

## **OCTAGRAM (7A)**

**Cicilia Rumanat Siahaan  
Wilson Firmanda  
Manurung  
Yansen Cristopel Sinaga  
Rahmat Diza Ramadhan  
Alfaroghi Swandi  
Arief Sulistyo Wibowo  
Yulia Rohmawati  
Marifatul Hasanah**

**Dokumen Laporan  
Final Project**



# Latar Belakang Masalah

Dataset: Banking

## Predicting Churn for Bank Customers

### Problem Statement:

Jumlah nasabah atau customer yang churn pada perusahaan perbankan (Banking) semakin meningkat. Hal ini menyebabkan pendapatan perusahaan menurun. Churn adalah istilah yang menggambarkan seorang customer berpindah ke layanan yang lain atau berhenti menggunakan layanan. Oleh karena itu perlu memprediksi nasabah atau customer bank yang akan churn berdasarkan berbagai atribut yang ada.

### Objective:

- Mengidentifikasi faktor-faktor yang memengaruhi nasabah bank churn.
- Melakukan segmentasi nasabah atau customer bank berdasarkan atribut yang tersedia
- Membuat dan mengembangkan model Machine Learning untuk memprediksi customer atau nasabah bank yang berpotensi churn.

### Role:

- Data Scientist yang bekerja di perusahaan perbankan yang bertugas untuk membuat, menguji, mengembangkan dan mengevaluasi model Machine Learning untuk memprediksi customer atau nasabah yang berpotensi churn.
- Data Analyst yang bekerja sebagai analis dataset yang ada.
- Data Engineer yang bekerja sebagai pengumpulan dataset serta manajemen dataset

### Goal:

Mempertahankan customer atau nasabah bank agar tidak jadi churn sehingga pendapatan perusahaan perbankan dapat meningkat kembali.

### Business Metrics:

#### Churn Rate

Tingkat pergantian (churn rate) dalam data ini merujuk pada tingkat di mana nasabah atau pelanggan bank berhenti menggunakan layanan atau menutup akun mereka. Churn rate ini bertujuan untuk memberikan wawasan tentang seberapa baik bank mempertahankan pelanggan dan mencegah mereka beralih ke pesaing. Semakin rendah Churn Rate, semakin baik kinerja bisnis dalam mempertahankan pelanggan.



# Dataset

## Sumber Data: Kaggle (Predicting Churn For Bank Customers)

Dataset memiliki 10.000 baris data yang terdiri dari profil nasabah bank yang berhubungan dengan potensi churn-nya dengan atribut kolom sebagai berikut.

<b>Customer ID</b>	Merepresentasikan identifikasi unik nasabah
<b>Surname</b>	Merepresentasikan last name nasabah
<b>CreditScore</b>	Merepresentasikan angka yang mengestimasi kemampuan nasabah dalam melunasi utangnya tepat waktu
<b>Geography</b>	Merepresentasikan wilayah nasabah
<b>Gender</b>	Merepresentasikan jenis kelamin nasabah (male or female)
<b>Age</b>	Merepresentasikan usia nasabah
<b>Tenure</b>	Merepresentasikan jangka waktu pinjaman nasabah
<b>Balance</b>	Merepresentasikan saldo pada akun nasabah
<b>NumOfProducts</b>	Merepresentasikan jumlah produk yang digunakan atau dimiliki oleh nasabah
<b>HasCrCard</b>	Merepresentasikan kepemilikan nasabah akan kartu kredit
<b>IsActiveMember</b>	Merepresentasikan apakah member nasabah aktif atau tidak
<b>EstimatedSalary</b>	Merepresentasikan estimasi gaji nasabah
<b>Exited</b>	Merepresentasikan nasabah churn atau tidak oleh nasabah

# Pre-processing

## A. Handle Missing Values

```
df.isnull().sum() #tidak ada data missing value
```

RowNumber	0
CustomerId	0
Surname	0
CreditScore	0
Geography	0
Gender	0
Age	0
Tenure	0
Balance	0
NumOfProducts	0
HasCrCard	0
IsActiveMember	0
EstimatedSalary	0
Exited	0
dtype:	int64

Dataset yang tersedia **tidak memiliki missing values** sehingga tidak memerlukan tindakan apapun terkait hal tersebut.

## B. Handle Duplicated Data & Redundant Data

```
df.duplicated() #tidak ada data duplikat
```

0	False
1	False
2	False
3	False
4	False
...	...
9995	False
9996	False
9997	False
9998	False
9999	False
Length:	10000, dtype: bool

```
[ ] df.duplicated().sum() #duplikat
```

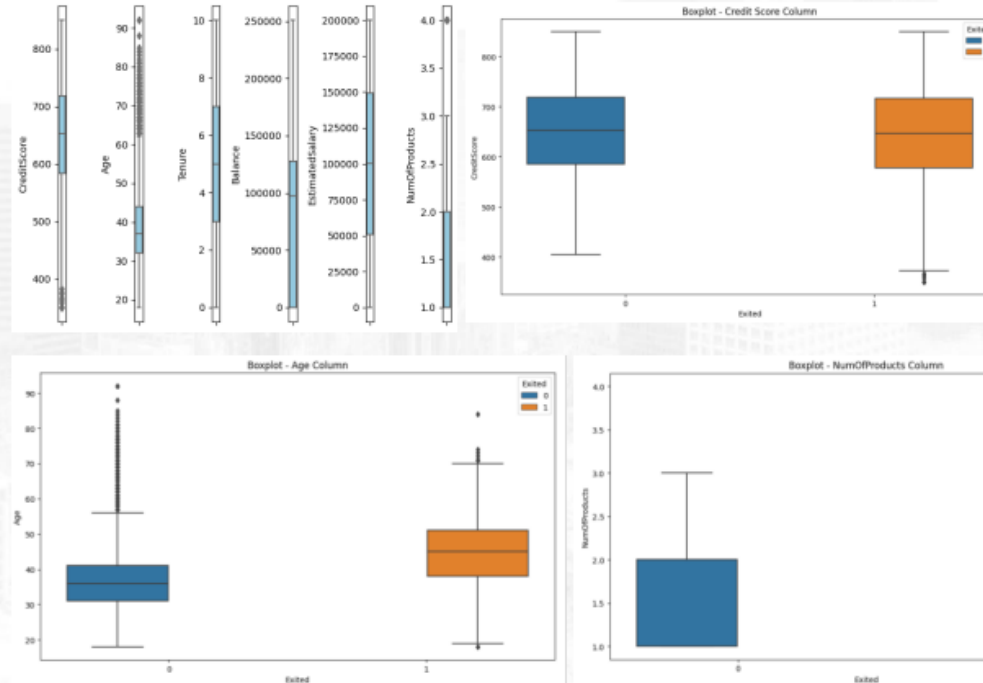
0	

```
[ ] df.T.duplicated().sum() #redundant
```

0	

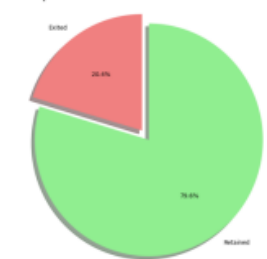
Dataset yang tersedia **tidak memiliki data yang dupikat & redundant**, sehingga tidak ada yang perlu dilakukan terkait hal tersebut.

## C. Handle Outliers



Kolom "Age", "Credit Score", dan kolom "Num of Products" memiliki data outlier. **Data outlier tersebut tetap digunakan** karena sebagian besar data outlier ada pada data nasabah yang churn, dimana nasabah churn memiliki jumlah data yang sedikit yaitu sekitar 20%. Apabila data outliers tersebut dihapus, kemungkinan dapat menghilangkan informasi data. Selain itu, metode yang digunakan untuk data ini adalah **metode yang robust terhadap outlier**.

Proporsi Customer Churn dan Tidak Churn



```
#Handle Outlier
# Kolom dengan outlier
kolom_outlier = ['CreditScore', 'Age', 'NumOfProducts']

# Mengatasi Outlier
for k in kolom_outlier:
    Q1 = x_ho[k].quantile(0.25)
    Q3 = x_ho[k].quantile(0.75)
    IQR = Q3 - Q1
    lower_bound = Q1 - 1.5 * IQR
    upper_bound = Q3 + 1.5 * IQR

    x_ho[k] = np.where(x_ho[k] > upper_bound, upper_bound, x_ho[k])
    x_ho[k] = np.where(x_ho[k] < lower_bound, lower_bound, x_ho[k])
```

```
# Drop Outlier
data_do = data.copy()

# Kolom yang memiliki outlier
kolom_outlier = ['CreditScore', 'Age', 'NumOfProducts']

# Function
def remove_outliers(df, column):
    Q1 = df[column].quantile(0.25)
    Q3 = df[column].quantile(0.75)
    IQR = Q3 - Q1
    lower_bound = Q1 - 1.5 * IQR
    upper_bound = Q3 + 1.5 * IQR
    return df[(df[column] >= lower_bound) & (df[column] <= upper_bound)]

# Menghapus outlier untuk setiap kolom dengan outlier
for j in kolom_outlier:
    data_do = remove_outliers(data_do, j)
```

Pada Clustering, modelling juga di-trial & error pada dataset yang outliernya **di-handle dengan IQR** dan dataset yang outliernya **dihapus**.

# Pre-processing

## D. Feature Transformation

```
Robust Scaling
from sklearn.preprocessing import RobustScaler

scaler = RobustScaler()
Numeric = ['CreditScore', 'Balance', 'EstimatedSalary', 'NumOfProducts', 'Tenure', 'Age']
data_RS = df2.copy()
data_RS[Numeric] = scaler.fit_transform(data_RS[Numeric])
data_RS
```

	CreditScore	Gender	Age	Tenure	Balance	NumOfProducts	HasCrCard	IsActiveMember	EstimatedSalary	Exited	France	Germany	Spain
0	-0.240268	0	0.416667	-0.75	-0.761480	0.0	1	1	0.011739	1	1	0	0
1	-0.320269	0	0.203333	-1.00	-0.104906	0.0	0	1	0.120112	0	0	0	1
2	-1.119403	0	0.416667	0.75	0.480346	2.0	1	0	0.139630	1	1	0	0
3	0.300746	0	0.166667	-1.00	-0.761480	1.0	0	0	-0.064717	0	1	0	0
4	1.477612	0	0.500000	-0.75	0.221806	0.0	1	1	-0.214661	0	0	0	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...
9995	0.880960	1	0.166667	0.00	-0.761480	1.0	1	0	-0.036676	0	1	0	0
9996	-1.014925	1	0.166667	1.25	-0.312091	0.0	1	1	0.015336	0	1	0	0
9997	0.425373	0	-0.083333	0.50	-0.761480	0.0	0	1	-0.060615	1	1	0	0
9998	0.895522	1	0.416667	-0.50	-0.173519	1.0	1	0	-0.074252	1	0	1	0
9999	1.044776	0	-0.750000	-0.25	0.258054	0.0	1	0	-0.630002	0	1	0	0

10000 rows x 13 columns

Metode *feature transformation* yang digunakan adalah metode **Robust Scalling**. Metode ini merupakan metode *feature transformation* yang robust terhadap data outlier.

## E. Feature Encoding

	CreditScore	Gender	Age	Tenure	Balance	NumOfProducts	HasCrCard	IsActiveMember	EstimatedSalary	Exited	France	Germany	Spain
0	619	0	42	2	0.00	1	1	1	101348.88	1	1	0	0
1	608	0	41	1	83807.86	1	0	1	112542.58	0	0	0	1
2	502	0	42	8	159660.80	3	1	0	113931.57	1	1	0	0
3	699	0	39	1	0.00	2	0	0	93826.63	0	1	0	0
4	850	0	43	2	125510.82	1	1	1	79084.10	0	0	0	1
5	645	1	44	8	113755.78	2	1	0	149756.71	1	0	0	1
6	822	1	50	7	0.00	2	1	1	10062.80	0	1	0	0

- Feature "Gender" dilakukan *feature encoding* dengan metode **label encoding** menjadi 0 dan 1 (0: Female; 1: Male)
- Feature "Geography" dilakukan *feature encoding* dengan metode **One Hot Encoding (OHE)**.

\*Tambahan pada Feature Transformation

### Feature Transformation : Standarization

```
[165] # Standarisasi
from sklearn.preprocessing import StandardScaler

xo_std = x_do.copy()
float_columns = x_do.select_dtypes(include=['float64']) #scaling data pada kolom numerik

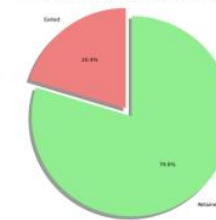
std_scaler = StandardScaler()
scaled_std = std_scaler.fit_transform(float_columns)

xo_std[float_columns.columns] = scaled_std
```

Pada Clustering, karena modelling *di-trial and error* pada dataset yang outliernya di-handle dengan outlier maupun dataset yang outliernya dihapus, maka *feature transformation* yang digunakan adalah **Standarization**.

## F. Handle Class Imbalance

Proporsi Customer Churn dan Tidak Churn



Dataset mentah memiliki proporsi data yang *imbalanced* yaitu pada kolom target 'Exited' jumlah nasabah yang churn jauh lebih sedikit dibandingkan dengan jumlah nasabah yang retain yaitu sebesar 20.4%. Oleh karena itu perlu **Handle Class Imbalance** dengan metode **SMOTE (Synthetic Minority Over-sampling Technique)**. Dengan metode SMOTE, jumlah akan bertambah yang sebelumnya 10000 menjadi 15926 dengan proporsi jumlah nasabah yang churn dan tidak churn sudah sama yaitu sama-sama sebesar 50% data.

```
#SMOTE
from imblearn.over_sampling import SMOTE

#feature X dan Y
X = data.drop('Exited', axis=1)
y = data['Exited']

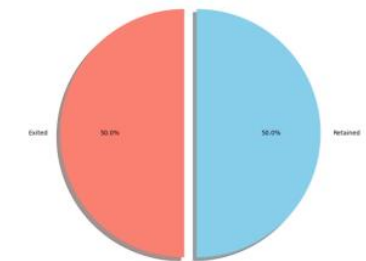
#Handling Class Imbalance
smote = SMOTE(random_state=1)
X_resampled, y_resampled = smote.fit_resample(X, y)

#Hasil SMOTE
df_SMOTE = pd.concat([pd.DataFrame(X_resampled), pd.DataFrame(y_resampled, columns=['Exited'])], axis=1)

exited_counts = df_SMOTE['Exited'].value_counts()
print("Jumlah data setelah SMOTE:", len(df_SMOTE))
print(exited_counts)

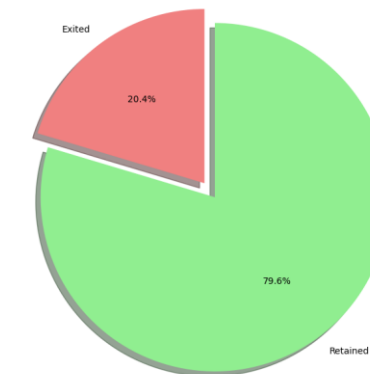
Jumlah data setelah SMOTE: 15926
1    7963
0    7963
Name: Exited, dtype: int64
```

Proporsi Customer Churn dan Tidak Churn Setelah Handle Class Imbalance



\*Tambahan pada Feature Transformation

Proporsi Customer Churn dan Tidak Churn

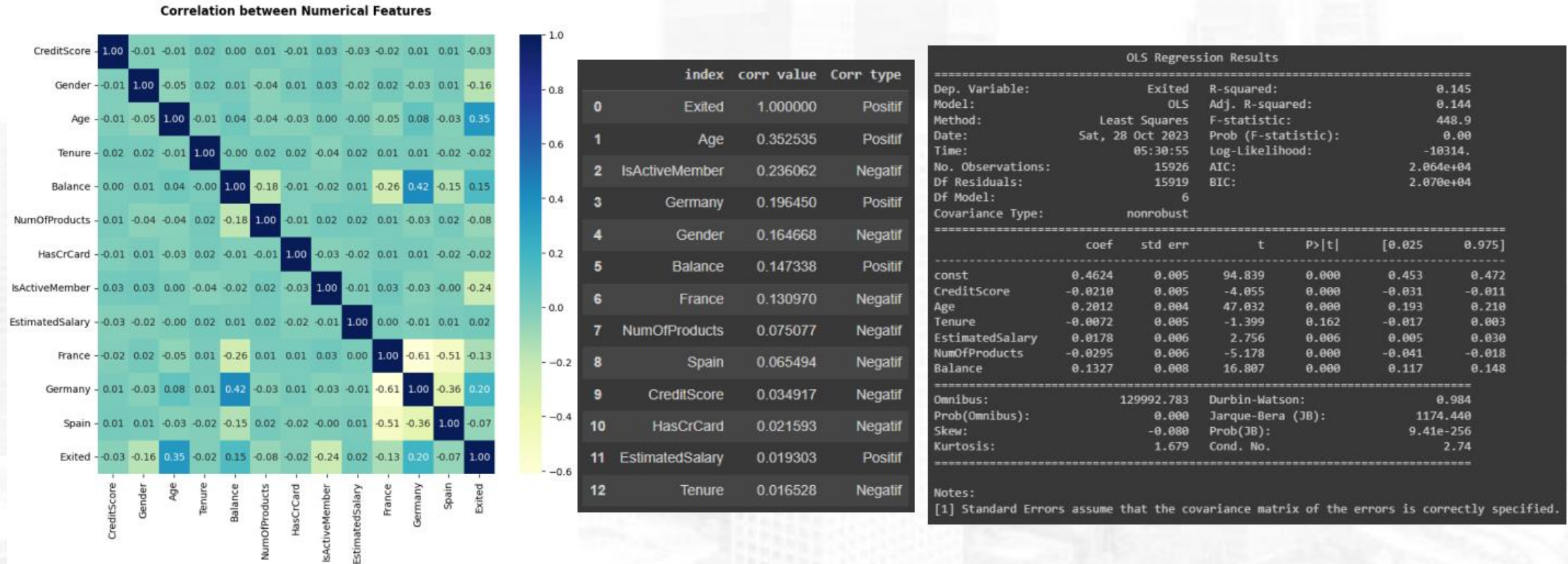


Pada Clustering, dataset yang digunakan tidak menggunakan label, sehingga handle class imbalance pada modelling Clustering tidak diperlukan.



# Pre-processing (Feature Engineering)

## A. Feature Selection



Untuk Data Numeric ('Credit Score', 'Age', 'Tenure', 'Estimated Salary', 'Num of Products', 'Balance'), **feature importance** dilihat dari **heatmap correlation** dan **Regression Linear (OLS)**. Berdasarkan hal tersebut, **kolom 'Tenure'** merupakan kolom dengan **corr value terkecil yaitu sebesar 0.016528**. Sehingga perlu pertimbangan untuk menghapus kolom tersebut.

# Pre-processing (Feature Engineering)

## B. Feature Extraction

	country_France	country_Germany	country_Spain
0	1	0	0
1	0	0	1
2	1	0	0
3	1	0	0
4	0	0	1
...	...	...	...
9995	1	0	0
9996	1	0	0
9997	1	0	0
9998	0	1	0
9999	1	0	0

10000 rows x 3 columns

**Feature Extraction** : Membuat feature baru dari feature yang sudah ada

- *One Hot Encoding (OHE)* pada kolom 'Geography'.
- Mengelompokkan nasabah berdasarkan kolom 'Age' dengan ketentuan klasifikasi menurut Departemen Kesehatan RI.
- Mengelompokkan nasabah berdasarkan kolom 'EstimatedSalary'.
- Mengelompokkan nasabah berdasarkan kolom 'IsActiveMember' dan kolom 'HasCrCard'

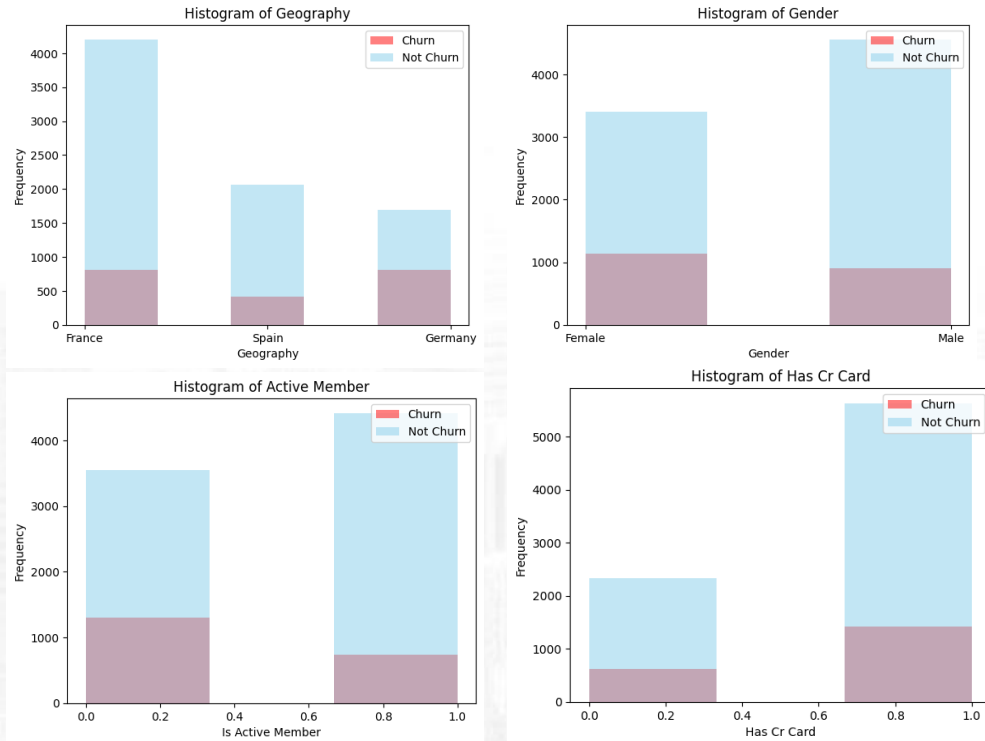
**Note s:** feature baru ini masih perlu pertimbangan apakah akan digunakan dalam model Machine Learning.

	CreditScore	Geography	Gender	Age	Tenure	Balance	NumOfProducts	HasCrCard	IsActiveMember	EstimatedSalary	Exited	Age_Category	Salary_Category	Aktif & Credit card
0	619	France	0	42	2	0.00	1	1	1	101348.88	1	Dewasa	Sedang	1
1	608	Spain	0	41	1	83807.86	1	0	1	112542.58	0	Dewasa	Sedang	0
2	502	France	0	42	8	159660.80	3	1	0	113931.57	1	Dewasa	Sedang	0
3	699	France	0	39	1	0.00	2	0	0	93826.63	0	Dewasa	Rendah	0
4	850	Spain	0	43	2	125510.82	1	1	1	79084.10	0	Dewasa	Rendah	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
9995	771	France	1	39	5	0.00	2	1	0	96270.64	0	Dewasa	Rendah	0
9996	516	France	1	35	10	57369.61	1	1	1	101699.77	0	Dewasa	Sedang	1
9997	709	France	0	36	7	0.00	1	0	1	42085.58	1	Dewasa	Sangat Rendah	0
9998	772	Germany	1	42	3	75075.31	2	1	0	92888.52	1	Dewasa	Rendah	0
9999	792	France	0	28	4	130142.79	1	1	0	38190.78	0	Dewasa	Sangat Rendah	0

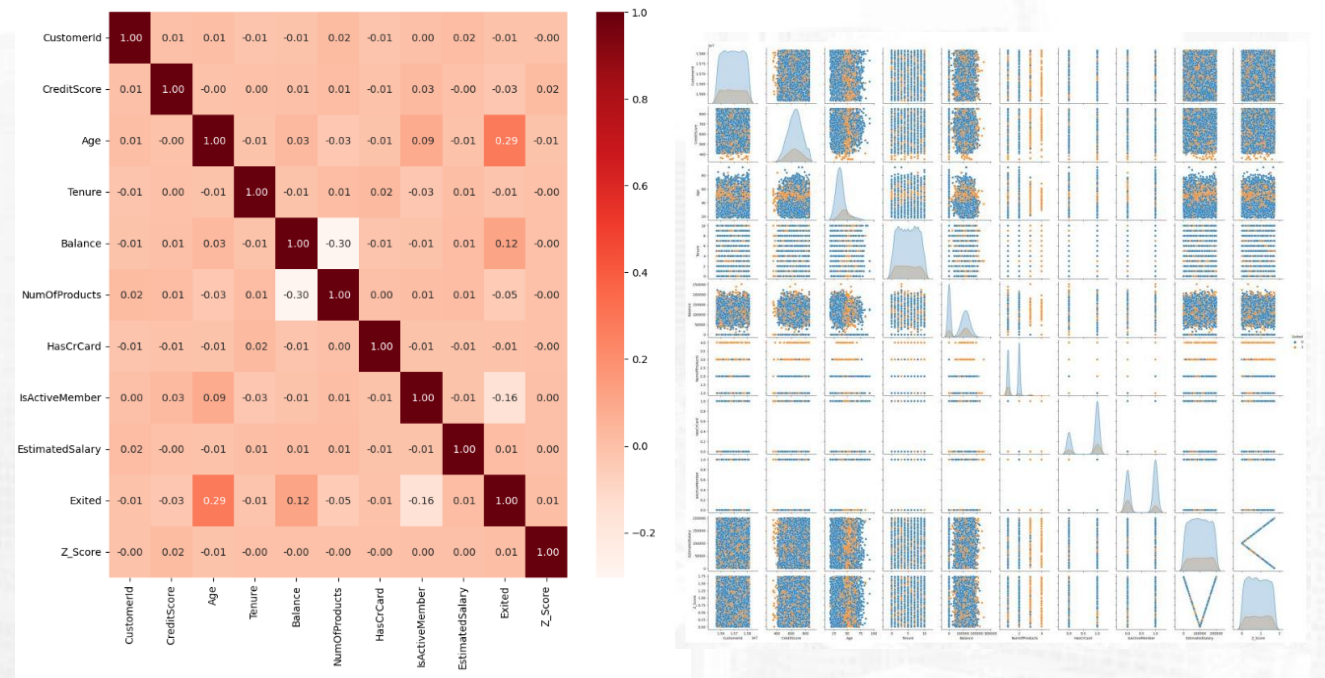
10000 rows x 14 columns



# STAGE 2 - Insights



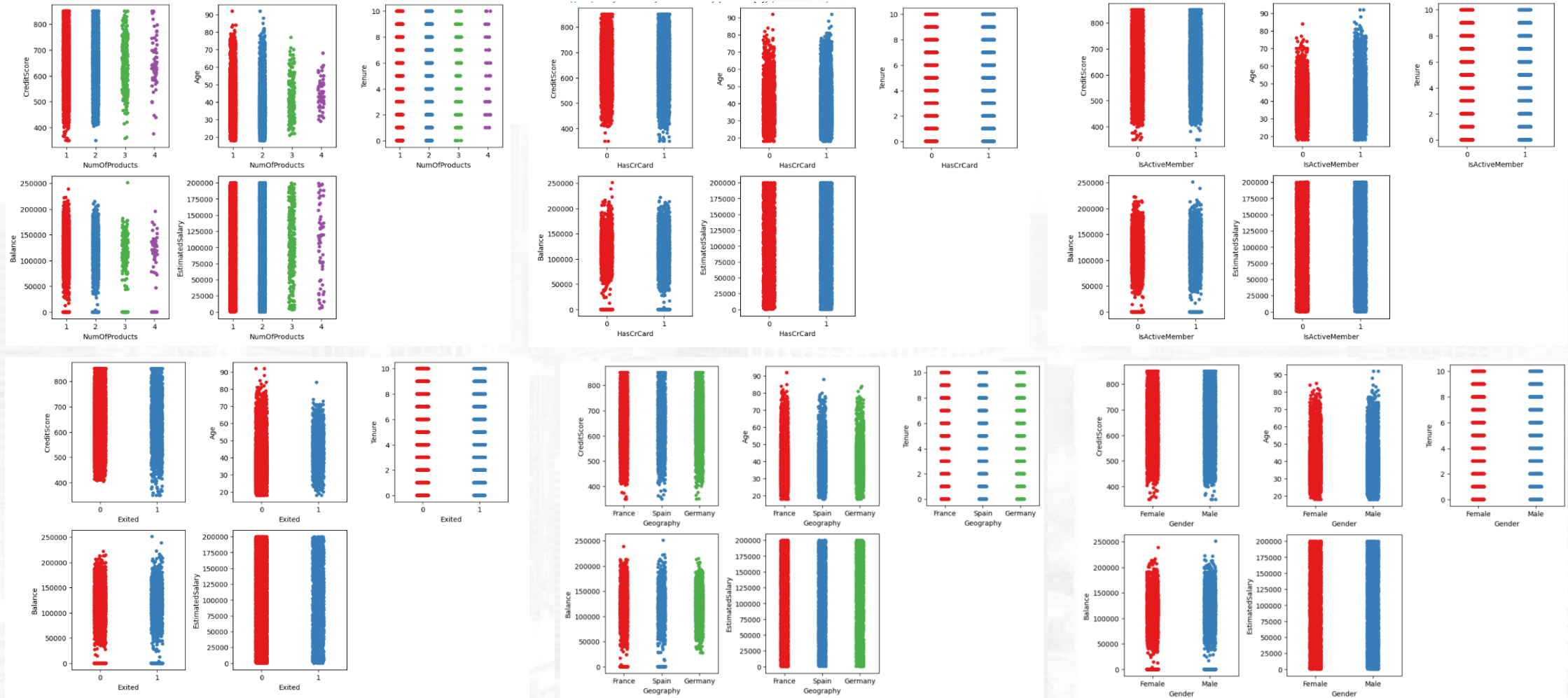
- Nasabah yang berasal dari Jerman adalah nasabah yang churn terbanyak sedangkan nasabah yang berasal dari Perancis adalah nasabah yang retention terbanyak.
- Nasabah berjenis kelamin perempuan lebih banyak yang churn dibandingkan dengan laki-laki.
- Nasabah yang merupakan member pasif cenderung churn dibandingkan dengan nasabah bank yang merupakan active member.
- Nasabah yang memiliki kartu kredit lebih banyak yang churn dibandingkan dengan yang tidak memiliki kartu kredit, akan tetapi nasabah yang tetap loyal (tidak churn) juga banyak yang memiliki kartu kredit.



- Berdasarkan heatmap yang dihasilkan, **terdapat korelasi antara feature (kolom) dengan label (Exited)** sehingga model linier bisa digunakan ke tahap selanjutnya.
- Kemudian pada korelasi tersebut ada dua dua *feature* yang berkorelasi positif (semakin tinggi nilai nya maka korelasi berbanding lurus), lalu ada tiga *feature* yang berkorelasi negatif (semakin rendah nilainya korelasi berbanding terbalik).
- Distribusi pada setiap kolom numerik seperti Age, Credit Score, Num of Product, Estimated Salary, Balance dan Tenure tidak berdistribusi normal.



# STAGE 2 - Insights



Berdasarkan kategorikal plot :

- Tidak terdapat korelasi yang signifikan pada *feature numerical* (credit score, balance, estimated salary, age dan tenure) dengan *feature kategori num of product* dan *has credit card*.
- Tidak terdapat korelasi yang signifikan pada *feature numerical* (credit score, balance, estimated salary, age dan tenure) dengan *feature kategori is active member* dan *exited*.
- Tidak terdapat korelasi yang signifikan pada *feature numerical* (credit score, balance, estimated salary, age dan tenure) dengan *feature kategori geography* dan *gender*.

# STAGE 3 - Modelling Experiments

## Classification

Dalam classification kami melakukan uji coba beberapa model, di antaranya adalah Random Forest, Decision Tree, SVM, dan Naïve Bayes. Berikut ini adalah hasil uji coba dari keempat model tersebut.

Model	Accuracy	Precision	Recall	F1	AUC
Random Forest	0,91	0,93	0,89	0,91	0,97
Descion Tree	0,83	0,83	0,82	0,83	0,91
SVM	0,87	0,88	0,84	0,86	0,94
Naive Bayes	0,81	0,80	0,80	0,80	0,88

- Berdasarkan uji coba, model dengan performa terbaik adalah **RANDOM FOREST**.
- RANDOM FOREST** memiliki nilai **F1-Score tertinggi** di antara model lainnya. F1-Score adalah rata-rata harmonis antara presisi (precision) dan recall (sensitivitas).
- Pemilihan **F1-Score** sebagai evaluation metric karena dataset mengalami ketidakseimbangan kelas sehingga perlu untuk mengidentifikasi nasabah yang akan churn dengan akurasi yang tinggi dan meminimalkan jumlah nasabah yang salah diklasifikasikan sebagai churn.



# STAGE 3 - Modelling Experiments

## Feature Importance (Classification)

Feature	Feature Importance
Age	0.2252352
Num of Products	0.18396855
Is Active Member	0.09250019
Balance	0.09296625
Estimated Salary	0.084213
Credit Score	0.07821962
Tenure	0.0627265
France	0.05348106
Gender	0.04822749
Spain	0.04100564
Germany	0.02391431
Has Cr Card	0.01354218

- **Feature Importance** secara berurutan dari yang tertinggi ke terendah beserta coefisien-nya adalah seperti pada gambar di samping.
- 7 Feature yang memiliki **importance tertinggi** adalah Age, Num of products, Is Active Member, Balance, Estimated Salary, Credit Score dan Tenure.
- 7 Feature tersebut akan dilakukan segmentasi (clustering) untuk melihat segmentasi profil nasabah bank.

# STAGE 3 - Modelling Experiments

## Clustering

Pada clustering kami juga melakukan uji coba beberapa model, di antaranya adalah K-Medoids, DBSCAN, GMM, Complete Linkage, K-Means, dan K-Prototype. Dalam modeling ini, kami menggunakan beberapa metode untuk outliers, yakni dengan without handle outlier, handle outlier with IQR, dan drop outlier.

Clustering (Data Without Handle Outlier)				
Evaluation Metrics	K-Medoids	DBSCAN	Complete Linkage	GMM
Silhouette Score	0.0725	-0.1761	0.1786	0.0993
Davies-Bouldin	3.5634	1.7588	1.940	2.7758
Calinski-Harabasz	738.5761	7.7491	913.0237	894.9819
Evaluation Metrics	Clustering (Handle Outlier With IQR)		Clustering (Drop Outlier)	
	K-Means	K-Prototype	K-Means	K-Prototype
Silhouette Score	0.1556	0.1557	0.1378	
Davies-Bouldin	1.9757	1.9740	2.2109	
Calinski-Harabasz	1458.3229	1454.9214	1513.0909	

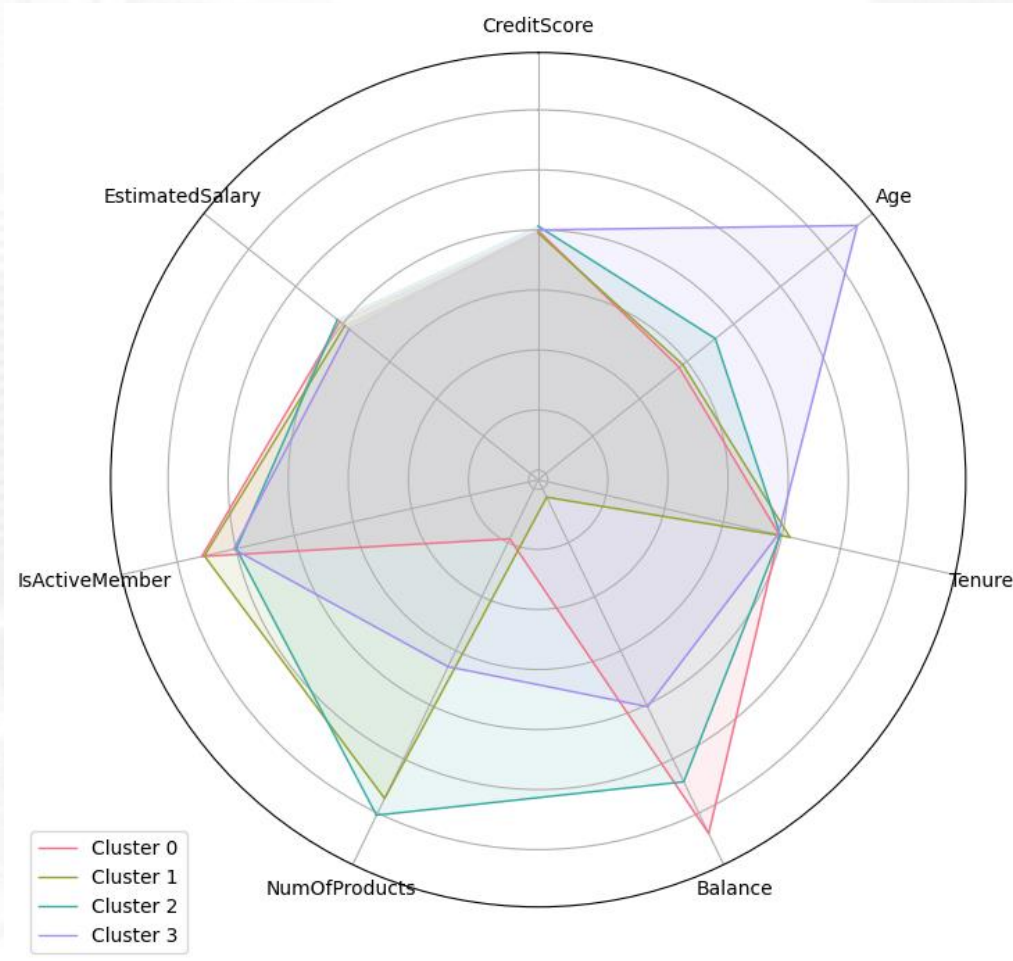
### MODEL TERBAIK

- Berdasarkan hasil uji coba, model dengan performa terbaik adalah **K-Prototype** dengan data outlier di-handle dengan IQR.
- K-Prototype** dipilih berdasarkan nilai **Silhouette Score** tertinggi di antara model lainnya.



# STAGE 3 - Modelling Experiments

## Feature Comparison (Clustering)



- Pada **Cluster 0** profil nasabah bank adalah nasabah yang aktif, dengan balance paling tinggi, tenure lebih lama, dan memiliki credit score serta estimated salary yang tinggi. Cluster ini memiliki rentang usia yang cenderung beragam. Sementara, untuk kepemilikan produk lebih rendah.
- Pada **Cluster 1** profil nasabah bank adalah nasabah aktif dengan kepemilikan jumlah produk yang tinggi, tenure sedang, estimated salary credit score sedang. Selain itu, nasabah pada cluster ini memiliki balance rendah dan usia nasabah cenderung beragam.
- Pada **Cluster 2** profil nasabah adalah nasabah pasif yang kepemilikan produknya tinggi dibandingkan dengan cluster lain. Nasabah pada cluster ini juga memiliki balance lebih tinggi dibanding cluster 1 dan 3, dengan tenure sedang, estimated salary dan credit score sedang. Untuk usia, nasabah pada cluster ini cenderung beragam.
- Pada **Cluster 3** nasabah cenderung pasif. Nasabah pada cluster ini memiliki usia lebih tua. Kepemilikan produk dan balance sedang. Sementara itu, lama tenure, estimated salary, dan credit score juga sedang.

# STAGE 4 – Business Recommendation



- Meningkatkan **komunikasi** antara bank dengan nasabah sehingga dapat meningkatkan keaktifan dan loyalitas nasabah.
- Meningkatkan **retensi pelanggan** dengan menawarkan insentif atau promosi pada nasabah yang memiliki banyak produk.
- Melakukan **survey kepuasan** atas produk yang dimiliki nasabah yang mencakup evaluasi produk, sehingga bank dapat mengembangkan produk sesuai dengan preferensi nasabah.
- Memberikan **limit kredit yang panjang** bagi nasabah yang membayar tepat waktu.
- Memberi **reward** kepada nasabah yang setia agar memperkuat loyalitas dan memotivasi nasabah untuk tetap setia.
- Menghadirkan **produk dan layanan sesuai umur** sehingga meningkatkan efektivitas pemasaran dan memastikan bahwa bank secara efisien memenuhi harapan nasabah.
- Melakukan **promosi beragam produk bank** serta manfaatnya sehingga dapat meningkatkan jumlah penggunaan produk bank oleh nasabah.
- Memberikan **layanan yang lebih mudah dan cepat**, seperti layanan perbankan digital yang lebih canggih dan efisien.

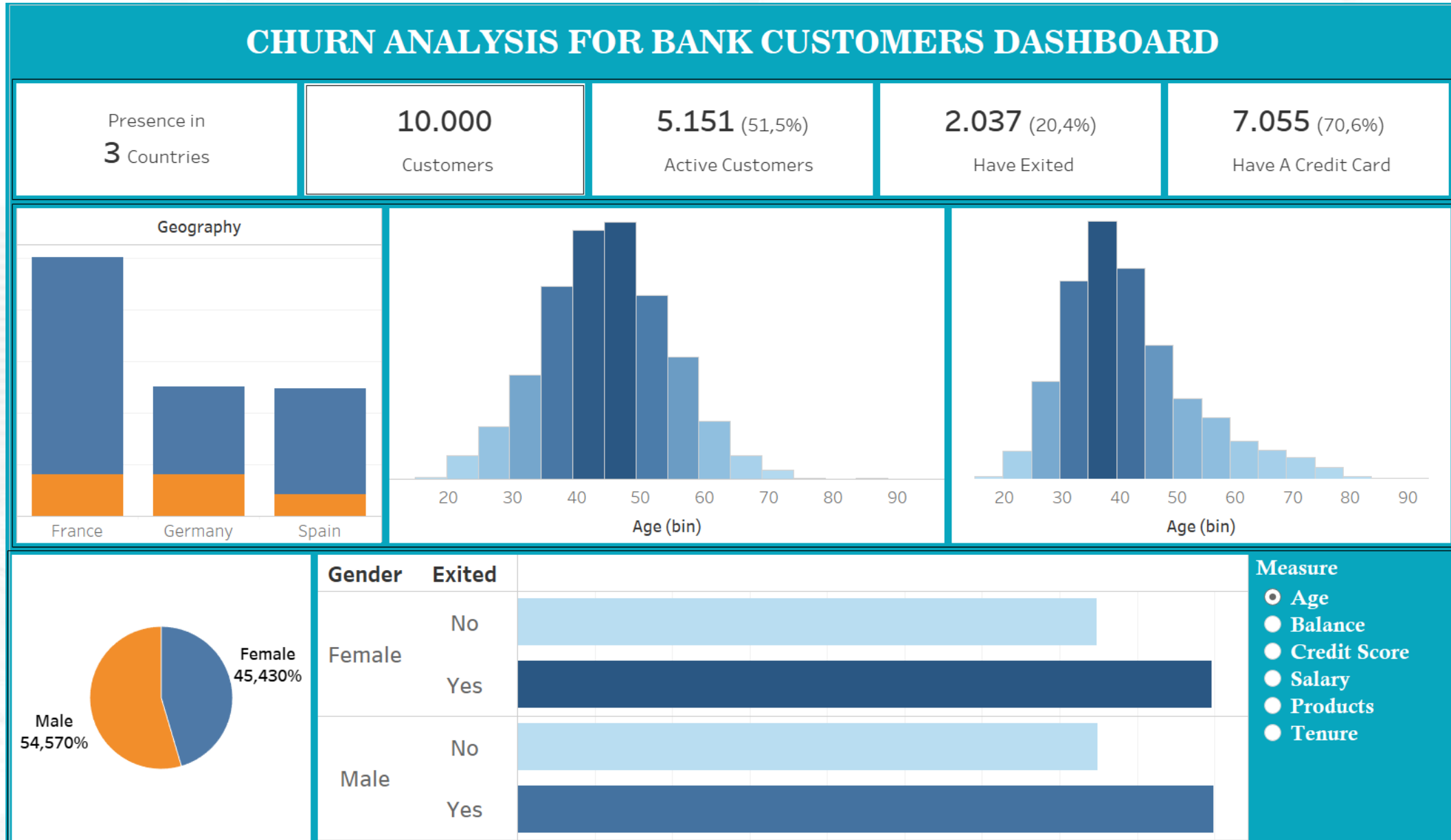
Setelah mencoba melakukan rekomendasi:

- **Evaluasi Kebijakan Penanganan Churn** Berdasarkan hasil model pihak bank bisa mengevaluasi kebijakan penindakan bagi nasabah yang cenderung akan churn.
- **Menyempurnakan Strategi Pemasaran** Berdasarkan hasil evaluasi kebijakan, Bank dapat menyempurnakan strategi pemasaran terkait nasabah yang memiliki kecenderungan akan churn.





# Output Tambahan : Dashboard



# Pembagian Tugas

Nama Lengkap	Stage 0	Stage 1	Stage 2	Stage 3	Stage 4
Cicilia Rumanat Siahaan	- Problem Statement	- Univariate Analysis	- Handling Class Imbalance - Feature Selection - Feature Transformation	- Classification - Clustering - Feature Importance - Business Insight - Recommendation	- PPT – Preprocessing - PPT – Clustering - PPT Laporan Akhir - Dashboard (Tableau)
Wilson Firmanda Manurung	- Role	- Descriptive Analysis	- Handling Class Imbalance	- Classification	- PPT – Modelling - PPT Laporan Akhir
Rahmat Diza Ramadhan	- Goals	- Univariate Analysis	- Feature Extraction	- Clustering	- PPT – EDA & Insight
Yansen Cristopel Sinaga	- Goals	- Git	- Feature Transformation	- Clustering	- PPT – Background
Yulia Rohmawati	- Objective	- Business Insight	- Git	- Classification - Business Insight - Recommendation	- PPT – Modelling - PPT – Flowchart - PPT Laporan Akhir
Arief Sulistyo Wibowo	- Objective	- Multivariate Analysis	- Feature Encoding	- Clustering	- PPT – Business Recommendation
Marifatul Hasanah	- Business Metrics	- Business Insight	- Feature Selection	- Classification - Business Insight - Recommendation	- PPT – Business Recommendation - PPT Laporan Akhir
Alfaroghi Swandi	- Business Metrics	- Business Insight	- Feature Extraction	- Clustering	- PPT – EDA & Insight





# TERIMA KASIH!