



**LEVAMTA**

Leadership, Vary,  
Multi Talent

# **Final Project : E-Commerce Shipping Data**



# Data Scientist PT DaHLah

**Agis Fauzi Rachman - Ketua**

**Karina Kristianti**

**Ahmad Mirza**

**Lina Utami Ningsih**

**Hendri Angkasa**

**Raisya Ayu Widiyanti**

**Special Thanks**

**Cimby Wicaksono Rezeki Nur Alam**



# Outline

- Business Understanding
- Data Understanding
- Data Preprocessing
- Modeling & Evaluation
- Business Recommendation
- Business Simulation

# Business Understanding

(Apa yang sedang terjadi ?)



# Latar Belakang

**LEVAMTA**  
Leadership, Vary,  
Multi Talent



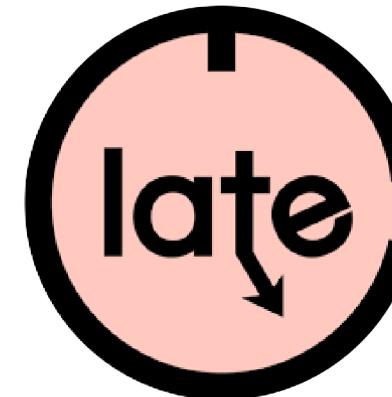
PT DaHLah merupakan sebuah perusahaan *E-commerce*.



Tim marketing telah mengeluarkan **biaya** besar untuk akuisisi pelanggan.



Kepercayaan customer menurun dan berpengaruh pada **turunnya revenue** perusahaan.



59.67% dari total pengiriman barang mengalami keterlambatan.



# Objective

- Membuat model untuk memprediksi suatu pengiriman terlambat atau tidak.

## Goals

- Menurunkan loss revenue akibat pembelian menurun karena keterlambatan pengiriman.

## Business Metrics

- On Time Delivery Ratio
- Customer Retention
- Revenue



# Data Understanding

(Bagaimana pola keterlambatannya ? )





# Exploratory Data Analysis

10.999

Pelanggan

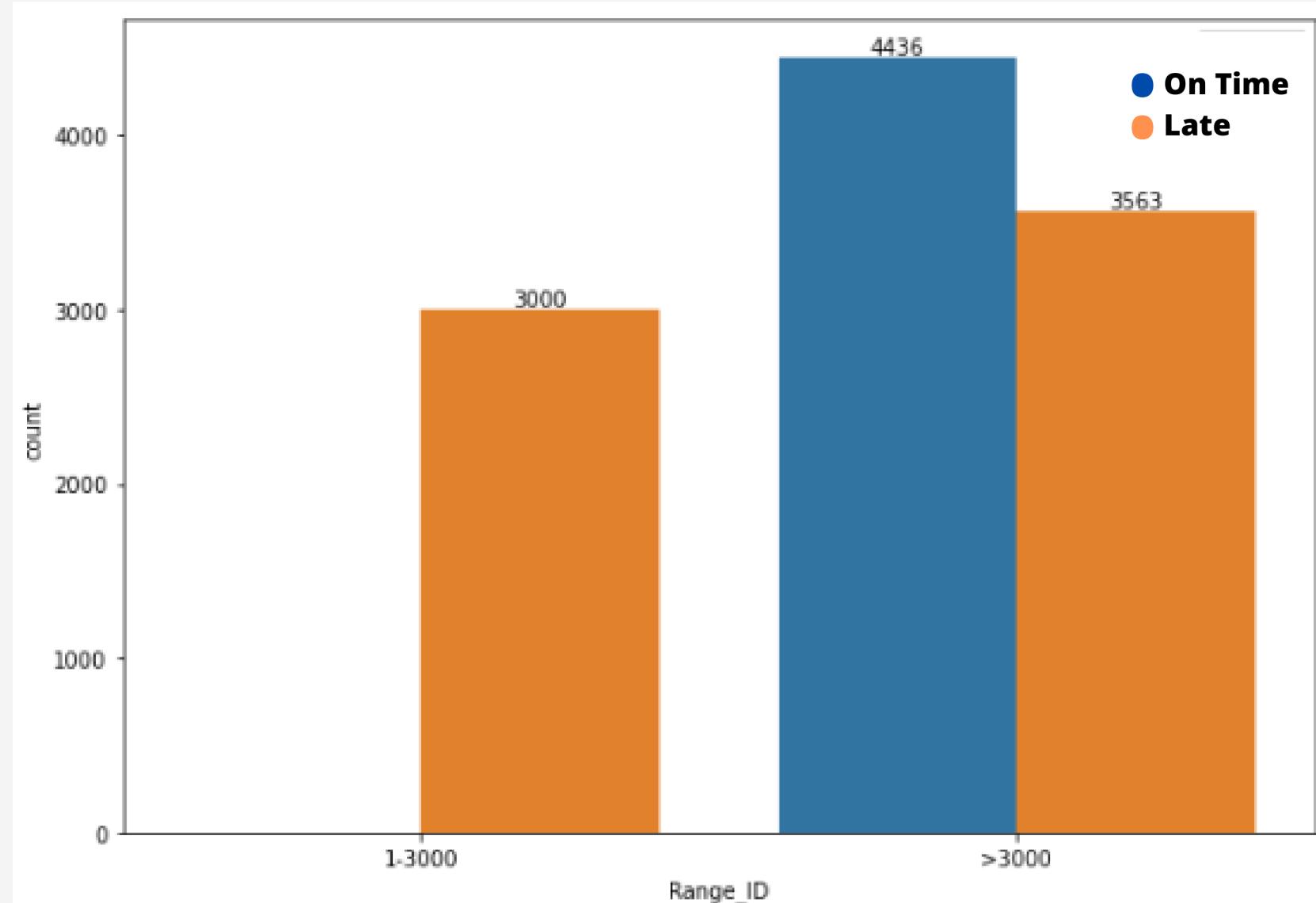
12

Kolom

Tidak ada data  
**KOSONG**

Tidak ada data  
**DUPLIKAT**

# ID Pelanggan



Kenapa 3.000 pelanggan awal seluruhnya  
mengalami keterlambatan?

# Customer Rating

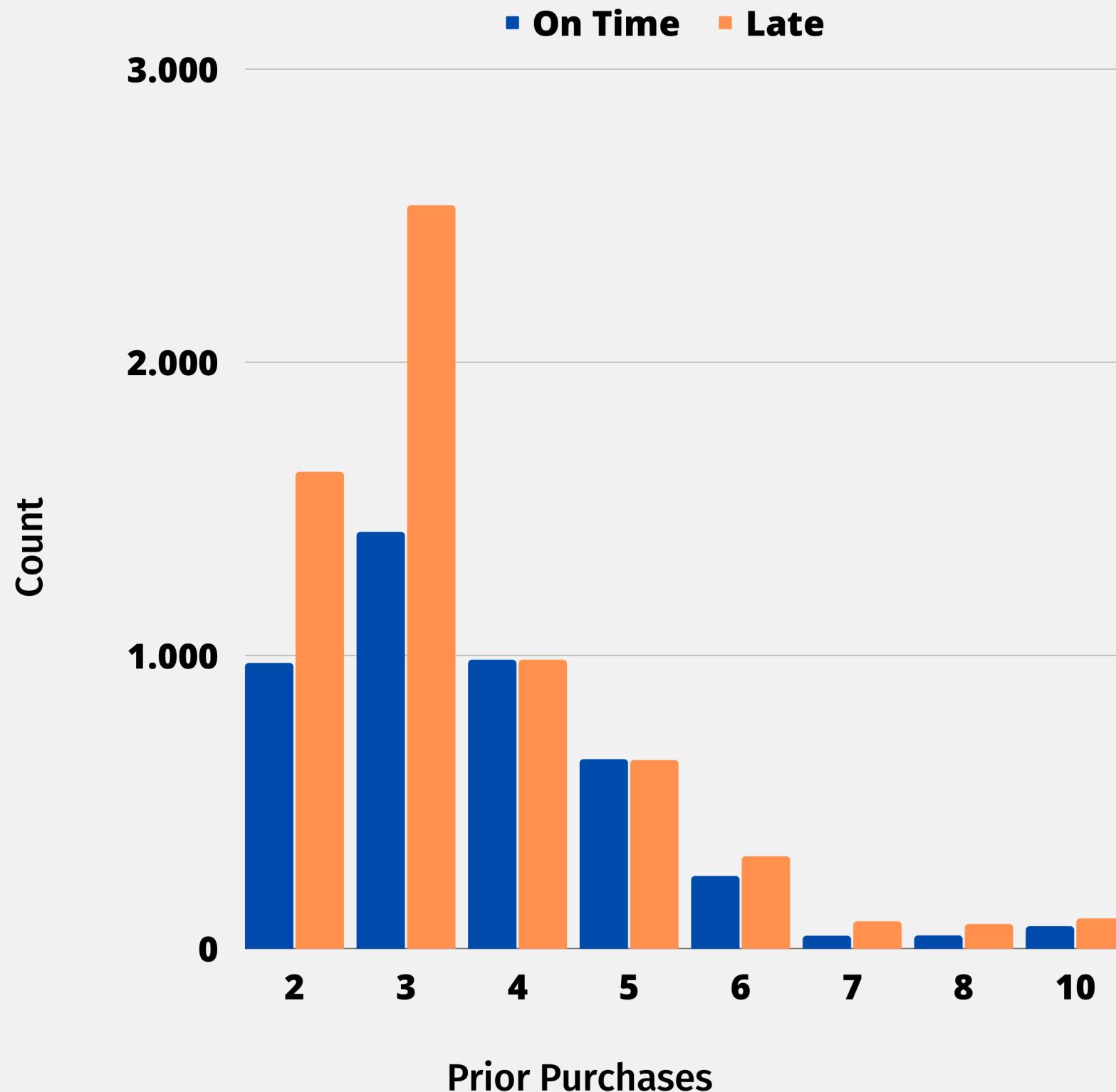


**LEVAMTA**  
Leadership, Vary,  
Multi Talent

Rating	Late/On Time	Count	Total
1	On Time	922	2235
	Late	1313	
2	On Time	892	2165
	Late	1273	
3	On Time	882	2239
	Late	1357	
4	On Time	886	2189
	Late	1303	
5	On Time	854	2171
	Late	1317	

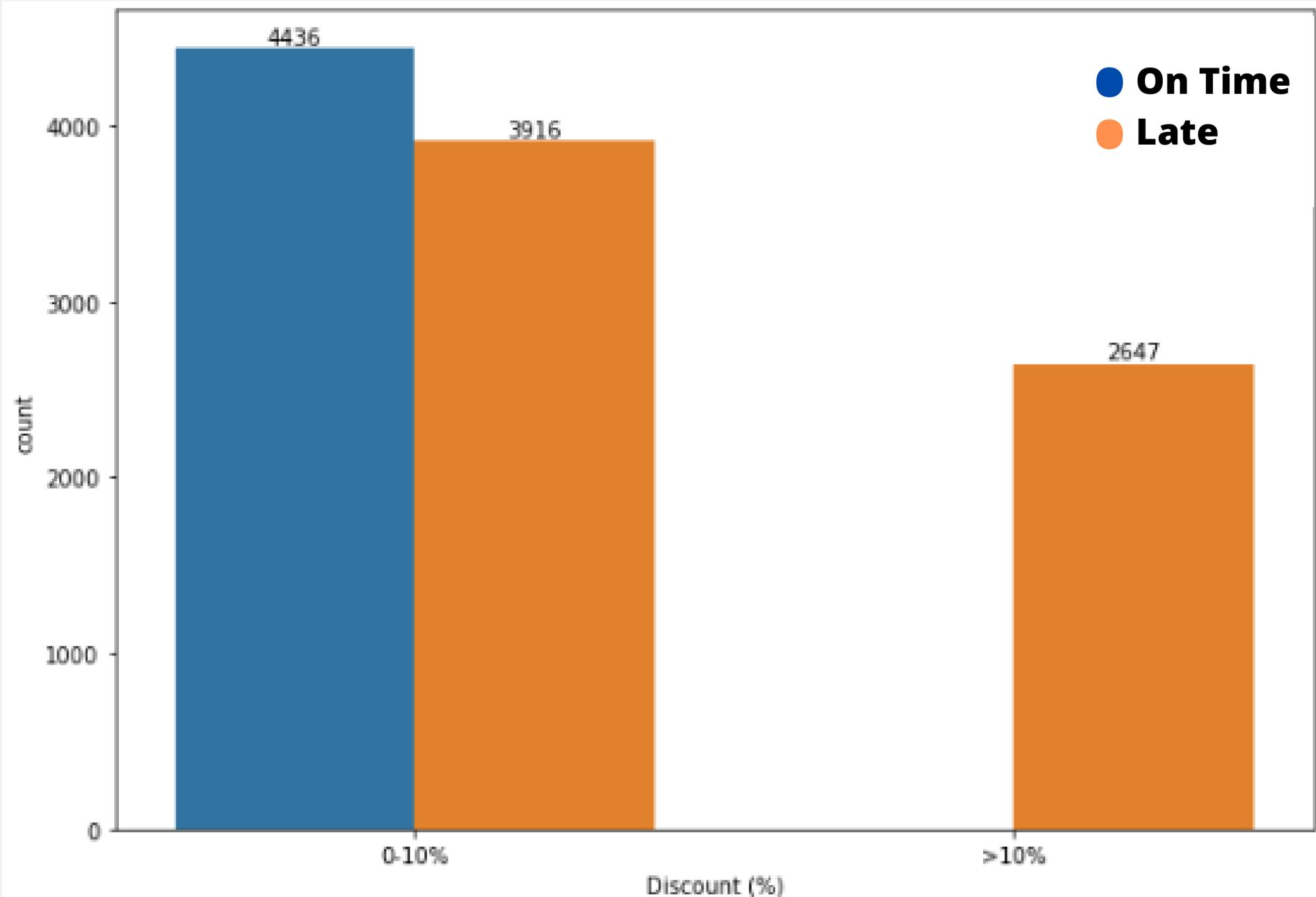


# Prior Purchases



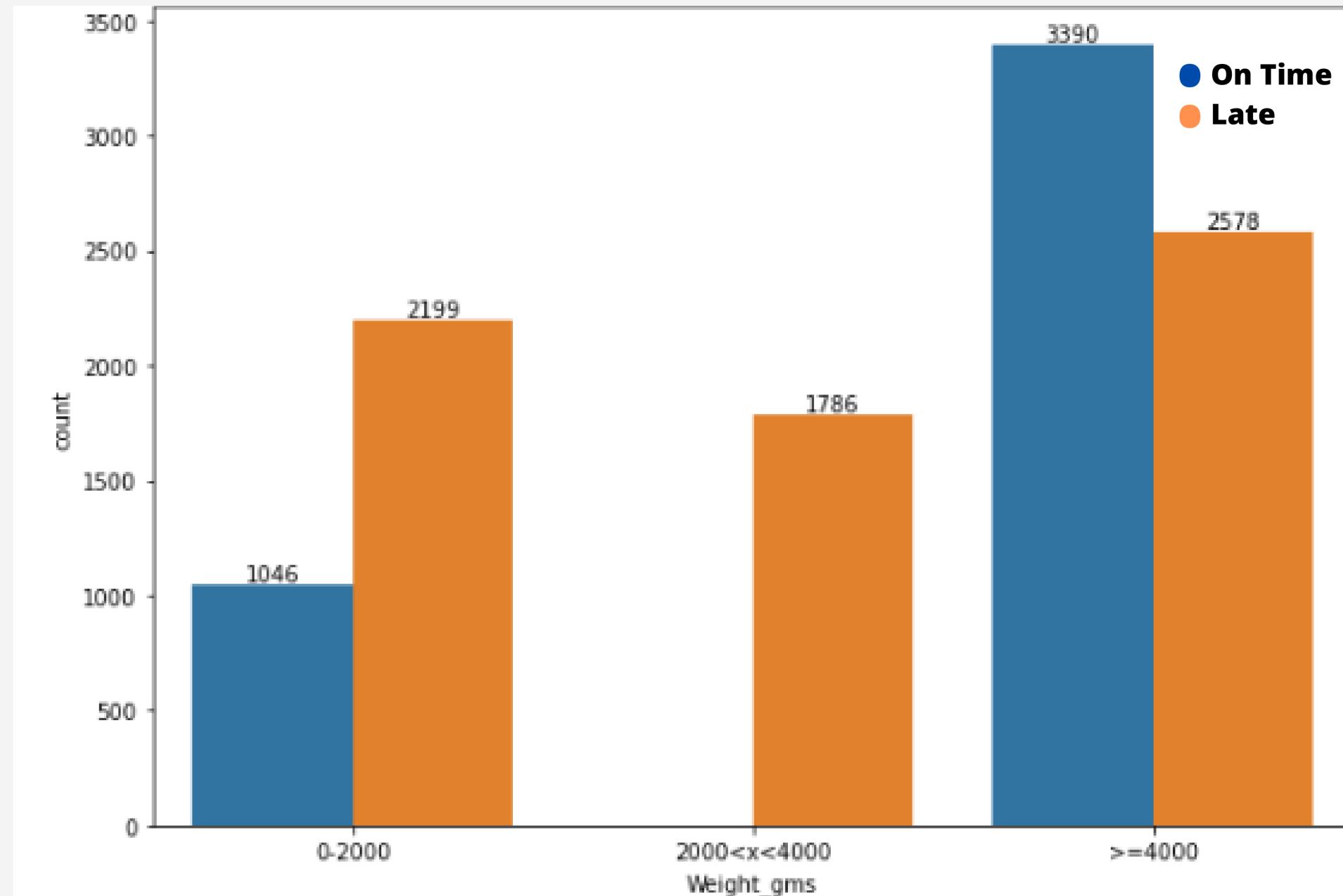
Pengiriman barang yang terlambat hampir paling banyak ada di prior purchase ke 3.

# Discount Offered



**100% keterlambatan**  
terjadi pada pembelian  
produk dengan diskon  
**lebih dari 10%**

# Weight (gms)

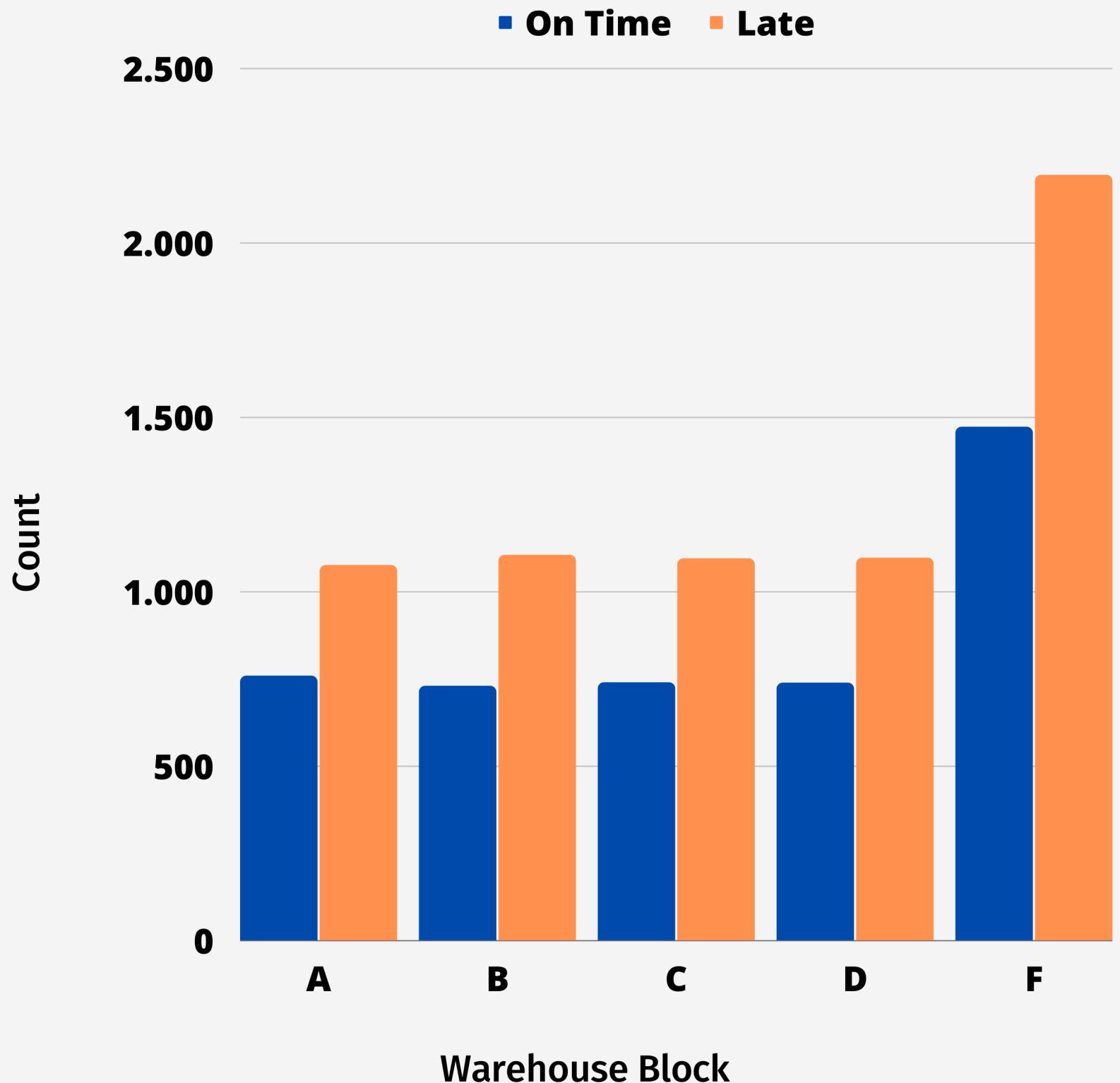


Kenapa barang dengan berat antara **2-4 kg** seluruhnya mengalami **keterlambatan**?

# Warehouse Block



**LEVAMTA**  
Leadership, Vary,  
Multi Talent

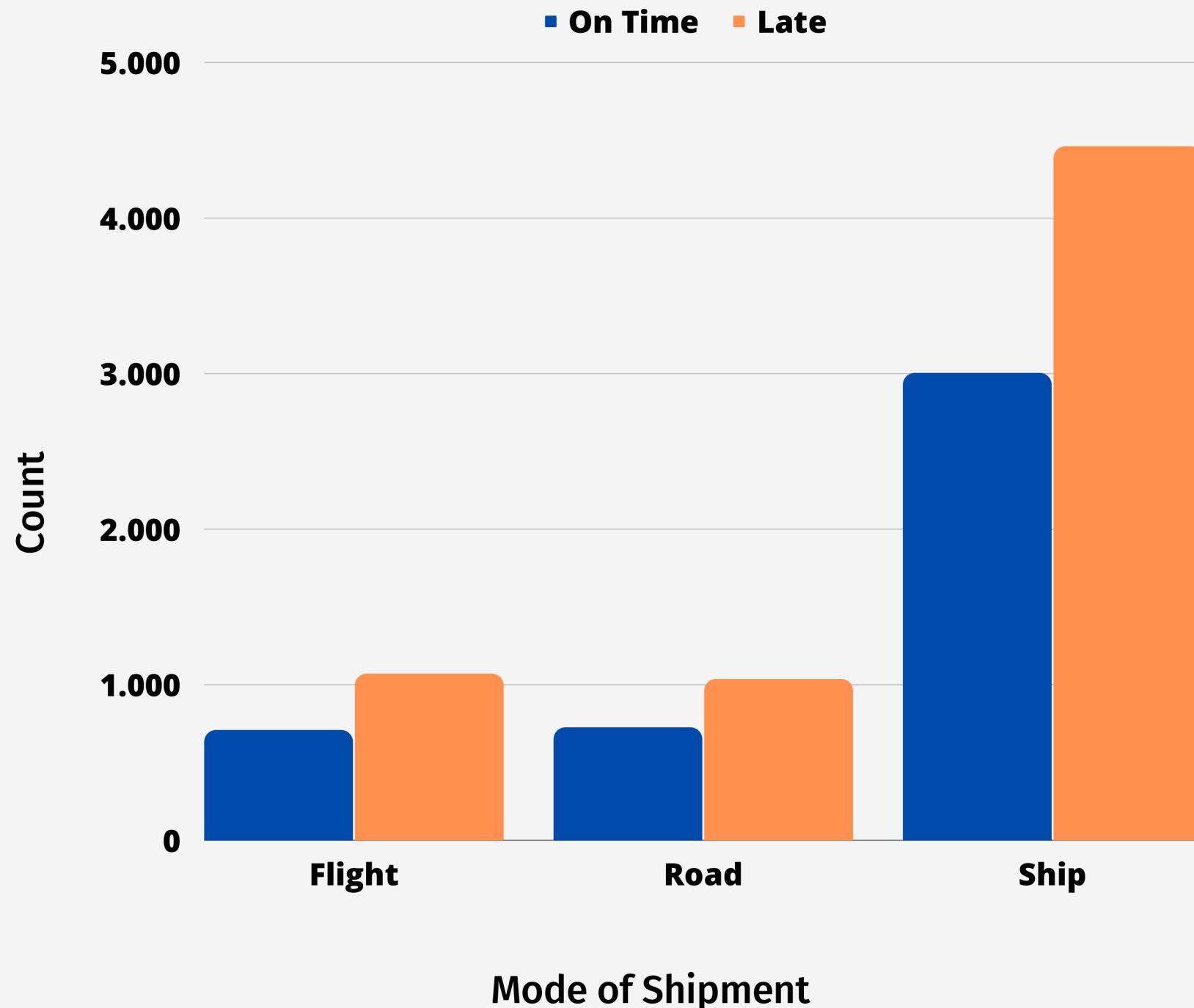


Pengiriman barang yang terlambat hampir selalu terjadi pada seluruh tipe Warehouse Block, namun paling banyak terjadi di Warehouse block F.

# Mode of Shipment



**LEVAMTA**  
Leadership, Vary,  
Multi Talent



Pengiriman barang yang terlambat terjadi pada seluruh transportasi pengiriman.



**Harga barang yang dibeli pelanggan kebanyakan  
245 USD**

---

**Pelanggan melakukan panggilan terbanyak 3-5x**

---

**Pengiriman menggunakan pesawat 60.2% merupakan  
yang terbanyak, namun tidak berpengaruh karena  
perbedaannya hanya 1% dengan transportasi lainnya**

---

**Sebanyak 50.4% pelanggan adalah perempuan  
dengan 30% diantaranya mengalami  
keterlambatan barang**

---

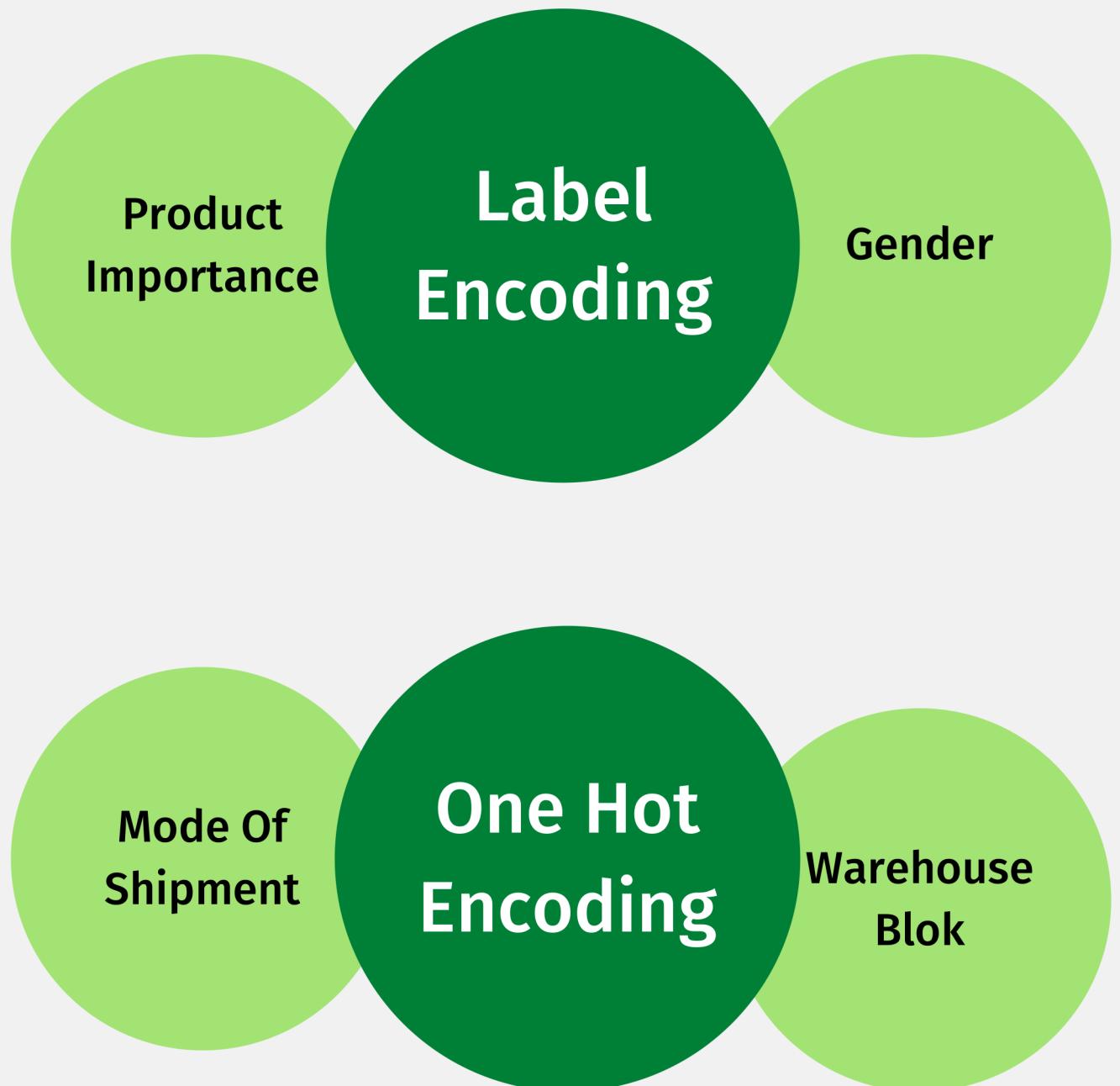


# Data Preprocessing





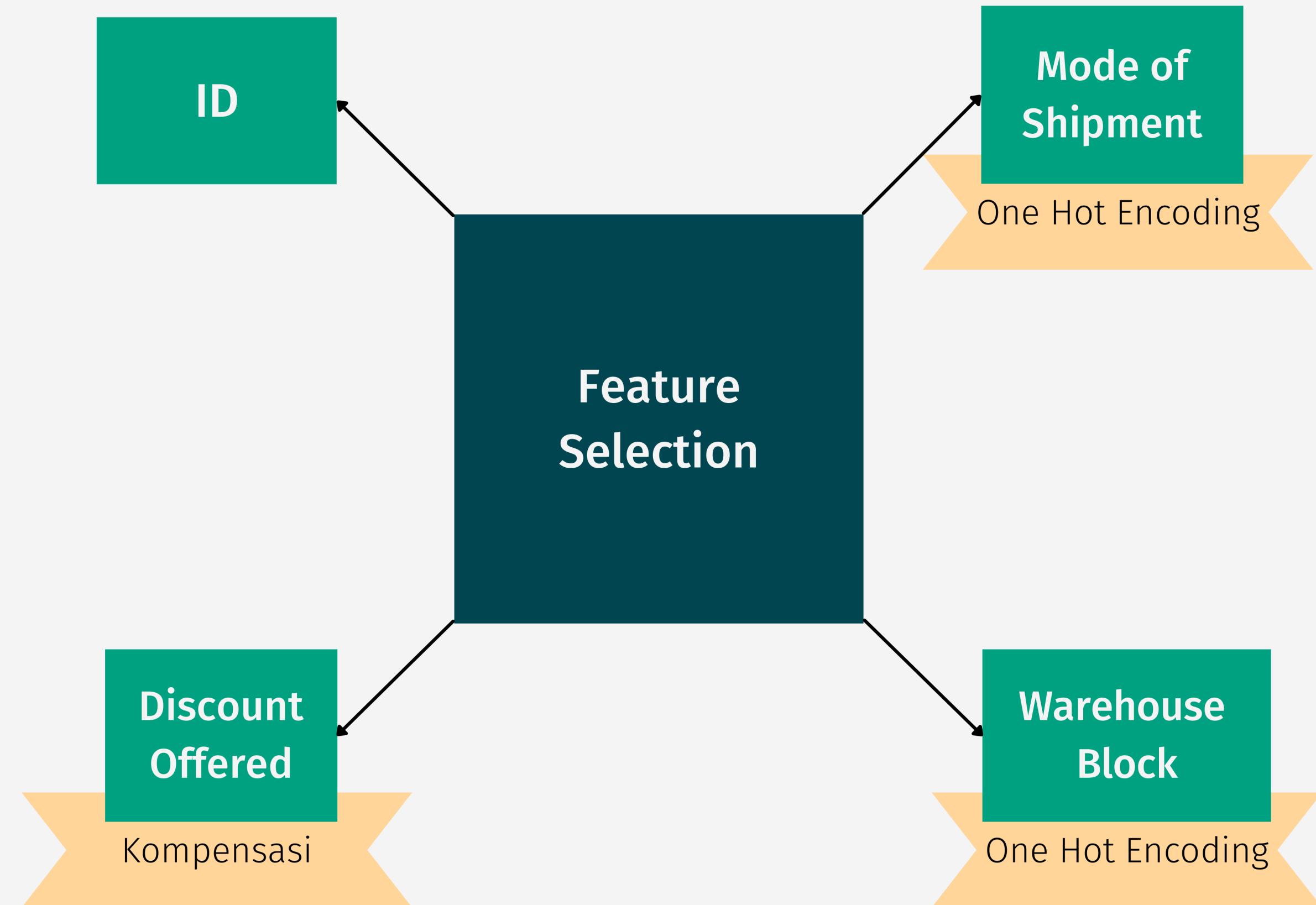
# Normalization & Feature Encoding





**LEVAMTA**

Leadership, Vary,  
Multi Talent





# Modeling & Evaluation





7.699  
Data Train

3.300  
Data Test

Recall

PREDIKSI

On Time



AKTUAL

Terlambat

Hal ini akan menyebabkan **customer** berharap barang sampai **tepat waktu**, padahal barang tersebut mengalami **keterlambatan**.

# Result

	Accuracy	Precision	Recall	F1-Score	roc_auc	roc_auc (Train)
Logistic Regression	0.65	0.67	0.83	0.74	0.64	0.62
k-Nearest Neighbor	0.65	0.72	0.68	0.7	0.73	0.87
Decision Tree	0.65	0.7	0.71	0.71	0.63	1
Random Forest	0.66	0.73	0.69	0.71	0.74	1
Adaboost	0.64	0.7	0.73	0.71	0.74	0.75
XGBoost	0.66	0.76	0.65	0.7	0.75	0.94

Memiliki nilai  
**Recall** terbesar



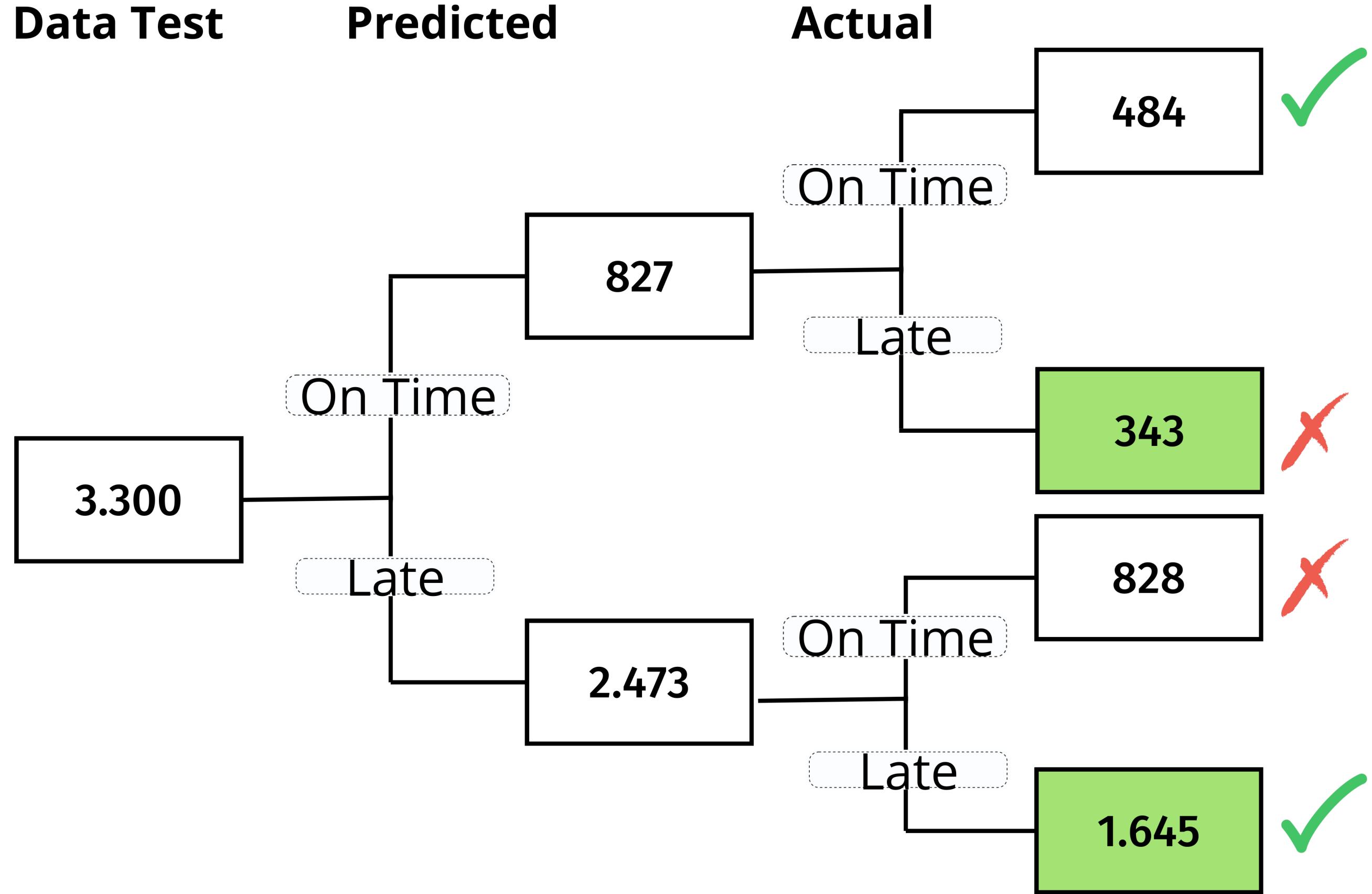
# Cross Validation

	Accuracy		Precision		Recall		F1-Score		roc_auc	
	CV (train)	Cv (test)	CV (train)	Cv (test)	CV (train)	Cv (test)	CV (train)	Cv (test)	CV (train)	Cv (test)
<b>Logistic Regression</b>	<b>0.64</b>	<b>0.58</b>	<b>0.67</b>	<b>0.62</b>	<b>0.8</b>	<b>0.73</b>	<b>0.73</b>	<b>0.66</b>	<b>0.66</b>	<b>0.62</b>
k-Nearest Neighbor	0.78	0.62	0.83	0.67	0.79	0.63	0.81	0.62	0.86	0.72
Decision Tree	0.99	0.62	1	0.68	0.99	0.67	0.99	0.66	0.99	0.61
Random Forest	0.99	0.62	0.99	0.68	0.99	0.64	0.99	0.63	0.99	0.73
Adaboost	0.66	0.61	0.71	0.65	0.73	0.66	0.72	0.64	0.75	0.71
XGBoost	0.86	0.61	0.92	0.67	0.84	0.59	0.88	0.58	0.95	0.73

# Evaluation



**LEVAMTA**  
Leadership, Vary,  
Multi Talent





www.PTDaHLah.com/pemesanan



## PERIKSA PESANAN KAMU!

### Alamat Pengiriman

Luiz | +54001002

jl. suarez X  
Argentina

EDIT

### Leamor\_co

Leamor\_co | +96621212

jl. abbas II  
Saudi Arabia



SUMBANG 261222 NEW

x1

\$174

Berat benda 1 kg (ringan)

Total Pesanan

\$174

### Layanan Pengiriman

Metode Pengiriman



Pesawat (Rec)

estimasi 3 - 4 hari

Darat

estimasi 5 - 7 hari

Kapal

estimasi 5 - 12 hari (! sedang penuh, berpotensi terlambat)



**PT.DaHLah**



Apabila memilih metode pengiriman selain yang direkomendasikan, akan menyebabkan keterlambatan pengiriman 2-3 hari

### Metode Pembayaran

EDIT

Bank Transfer

**KONFIRMASI**



# Business Recommendation



# Real Time

Adanya pelacakan secara real time atas posisi barang dan informasi keberadaan paket kiriman.

# Estimasi

Pemberian waktu estimasi tambahan untuk pengiriman yang memiliki potensi terlambat.

# Supply Chain Management

Melakukan audit untuk permasalahan ID pelanggan dan berat yang terjadi pada pelanggan awal dan berat antara 2-4kg, serta untuk mengatur sistem pergudangan agar efektif.

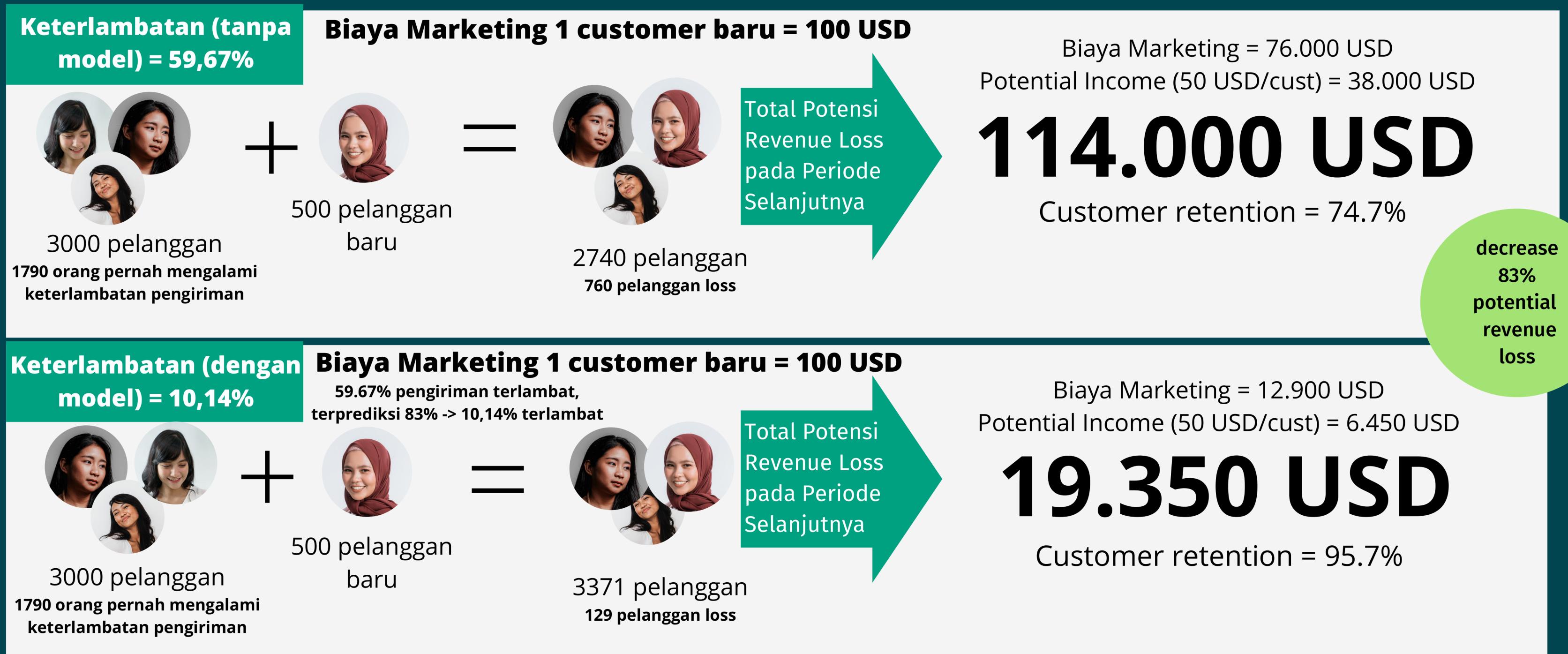


# Business Simulation



# Business Simulation

- Dalam memperoleh customer baru, perusahaan berpotensi menghabiskan **revenue lebih banyak** dibandingkan mempertahankan customer lama (sumber: Tidio.com)
- **17% customer** tidak akan berbelanja kembali setelah menerima pengiriman **terlambat 1x** (sumber: hollingsworthllc.com)
- **55% customer** tidak akan berbelanja kembali setelah menerima pengiriman **terlambat 2-3x** (sumber: hollingsworthllc.com)



# Business Simulation



Poin Permasalahan	Tanpa Model	Dengan Model
Total Pelanggan pada Awal Periode 1	3500	3500
Total Pelanggan pada Awal Periode Selanjutnya	2740	3371
Customer Retention	74,7%	95,7%
Revenue (Loss)	114.000 USD	19.350 USD

decrease 83%  
potential  
revenue loss  
!!!

# TERIMA KASIH

