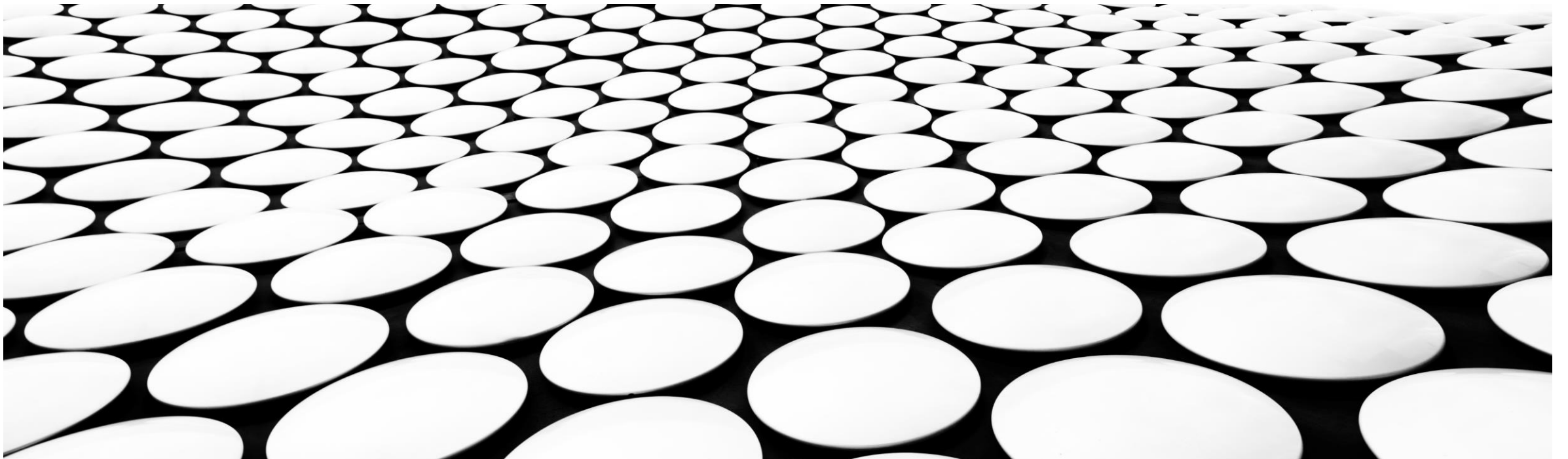

PREDICTING THE BEST OFFER FOR COVID-19 SWAB TEST

OSVALDO FIGO



- Berbagai cara sudah dilakukan untuk memberantas wabah COVID-19.
- Kurangnya edukasi dan harga swab test yang sering dirasa masyarakat **terlalu tinggi**.
- Banyak masyarakat yang menganggap sepele COVID-19 dan mengabaikan swab test yang sudah disediakan pemerintah.
- Apa ada cara yang dapat **menarik** perhatian dan tidak memberatkan masyarakat dari segi **biaya**?
 1. Harga terbaik dan **Diskon-diskon** menarik untuk masyarakat.
 2. Menggunakan algoritma *Machine Learning* untuk dapat **mempersingkat waktu** dalam **menentukan harga** dan **diskon**

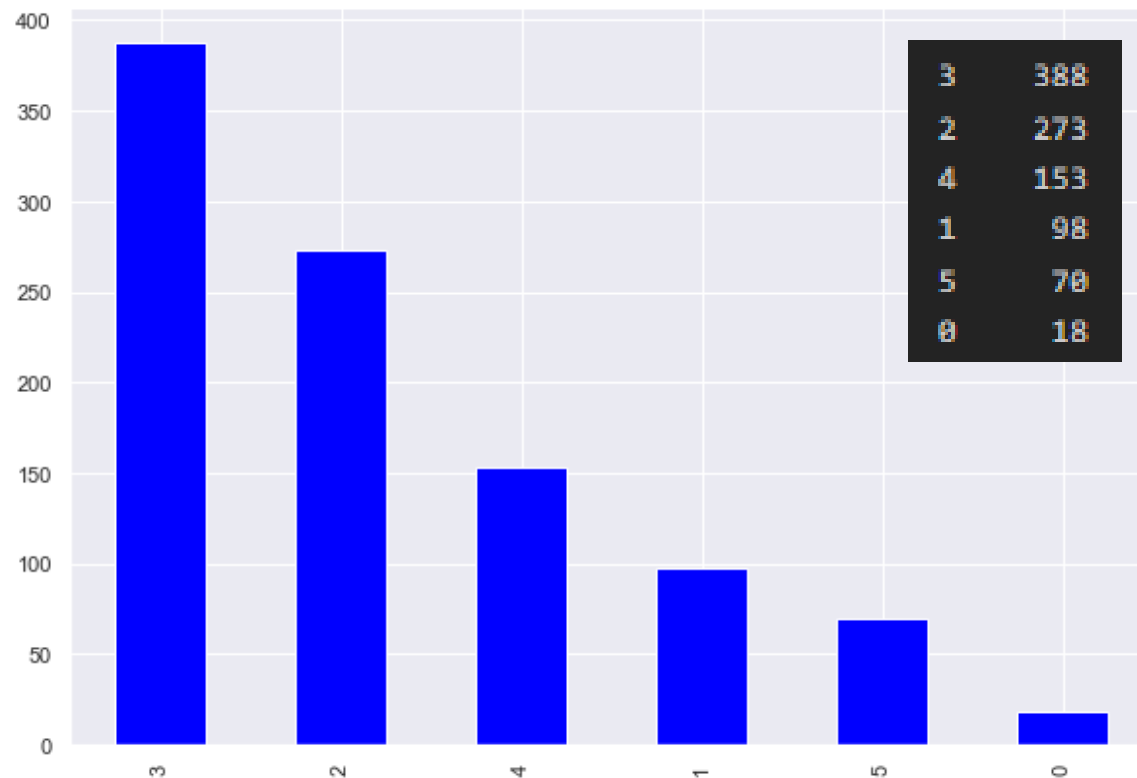
LATAR BELAKANG & RUMUSAN MASALAH

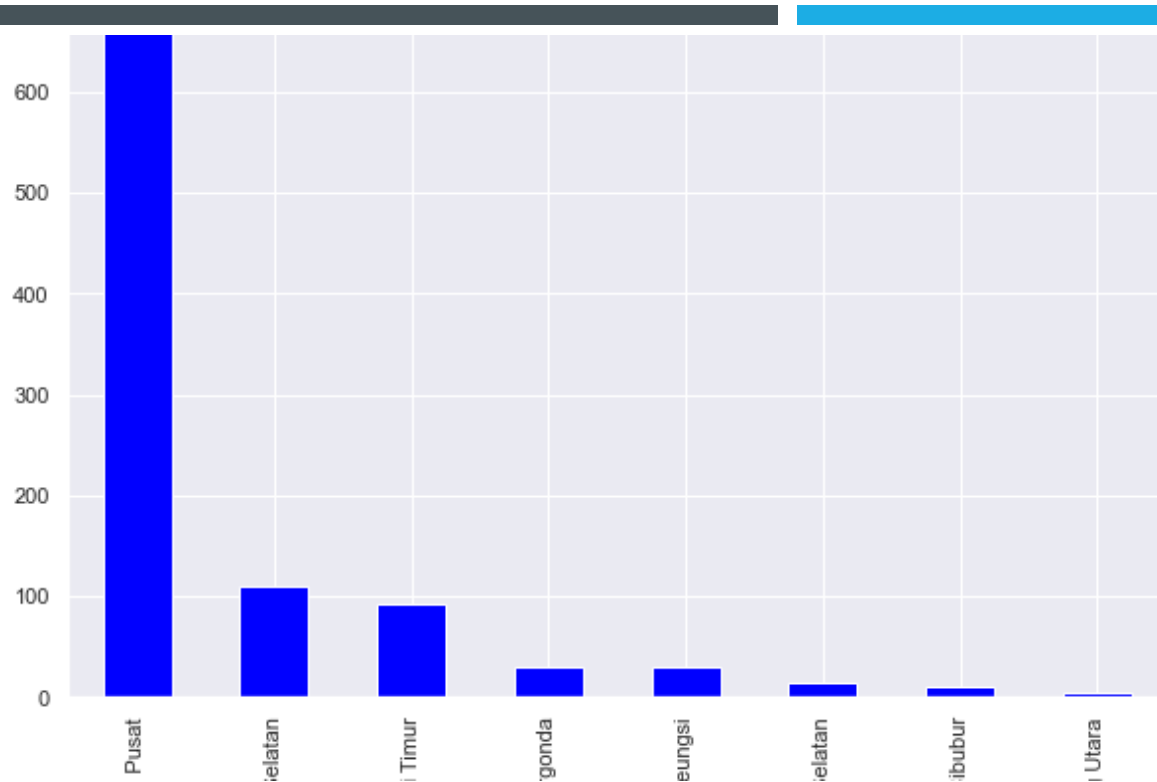
TUJUAN PENELITIAN

- Membantu pihak administrasi klinik untuk menentukan **Harga** dan **Diskon** terbaik yang dapat diberikan kepada calon pasien yang ingin melakukan Swab Test.
- Mencari tahu ***Insights*** apa saja yang bisa diperoleh dari *dataset* yang didapat, seperti:
 1. Mana klinik yang paling banyak dikunjungi?
 2. Mana klinik dengan kasus **positif** paling banyak?
 3. Berapa kelompok umur yang paling banyak melakukan swab test?
 4. DSB. (nanti akan dijelaskan pada proses EDA)

EXPLORATORY DATA ANALYSIS

Kelompok umur yang paling banyak melakukan tes





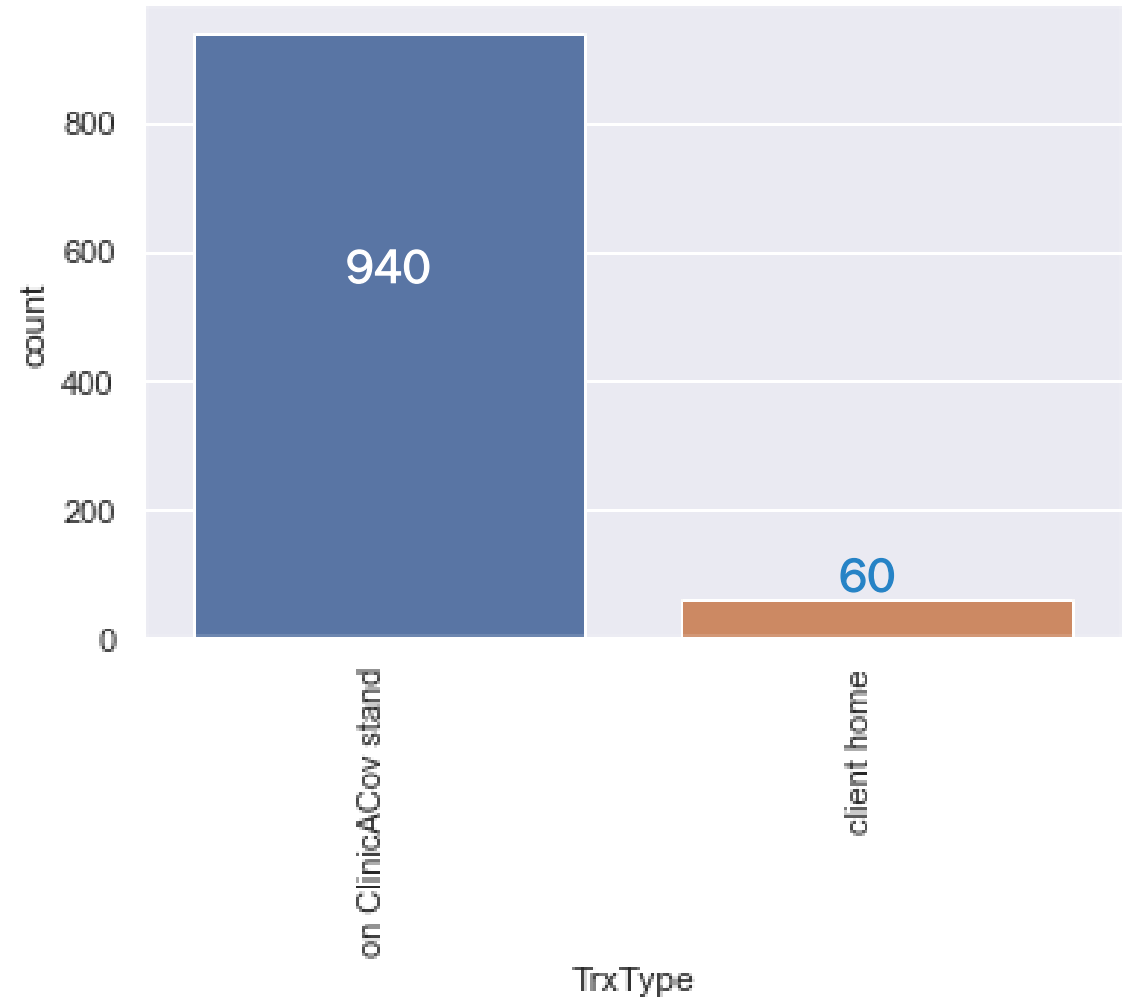
ClinicACov	Cikarang Pusat	707
ClinicACov	Cikarang Selatan	111
ClinicACov	Bekasi Timur	93
ClinicACov	Margonda	31
ClinicACov	Cileungsi	30
ClinicACov	Bekasi Selatan	14
ClinicACov	Cibubur	10
ClinicACov	Cikarang Utara	4

EXPLORATORY DATA ANALYSIS

KLINIK YANG PALING BANYAK DIKUNJUNGI MASYARAKAT

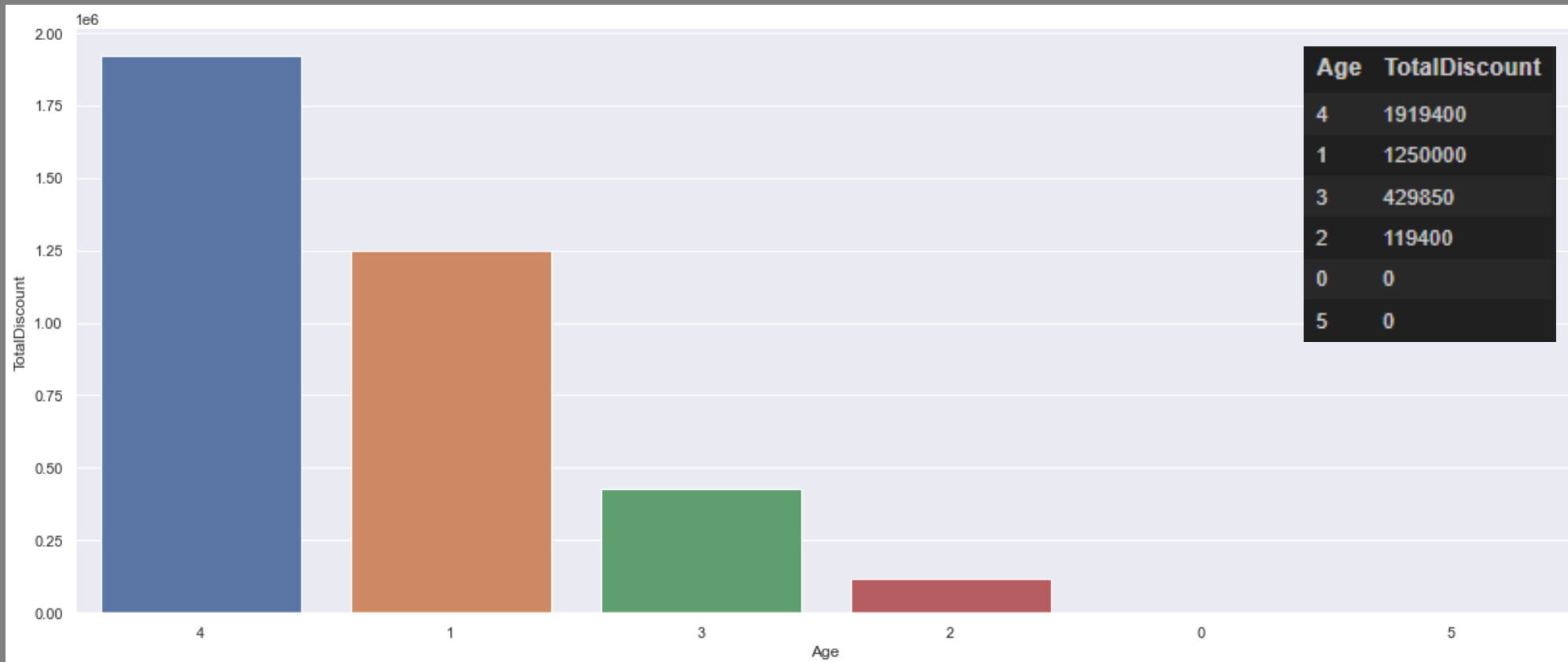
EXPLORATORY DATA ANALYSIS

TIPE TRANSAKSI YANG DILAKUKAN MASYARAKAT

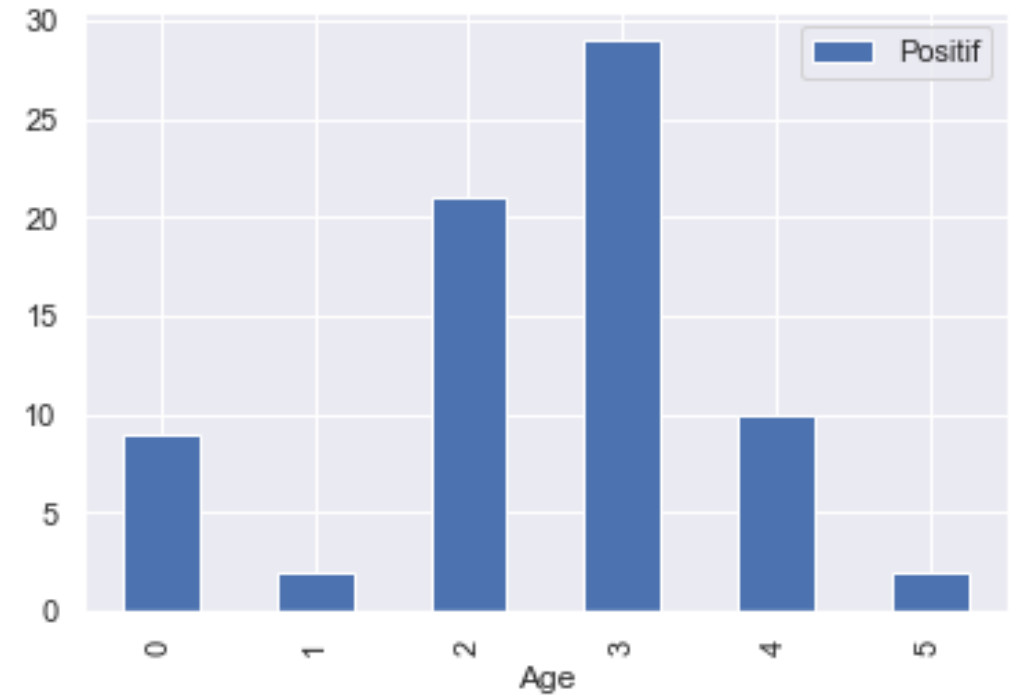


EXPLORATORY DATA ANALYSIS

Kelompok umur dengan diskon paling besar yang diterima



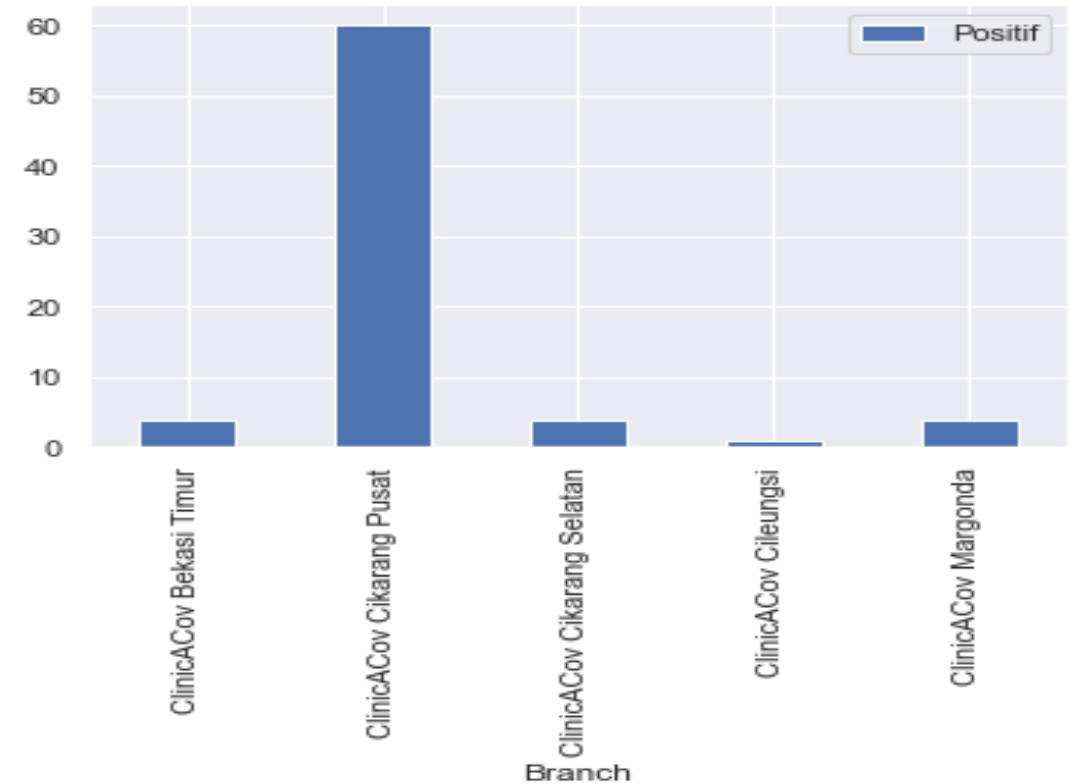
Age	Positif	Total Tes	Persentase
0	9	273	3.30
1	2	388	0.52
2	21	98	21.43
3	29	153	18.95
4	10	70	14.29
5	2	18	11.11



EXPLORATORY DATA ANALYSIS

KELOMPOK UMUR DENGAN KASUS POSITIF TERTINGGI

	Branch	Positif	Total Tes	Persentase
	ClinicACov Cikarang Selatan	4	111	3.60
	ClinicACov Cikarang Pusat	60	707	8.49
	ClinicACov Cileungsi	1	30	3.33
	ClinicACov Bekasi Timur	4	93	4.30
	ClinicACov Bekasi Selatan	0	14	0.00
	ClinicACov Margonda	4	31	12.90
	ClinicACov Cibubur	0	10	0.00
	ClinicACov Cikarang Utara	0	4	0.00



EXPLORATORY DATA ANALYSIS

KLINIK DENGAN KASUS POSITIF TERTINGGI

MACHINE LEARNING MODELING

- Discount Prediction (*diskon yang akan diberikan*)

Features(Independent Variable):

1. **TrxType** ==> Tipe transaksi yang dilakukan.
2. **Branch** ==> Klinik tempat melakukan tes.
3. **Package** ==> Paket tes yang diambil.
4. **PatientGender** ==> Jenis kelamin pasien.
5. **OrderCreated** ==> Kapan terjadi transaksi tes.
6. **Age** ==> Umur Pasien.
7. **Month** ==> Bulan tes dilakukan.

TARGET VARIABLE (Y) ==> **TotalDiscount**

- Grand Total Prediction (*Harga tes yang akan diberikan*)

Features(Independent Variable):

1. **TrxType** ==> Tipe transaksi yang dilakukan.
2. **Branch** ==> Klinik tempat melakukan tes.
3. **Package** ==> Paket tes yang diambil.
4. **PatientGender** ==> Jenis kelamin pasien.
5. **OrderCreated** ==> Kapan terjadi transaksi tes.
6. **Age** ==> Umur Pasien.
7. **Month** ==> Bulan tes dilakukan.
8. **Date** ==> Tanggal/Hari tes dilakukan.

TARGET VARIABLE (Y) ==> **Grand Total**

CONFUSION MATRIX

LOGISTIC REGRESSION

Total Discount Confusion Matrix

Actual	0	192	0	1	0
	1	2	0	0	0
	2	3	0	0	0
	3	2	0	0	0
		0	1	2	3
		Predicted			

SVM

Total Discount Confusion Matrix

Actual	0	193	0	0	0
	1	2	0	0	0
	2	3	0	0	0
	3	2	0	0	0
		0	1	2	3
		Predicted			

RANDOM FOREST

Total Discount Confusion Matrix

Actual	0	193	0	0	0
	1	0	2	0	0
	2	0	0	3	0
	3	0	0	0	2
		0	1	2	3
		Predicted			

Scaled

Total Discount Confusion Matrix

Actual	0	193	0	0	0
	1	2	0	0	0
	2	3	0	0	0
	3	0	0	0	2
		0	1	2	3
		Predicted			

Total Discount Confusion Matrix

Actual	0	193	0	0	0
	1	2	0	0	0
	2	3	0	0	0
	3	0	0	0	2
		0	1	2	3
		Predicted			

Total Discount Confusion Matrix

Actual	0	193	0	0	0
	1	1	1	0	0
	2	3	0	0	0
	3	0	0	0	2
		0	1	2	3
		Predicted			

KESIMPULAN DAN SARAN

- Memberikan lebih banyak **supply peralatan medis** dan **obat-obatan** ke klinik-klinik dengan persentase positif COVID-19 yang tinggi dan ramai pasien yang tinggi.
- Melakukan sosialisasi dan edukasi tentang bahaya COVID-19.
- Memakai *prototype machine learning* ini di klinik-klinik yang kurang kunjungan pasien.
- Menawarkan **diskon-diskon** menarik kepada orang-orang dengan kelompok umur **19 sampai 45** tahun, serta memberikan **edukasi** kepada orang-orang dengan kelompok umur **19 sampai 30**.
- Model terbaik yang didapat adalah **Random Forest** tanpa scaling untuk memprediksi diskon, sementara untuk memprediksi harga Swab Test digunakan algoritma **Decision Tree**.

KESIMPULAN DAN SARAN

01

Menyediakan data *real* yang tidak banyak *duplicate*.

02

Atribut terlalu sedikit.

03

Varibel target yang akan diprediksi memiliki **ketidakseimbangan** data, dapat dicari data yang *balance* atau dilakukan *Handling Imbalance Data*.