1. Project Overview

This project analyzes customer churn behavior in a telecommunications company using customer-level service, contract, and billing data. The goal is to identify churn patterns, high-risk customer segments, and the key factors that drive customers to leave, in order to support data-driven retention strategies.

2. Dataset Summary

- **Rows:** 7,043
- **Columns:** 33

Key Features:

- Customer profile (gender, senior citizen, partner, dependents, city)
- Account information (contract type, payment method, paperless billing, tenure)
- Service usage (internet service, phone service, streaming, security, tech support)
- Financials (monthly charges, total charges, CLTV)
- Churn indicators (churn value, churn reason, churn score)

Missing Data:

Some customers have missing values in total_charges and churn_reason, which were handled during data preparation.

3. Exploratory Data Analysis using Python

Began with data preparation and cleaning in Python to ensure the Telco customer dataset was ready for analysis.

- **Data Loading**
  Imported the Telco customer churn dataset into Python using pandas.
- **Initial Exploration**
  Used df.info() to inspect the structure of the dataset and df.describe() to review summary statistics such as tenure, monthly charges, total charges, and CLTV.

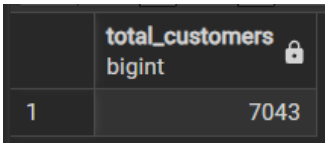
	CustomerID	Count	Country	State	City	Zip Code	Lat Long	Latitude	Longitude	Gender	...	Contract	Paperless Billing	Payment Method	Monthly Charges	Total Charges	Churn Label	Churn Value	Churn Score	CLTV	Churn Reason
count	7043	7043.0	7043	7043	7043	7043.000000	7043	7043.000000	7043.000000	7043	...	7043	7043	7043	7043.000000	7043	7043	7043.000000	7043.000000	7043.000000	1869
unique	7043	NaN	1	1	1129	NaN	1652	NaN	NaN	2	...	3	2	4	NaN	6531	2	NaN	NaN	NaN	20
top	3186-AJIEK	NaN	United States	California	Los Angeles	NaN	34.02381, -118.156582	NaN	NaN	Male	...	Month-to-month	Yes	Electronic check	NaN		No	NaN	NaN	NaN	Attitude of support person
freq	1	NaN	7043	7043	305	NaN	5	NaN	NaN	3555	...	3875	4171	2365	NaN	11	5174	NaN	NaN	NaN	192
mean	NaN	1.0	NaN	NaN	NaN	93521.964646	NaN	36.282441	-119.798880	NaN	...	NaN	NaN	NaN	64.761692	NaN	NaN	0.265370	58.699418	4400.295755	NaN
std	NaN	0.0	NaN	NaN	NaN	1865.794555	NaN	2.455723	2.157889	NaN	...	NaN	NaN	NaN	30.090047	NaN	NaN	0.441561	21.525131	1183.057152	NaN
min	NaN	1.0	NaN	NaN	NaN	90001.000000	NaN	32.555828	-124.301372	NaN	...	NaN	NaN	NaN	18.250000	NaN	NaN	0.000000	5.000000	2003.000000	NaN
25%	NaN	1.0	NaN	NaN	NaN	92102.000000	NaN	34.030915	-121.815412	NaN	...	NaN	NaN	NaN	35.500000	NaN	NaN	0.000000	40.000000	3469.000000	NaN
50%	NaN	1.0	NaN	NaN	NaN	93552.000000	NaN	36.391777	-119.730885	NaN	...	NaN	NaN	NaN	70.350000	NaN	NaN	0.000000	61.000000	4527.000000	NaN
75%	NaN	1.0	NaN	NaN	NaN	95351.000000	NaN	38.224869	-118.043237	NaN	...	NaN	NaN	NaN	89.850000	NaN	NaN	1.000000	75.000000	5380.500000	NaN
max	NaN	1.0	NaN	NaN	NaN	96161.000000	NaN	41.962127	-114.192901	NaN	...	NaN	NaN	NaN	118.750000	NaN	NaN	1.000000	100.000000	6500.000000	NaN

- **Missing Data Handling**
  Checked for missing values in the dataset. Missing values in the total_charges column were converted to numeric format and filled using the median. Missing values in the churn_reason column were replaced with "Not Churned" to ensure all customers had a valid churn category.
- **Column Standardization**
  Renamed all columns into snake_case format to improve readability and ensure consistency across Python, SQL, and Power BI.
- **Feature Engineering**
  Created a tenure_group column by grouping tenure_months into five segments (0–1y, 1–2y, 2–4y, 4–5y, and 5y+) to support churn analysis across different customer lifecycle stages.
- **Data Consistency Check**
  Verified that churn_label and churn_value represented the same churn outcome. Since both columns contained identical information, churn_label was removed to avoid redundancy.
- **Database Integration**
  Loaded the cleaned and prepared dataset into PostgreSQL, making it available for SQL analysis and Power BI visualization.

4. Data Analysis Using SQL

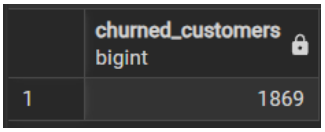
Performed structured analysis in PostgreSQL to answer key business questions:

1. Total Customers : How many active customers are in the dataset?



	total_customers bigint
1	7043

2. Churned Customers : How many customers have actually churned?

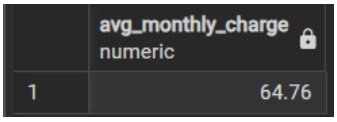


	churned_customers bigint
1	1869

3. Churn Rate : Out of all customers, what percentage have churned?

	churn_rate numeric
1	26.54

4. Average Monthly Charge : What is the average amount customers pay per month?



	avg_monthly_charge numeric
1	64.76

5. Churned Customers by Gender : Does churn occur more among men or women?

	gender text	churned_customers bigint
1	Female	939
2	Male	930

6. Churned Customers by Senior Citizen : Does churn occur more among senior citizens or non-senior customers?

	senior_citizen text	churned_customers bigint
1	No	1393
2	Yes	476

7. Top 5 Cities by Churned Customers : Which city has the highest number of churned customers?

	city text	churned_customers bigint
1	Los Angeles	90
2	San Diego	50
3	San Francis...	31
4	San Jose	29
5	Sacramento	26

8. Churned Customers by Internet Service : Which internet service type contributes the most to churn?

	internet_service text	churned_customers bigint
1	Fiber optic	1297
2	DSL	459
3	No	113

9. Churn Rate by Payment Method : Which payment method has the highest churn risk?

	payment_method text	churn_rate numeric
1	Electronic check	45.29
2	Mailed check	19.11
3	Bank transfer (automat...	16.71
4	Credit card (automatic)	15.24

10. Churn Rate by Contract : Which contract type has the highest churn risk?

	contract text	churn_rate numeric
1	Month-to-mon...	42.71
2	One year	11.27
3	Two year	2.83

11. Churned Customers & Churn Rate by Tenure Group : How does churn change as customer tenure increases?

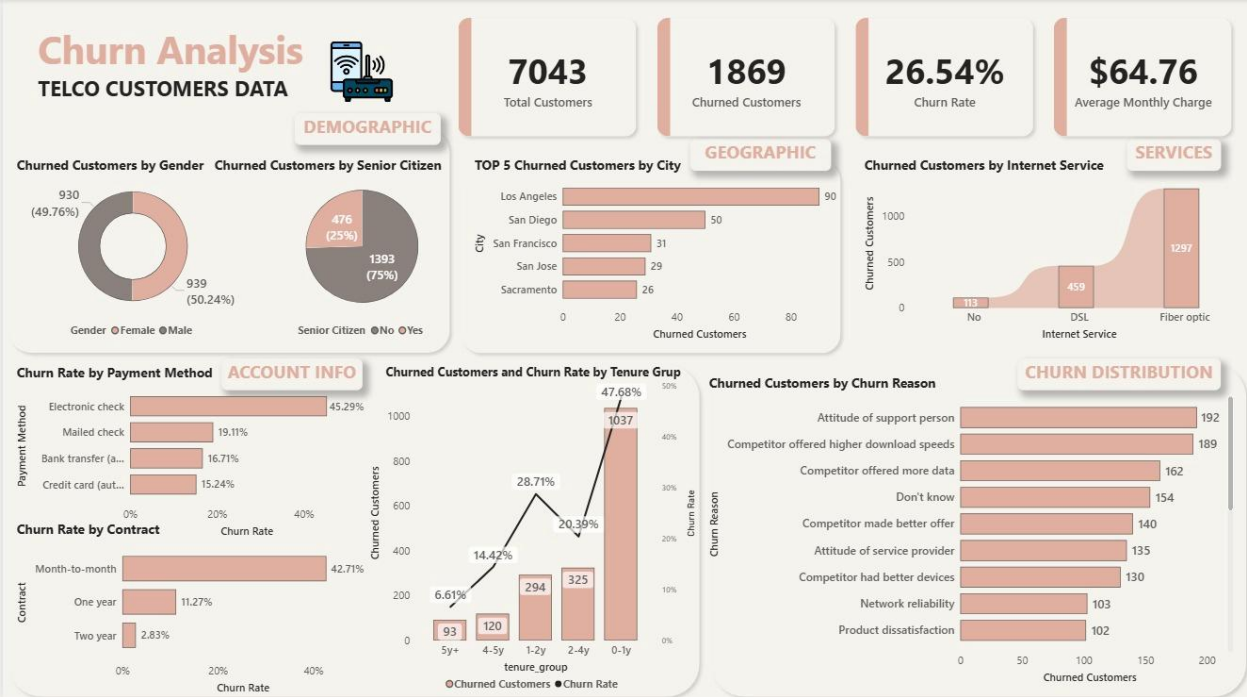
	tenure_group text	total_customers bigint	churned_customers bigint	churn_rate numeric
1	0-1y	2175	1037	47.68
2	1-2y	1024	294	28.71
3	2-4y	1594	325	20.39
4	4-5y	832	120	14.42
5	5y+	1407	93	6.61

12. Churned Customers by Churn Reason : Why do customers leave the company?

	churn_reason text	churned_customers bigint
1	Attitude of support person	192
2	Competitor offered higher download spe...	189
3	Competitor offered more data	162
4	Don't know	154
5	Competitor made better offer	140
6	Attitude of service provider	135
7	Competitor had better devices	130
8	Network reliability	103
9	Product dissatisfaction	102
10	Price too high	98
11	Service dissatisfaction	89
12	Lack of self-service on Website	88
13	Extra data charges	57
14	Moved	53
15	Long distance charges	44
16	Limited range of services	44
17	Lack of affordable download/upload spe...	44
18	Poor expertise of phone support	20
19	Poor expertise of online support	19
20	Deceased	6

5. Dashboard in Power BI

Built an interactive dashboard in Power BI to present insights visually.



5. Business Recomendation

• Convert Month-to-Month Customers to Long-Term Contracts

The churn rate for month-to-month customers is dramatically higher (over 40%) compared to one-year and two-year contracts, which show very low churn. The company should aggressively promote long-term contracts by offering discounts, loyalty rewards, or bundled services to reduce churn in this high-risk segment.

• Reduce Churn Among Electronic Check Users

Customers who pay using electronic check have the highest churn rate among all payment methods. This suggests lower engagement and weaker customer commitment. The company should encourage these customers to switch to automatic payments such as credit card or bank transfer by offering incentives like small monthly discounts or bonus data.

• Prioritize Retention for Fiber Optic Customers

Fiber optic users account for the largest number of churned customers. Since this group typically represents premium and high-usage customers, losing them has a significant impact on the business. The company should improve service reliability, pricing transparency, and customer support specifically for fiber optic subscribers.

• **Focus Retention Efforts on New Customers (0–1 Year Tenure)**
Customers in the first year of subscription have the highest churn rate, which drops significantly as tenure increases. This indicates that the early customer experience is critical. Onboarding programs, proactive support, and early-stage promotions should be used to stabilize new customers and prevent early churn.

• **Target High-Churn Cities with Local Retention Campaigns**
Cities such as Los Angeles and San Diego contribute the highest number of churned customers. Localized marketing, targeted promotions, or service quality improvements in these cities could have a large impact on reducing overall churn.

• **Improve Customer Service and Competitive Positioning**
The top churn reasons are related to competitor offers, better download speeds, more data, and dissatisfaction with service attitude. This indicates that both pricing competitiveness and customer experience need to be improved. The company should review its pricing plans, upgrade network performance, and provide additional training to customer support teams to reduce churn driven by competition and poor service perception.