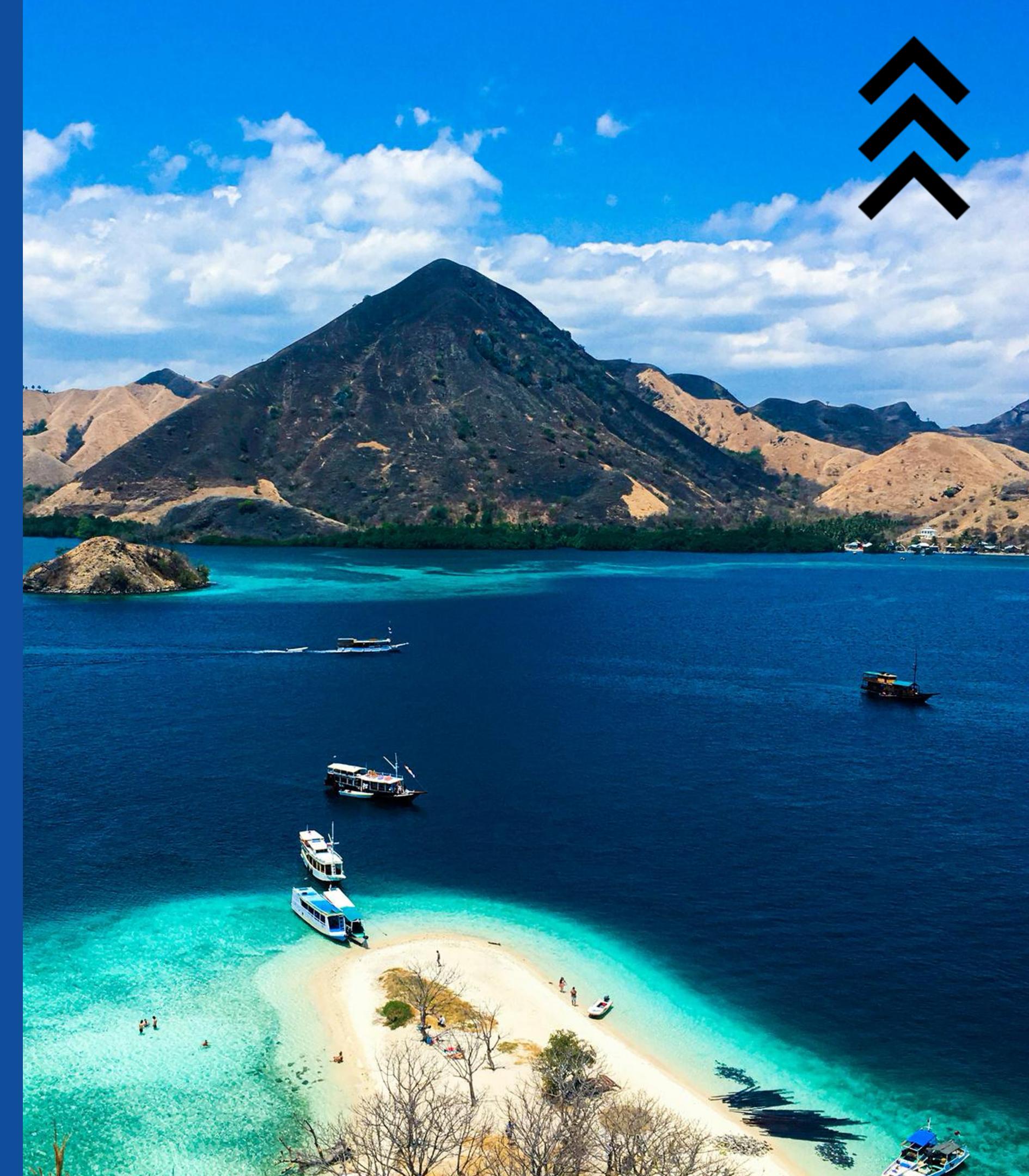
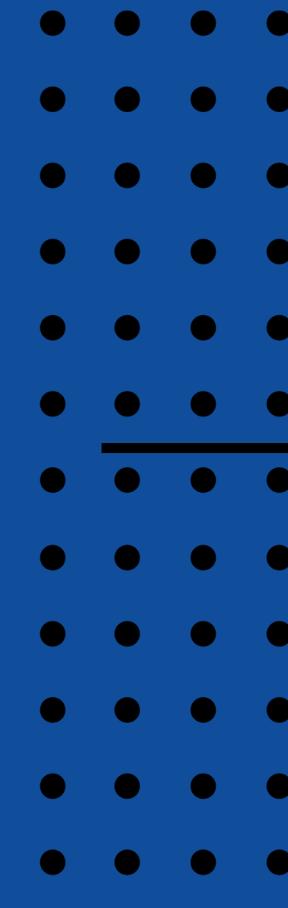


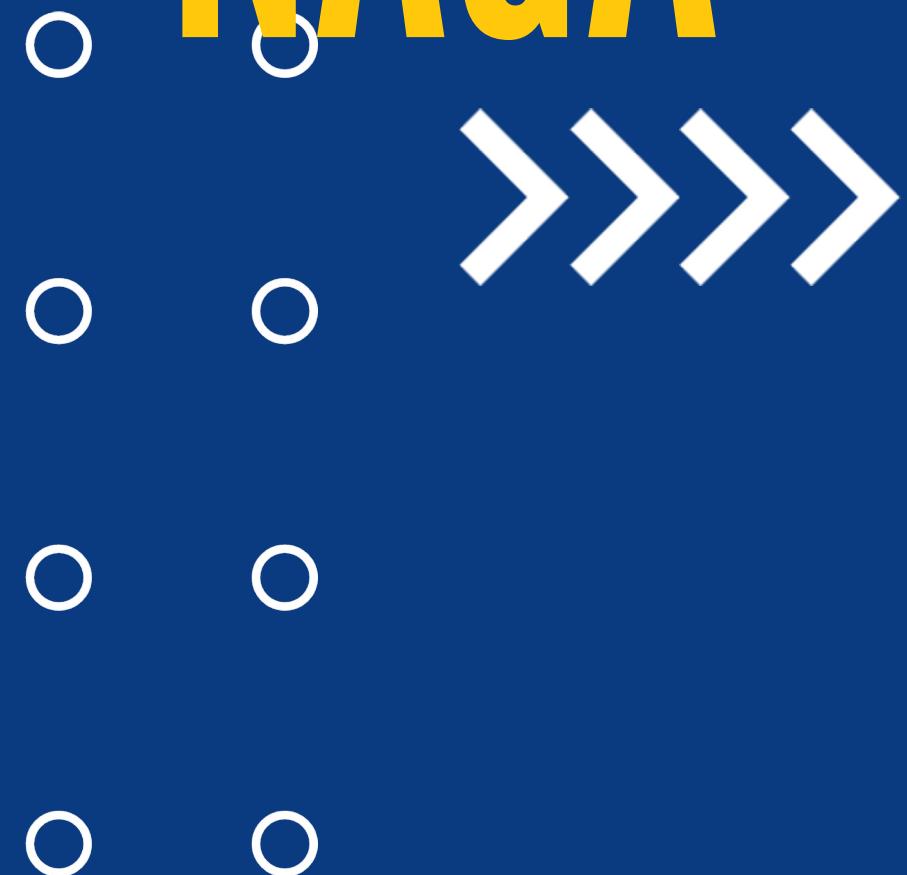


 TRIPS&TRAVEL.COM
7 BOLA NAGA

HOLIDAY PACKAGE PREDICTION



7 BOLA NAGA





ABOUT US

Perusahaan Trips&Travel.com merupakan suatu perusahaan dengan model bisni berbasis customer. Saat ini kami memiliki lima tipe paket tersedia; Basic, Standard, Deluxe, Super Deluxe, dan King. Bersamaan dengan paket yang sudah ada kami juga berencana untuk meluncurkan paket baru yang bernama 'Wellness Tourism Package'.

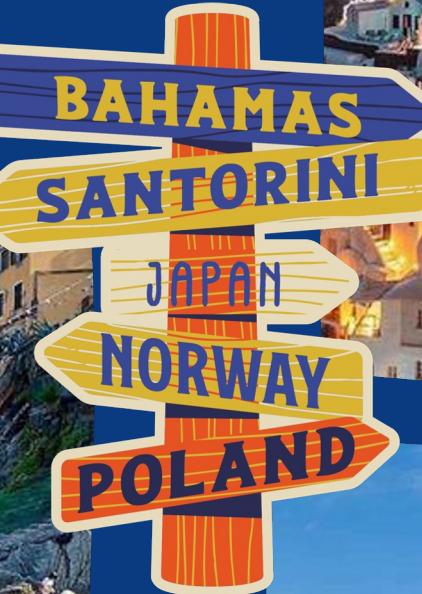


TABLE OF CONTENTS



Business Statement



Preprocessing



EDA



Modeling



Business Recomendation

BUSINESS STATEMENT

Understanding Business Problem



TRIPS&TRAVEL.CO

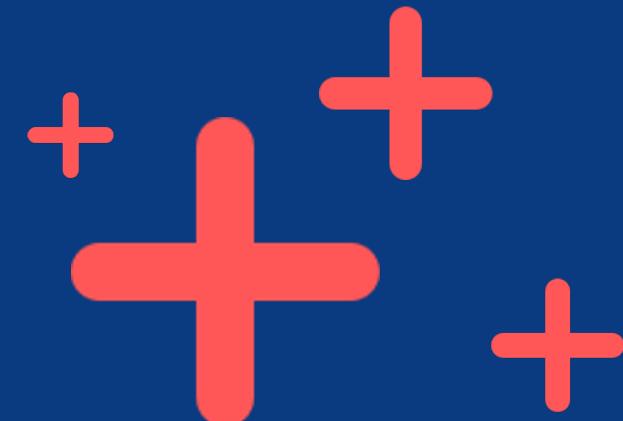
M



Problem Statement

Hanya 19% customer yang melakukan pembelian package pada tahun lalu.

Ini menunjukkan bahwa targeting pelanggan yang tidak tepat atau bahkan tidak ada targeting pelanggan yang dilakukan, mengakibatkan tingginya pengeluaran biaya marketing



Conversion Rate

19%



Goal

- effective customer targeting berdasarkan karakter customer yang dapat menurunkan biaya marketing

Objective

Membuat model machine learning yang dapat memprediksi customer yang berpotensi membeli package dengan memanfaatkan data pelanggan yang tersedia di perusahaan.



Business metric

★ Conversion Rate

EDA

Exploratory Data Analysis



TRIPS&TRAVEL.COM

EXPLORATORY DATA ANALYSIS

```
df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4888 entries, 0 to 4887
Data columns (total 20 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   CustomerID      4888 non-null    int64  
 1   ProdTaken        4888 non-null    int64  
 2   Age              4662 non-null    float64 
 3   TypeofContact   4863 non-null    object  
 4   CityTier         4888 non-null    int64  
 5   DurationOfPitch 4637 non-null    float64 
 6   Occupation       4888 non-null    object  
 7   Gender            4888 non-null    object  
 8   NumberOfPersonVisiting 4888 non-null  int64  
 9   NumberofFollowups 4843 non-null    float64 
 10  ProductPitched   4888 non-null    object  
 11  PreferredPropertyStar 4862 non-null  float64 
 12  MaritalStatus     4888 non-null    object  
 13  NumberOfTrips     4748 non-null    float64 
 14  Passport          4888 non-null    int64  
 15  PitchSatisfactionScore 4888 non-null  int64  
 16  OwnCar            4888 non-null    int64  
 17  NumberOfChildrenVisiting 4822 non-null  float64 
 18  Designation        4888 non-null    object  
 19  MonthlyIncome      4655 non-null    float64 
dtypes: float64(7), int64(7), object(6)
memory usage: 763.9+ KB
```

4888 rows
20 columns ➤



Categorical (object):

- **TypeofContact**
- **Occupation**
- **Gender**
- **ProductPitched**
- **MaritalStatus**

Numerical (Int64, Float64):

- **CustomerID**
- **ProdTaken (Target)**
- **Age**
- **CityTier**
- **DurationOfPitch**
- **NumberOfPersonVisiting**
- **NumberofFollowups**
- **PreferredPropertyStar**
- **NumberOfTrips**
- **Passport**
- **PitchSatisfactionScore**
- **OwnCar**
- **NumberOfChildrenVisiting**
- **MonthlyIncome**

EDA

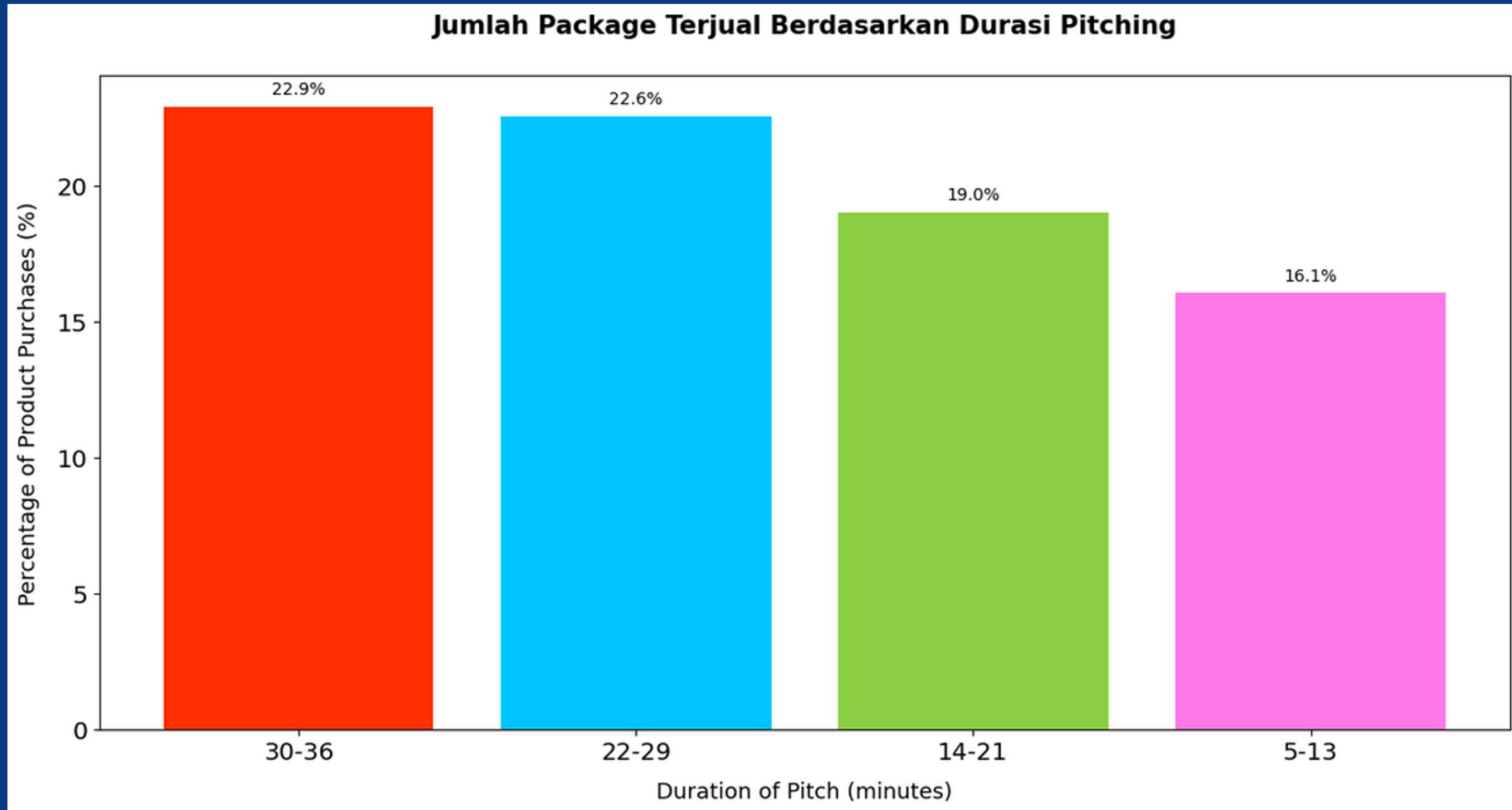
Business Insight



TRIPS&TRAVEL.COM

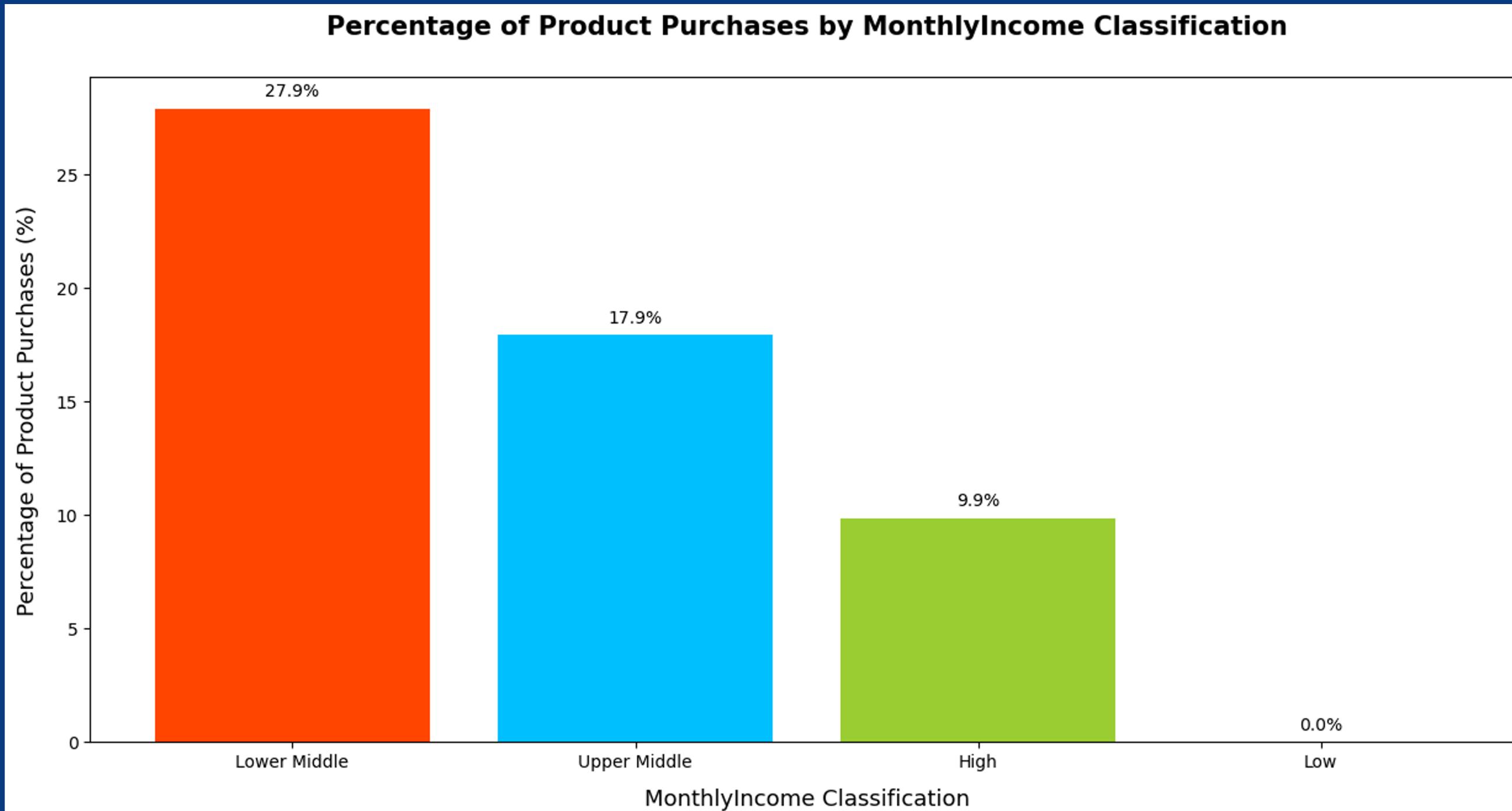
BUSINESS INSIGHT

Jumlah Package Terjual Berdasarkan Durasi Pitching



Customer yang memiliki duration of pitch > 22 menit memiliki
conversion rate yang lebih tinggi

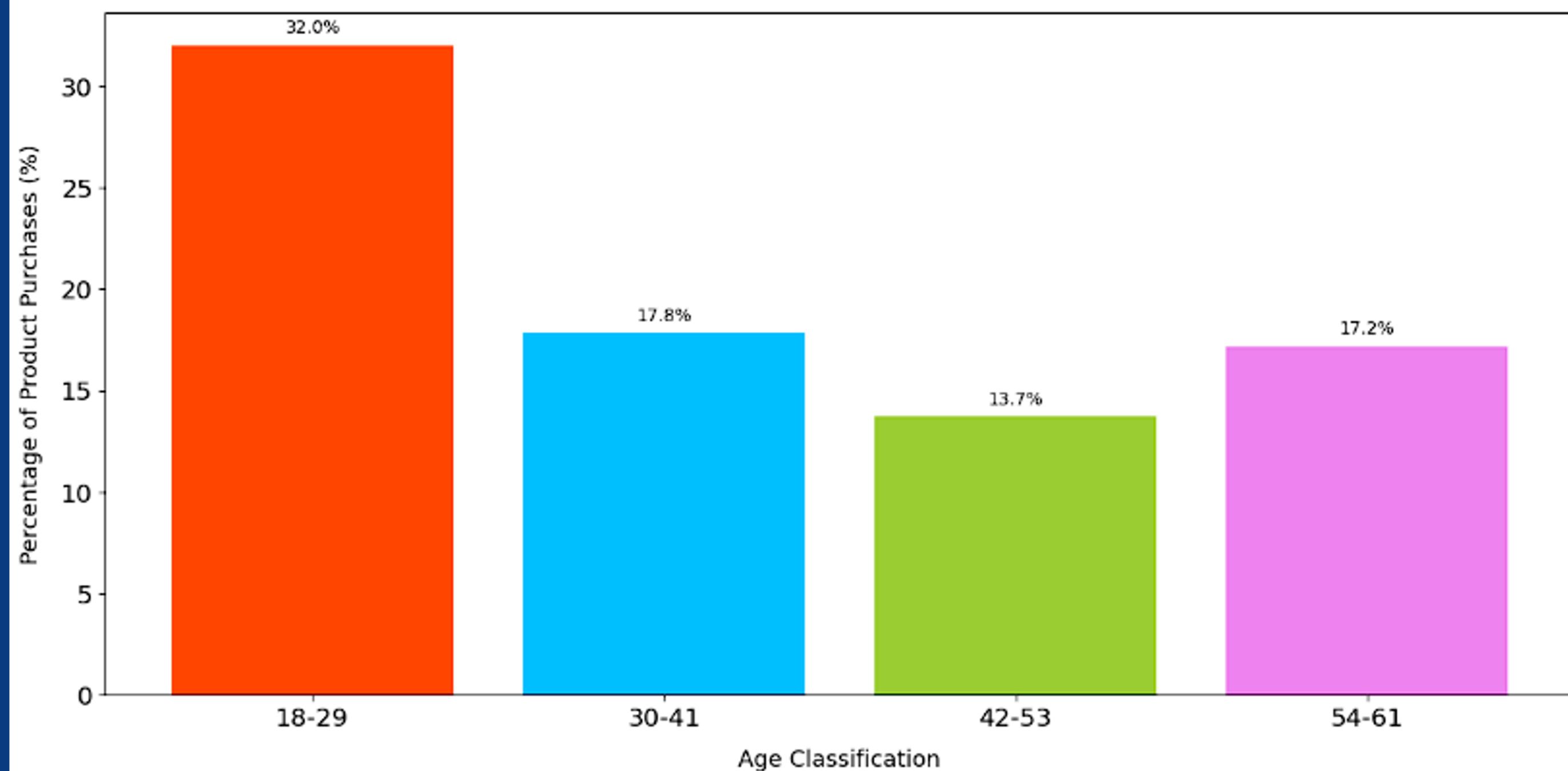
BUSINESS INSIGHT



Mayoritas customer yang Memiliki Monthly Income Lower Middle (\$11.000 - \$20.000) memiliki conversion rate paling besar

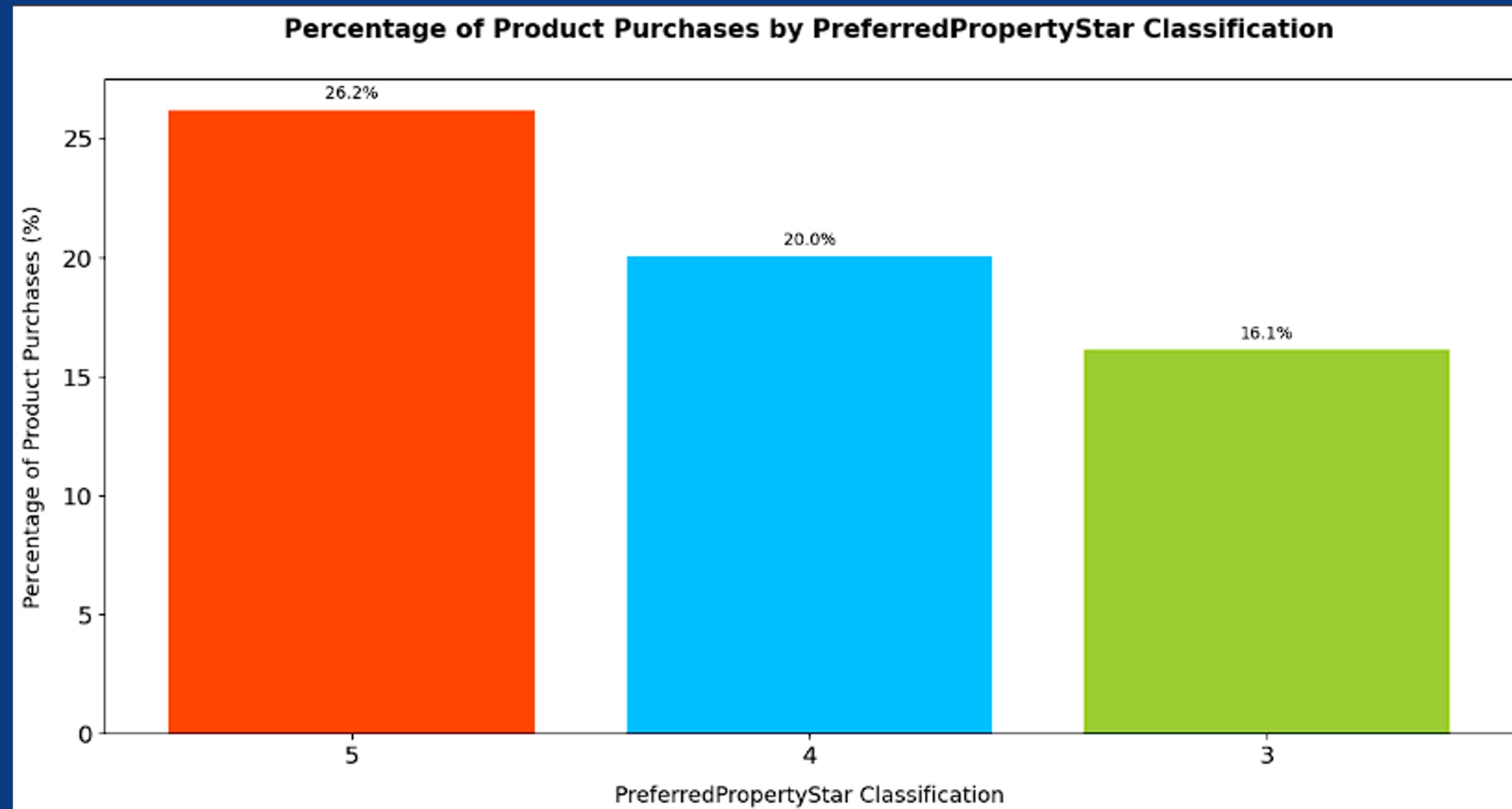
BUSINESS INSIGHT

Percentage of Product Purchases by Age Classification



customer yang memiliki nilai conversion rate tertinggi terdapat pada
umur 18-29 tahun dengan conversion rate sebesar 32%

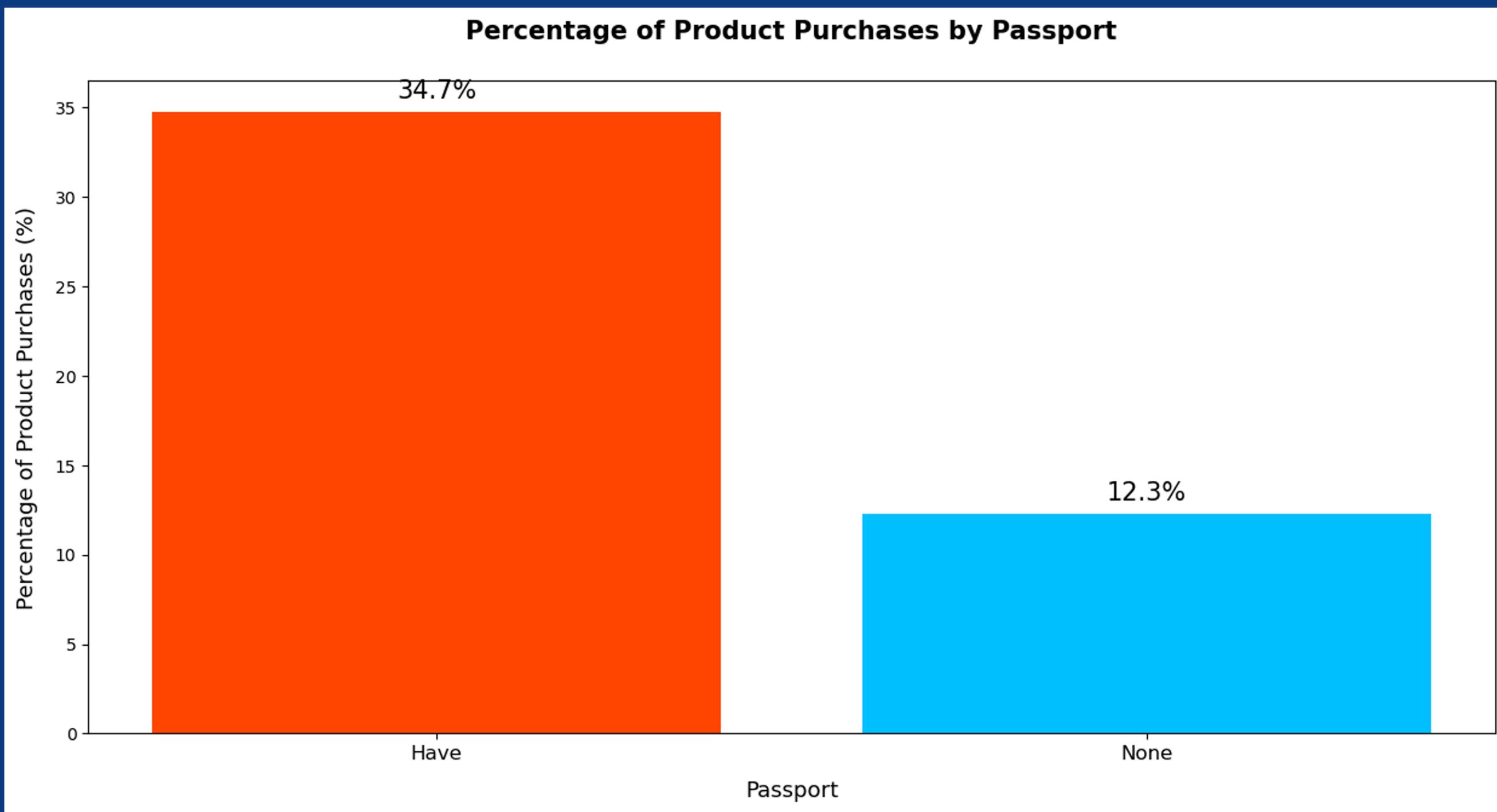
BUSINESS INSIGHT



Kelompok PreferredPropertyStar paling tinggi dimiliki oleh kelompok
5.0 dengan presentase 26.2%

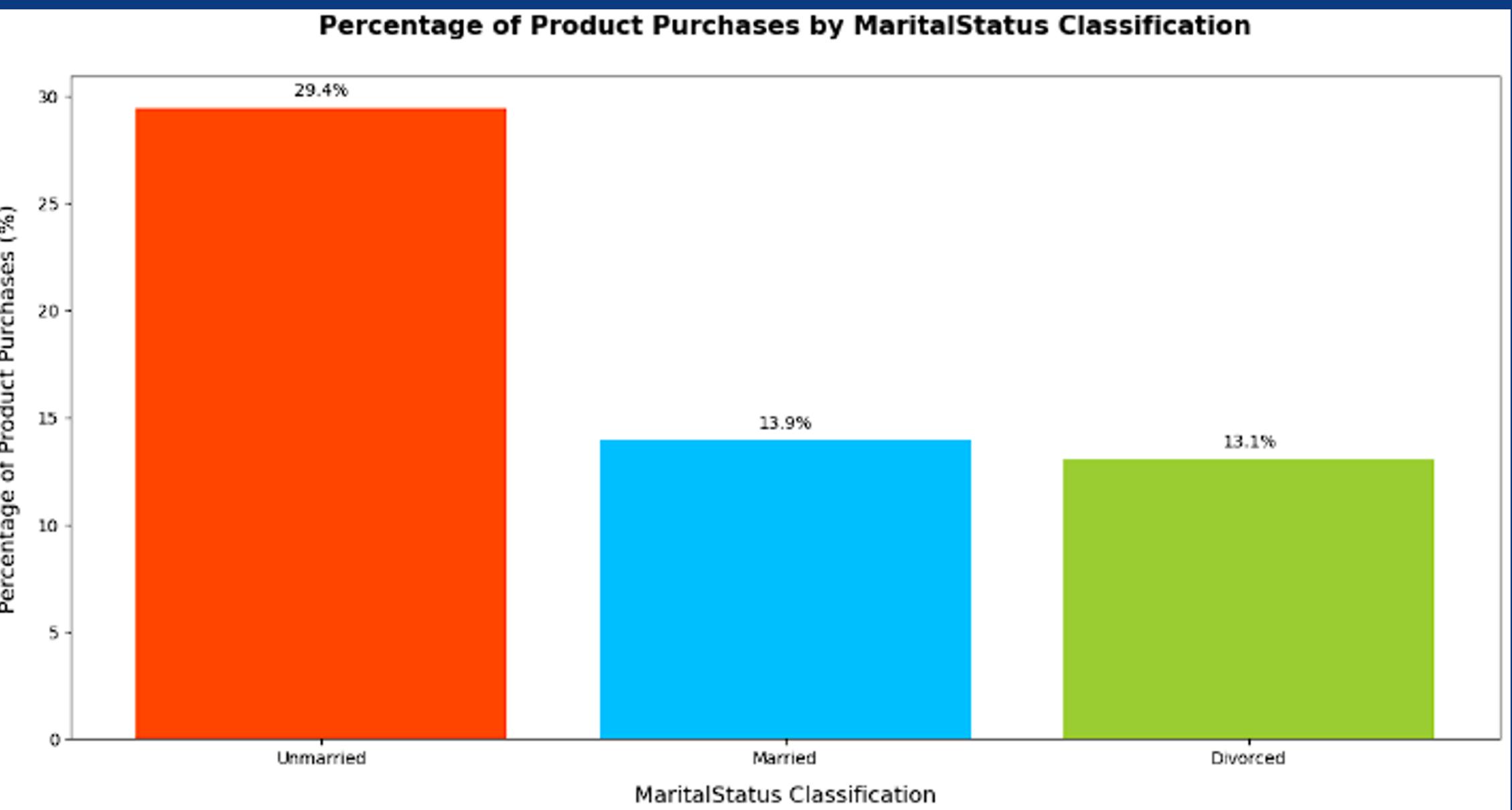
BUSINESS INSIGHT

Percentage of Product Purchases by Passport



Pembelian produk berdasarkan kepemilikan Passport paling tinggi dimiliki oleh pemegang Passport yaitu sebesar 34.7%

BUSINESS INSIGHT



- Pelanggan yang belum ada status menikah (**unmarried**) cenderung lebih tertarik dan lebih memungkinkan untuk membeli package/produk

PREPROCESSING

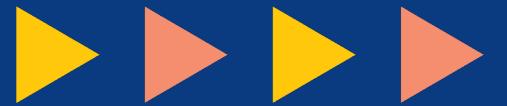
process data before modelling



TRIPS&TRAVEL.COM



DATA CLEANSING



Handling Missing Values

- Age : Mean
- TypeOfContact : Self enq
- DurationOfPitch : Median
- NumberOfFollowups : Median
- PreferredProperty : Modus
- NumberOfTrips : Median
- NumberOfChildren : Modus
- MonthlyIncome : Median

Handling Inconsistent Data

- Gender : Fe Male -> Female
- MaritalStatus : Single -> Unmarried
- Age : dtype = Float -> Int

Handling Duplicate Data

Pada dataset ini tidak temukannya
duplicate data

FEATURE ENGINEER

outlier & Feature Transformation

- Outlier menggunakan z-score

- Age
 - DurationOfpitch
 - NumberOfTrips
 - MonthlyIncome

- Feature Transformation

- log: DOP_norm, MI_std,

- NumberOfTrips

- Normalisasi: Age

- Standardisasi: DOP_norm, MI_std

Feature Encoding

- Label Encoding
 - TypeofContract
 - Gender
 - ProductPitched
 - Designation
- ONEHOT Encoding
 - Occupation
 - MaritalStatus

Feature Selection

- Fclassif -> Categorical
- ChiSquare-> Numerical



- Split Data

Split data dilakukan dengan ratio 80:20

Split Data & Imbalance Class

- Imbalance Class

SMOTE



MODELLING

Supervised Model to predict prospective customer



TRIPS&TRAVEL.COM

MODEL COMPARISON

Model	Accuracy	Precision	Recall	F1-Score
Decison Tree	0.90	0.95	0.92	0.94
XGBoost	0.90	0.91	0.97	0.94
Logistic Regression	0.83	0.88	0.91	0.90
Random Forest	0.90	0.91	0.97	0.94
Support Vector Machine	0.85	0.88	0.94	0.91

- Dengan mencoba 5 algoritma model di atas, nilai akurasi dan precision paling tinggi diperoleh oleh Decision Tree dengan akurasi 90% dan precision 95%

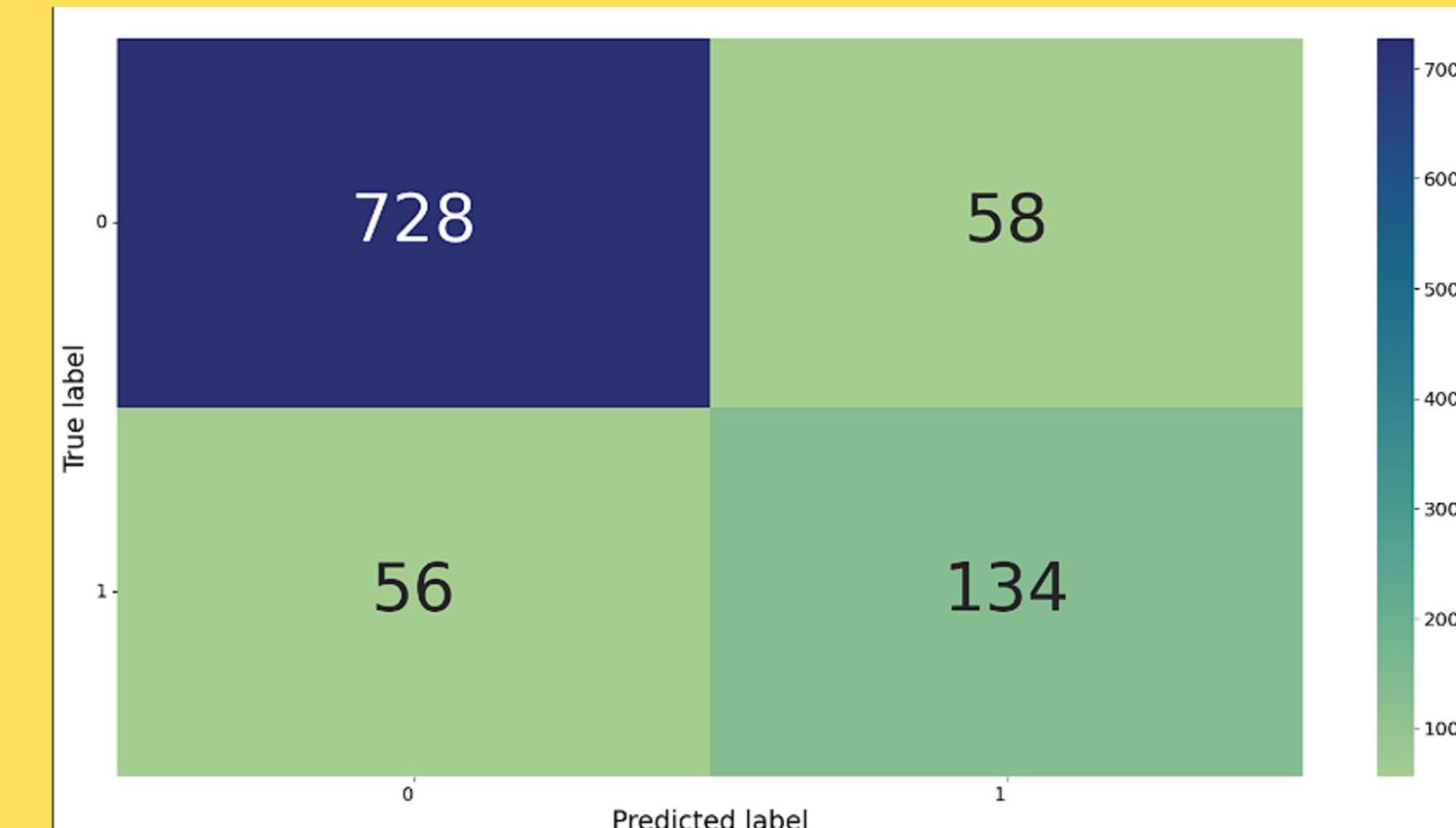
DECISION TREE & HYPERPARAMETER TUNING

Decision Tree Model Before & After
Hyperparameter tuning

Decision Tree Score	Before Tuning	After Tuning
Training Accuracy	1.0	0.93
Testing Accuracy	0.89	0.88
Precision	0.95	0.93
Recall	0.92	0.92
F1-Score	0.94	0.93

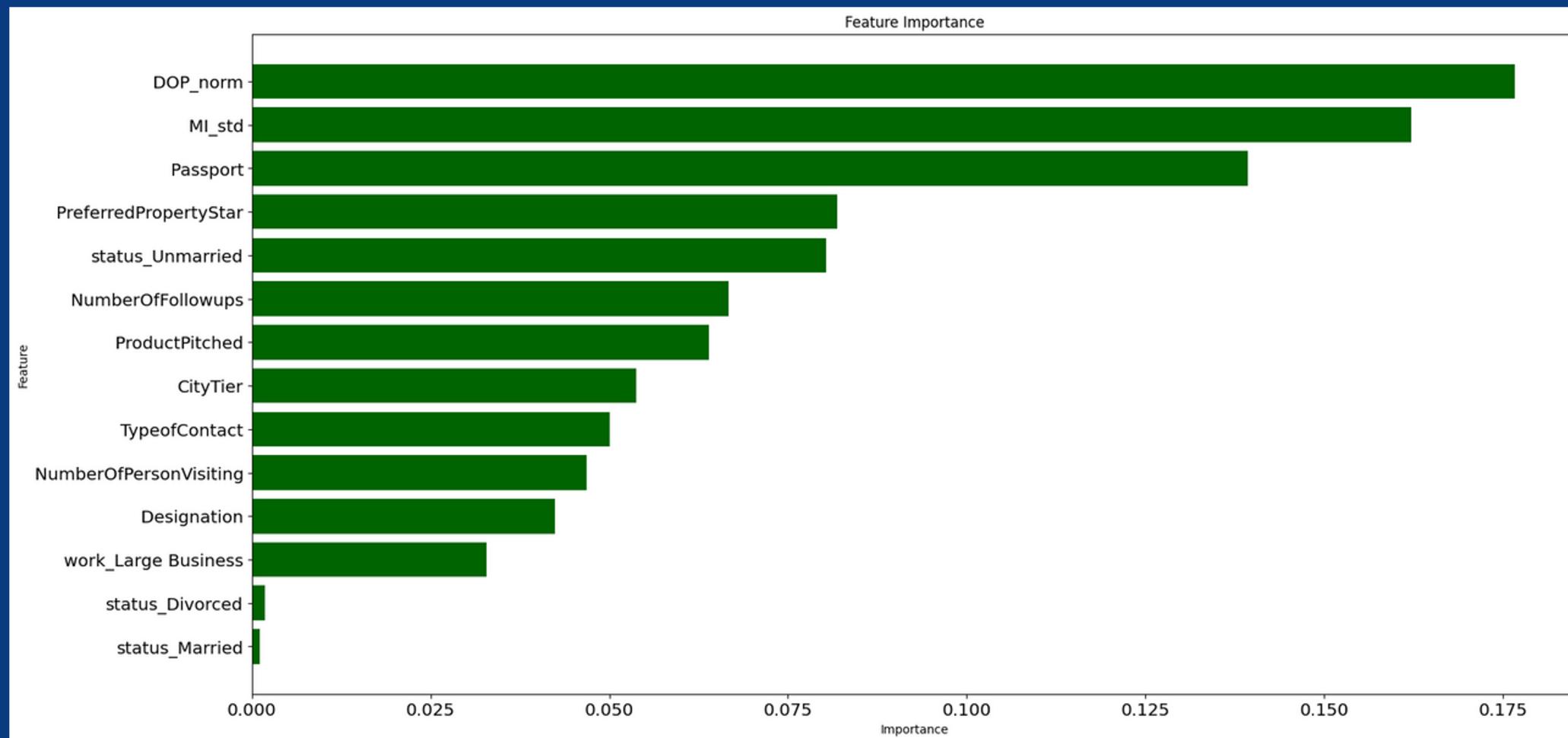
- Hyperparameter tuning yang dilakukan menjadikan model decision tree tidak overfitting

Confusion Matrix



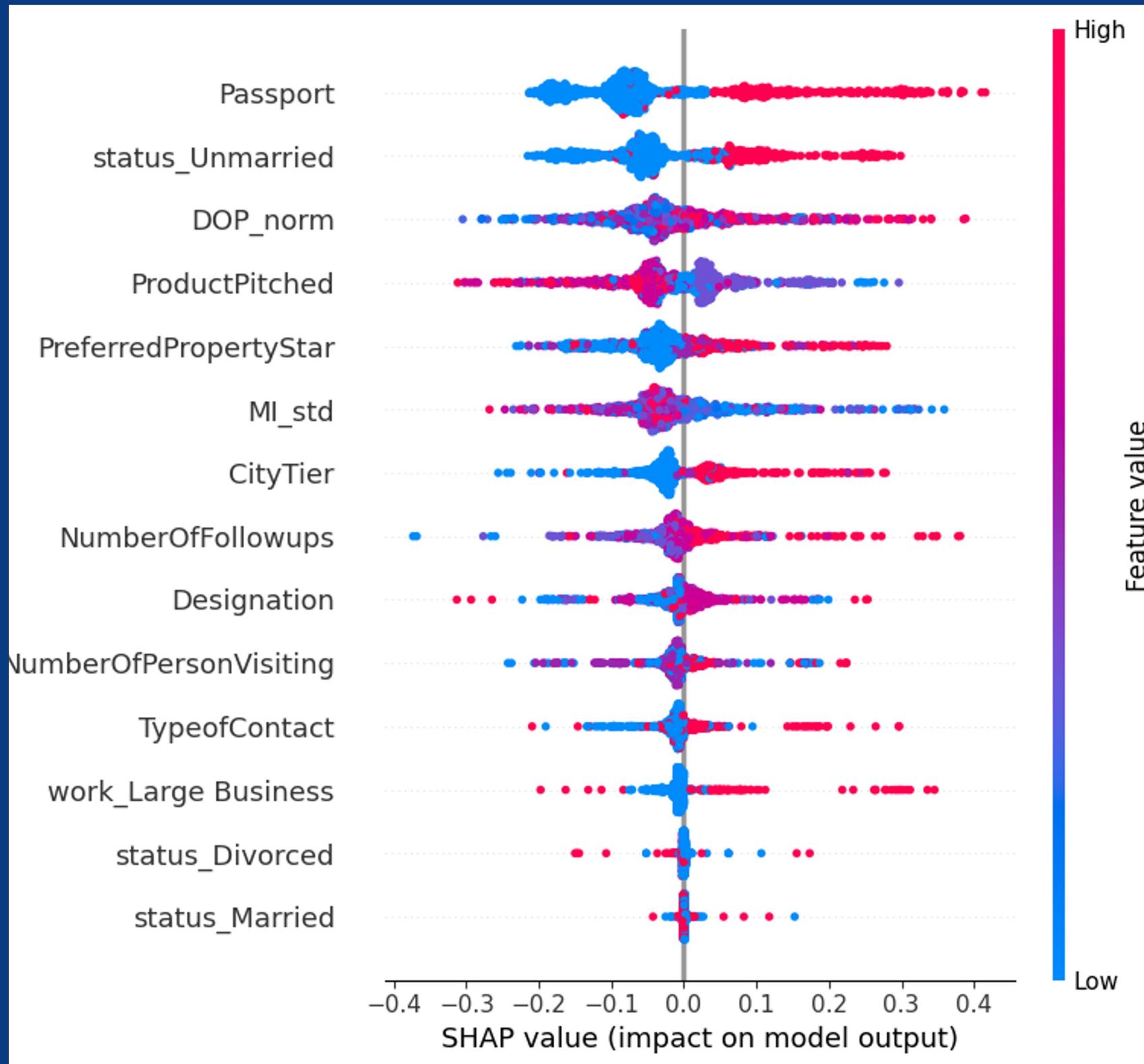
- True Positive: 134
- True Negative: 728
- False Positive: 58
- False Negative: 56

FEATURE IMPORTANCE



Jumlah penghasilan berpengaruh terhadap keputusan customer untuk membeli package.
Calon customer yang sudah memiliki paspor atau yang duration of pitch nya lebih lama
memiliki probabilitasnya yg lebih tinggi untuk membeli package.

SHAP VALUES



Fitur Passport memiliki kontribusi yang tinggi dalam membuat prediksi. Hal ini menunjukkan bahwa status memiliki paspor dapat mempengaruhi hasil dari pitch penjualan. Mungkin ada faktor-faktor terkait perjalanan atau mobilitas yang mempengaruhi minat atau kemampuan seseorang untuk membeli produk.



BUSINESS RECOMMENDATION

Business Recommendation for better Company



TRIPS&TRAVEL.COM



RECOMMEND CUSTOMER TYPE

Berdasarkan Feature Importance

>22 menit Duration Of Pitch

Lower Middle Monthly Income
(\$11.000 - \$20.000)

Punya Passport

Preferred Property 5.0 Star

Belum Menikah

CONVERSION RATE

Before Modelling	After Modelling
Sebelum dilakukan modelling, presentase conversion rate sebesar 19.4%	Sesudah dilakukan modelling dengan menggunakan Decision Tree, presentase model menjadi 69.7%. Dengan perhitungan : $132 / 191 = 0.697$

Perhitungan Conversion Rate menggunakan hasil TP dan FP dari confusion matrix yang digunakan setelah melakukan tuning hyperparameter. Disimpulkan dari hasil conversion rate sebelum dan sesudah terdapat kenaikan sebesar 50.3%.

APPLICATION OF CONVERSION RATE

Asumsi paket akan di tawarkan kepada 1000 customer dengan harga paket yang sama dan busdget marketing yang sama

	Jumlah Customer Convert	Gross Sales
Without Modelling	194 Orang (1000 x 19%)	\$19,400 (190 x \$100*)
With Modelling	697 orang (1000 x 69%)	\$69,700 (690 x \$100)

259% 

Dengan model yang telah dibuat, dapat menaikan Gross Sales sampai \$50,300 atau sebesar 259%. Oleh karena itu model ini dapat direkomendasikan untuk mencari pelanggan baru

*menggunakan paket Basic (\$100)

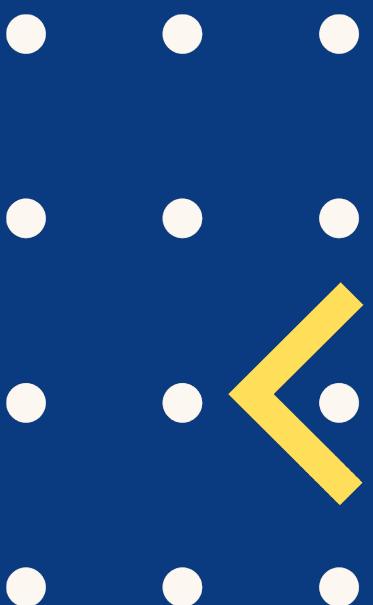
MARKETING COSTS

existing Customer

	Without Modelling	With Modelling
Total Customer yang dihubungi	976 orang	192 orang
Marketing Costs	\$4050,4 (976 x \$4.15*)	\$796.8 (121 x \$4.15)

Dengan model yang telah dibuat, dapat mengurangi biaya marketing sebesar **\$3,253 .6** atau sebesar **80.33%**. Oleh karena itu model ini dapat direkomendasikan untuk melakukan pitching produk terhadap customer yang sudah ada.

THANK YOU!



DISCUSSION SESSION

Let's Discuss It Together!

A graphic element consisting of three light-colored wooden blocks arranged horizontally. The first block contains the letter 'Q', the second block contains the symbol '&', and the third block contains the letter 'A'. The blocks are set against a dark green circular background, which is itself centered within a larger circular frame composed of two concentric arcs in yellow and light blue.

TRIPS&TRAVEL.COM