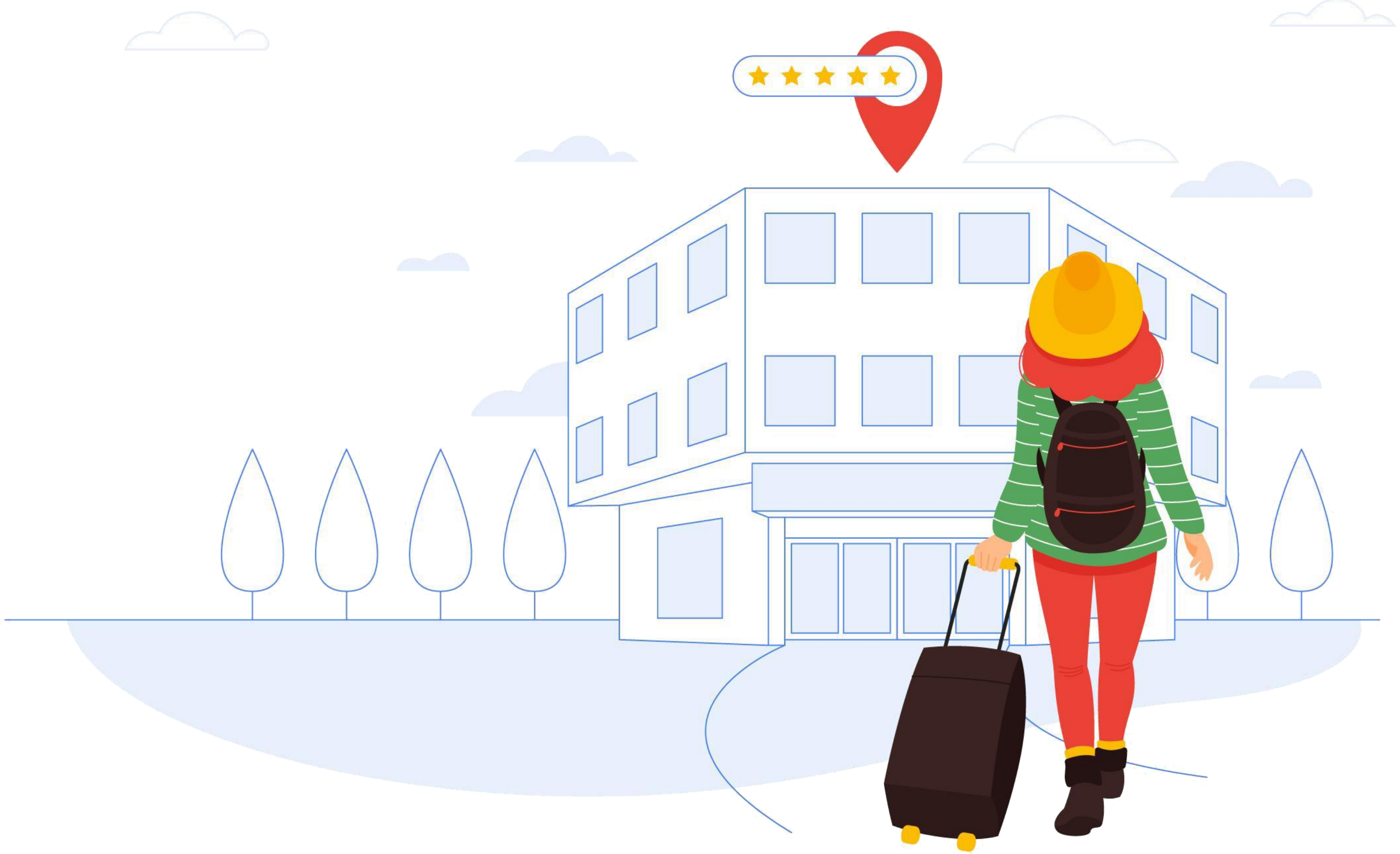
Kitasabi.com

Kitasabi.com

HOTEL BOOKING CANCELLATION



Kitasabi.com

sebagai

TIM ANALYST

Kitasabi.com team

M Choiril Iman

Celestial Randy

Ahmad Reza

Yehezkiel N

Risca Naquitasia

Sonia Epifany

Oky Hariawan

Latar Belakang Masalah

40%

Tingkat pembatalan hotel secara global

Penurunan revenue

Kebiasaan cancel

37%

Tingkat pembatalan hotel **di Portugal** (City Hotel & Resort Hotel)

Asumsi kerugian per tahun

Resort Hotel

25%

€ 484k

*Diambil dari kolom ADR

7 Miliar Rupiah

City Hotel

30%

€ 1,500k

*Diambil dari kolom ADR

23 Miliar Rupiah

Solusi?

CANCELLATION RATE PREDICTION

Target

Menurunkan cancellation rate

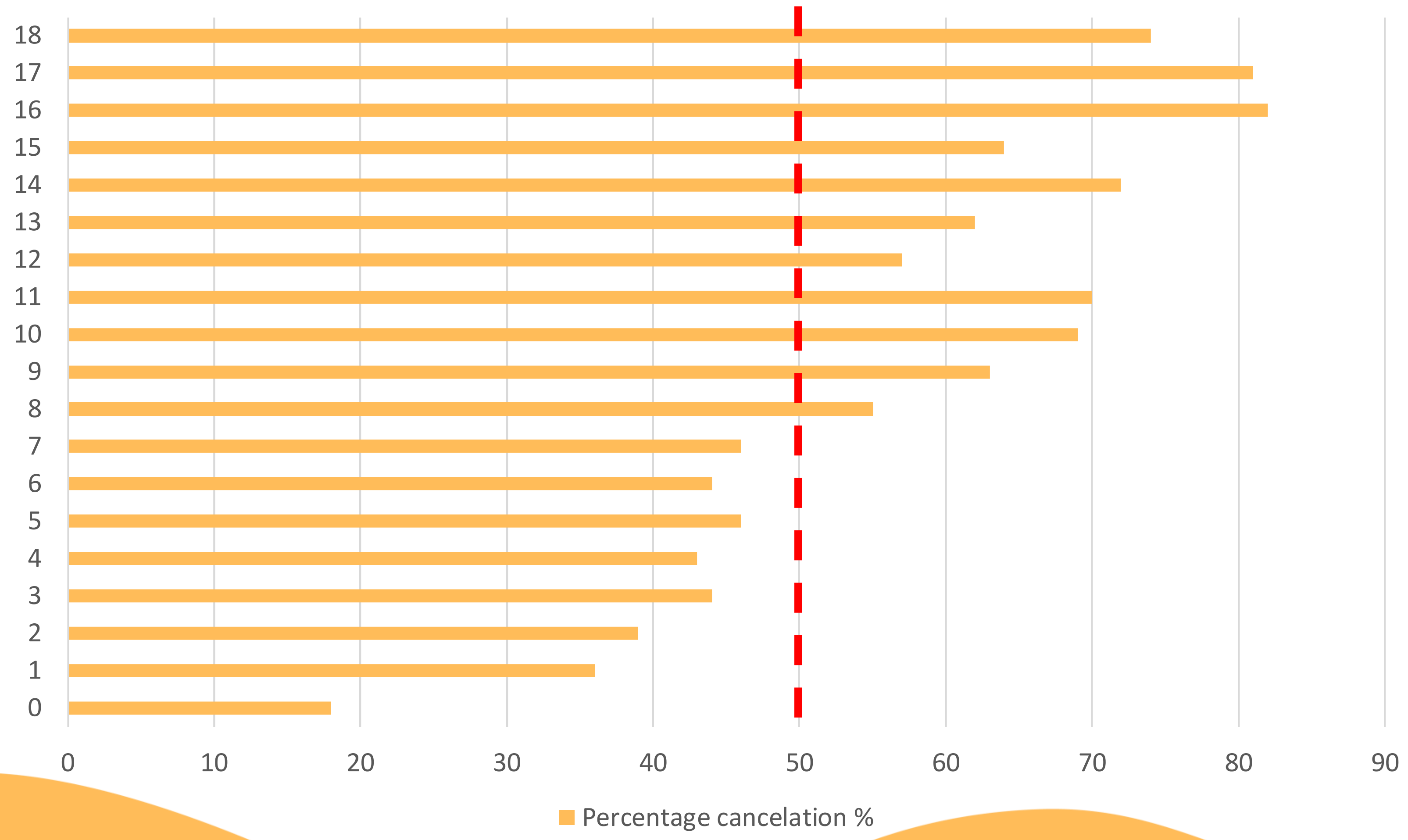
Cancellation rate

Business Metrics

Revenue

Exploratory Data Analyst

Lead time

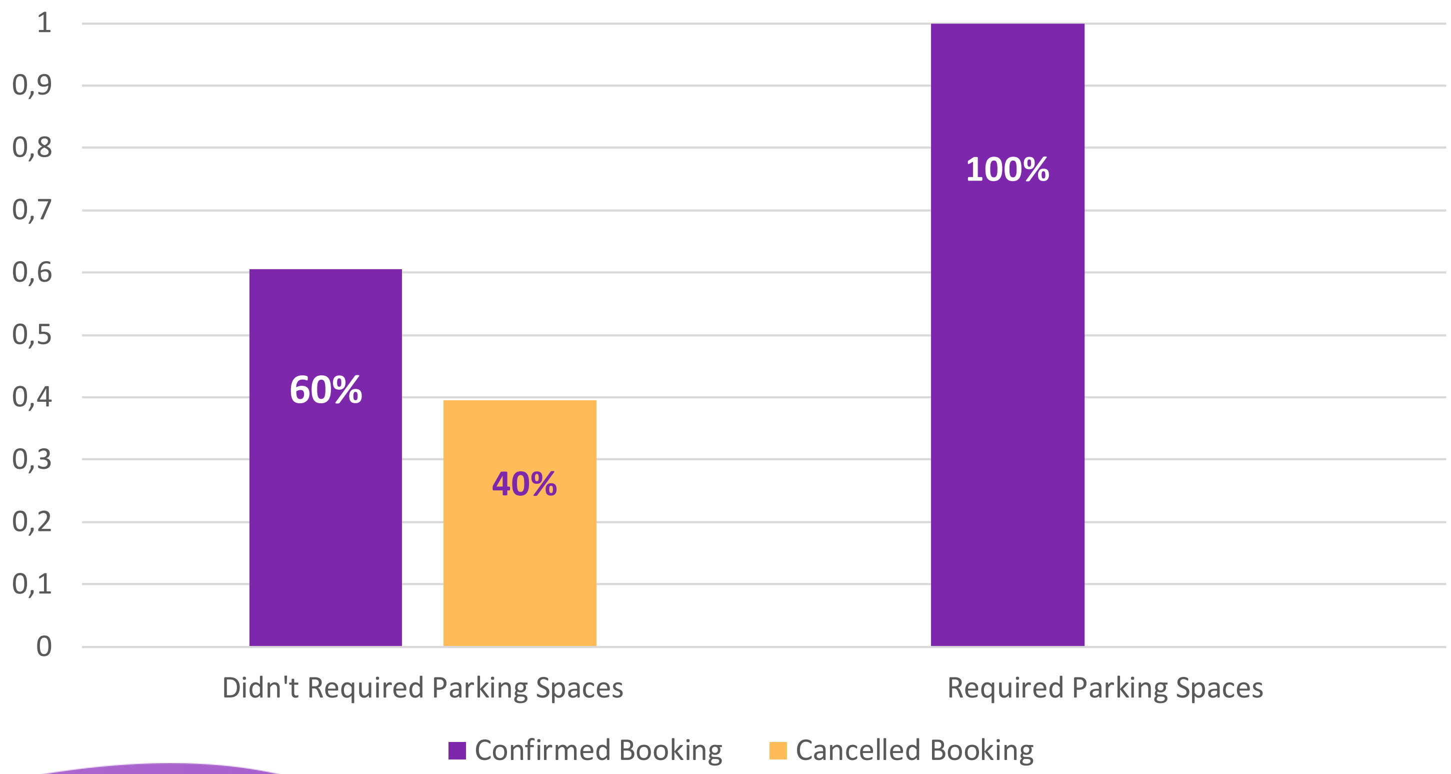


Pemesanan hotel dengan lead time lebih dari **7 bulan** memiliki peluang pembatalan lebih besar **(>50%)**

Semakin lama Lead time semakin tinggi kemungkinan pemesanan dibatalkan

Exploratory Data Analyst

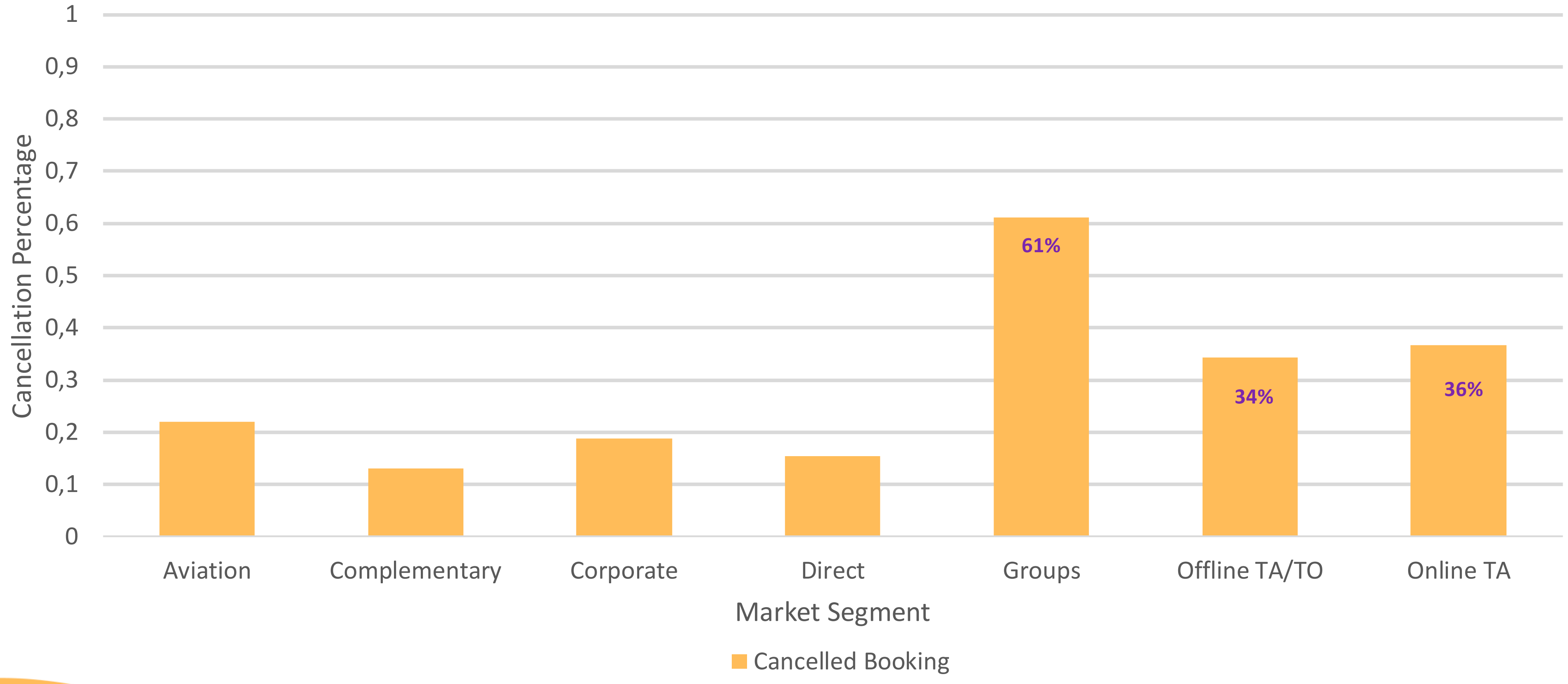
Parking Space & Cancelled Rate



Pemesanan yang memerlukan tempat parkir tidak melakukan membatalkan pesanan

Exploratory Data Analyst

Cancellation rate by market segment

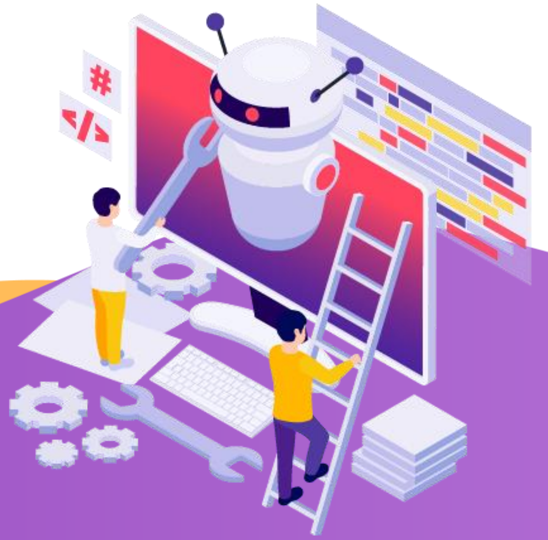


Berikut persentase tingkat pembatalan :

Groups 61% (tertinggi)

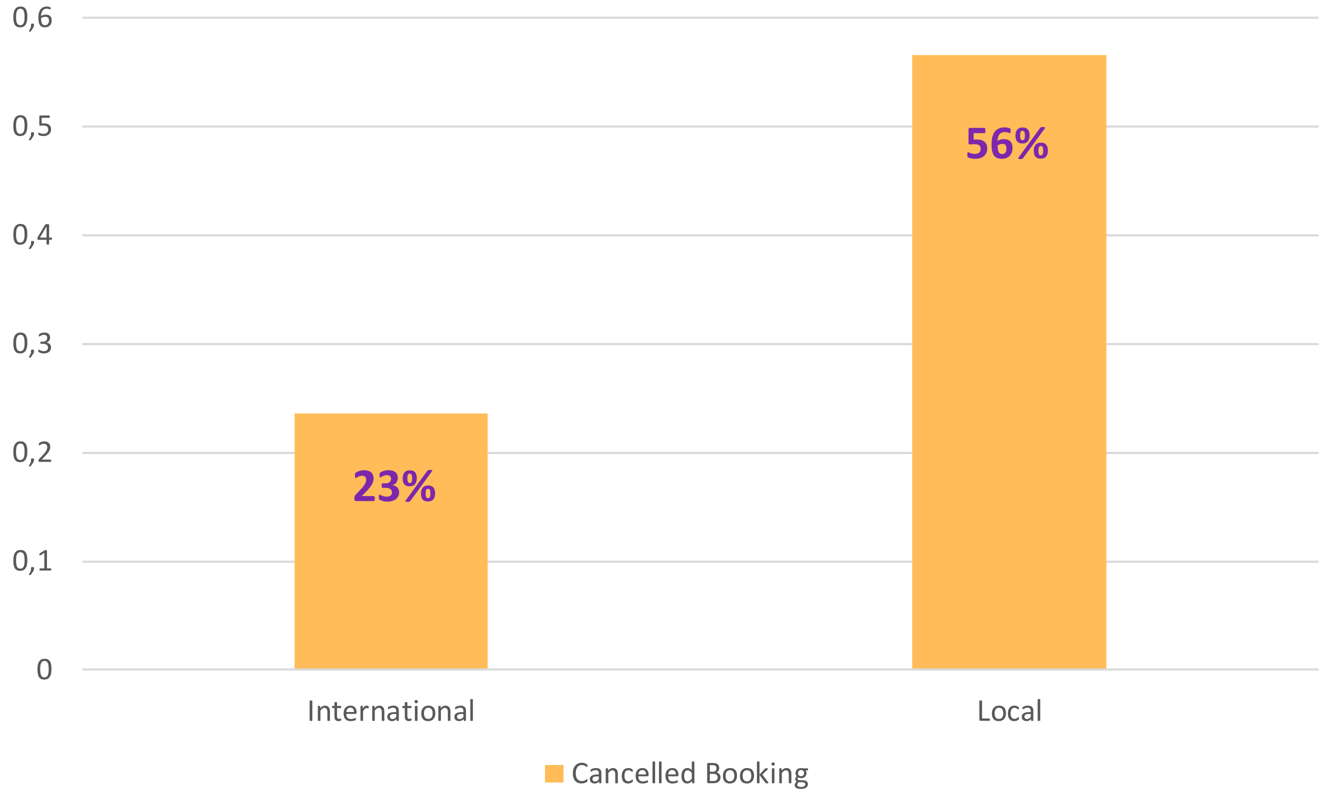
Aviation, Complementary, Corporate, Direct 18 - 22%

Travel Agent 34 - 36 %



Exploratory Data Analyst

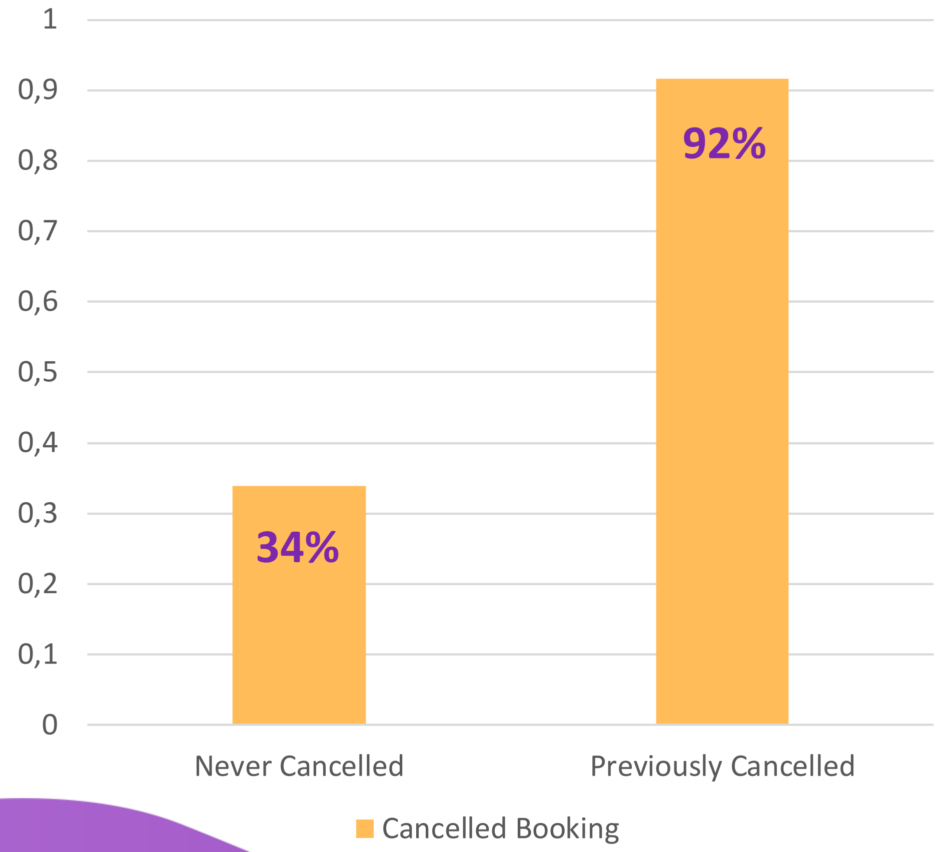
Guest location



Pemesan turis lokal memiliki kemungkinan lebih tinggi untuk pembatalan pemesanan

Exploratory Data Analyst

Cancellation rate for previously cancelled booking



Pemesan yang sebelumnya pernah melakukan pembatalan pemesanan, memiliki kemungkinan lebih tinggi untuk membatalkan pesanan nya

Data Cleansing and Pre Processing

Handle Missing Value

Company Drop
Agent Ada ID = 1, tidak ada = 0
Country Isi dengan nilai modus
Children Hapus baris yang tidak memiliki value

Handle Duplicated Value

Tidak ada duplicated value

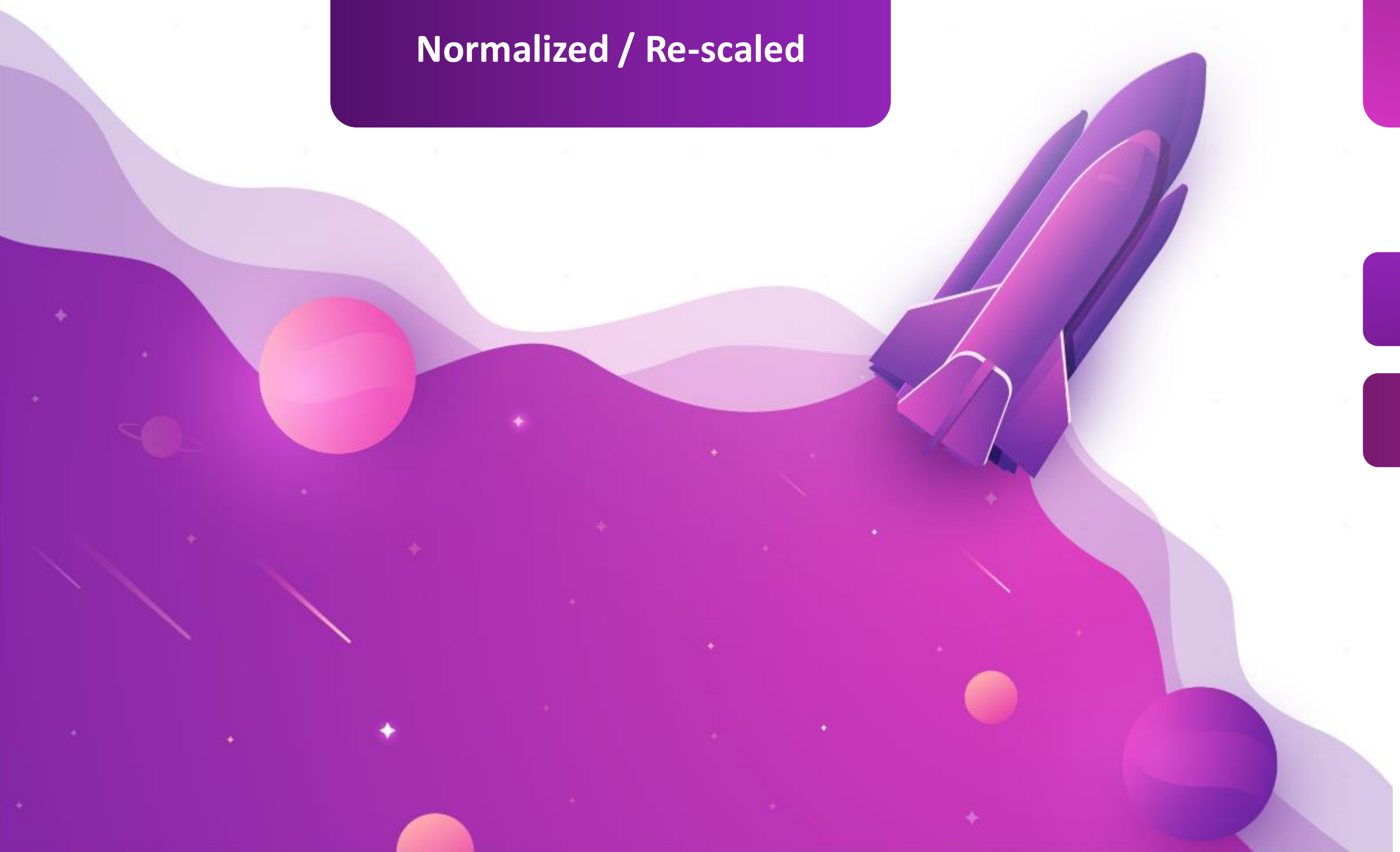
Handle Outlier

Mengubah outlier menjadi batas atas (High Limit)
Ganti nilai outlier dari feature yang memiliki nilai low limit dan high limit sama menjadi nilai terdekat

Feature Engineering

Feature Transformation

Normalized / Re-scaled



Categorical Encoding

- Label encoding
- One hot encoding

Feature Engineering

Feature Extraction

stays in weekend night + stays in week night = **Total stays**

Adults + Children + Babies = **Total guest**

Children + Babies = **Kids**

Guest Location = Value menjadi local and internasional
Meal = Undifined diubah jadi SC
Market Segment = Undifined diganti jadi nilai modus

Feature Engineering

Feature Selection

- Name
- Email
- Phone Number
- Credit card

- Children
- Babies
- Stay in weekend nights
- Stay in week nights

119.390 rows
36 columns

BEFORE

119.386 rows
44 columns

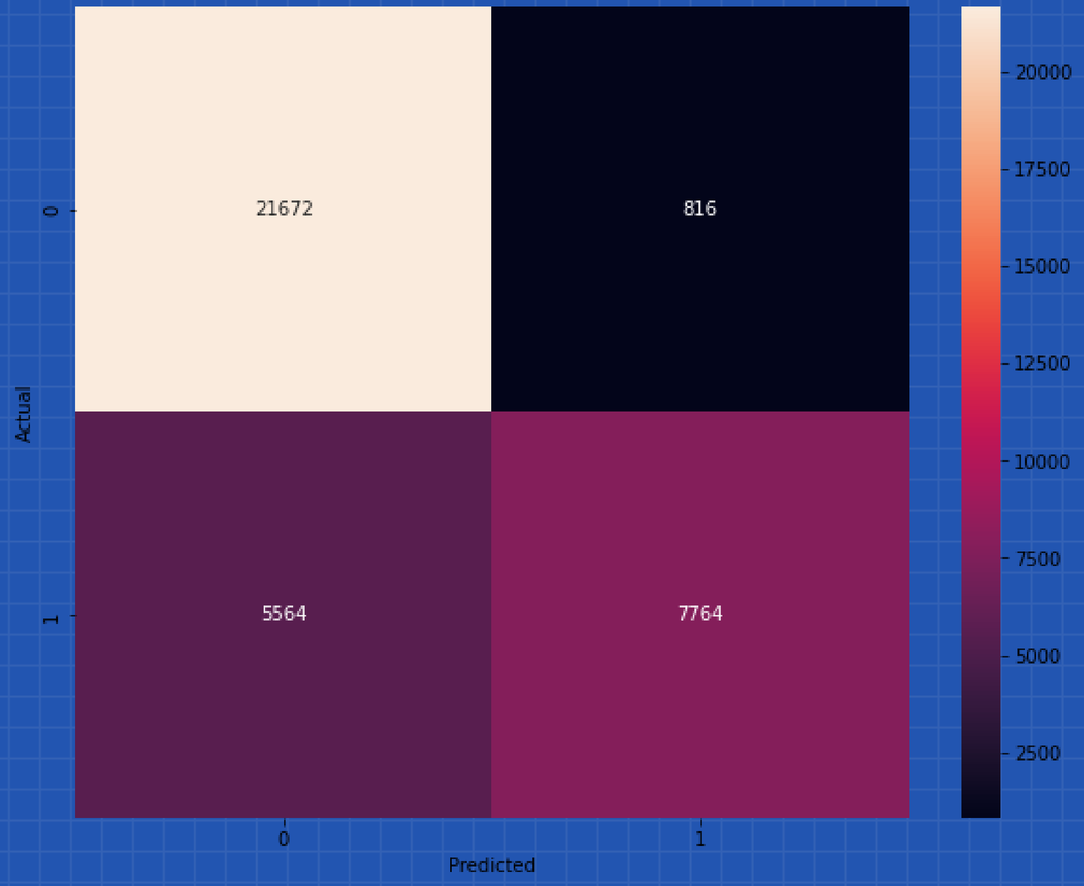
AFTER

Modelling

Modelling Selection

Algoritma	Accuracy	Precision	Recall	F1 – Score	AUC
Logistic Regression	80%	81%	62%	70%	88%
XGBoost	84%	83%	71%	77%	92%
KNN	84%	79%	76%	78%	90%
Adaboost	82%	81%	68%	74%	90%
LightGBM	82%	**91%**	58%	71%	90%
Random Forest	89%	88%	81%	84%	95%
Decision Tree	84%	79%	80%	79%	84%

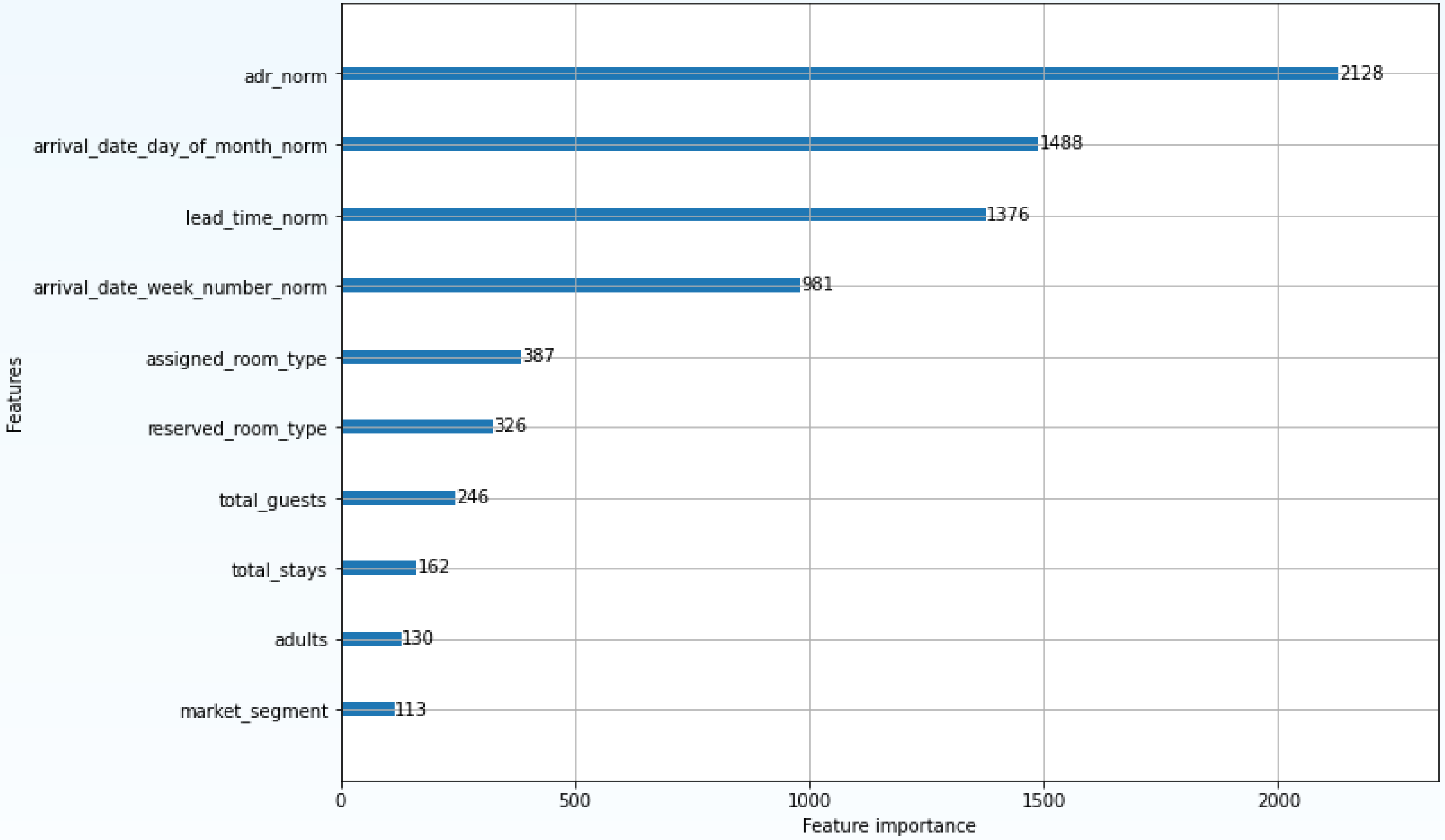
Confusion Matrics



Fokus pada predictive model yang paling bagus dalam menurunkan false positif

False Positif: Hasil prediksi customer melakukan cancellation tetapi kenyataanya customer tetap menginap

Feature importance



Feature yang paling berpengaruh terhadap cancelation adalah adr

Implementasi

Customer Booking and input data → Check via Model → Customer Predict Cancel?

- No → Assign Room for Customer
- Yes → Sent Notification to Customer → Customer Cancel Confirmation (24h)?
  - Yes → Room Available for Other Customer
  - No → Assign Room for Customer

*Free upgrade room jika customer tidak jadi melakukan cancel

Business recommendation

- Non refund policy applied if lead time > 210 days
- Check in Online
- VIP Parking/Vallet
- Phone reminder
- Special offer for local customer
- Hotel partnership
- Customer loyalty point
- Travel agent partnership

Simulation

Cancelation rate

37%

20%

Mengurangi potensi penurunan revenue sebesar

25%

16%

9%

Kitasabi.com

THANKYOU