

JB Link Telco Churn Analysis & Prediction

Final Project Kampus Merdeka



Profile



Dewa Adji Kusuma
Project Lead

Jurusan : Teknik Informatika
Kampus : Institut Teknologi
Telkom Purwokerto



Siti Istiqomah
Visualization

Jurusan : Komputerisasi
Akuntansi
Kampus : Politeknik
Negeri Semarang



Aqiela Fadia Haya
Analysis

Jurusan : Pendidikan
Matematika
Kampus : Universitas
Muhammadiyah Prof. Dr.
Hamka



Nurlina
Analysis

Jurusan : D3-Teknik
Telekomunikasi
Kampus : Politeknik
Negeri Ujung Pandang



Ratu Ryza Amira
Visualization

Jurusan : Teknik Industri
Kampus : ITS

Outline



1

Business Understanding, Memahami bisnis perusahaan dan menetapkan Problem Statement, Objective, Goals, dan Hipotesis.

2

Data Understanding, Memahami scope data yang diberikan dari tiap kolomnya.

3

Data Preprocessing, Melihat elemen data tiap kolom dan melakukan transformasi data.

4

Exploratory Data Analysis, Membuat visualisasi dari data yang sudah di transformasi.

5

Modelling, Melakukan prediksi dari data untuk menemukan pola churn.

6

Kesimpulan dan Rekomendasi, Memberikan kesimpulan dan rekomendasi untuk mengatasi customer churn.

Business Understanding



Background

JB Link adalah sebuah perusahaan yang bergerak dalam bidang telekomunikasi yang berlokasi di California. Pada saat ini perusahaan JB Link sudah menyediakan layanan telepon dan internet kepada pelanggannya lebih dari 1000 kota dan 1600 kode pos. Sudah 6 tahun Perusahaan JB Link berada dan telah berkembang pesat dengan berinvestasi pada infrastruktur untuk membawa jaringan internet dan telepon ke daerah-daerah yang memiliki jangkauan yang buruk atau tidak ada sama sekali.



Problem Statement

- Kategorisasi usia apa dengan churn rate tertinggi?
- Jenis layanan apa yang paling menguntungkan?
- Apakah extra service dapat mempengaruhi churn?

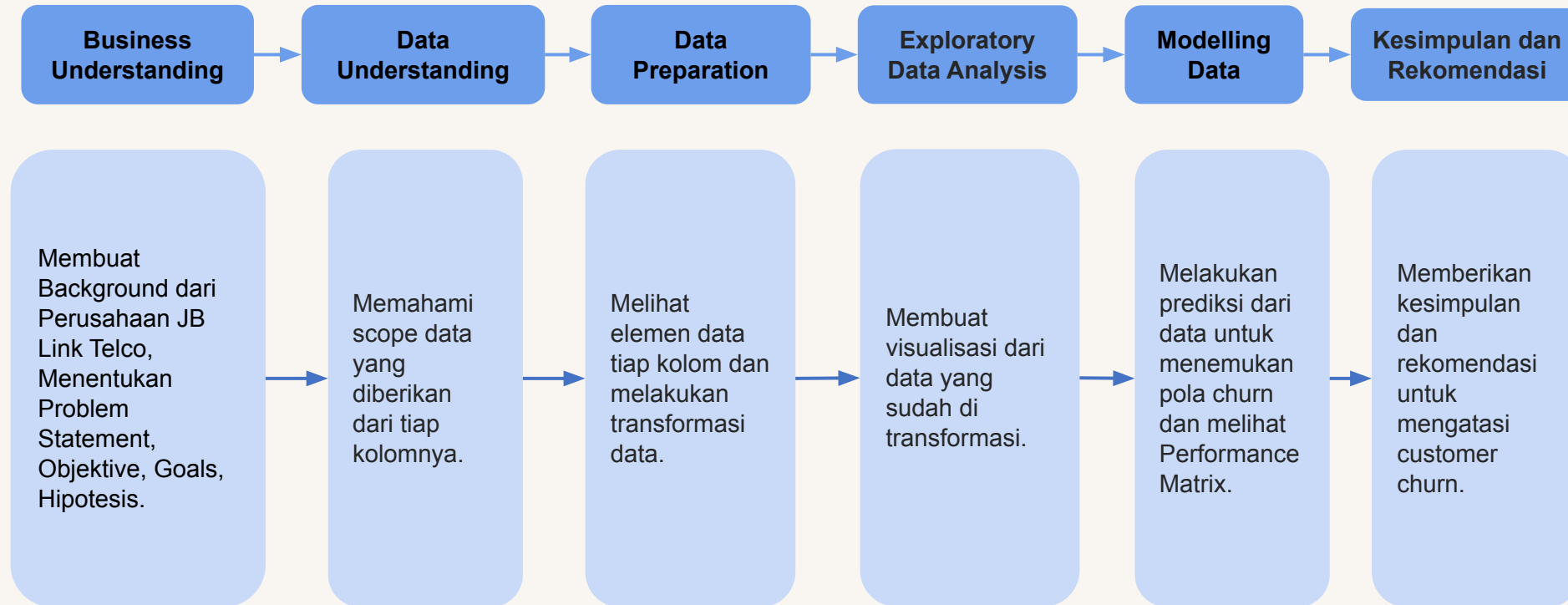
Hipotesis

- Middle citizen merupakan kelompok usia yang paling tinggi persentase terjadinya churn karena jumlah populasinya yang tinggi.
- Phone Service adalah layanan yang paling menguntungkan karena merupakan business core perusahaan JB Link Telco.
- Extra Service tidak dapat mempengaruhi churn karena tidak terlihat pola churnnya.

Goals

Menemukan pola dari customer churn, dan memberikan rekomendasi untuk mengurangi customer churn.

Flow / Metodologi



Data Understanding



Column	Deskripsi
Customer ID [Nominal]	Unique customer id
Referred a Friend [Nominal] [Biner]	Pernah kasih referral code ke teman/keluarga
Gender [Nominal]	Jenis kelamin customer
Under 30 [Nominal] [Biner]	Apakah umur customer dibawah 30, atau tidak
Senior Citizen [Nominal] [Biner]	Jika umur customer diatas 65 tahun
Married [Nominal] [Biner]	Customer sudah menikah, atau belum
Dependents [Nominal] [Biner]	jika pelanggan tinggal dengan tanggungan: Ya, Tidak. Tanggungan bisa anak-anak, orang tua, dll
Offers [Nominal]	Marketing offer terakhir yang diterima
Phone Service [Nominal] [Biner]	Customer berlangganan layanan telepon perusahaan atau tidak
Multiple Lines [Nominal] [Biner]	Customer berlangganan beberapa saluran telepon
Internet Service [Nominal] [Biner]	Customer berlangganan internet perusahaan atau tidak
Internet Type [Nominal]	Tipe layanan internet jika berlangganan
Online Security [Nominal] [Biner]	Customer berlangganan layanan keamanan online tambahan dari perusahaan atau tidak
Online Backup [Nominal] [Biner]	Customer berlangganan layanan Online Backup dari perusahaan atau tidak
Device Protection Plan [Nominal] [Biner]	Customer berlangganan layanan paket Device Protection untuk alat internet mereka dari perusahaan atau tidak

Premium Tech Support [Nominal] [Biner]	jika pelanggan berlangganan paket dukungan teknis tambahan dari perusahaan dengan waktu tunggu yang berkurang
Streaming TV [Nominal] [Biner]	Jika customer menggunakan layanan internet untuk streaming program TV dari penyedia pihak ke3
Streaming Movies [Nominal] [Biner]	Jika customer menggunakan layanan internet untuk streaming film dari penyedia pihak ke3
Streaming Music [Nominal] [Biner]	Jika customer menggunakan layanan internet untuk streaming music dari penyedia pihak ke3
Unlimited Data [Nominal] [Biner]	Customer sudah bayar bulanan tambahan untuk unlimited download/upload
Contract [Nominal]	Jenis langganan customer saat ini
Paperless Billing [Nominal] [Biner]	Customer sudah memilih billing tanpa kertas atau tidak
Payment Method [Nominal]	Metode pembayaran yang digunakan customer
City [Nominal]	kota customer tinggal
Churn Category [Nominal]	Kategori Alasan customer churn
Churn Reason [Nominal]	Alasan khusus pelanggan untuk meninggalkan perusahaan

Deskripsi :

- 26 kolom
- Tipe data kolom Nominal
- Beberapa diantaranya bersifat Biner

Data Understanding



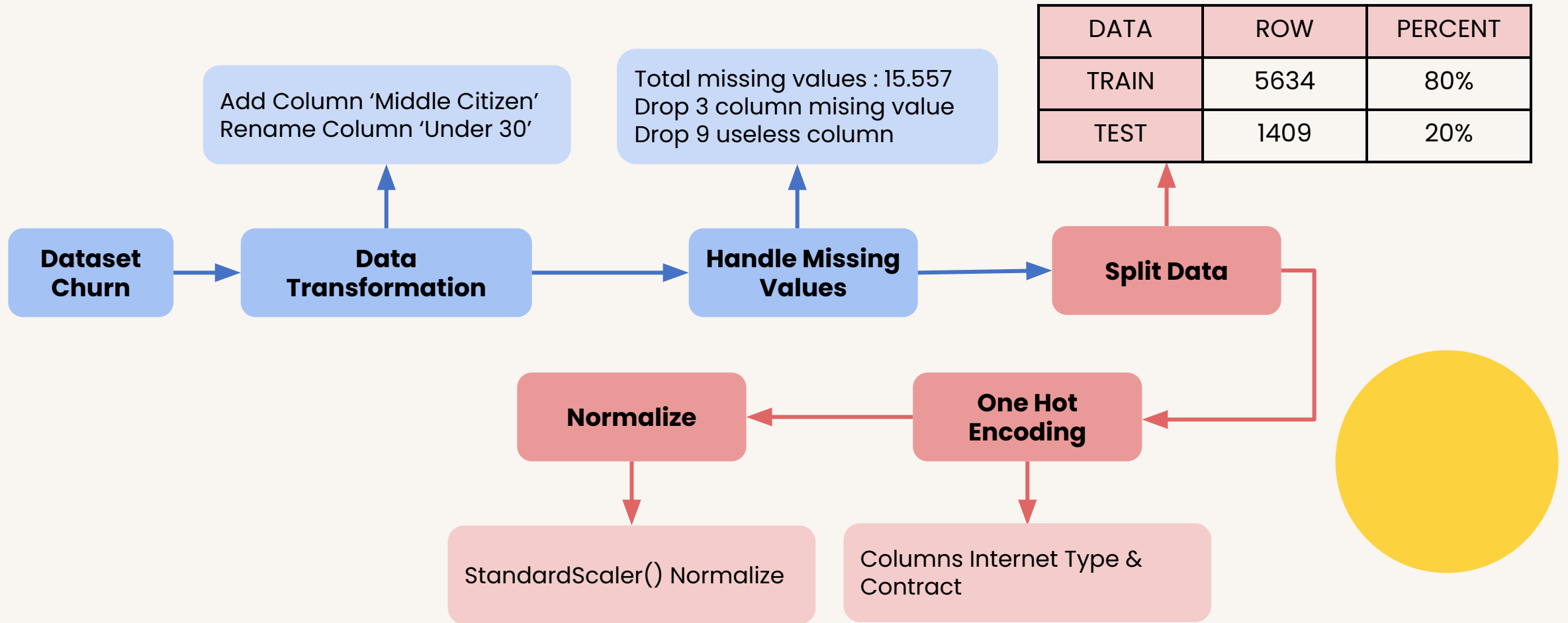
Number of Referrals [Numerik]	Jumlah referral code yang dishare
Tenure in Months [Numerik]	Menunjukkan jumlah total bulan pelanggan telah bersama perusahaan pada akhir kuartal yang ditentukan di atas
Avg Monthly Long Distance Charges [Numerik]	Biaya jarak rata-rata, dihitung hingga akhir kuartal
Avg Monthly GB Download [Numerik]	Volume download(GB) rata-rata, dihitung hingga akhir kuartal
Monthly Charge [Numerik]	Total tagihan bulanan saat ini dari semua layanan mereka dari perusahaan
Total Regular Charges [Numerik]	Total tagihan reguler customer tanpa layanan tambahan
Total Refunds [Numerik]	Total pengembalian uang customer
Total Extra Data Charges [Numerik]	total biaya pelanggan untuk unduhan data tambahan di atas yang ditentukan dalam paket mereka
Total Long Distance Charges [Numerik]	total biaya pelanggan untuk jarak jauh di atas yang ditentukan dalam paket mereka
Age [Numerik]	Umur customer
Number of Dependents [Numerik]	Jumlah tanggungan
Zip Code [Numerik]	Kode post lokasi customer tinggal
Latitude [Numerik]	
Longitude [Numerik]	

Population [Numerik]	Perkiraan populasi saat ini untuk seluruh area Kode Pos
Churn Value [Numerik] [Biner]	1 untuk pelanggan meninggalkan perusahaan kuartal ini. 0 untuk pelanggan tetap bersama perusahaan
CLTV [Numerik]	Nilai Seumur Hidup Pelanggan. CLTV yang diprediksi dihitung menggunakan rumus perusahaan dan data yang ada. Semakin tinggi nilainya, semakin berharga pelanggan
Total Customer Svc Requests [Numerik]	Berapa kali pelanggan menghubungi layanan pelanggan dalam kuartal terakhir
Product/Service Issues Reported [Numerik]	Berapa kali pelanggan melaporkan masalah dengan produk atau layanan dalam kuartal terakhir
Customer Satisfaction [Numerik]	Peringkat kepuasan pelanggan perusahaan secara keseluruhan dari 1 (Sangat Tidak Puas) hingga 5 (Sangat Puas) yang dikumpulkan berdasarkan permintaan layanan pelanggan

Deskripsi :

- 20 kolom
- Tipe data kolom Numerik
- Satu kolom diantaranya bersifat Biner

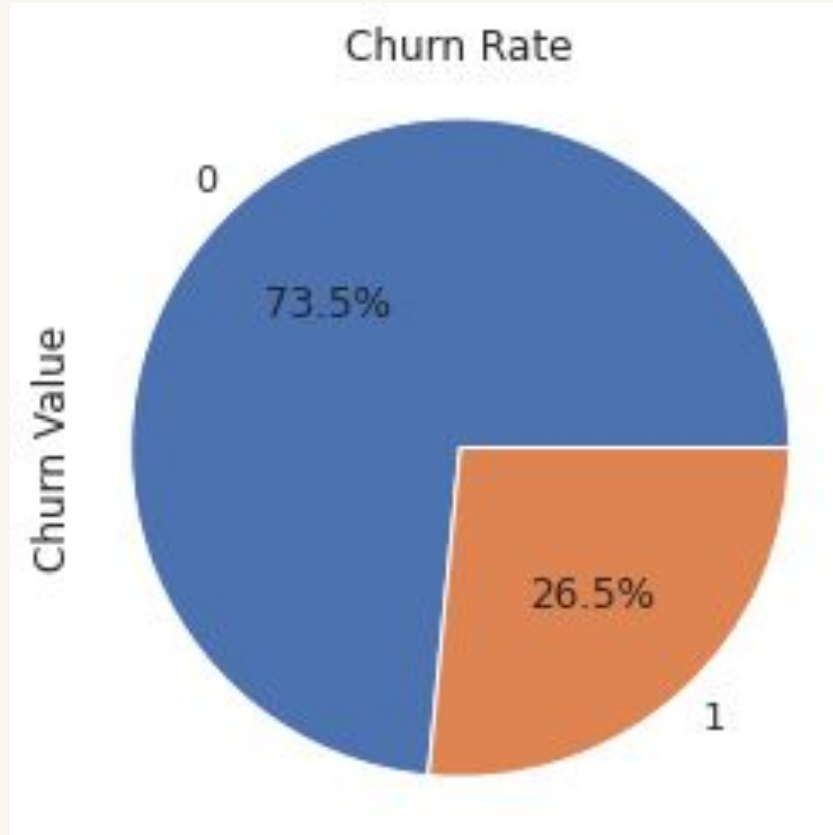
Data Preparation & Preprocessing



Exploratory Data Analysis

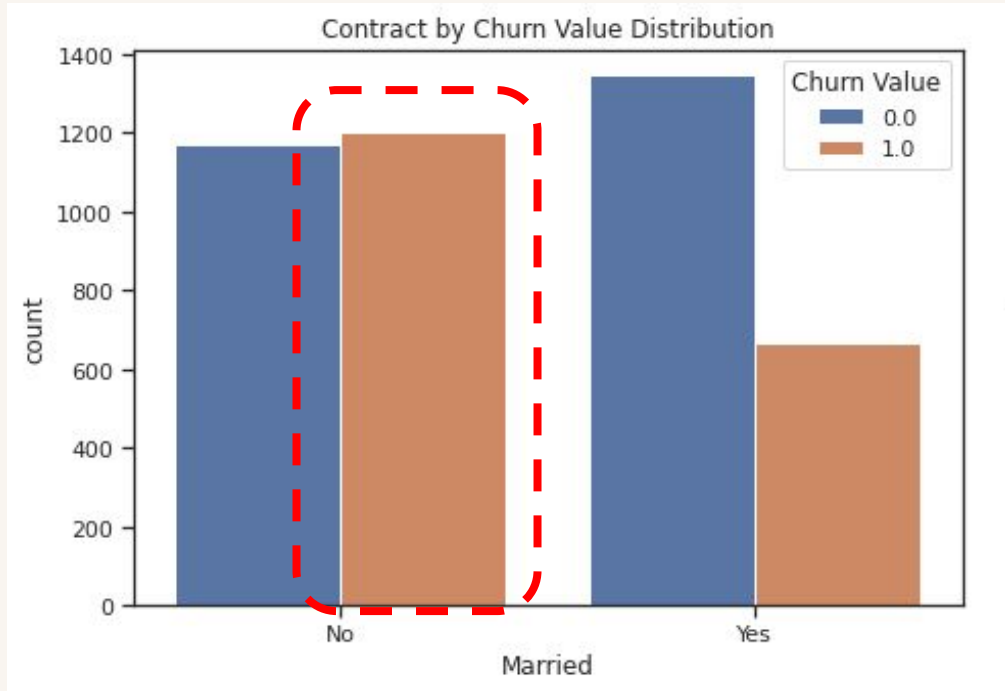


Churn Value Distribution

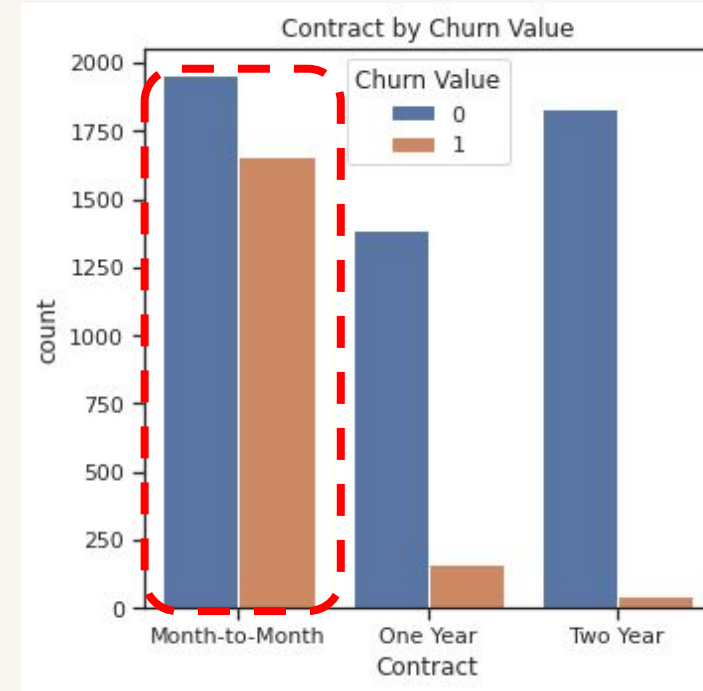


Churn rate	
1	0
1869	5174

Churn distribution by Married & Contract

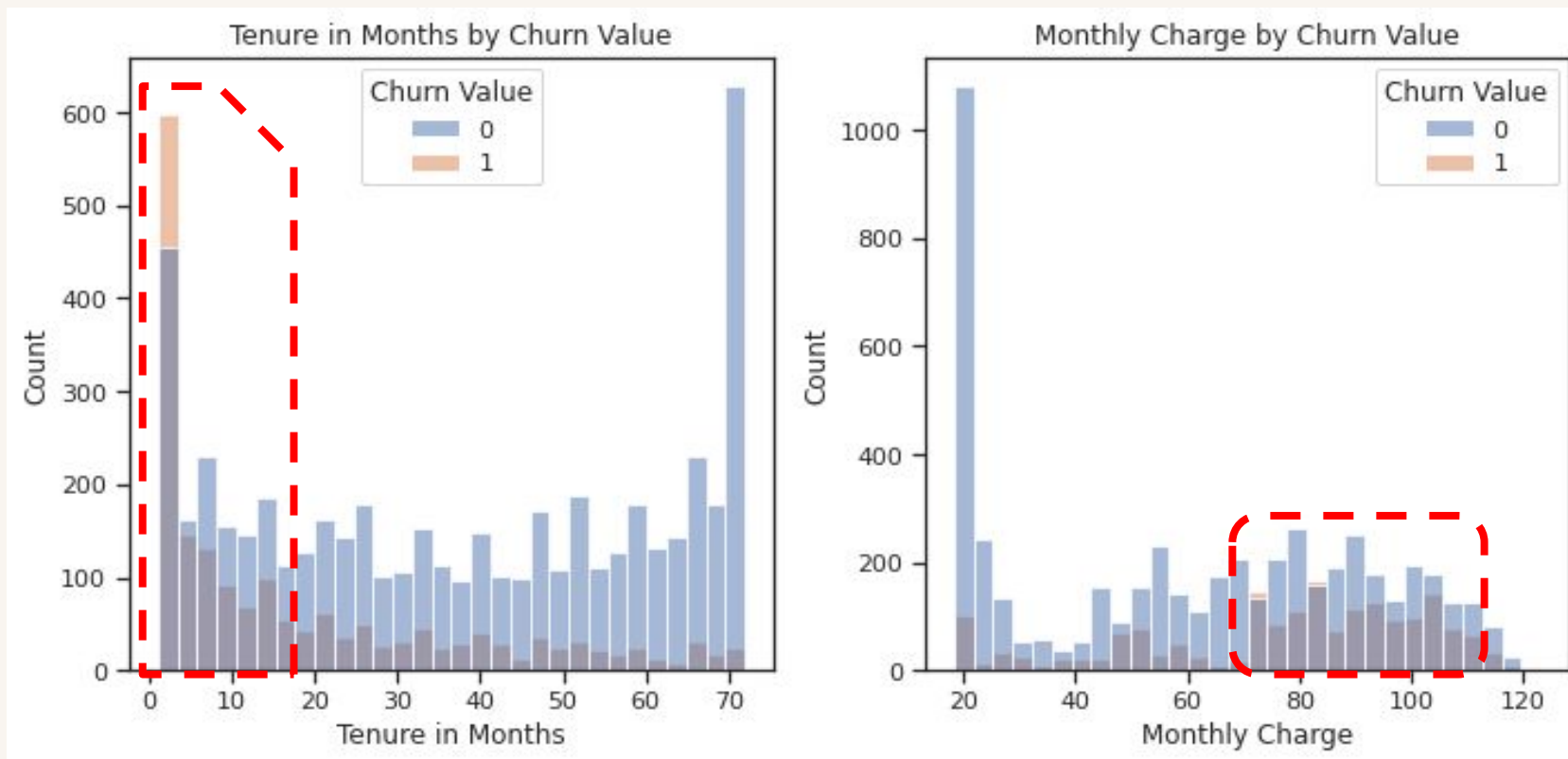


Customer yang **belum menikah**, memiliki potensi Churn yang lebih tinggi dibanding yang sudah menikah.



Customer dengan Contract **Month-to-Month**, memiliki potensi Churn lebih tinggi dibanding Contract One Year atau Two Year.

Churn distribution by Tenure in Months & Monthly Charge

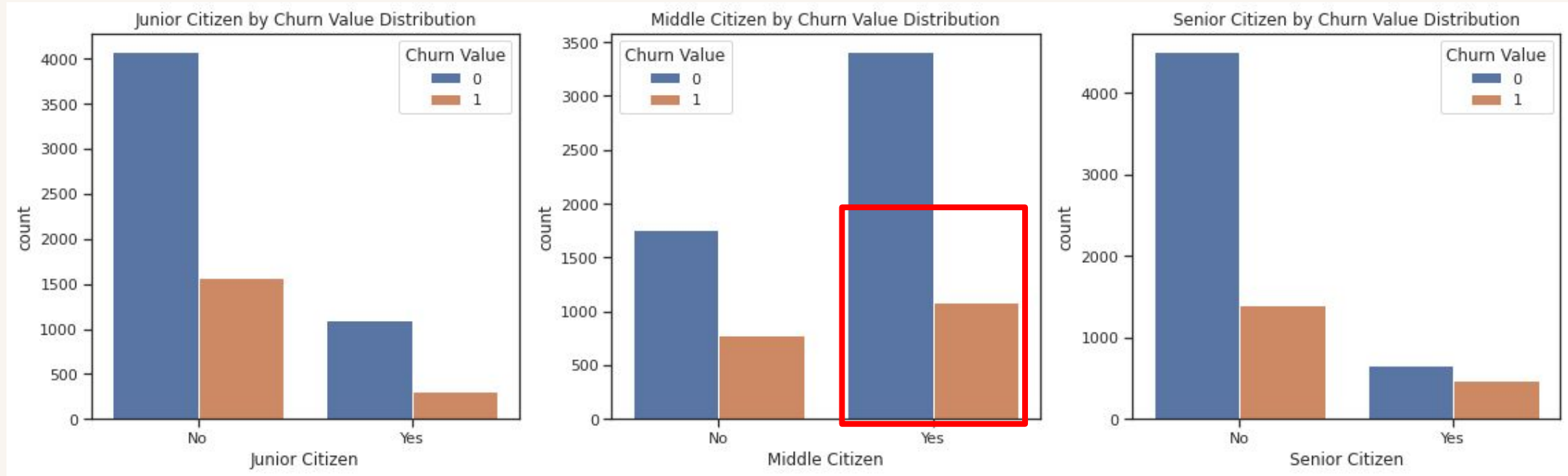


Customer dengan Tenure **dibawah 12 bulan**, cenderung terjadi Churn yang lebih tinggi dibandingkan Tenure diatas 12 bulan.

Customer dengan Monthly Charge **diantara \$80-100**, memiliki potensi terjadinya Churn yang tinggi.

Answering the Problem Statement 1

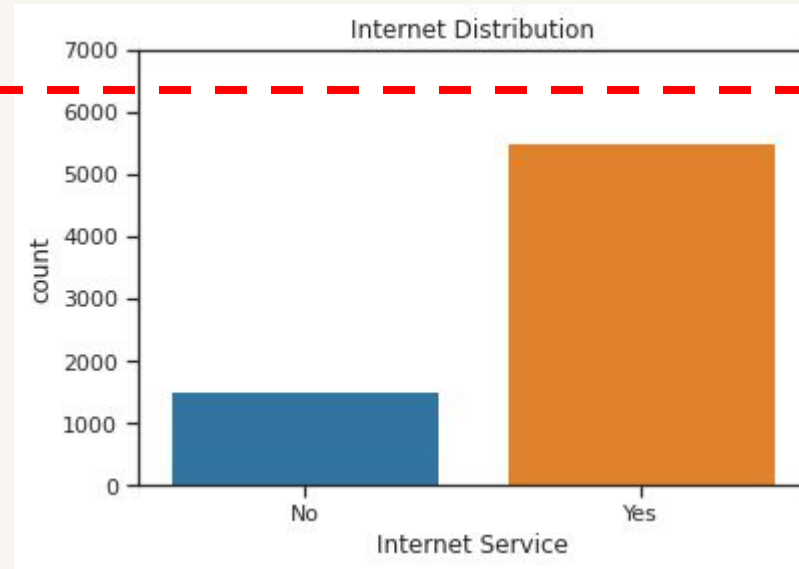
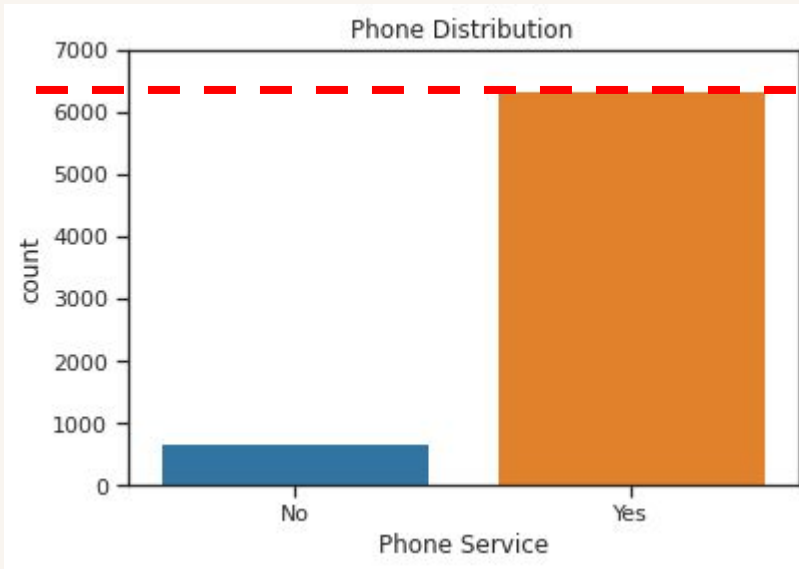
Kategorisasi usia apa dengan churn rate tertinggi?



Customer dengan **usia antara 30–65 tahun** memiliki kemungkinan paling tinggi terjadi Churn.

Answering the Problem Statement 2

Jenis layanan apa yang paling menguntungkan?

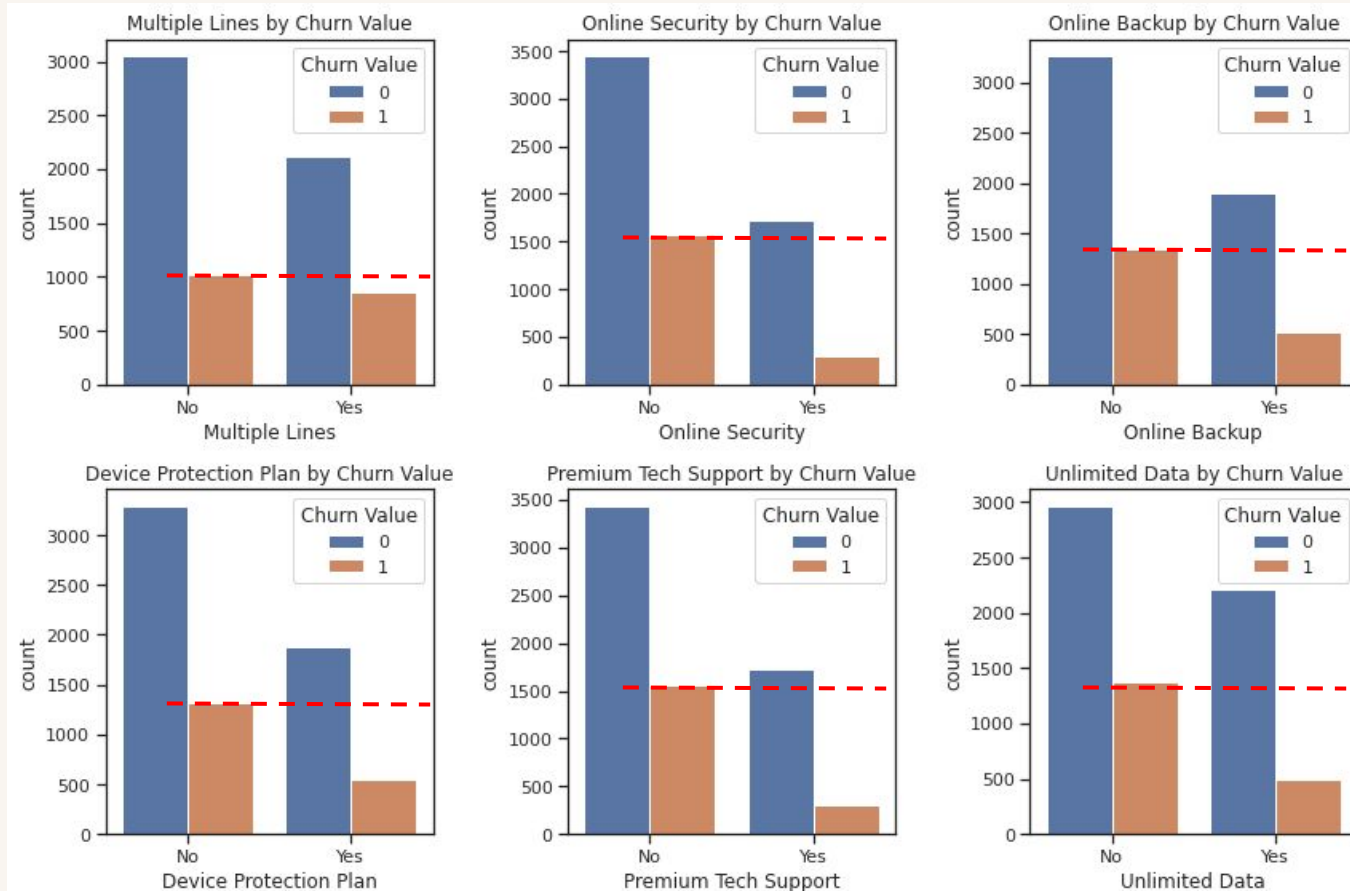


	Yes	No
Phone Service	6361	682
Internet Service	5517	1526

Layanan yang paling **menguntungkan ialah Phone Service**, karena Customer lebih banyak menggunakan layanan Phone Service Dibanding Internet Service.

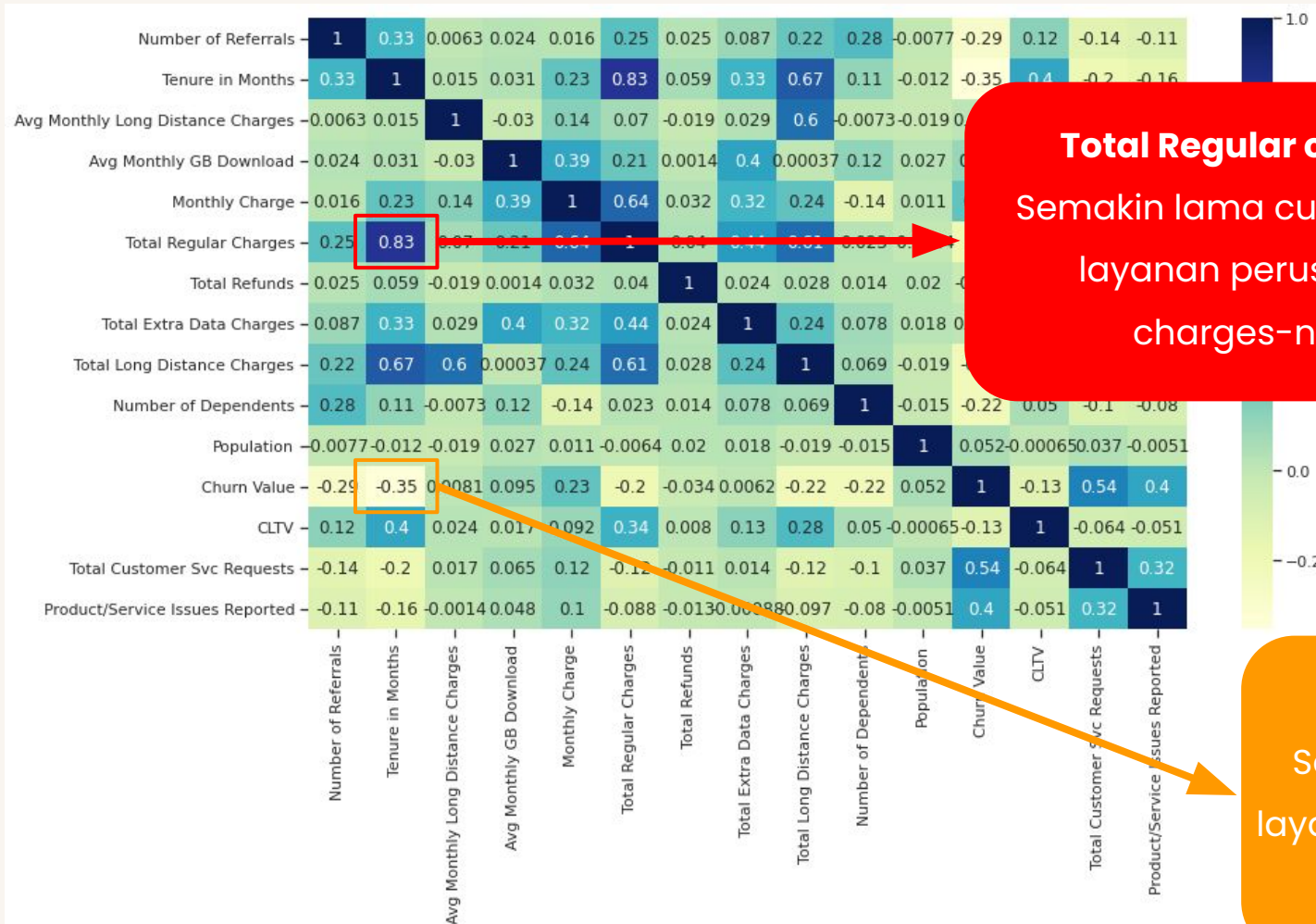
Answering the Problem Statement 3

Apakah extra service dapat mempengaruhi churn?



Berdasarkan hasil analisis, kelompok layanan Extra Service yang diberikan oleh perusahaan ternyata **tidak mempengaruhi churn**.

Heatmap Correlation



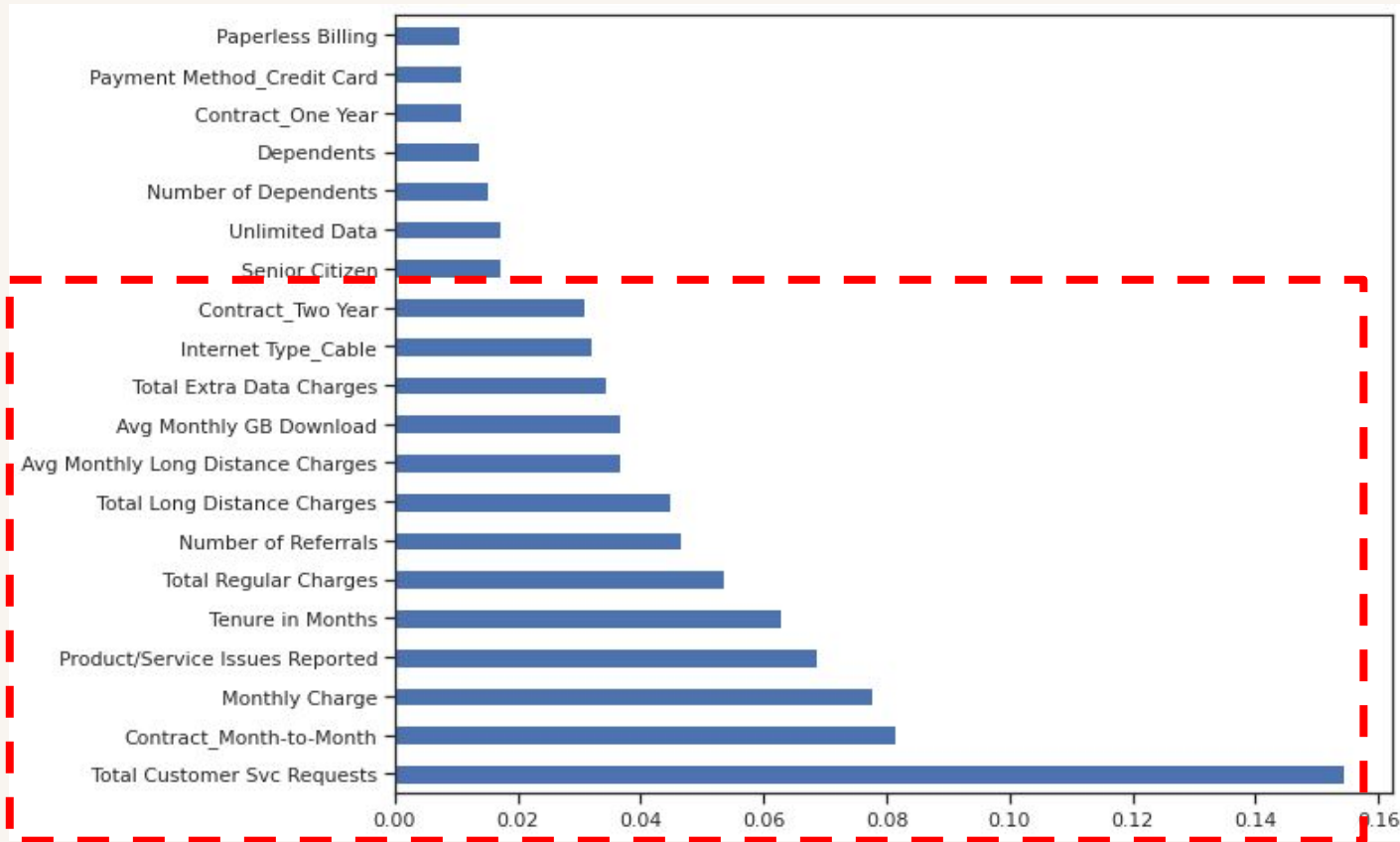
Total Regular charges & Tenure in Months

Semakin lama customer berlangganan dengan layanan perusahaan, maka total regular charges-nya akan semakin tinggi.

Churn Value & Tenure in Months

Semakin lama customer berlangganan dengan layanan perusahaan, maka tidak akan terjadi churn pada customer tersebut.

Modelling



Menggunakan Feature Importance dari semua feature yang ada pada dataset, dipilih 13 Feature teratas karena skor yang lebih signifikan sangat berpengaruh terhadap hasil prediksi Customer Churn.

Hasil Modelling

MODEL	RECALL		F1-SCORE		ACCURACY		PRECISION	
	TRAIN	TEST	TRAIN	TEST	TRAIN	TEST	TRAIN	TEST
K Nearest Neighbor	0.75	0.71	0.80	0.77	90.2	88.5	0.87	0.83
Decision Tree	0.59	0.61	0.70	0.70	86.81	86.37	0.87	0.83
Support Vector Machine	0.74	0.73	0.80	0.80	90.34	90.13	0.88	0.88
Logistic Regression	0.88	0.87	0.76	0.75	85.46	84.74	0.67	0.66
Gaussian NB	0.89	0.88	0.69	0.68	78.66	77.85	0.56	0.55
Random forest Classifier	1.00	0.75	1.00	0.79	100.0	89.63	1.00	0.84
Gradient Boosting Classifier	0.81	0.75	0.86	0.81	92.72	90.56	0.91	0.87

Berdasarkan hasil modelling kami, Logistic Regression dapat memprediksi **Churn** cukup baik dengan **Recall sebesar 87%**.

Logistic Regression juga dipilih dengan pertimbangan F1-Score yang cukup baik sebesar 75%

	Predicted Not Churn	Predicted Churn
Actual Not Churn	868 True Negative(TN)	168 False Positive(FP)
Actual Churn	47 False Negative(FN)	326 True Positive(TP)

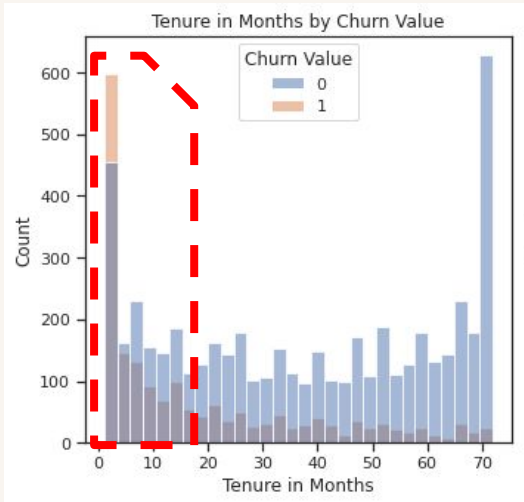
WHY USING RECALL ?

Tujuannya untuk melihat kemampuan model dalam menemukan **Churn** sebanyak mungkin.

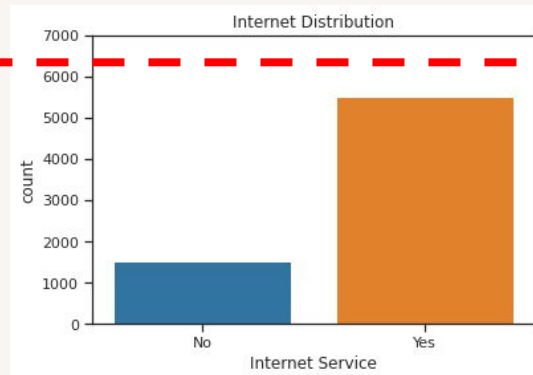
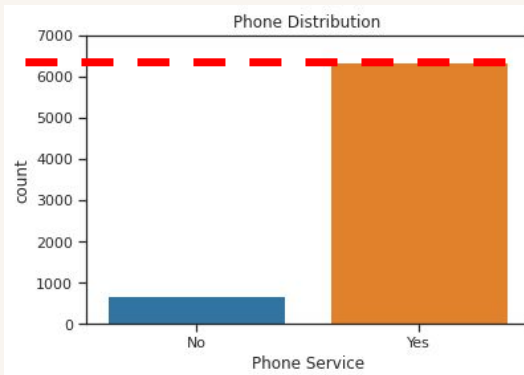
SECONDARY MATRIX : F1-SCORE

Untuk membandingkan rata-rata precision dan recall karena bobot False Negatif dan False Positif tidak mendekati(*symmetric*).

Recommendation

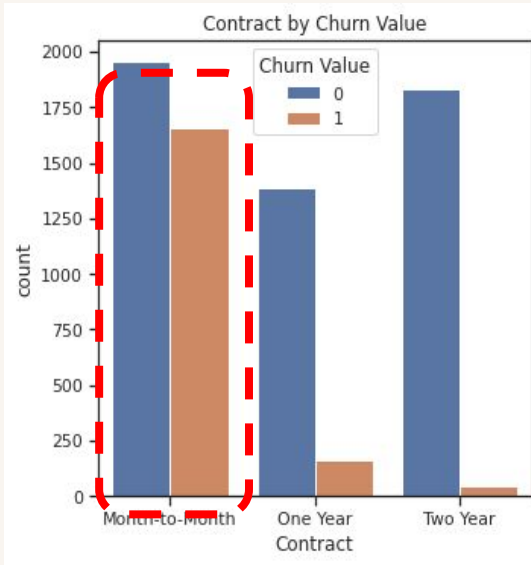


Memberikan special treatment kepada customer dengan **Tenure <12 bulan** seperti dicount 30% untuk langganan 1 tahun pertama untuk menjaga customer tetap berlangganan pada layanan perusahaan



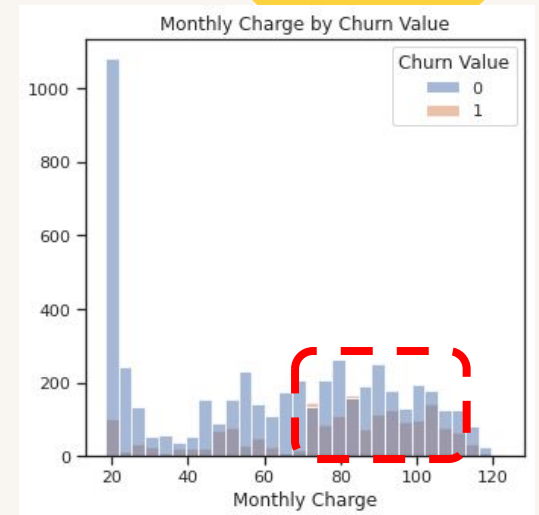
Meningkatkan layanan pada **Internet Service** seperti peningkatkan pada kualitas jaringan untuk menarik Customer berlangganan, serta memberi **bonus Device Protection Plan** jika Customer berlangganan pada Internet Type Fiber Optic.

Recommendation



Memberikan penawaran kepada customer dengan tipe **Contract Month-to-Month** untuk beralih ke **One-Year Contract** dengan bonus salah satu Extra service pilihan Customer untuk menghindari Churn.

Menjadikan Customer dengan **Monthly Charge diatas \$70** menjadi **High Value Customer**, special treatment yang didapatkan seperti respon cepat terhadap issue yang dilaporkan atau menjadi Customer pertama yang ditawarkan promo Extra Service agar Customer tetap royal kepada Perusahaan.



Meningkatkan responsibilitas terhadap laporan issue dari Customer karena **Customer Svc Requests** sangat mempengaruhi Churn.



Thank you!