

Dashboard Analisis Penyebab Churn dan Rekomendasi Aksi untuk Perusahaan Telekomunikasi



Laporan ini Disusun untuk Memenuhi Tugas Akhir Program MSIB Kampus
Merdeka Batch 3
Studi Independen Junior Data Scientist - VINIX

Disusun oleh:
Kelompok 42
[Ryan Besto Saragih]
[Ulfa Dwi Ariani]
[Hasna Aisy Yumna]
[Wisnu Hafid Firdaus Oktrobian]

PT VINIX SEVEN AURUM

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Industri telekomunikasi modern ditandai oleh tingkat persaingan yang sangat tinggi dan pasar yang jenuh. Dalam kondisi seperti ini, kemampuan sebuah perusahaan untuk mempertahankan pelanggan yang sudah ada (retensi pelanggan) menjadi faktor krusial yang menentukan keberlanjutan bisnis. Kehilangan pelanggan, atau yang lebih dikenal dengan istilah customer churn, tidak hanya berdampak pada hilangnya pendapatan berulang (recurring revenue), tetapi juga menuntut biaya yang signifikan untuk mengakuisisi pelanggan baru sebagai penggantinya. Biaya akuisisi pelanggan baru seringkali jauh lebih mahal dibandingkan biaya untuk mempertahankan pelanggan yang loyal.

Menghadapi tantangan ini, PT Phoenix 7 Aurum sebagai stakeholder dalam proyek ini menyadari pentingnya strategi retensi yang proaktif dan berbasis data. Pengambilan keputusan yang hanya berdasarkan intuisi atau tren pasar secara umum tidak lagi memadai. Diperlukan sebuah analisis mendalam terhadap data historis pelanggan untuk mengidentifikasi pola, perilaku, serta karakteristik yang membedakan antara pelanggan yang loyal dengan pelanggan yang berpotensi churn. Proyek ini, yang merupakan bagian dari program MSIB Kampus Merdeka Batch 3 bersama PT VINIX SEVEN AURUM, bertujuan untuk menjawab tantangan tersebut dengan melakukan analisis data end-to-end, mulai dari eksplorasi data hingga penyajian rekomendasi strategis yang dapat ditindaklanjuti.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka dirumuskan serangkaian pertanyaan analisis yang menjadi fokus utama dalam proyek ini:

- Seberapa besar tingkat churn secara keseluruhan, dan bagaimana karakteristik dasar pelanggan yang churn dari segi rata-rata tenure (lama berlangganan) dan Monthly Charges (tagihan bulanan)?
- Seberapa signifikan pengaruh tipe Contract terhadap churn? Secara spesifik, berapa besar tingkat churn pada pelanggan Month-to-month dibandingkan kontrak tahunan, dan seberapa besar porsi mereka dari total pelanggan yang churn?
- Seberapa efektif layanan pendukung seperti Tech Support dan Online Security dalam menurunkan tingkat churn saat dimiliki oleh pelanggan?
- Apakah churn rate menurun secara signifikan seiring dengan bertambahnya jumlah layanan tambahan yang dimiliki oleh seorang pelanggan?

- Berapa estimasi kerugian pendapatan bulanan (Monthly Recurring Revenue) akibat churn, dan berapa potensi penurunannya jika kita berhasil memigrasikan sebagian pelanggan Month-to-month ke kontrak tahunan?
- Di antara pelanggan berisiko tinggi (kontrak bulanan, tenure pendek), siapakah segmen yang paling bernilai tinggi (high-value) untuk diprioritaskan dalam program retensi?

1.3. Tujuan Proyek

Selaras dengan rumusan masalah yang telah ditetapkan, tujuan utama dari pelaksanaan proyek analisis data ini adalah sebagai berikut:

- Mengidentifikasi dan mengukur faktor-faktor utama yang secara signifikan mempengaruhi kecenderungan churn pelanggan di perusahaan.
- Mengkuantifikasi dampak finansial dari customer churn dalam bentuk estimasi Monthly Recurring Revenue (MRR) yang hilang untuk memberikan gambaran skala masalah kepada stakeholder.
- Mengembangkan sebuah dashboard interaktif menggunakan Google (Looker) Studio yang menyajikan temuan-temuan kunci secara visual dan dapat digunakan sebagai alat bantu monitoring oleh tim terkait.
- Memberikan rekomendasi strategis yang terukur dan dapat ditindaklanjuti (actionable) kepada stakeholder untuk membantu menyusun program retensi pelanggan yang lebih efektif dan efisien.

BAB II

METODOLOGI DAN PERSIAPAN DATA

2.1. Kerangka Kerja Proyek: CRISP-DM

Untuk memastikan proyek ini berjalan secara sistematis dan terstruktur, kami mengadopsi metodologi CRISP-DM (Cross-Industry Standard Process for Data Mining). CRISP-DM adalah kerangka kerja standar industri yang membagi siklus hidup proyek analisis data menjadi enam fase utama. Pendekatan ini memastikan bahwa setiap langkah, mulai dari pemahaman masalah bisnis hingga penyajian hasil, dilakukan secara logis dan menyeluruh.

Adapun keenam fase tersebut yang diterapkan dalam proyek ini adalah:

- **Business Understanding:** Memahami tujuan dan persyaratan proyek dari sudut pandang bisnis, kemudian merumuskannya ke dalam definisi masalah analisis data.
- **Data Understanding:** Melakukan eksplorasi data awal untuk membiasakan diri dengan data, mengidentifikasi kualitas data, dan menemukan wawasan pertama.
- **Data Preparation:** Mencakup semua aktivitas untuk membangun dataset final dari data mentah awal, seperti pembersihan data, penanganan nilai yang hilang, dan rekayasa fitur.
- **Modeling (Analysis):** Pada tahap ini, berbagai teknik analisis diterapkan untuk menemukan pola dan wawasan yang menjawab rumusan masalah.
- **Evaluation:** Mengevaluasi hasil analisis untuk memastikan temuan yang didapat valid dan menjawab tujuan bisnis yang telah ditetapkan di awal.
- **Deployment:** Menyajikan hasil akhir proyek, yang dalam konteks ini berupa laporan analisis, materi presentasi, dan sebuah dashboard interaktif.

2.2. Data Understanding

2.2.1. Deskripsi Sumber Data

Data yang digunakan dalam analisis ini adalah dataset "Telco Customer Churn" yang bersumber dari platform Kaggle. Dataset ini berisi 7.043 baris data yang merepresentasikan pelanggan individual dari sebuah perusahaan telekomunikasi fiktif, dengan 21 kolom yang mendeskripsikan berbagai atribut demografis, layanan yang digunakan, informasi akun, dan status churn setiap pelanggan.

2.2.2. Struktur dan Karakteristik Data

Langkah awal dalam memahami data adalah dengan memeriksa struktur dan tipe data dari setiap kolom. Hal ini dilakukan untuk mengidentifikasi potensi masalah, seperti tipe data yang tidak sesuai, yang dapat menghambat proses analisis.

```
Informasi Dataset:  
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 7043 entries, 0 to 7042  
Data columns (total 21 columns):  
 #   Column           Non-Null Count  Dtype     
---  --    
 0   customerID      7043 non-null    object    
 1   gender          7043 non-null    object    
 2   SeniorCitizen   7043 non-null    int64    
 3   Partner          7043 non-null    object    
 4   Dependents      7043 non-null    object    
 5   tenure          7043 non-null    int64    
 6   PhoneService     7043 non-null    object    
 7   MultipleLines    7043 non-null    object    
 8   InternetService  7043 non-null    object    
 9   OnlineSecurity   7043 non-null    object    
 10  OnlineBackup     7043 non-null    object    
 11  DeviceProtection 7043 non-null    object    
 12  TechSupport      7043 non-null    object    
 13  StreamingTV      7043 non-null    object    
 14  StreamingMovies   7043 non-null    object    
 15  Contract          7043 non-null    object    
 16  PaperlessBilling 7043 non-null    object    
 17  PaymentMethod     7043 non-null    object    
 18  MonthlyCharges   7043 non-null    float64   
 19  TotalCharges     7043 non-null    object    
 20  Churn            7043 non-null    object    
 dtypes: float64(1), int64(2), object(18)  
 memory usage: 1.1+ MB
```

Dari gambar di atas, dapat dilihat bahwa dataset terdiri dari tipe data integer, float, dan object. Sebuah temuan penting pada tahap ini adalah kolom TotalCharges (total tagihan) yang seharusnya bersifat numerik, namun terdeteksi sebagai tipe data object. Hal ini mengindikasikan adanya nilai non-numerik dalam kolom tersebut yang perlu ditangani pada tahap persiapan data.

2.3. Data Preparation

Tahap persiapan data adalah fase kritis untuk memastikan kualitas dan keandalan data sebelum dilakukan analisis. Data mentah jarang sekali sempurna, sehingga serangkaian proses pembersihan dan transformasi perlu dilakukan.

2.3.1. Penanganan Tipe Data yang Salah

Sesuai temuan pada tahap sebelumnya, langkah pertama adalah mengubah tipe data kolom TotalCharges dari object menjadi numerik. Proses ini dilakukan menggunakan fungsi pandas.to_numeric. Parameter errors='coerce' digunakan untuk secara otomatis mengubah nilai apa pun yang tidak dapat dikonversi menjadi angka (seperti spasi kosong) menjadi nilai NaN (Not a Number).

2.3.2. Penanganan Nilai Hilang (Missing Values)

Proses konversi pada langkah sebelumnya menghasilkan 11 nilai NaN pada kolom TotalCharges. Investigasi lebih lanjut dilakukan untuk memahami konteks dari nilai yang hilang ini.

```
Jumlah Missing Values Setelah Konversi TotalCharges:  
customerID      0  
gender          0  
SeniorCitizen   0  
Partner         0  
Dependents     0  
tenure          0  
PhoneService    0  
MultipleLines   0  
InternetService 0  
OnlineSecurity  0  
OnlineBackup    0  
DeviceProtection 0  
TechSupport     0  
StreamingTV    0  
StreamingMovies 0  
Contract        0  
PaperlessBilling 0  
PaymentMethod   0  
MonthlyCharges  0  
TotalCharges    11  
Churn           0  
dtype: int64
```

Setelah diselidiki, ditemukan bahwa semua 11 baris data dengan TotalCharges yang hilang memiliki nilai tenure (lama berlangganan) sama

dengan 0 bulan. Hal ini memberikan justifikasi yang kuat bahwa pelanggan tersebut adalah pelanggan yang sangat baru sehingga belum memiliki riwayat tagihan sama sekali. Berdasarkan logika ini, keputusan diambil untuk mengisi nilai NaN tersebut dengan angka 0, karena secara bisnis, total tagihan mereka memang nol.

2.3.3. Rekayasa Fitur (Feature Engineering)

Untuk mempermudah proses analisis dan kalkulasi, beberapa fitur baru dibuat dari kolom yang sudah ada:

- Churn_numeric: Sebuah kolom baru dibuat untuk merepresentasikan status churn secara numerik, di mana nilai 'Yes' diubah menjadi 1 dan 'No' diubah menjadi 0. Ini memungkinkan perhitungan tingkat churn secara langsung dengan menghitung rata-rata (mean) dari kolom ini.
- Num_Additional_Services: Untuk menganalisis pengaruh ekosistem layanan, sebuah kolom baru dibuat untuk menghitung jumlah layanan tambahan (seperti OnlineSecurity, TechSupport, dll.) yang dimiliki oleh setiap pelanggan.

BAB III

HASIL ANALISIS DAN PEMBAHASAN

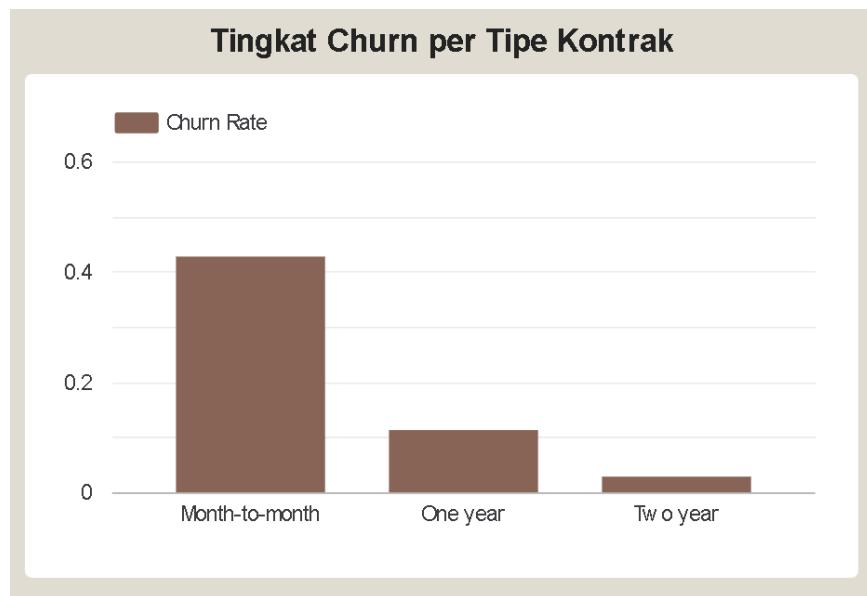
3.1. Gambaran Umum dan Skala Masalah Churn

Langkah pertama dalam analisis adalah mengukur skala dan dampak dari customer churn. Ditemukan bahwa tingkat churn secara keseluruhan berada pada angka 26.54%, yang berarti lebih dari seperempat total pelanggan berhenti berlangganan. Dalam angka absolut, ini setara dengan 1.869 pelanggan yang hilang. Dampak finansial dari churn ini sangat signifikan, dengan estimasi kerugian pendapatan bulanan (Monthly Recurring Revenue - MRR) mencapai **\$139,130.85**.



3.2. Faktor Kontrak sebagai Pendorong Utama Churn

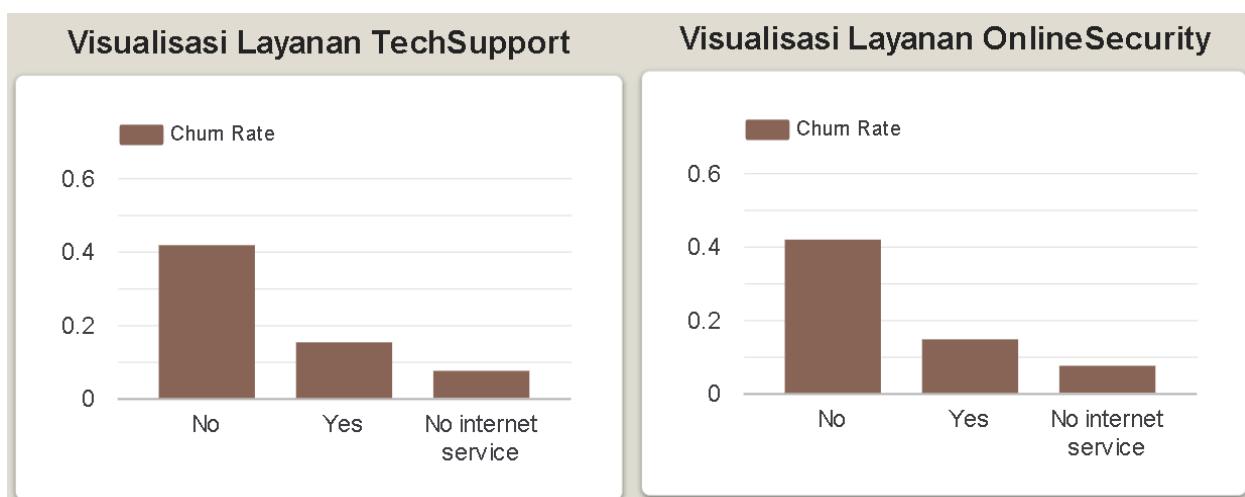
Setelah memahami skala masalah, analisis dilanjutkan untuk mengidentifikasi akar penyebabnya. Ditemukan bahwa faktor tunggal yang paling signifikan dalam mendorong churn adalah tipe kontrak pelanggan. Pelanggan dengan kontrak fleksibel Month-to-month (bulanan) menunjukkan tingkat churn yang sangat tinggi sebesar **42.71%**. Angka ini sangat kontras dengan pelanggan kontrak One year (**11.27%**) dan Two year (**2.83%**) yang jauh lebih loyal.



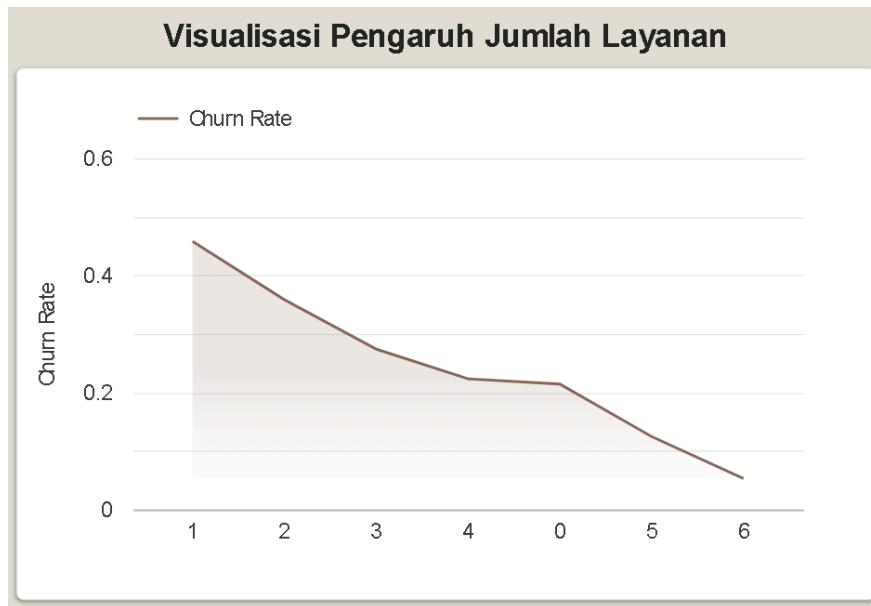
Krusialnya faktor ini diperkuat lebih lanjut saat melihat komposisi dari total pelanggan yang churn. Dari 1.869 pelanggan yang churn, sebanyak 88.6% di antaranya berasal dari segmen pelanggan Month-to-month. Temuan ini menegaskan bahwa strategi retensi harus memprioritaskan penanganan pelanggan dalam segmen kontrak bulanan ini.

3.3. Pengaruh Layanan Dukungan dan Ekosistem Layanan

Selain tipe kontrak, analisis juga menunjukkan bahwa ketersediaan layanan dukungan teknis dan keamanan daring memainkan peran penting dalam menjaga loyalitas pelanggan. Tingkat churn pada pelanggan yang tidak memiliki TechSupport (**41.64%**) dan OnlineSecurity (**41.77%**) hampir tiga kali lipat lebih tinggi dibandingkan pelanggan yang memiliki layanan tersebut (sekitar **15%**).



Secara lebih luas, analisis menunjukkan bahwa semakin banyak layanan tambahan yang diadopsi oleh pelanggan, semakin rendah kecenderungan mereka untuk churn. Hal ini mengindikasikan bahwa semakin dalam seorang pelanggan terintegrasi ke dalam ekosistem layanan perusahaan, semakin tinggi "biaya beralih" (switching cost) yang mereka rasakan, sehingga meningkatkan loyalitas.



3.4. Analisis Dampak Finansial dan Segmentasi Prioritas

Untuk memberikan perspektif solusi yang terukur, sebuah skenario bisnis disimulasikan. Jika perusahaan berhasil meluncurkan program retensi yang dapat menurunkan jumlah pelanggan Month-to-month yang churn sebesar **25%**, maka ada potensi penyelamatan pendapatan bulanan sebesar **\$30,230**.

Namun, untuk efisiensi sumber daya, tidak semua pelanggan berisiko perlu ditargetkan dengan cara yang sama. Oleh karena itu, dilakukan segmentasi untuk mengidentifikasi kelompok pelanggan "High-Value, High-Risk". Segmen ini didefinisikan sebagai pelanggan aktif dengan kontrak Month-to-month, masa berlangganan pendek (≤ 12 bulan), dan memiliki tagihan bulanan di atas rata-rata. Kelompok inilah yang harus menjadi prioritas utama tim retensi karena kehilangan mereka akan memberikan dampak finansial terbesar.

Top High-Value, High-Risk Customers to Contact

	customerID	tenure	MonthlyCharges	InternetService
1.	3292-PBZEJ	11	\$111.40	Fiber optic
2.	5760-IFJOZ	3	\$107.95	Fiber optic
3.	2081-VEYBH	3	\$107.95	Fiber optic
4.	7145-FEJWU	12	\$105.30	Fiber optic
5.	6734-GMPVK	5	\$105.30	Fiber optic
6.	2018-PZKMU	9	\$103.10	Fiber optic
7.	9547-ITEFG	9	\$102.60	Fiber optic
8.	8118-TJAEG	9	\$101.50	Fiber optic
9.	4132-KALRO	4	\$100.85	Fiber optic
10.	0722-SVSFK	7	\$100.40	Fiber optic

Analisis lebih lanjut pada segmen ini mengungkapkan sebuah insight penting: meskipun segmen prioritas ini hanya mewakili sebagian kecil dari total pelanggan aktif, mereka menyumbang porsi yang tidak proporsional dari total pendapatan bulanan yang berisiko hilang. Ini menegaskan bahwa pendekatan yang tertarget pada segmen ini akan memberikan Return on Investment (ROI) yang tertinggi.

BAB IV

REKOMENDASI STRATEGIS

4.1. Prioritas Utama: Meluncurkan Program Migrasi Kontrak Fleksibel

Temuan pada Bab 3.2 secara tegas menunjukkan bahwa pelanggan dengan kontrak Month-to-month adalah sumber utama churn, dengan tingkat churn mencapai 42.71% dan menyumbang 88.6% dari total pelanggan yang hilang.

Rekomendasi Aksi:

Meluncurkan kampanye pemasaran dan retensi yang proaktif dengan tujuan utama memigrasikan pelanggan dari kontrak Month-to-month ke kontrak One year atau Two year. Beberapa inisiatif yang dapat dilakukan antara lain:

- **Penawaran Incentif Finansial:** Memberikan diskon khusus pada tagihan bulanan (misalnya, potongan harga 10-15%) bagi pelanggan yang bersedia meningkatkan komitmen kontraknya ke 1 tahun.
- **Penawaran Nilai Tambah (Value Add):** Menawarkan benefit non-finansial seperti gratis satu layanan tambahan (misalnya, OnlineSecurity) selama 3 bulan pertama jika pelanggan beralih ke kontrak tahunan.
- **Komunikasi Tersegmentasi:** Mengirimkan penawaran ini melalui email atau notifikasi aplikasi yang dipersonalisasi, terutama kepada pelanggan Month-to-month yang telah berlangganan lebih dari 6 bulan, karena mereka sudah menunjukkan tingkat loyalitas awal.

4.2. Peningkatan Nilai: Mengembangkan Strategi Bundling Layanan Dukungan

Analisis pada Bab 3.3 membuktikan bahwa pelanggan yang memiliki layanan dukungan seperti TechSupport dan OnlineSecurity memiliki tingkat churn yang jauh lebih rendah. Selain itu, loyalitas meningkat seiring dengan bertambahnya jumlah layanan yang dimiliki.

Rekomendasi Aksi:

Memperkuat strategi cross-selling dan bundling untuk meningkatkan adopsi layanan tambahan, khususnya layanan dukungan, sehingga pelanggan lebih terintegrasi dalam ekosistem perusahaan.

- **Paket Selamat Datang (Welcome Package):** Saat pelanggan baru mendaftar, terutama yang memilih kontrak Month-to-month, tawarkan paket bundling yang sudah mencakup TechSupport atau OnlineSecurity dengan harga yang lebih menarik dibandingkan membeli secara terpisah.
- **Uji Coba Gratis (Free Trial):** Memberikan periode uji coba gratis selama 1-2 bulan untuk layanan dukungan kepada pelanggan yang belum memilikinya.

Setelah masa uji coba berakhir, tawarkan harga berlangganan khusus untuk melanjutkan layanan.

4.3. Efisiensi Sumber Daya: Fokus pada Segmen Prioritas (High-Value, High-Risk)

Analisis segmentasi pada Bab 3.4 berhasil mengidentifikasi kelompok pelanggan yang paling kritis: mereka yang bernilai tinggi bagi perusahaan (tagihan bulanan besar) namun juga sangat berisiko untuk churn (kontrak bulanan, tenure pendek).

Rekomendasi Aksi:

Mengalokasikan sumber daya tim retensi secara efisien dengan memprioritaskan penanganan segmen "High-Value, High-Risk" yang telah teridentifikasi dalam dashboard.

- **Pembentukan Tim Retensi VIP:** Menugaskan anggota tim retensi khusus (atau dedicated account manager) untuk secara personal menghubungi pelanggan yang masuk dalam daftar teratas segmen ini. Pendekatan personal ini dapat meningkatkan peluang keberhasilan retensi.
- **Penawaran Retensi Eksklusif:** Memberdayakan tim ini untuk memberikan penawaran "di luar skrip" yang lebih superior, seperti diskon personal yang lebih besar, upgrade layanan gratis, atau penyesuaian paket sesuai kebutuhan pelanggan.
- **Pemanfaatan Dashboard sebagai Alat Operasional:** Tim retensi harus menggunakan dashboard yang telah dibuat sebagai alat kerja harian untuk memonitor daftar pelanggan prioritas ini secara dinamis.

BAB V

KESIMPULAN

Proyek analisis customer churn ini telah berhasil mengidentifikasi akar permasalahan, mengkuantifikasi dampak bisnis, dan merumuskan strategi retensi berbasis data. Berdasarkan analisis yang telah dilakukan terhadap dataset pelanggan telekomunikasi, dapat ditarik beberapa kesimpulan kunci. Tingkat churn perusahaan sebesar 26.54% merupakan angka yang signifikan dan didorong secara dominan oleh pelanggan dengan kontrak Month-to-month, yang memiliki tingkat churn empat kali lebih tinggi dibandingkan pelanggan kontrak tahunan dan menyumbang 88.6% dari total pelanggan yang hilang. Selain itu, ditemukan bahwa adopsi layanan tambahan, khususnya layanan dukungan seperti TechSupport dan OnlineSecurity, berkorelasi kuat dengan peningkatan loyalitas pelanggan.

Menjawab temuan tersebut, laporan ini mengajukan tiga pilar rekomendasi strategis yang dapat ditindaklanjuti. Pertama, meluncurkan program migrasi kontrak dengan insentif untuk mendorong pelanggan Month-to-month beralih ke komitmen jangka panjang. Kedua, memperkuat strategi bundling untuk meningkatkan integrasi pelanggan ke dalam ekosistem layanan perusahaan. Ketiga, dan yang paling mendesak, adalah mengimplementasikan program retensi VIP yang secara efisien memfokuskan sumber daya pada segmen pelanggan "High-Value, High-Risk" untuk memberikan ROI maksimal.

Sebagai hasil akhir, proyek ini tidak hanya menghasilkan serangkaian insight statis, tetapi juga sebuah dashboard interaktif di Looker Studio yang dapat berfungsi sebagai alat monitoring berkelanjutan. Dengan mengimplementasikan rekomendasi yang didasarkan pada bukti data ini, perusahaan diharapkan dapat secara efektif menekan tingkat churn, meningkatkan loyalitas pelanggan, dan pada akhirnya mengamankan pendapatan jangka panjang secara lebih berkelanjutan.