

CUSTOMER CHURN PREDICTION

Phase 5: Submission Document

Project Title: *Customer Churn Prediction*

Phase 4: Project Documentation & Submission

Topic: In this section, we will document the complete project and prepare it for submission.



CUSTOMER CHURN PREDICTION

INTRODUCTION:

- ❖ Customer churn prediction is a vital component of customer relationship management and business strategy. It involves the use of data and predictive analytics to identify and forecast the likelihood of customers discontinuing their relationship with a company or business.
- ❖ Churn, often referred to as customer attrition or customer turnover, is a significant concern for businesses across various industries, as retaining existing customers is generally more cost-effective than acquiring new ones.
- ❖ Customer churn can result in significant revenue loss for businesses. Identifying potential churners allows companies to take targeted actions to retain these customers and, in turn, protect their revenue streams.
- ❖ Understanding why customers are leaving enables companies to address underlying issues, improve customer satisfaction, and enhance the overall customer experience.
- ❖ The primary goal of customer churn prediction is to identify and understand the factors and patterns that

contribute to customer attrition. By doing so, businesses can take proactive measures to reduce churn and improve customer retention.

GIVEN DATASET:

CustomerID	Gender	SeniorCitizen	Partner	Dependents	Tenure	PhoneService	MultipleLines	InternetService
7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL
5575-GNVDE	Male	0	No	No	34	Yes	No	DSL
3668-QPYBK	Male	0	No	No	2	Yes	No	DSL
7795-CFOCW	Male	0	No	No	45	No	No phone service	DSL
9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic
9305-CDSKC	Female	0	No	No	8	Yes	Yes	Fiber optic
1452-KIOVK	Male	0	No	Yes	22	Yes	Yes	Fiber optic
6713-OKOMC	Female	0	No	No	10	No	No phone service	DSL
7892-POOKP	Female	0	Yes	No	28	Yes	Yes	Fiber optic
6388-TABGU	Male	0	No	Yes	62	Yes	No	DSL
9763-GRSKD	Male	0	Yes	Yes	13	Yes	No	DSL
7469-LKBCI	Male	0	No	No	16	Yes	No	No
8091-TTVAX	Male	0	Yes	No	58	Yes	Yes	Fiber optic
0280-XJGEX	Male	0	No	No	49	Yes	Yes	Fiber optic
5129-JLPIS	Male	0	No	No	25	Yes	No	Fiber optic
3655-SNQYZ	Female	0	Yes	Yes	69	Yes	Yes	Fiber optic
8191-XWSZG	Female	0	No	No	52	Yes	No	No
9959-WOFKT	Male	0	No	Yes	71	Yes	Yes	Fiber optic
4190-MFLUW	Female	0	Yes	Yes	10	Yes	No	DSL

Here's a list of tools and software commonly used in the process:

➤ **Model Evaluation:**

Assess the model's performance using evaluation metrics such as accuracy, precision, recall, F1-score, and the Receiver Operating Characteristic Area Under the Curve (ROC AUC).

➤ **Tune and Iterate:**

Fine-tune your model based on evaluation results. Adjust hyper parameter, try different algorithms, or consider using ensemble methods to improve performance.

➤ **Validation:**

Use the validation set to check how well your model generalizes to new, unseen data. Make necessary adjustments based on validation results.

➤ **Testing:**

Finally, evaluate your model using the test dataset to obtain an accurate assessment of its predictive power. This step simulates how the model will perform in real-world scenarios.

➤ **Data Collection:**

Gather historical customer data, which may include customer demographics, transaction history, customer interactions, and any other relevant information. This data is essential for training and testing your churn prediction model.

➤ **Data Preprocessing:**

- Handle missing values by imputing or removing them.
- Encode categorical variables into numerical format.
- Normalize or scale numerical features to ensure they have a similar impact on the model.

➤ **Exploratory Data Analysis (EDA):**

- Conduct data visualization and statistical analysis to understand your data.
- Identify patterns and correlations within the data that may be associated with churn.

➤ **Resource Allocation:**

Instead of applying retention strategies to all customers, companies can allocate resources more efficiently by focusing on those customers with a high likelihood of churning.



1. DESIGN THINKING AND PRESENT IN FORM OF DOCUMENT

1. Empathize:

- Understand the needs, concerns, and pain points of your customers who might churn.
- Conduct interviews, surveys, or analyze customer feedback to gather insights.

2. Define:

- Clearly define the problem you're trying to address, such as identifying factors leading to customer churn.

- Create a problem statement that focuses on customer needs and business objectives.

3. Ideate:

- Brainstorm potential solutions and strategies for predicting and preventing customer churn.
- Encourage cross-functional collaboration among teams.

4. Prototype:

- Develop a prototype or model for your customer churn prediction system.
- Use data analysis and machine learning techniques to create a predictive model.

5. Test:

- Evaluate the prototype's accuracy in predicting churn by using historical data.
- Refine the model as needed, considering feedback from stakeholders.

6. Implement:

- Roll out the final customer churn prediction system into your business operations.
- Train staff and implement processes to act on churn predictions.

7. Iterate:

- Continuously gather data and feedback to refine the churn prediction model.
- Stay responsive to changing customer behavior and market conditions.

8. Document:

- Create a comprehensive document outlining each step in the design thinking process.
- Include details of the research, models used, and the iterative approach.

9. Present:

- Communicate your findings, model performance, and the value of churn prediction to stakeholders.
- Highlight how this approach aligns with customer-centric goals and business growth.

10. Future Recommendations:

- Suggest future enhancements and improvements to the churn prediction system.
- Emphasize the importance of ongoing customer engagement and feedback collection.

2. DESIGN INTO INNOVATION

1. Advanced Data Analytics:

- Utilize advanced data analytics techniques, such as machine learning and deep learning, to improve the accuracy of churn prediction models.

2. Real-time Data Integration:

- Incorporate real-time data sources, like social media feeds or customer interactions, to make predictions more dynamic and responsive to changing customer behaviour.

3. Feature Engineering:

- Explore new and relevant features that can enhance the predictive power of your models. This could include sentiment analysis of customer feedback or geospatial data.

4. Predictive Maintenance:

- Use predictive analytics not only to identify potential churners but also to predict when

they might churn, allowing for proactive retention efforts.

5. Segmentation:

- Segment customers based on different characteristics and apply customized churn prediction models and retention strategies to each segment.

6. Explainability and Interpretability:

- Develop models that provide insights into why a customer is likely to churn, which can guide your innovation in retention strategies.

7. A/B Testing:

- Continuously test and optimize different churn prevention strategies to identify what works best for your specific customer base.

8. Customer Engagement Platforms:

- Integrate customer engagement platforms that enable personalized interactions with at-risk customers through various channels like email, SMS, or chatbots.

9. Predictive Analytics Tools:

- Implement predictive analytics tools that allow non-technical staff to leverage the power of churn prediction, making it accessible to various departments within the organization.

10. Feedback Loops:

- Establish feedback loops where the results of churn prediction models inform product and service improvements.

11. Ethical Considerations:

- Ensure that your churn prediction models are built and used ethically, respecting customer privacy and avoiding biases.

12. Cross-functional Collaboration:

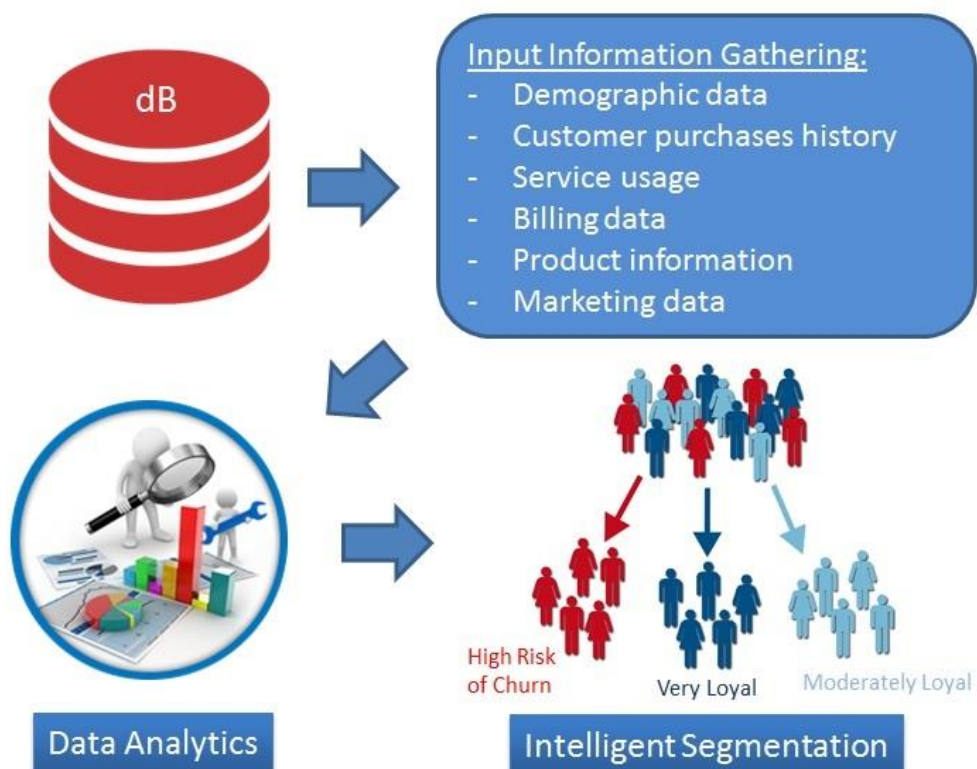
- Foster collaboration between data scientists, marketing, customer service, and other relevant departments to leverage collective knowledge and drive innovation.

13. Data Enrichment:

- Enhance your customer data with external sources like demographics, market trends, or social data to gain a more comprehensive view of your customers.

14. Continuous Learning:

- Keep your models updated and continuously learn from customer behaviour and model performance to adapt to changing customer dynamics.



PROGRAM:

In [1]:

#Import necessary libraries

```
import pandas as pd
```

```
import numpy as np
```

```
import missingno as msno
```

```
import matplotlib.pyplot as plt
```

```
import seaborn as sns
```

```
import plotly.express as px
```

```
import plotly.graph_objects as go
```

```
from plotly.subplots import make_subplots
```

```
import warnings
```

```
warnings.filterwarnings('ignore')
```

In[2]:

```
from sklearn.preprocessing import StandardScaler
```

```
from sklearn.preprocessing import LabelEncoder
```

```
from sklearn.tree import DecisionTreeClassifier
```

```
from sklearn.ensemble import RandomForestClassifier
```

```
from sklearn.naive_bayes import GaussianNB
```

```
from sklearn.neighbors import KNeighborsClassifier
```

```
from sklearn.svm import SVC
```

```
from sklearn.neural_network import MLPClassifier
from sklearn.ensemble import AdaBoostClassifier
from sklearn.ensemble import GradientBoostingClassifier
from sklearn.ensemble import ExtraTreesClassifier
from sklearn.linear_model import LogisticRegression
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
from xgboost import XGBClassifier
from catboost import CatBoostClassifier
from sklearn import metrics
from sklearn.metrics import roc_curve
from sklearn.metrics import recall_score, confusion_matrix,
precision_score, f1_score, accuracy_score,
classification_report
```

In [3]:

```
#loading data
```

```
df = pd.read_csv('../input/telco-customer-churn/WA_Fn-
UseC_-Telco-Customer-Churn.csv')
```

In [4]:

```
df.head()
```

out[4]:

	customerID	gender	SeniorCitizen	Partner	Dependents	Tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	. . .	DeviceProtection	TechSupport	StreamingTV	StreamingMovies	Contract	PaperlessBilling	PaymentMethod	MonthlyCharges	TotalCharges	Churn
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	. . .	No	No	No	No	Month-to-month	Yes	Electronic check	29.85	29.85	No
1	5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	. . .	Yes	No	No	No	One year	No	Mailed check	56.95	1889.5	No
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	. . .	No	No	No	No	Month-to-month	Yes	Mailed check	53.85	108.15	Yes
3	7795-CFOW	Male	0	No	No	45	No	No phone service	DSL	Yes	. . .	Yes	Yes	No	No	One year	No	Bank transfer (automatic)	42.30	1840.75	No

	customerID	gender	SeniorCitizen	Partner	Dependents	Tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	DeviceProtection	TechSupport	StreamingTV	StreamingMovies	Contract	PaperlessBilling	PaymentMethod	MonthlyCharges	TotalCharges	Churn	
4	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	..	No	No	No	No	Month-to-month	Yes	Electronic check	70.70	151.65	Yes

In [5]:

```
df.shape
```

Out [5]:

```
(7043, 21)
```

In [6]:

```
df.info()
```

```
<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)
```

In [7]:

```
df.columns.values
```

Out[7]:

```
array(['customerID', 'gender', 'SeniorCitizen', 'Partner', 'Dependents', 'tenure', 'PhoneService', 'MultipleLines', 'InternetService', 'OnlineSecurity', 'OnlineBackup', 'DeviceProtection', 'TechSupport', 'StreamingTV', 'StreamingMovies', 'Contract', 'PaperlessBilling', 'PaymentMethod', 'MonthlyCharges', 'TotalCharges', 'Churn'], dtype=object)
```

In [8]:

```
df.dtypes
```

Out[8]:

customerID	object
gender	object
SeniorCitizen	int64
Partner	object
Dependents	object
tenure	int64

```

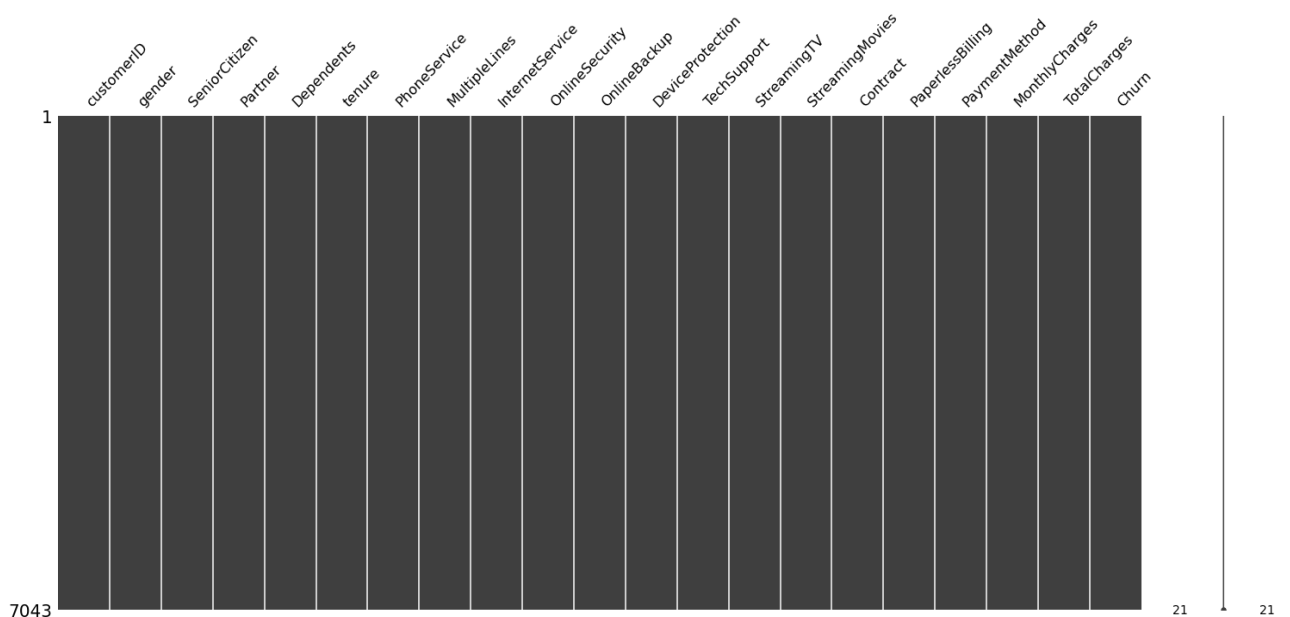
PhoneService      object
MultipleLines     object
InternetService   object
OnlineSecurity     object
OnlineBackup      object
DeviceProtection  object
TechSupport       object
StreamingTV       object
StreamingMovies   object
Contract          object
PaperlessBilling  object
PaymentMethod     object
MonthlyCharges    float64
TotalCharges      object
Churn             object
dtype: object

```

In [9]:

Visualize missing values as a matrix

```
msno.matrix(df);
```



In [10]:

```
df = df.drop(['customerID'], axis = 1)
df.head()
```

Out[10]:

	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	OnlineBackup	DeviceProtection	TechSupport	StreamingTV	StreamingMovies	Contract	PaperlessBilling	PaymentMethod	MonthlyCharges	TotalCharges	Churn
0	Female	0	Yes	No	1	No	No phone service	DSL	No	Yes	No	No	No	No	Month-to-month	Yes	Electronic check	29.85	29.85	No
1	Male	0	No	No	34	Yes	No	DSL	Yes	No	Yes	No	No	No	One year	No	Mailed check	56.95	1889.5	No
2	Male	0	No	No	2	Yes	No	DSL	Yes	Yes	No	No	No	No	Month-to-month	Yes	Mailed check	53.85	108.15	Yes
3	Male	0	No	No	45	No	No phone service	DSL	Yes	No	Yes	Yes	No	No	One year	No	Bank transfer (auto)	42.30	1840.75	No

gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	OnlineBackup	DeviceProtection	TechSupport	StreamingTV	StreamingMovies	Contract	PaperlessBilling	PaymentMethod	MonthlyCharges	TotalCharges	Churn
						vic e								ear		omatic)			
4 Female	0	No	No	2	Yes	No	Fiber optic	No	No	No	No	No	No	Month- to- month	Yes	Electronic check	70.70	151.65	Yes

In [11]:

```
df['TotalCharges'] = pd.to_numeric(df.TotalCharges, errors='coerce')
df.isnull().sum()
```

Out[11]:

```
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

```

In [12]:

```
df[np.isnan(df['TotalCharges'])]
```

Out[12]:

	gender	Senior Citizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	OnlineBackup	DeviceProtection	TechSupport	StreamingTV	StreamingMovies	Contract	PaperlessBilling	PaymentMethod	MonthlyCharges	TotalCharges	Churn
488	Female	0	Yes	Yes	0	No	No phone service	DSL	Yes	No	Yes	Yes	Yes	No	Two year	Yes	Bank transfer (automatic)	52.55	NaN	No
753	Male	0	No	Yes	0	Yes	No	No	No internet service	No internet service	No internet service	No internet service	No internet service	No internet service	Two year	No	Mail check	20.25	NaN	No

	gender	Senior Citizen	Partner	Dependents	tenure	Phone service	Multiple lines	Internet service	Online security	Online Backup	Device Protection	Tech Support	Streaming TV	Streaming Movies	Contract	Paperless Billing	Payment Method	Monthly Charges	Total Charges	Churn
936	Female	0	Yes	Yes	0	Yes	No	DSL	Yes	Yes	Yes	No	Yes	Yes	Two year	No	Mail check	80.85	NAN	No
1082	Male	0	Yes	Yes	0	Yes	Yes	No	No internet service	No internet service	No internet service	No internet service	No internet service	No internet service	Two year	No	Mail check	25.75	NAN	No
1340	Female	0	Yes	Yes	0	No	No phone service	DSL	Yes	Yes	Yes	Yes	Yes	No	Two year	No	Credit card (automatic)	56.05	NAN	No
3331	Male	0	Yes	Yes	0	Yes	No	No	No internet service	No internet service	No internet service	No internet service	No internet service	No internet service	Two year	No	Mail check	19.85	NAN	No
38	Male	0	Yes	Yes	0	Yes	Yes	No	No internet	No internet	No internet	No internet	No internet	No internet	Two y	No	Mail ch	25.35	NAN	No

	gender	Senior Citizen	Partner	Dependents	tenure	Phone Service	Multiple Lines	Internet Service	Online Security	Online Backup	Device Protection	Tech Support	Streaming TV	Streaming Movies	Contract	Paperless Billing	Payment Method	Monthly Charges	Total Charges	Churn
26	Male								service	net service	service	net service	net service	service	ear		check			
4380	Female	0	Yes	Yes	0	Yes	No	No	No internet service	No internet service	No internet service	No internet service	No internet service	No internet service	Two years	No	Mail check	20.00	NAN	No
5218	Male	0	Yes	Yes	0	Yes	No	No	No internet service	No internet service	No internet service	No internet service	No internet service	No internet service	One year	Yes	Mail check	19.70	NAN	No
6670	Female	0	Yes	Yes	0	Yes	Yes	DSL	No	Yes	Yes	Yes	Yes	No	Two years	No	Mail check	73.35	NAN	No
6754	Male	0	No	Yes	0	Yes	Yes	DSL	Yes	Yes	No	Yes	No	No	Two years	Yes	Bank transfer (61.90	NAN	No

In [13]:

```
df[df['tenure'] == 0].index
```

Out[13]:

```
Int64Index([488, 753, 936, 1082, 1340, 3331, 3826, 4380, 5218, 670, 6754], dtype='int64')
```

In [14]:

```
df.drop(labels=df[df['tenure'] == 0].index, axis=0, inplace=True)
df[df['tenure'] == 0].index
```

Out[14]:

```
Int64Index([], dtype='int64')
```

To solve the problem of missing values in TotalCharges column, I decided to fill it with the mean of TotalCharges values.

In [15]:

```
df.fillna(df["TotalCharges"].mean())
```

Out[15]:

	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	OnlineBackup	DeviceProtection	TechSupport	StreamingTV	StreamingMovies	Contract	PaperlessBilling	PaymentMethod	MonthlyCharges	TotalCharges	Churn
0	Female	0	Yes	No	1	No	No phone service	DSL	No	Yes	No	No	No	No	Month-to-month	Yes	Electronic check	29.85	29.85	No

	gender	Senior Citizen	Partner	Dependents	tenure	Phone Service	Multiple Lines	Internet Service	Online Security	Online Backup	Device Protection	Tech Support	Streaming TV	Streaming Movies	Contract	Paperless Billing	Payment Method	Monthly Charges	Total Charges	Churn
1	Male	0	No	No	34	Yes	No	DSL	Yes	No	Yes	No	No	No	One year	No	Mail check	56.95	1889.50	No
2	Male	0	No	No	2	Yes	No	DSL	Yes	Yes	No	No	No	No	Month-to-month	Yes	Mail check	53.85	108.15	Yes
3	Male	0	No	No	45	No	No phone service	DSL	Yes	No	Yes	Yes	No	No	One year	No	Bank transfer (automatic)	42.30	1840.75	No
4	Female	0	No	No	2	Yes	No	Fiber optic	No	No	No	No	No	No	Month-to-month	Yes	Electronic check	70.70	151.65	Yes

	gender	Senior Citizen	Partner	Dependents	tenure	Phone service	Multiple lines	Internet service	Online security	Online Backup	Device Protection	Tech Support	Streaming TV	Streaming Movies	Contract	Paperless Billing	Payment Method	Monthly Charges	Total Charges	Churn
															th					

7038	Male	0	Yes	Yes	24	Yes	Yes	DSL	Yes	No	Yes	Yes	Yes	Yes	One year	Yes	Mail check	84.80	1990.50	No
7039	Female	0	Yes	Yes	72	Yes	Yes	Fiber optic	No	Yes	Yes	No	Yes	Yes	One year	Yes	Credit card (automatic)	103.20	7362.90	No
7040	Female	0	Yes	Yes	11	No	No phone service	DSL	Yes	No	No	No	No	No	Month-to-month	Yes	Electronic check	29.60	346.45	No

	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	OnlineBackup	DeviceProtection	TechSupport	StreamingTV	StreamingMovies	Contract	PaperlessBilling	PaymentMethod	MonthlyCharges	TotalCharges	Churn
7041	Male	1	Yes	No	4	Yes	Yes	Fiber optic	No	No	No	No	No	No	Month-to-month	Yes	Mail check	74.40	306.60	Yes
7042	Male	0	No	No	66	Yes	No	Fiber optic	Yes	No	Yes	Yes	Yes	Yes	Two year	Yes	Bank transfer (automatic)	105.65	6844.50	No

In [16]:

```
df.isnull().sum()
```

Out[16]:

```
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      0
Churn             0
dtype: int64

```

In [17]:

```

df["SeniorCitizen"] = df["SeniorCitizen"].map({0: "No", 1: "Yes"})
df.head()

```

Out[17]:

	gender	Senior Citizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	OnlineBackup	DeviceProtection	TechSupport	StreamingTV	StreamingMovies	Contract	PaperlessBilling	PaymentMethod	MonthlyCharges	TotalCharges	Churn
0	Female	No	Yes	No	1	No	No phone service	DSL	No	Yes	No	No	No	No	Month-to-month	Yes	Electronic check	29.85	29.85	No

	gender	Senior Citizen	Partner	Dependents	tenure	Phone service	Multiple lines	Internet service	Online security	Online backup	Device Protection	Tech Support	Streaming TV	Streaming Movies	Contract	Paperless Billing	Payment Method	Monthly Charges	Total Charges	Churn
1	Male	No	No	No	34	Yes	No	DSL	Yes	No	Yes	No	No	No	One year	No	Mail check	56.95	1889.50	No
2	Male	No	No	No	2	Yes	No	DSL	Yes	Yes	No	No	No	No	Month-to-month	Yes	Mail check	53.85	108.15	Yes
3	Male	No	No	No	45	No	No phone service	DSL	Yes	No	Yes	Yes	No	No	One year	No	Bank transfer (automatic)	42.30	1840.75	No
4	Female	No	No	No	2	Yes	No	Fiber optic	No	No	No	No	No	No	Month-to-month	Yes	Electronic check	70.70	151.65	Yes

	g e n d e r	Se ni or Cit ize n	P a r t n e r	D e p e n d e n t s	t e n u r e	Ph o n e S e r v i c e	M u l t i p l e L i n e s	Int er ne t S e r v i c e	On lin e S e c u r i t y	On lin e B a c k u p	De vic e P r o t e c t i o n	Te ch S u p p o r t	St re a m i n g T V	Str ea m i n g M o v i e s	C o n t r a c t	Pa p e r l e s s B i l l i n g	Pay m e n t M e t h o d	Mo n t h l y C h a r g e s	To t a l C h a r g e s	C h u r n
															n t h					

In [18]:

```
df["InternetService"].describe(include=['object', 'bool'])
```

Out[18]:

```
count      7032
unique       3
top      Fiber optic
freq      3096
```

Name: InternetService, dtype: object

In [19]:

```
numerical_cols = ['tenure', 'MonthlyCharges', 'TotalCharges']
df[numerical_cols].describe()
```

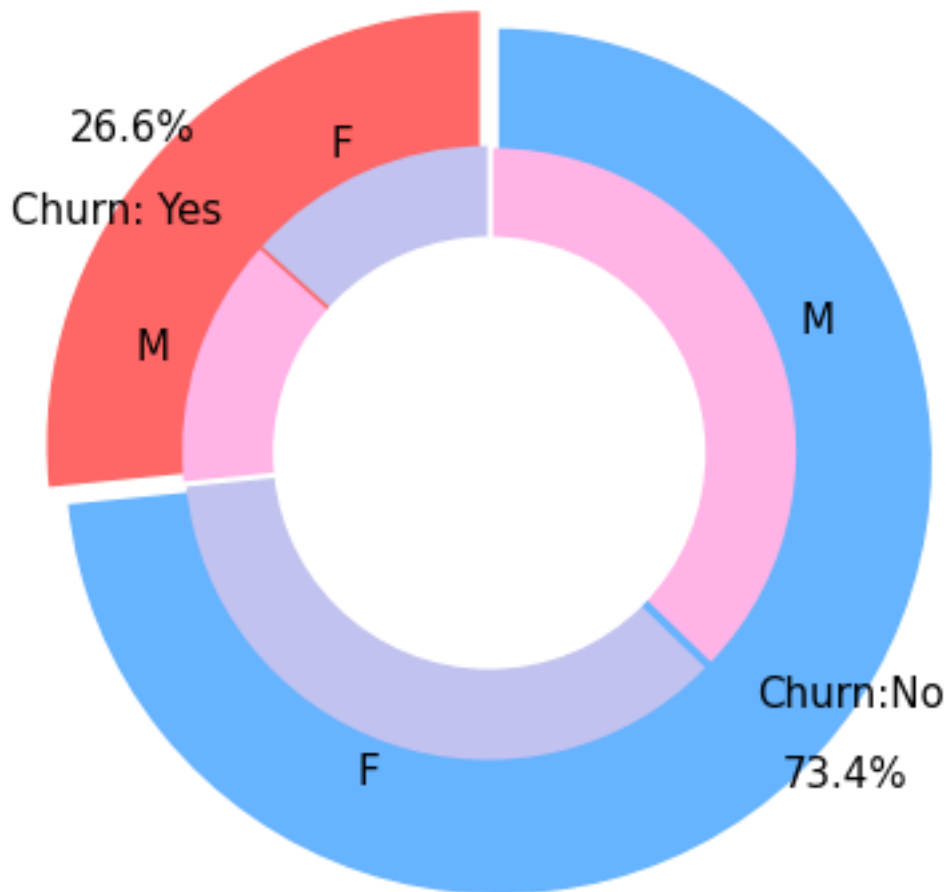
Out[19]:

	tenure	MonthlyCharges	TotalCharges
count	7032.000000	7032.000000	7032.000000
mean	32.421786	64.798208	2283.300441

	tenure	MonthlyCharges	TotalCharges
std	24.545260	30.085974	2266.771362
min	1.000000	18.250000	18.800000
25%	9.000000	35.587500	401.450000
50%	29.000000	70.350000	1397.475000
75%	55.000000	89.862500	3794.737500
max	72.000000	118.750000	8684.800000

Data Visualization:

Churn Distribution w.r.t Gender: Male(M), Female(F)



ADVANTAGES:

- ❖ **Revenue Retention:** Identifying and retaining customers who might otherwise churn can lead to increased revenue. It's generally more cost-effective to keep existing customers than to acquire new ones.

- ❖ **Cost Savings:** Reducing churn can lead to cost savings. Acquiring new customers is typically more expensive than retaining existing ones. By focusing on customer retention, you can save on marketing and acquisition costs.
- ❖ **Customer Lifetime Value:** By preventing churn, you can maximize the customer lifetime value (CLV). Loyal and long-term customers tend to spend more over time, making them more profitable.
- ❖ **Improved Customer Satisfaction:** Churn prediction can help identify areas where customers are dissatisfied or dissuaded from continued engagement. This information can be used to improve products or services, leading to increased customer satisfaction.
- ❖ **Personalized Marketing:** Churn prediction enables businesses to create personalized marketing and engagement strategies. Tailoring messages, offers, and recommendations to individual customers increases the likelihood of retaining them.
- ❖ **Customer Segmentation:** Churn prediction models can help identify customer segments with different churn risks. This information can guide marketing, product development, and customer support strategies tailored to each segment.
- ❖ **Continuous Improvement:** Churn prediction models can be refined and improved over time as more data becomes available and as machine learning techniques advance.

DISADVANTAGES:



7 REASONS YOU ARE LOSING CUSTOMERS

- 1 The overall customer experience is lacking
- 2 Customers don't get the value they expected from your product
- 3 Lack of robust onboarding process to help users start using the offering
- 4 Slow load time or processing time of your product/website
- 5 Your product is too expensive compared to competitors.
- 6 UI/UX is not intuitive and fluid
- 7 Bugs and glitches

Gramener
Insights as Stories

CONCLUSION:

Customer Churn Prediction is a vital tool for businesses to retain customers and reduce revenue loss. Through the analysis of historical customer data and the

application of machine learning models, we can identify patterns and factors that contribute to customer churn.

By proactively identifying at-risk customers and implementing targeted retention strategies, businesses can significantly improve customer retention rates and ultimately boost their bottom line.

The accuracy and effectiveness of the churn prediction model will play a critical role in achieving these objectives, making it a valuable asset for customer-centric organizations.

Customer churn analysis allows to minimize acquisition costs and increase marketing efficiency, preparing a solid base for future marketing analysis and campaigns. Customer churn analysis opens new opportunities for cross-selling and upselling and serves as one of the starting points for customer-driven product development, keeping customers engaged and loyal over time.