

# TRAVEL INSURANCE

## PREDICTION



**BYTE BLAZERS**

Source data : [dataset](#)

# Byte Blazers

Sebagai tim data scientist perusahaan PT Byte Blazers, kami diminta untuk **memberikan rekomendasi Travel Insurance** berdasarkan data yang tersedia untuk **memberikan wawasan lebih dalam membimbing strategi pemasaran di masa depan.**

## Meet The Team

### Project Team Leader



M. Faiz



Chianti Ridwan



Retno Debbi



Imam Luthfi



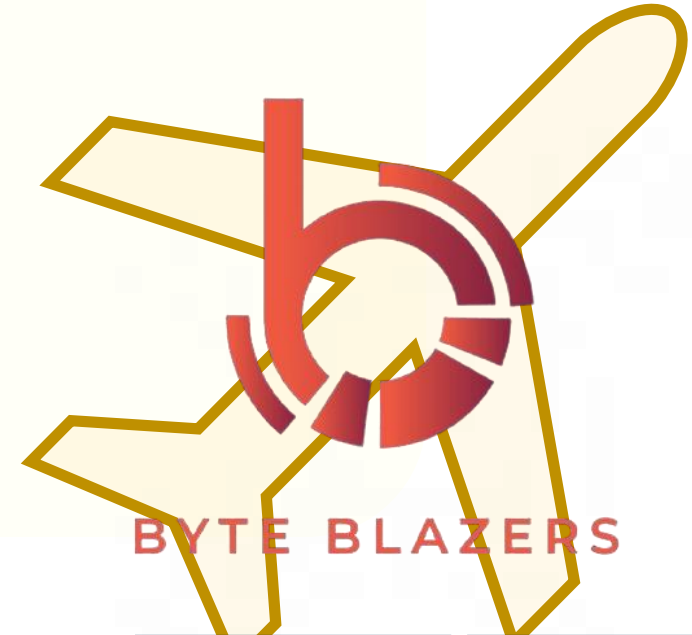
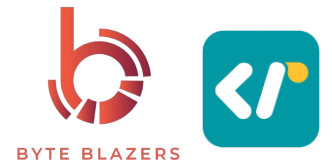
Lulu Safira



Robiatul Adawiyah



Melliza Nastasia



# About The Client

## TravelGuard+

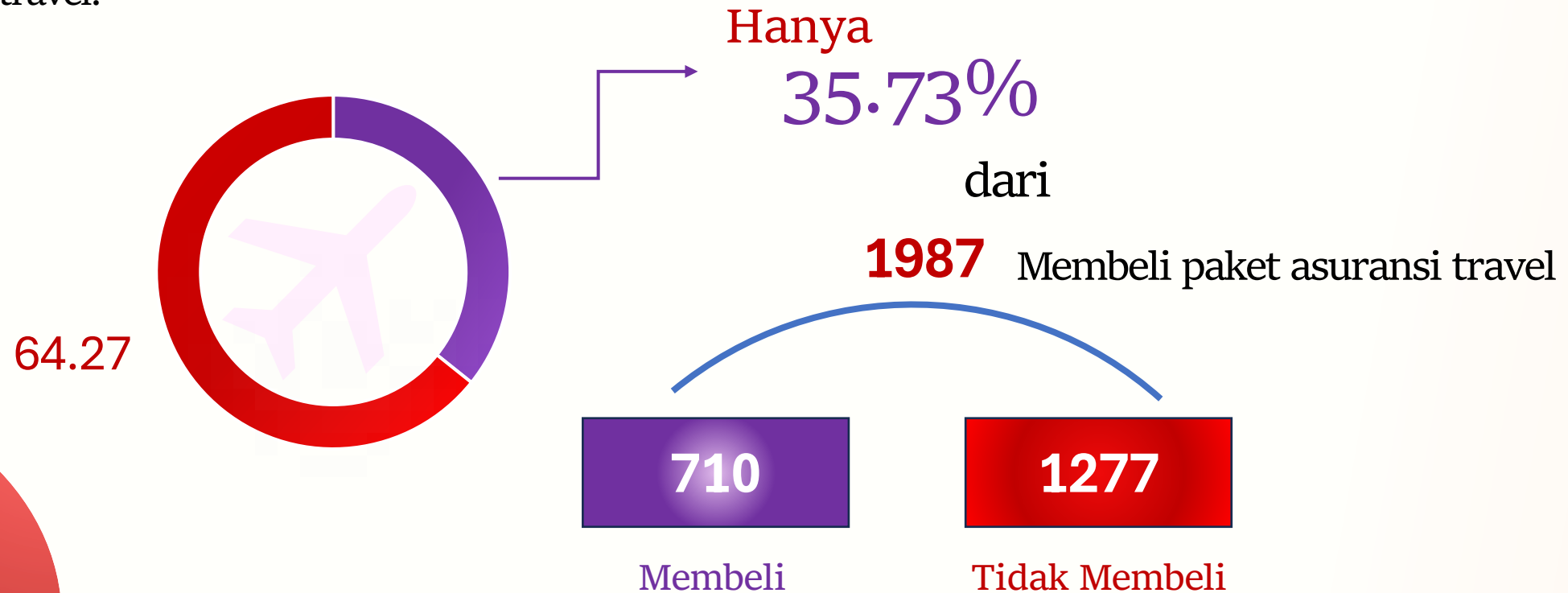
**TravelGuard+** adalah Perusahaan yang menawarkan **paket asuransi**

Berdasarkan data yang diambil pada tahun 2019, **TravelGuard+** ingin mengidentifikasi pelanggan yang sudah pasti akan membeli **paket asuransi travel**

**TravelGuard+** meminta **Byte Blazers** untuk membantu mengidentifikasi pelanggan yang akan membeli **paket asuransi travel** yang mereka tawarkan

# What is The Problem

Adanya **penurunan minat** pelanggan untuk membeli asuransi travel perjalanan ke luar negeri akibat adanya pandemi COVID-19. Bahkan dari total 1987 pelanggan, **hanya 710 orang** (35,73%) yang memiliki asuransi travel.





# What is **The Problem**

## Background

TravelGuard+ ingin mengetahui pelanggan mana yang sudah pasti akan membeli **paket asuransi travel**

## Goals

Meningkatkan penjualan paket asuransi perjalanan baru dalam satu tahun ke depan berdasarkan analisis pola pembelian pelanggan (pada dataset).

## Objective

Membuat model machine learning yang dapat membantu bisnis TravelGuard+ ini memprediksi pelanggan yang akan membeli paket asuransi travel.

## Business Metrics

Jumlah transaksi user.

# Exploratory Data Analysis

Pendeteksian data bermasalah dan insight yang bermanfaat.



**About Dataset**



**Preliminary Insight**

# Datasets



Unnamed : 0  
Sebagai ID



Employment Type



GraduateOrNot



FrequentFlyer



EverTravelAbroad



Age



AnnualIncome



FamilyMembers



ChronicDiseases



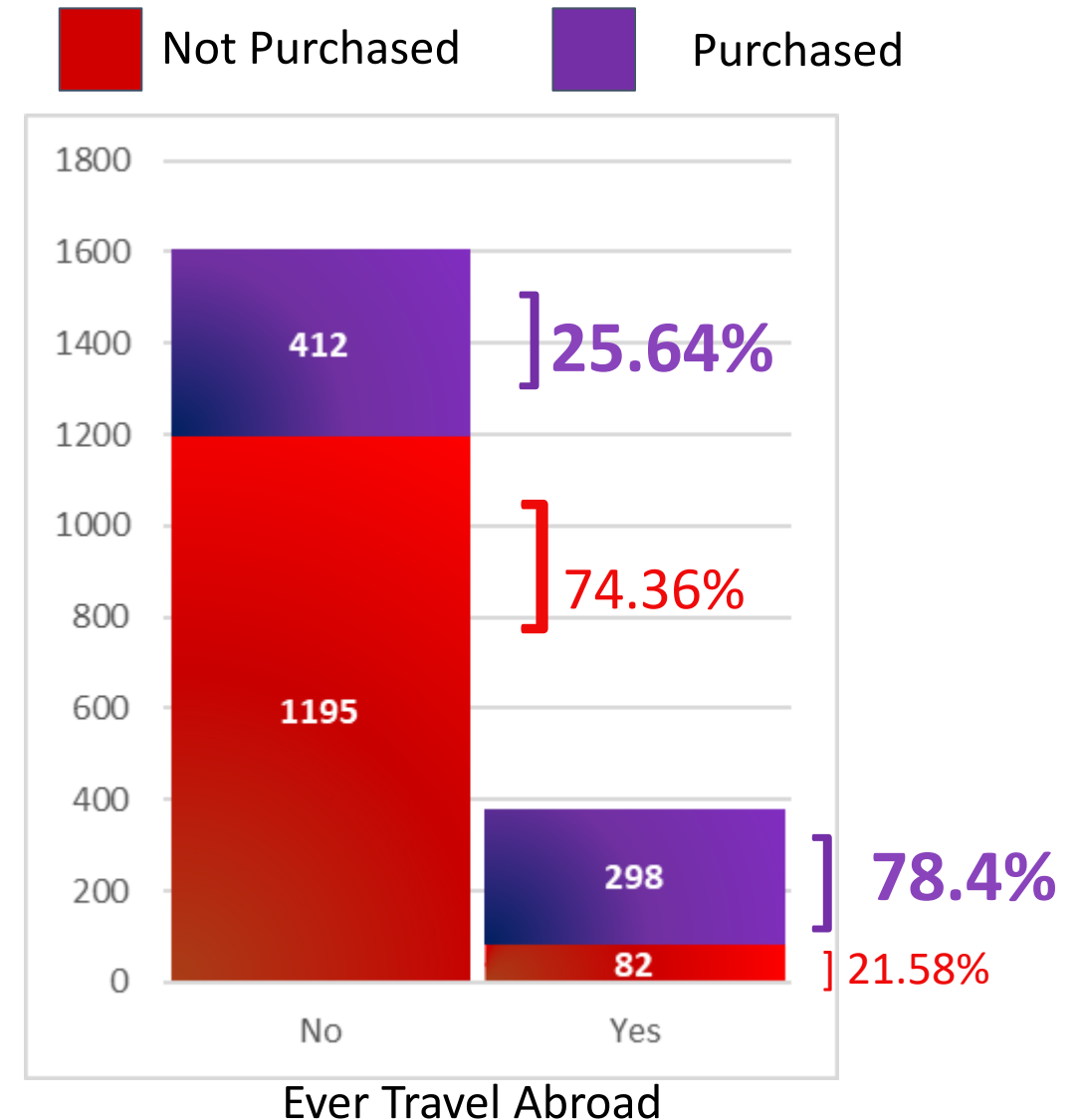
TravelInsurance

Dataset terdiri dari **1987 baris** dan **10 kolom**.

Tidak terdapat missing values,  
tidak terdapat duplicate data, dan tipe data sudah  
sesuai

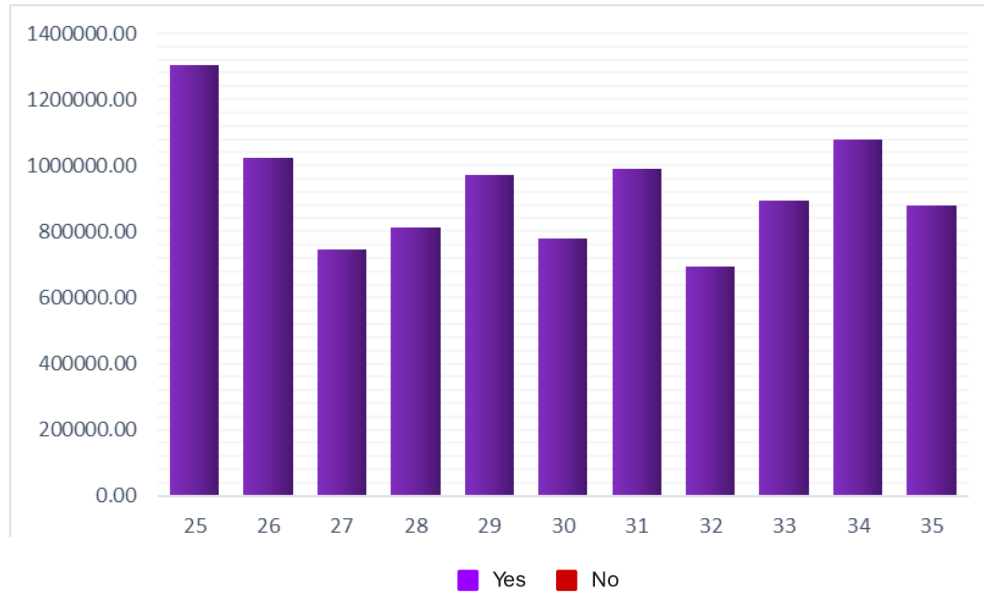
# Ever Travel Abroad To Travel Insurance

Mayoritas pelanggan tidak pernah keluar negeri, akan tetapi 78.4% dari pembeli asuransi adalah pelanggan yang pernah keluar negeri

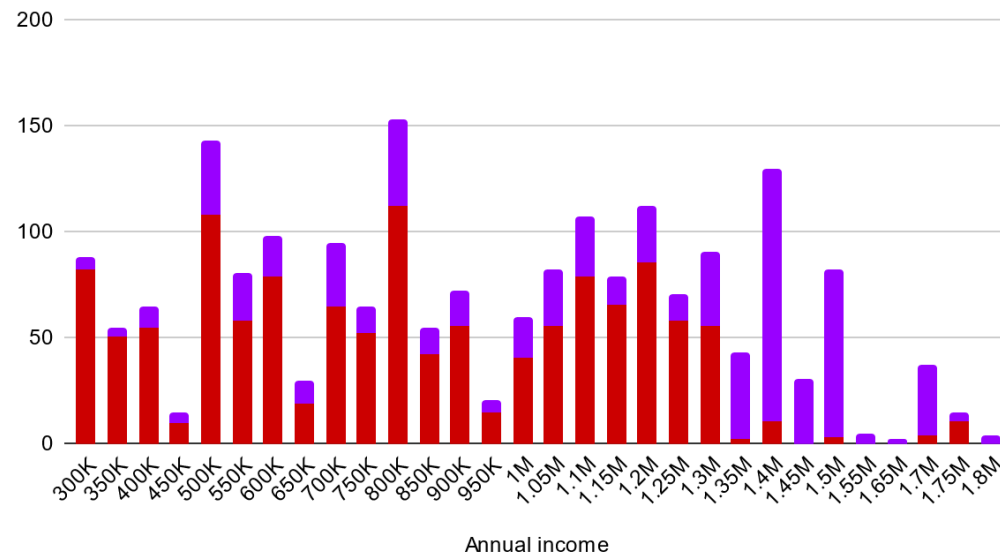




# Annual Income



Rata-rata **income tertinggi** berada di customer dengan umur 25 tahun yaitu 1.303.424

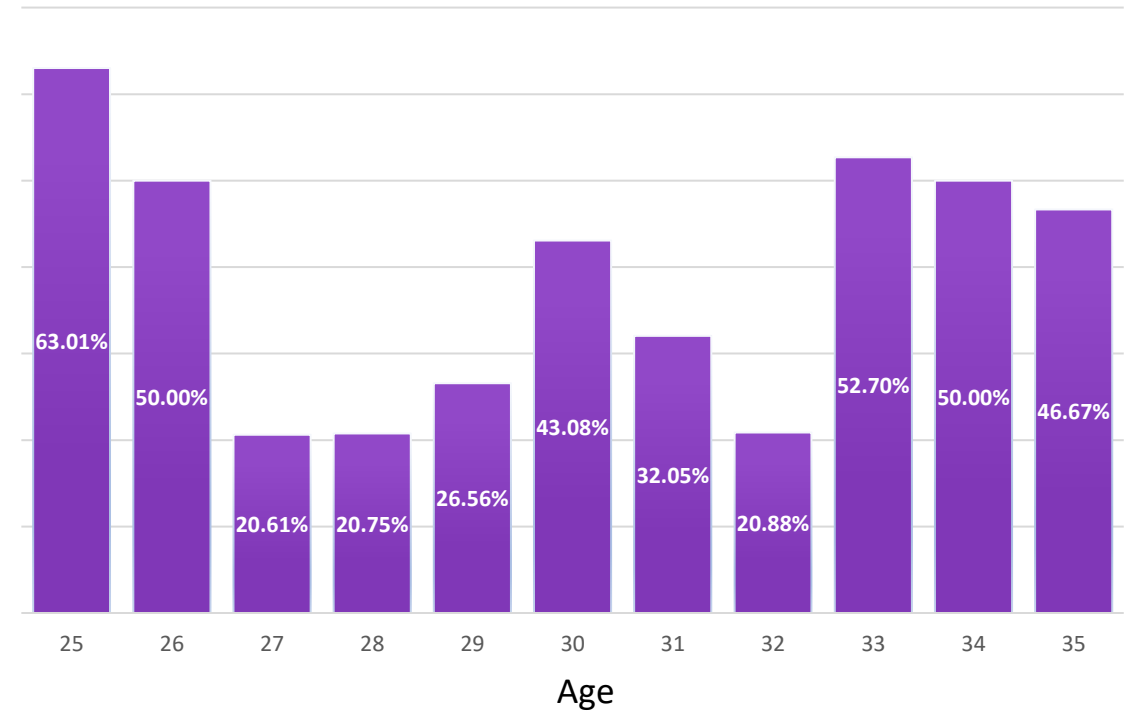


Pelanggan dengan Annual Income **diatas 1.3M** memiliki **potensi tinggi** untuk membeli paket asuransi

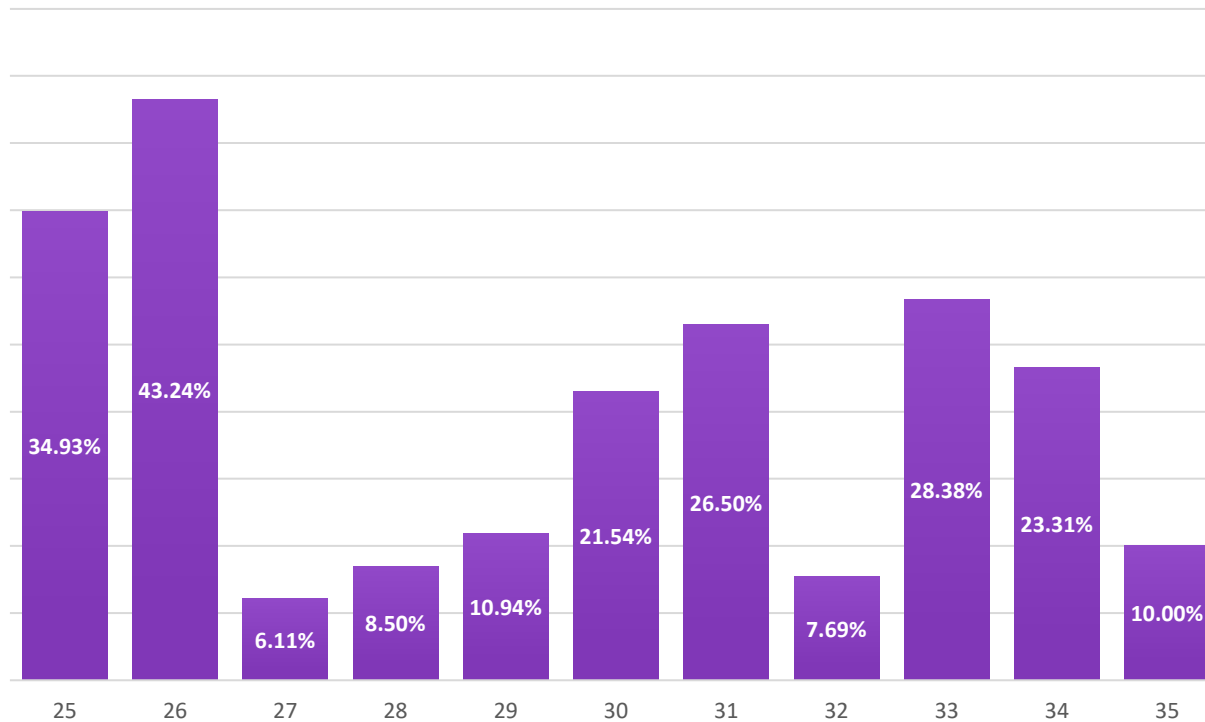
Sementara pelanggan dengan annual income 300k - 400k memiliki **potensi terendah** untuk membeli asuransi

# Age To Travel Insurance

Mayoritas yang membeli berada di usia 25 tahun.



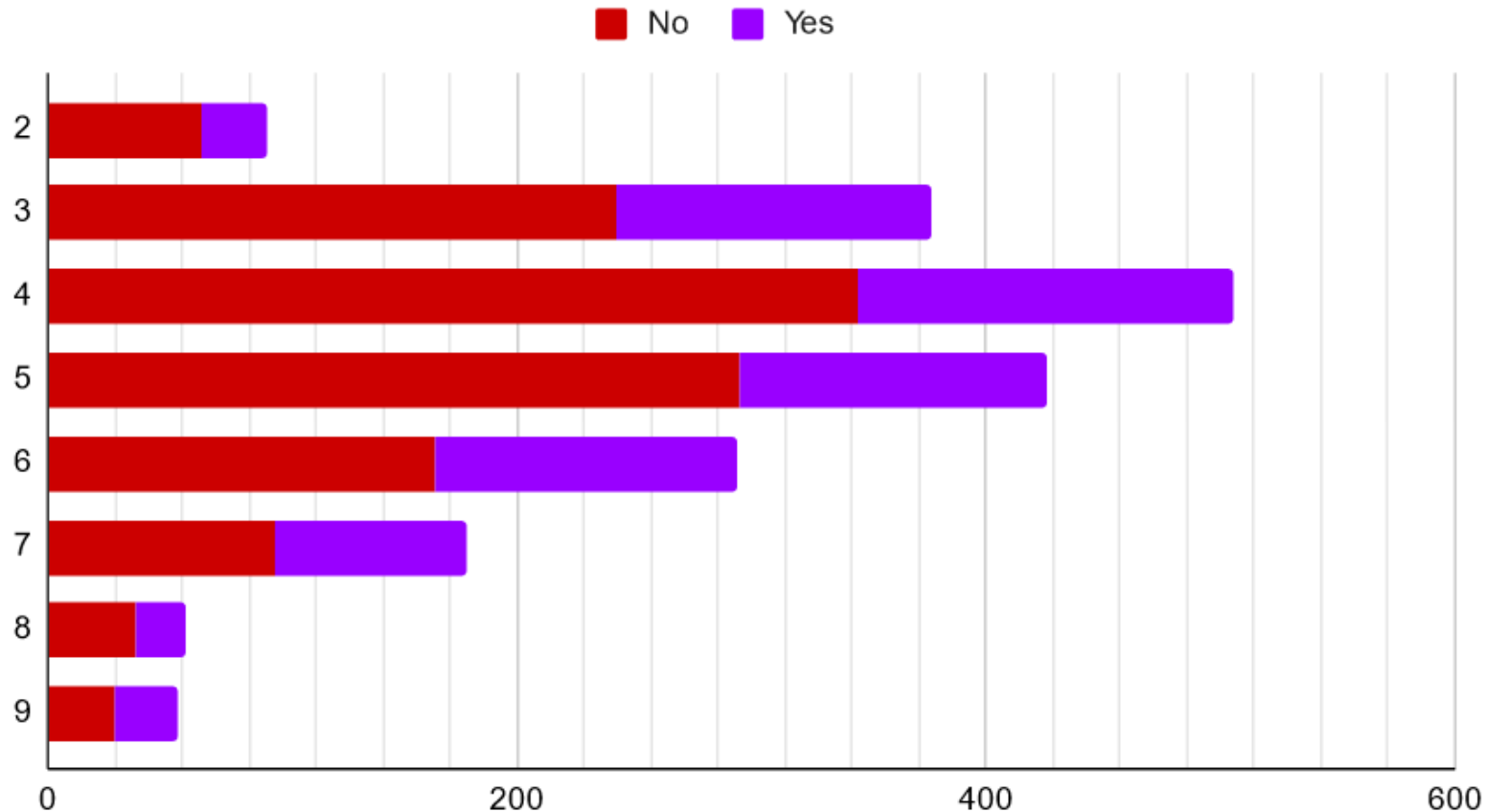
# Age To Ever Travel Abroad



Mayoritas pelanggan yang berada di usia 27 tahun **tidak pernah** berpergian jauh.

# Family Members To Travel Insurance

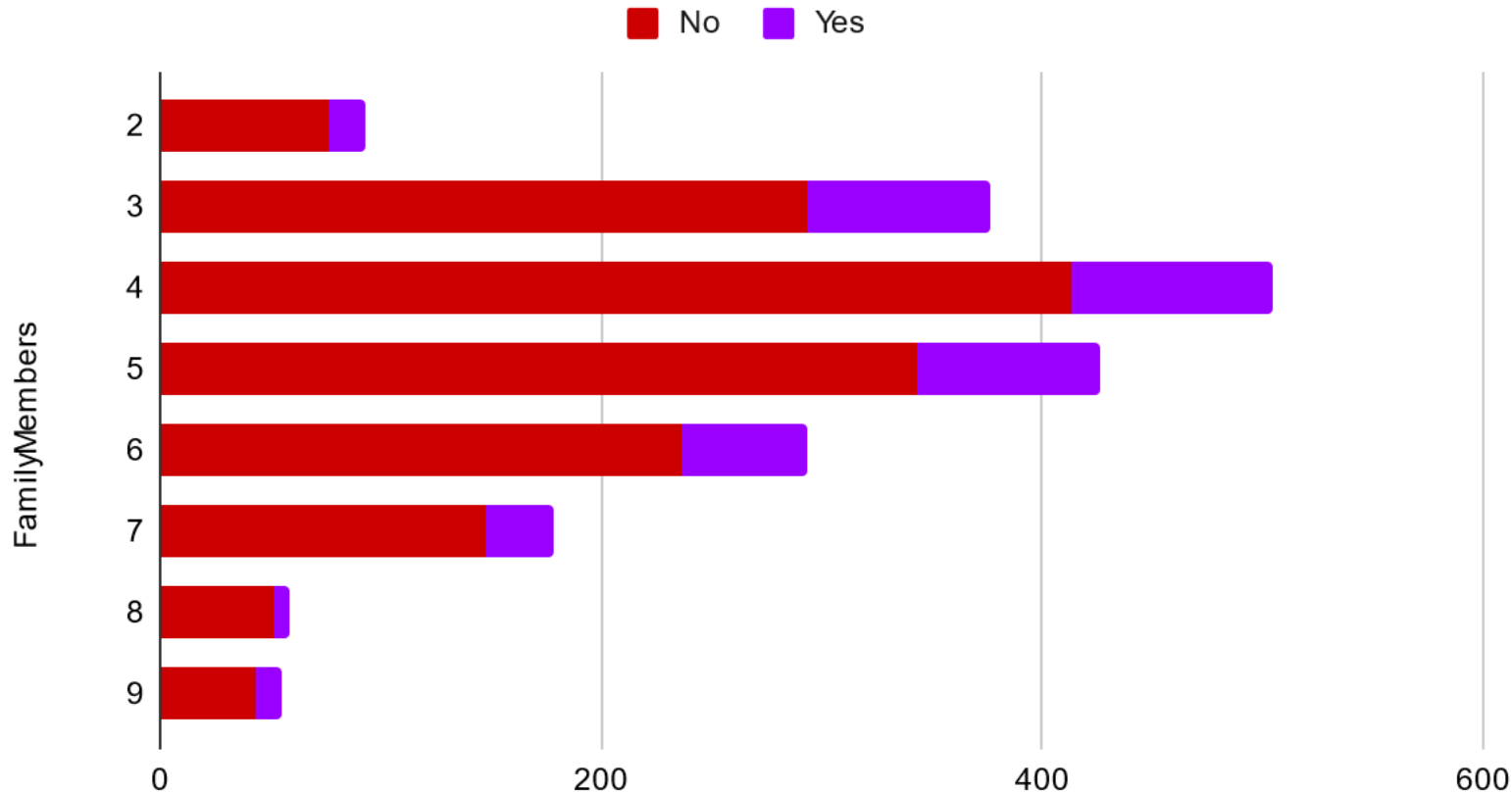
FamilyMembers To TravelInsurance



Dari semua kelompok Family Member, transaksi **tidak membeli paket** asuransi **lebih tinggi** dibandingkan transaksi yang **membeli paket**.

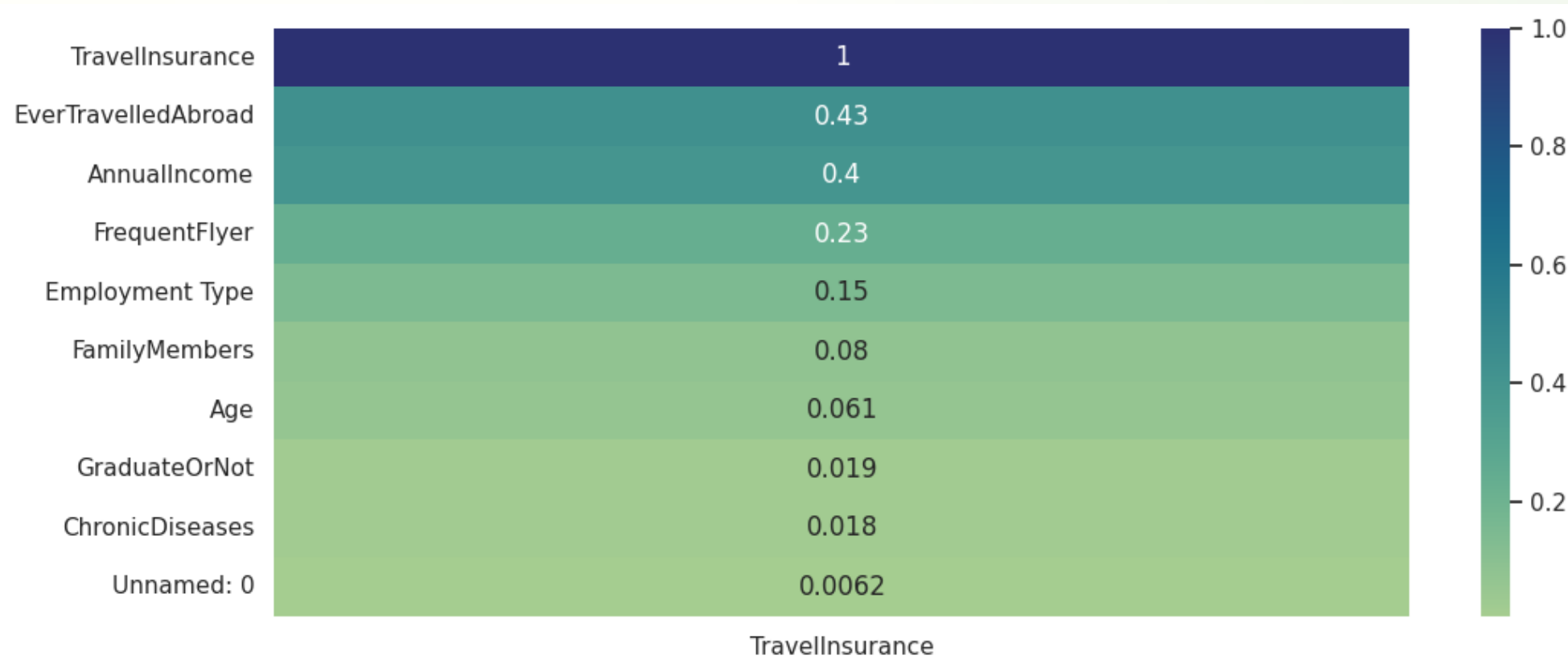
# Family Members To Travel Abroad

FamilyMembers to TravelAbroad



Transaksi pelanggan yang **tidak pernah** Travel Abroad lebih **tinggi** dibandingkan yang **pernah** melakukan Travel Abroad

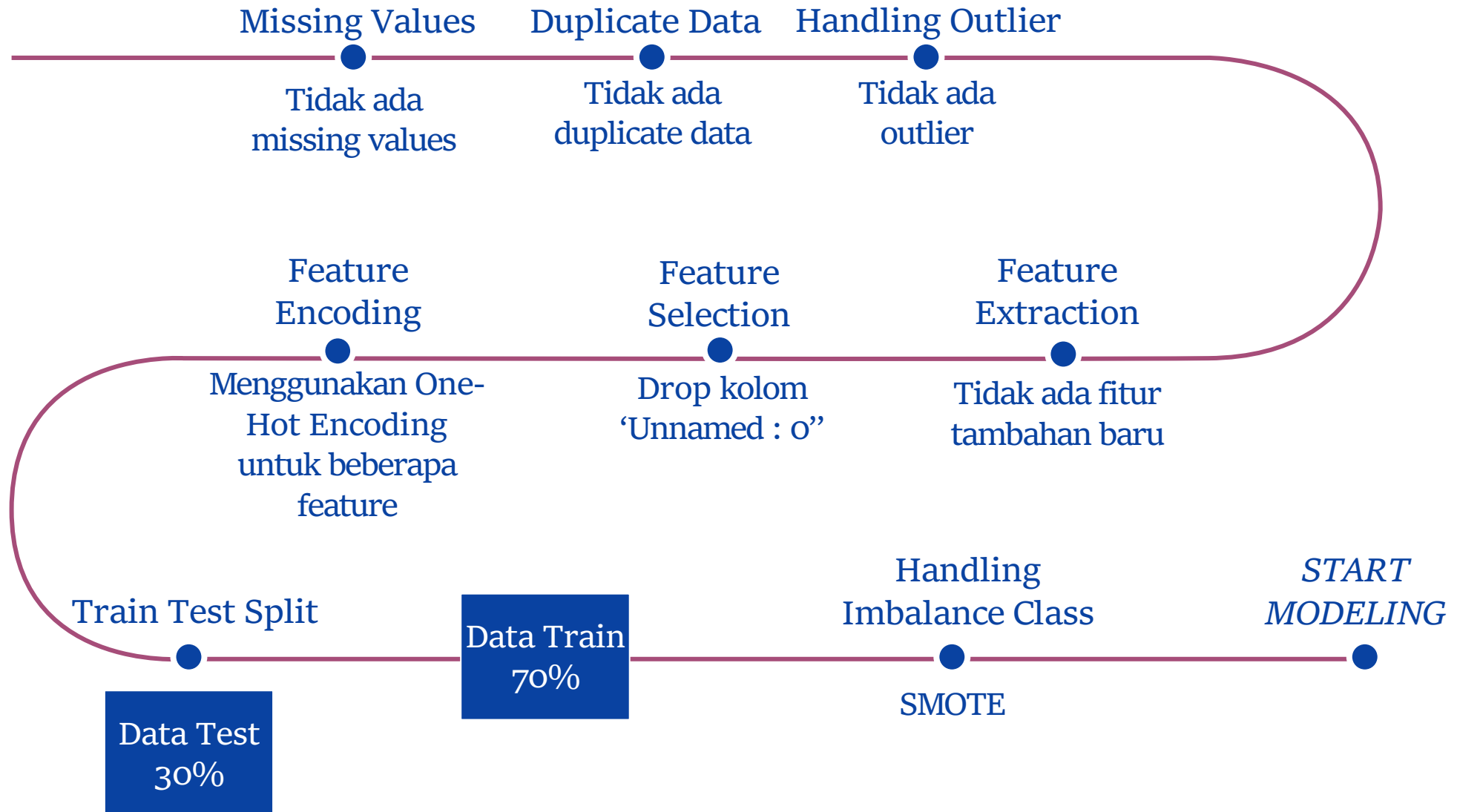
## Korelasi masing-masing Feature Terhadap target (Travel Insurance)



EverTravelAbroad memiliki korelasi paling tinggi disusul dengan AnnualIncome



# Pre-Processing Flow



# Machine Learning Modelling



**Model Evaluation**



**Feature Importance**

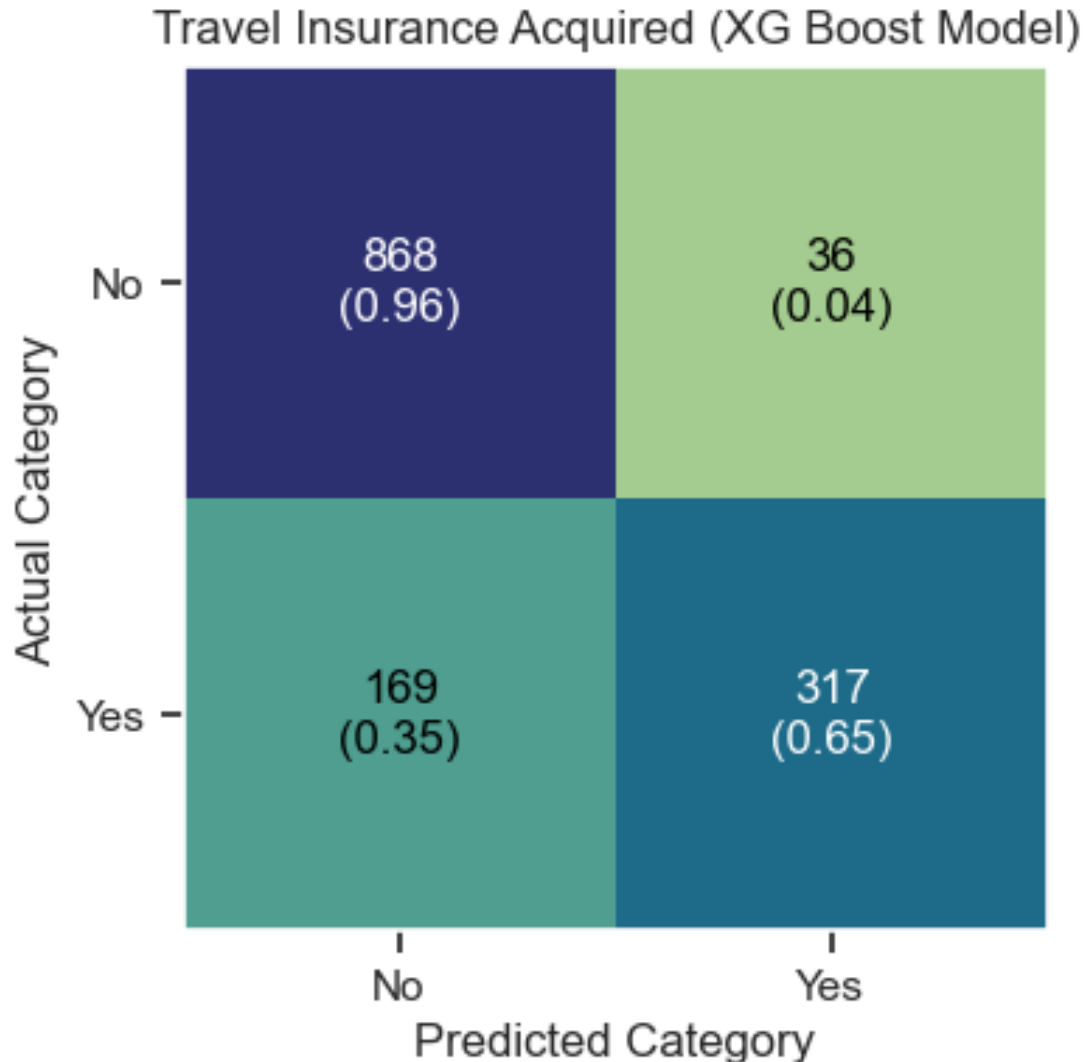


# Model Evaluation

| Model               | Before Hyperparameter Tuning |               | After Hyperparameter Tuning |               |
|---------------------|------------------------------|---------------|-----------------------------|---------------|
|                     | Recall (Train)               | Recall (Test) | Recall (Train)              | Recall (Test) |
| Logistic Regression | 0.55                         | 0.58          | 0.47                        | 0.52          |
| Random Forest       | 0.86                         | 0.67          | 0.59                        | 0.60          |
| Gradient Boosting   | 0.65                         | 0.63          | 0.58                        | 0.59          |
| XG Boost            | 0.81                         | 0.65          | 0.65                        | 0.61          |

Berdasarkan analisis terhadap metrik recall dan metrik lainnya, model XG Boost menjadi model terbaik di antara keempat model yang telah di evaluasi.

## Confusion Matrix



Kesimpulan:

- Model tampaknya **memiliki performa yang cukup baik** pada data pelatihan dan data pengujian, dengan akurasi sekitar 85.25% pada data pelatihan dan 80.90% pada data pengujian.

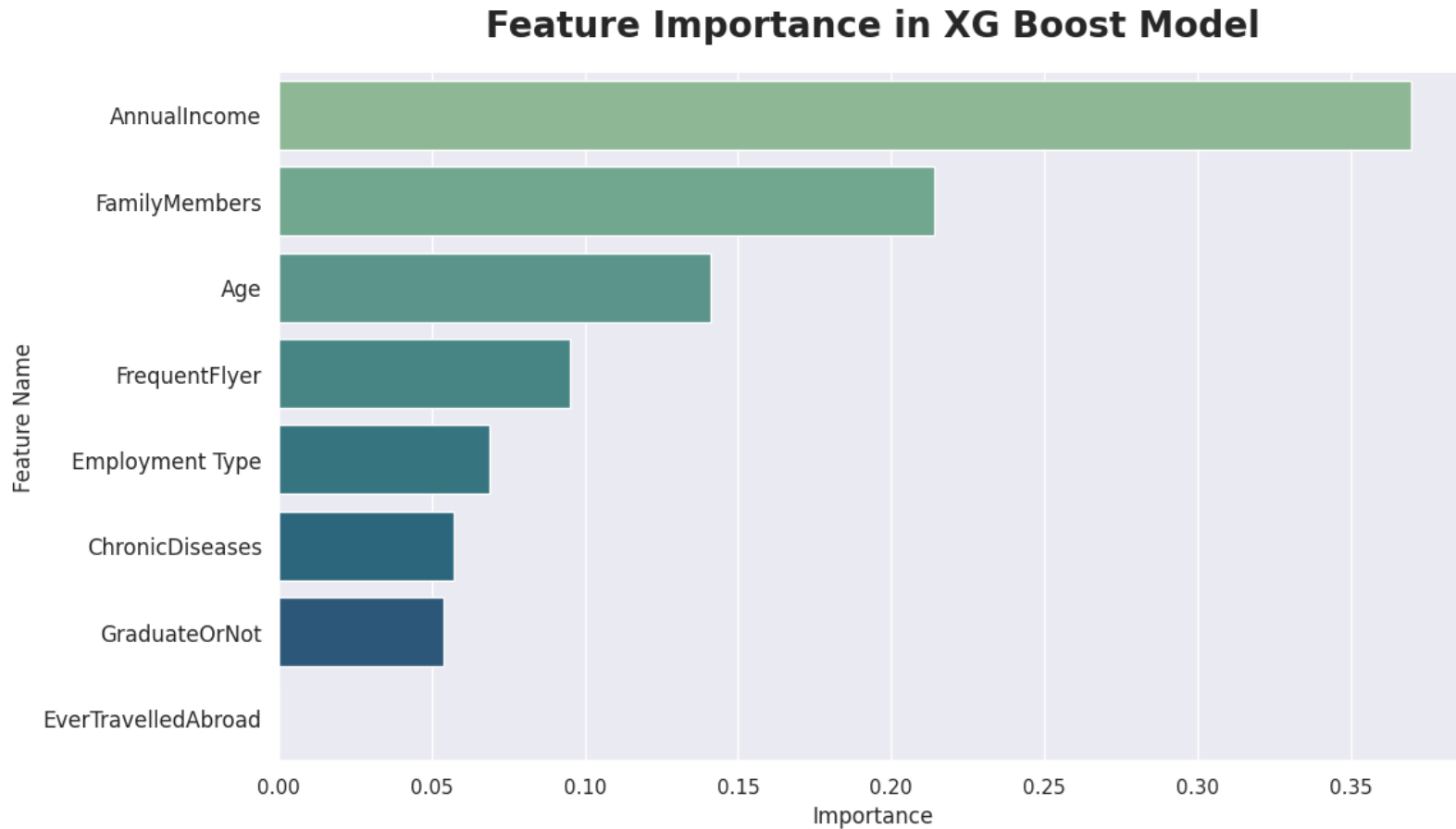
$$\text{Recall} = \text{TP} / \text{TP} + \text{FN}$$

$$= 317 / (317 + 169) * 100$$

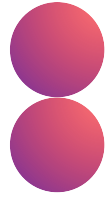
$$= 65,22\%$$

**from 35.73% to 65.22% → Increase 29.49%**

# Feature Importance



Berdasarkan feature importance score disamping, dapat dilihat feature [AnnualIncome](#), [FamilyMembers](#), dan [Age](#) merupakan **top feature importance** yang dapat kita jadikan fokus untuk mendapatkan hasil prediksi yang akan membeli asuransi travel.



# Business Recommendation

untuk TravelGuard+







# 1

Pelanggan dengan kategori **mid to high annual income** berpotensi untuk membeli paket asuransi travel. Oleh karena itu tim marketing dapat membuat pendekatan strategis untuk menarik segmen pelanggan ini dengan **memberikan perlakuan khusus yang bisa didapatkan oleh mereka (VIP Poin)**

Membuat pilihan paket untuk perjalanan domestic dan internasional

# 2

# 3

Mempertimbangkan pembuatan bundling promo untuk segmentasi pelanggan yang bepergian secara berkelompok (baik keluarga maupun instansi pemerintahan/swasta)



4

Melakukan **Strategi Pemasaran** untuk menarik pelanggan dalam kelompok Usia 25 tahun, karena mereka menunjukkan kemungkinan tertinggi untuk membeli paket Travel Insurance. Oleh sebab itu **kelompok tersebut dilibatkan iklan dan promosi sebagai strategi pemasaran.**

Mempertimbangkan **pengelompokkan** pelanggan berdasarkan **riwayat perjalanan, pendapatan tahunan** dan **anggota keluarga** dalam paket Travel Insurance

5



# Terima Kasih