

TUGAS MISI 1 PERGUDANGAN DATA
**“Perancangan *Data Warehouse* untuk Analisis Customer Churn
pada Industri Telekomunikasi”**



KELOMPOK 8 DW RA

ANGGOTA KELOMPOK :

EKSANTY F SUGMA ISLAMIATY	122450001
RESIDEN NUSANTARA R M	122450080
AISYAH TIARA PRATIWI	121450074
UKASYAH MUNTAHA	122450028
RENDRA EKA PRAYOGA	122450112

PROGRAM STUDI SAINS DATA
FAKULTAS SAINS
INSTITUT TEKNOLOGI SUMATERA
LAMPUNG SELATAN
2025

A. Profil Industri & Masalah Bisnis

Industri telekomunikasi di California merupakan sektor yang berfokus pada layanan pelanggan dan kekuatan perusahaan telekomunikasi untuk mempertahankan pelanggan yang sudah ada agar terus menggunakan layanan mereka dalam jangka waktu yang lama (retensi). Perusahaan ini menawarkan berbagai layanan seperti telepon, internet, streaming, serta dukungan teknis premium, dan mengalami tantangan dalam mempertahankan pelanggan akibat churn (berhentinya pelanggan dari layanan). Masalah utama yang dihadapi adalah kurangnya pemahaman mendalam terhadap faktor-faktor yang menyebabkan pelanggan berhenti berlangganan. Hal ini berdampak langsung pada pendapatan dan pertumbuhan bisnis. Oleh karena itu, perusahaan membutuhkan sistem Data Warehouse yang mampu mengintegrasikan data demografi pelanggan, lokasi, layanan yang digunakan, serta data status pelanggan untuk mendukung analisis churn secara menyeluruh. Dengan demikian, perusahaan dapat merancang strategi retensi pelanggan yang lebih efektif dan meningkatkan loyalitas pelanggan di masa mendatang.

B. Daftar Stakeholder dan Tujuan Bisnis

Stakeholder	Tujuan Bisnis
Direktur Utama	1. Meningkatkan efektivitas kampanye pemasaran untuk menurunkan tingkat churn pelanggan 2. Meningkatkan loyalitas pelanggan melalui program penawaran yang disesuaikan berdasarkan analisis data pelanggan
Manajer Layanan Pelanggan	1. Meningkatkan kepuasan pelanggan dengan mempercepat respons terhadap keluhan dan masalah teknis 2. Mengurangi churn akibat layanan buruk dengan menerapkan solusi berbasis data dari pengaduan pelanggan
Analisis Data	1. Mengidentifikasi pola perilaku pelanggan yang berisiko churn melalui model prediktif 2. Menyediakan wawasan berbasis data untuk mendukung keputusan strategis di semua departemen
Tim Produk	1. Mengembangkan layanan atau fitur baru berdasarkan feedback pelanggan yang berpotensi meningkatkan retensi 2. Memperbaiki layanan yang memiliki korelasi tinggi dengan churn, seperti internet atau layanan streaming

Manajer Keuangan	<ol style="list-style-type: none"> 1. Mengoptimalkan pendapatan bulanan rata-rata per pelanggan (ARPU) dengan penawaran layanan tambahan 2. Mengontrol biaya akibat churn pelanggan, termasuk biaya pemasaran untuk akuisisi pelanggan baru
------------------	---

Stakeholder memainkan peran penting dengan memberikan pandangan yang beragam dan berharga. Dengan memahami tujuan dan kebutuhan masing-masing stakeholder, perusahaan telekomunikasi dapat memastikan bahwa analisis pelanggan yang dilakukan relevan, komprehensif, dan menghasilkan tindakan yang sesuai dengan tujuan bisnis secara keseluruhan, seperti meningkatkan kepuasan dan loyalitas pelanggan.

C. Tujuan Bisnis

Tujuan bisnis utama yang ingin dicapai dalam perancangan *data warehouse* ini antara lain sebagai berikut:

1. Meningkatkan Retensi Pelanggan
2. Mengurangi Alasan Churn karena Layanan
3. Meningkatkan Efektivitas Kampanye Penawaran
4. Optimasi Biaya Operasional
5. Memaksimalkan Pendapatan per Pelanggan (ARPU)

D. Interview Simulasi

1. Kepada Manajer Pemasaran:
“Apa indikator utama yang Anda gunakan untuk menilai keberhasilan kampanye loyalitas pelanggan?”
2. Kepada Manajer Layanan Pelanggan:
“Apa saja keluhan pelanggan yang paling sering muncul sebelum mereka memutuskan untuk churn?”
3. Kepada Analis Data:
“Bagaimana Anda mengidentifikasi pelanggan yang berisiko churn dalam 3 bulan ke depan?”
4. Kepada Tim Produk:
“Apakah ada fitur atau layanan tertentu yang terbukti mendorong pelanggan untuk tetap menggunakan layanan kami?”
5. Kepada Manajer Keuangan:
“Seberapa besar pengaruh churn terhadap proyeksi pendapatan perusahaan per kuartal?”

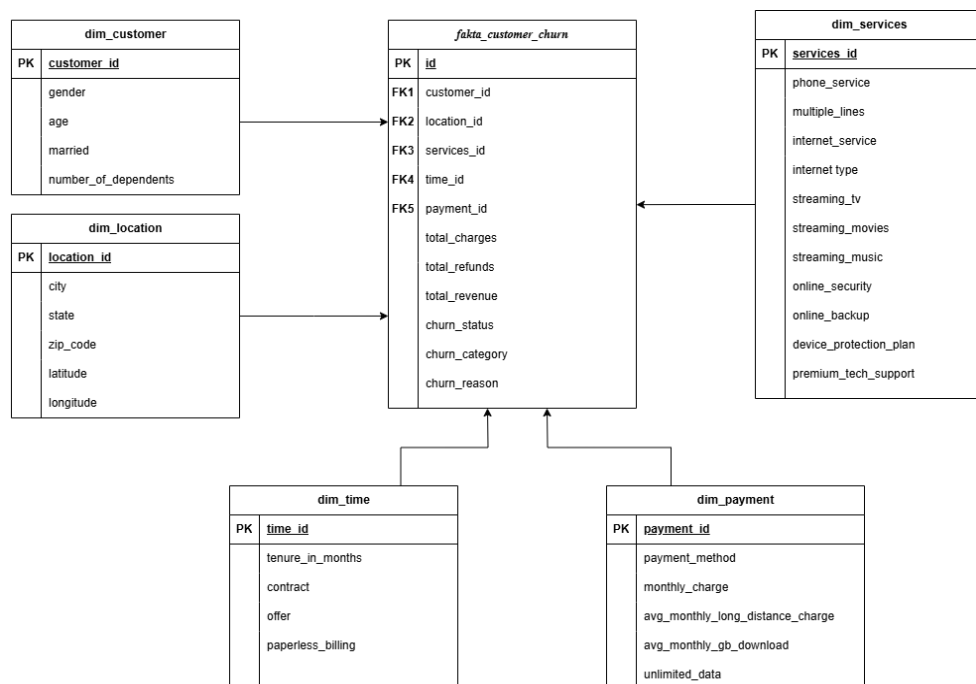
E. Studi Kasus

Banyak perusahaan telekomunikasi menghadapi tantangan besar dalam mempertahankan pelanggan yang telah berlangganan selama 6–12 bulan. Data dari berbagai penyedia jasa menunjukkan bahwa pelanggan cenderung churn saat merasa layanan tidak sepadan dengan biaya, terutama jika terdapat gangguan internet, layanan pelanggan lambat, atau tagihan tidak transparan. Dalam konteks

ini, perusahaan perlu menggunakan data warehouse untuk mengintegrasikan data dari layanan, keluhan, dan transaksi agar bisa melihat korelasi antar faktor.

Misalnya, pelanggan yang menggunakan layanan streaming namun tidak memiliki dukungan teknis premium menunjukkan tingkat churn yang lebih tinggi. Dengan mengandalkan data warehouse, perusahaan bisa melakukan analisis prediktif dan segmentasi pelanggan untuk melakukan tindakan proaktif seperti mengirim diskon atau peningkatan layanan otomatis kepada segmen berisiko. Ini menjadi dasar strategi pengambilan keputusan berbasis data yang lebih tajam dan efektif.

F. Diagram Skema Bintang



G. Penjelasan Tabel Fakta dan Dimensi

Dalam perancangan Data Warehouse untuk industri telekomunikasi, tabel fakta merupakan pusat dari skema bintang yang menyimpan data kuantitatif atau metrik-metrik yang dapat dihitung dan dianalisis. Pada kasus Telecom Customer Churn, tabel fakta bernama `fakta_customer_churn` mencatat metrik penting seperti total charges, total revenue, total refunds, dan status churn pelanggan. Tabel ini menjadi titik utama untuk menganalisis perilaku pelanggan, pendapatan, serta alasan dan kategori churn yang terjadi selama periode tertentu.

Sementara itu, tabel dimensi menyediakan konteks terhadap data yang disimpan dalam tabel fakta. Misalnya, `dim_customer` menyimpan informasi demografis pelanggan seperti jenis kelamin, usia, status pernikahan, dan jumlah tanggungan; `dim_location` berisi informasi geografis seperti kota, provinsi, dan koordinat. Tabel Fakta dan dimensi dapat meningkatkan efisiensi

penyimpanan dan kecepatan query. Selain itu, pendekatan ini mendukung fleksibilitas dalam drill down atau slice and dice data untuk berbagai kebutuhan bisnis. Analisis churn yang mendalam bisa dilakukan secara cepat berdasarkan waktu (melalui *dim_time*), jenis layanan (melalui *dim_services*), dan metode pembayaran (melalui *dim_payment*). Hal ini menjadi landasan penting dalam pengambilan keputusan strategis bagi perusahaan telekomunikasi.

H. Deskripsi Sumber Data

Dataset Telecom Customer Churn dalam file berformat csv merupakan kumpulan data dari sebuah perusahaan telekomunikasi di California yang didapat dari *Kaggle* yang mencakup informasi lengkap mengenai 7.043 pelanggan selama kuartal kedua (April-Juni) tahun 2022. Setiap baris merepresentasikan satu pelanggan, berisi atribut demografis (seperti usia, jenis kelamin, status pernikahan), lokasi geografis, detail layanan langganan (seperti internet, telepon, dan streaming), status dan alasan churn, serta informasi keuangan seperti total pembayaran, pengembalian dana, dan pendapatan. Dataset ini digunakan untuk memahami perilaku pelanggan, mengidentifikasi faktor penyebab churn, dan mendukung pengambilan keputusan strategis berbasis data.

I. Metadata

Sebanyak 38 atribut data yang tersedia dalam dataset adalah sebagai berikut:

No	Nama Atribut	Tipe Data	Deskripsi
1	Customer ID	String	ID unik untuk setiap pelanggan
2	Gender	String (Male/Female)	Jenis kelamin pelanggan
3	Age	Integer	Usia pelanggan
4	Married	String (Yes/No)	Status pernikahan pelanggan
5	Number of Dependents	Integer	Jumlah tanggungan (anak/keluarga)
6	City	String	Kota tempat tinggal pelanggan
7	Zip Code	Integer	Kode pos tempat tinggal
8	Latitude	Float	Koordinat lintang lokasi pelanggan
9	Longitude	Float	Koordinat bujur lokasi pelanggan
10	Number of Referrals	Integer	Jumlah referensi yang diberikan pelanggan
11	Tenure in Months	Integer	Lama berlangganan dalam bulan
12	Offer	String	Nama penawaran khusus yang diterima pelanggan
13	Phone Services	String (Yes/No)	Apakah pelanggan menggunakan layanan telepon
14	Avg Monthly Long Distance Charges	Float	Rata-rata biaya telepon jarak jauh per bulan
15	Multiple Lines	String (Yes/No/No phone service)	Apakah memiliki lebih dari satu saluran telepon

16	Internet Services	String (Yes/No)	Apakah menggunakan layanan internet
17	Internet Type	String	Tipe layanan internet (Fiber Optic, DSL, dsb.)
18	Avg Monthly GB Download	Float	Rata-rata penggunaan data internet per bulan
19	Online Security	String (Yes/No)	Layanan keamanan online tersedia atau tidak
20	Online Backup	String (Yes/No)	Layanan pencadangan data online
21	Device Protection Plan	String (Yes/No)	Apakah pelanggan memiliki proteksi perangkat
22	Premium Tech Support	String (Yes/No)	Apakah pelanggan memiliki dukungan teknis premium
23	Streaming TV	String (Yes/No)	Layanan streaming TV tersedia atau tidak
24	Streaming Movies	String (Yes/No)	Layanan streaming film tersedia atau tidak
25	Streaming Music	String (Yes/No)	Layanan streaming musik tersedia atau tidak
26	Unlimited Data	String (Yes/No)	Apakah pelanggan memiliki data tanpa batas
27	Contract	String	Jenis kontrak (Bulanan, Tahunan, dsb.)
28	Paperless Billing	String (Yes/No)	Metode tagihan tanpa kertas
29	Payment Method	String	Metode pembayaran (Credit card, Bank transfer, dsb.)
30	Monthly Charge	Float	Biaya langganan per bulan
31	Total Charges	Float	Total biaya yang dibayarkan oleh pelanggan
32	Total Refunds	Float	Total pengembalian dana ke pelanggan
33	Total Extra Data Charges	Float	Total biaya tambahan untuk data
34	Total Long Distance Charges	Float	Total biaya telepon jarak jauh
35	Total Revenue	Float	Total pendapatan yang diperoleh dari pelanggan
36	Customer Status	String	Status pelanggan saat ini (Joined, Stayed, Churned)
37	Churn Category	String	Kategori penyebab churn (Harga, Layanan, Kompetitor, dll.)
38	Churn Reason	String	Alasan spesifik pelanggan berhenti berlangganan

REFERENSI

- [1] Zhuang, S. Telecom Customer Churn Prediction [Dataset]. Maven Analytics, 2021. [Online]. Available: <https://www.kaggle.com/datasets/shilongzhuang/telecom-customer-churn-by-maven-analytics>
- [2] Ahmad, A. K., Jafar A., and Aljoumaa K. "Customer churn prediction in telecom using machine learning and social network analysis in big data platform," arXiv preprint arXiv:1904.00690, 2019. [Online]. Available: <https://arxiv.org/abs/1904.00690>