CHURN RISK ANALYSIS USING PYTHON FOR A B2B SAAS SUBSCRIPTION PLATFORM IN THE U.S.

1. Background

A B2B SaaS company offering project management software to small and mid-sized U.S. businesses observed an uptick in customer cancellations over a six-month period. Their internal team had user logs and payment data but lacked the bandwidth to analyze patterns driving churn.

We were brought in to conduct a detailed data analysis using Python tools (Jupyter Notebook and VS Code) to uncover churn signals, identify at-risk customer segments, and deliver a structured report usable by marketing, product, and customer success teams.

2. Objective

- To identify behavior-based, demographic, and usage-driven predictors of churn
- To create a Python-driven churn scoring logic based on customer-level features
- To structure findings in a clean, visually rich report suitable for executive and operational decision-makers
- To provide reusable code for future churn model updates and customer tagging

3. Data Used

Source: Internal customer database and usage logs

Time Frame: April 2022 – March 2023

Dataset Size: 12,500 customer accounts

Key Variables:

- Customer_ID
- Plan_Type (Basic, Pro, Enterprise)
- Join_Date, Churn_Date
- Avg_Weekly_Logins
- Support_Tickets_Last_90_Days
- Time_On_Platform_Per_Session

- Contract_Length_Months
- Account_Manager_Assigned (Yes/No)
- Is_Churned (1/0)

4. Methodology

4.1 Data Preparation

- Merged data from multiple sources (CRM exports, product analytics, and Stripe billing reports)
- Cleaned and standardized missing entries and anomalies (e.g., negative usage time)
- Created derived fields:
    - Tenure_Months = difference between Join_Date and last activity
    - Churned_Within_6_Months flag
    - Normalized login and support ticket activity using Min-Max scaling

4.2 Analysis Performed

- Univariate and bivariate analysis: Churn_Rate by Plan_Type, Churn vs Support Interaction
- Correlation heatmap and churned customer profile segmentation
- Grouped customer behavior into quartiles to isolate high-risk behavioral patterns
- Built a basic churn risk score using weighted combinations of:
    - Low usage
    - High support ticket count
    - Short contract length
    - Lack of account manager assignment

5. Key Results

- **Overall churn rate**: 19.8%
- **Customers on Basic plans** were 3.1× more likely to churn than Enterprise clients
- Churn rate for customers without a **dedicated account manager** was 27.6% compared to 11.9%

- **Logins per week < 2.5** combined with **support tickets > 3** led to 48% churn likelihood
- High churn concentration observed in first **90 days of platform usage**

6. Insights Delivered

- Identified a group of 1,420 active customers with **early churn signals**
- Provided list of top 50 accounts with churn score > 0.75 (flagged for CS outreach)
- Suggested implementation of a **"90-Day Health Check Campaign"** for new users
- Recommended product UI changes based on session length and feature abandonment data

7. Reporting Output

- **Interactive Jupyter Notebook**:
  - Fully annotated, covering data cleaning, analysis, charts, and scoring logic
  - Dynamic cell filters to adjust churn score threshold and extract segments
  - Visual outputs: boxplots, bar charts, stacked churn ratios by plan and activity quartiles
- **PDF Report (Professional Style)**:
  - Executive summary
  - 7 key findings
  - Visual analysis of top churn predictors
  - Appendix with scoring formula and implementation guide
- **Excel File**:
  - Customer_ID with churn score, segment label, and priority tag
  - Sheet for CS team: "Accounts to Contact – Week 1, 2, 3"

8. Business Impact

- Client's CS team launched a proactive outreach strategy using the churn score output
- Shortened response time to usage drops and increased win-back campaign ROI
- Within 2 months, retention improved by **6.2%** for new Basic plan customers

- Python-based scoring logic was integrated into their CRM for real-time monitoring