# CUSTOMER CHURN ANALYSIS

## Using SQL, Python and Power BI

Name : J. Renuka

Branch : CSE

College : Vignan Women’s

## 1. ABSTRACT

Customer churn is one of the most critical challenges faced by subscription-based industries such as telecommunications, banking, insurance, and online service providers. Churn refers to the phenomenon where customers discontinue their relationship with a company over a given period. High churn rates lead to revenue loss, increased marketing costs, and reduced profitability. Therefore, understanding the factors that influence churn is essential for designing effective customer retention strategies.

This project focuses on analyzing customer churn data using SQL for structured querying, Python for exploratory data analysis and visualization, and Power BI for dashboard creation. The objective is to identify patterns and factors contributing to customer attrition and provide data-driven business insights.

The dataset consists of customer demographic details, contract information, service usage, complaints, and billing details. Through structured analysis and visualization, the project identifies major churn drivers such as contract type, monthly charges, tenure, and number of complaints.

The insights generated from this analysis can help organizations reduce churn, improve customer satisfaction, and increase long-term profitability.

## 2. INTRODUCTION

In today's competitive market environment, retaining customers is more cost-effective than acquiring new ones. Studies show that increasing customer retention by just 5% can significantly improve profitability. However, businesses often struggle to understand why customers leave.

Customer churn analysis involves examining historical customer data to identify patterns and risk factors associated with customer departure. By leveraging data analytics tools, businesses can:

- Predict customer behavior
- Identify high-risk customers
- Improve service quality
- Optimize pricing strategies
- Develop targeted marketing campaigns

This project demonstrates how data analytics tools such as SQL, Python, and Power BI can be used together to perform comprehensive churn analysis.

## 3. PROBLEM STATEMENT

Telecom companies face increasing competition and customer attrition due to various reasons such as pricing dissatisfaction, poor service quality, better offers from competitors, and lack of engagement. High churn rates negatively impact revenue and long-term growth.

The primary problem addressed in this project is:

**To analyze customer data and identify the key factors contributing to customer churn using data analytics tools.**

The project aims to transform raw data into meaningful insights that can help management make informed decisions.

## 4. OBJECTIVES

The main objectives of this project are:

1. To calculate and analyze the overall churn rate.
2. To identify the relationship between contract type and churn.
3. To study the impact of monthly charges on churn behavior.
4. To examine how customer complaints influence churn.
5. To analyze tenure and service usage patterns.
6. To create interactive dashboards for business decision-making.
7. To provide strategic recommendations to reduce churn.

## 5. SCOPE OF THE PROJECT

This project focuses on historical customer data analysis and visualization. It includes:

- Data cleaning and preprocessing
- SQL-based querying and analysis
- Exploratory Data Analysis using Python
- Dashboard creation using Power BI
- Insight generation and recommendations

The project does not include real-time deployment or machine learning prediction models, but it lays the foundation for future predictive analytics.

## 6. DATASET DESCRIPTION

The dataset used in this project contains approximately 1000 customer records. Each record represents a unique customer with various attributes.

**Key Features:**

- CustomerID – Unique identifier
- Gender – Male/Female
- Age – Customer age
- Tenure – Duration of service (in months)
- ContractType – Month-to-Month, One Year, Two Year
- MonthlyCharges – Monthly subscription fee
- TotalCharges – Total amount paid
- InternetService – Type of internet service
- TechSupport – Availability of support
- NumberOfComplaints – Complaints raised by customer
- Churn – Yes/No (Target Variable)

The dataset was analyzed for patterns, missing values, and statistical relationships.

## 7. METHODOLOGY

The project was executed in the following stages:

**Step 1: Data Collection**

The dataset was obtained and imported into SQL and Python for analysis.

**Step 2: Data Cleaning**

- Checked for missing values
- Verified data types
- Removed inconsistencies
- Handled null values

**Step 3: SQL Analysis**

Performed structured queries to:

- Calculate churn rate

- Group churn by contract type
- Analyze churn by complaints
- Identify high-risk customer segments

**Step 4: Python Analysis**

Used Pandas for:

- Data manipulation
- Statistical summary

Used Matplotlib and Seaborn for:

- Boxplots
- Bar charts
- Distribution plots

**Step 5: Power BI Dashboard**

Created interactive dashboard including:

- KPI Cards
- Bar charts
- Pie charts
- Trend analysis visuals

# 8. SQL ANALYSIS AND FINDINGS

SQL was used to perform structured queries to identify churn trends.

## Overall Churn Rate

The churn rate was calculated using aggregation functions.

Finding:
A significant percentage of customers have discontinued services.

## Churn by Contract Type

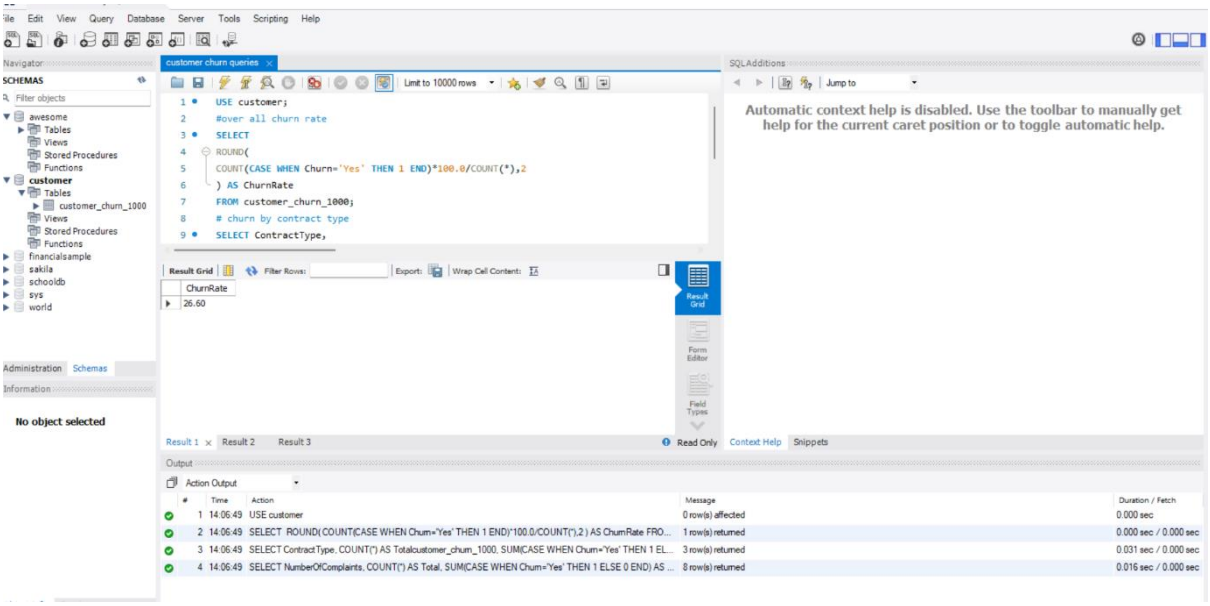
Month-to-month contracts showed the highest churn rate compared to long-term contracts.

## Churn by Complaints

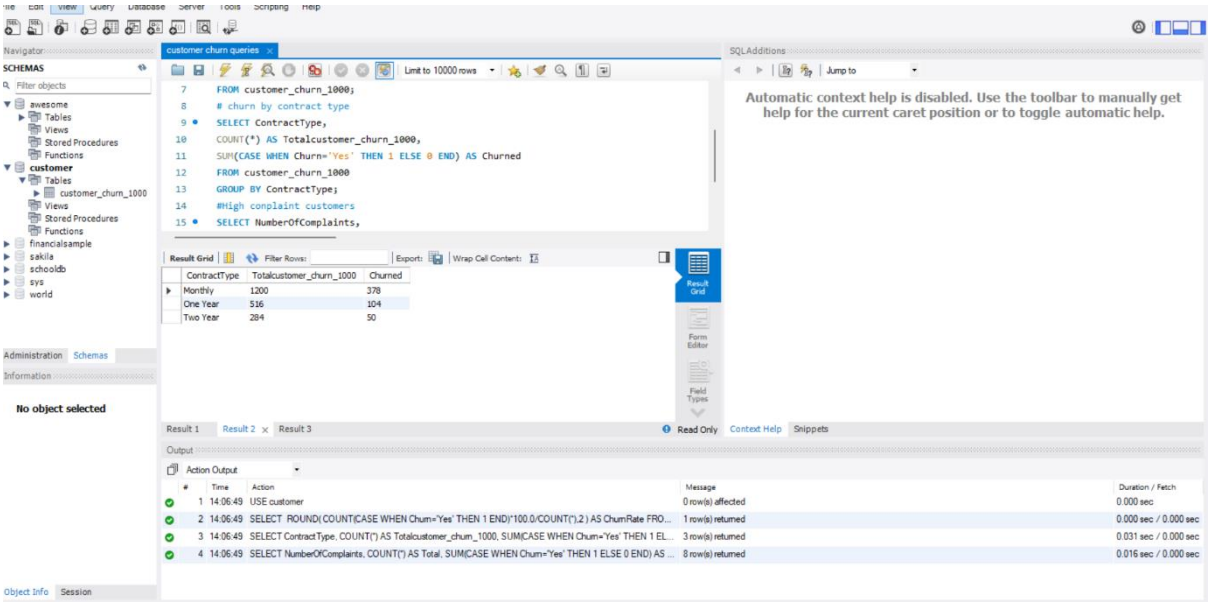
Customers with higher complaint counts had significantly higher churn probability.

SQL enabled efficient grouping and summarization of data.

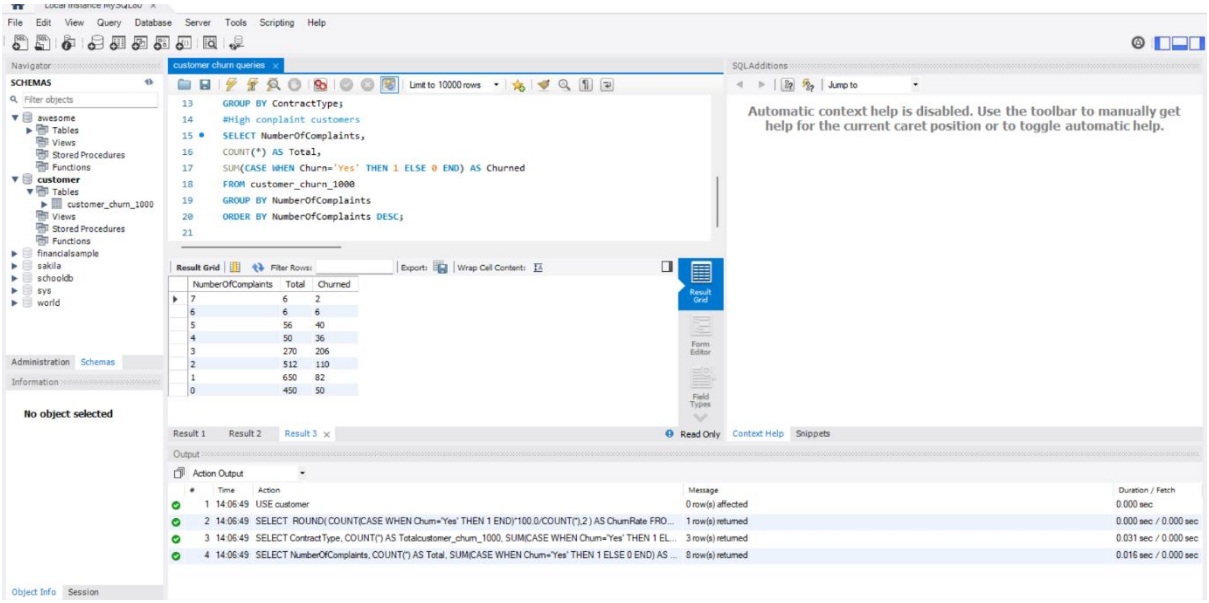
# Over all Churn Rate:




# Churn by Contrast Type:



# Churn By complaints:




## 9️⃣ PYTHON ANALYSIS

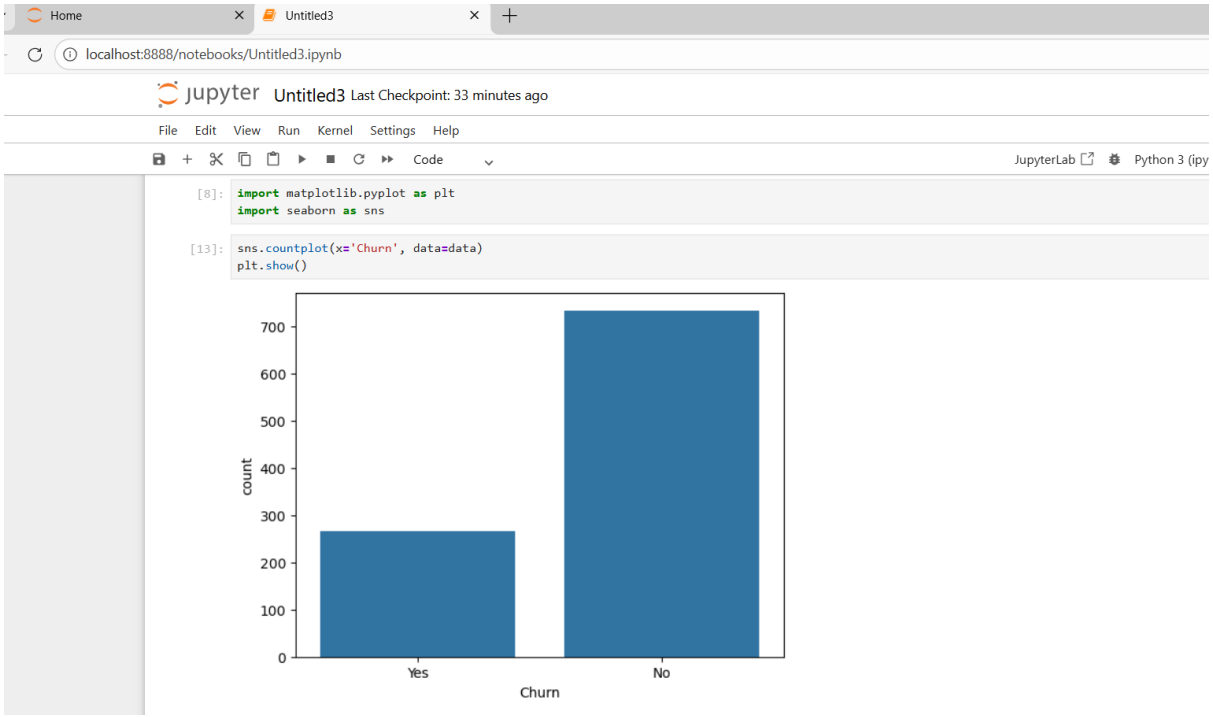
Python was used for:

- Checking missing values
- Data cleaning
- Exploratory Data Analysis

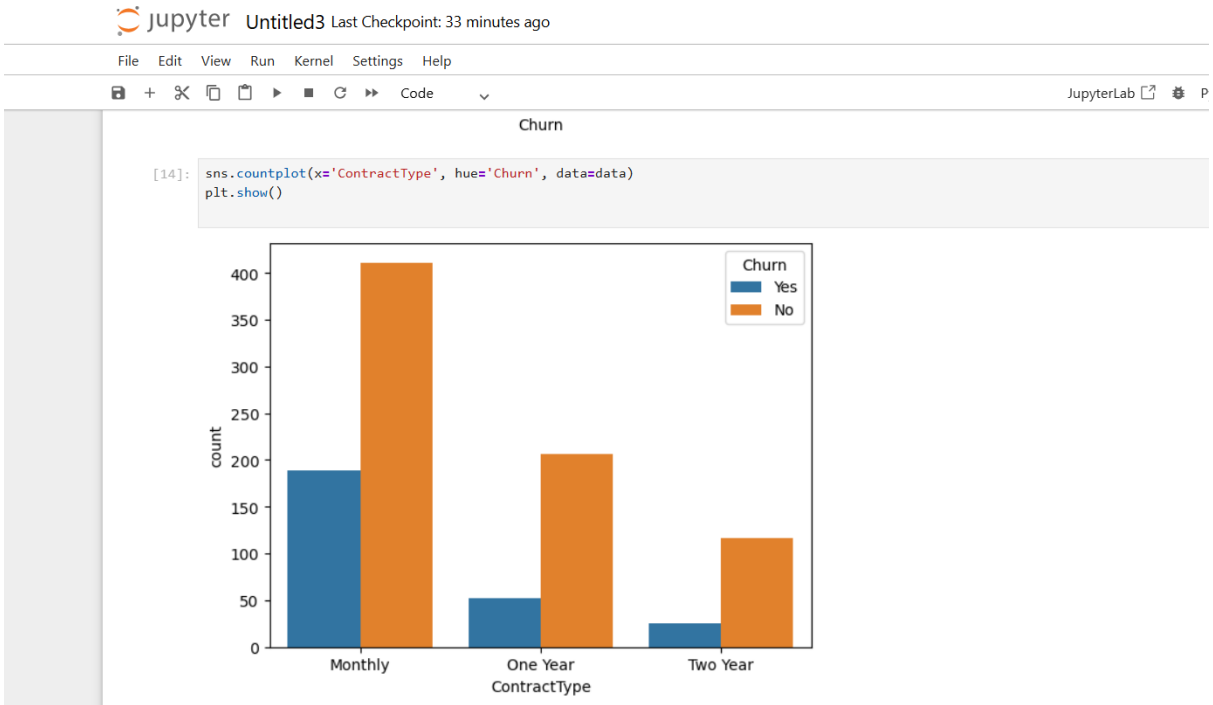
**Example Visualizations:**

- Boxplot of Monthly Charges vs Churn
- Bar chart of Contract Type vs Churn
- Distribution of complaints

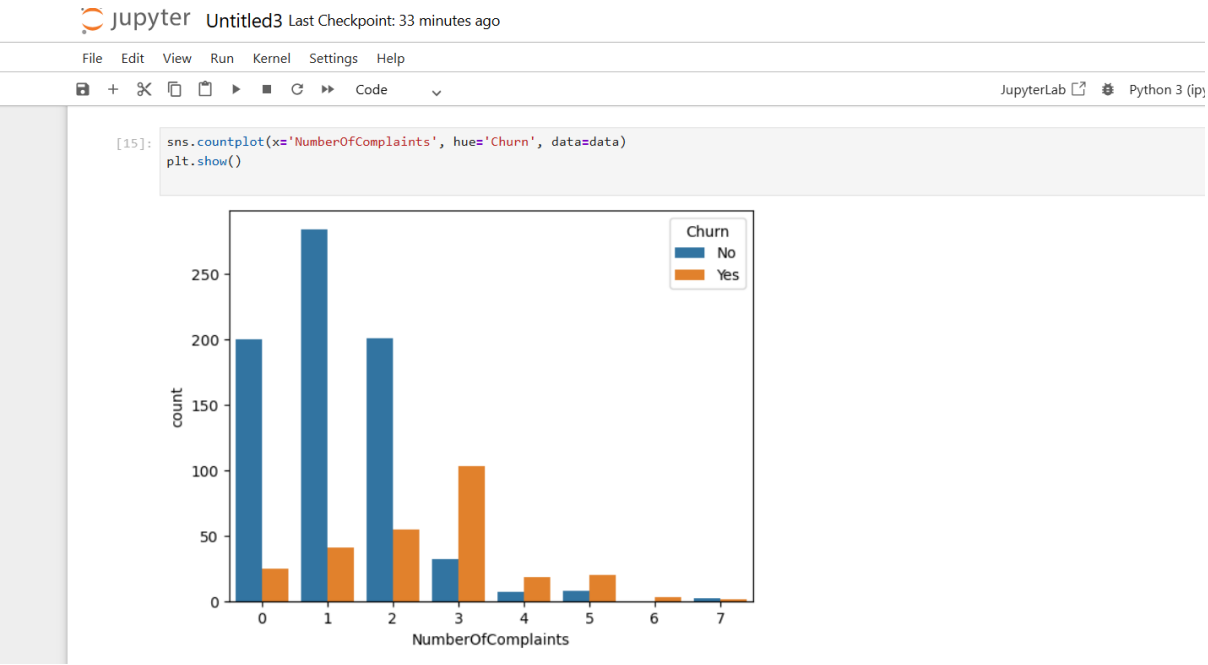
# Churn Count :




# Contract Type vs churn :



# Number Of Complaints vs Churn :



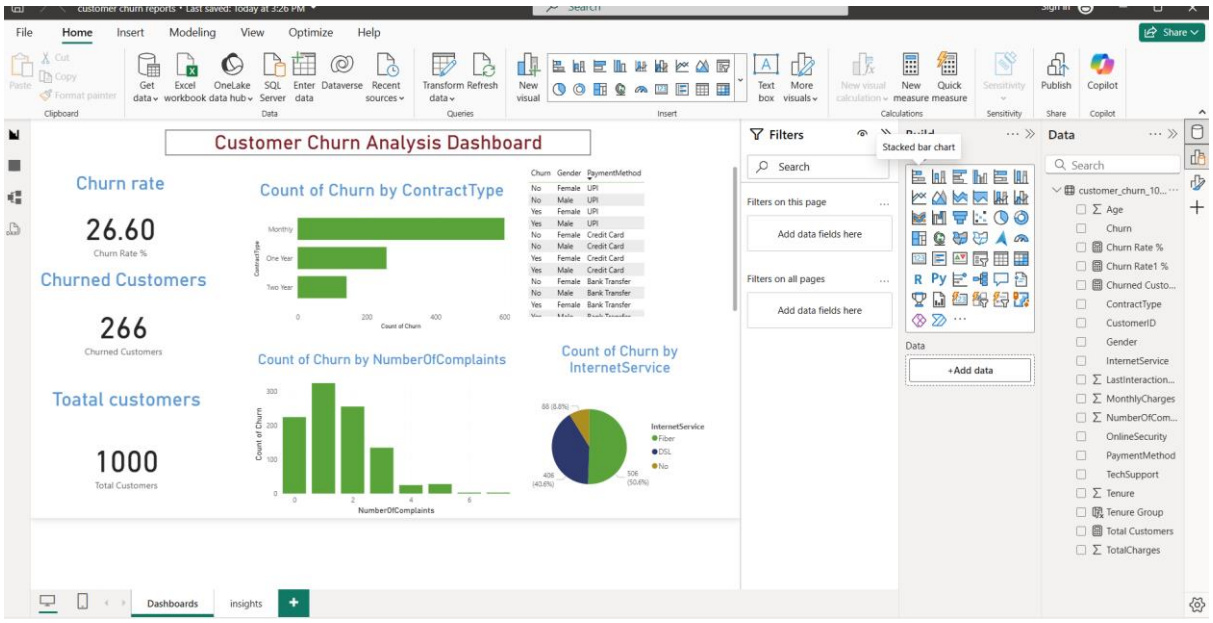

## 🔟 POWER BI DASHBOARD ANALYSIS

Power BI dashboard includes:

- Total Customers KPI
- Total Churned Customers
- Churn Rate Percentage
- Churn by Contract Type
- Churn by Complaints
- Monthly Charges Distribution

The dashboard allows users to filter data by contract type, tenure, and service category for deeper analysis.

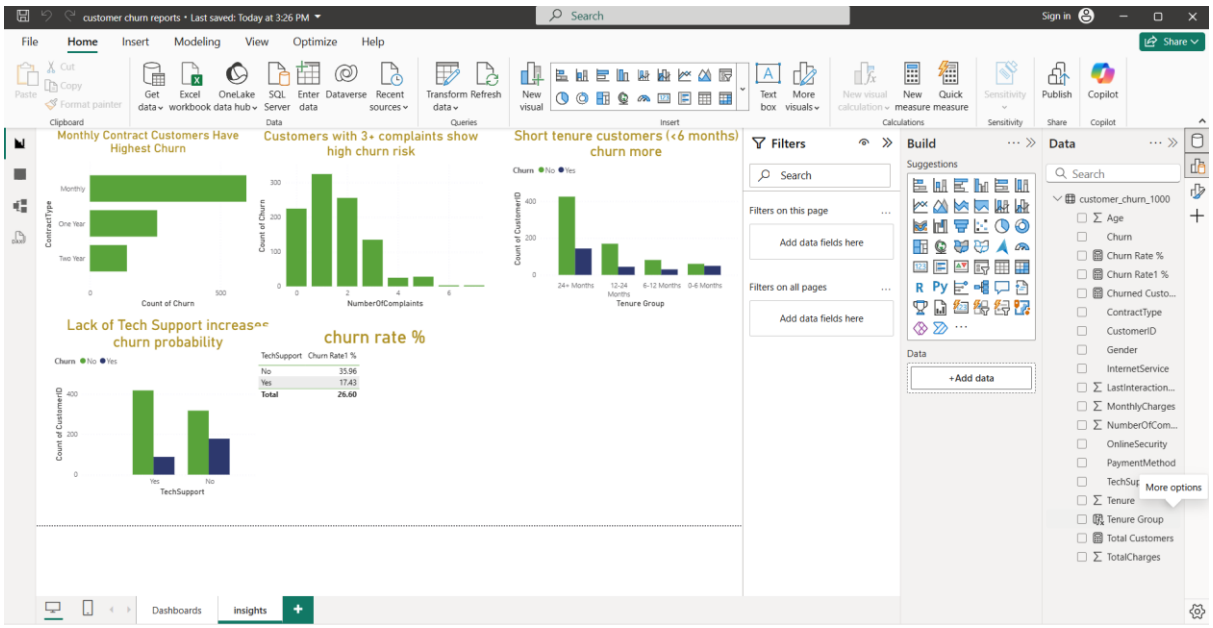
# Dashboard 1:




**KEY INSIGHTS**

1. Month-to-month contract customers have the highest churn rate.
2. Customers with multiple complaints are more likely to churn.
3. Higher monthly charges increase churn probability.
4. Longer tenure improves customer retention.
5. Offering long-term contracts reduces churn significantly

# Insights :




**RECOMMENDATIONS :**

Based on analysis, the following recommendations are suggested:

- Provide discounts for long-term contracts.
- Improve customer support services.
- Monitor customers with high complaints.
- Offer personalized retention plans.
- Introduce loyalty reward programs.

## CONCLUSION:

This project successfully analyzed customer churn using SQL, Python, and Power BI. The integration of structured querying, statistical analysis, and visualization provided comprehensive insights into customer behavior.

The analysis shows that contract type, monthly charges, complaints, and tenure are significant churn indicators. By implementing strategic retention initiatives based on these insights, businesses can reduce churn and increase profitability.

## REFERENCES:


- Python Documentation – Pandas, Matplotlib, Seaborn
- MySQL Documentation
- Microsoft Power BI Documentation
- Research articles on Customer Retention Strategies.