



ANALYTIC ADVENTURER

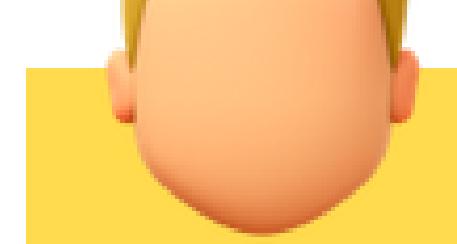
FINAL PROJECT

Holiday Package Prediction
Rakamin Academy

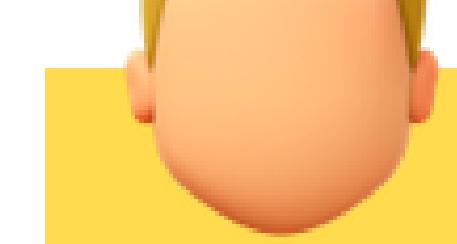
Our Team



**MUHAMAD
RAJWA
ATHORIQ**



**REYNALDI
SILALALAHI**



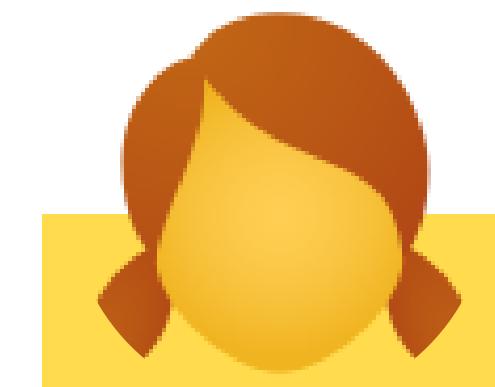
**DANU SATRIA
WIRATAMA**



**NOFRI
RHAMAWATI**



**WAHYU ISMI
AZIZAH**



**WIKUDHARA
SHOFRINA
WIJAYA**



Who We are?

Sebuah team Data Scientist yang bertanggung jawab untuk menganalisis data pelanggan dan mengembangkan model prediktif yang akurat dengan menentukan pelanggan yang paling potensial membeli paket Wellness Tourism.

Table of Content



BUSINESS UNDERSTANDING



EDA & INSIGHT



DATA PRE-PROCESSING



MODELING



BUSINESS RECOMENDATION



BUSINESS UNDERSTANDING



Background

"Trips & Travel.Com" berencana memperkenalkan paket baru bernama Wellness Tourism Package untuk memperluas basis pelanggan. Saat ini, perusahaan menawarkan lima paket, dan memiliki 18% pelanggan yang membeli paket tersebut tahun lalu. Tingginya biaya pemasaran disebabkan oleh pendekatan acak dalam menghubungi pelanggan. Untuk peluncuran paket baru, perusahaan ingin memanfaatkan data pelanggan yang ada dan potensial agar pengeluaran pemasaran lebih efisien dan tepat sasaran



Problem

Meskipun Trips & Travel.Com memiliki conversion rate yang tinggi, yaitu mencapai 18% (di atas rata-rata industri travel yang hanya 11,9%), tetapi masalah biaya pemasaran masih menjadi kendala. Hal ini disebabkan oleh pemasaran yang dilakukan secara acak tanpa mempertimbangkan potensi pelanggan yang sudah ada.

Goal

Mengefisiensikan cost marketing yang akan dikeluarkan oleh tim marketing perusahaan Trips & Travel.Com

Objective

- **Efisiensi Anggaran Pemasaran**

Memangkas pengeluaran biaya marketing sebesar 60% dengan cara menawarkan package kepada customer yg memiliki potensi pembelian yang tinggi

- **Meningkatkan Conversion Rate**

Meningkatkan tingkat konversi pembelian melalui pitching lebih dari 18%.

Business Metrics

MARKETING EFFICEINCY RATIO

Pendapatan kotor yang diterima oleh perusahaan dibagi dengan biaya marketing yang dikeluarkan

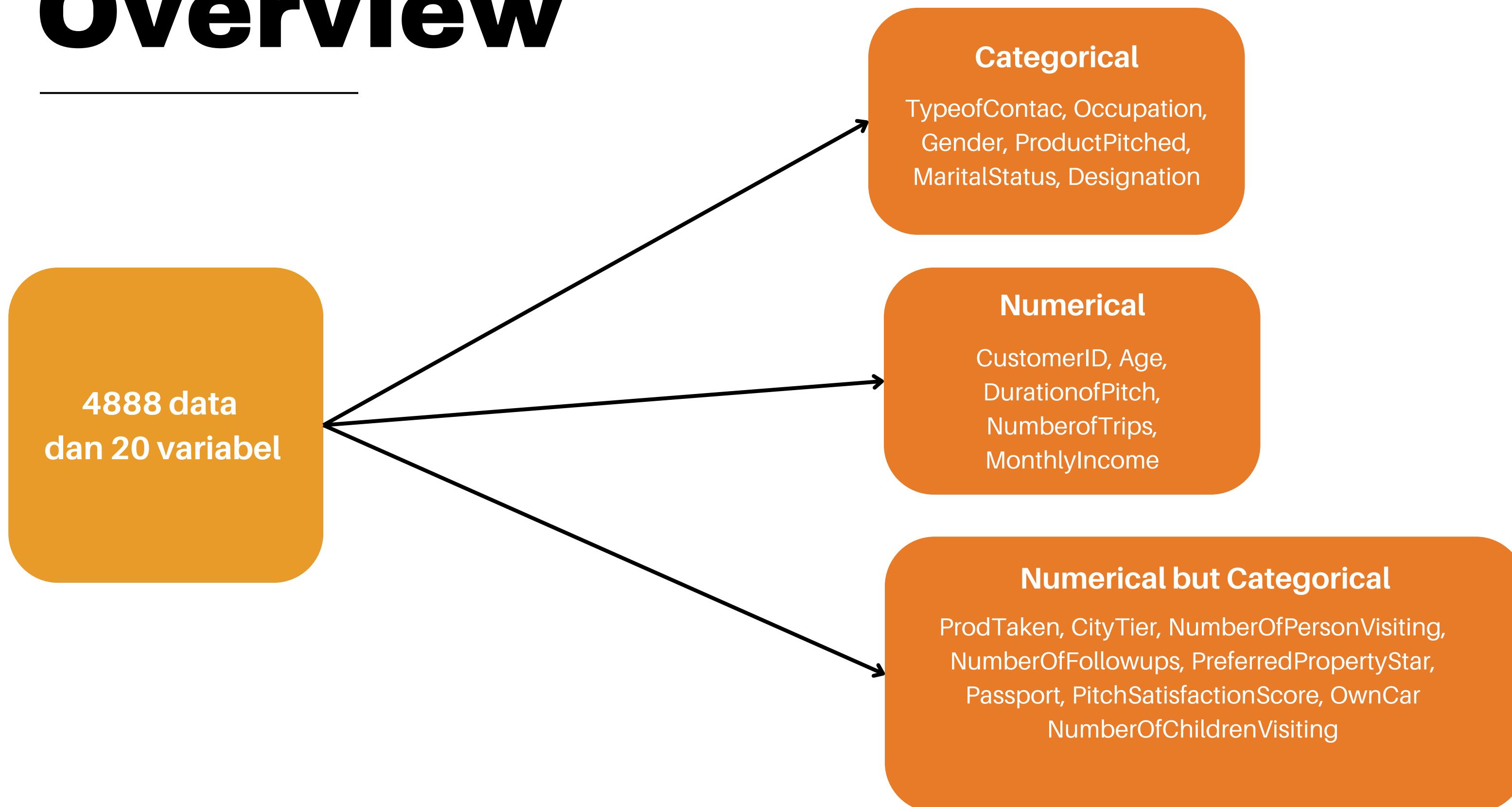
CONVERSION RATE

Mengukur tingkat konversi customer yang menerima tawaran dari pihak perusahaan

ANALYTIC ADVENTURER

EDAG INSIGHT

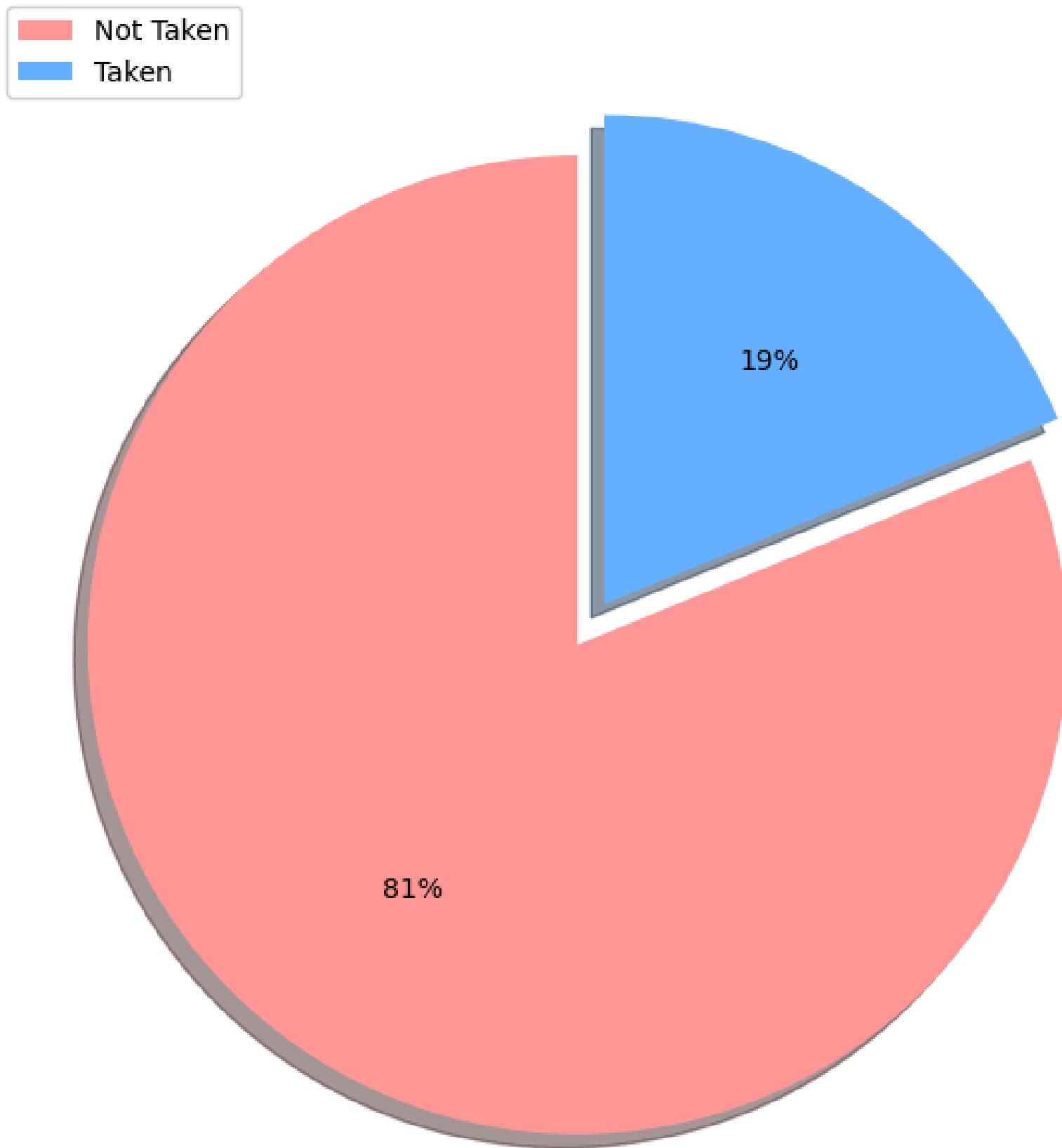
Dataset Overview



Checking Missing Value

Variabel	Missing Value
Age	226
MonthlyIncome	233
NumberOfFollowups	45
DurationOfPitch	251
TypeofContact	25
PreferredPropertyStar	26

Komposisi ProdTaken



Univariate Analysis

Data status:
Imbalance

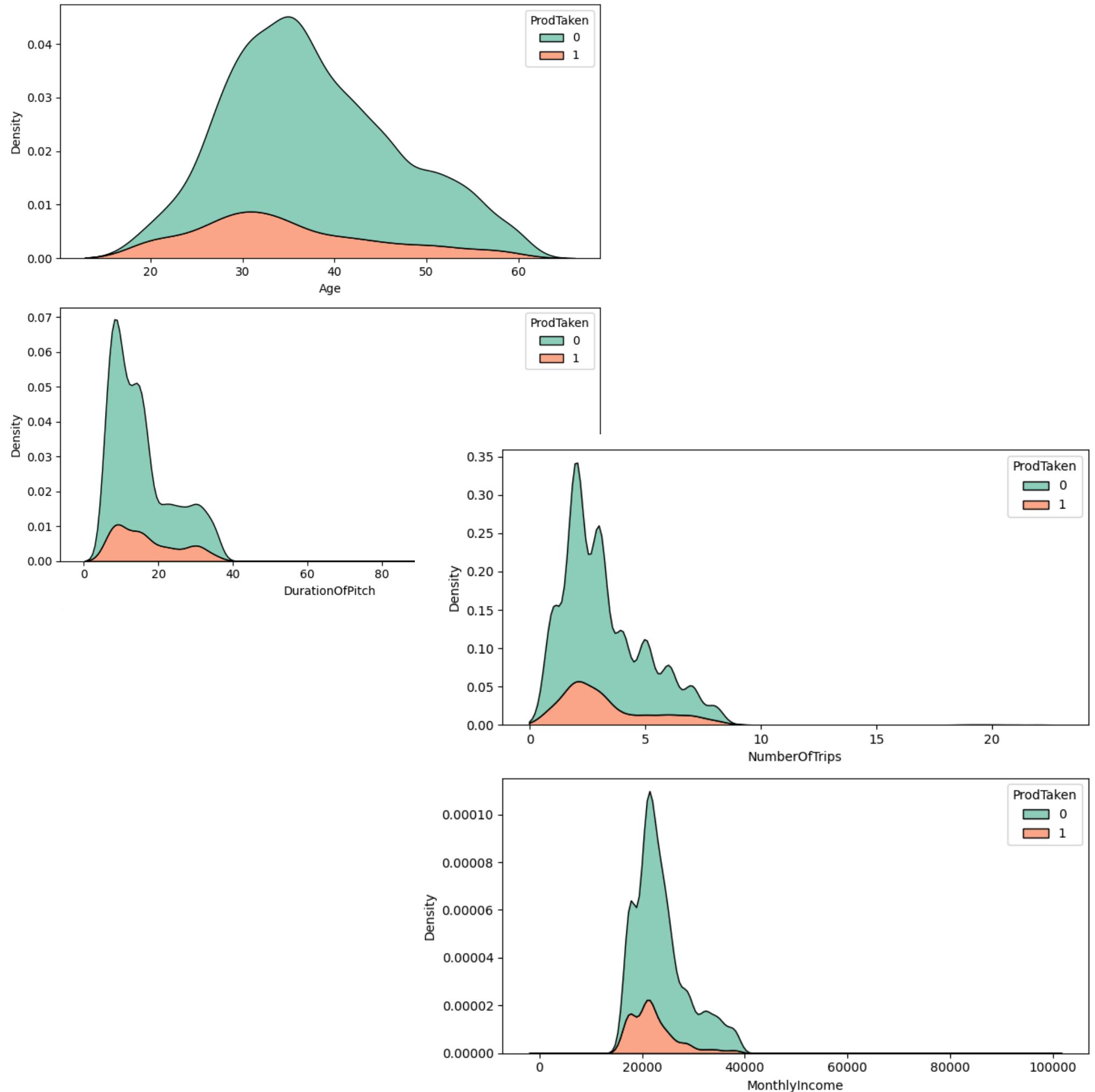
Bivariate Analysis

Numerical Columns

Faktor dominan dalam pembelian paket

- Umur
- DurationOfPitch
- MonthlyIncome
- NumberOfTrips

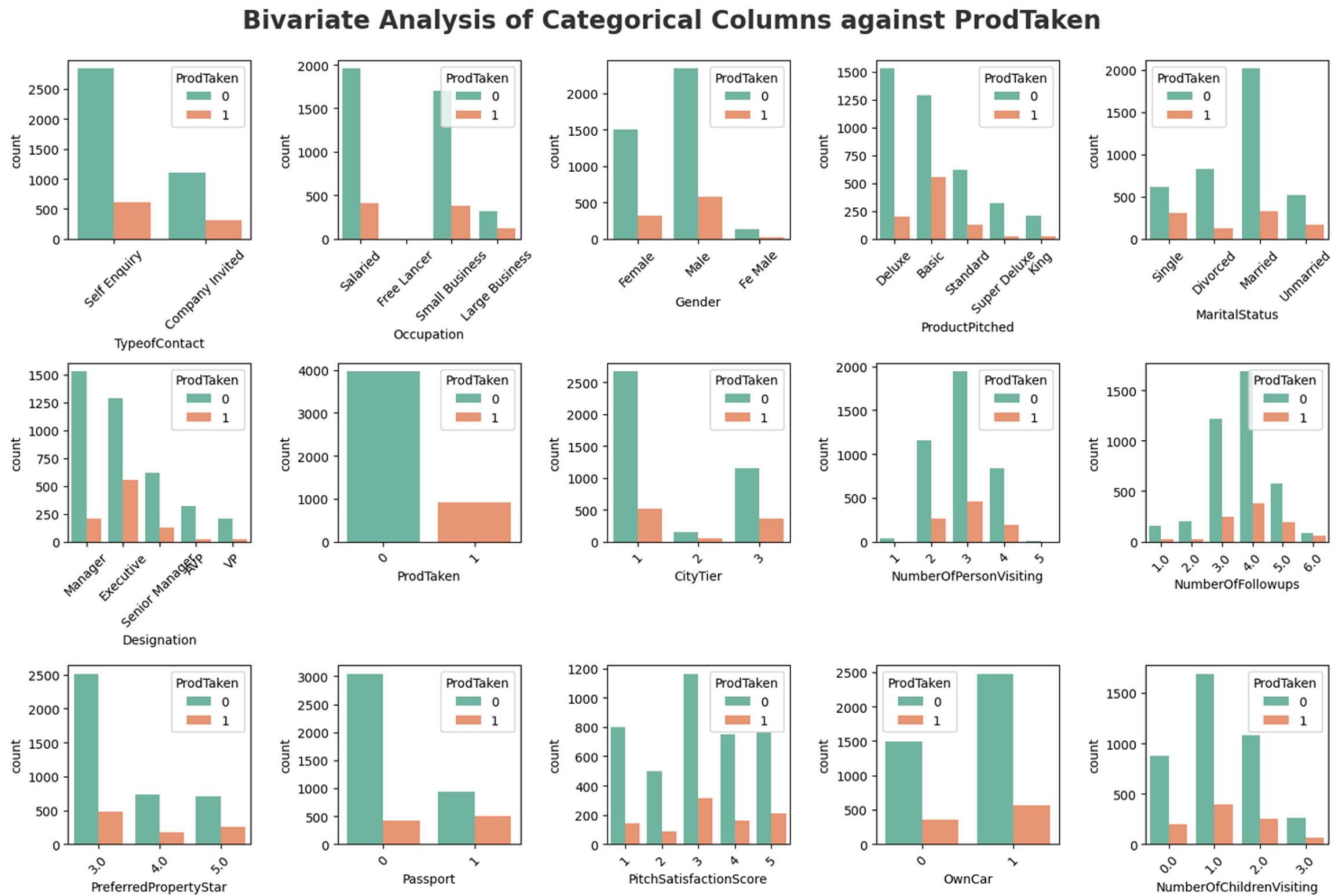
29 s.d. 40 tahun
20 menit
15000 s.d. 25000
 ≤ 5 trips



Bivariate Analysis

Categorical Columns

Faktor dominan dalam pembelian paket



Self Enquiry

salaried

Male

Deluxe

married

ManagerSenior

citytier3

numberofperson3

numberoffollowups4

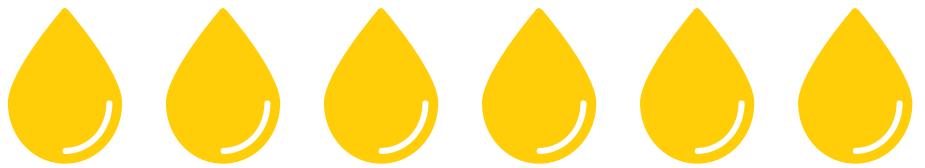
preferedpropertystar3

yang memiliki pasport

dengan memiliki owncar

serta yang membawa anak 1

Bivariate Analysis



uji chi-square

kolom Gender, OwnCar, dan
NumberOfChildrenVisiting tidak akan
dimasukkan kedalam feature karena tidak
berkorelasi dengan target (ProdTaken)

uji t-test

Semua fitur memiliki perbedaan yang signifikan
terhadap ProdTaken dibandingkan dengan fitur
NumberOfTrips. perbedaan tersebut
mempengaruhi customer dalam pembelian

Bivariate Analysis

	Column	Hypothesis Result
0	TypeofContact	H1 diterima: Berkorelasi dengan target
1	Occupation	H1 diterima: Berkorelasi dengan target
2	Gender	H0 diterima: Tidak berkorelasi dengan target
3	ProductPitched	H1 diterima: Berkorelasi dengan target
4	MaritalStatus	H1 diterima: Berkorelasi dengan target
5	Designation	H1 diterima: Berkorelasi dengan target
6	CityTier	H1 diterima: Berkorelasi dengan target
7	NumberOfPersonVisiting	H1 diterima: Berkorelasi dengan target
8	NumberOfFollowups	H1 diterima: Berkorelasi dengan target
9	PreferredPropertyStar	H1 diterima: Berkorelasi dengan target
10	Passport	H1 diterima: Berkorelasi dengan target
11	PitchSatisfactionScore	H1 diterima: Berkorelasi dengan target
12	OwnCar	H0 diterima: Tidak berkorelasi dengan target
13	NumberOfChildrenVisiting	H0 diterima: Tidak berkorelasi dengan target

	Feature	t-statistic	p-value	Hypothesis Result
0	Age	10.162969	5.155096e-24	H1 diterima: Ada perbedaan signifikan
1	DurationOfPitch	-5.344214	9.516584e-08	H1 diterima: Ada perbedaan signifikan
2	NumberOfTrips	-1.302106	1.929435e-01	H0 diterima: Tidak cukup bukti untuk menolak h...
3	MonthlyIncome	8.984508	3.705260e-19	H1 diterima: Ada perbedaan signifikan

Multivariate Analysis

Lemah - Negatif

Age

MonthlyIncome

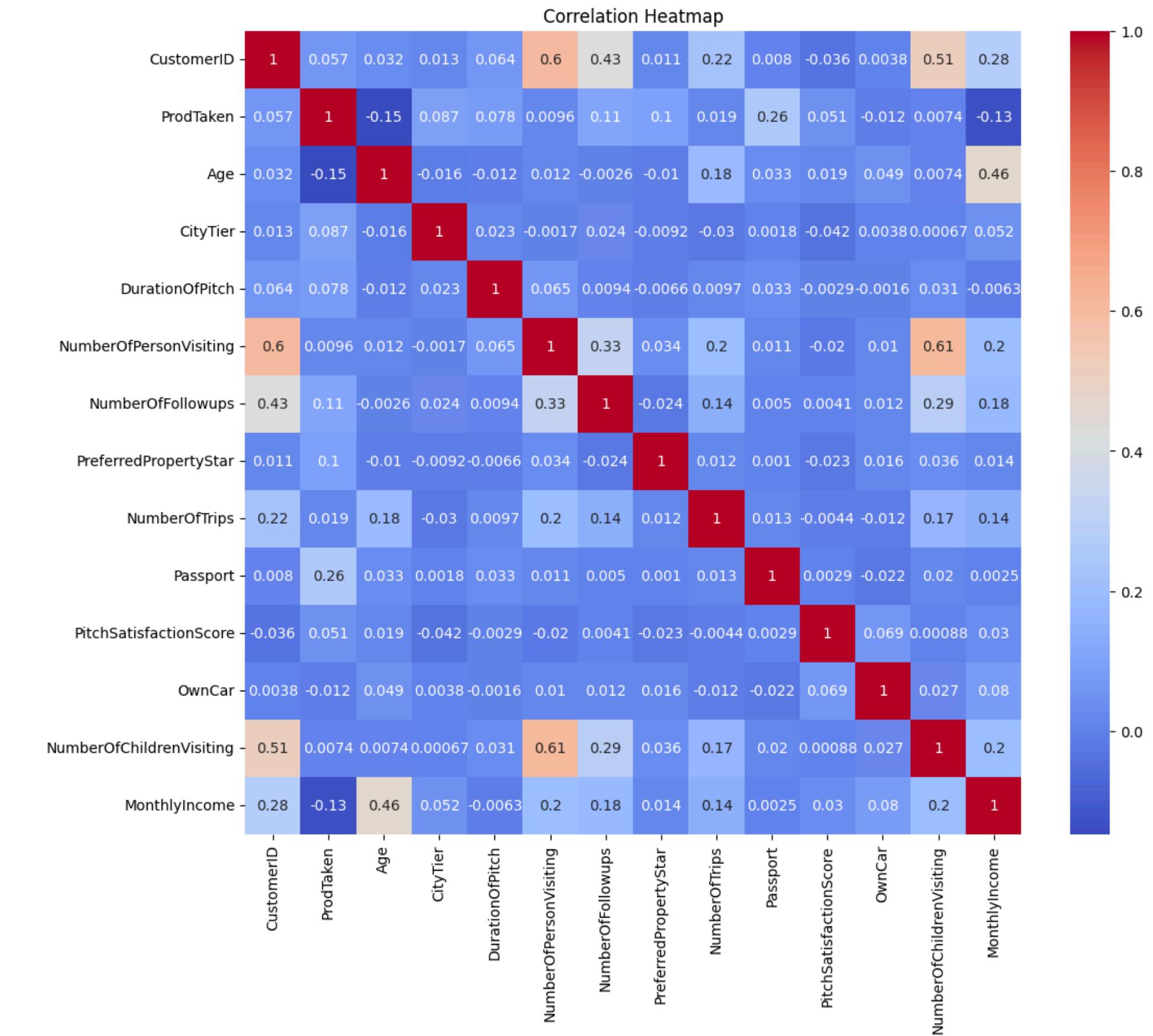
Kuat - Positif

Passport

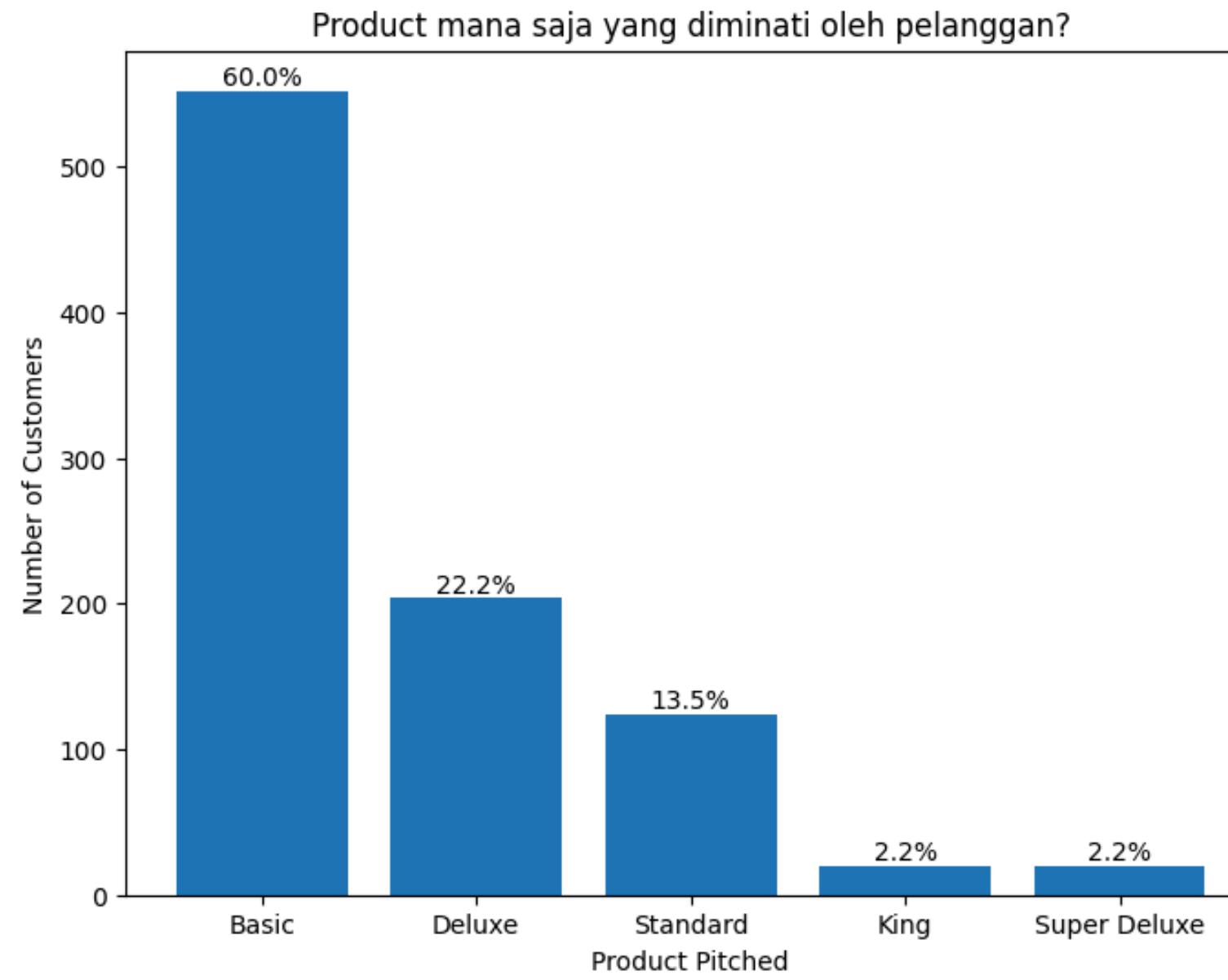
NumberOfTrips

NumberOfFollowups

PreferredPropertyStar

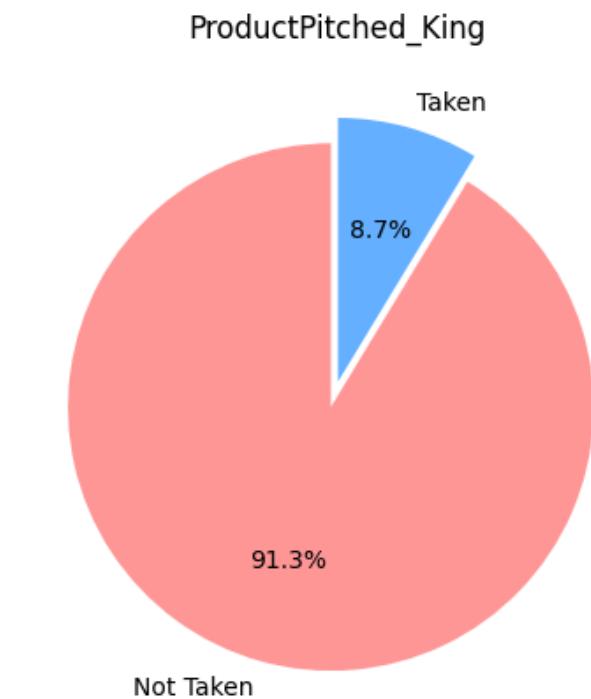
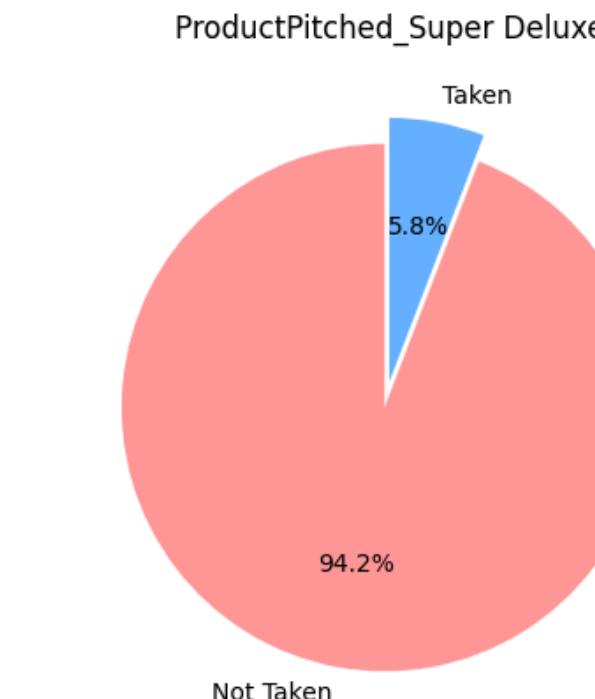
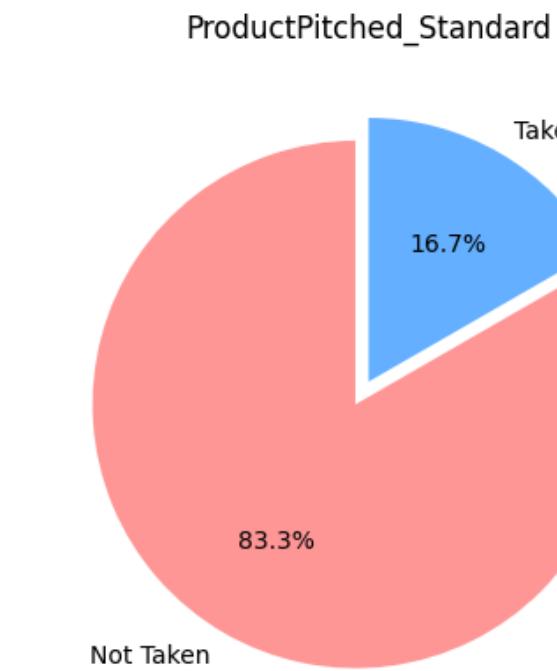
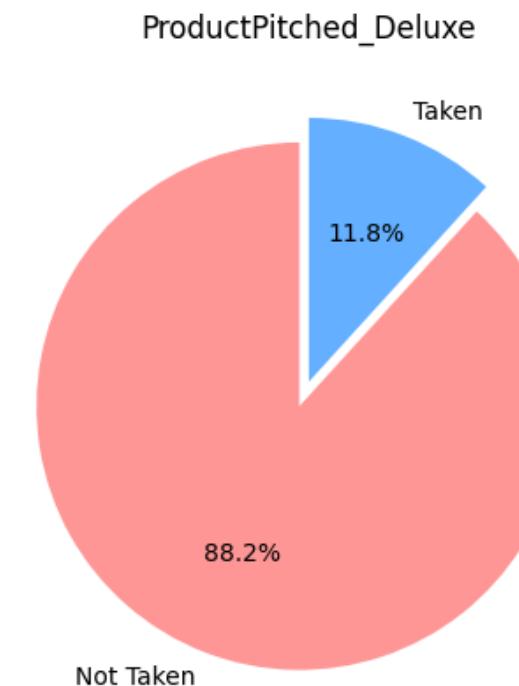
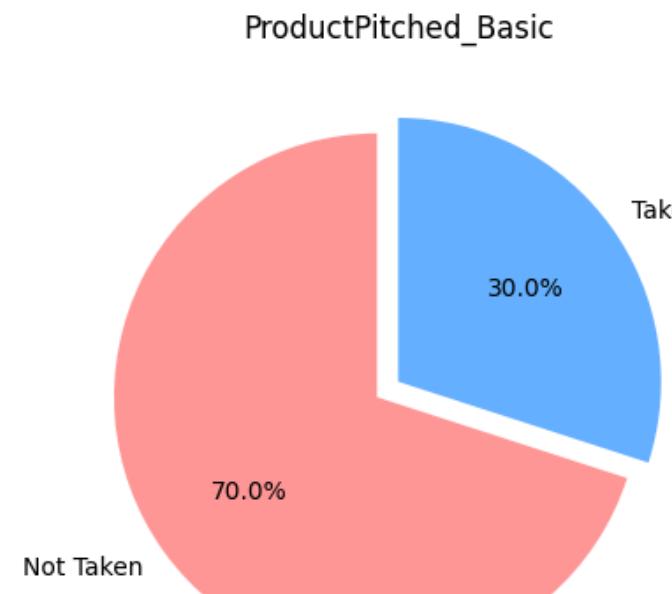


Insight



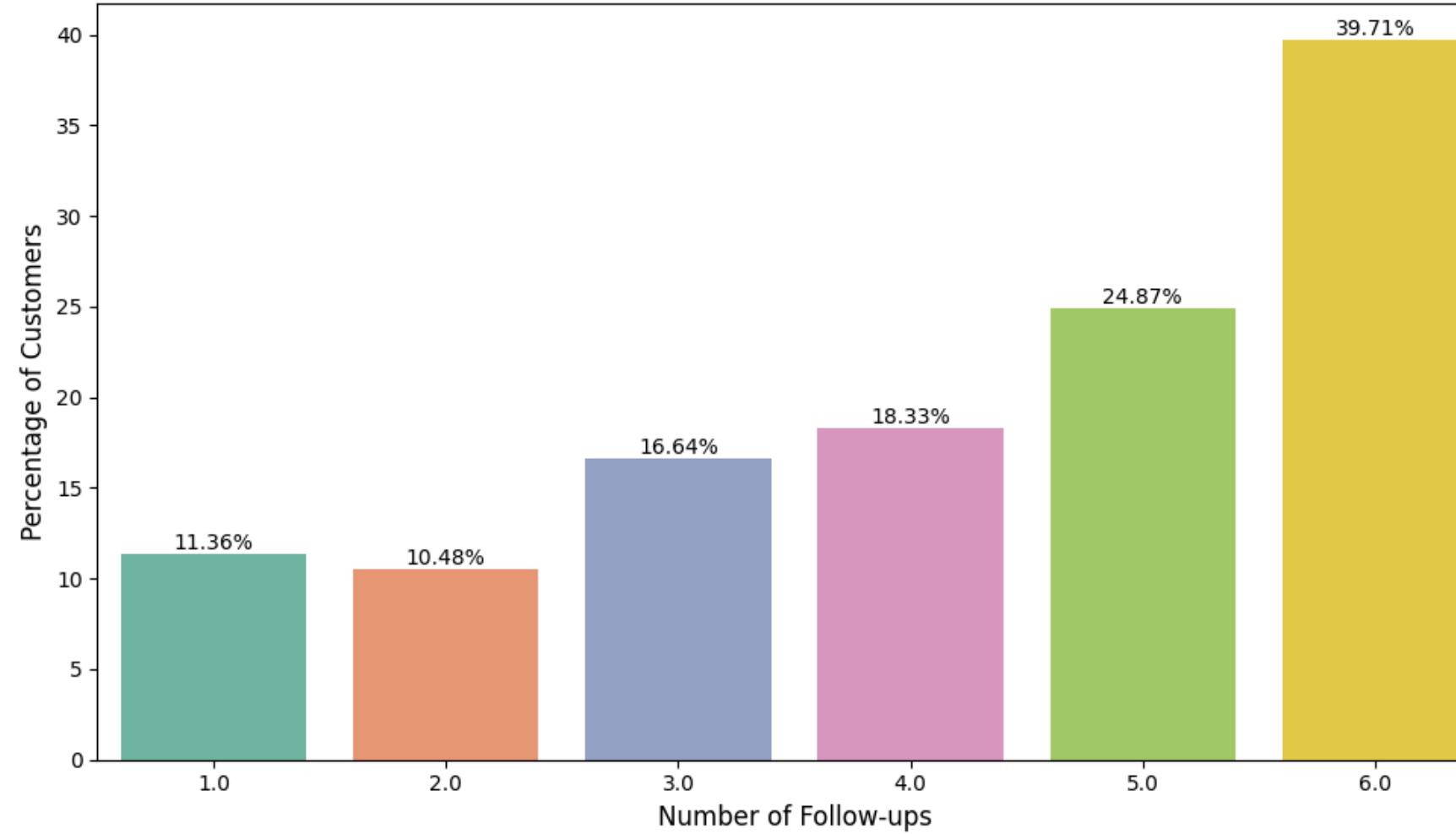
Product mana saja yang diminati oleh pelanggan?

1. Basic
2. Standard
3. Deluxe
4. King
5. Super Deluxe



Insight

Apakah jumlah followups dari sales mempengaruhi Customers untuk membeli paket perjalanan?



**Apakah jumlah follow-ups
dari sales mempengaruhi
customer untuk membeli
paket perjalanan?**

1 Follow up

6 Follow up

11.36% pembelian

39.71% pembelian

Intensif Follow up

Peningkatan konversi
pelanggan

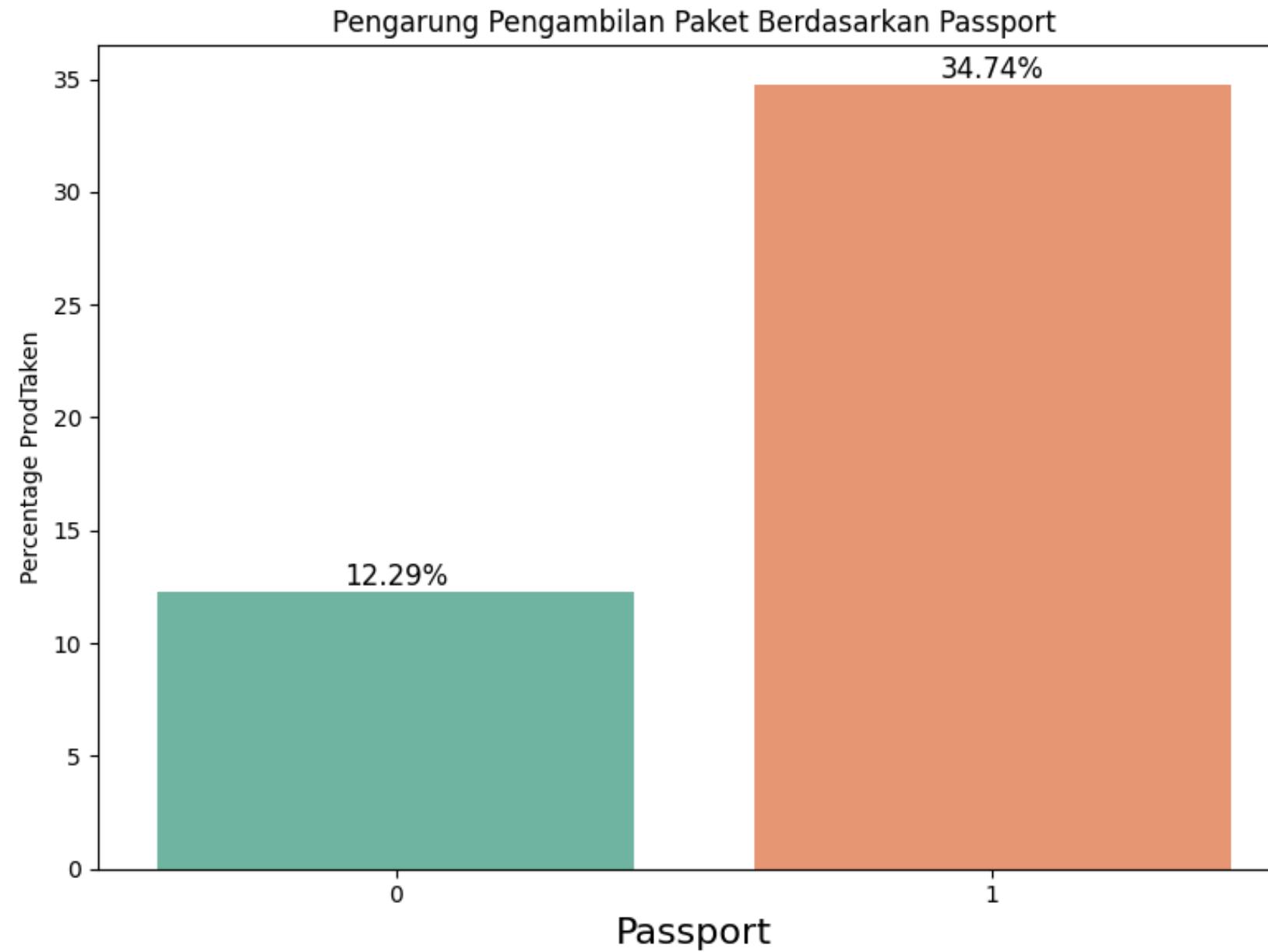
Insight

Jenis kota apa yang memiliki penjualan paket perjalanan terbanyak?

1. City Tier 3
2. City Tier 2
3. City Tier 1



Insight

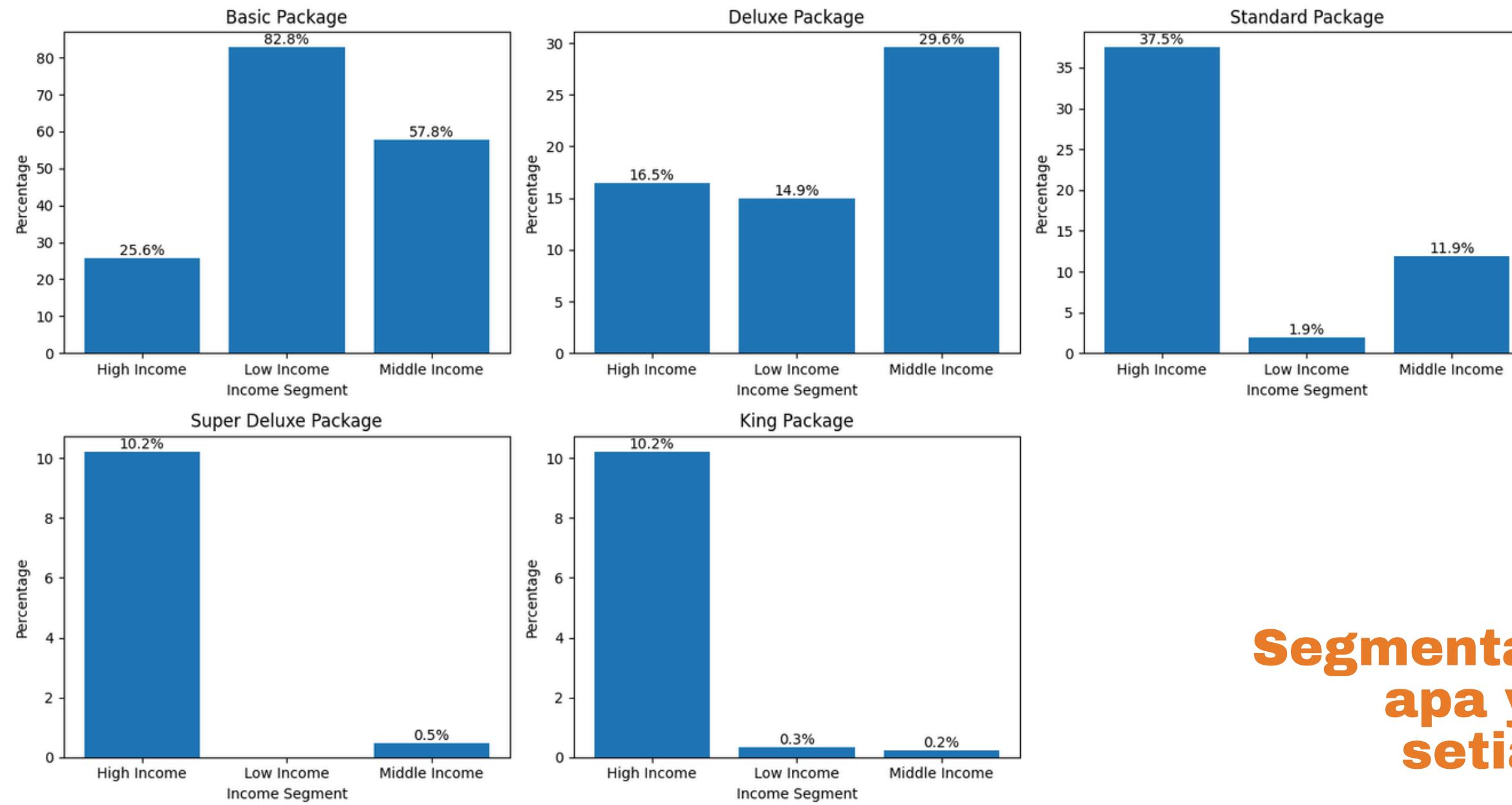


Apakah mempunyai passport berpengaruh pada pengambilan paket?

Memiliki passport



Kecenderungan untuk
membeli paket perjalanan



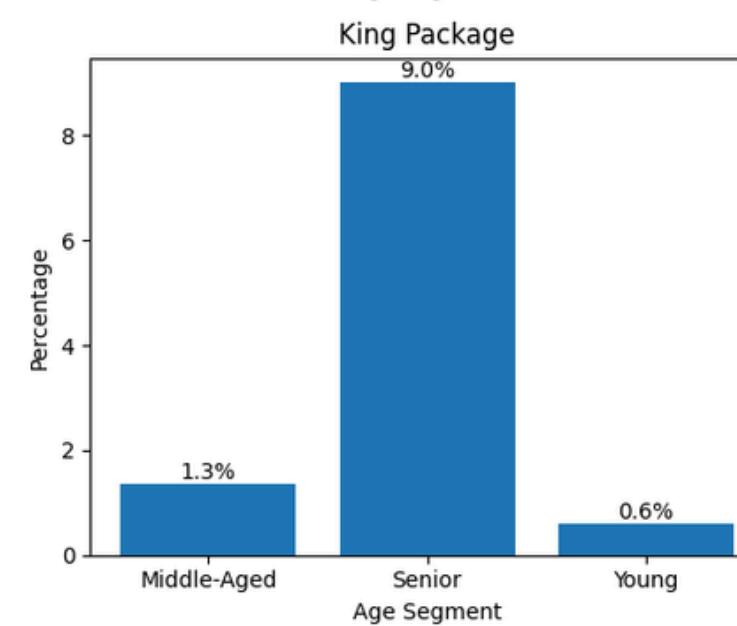
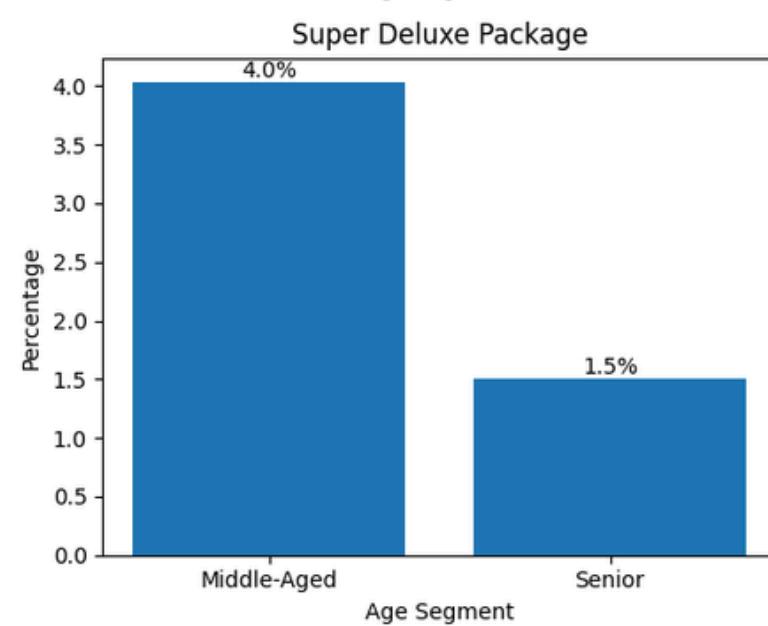
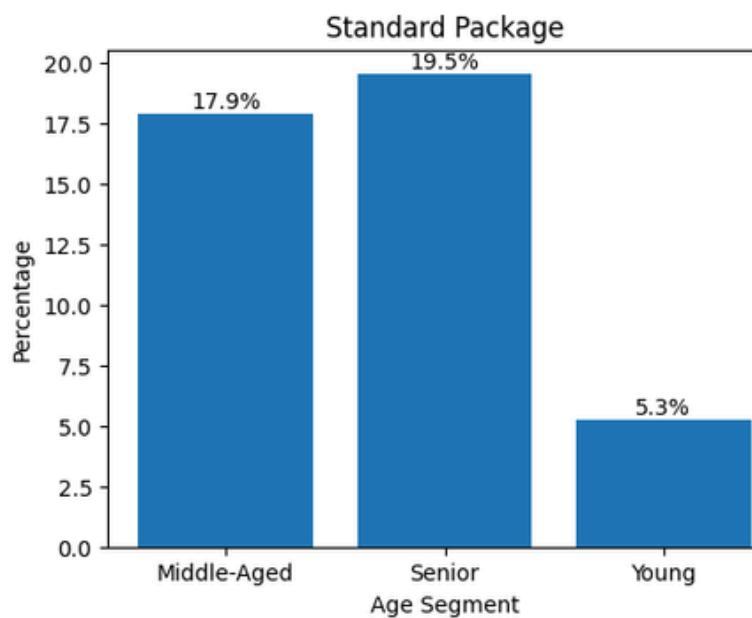
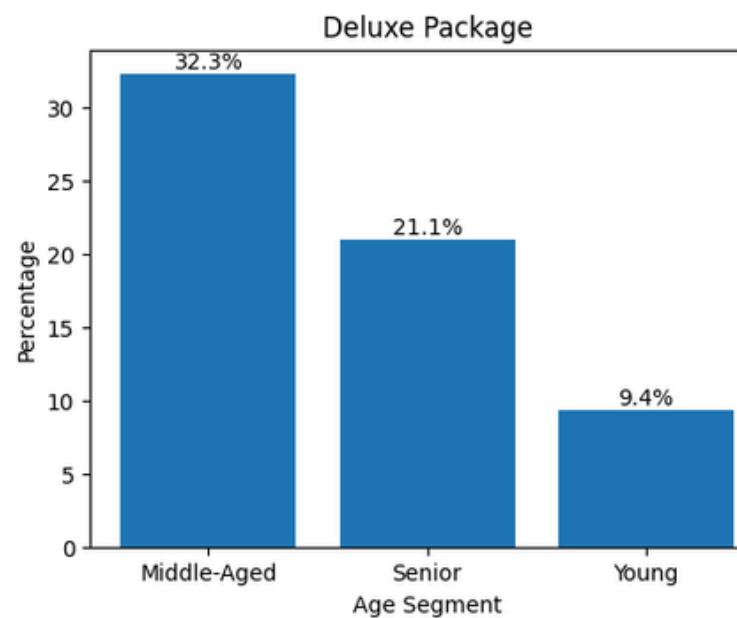
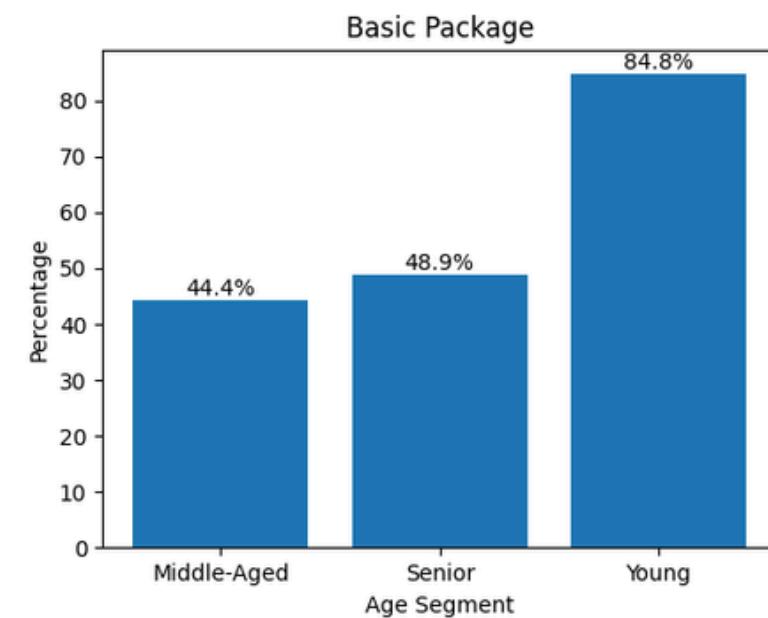
Low Income
Middle Income
High Income



Basic
Deluxe
Standard, Super Deluxe, King

Insight

Segmentasi Umur dan Income apa yang mendominasi di setiap paket perjalanan?



Young Age, Senior Age
Middle Age
Senior Age



Insight

**Segmentasi Umur dan Income
apa yang mendominasi di
setiap paket perjalanan?**

Deluxe
Super Deluxe
Standard, King

DATA PRE-PROCESSING

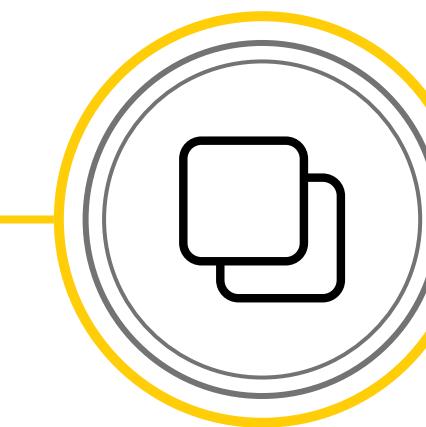
Data Pre-processing



Handling Missing Value

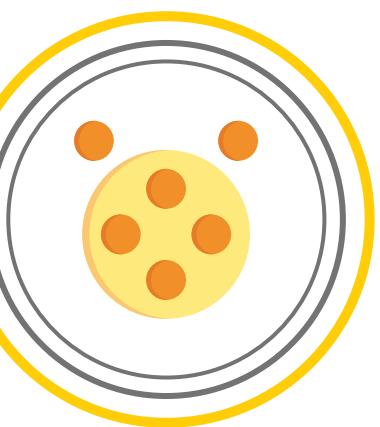
Mengisi missing values untuk kolom **numerik** dengan nilai **median**

Mengisi missing values untuk kolom **kategorikal** dengan nilai **mode**



Handling Duplicate

Terdapat **596 data duplikat**. Kemudian data tersebut dihapus dan menyisakan satu data.

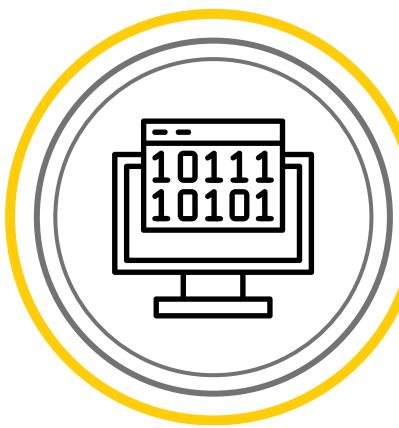


Handling Outliers

Terdapat outliers pada kolom 'Age', 'MonthlyIncome', 'NumberOfFollowups', 'DurationOfPitch'

Didapatkan total **687 data outliers** dan dilakukan penghapusan.

Data Pre-processing



Encoding Value



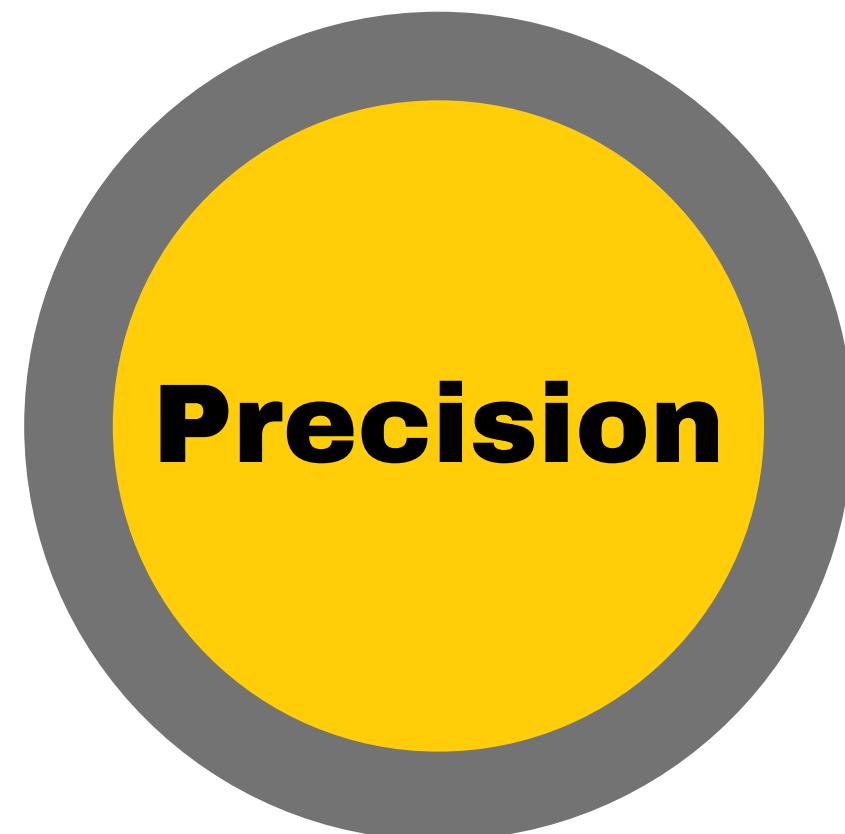
Scaling

Melakukan **label encoding** pada kolom 'Designation'. Dengan nilai Executive: 1, Manager: 2, Senior Manager: 3, AVP: 4, VP: 5

Melakukan **one hot encoding** pada kolom 'MaritalStatus','Occupation', 'ProductPitched', 'TypeofContact'

Digunakan untuk menyesuaikan rentang atau distribusi fitur dalam dataset agar memiliki **skala yang seragam**

MODEL EVALUATION



	Predict Yes	Predict No
Actual No	False Positif	True Negatif
Actual Yes	True Positif	False Negatif



DEFINISI

Precision menunjukkan seberapa banyak dari semua prediksi positif yang benar-benar merupakan positif aktual.



PERHITUNGAN

$\frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Positives (FP)}}$



TUJUAN

Metrik yang digunakan untuk mengukur akurasi model klasifikasi dalam hal positif yang diprediksi

Result Best Model

	Score Precision Training	Score Precision Testing
AdaBoost	0.65	0.65
Logistic Regression	0.68	0.67
Decision Tree	1.0	0.71
Random Forest	1.0	0.86
Gradient Boosting	0.85	0.76
XGBoost	1.0	0.87

HYPERPARAMETER TUNING

	Score Precision Training	Score Precision Testing
AdaBoost	0.79	0.83

Feature Importance



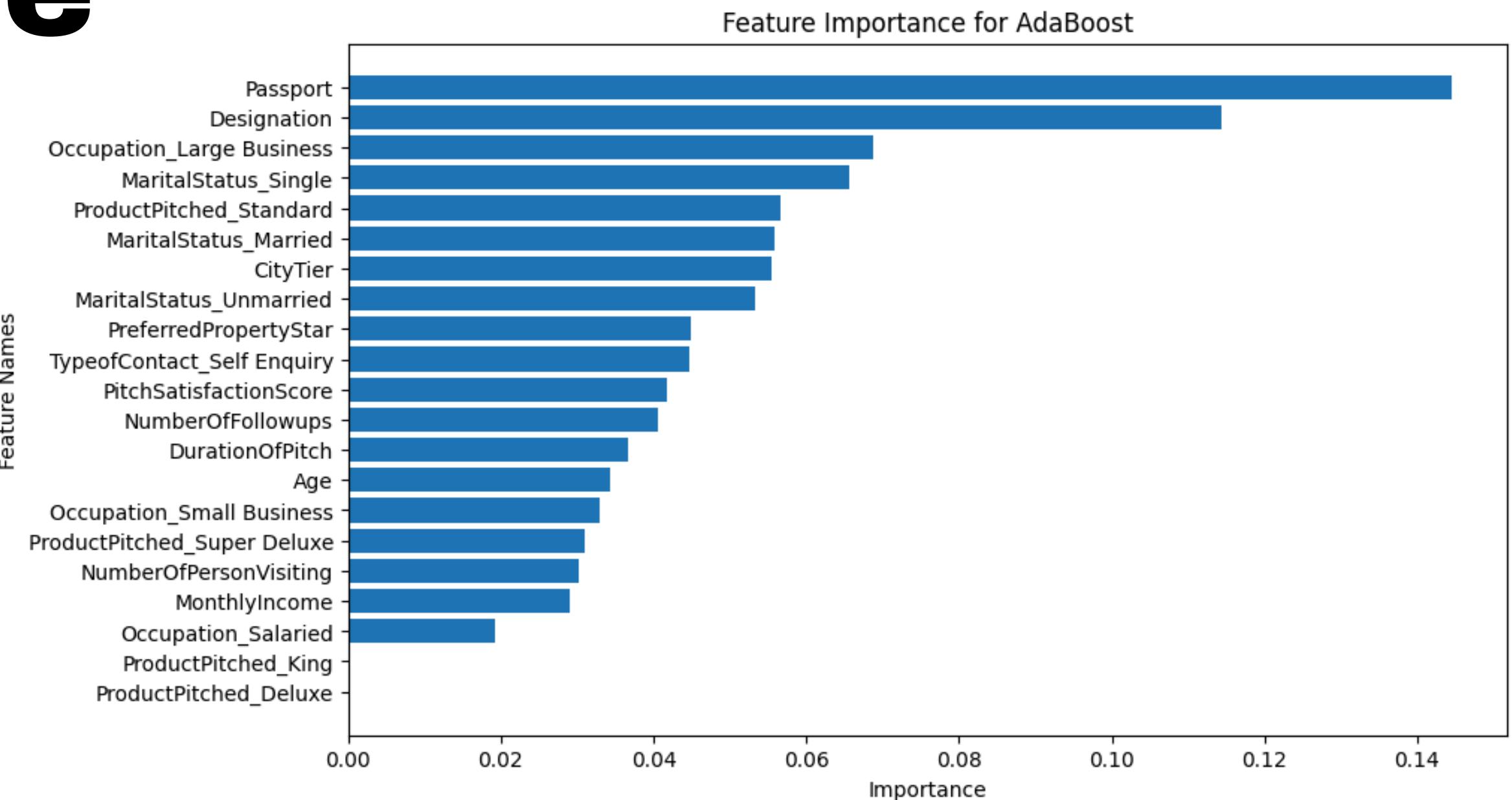
Pelanggan yang memiliki passport memiliki potensial untuk membeli package yang ditawarkan



Custmer yang memiliki bisnis yang terbilang besar memiliki potensial untuk membeli package yang ditawarkan



Customer dengan jabatan Executive memiliki potensi yang besar pula untuk melakukan pembelian terhadap Wellness Tourism Package saat ditawarkan



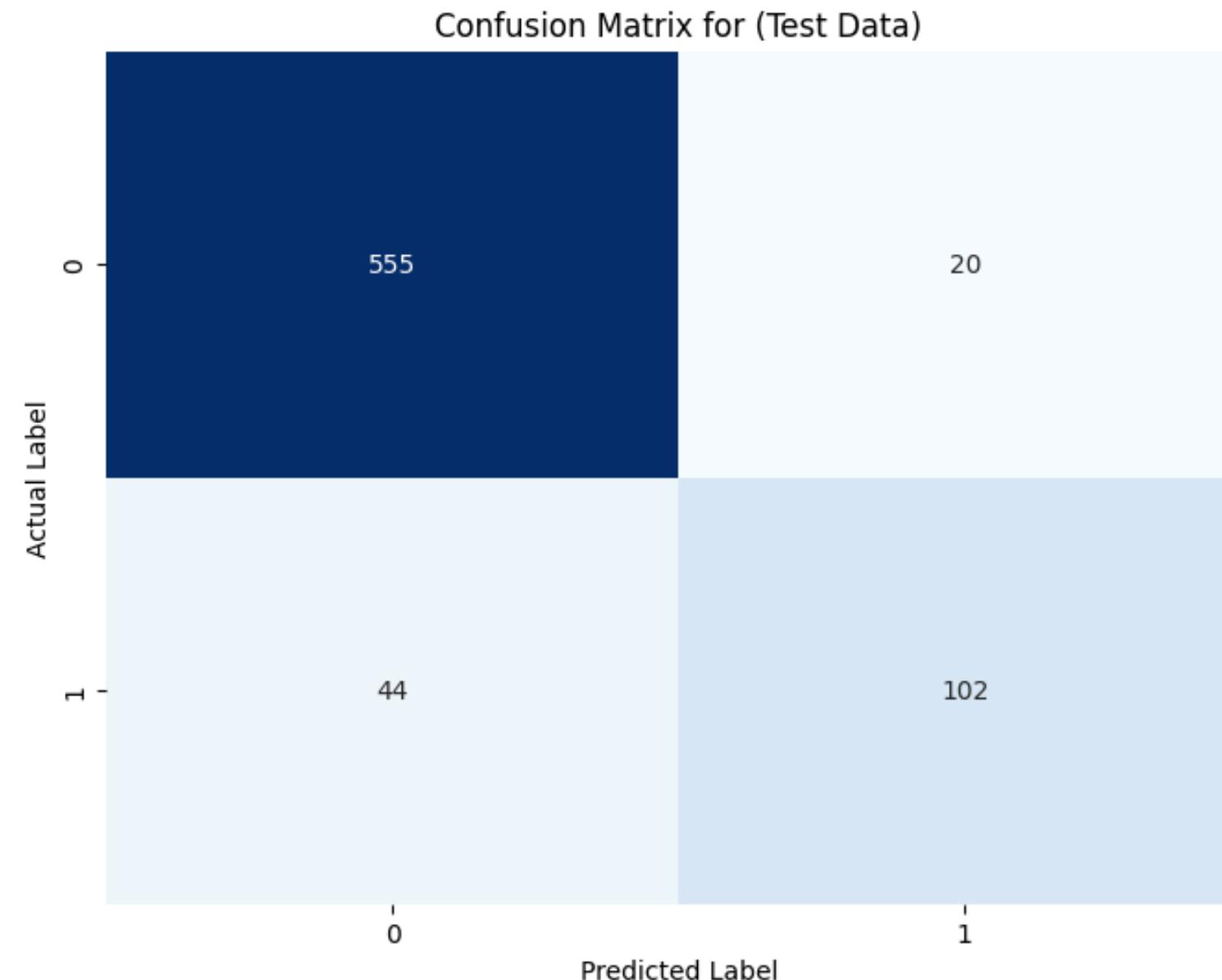
Business Simulation



- Meningkatnya probabilitas customer melakukan pembelian
- Mengefisiensi Cost pada sisi marketing

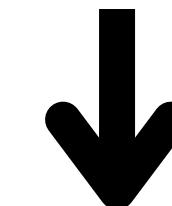
Example Case

ANALYTIC ADVENTURER



Tanpa Model

- Conversion Rate : 18%
- Target : 721 Customer (Seluruh Customer)
- Cost Marketing per Customer : Rp 300.000
- Total Cost : $721 \times 300.000 = \text{Rp } 216.300.000$



Dengan Model

- Conversion Rate : 20%
- Target : 146 Customer
- Cost Marketing per Customer : Rp 300.000
- Total Cost : $146 \times 300.000 = \text{Rp } 43.800.000$

COST EFFECTIVENESS : $216.300.000 - 43.800.000 = 178.500.000$

Note*

- berdasarkan research cost marketing = 5% pendapatan
- berdasarkan artikel moneygeek rata rata penjualan paket penerbangan di US adalah 6jt



81%

BUSINESS INSIGHT & RECOMMENDATION

Insight

Pelanggan yang memiliki **passport**, menjabat sebagai **executive** atau **pengusaha skala besar**, serta berstatus **single/unmarried** memiliki potensi tinggi untuk melakukan pembelian terhadap paket yang ditawarkan, yaitu Wellness Tourism Package.

Pelanggan yang **tinggal di perkotaan besar** memiliki kemungkinan lebih besar untuk melakukan pembelian dibandingkan dengan pelanggan dari kota-kota lainnya. Faktor-faktor ini menekankan pentingnya memahami demografi dan karakteristik pelanggan untuk merancang strategi pemasaran yang lebih efektif dan menargetkan segmen pasar yang tepat.

Recommendation

Fokus Pemasaran

Fokuskan penawaran terhadap customer yang memiliki passport

Mengadakan Payday Sale

diharapkan dapat menarik perhatian customer dengan posisi executive dan meningkatkan penjualan secara signifikan pada periode gaji

Special Offer Price

Memberikan harga khusus untuk pembelian secara massive (untuk menarik customer yang memiliki bisnis berskala besar, sehingga dapat mengajak keluarga ataupun para staff-nya)

Memberikan Perlakuan Khusus

Memberikan perlakuan khusus kepada pelanggan **single/unmarried** dapat menjadi strategi efektif untuk menarik minat mereka, karena mereka cenderung lebih responsif terhadap penawaran. Misalnya, memberikan **merchandise** untuk pelanggan **single/unmarried** yang membeli paket pada hari **Valentine**.

Perbanyak Promosi

Perbanyak promosi di daerah dengan tingkat perkembangan yang lebih tinggi

Analytic  Adventurers