IBM Customer Churn Analysis and Prediction Report

Telco customer churn: IBM dataset:

https://www.kaggle.com/datasets/yeanzc/telco-customer-churn-ibm-dataset

Dataset Shape: (7043, 33)

Data Inspection

Steps:

1. Checking shape of data
2. Checking dtypes of all columns using info()
3. Checking Stats summary
4. Checking missing values

    only "Churn Reason" has missing values (1574 missing)

5. Checking Churn reasons based on data

    Top churn reasons are:

    1) Attitude of support person (~10%)

    2) Competitor offered higher download speeds (~10%)

6. Checking duplicate rows

    No duplicate rows in dataset

7. Checking initial churn

    Initial analysis showed around 26.5% churn, indicating a significant business risk.

Data Cleaning & Preprocessing

1. Handling inconsistent data types

    Converting Total Charges from object to numeric type

2. Handling missing values

    Filling nan in "Total Charges" with 0 since those customers have "Tenure Months" = 0 but "Contract" = Two Years

3. No duplicate rows present
4. No inconsistent values present in any column

```
for col in df.columns:
   print(col,'\n', df[col].unique(),'\n\n')
```

Exploratory Data Analysis (EDA)

1. Target Variable Analysis (Churn Rate)

    26.5% of customers churned, indicating a high retention risk.

	proportion
Churn Label	
No	73.5
Yes	26.5

    dtype: float64

2. Demographic Analysis vs Churn

    *1. **Senior Citizens are High-Risk***
    While churn for non-seniors is relatively low (approx. 23%), it jumps significantly to over 40% for Senior Citizens. This is the most dramatic churn difference among the categories shown.

    *2. **Family Stability Reduces Churn***
    Customers with deeper household ties are much more likely to stay:

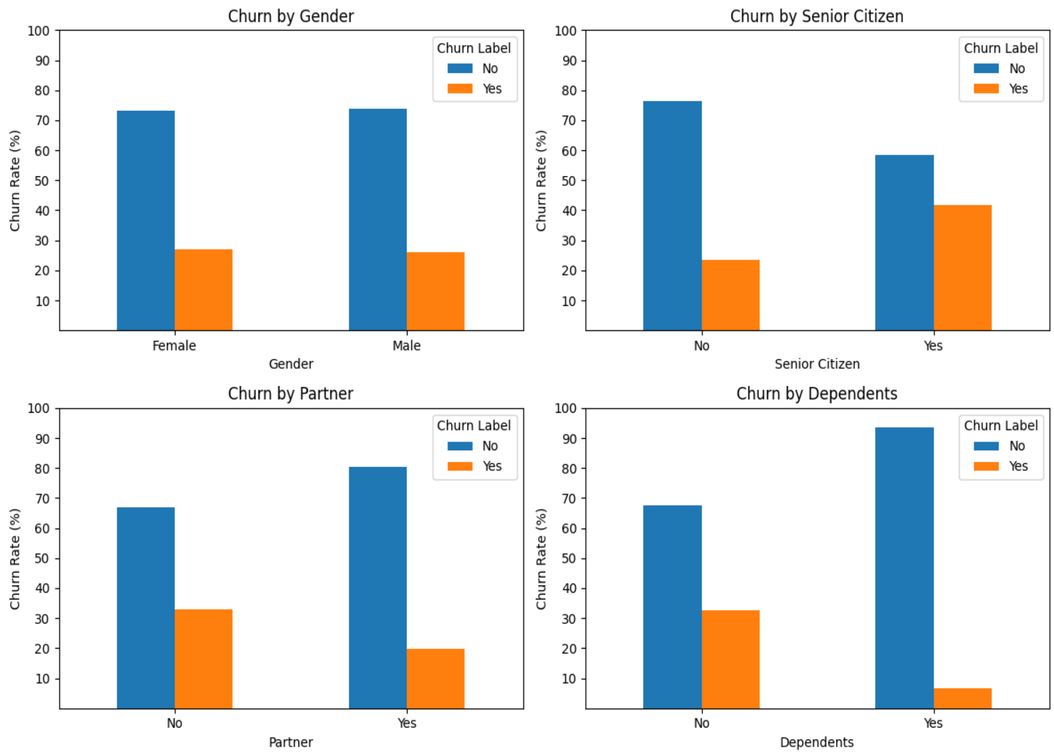
    Dependents: Those without dependents churn at a rate of roughly 33%, whereas those with dependents have a very low churn rate (approx. 7%).

    Partners: Customers without a partner churn at around 33%, while those with a partner churn at only 20%.

    *3. **Gender is Negligible***
    The churn rates for Females and Males are almost identical (approx. 26–27%). This suggests that gender-based targeting is likely not a primary lever for reducing churn in this specific dataset.

Category	High Churn Segment	Low Churn Segment	Difference (Gap)
Senior Citizen	Yes (~42%)	No (~23%)	**~19%**
Dependents	No (~33%)	Yes (~7%)	**~26%**

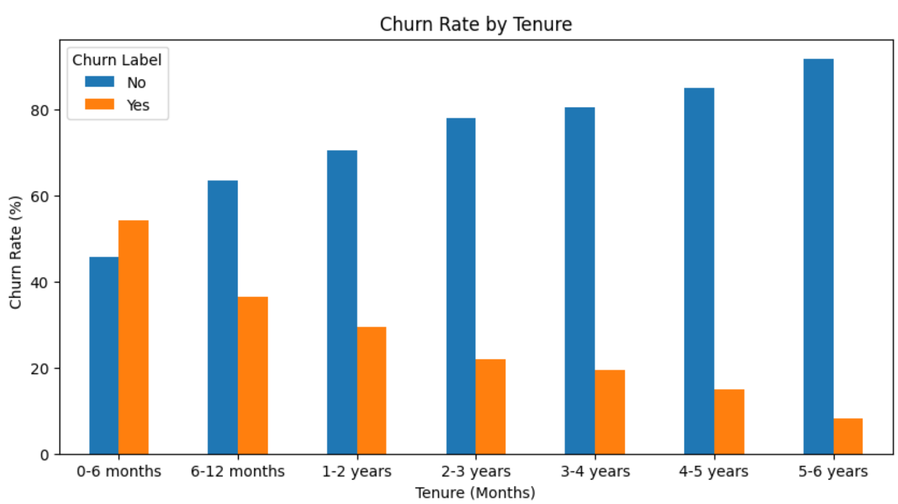
Partner	No (~33%)	Yes (~20%)	**~13%**
Gender	Female (~27%)	Male (~26%)	**~1%**



3. Tenure Analysis

Churn Rate is at peak (~55%) in initial 6 months.

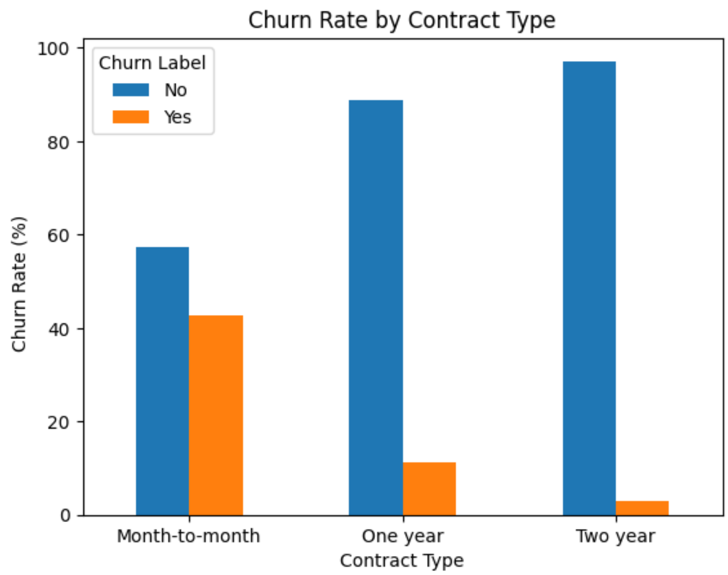
Company should prioritize them and offer rewards for long-term loyalty.



4. Contract Type vs Churn

- *Similar observations as in tenure.*

- *Churn Rate is ~40% in month-to-month contract type.*
- *Incentivize Long-Term Contracts: To reduce the high churn seen in month-to-month plans, offer discounts or exclusive perks to encourage customers to switch to 1 or 2- year commitments*



5. Billing & Charges Analysis

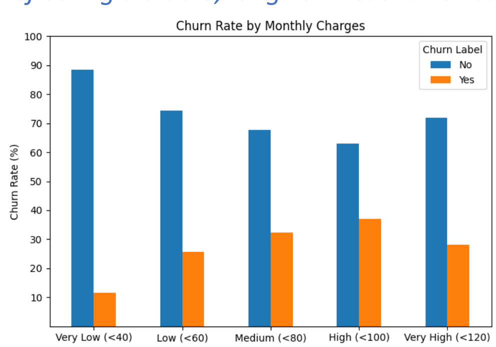
   Monthly Charges:

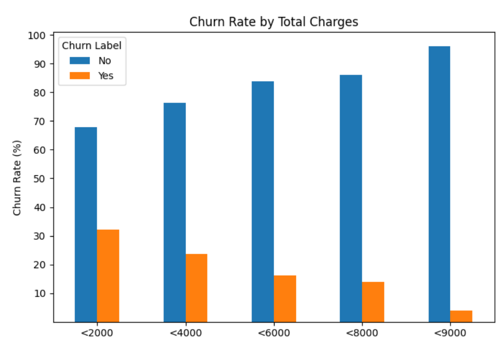
   Churn increases as monthly charges rise, peaking in the "High (<100)" bracket at nearly 40%.

   Total Charges:

   Accumulated Value: The churn rate is highest for those with the lowest total lifetime spend (under 2000).

   Inverse Relationship: As total charges accumulate over time, the churn rate drops consistently. Customers with over 9000 in total charges show the highest retention, reflecting a stable, long-term customer base.

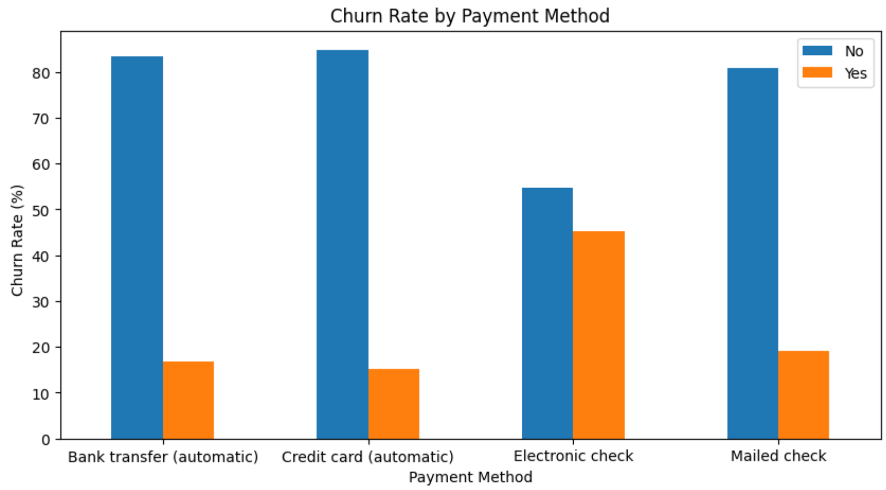




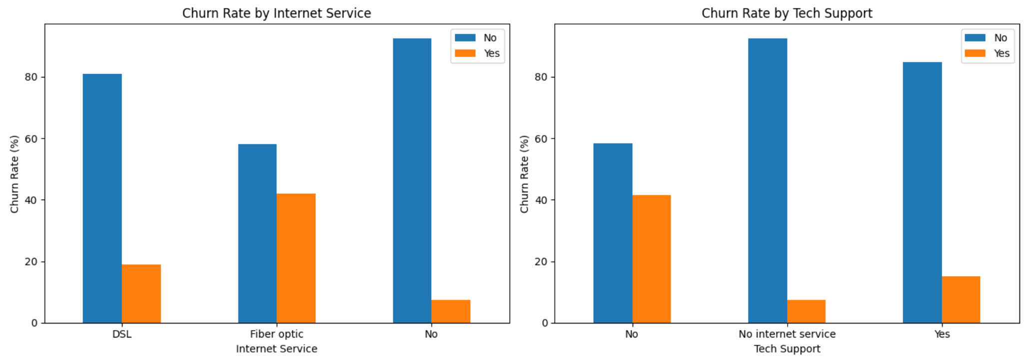
6. Payment Method vs Churn

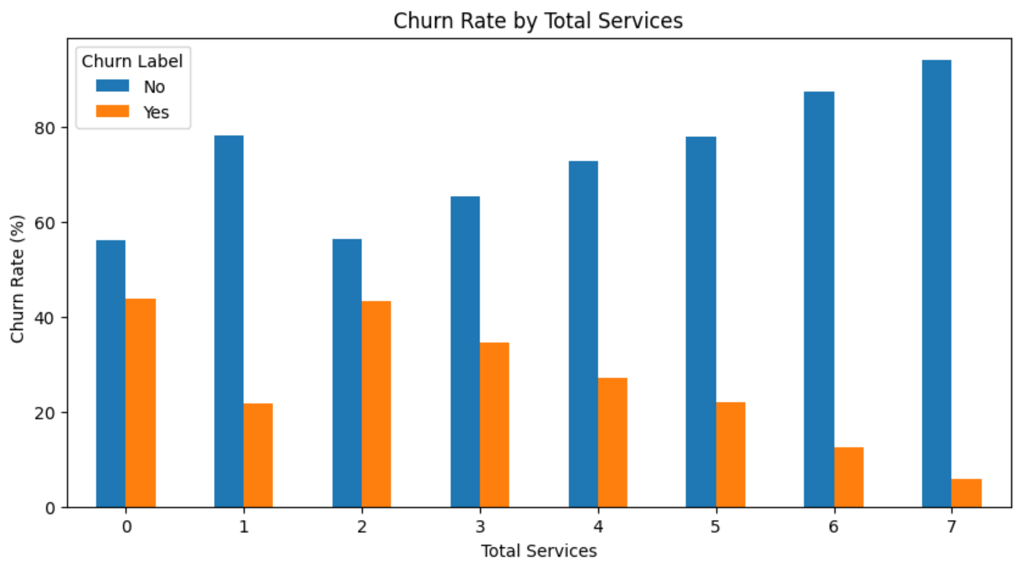
   Payment friction or behavior appears to be a major signal for attrition.

- *Manual vs. Automatic: Customers using Electronic checks show a disproportionately high churn rate of approximately ~45%.*
- *Stability in Automation: Those on Credit card or Bank transfer (automatic) are significantly more loyal, with churn rates dropping below ~20%.*



7. Service Usage vs Churn
    - *Customers using Fiber optic internet churn at a much higher rate (over 40%) compared to DSL users.*
    - *Those without Tech Support are significantly more likely to leave than those who utilize it.*
    - *There is a clear inverse relationship between the number of services a customer uses and their likelihood to leave*
      churn drops from over 40% for those with 0–2 services to roughly 5% for those with 7 services.





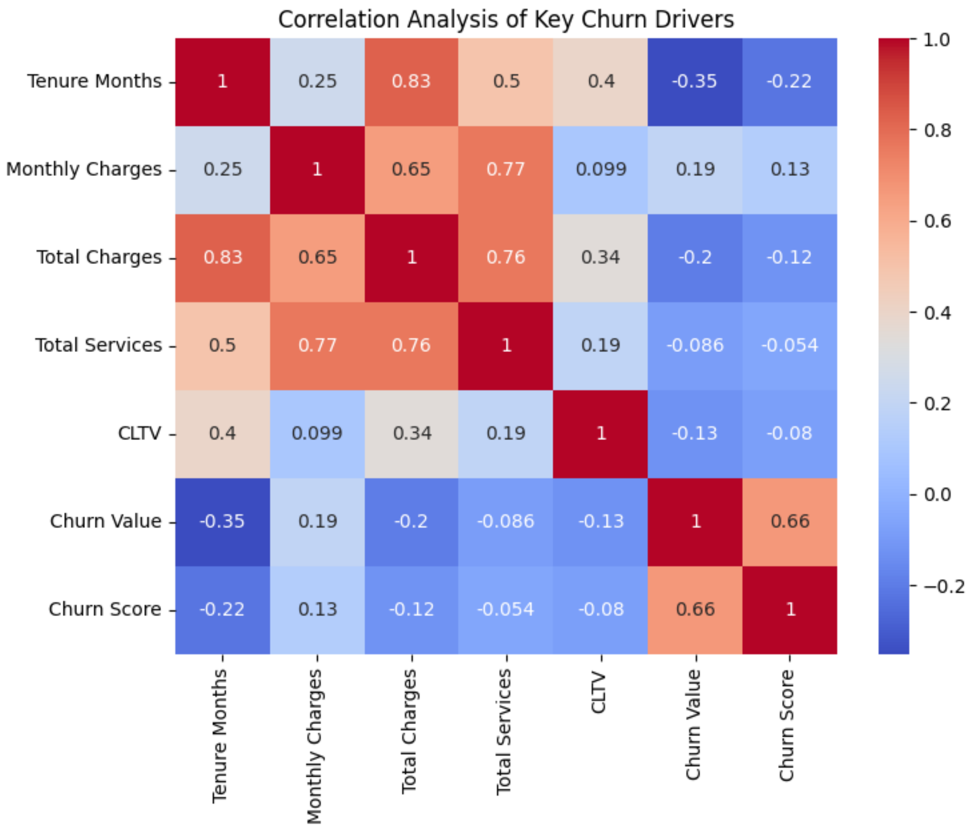
8. Correlation Analysis

*1) **Tenure vs. Total Charges** :0.83 (Strong +)*
Long-term customers are the primary revenue drivers.

*2) **Monthly Charges vs. Total Services**: 0.77 (Strong +)*
Upselling more services directly increases monthly revenue but may increase churn risk.

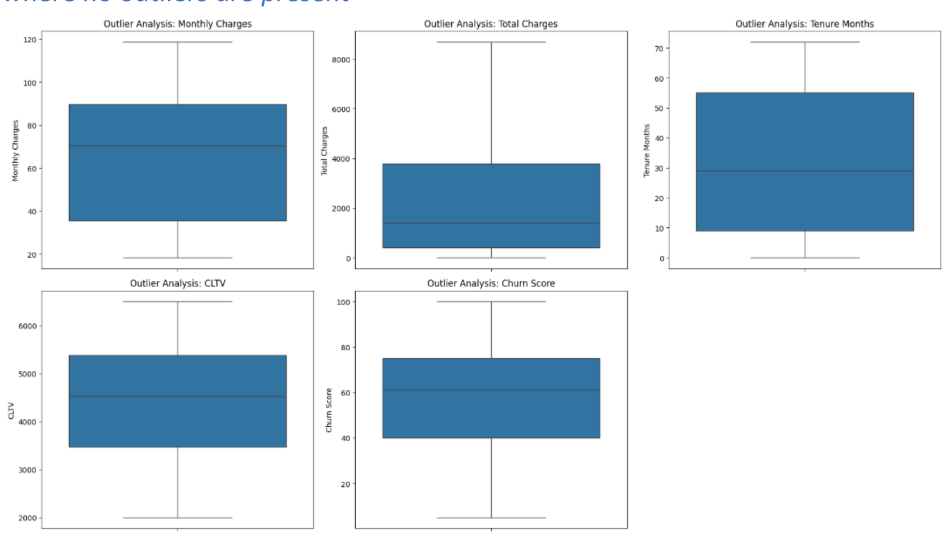
*3) **Churn Score vs. Churn Value**: 0.66 (Strong +)*
Churn prediction model is reliable and should be used for proactive outreach.

*4) **Tenure vs. Churn Value** : -0.35 (Moderate -)*
The risk of churn drops as a customer stays longer; focus on the onboarding phase.



9. Outlier Detection

High impact columns include ['Monthly Charges','Total Charges','Tenure Months','CLTV','Churn Score'] where no outliers are present



ML Modeling

1. Drop Identifier & Leakage Columns
2. Define Target & Features
3. Define Column Groups
4. Train-Test Split
5. Preprocessing Using ColumnTransformer
6. Create Pipeline with Model
7. Train Model
8. Model Evaluation
9. Random Forest Pipeline

    After splitting the data, I used a ColumnTransformer inside a pipeline to apply scaling to numerical features and encoding to categorical features. This ensured no data leakage and made the workflow production-ready. The model was trained and evaluated using recall and ROC-AUC since identifying churners is more critical than overall accuracy.

10. **Model Performance Insights (Customer Churn Prediction)**

    - The churn prediction model achieved 81% accuracy, indicating strong overall predictive capability.

    - A ROC-AUC score of 0.85 confirms that the model effectively distinguishes between churn and non-churn customers and performs significantly better than random guessing, making it reliable for decision-making.

    - For non-churn customers (Class 0), the model shows high effectiveness with 85% precision, 89% recall, and an F1-score of 87%, ensuring loyal customers are correctly identified and unnecessary retention efforts are minimized.

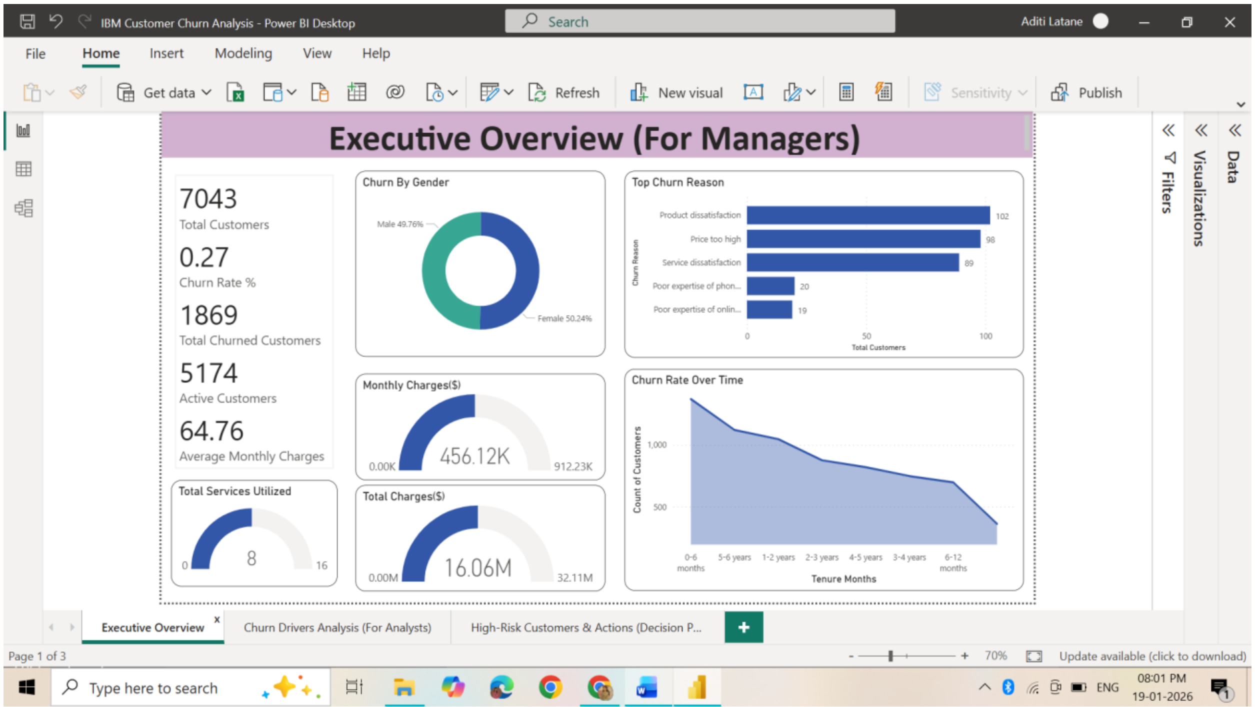
    - For churn customers (Class 1 – critical class), the model achieves 66% precision, 57% recall, and an F1-score of 61%. While the model successfully identifies high-risk customers, improving recall could help capture more potential churners.

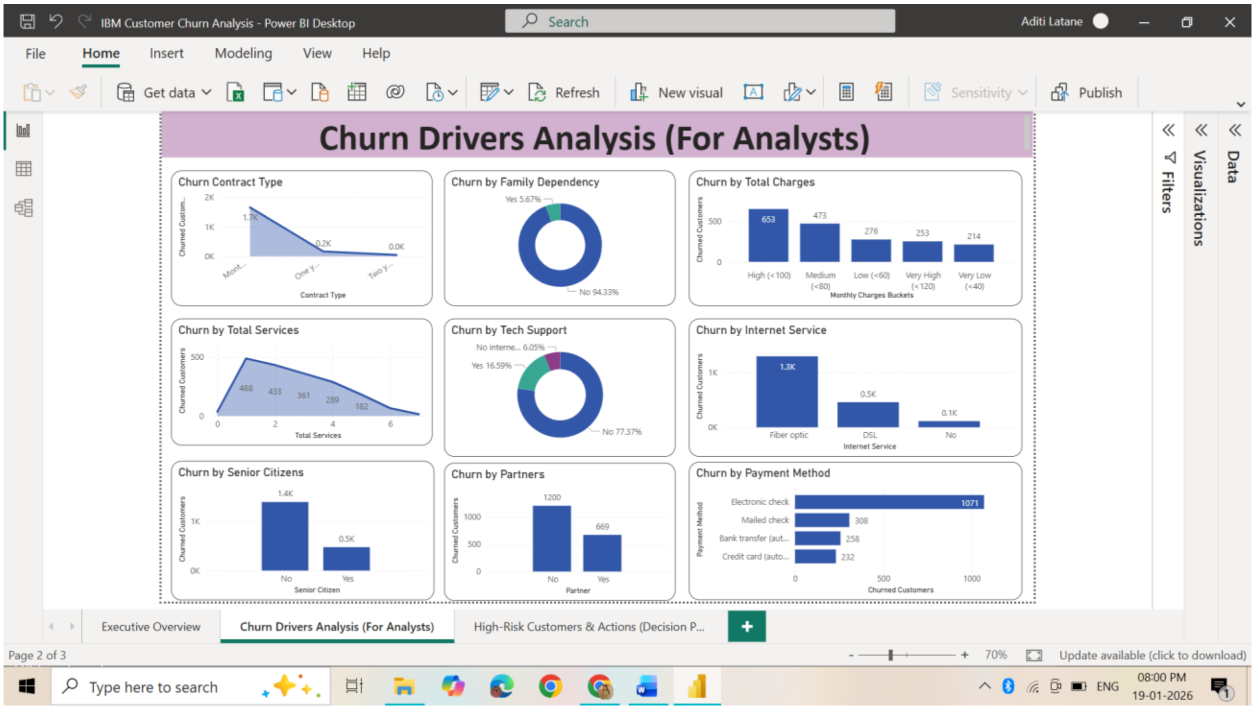
    - The dataset shows class imbalance, with significantly more non-churn customers than churn customers. This explains the stronger performance on non-churn predictions and the comparatively lower recall for churn cases, which is common in real-world business data.

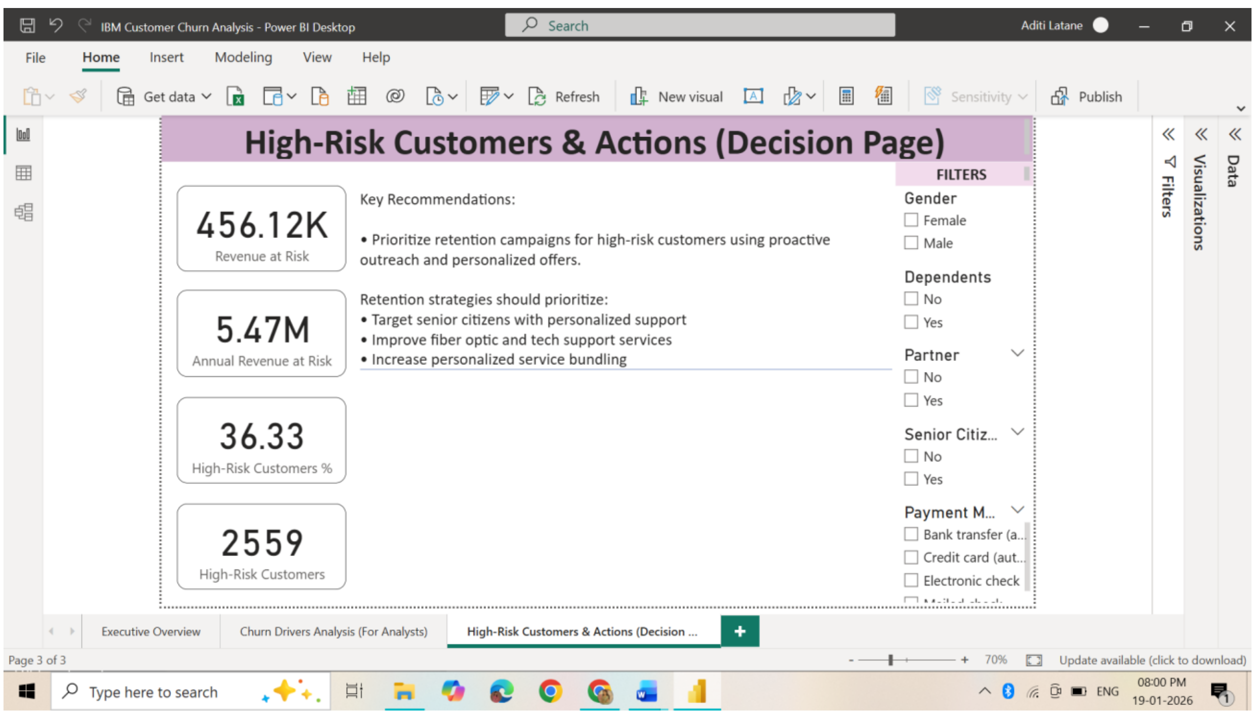
Business Recommendations

Sr. No.	Recommendation	Reason	Actions
1	Strengthen early-tenure retention (0–6 months)	Churn peaks at ~55% during the first 6 months, indicating high onboarding risk	Launch structured onboarding programs, offer early loyalty rewards, provide proactive support during initial months
2	Incentivize long-term contracts	Month-to-month customers show ~40% churn compared to lower churn in long-term contracts	Provide discounts or perks for switching to 1- or 2-year contracts, target high-risk month-to-month users
3	Target senior citizens with personalized support	Senior citizens have the highest churn rate (~42%) among demographics	Offer simplified plans, dedicated support, and senior-friendly communication
4	Promote auto-pay payment methods	Electronic check users churn at ~45%, while auto-pay users churn below 20%	Encourage credit card/bank auto-debit, offer incentives for auto-pay enrollment
5	Improve fiber optic and tech support services	Fiber optic users churn over 40%, and lack of tech support increases churn	Enhance service quality, bundle free or discounted tech support, proactively assist fiber users
6	Increase service bundling and cross-selling	Churn drops from >40% (0–2 services) to ~5% (7 services)	Create personalized service bundles, offer bundle discounts instead of single-service plans
7	Use churn prediction model for proactive retention	Model has ROC-AUC of 0.85, making it reliable for identifying churn risk	Prioritize retention campaigns for predicted churners, reduce unnecessary offers to loyal customers

Visualization & Dashboard




IBM Customer Churn Analysis - Power BI Desktop
Search
Aditi Latane
File
Home
Insert
Modeling
View
Help
Get data
Refresh
New visual
Sensitivity
Publish
Churn Drivers Analysis (For Analysts)
Churn Contract Type
Contract Type
Churn by Family Dependency
Yes 5.67%
No 94.33%
Churn by Total Charges
653
473
276
253
214
High (<100)
Medium (<80)
Low (<60)
Very High (<120)
Very Low (<40)
Monthly Charges Buckets
Churn by Total Services
Total Services
Churn by Tech Support
No interne... 6.05%
Yes 16.59%
No 77.37%
Churn by Internet Service
1.3K
0.5K
0.1K
Fiber optic
DSL
No
Internet Service
Churn by Senior Citizens
1.4K
0.5K
No
Yes
Senior Citizen
Churn by Partners
1200
669
No
Yes
Partner
Churn by Payment Method
Electronic check
Mailed check
Bank transfer (aut...
Credit card (auto...
1071
308
258
232
Churned Customers
Filters
Visualizations
Data
Executive Overview
Churn Drivers Analysis (For Analysts)
High-Risk Customers & Actions (Decision P...
Page 2 of 3
70%
Update available (click to download)
Type here to search
ENG
08:00 PM
19-01-2026


IBM Customer Churn Analysis - Power BI Desktop
Search
Aditi Latane
File
Home
Insert
Modeling
View
Help
Get data
Refresh
New visual
Sensitivity
Publish
High-Risk Customers & Actions (Decision Page)
456.12K
Revenue at Risk
5.47M
Annual Revenue at Risk
36.33
High-Risk Customers %
2559
High-Risk Customers
Key Recommendations:
• Prioritize retention campaigns for high-risk customers using proactive outreach and personalized offers.
Retention strategies should prioritize:
• Target senior citizens with personalized support
• Improve fiber optic and tech support services
• Increase personalized service bundling
FILTERS
Gender
Female
Male
Dependents
No
Yes
Partner
No
Yes
Senior Citiz...
No
Yes
Payment M...
Bank transfer (a...
Credit card (aut...
Electronic check
Filters
Visualizations
Data
Executive Overview
Churn Drivers Analysis (For Analysts)
High-Risk Customers & Actions (Decision ...
Page 3 of 3
70%
Update available (click to download)
Type here to search
ENG
08:00 PM
19-01-2026