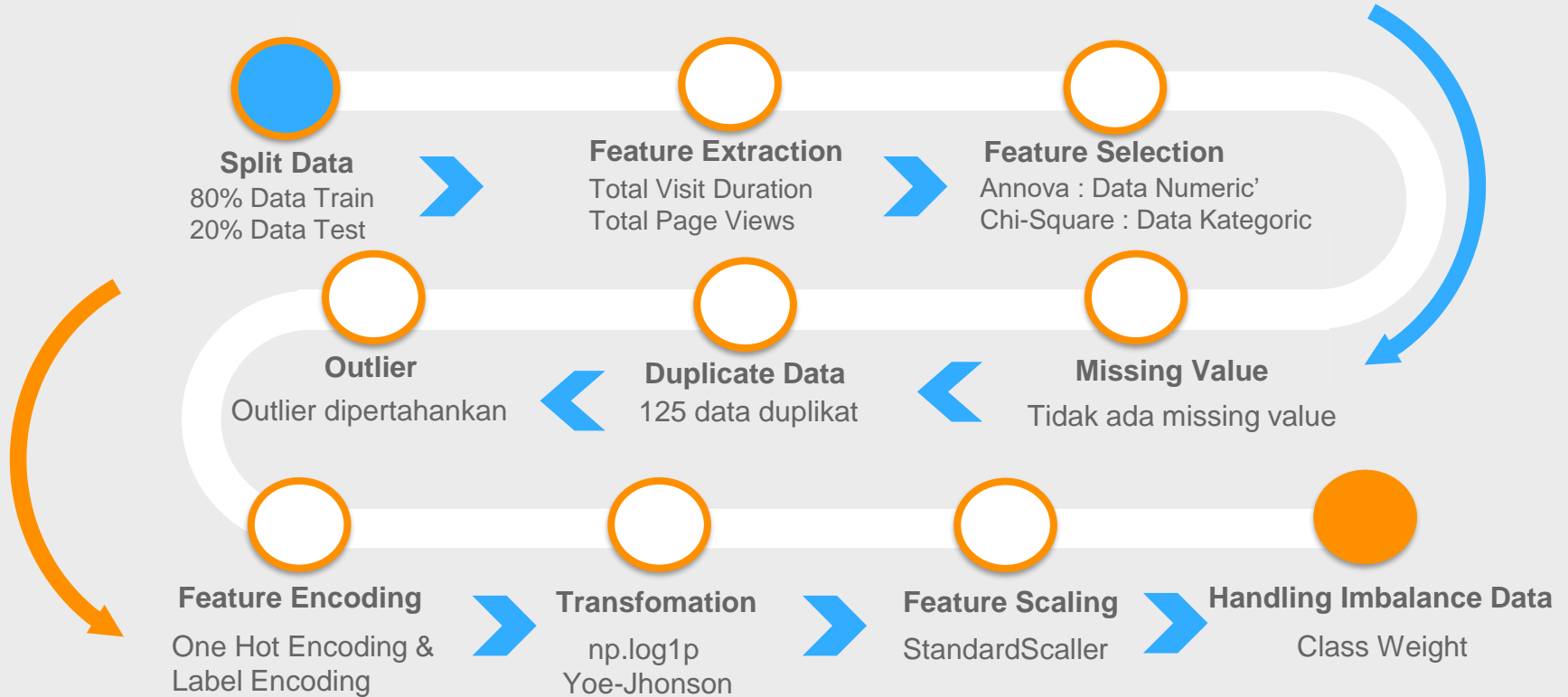


# Data Pre-Processing





# Data Pre-Processing

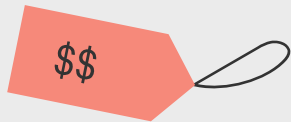


04

# Modelling



# Modelling



**Logistic Regression**

**Decision Tree**

**Random Forest**

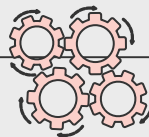
**KNN**

**SVC**

**Ada Boost**

**Gradient Boosting**

**XGBoost**





# Model Evaluation

Model	Recall		Precision		F1 Score		ROC-AUC	
	Train	Test	Train	Test	Train	Test	Train	Test
Logistic Regression	0.779	0.764	0.525	0.515	0.628	0.615	0.911	0.901
Decision Tree	1.000	0.574	1.000	0.564	1.000	0.569	1.000	0.746
Random Forest	1.000	0.499	1.000	0.768	1.000	0.605	1.000	0.923
KNN	0.489	0.387	0.800	0.659	0.607	0.487	0.940	0.794
SVC	0.841	0.759	0.608	0.550	0.706	0.637	0.946	0.905
Ada Boost	0.596	0.580	0.693	0.678	0.641	0.625	0.929	0.915
<b>Gradient Boosting</b>	<b>0.663</b>	<b>0.604</b>	<b>0.807</b>	<b>0.728</b>	<b>0.728</b>	<b>0.660</b>	<b>0.953</b>	<b>0.931</b>
XG Boost	0.945	0.589	0.998	0.701	0.971	0.640	0.999	0.924

Model **Gradient Boosting Classifier** lebih bagus berdasarkan ROC-AUC score



# Hyperparameter Tuning : Model Gradient Boosting Classifier



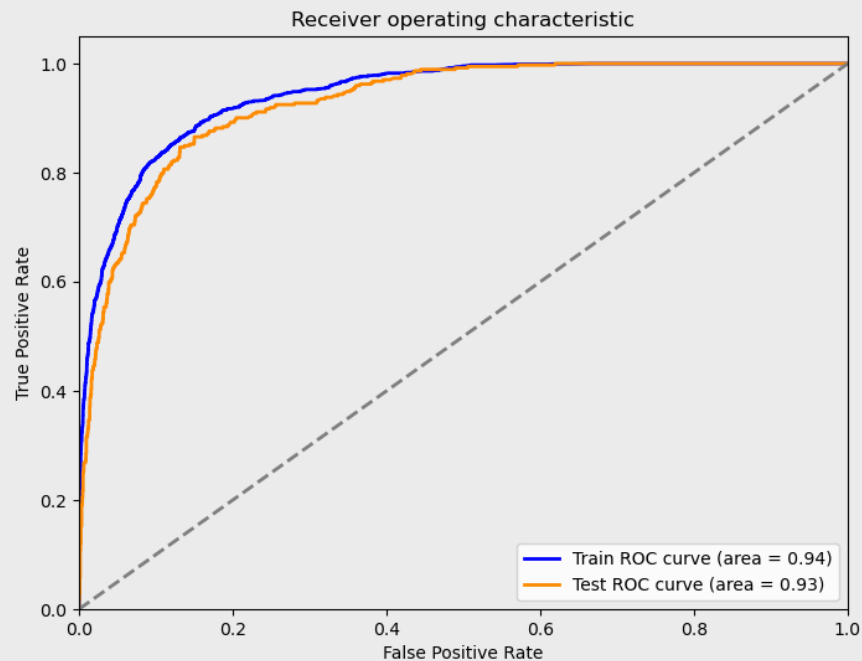
	Recall		Precision		F1 Score		ROC-AUC	
	Train	Test	Train	Test	Train	Test	Train	Test
Before Tuning	0.663	0.604	0.807	0.728	0.728	0.660	0.953	0.931
After Tuning	0.639	0.609	0.783	0.738	0.704	0.668	0.945	0.932

Model Gradient Boosting Classifier menjadi model yang best fit

---

# ROC AUC Curve

## Train and Data Test



Curve ROC AUC

# Confusion Matrix

Before Adjustment

Actual	No	Yes
	Predict	Predict
No	True Negative 2007	False Positive 87
Yes	False Negative 148	True Positive 224



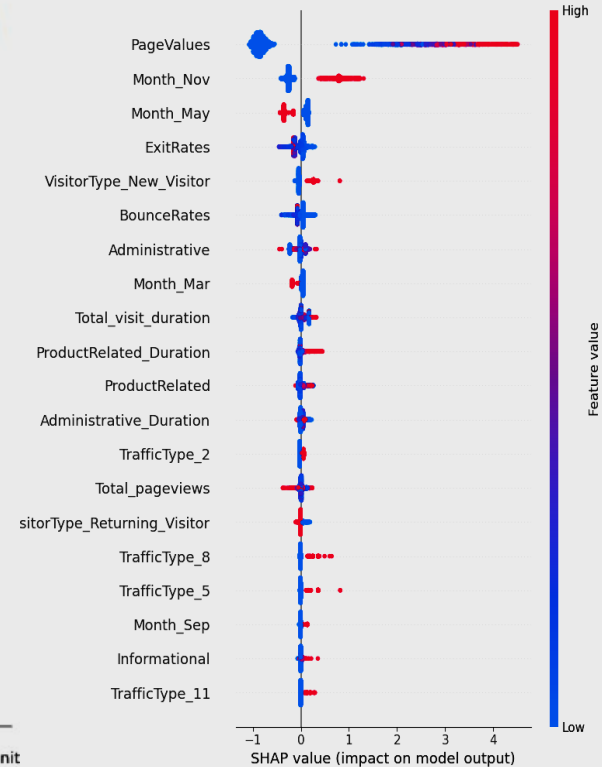
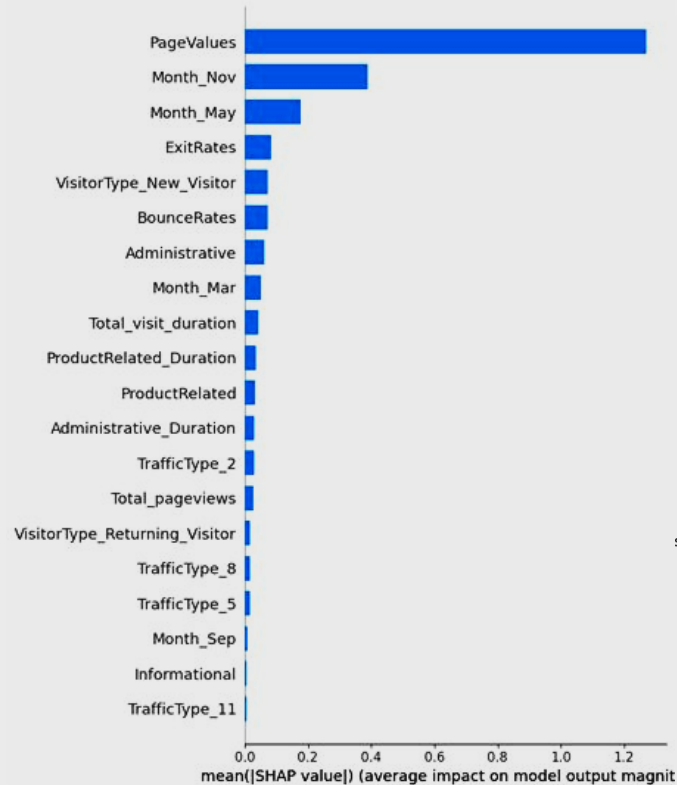
After Adjustment

Actual	No	Yes
	Predict	Predict
No	True Negative 1323	False Positive 771
Yes	False Negative 14	True Positive 358



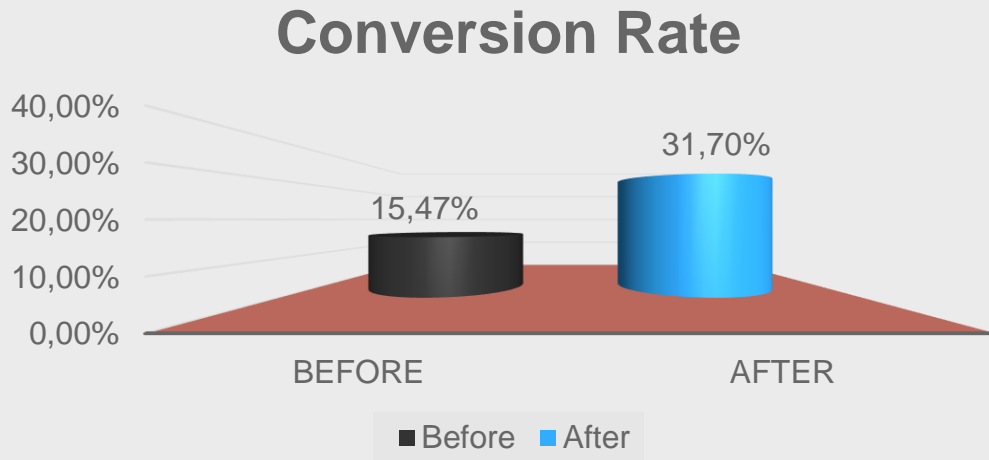
# Feature Importance

- Feature **Page Values** memiliki pengaruh paling besar terhadap hasil prediksi model.
- Feature **Month Nov & May**, **Exit Rates** dan **Bounce Rates** juga merupakan fitur penting yang cukup memiliki pengaruh yang signifikan.





# Potential Conversion Rate



15,47% Conversion Rate



Conversion Rate 31.7%



# Simulasi Bisnis



15,47% Conversion Rate



Conversion Rate 31.7%

- Total pengunjung dalam setahun: 12.330
- Rata-rata per transaksi: Rp 200.000
- Tingkat konversi sebelum menggunakan model: 15,47%

- Tingkat konversi setelah menggunakan model: 31,7%
- Biaya penggunaan operasional model per tahun: Rp 50.000.000
- (-) **Biaya investasi awal** untuk membeli dan mengimplementasikan model:  
Rp 200.000.000
- (-) Biaya cadangan atau darurat: Rp 20.000.000 (10% dari biaya investasi awal)
- (-) **Total biaya:** Rp 50.000.000 + Rp 200.000.000 + Rp 20.000.000 = Rp 270.000.000

$$\frac{\text{Rp.200.000} * 1908}{\text{Rp 381.600.000/thn}} \times$$

$$\frac{\text{Rp.200.000} * 3908}{\text{Rp 781.600.000/thn}} \times$$

$$\begin{aligned} \text{Keuntungan bersih} &= \text{Rp 781.600.000} - \text{Rp 270.000.000} \\ &= \text{Rp 511.600.000} \end{aligned}$$

# Simulasi Bisnis



## ROI (Return on Investment)

$$\begin{aligned}\text{ROI} &= (\text{Keuntungan Bersih} / \text{Biaya Investasi Awal}) \times 100\% \\ &= (\text{Rp } 511.600.000 / 200.000.000) \times 100\% \\ &= 2558\%\end{aligned}$$

## BEP (Break Even Point)

$$\begin{aligned}\text{BEP}_{\text{jumlah transaksi}} &= \text{Biaya Tetap} / (\text{Harga Jual per unit} - \text{Biaya Variabel per unit}) \\ &= \text{Rp } 270.000.000 / \text{Rp } 200.000 \\ &= 1325 \text{ Transaksi}\end{aligned}$$

$$\begin{aligned}\text{BEP}_{\text{pendapatan}} &= \text{BEP}_{\text{jumlah transaksi}} \times \text{Harga Jual per unit} \\ &= 1325 \times \text{Rp } 200.000 \\ &= \text{Rp } 270.000.000\end{aligned}$$

## PP (Payback Period)

$$\begin{aligned}\text{PP}_{\text{tahun}} &= \text{Biaya Investasi Awal} / \text{Keuntungan Bersih per tahun} \\ \text{PP}_{\text{bulan}} &= \text{PP}_{\text{tahun}} \times 12 \text{ bulan} \\ \text{PP}_{\text{hari}} &= \text{PP}_{\text{bulan}} \times 30 \text{ hari}\end{aligned}$$

$$\begin{aligned}\text{PP}_{\text{tahun}} &= \text{Rp } 200.000.000 / \text{Rp } 511.600.000 \\ &= 0,39 \text{ tahun} \\ \text{PP}_{\text{bulan}} &= 0,39 \times 12 \text{ bulan} \\ &= 4,7 \text{ bulan} \\ \text{PP}_{\text{hari}} &= 4,7 \times 30 \text{ hari} \\ &= 140 \text{ hari}\end{aligned}$$



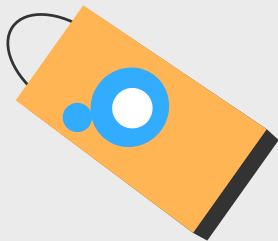


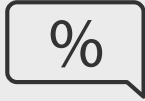
05

# Recommendation

Rakamin Academy Data Science | Batch 30

Lalu, rekomendasi bisnis apa yang bisa ditawarkan?





# Recommendation



## Page Values

Page Values adalah nilai rata-rata untuk halaman yang dikunjungi pengguna sebelum menyelesaikan konversi atau transaksi e-Commerce



SEO

(Search Engine Optimization)

**50% OFF**  
COUPON

**30% OFF**  
COUPON

**10% OFF**  
COUPON

**50% OFF**  
COUPON

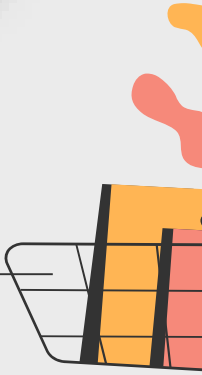
**30% OFF**  
COUPON

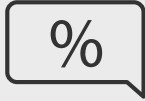
**10% OFF**  
COUPON

Voucher per Kategori

Page Values = ((Transaksi Revenue + Total Goal Revenue) / Unique Pageviews)

sumber data: <https://www.datadriven.com/page-value-google-analytics/>





# Recommendation



## Month

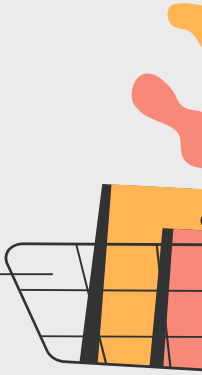
Month adalah Bulan ketika pengguna mengunjungi situs web.

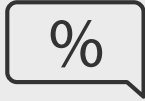


Promo Black Friday Bulan November



Promo Event Akhir Tahun



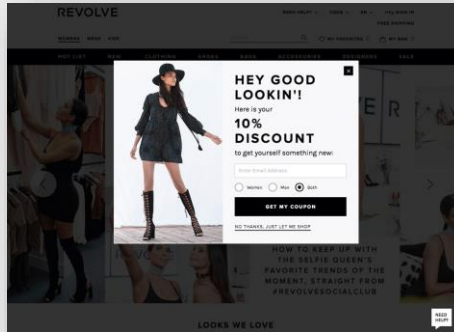


# Recommendation



## Exit Rates

Exit Rates adalah persentase pengunjung yang meninggalkan situs dari halaman yang sedang dikunjungi.



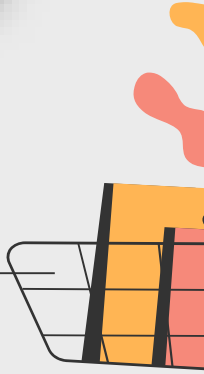
Pop-up Exit Intent



Live Chat

Exit rate = exits / pageviews

sumber data: <https://www.hotjar.com/google-analytics/glossary/exits/>



%

# Recommendation



## Bounce Rates

Bounce Rates adalah persentase pengunjung yang masuk ke situs dan kemudian keluar daripada terus melihat halaman lain dalam situs yang sama.



Meningkatkan keamanan situs

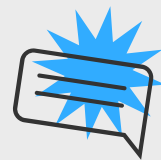


Tampilan situs web yang menarik dan mudah digunakan

Bounce Rates =  $\frac{\text{jumlah kunjungan halaman tunggal}}{\text{jumlah visitor}} \times 100\%$

sumber data: <https://redcomm.co.id/knowledges/bounce-rate-dan-seberapa-penting-dampaknya-bagi-website-bisnis?readmore=true>





THANK YOU

---