



Uno Inc.

POTENTIAL CUSTOMER ANALYSIS AND PREDICTION

Analysis Report



MEET THE TEAM

Uno Inc.



**Devriansyah
Sya'ban**

DATA SCIENCE
LEADER



Vira Diana

DATA ANALYST



Adinda Dita

DATA ANALYST



Kun Aggiar

STATISTICIAN



Rahmadani A

STATISTICIAN



Laurenzius Julio

ML ENGINEER



**Farih Afdhalul
Ihsan**

ML ENGINEER



Dignu Akbar

DATA SCIENTIST

Who We Are?

Konsultan yang bergerak di bidang data untuk membantu dalam melakukan analisis dan pemecahan masalah perusahaan Anda.

OUTLINE

-  **PROBLEM STATEMENT**
-  **DATA OVERVIEW & INSIGHT**
-  **DATA PRE-PROCESSING**
-  **MODEL & EVALUATION**
-  **BUSINESS RECOMMENDATION & IMPACT**

OUTLINE



PROBLEM STATEMENT



DATA OVERVIEW & INSIGHT



DATA PRE-PROCESSING



MODEL & EVALUATION



BUSINESS RECOMMENDATION & IMPACT

Business Understanding

Perusahaan X merupakan perusahaan yang bergerak dibidang **Tour and Travel**

Guna **meningkatkan profit** perusahaan, selain memiliki produk utama berupa layanan/jasa travel, Perusahaan X juga mengambil peluang kesempatan dengan **menambahkan segmen bisnis asuransi (*travel insurance*)** yang berupa penyediaan layanan asuransi (mencakup perlindungan terhadap covid 19) bagi para *customer travel* nya.



Uno Inc.

Dalam melakukan campaign terhadap produk asuransinya,
**Perusahaan X selalu memberikan campaign menyeluru
kehada seluruh customernya**

Diasumsikan bahwa ada biaya yang perlu dikeluarkan oleh perusahaan pada setiap aksi campaign (jenis apapun dan cara apapun) yang diberikan / dilakukan terhadap customer



Problem Statement



Campaign Cost Tidak Efisien dikarenakan campaign produk asuransi diberikan **secara menyeluruh** ke semua customer



Conversion Rate saat ini (35,7%) masih di bawah target perusahaan (60%) untuk product travel insurance yang ditawarkan kepada para customer



GOALS



Meningkatkan
conversion rate



Optimasi campaign cost

Objectives

Uno Inc.



Membangun model machine learning yang bisa memprediksi customer potensial untuk membeli paket travel insurance.

Menemukan faktor penting dan karakteristik utama customer yang berpotensi membeli product travel insurance terbaru.

Memberikan insight dan rekomendasi untuk menurunkan kemungkinan perusahaan merugi dalam campaign berdasarkan hasil yang didapatkan dari proses pengolahan data yang dimiliki oleh perusahaan.

Business Metrics



Conversion Rate

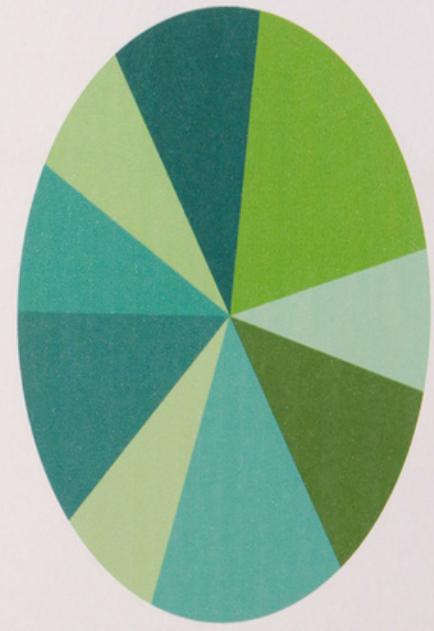


Revenue

Budget Actual Difference



SALE REPORT



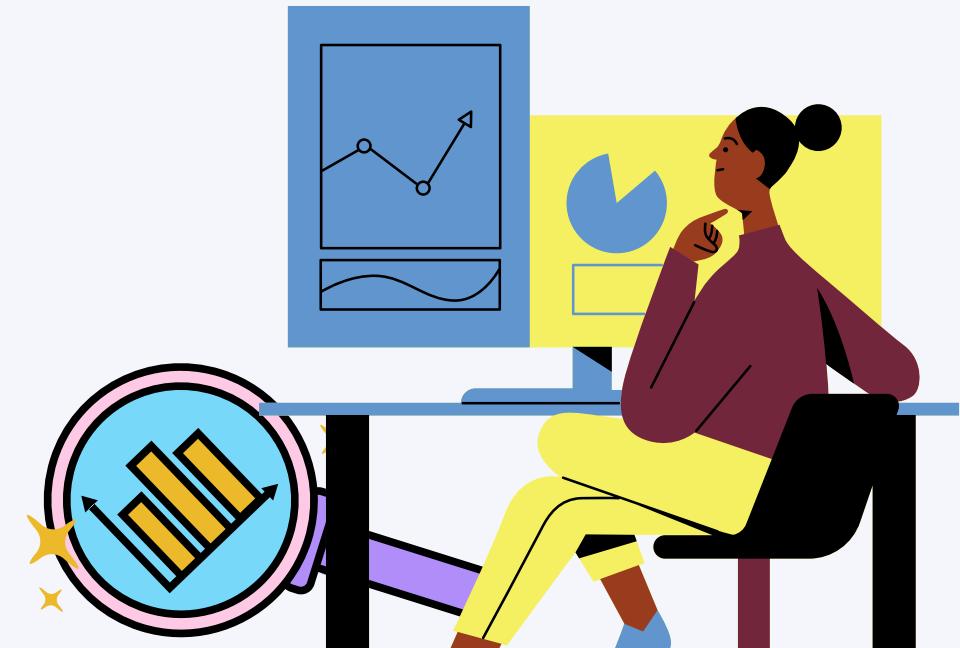
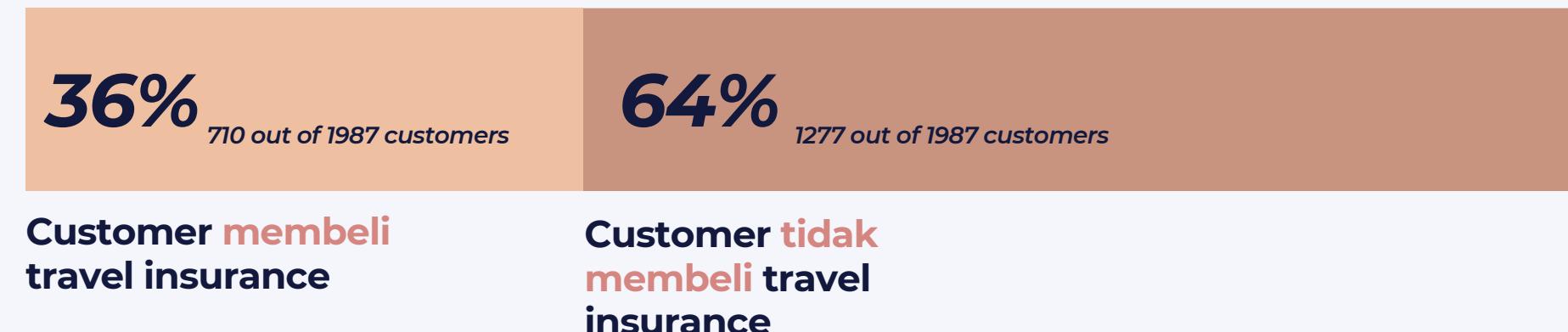
OUTLINE

-  **PROBLEM STATEMENT**
-  **DATA OVERVIEW & INSIGHT**
-  **DATA PRE-PROCESSING**
-  **MODEL & EVALUATION**
-  **BUSINESS RECOMMENDATION & IMPACT**

Data Overview

- 1 Dataset merupakan data penjualan travel insurance sejak 2019
- 2 Dataset terdiri dari **1.986 baris**, **9 kolom feature**, dan **1 kolom target**
- 3 Terdapat **6 feature data numerik** dan **4 feature data kategorikal**
- 4 Terdapat **738 Data Duplikat** dan **Tidak terdapat data null atau kosong**

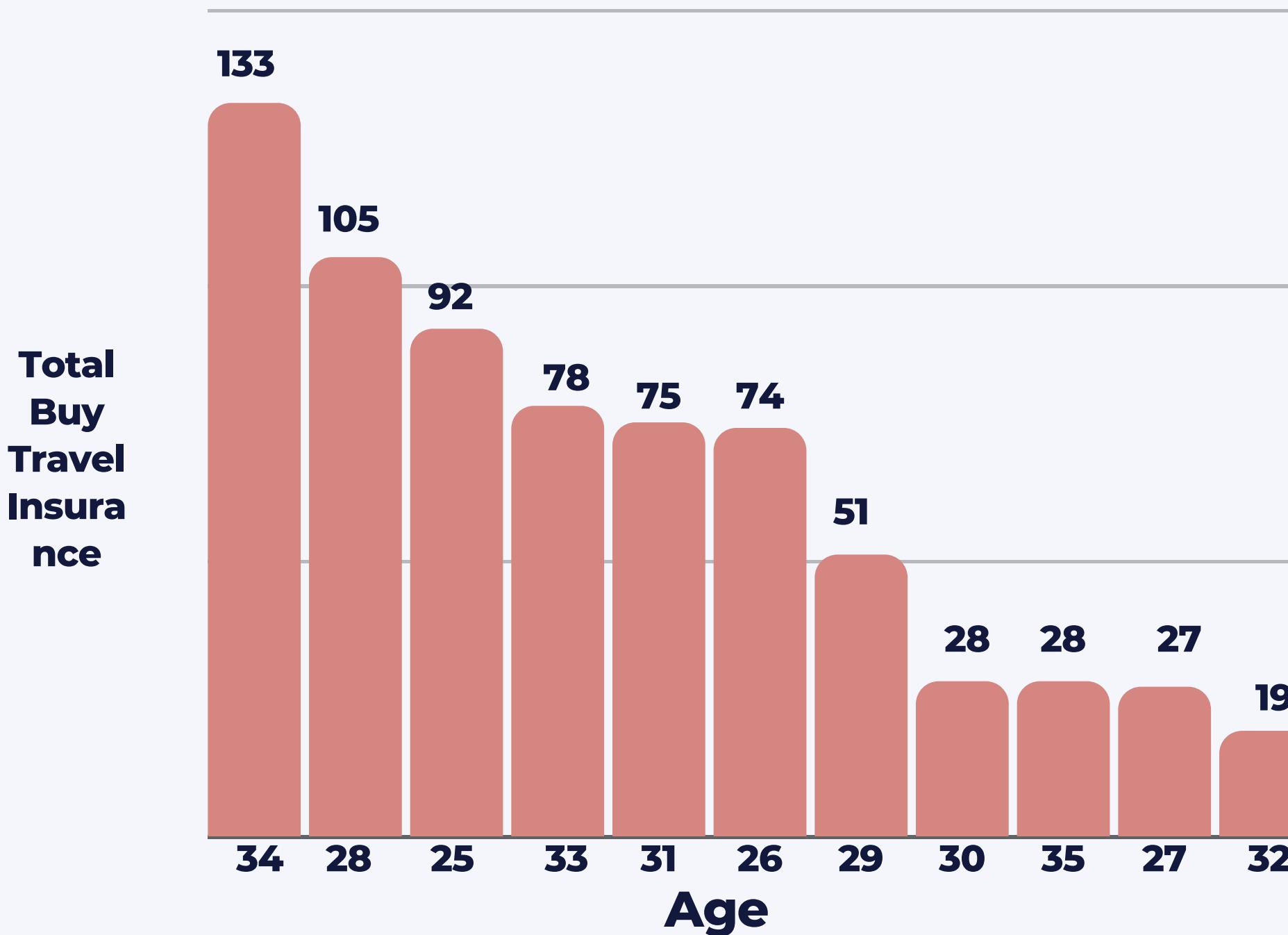
Target Feature



Variabel	Definisi
Age	umur customer
Employment Type	apakah pelanggan bekerja di sektor swasta atau pemerintahan
GraduateOrNot	apakah pelanggan adalah lulusan perguruan tinggi atau tidak
Annual Income	pendapatan tahunan customer
FamilyMembers	Jumlah anggota keluarga yang tinggal bersama customer
ChronicDiseases	apakah customer menderita penyakit kronis atau kondisi seperti diabetes / bp tinggi atau asma, dll.
FrequentFlyer	apakah customer sering menggunakan layanan jasa penerbangan (>4x)
EverTravelledAbroad	apakah customer pernah berpergian ke luar negeri
TravellInsurance	apakah customer membeli asuransi perjalanan atau tidak

Data Insight

Pengaruh **Age** terhadap Pembelian Travel Insurance

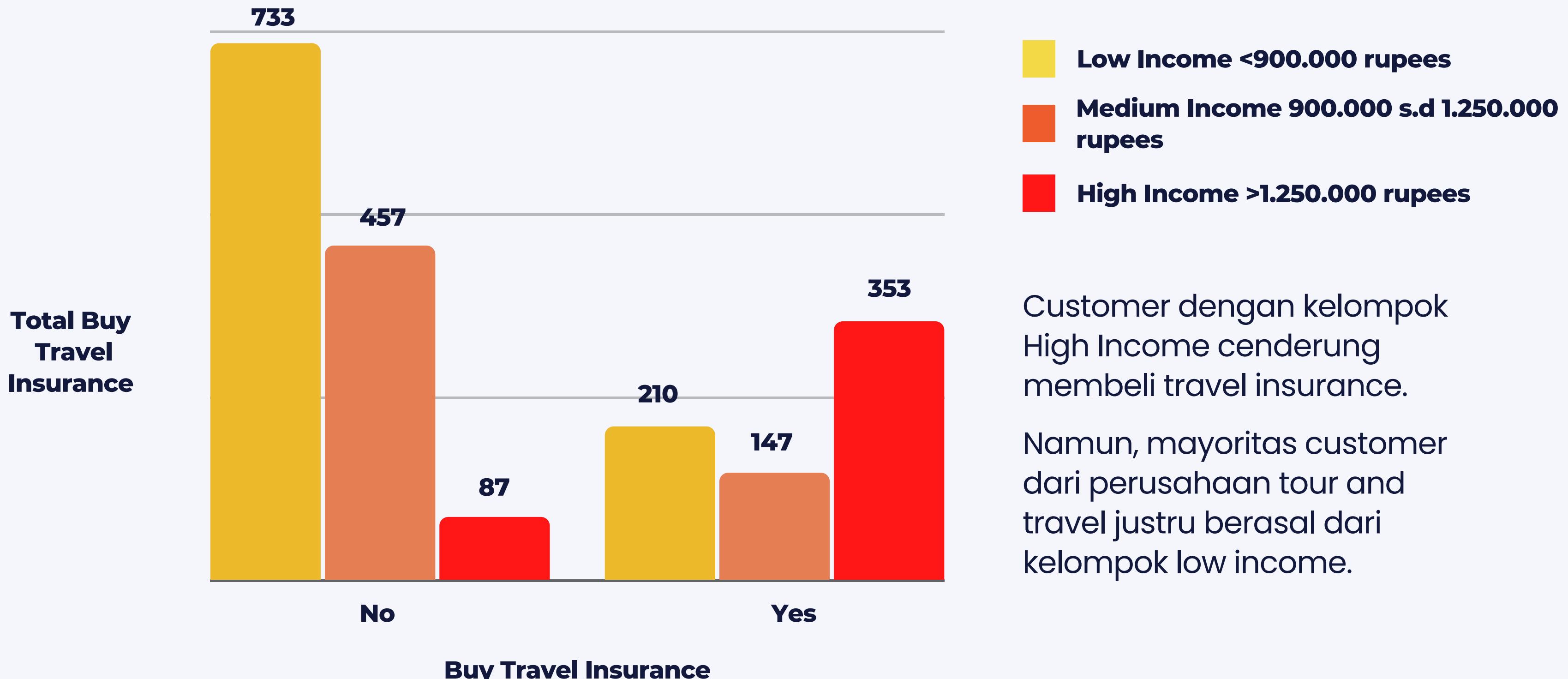


Lima usia tertinggi customer yang memilih untuk membeli paket travel Insurance adalah usia 25, 28, 31, 33, dan 34 dengan pembeli paket Travel Insurance tertinggi adalah customer berusia 34 tahun.

Data juga menunjukkan adanya **dua kelompok usia** customer yaitu **generasi milenial (34 - 28)** dan **generasi Z (27-14)**

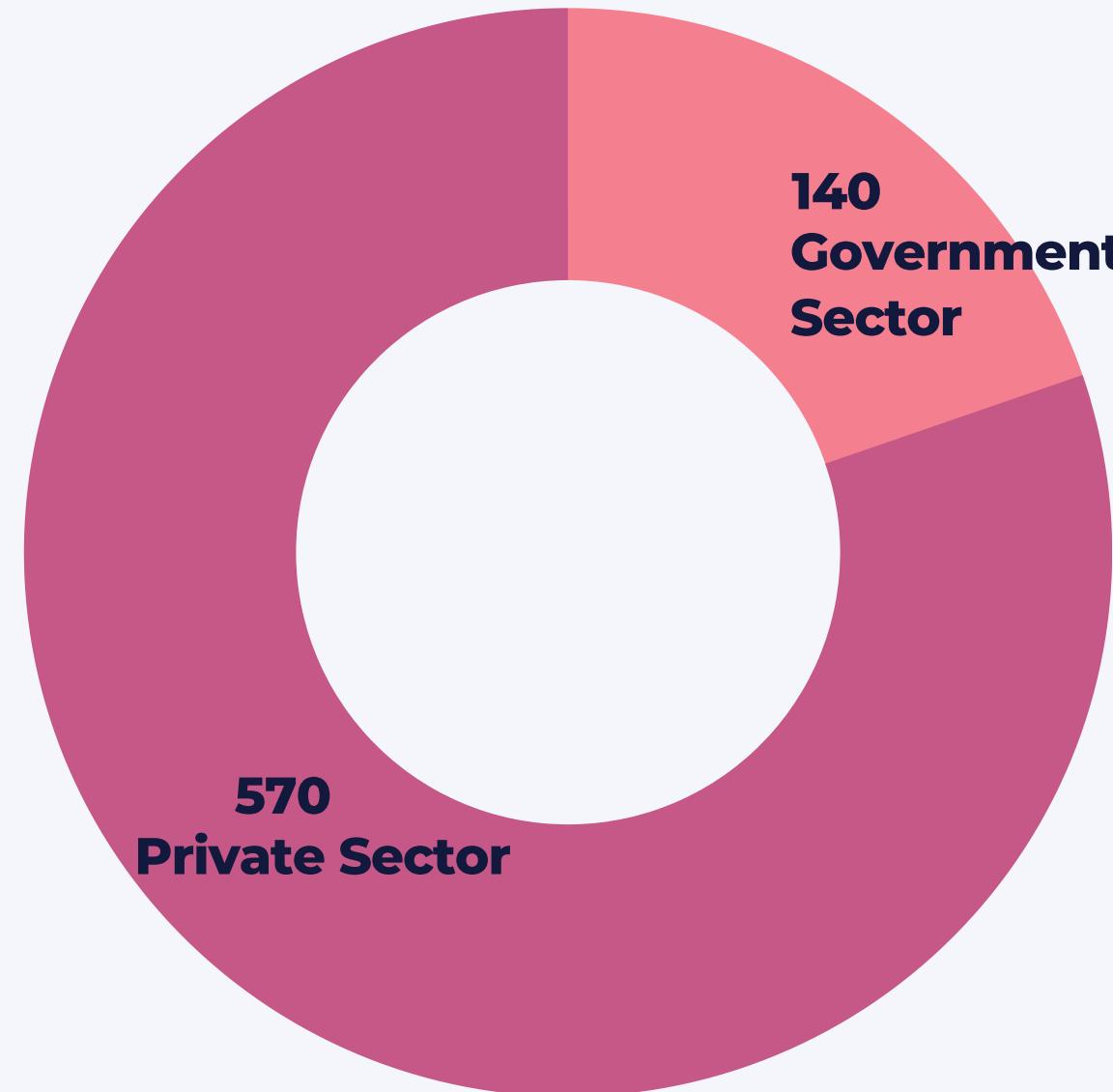
Data Insight

Pengaruh **Annual Income** terhadap Pembelian Travel Insurance



Data Insight

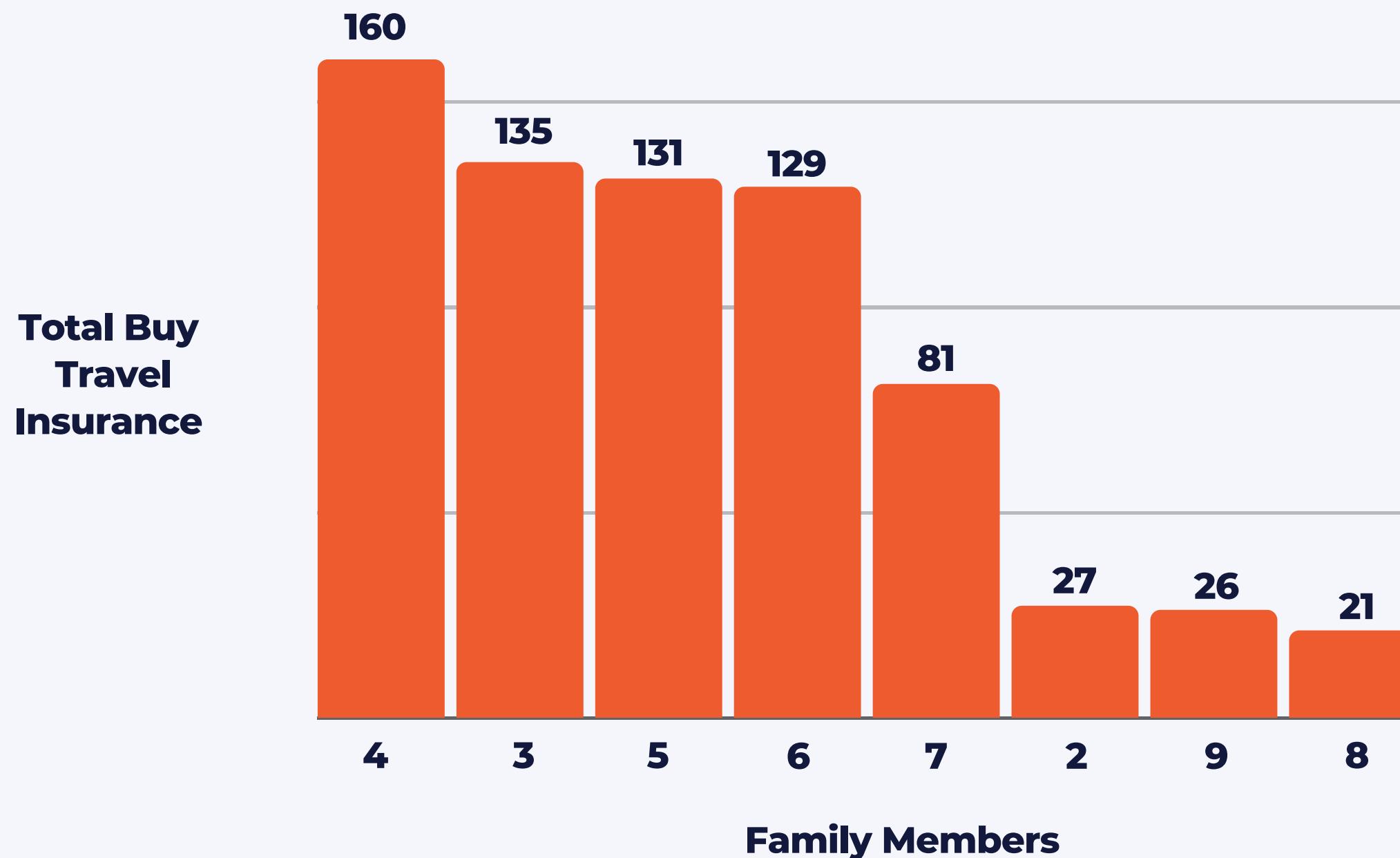
Pengaruh **Employment Type** terhadap Pembelian Travel Insurance



Mayoritas customer yang
membeli travel insurance
bekerja di **Private
Sector/Self Employed**

Data Insight

Pengaruh **Family Member** terhadap Pembelian Travel Insurance



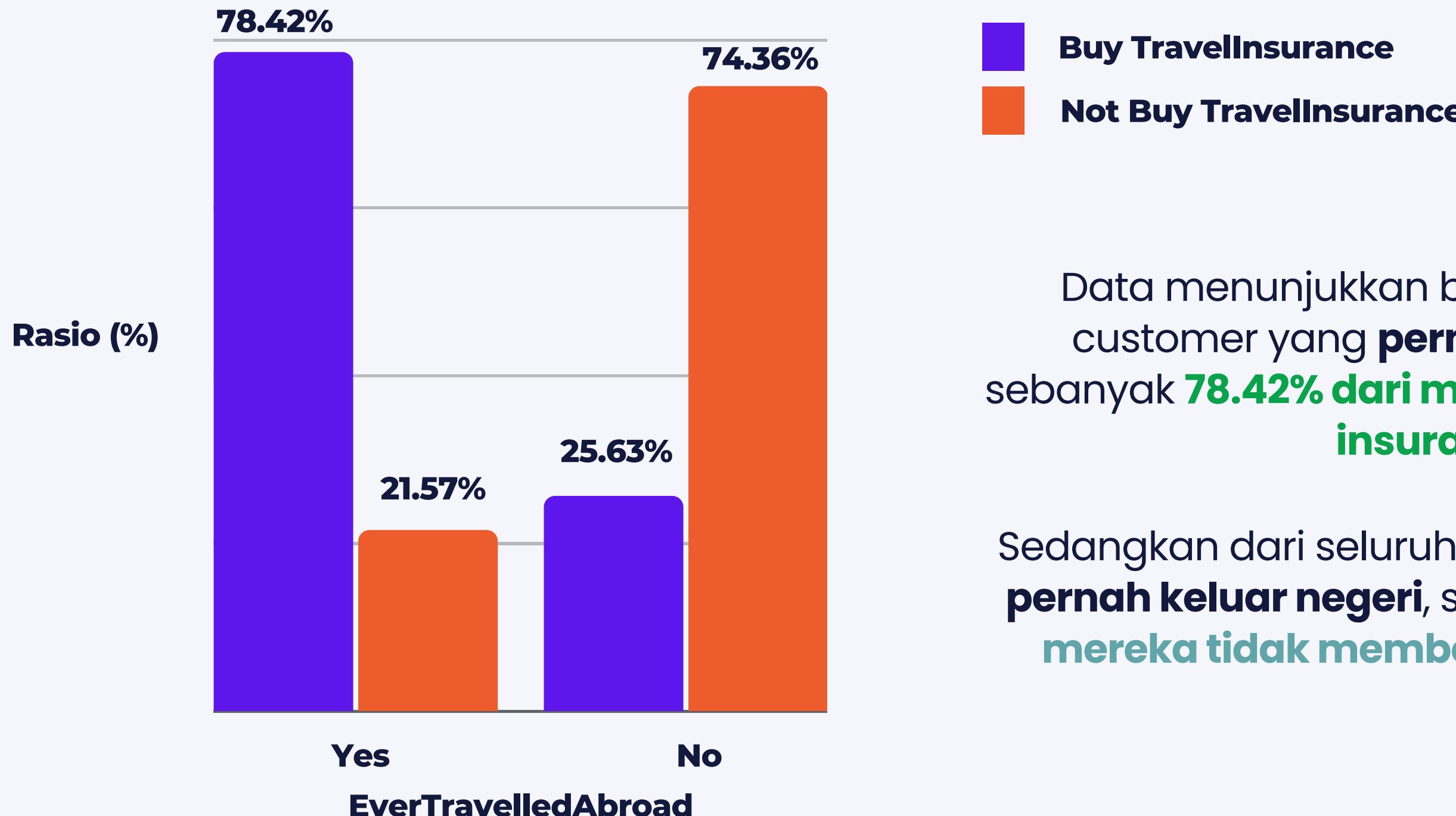
Mayoritas customer yang membeli travel insurance memiliki jumlah anggota keluarga 4 orang.

Dapat disimpulkan bahwa jumlah anggota keluarga tidak terlalu berpengaruh terhadap travel insurance.

Customer dengan 8 hingga 9 keluarga tidak membeli banyak travel insurance dibandingkan keluarga lainnya

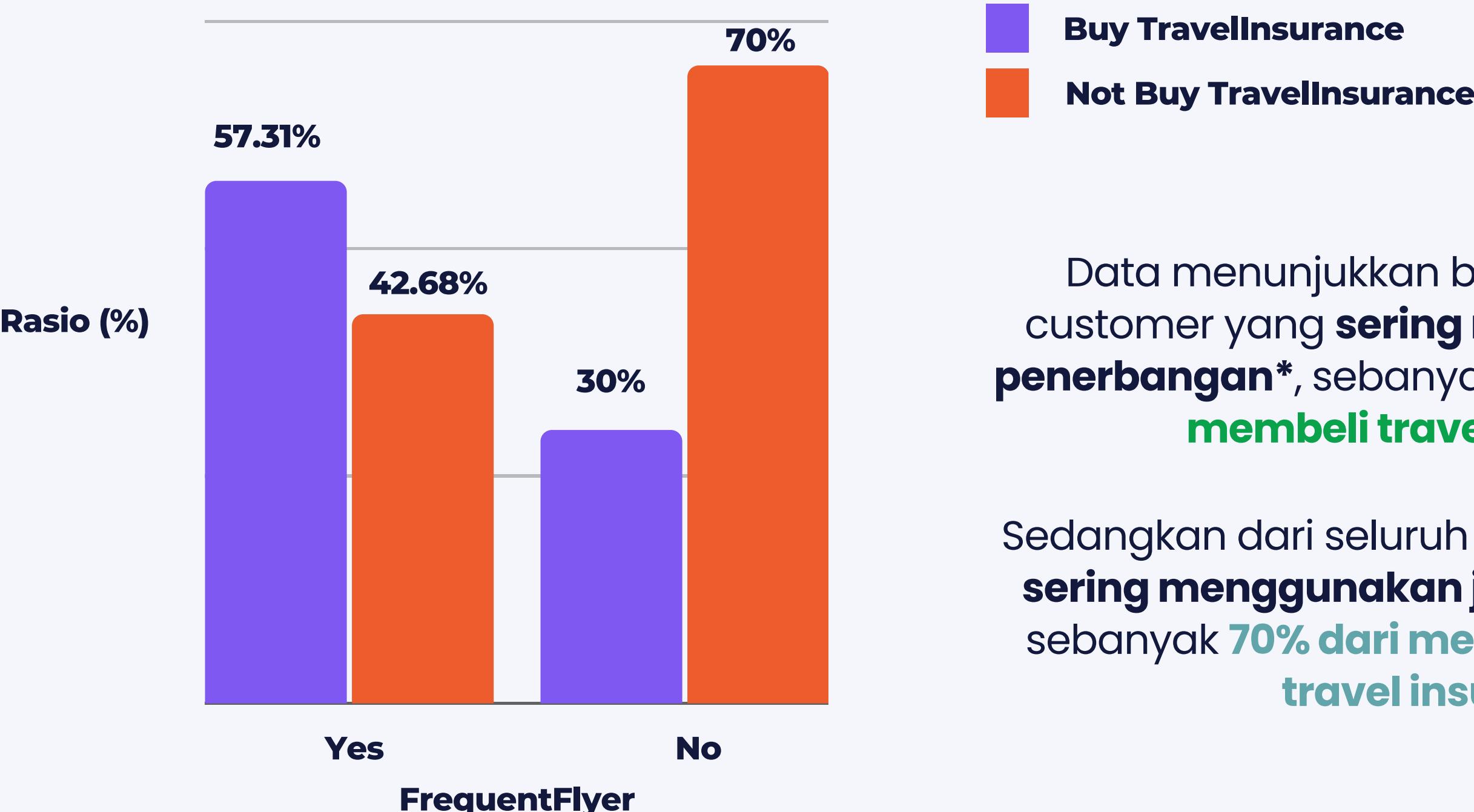
Data Insight

Pengaruh **Ever Travelled Abroad** terhadap Pembelian Travel Insurance



Data Insight

Pengaruh **FrequentFlyer** terhadap Pembelian Travel Insurance

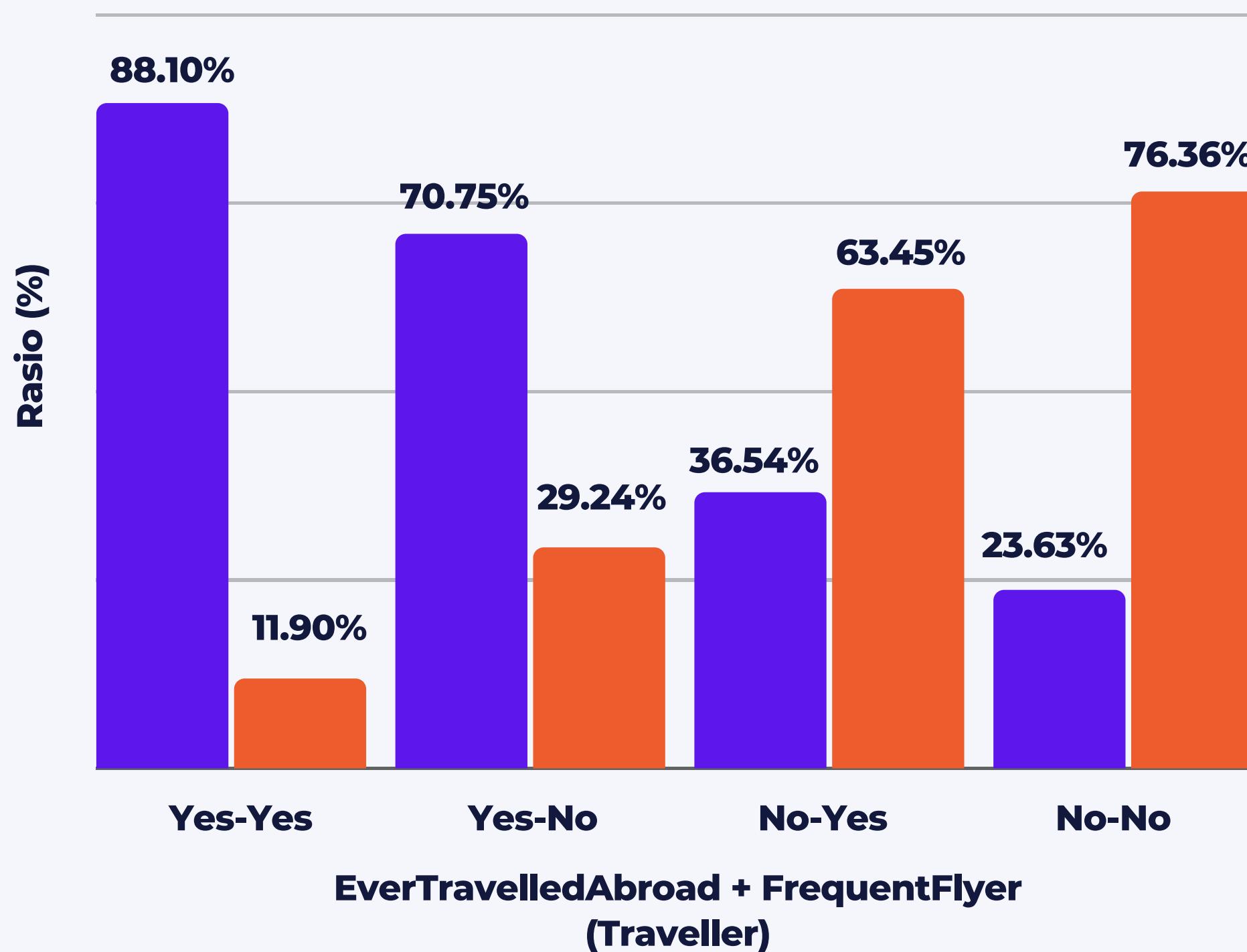


Data menunjukkan bahwa dari seluruh customer yang **sering menggunakan jasa penerbangan***, sebanyak **57.31% dari mereka membeli travel insurance**

Sedangkan dari seluruh customer yang **tidak sering menggunakan jasa penerbangan***, sebanyak **70% dari mereka tidak membeli travel insurance**

Data Insight

Pengaruh **EverTravelledAbroad + FrequentFlyer (Traveller)** terhadap Pembelian Travel Insurance



- █ Buy TravellInsurance
- █ Not Buy TravellInsurance

Data menunjukkan bahwa dari seluruh customer yang **pernah keluar negeri + sering menggunakan jasa penerbangan***, maka sebanyak **88.10% dari mereka membeli travel insurance!**

Sedangkan dari seluruh customer yang **tidak pernah keluar negeri + tidak sering menggunakan jasa penerbangan***, maka sebanyak **76.36% dari mereka tidak membeli travel insurance**

OUTLINE

-  **PROBLEM STATEMENT**
-  **DATA OVERVIEW & INSIGHT**
-  **DATA PRE-PROCESSING**
-  **MODEL & EVALUATION**
-  **BUSINESS RECOMMENDATION & IMPACT**

Data Pre-Processing.

Uno Inc.

Data Split
Data Train : Data Test
80 : 20



Data Cleansing

Drop 738 Data
Duplicate

From 1987 Row
To 1249 Row

Feature Encoding

Feature	0	1
Emmployment Type	Government Sector	Private Sector
Graduate or Not	No	Yes
Frequent Flyer	No	Yes
Ever Travelled Abroad	No	Yes

Feature Extraction

Feature	0	1
Age_Bracket	≤ 30	> 30
Income_Bracket	$< 1.300.001$	$> 1.300.000$
Traveller	Else	Frequent Flyer dan Ever Travelled Abroad

OUTLINE

-  **PROBLEM STATEMENT**
-  **DATA OVERVIEW & INSIGHT**
-  **DATA PRE-PROCESSING**
-  **MODEL & EVALUATION**
-  **BUSINESS RECOMMENDATION & IMPACT**

Model Evaluation.

		PREDICT	
		Buy Travel Insurance	Not Buy Travel Insurance
ACTUAL	Buy Travel Insurance	TRUE POSITIVE (TP)	FALSE NEGATIVE (FN)
	Not Buy Travel Insurance	FALSE POSITIVE (FP)	TRUE NEGATIVE (TN)

Distribusi data target

36%	64%
-----	-----

Customer membeli travel insurance Customer tidak membeli travel insurance

$$Precision = \frac{TP}{TP + FP}$$

Memperhitungkan nilai **Precision** untuk mengurangi angka **False Positive** (prediksi setuju, tetapi kenyataan menolak) yaitu memperkecil kemungkinan munculnya prediksi customer membeli insurance travel sedangkan data actual menunjukkan customer tidak membeli insurance travel. **Hal ini dilakukan untuk memperkecil cost marketing.**

Result Modeling.

DATA TRAIN RESULT

MODEL	PRECISION
LOGISTIC REGRESSION	0.84
KNN	0.51
DECISION TREE	0.85
RANDOM FOREST	0.85
ADABOOST	0.84

Berdasarkan hasil modeling, decision tree dan random forest menunjukkan hasil **precision** terbaik sebesar 85%

Di antara algoritma decision tree dan random forest tersebut, dipilih algoritma **Decision Tree** sebagai algoritma terbaik dengan pertimbangan bahwa **algoritma decision tree lebih sederhana, interpretable, dan mudah dipahami dibandingkan dengan algoritma random forest yang secara model lebih rumit.**



Result Modeling.

DATA TEST RESULT

MODEL	PRECISION
DECISION TREE	0.89

Skor hasil pengujian model menggunakan data test menghasilkan **Precision Score 89%**



Rakamin Final Project

CONFUSION MATRIX DECISION TREE

AKTUAL	PREDICT BUY TRAVEL INSURANCE	PREDICT NOT BUY TRAVEL INSURANCE
	ACTUAL BUY TRAVEL INSURANCE	ACTUAL NOT BUY TRAVEL INSURANCE
ACTUAL BUY TRAVEL INSURANCE	224	7
ACTUAL NOT BUY TRAVEL INSURANCE	89	55

PREDIKSI

Result Modeling.

Model	Cross Validation (Train)	Cross Validation (Test)
Decission Tree	0.86	0.85

Pengecekan Overfitting dan Underfitting

Dari hasil k-fold cross validation bisa diketahui bahwa untuk model Decision Tree sudah **best fit** di mana crossval train lebih besar dari crossval test dengan selisih yang tidak terlalu jauh.



OUTLINE

-  **PROBLEM STATEMENT**
-  **DATA OVERVIEW & INSIGHT**
-  **DATA PRE-PROCESSING**
-  **MODEL & EVALUATION**
-  **BUSINESS RECOMMENDATION & IMPACT**

Business Recomendation

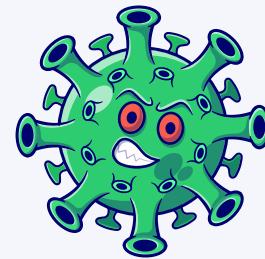
Uno Inc.



Tim marketing hendaknya dapat menyesuaikan harga travel insurance yang akan diluncurkan agar secara psikologis harga yang terlihat bisa lebih menarik perhatian mereka yang berasal dari segmen "High Life" dan membuat mereka merasa bahwa hal tersebut adalah sebuah "good deal"



Membuat campaign atau konten yang bisa mengedukasi, meningkatkan awareness dan lebih menarik perhatian bagi target agar tertarik untuk membeli travel insurance, disesuaikan berdasarkan data-data potensial yang telah dimiliki dan dikelola, seperti umur, pendapatan, pekerjaan, dan apakah dia seorang traveller atau bukan.



Diasumsikan bahwa campaign diluncurkan pada 2020 saat awal pandemi yang berarti ketakutan dari masyarakat terhadap covid19 masih cukup tinggi, sehingga launching dari travel insurance terbaru yang mengcover perlindungan terhadap covid19 bisa dilakukan secepatnya.

Business Recomendation

Uno Inc.



Campaign yang diluncurkan untuk product ini juga bisa dengan menyisipkan "rasa aman" untuk berpergian bahkan ketika pandemi sedang berlangsung karena kelebihan product tersebut yang juga termasuk perlindungan biaya perawatan dan pengobatan apabila customer terkena covid19 sehingga orang-orang yang berpergian akan merasa lebih aman.



Offering dapat dilakukan tim marketing hanya ke pelanggan potensial saja (terprediksi membeli oleh model), yang artinya termasuk konsumen yang tidak membeli travel insurance tapi memiliki ciri – ciri seperti customer yang sudah pasti membeli travel insurance berdasarkan model yang dibuat, sehingga diharapkan dengan diberikannya offering hanya kepada customer jenis ini maka campaign cost bisa dioptimasi dan conversion rate bisa meningkat.

Business Impact

Uno Inc.

Matriks = Conversion Rate

PREDIKSI SETELAH MODELLING (DATA TEST)

PREDICT BUY TRAVEL INSURANCE	
TP	224
FP	89
313	

Jika diasumsikan sebanyak **25% dari False Positif** kemudian menjadi beli travel insurance setelah **diberikan campaign dan penerapan business recommendation**:

$$25\% \times 89 = 22 \text{ orang}$$

*Asumsi revenue yang diperoleh dari setiap customer yang melakukan pembelian = \$5

*Asumsi biaya yang dikeluarkan untuk melakukan campaign terhadap setiap customer = \$0.1

Conversion Rate = jumlah customer yg beli / jumlah customer yg diberikan campaign

Before

$$\text{Conversion Rate} = 710 / 1986 = 35.7\%$$

After

$$\text{Conversion Rate} = 224 / 313 = 71.5\% \uparrow$$

$$\text{Conversion Rate} = (224 + 22) / 313 = 78.59\% \uparrow \text{ (asumsi ditambah keberhasilan dari false positif)}$$

Conversion Rate Increased by **35.8% - 42.89%**!

Rakamin Final Project

Business Impact

Uno Inc.

Matriks = Campaign Cost (optimasi)

		PREDIKSI	
		PREDICT BUY TRAVEL INSURANCE	PREDICT NOT BUY TRAVEL INSURANCE
AKTUAL	ACTUAL BUY TRAVEL INSURANCE	224	7
	ACTUAL NOT BUY TRAVEL INSURANCE	89	55

Jika diasumsikan sebanyak 25% dari False Positif kemudian menjadi beli travel insurance setelah diberikan campaign dan penerapan business recommendation :

$$25\% \times 89 = 22 \text{ orang}$$

*Asumsi revenue yang diperoleh dari setiap customer yang melakukan pembelian = \$5

*Asumsi biaya yang dikeluarkan untuk melakukan campaign terhadap setiap customer = \$0.1

Before

$$\begin{aligned}\text{Campaign Cost} &= 375 \times \$0.1 &=& \$ 37.5 \\ \text{Revenue} &= 224 \times \$5 &=& \$ 1120 \\ \text{ROAS} &= \$ 1120 / \$ 37.5 &=& 29,86x\end{aligned}$$

After

$$\begin{aligned}\text{Campaign Cost} &= 313 \times \$0.1 &=& \$ 31.3 \\ \text{Revenue} &= 246 \times \$5 &=& \$ 1230 \\ \text{ROAS} &= \$ 1230 / \$ 31.3 &=& 39,29x\end{aligned}$$

(Cost Saving = 16.5%!) ↑
(Revenue increased = 9.8%!) ↑
(ROAS increased 31.5%!) ↑

Campaign Cost menjadi lebih efisien dan optimized + Revenue mengalami peningkatan!

Rakamin Final Project

Simulasi Business Impact

Uno Inc.

Apabila diasumsikan ada 1000 data baru yang masuk dan diprediksi bahwa ada 500 orang customer potensial. Dengan akurasi model sebesar 89% berarti data yang diprediksi dikurangi 11% atau 55 orang sebagai margin error yang berarti ada 445 orang yang diprediksi pasti membeli travel insurance.

Simulasi:

Travel insurance price	= \$5
Campaign cost per person = \$0.1	= \$100/1000 target
445 x \$5 - \$50	= \$2175 (gross profit)

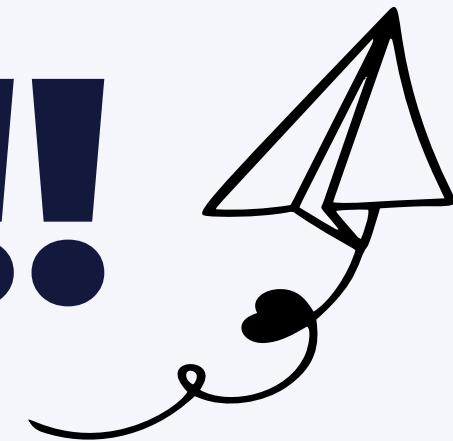
Jika ada tambahan dari false positif yang membeli travel insurance sebanyak:

1. 25% maka = $(445 + 13) \times 5 - 50 = \2240 (gross profit) (**gross profit meningkat 2.98%**)
2. 50% maka = $(445 + 26) \times 5 - 50 = \2305 (gross profit) (**gross profit meningkat 5.97%**)
3. 75% maka = $(445 + 39) \times 5 - 50 = \2370 (gross profit) (**gross profit meningkat 8.96%**)

Dengan menerapkan hal ini atau dengan fokus hanya memberikan campaign kepada mereka yang potensial bisa memaksimalkan campaign cost perusahaan.

Uno Inc.

THANK YOU!!!



Rakamin Final Project