SUBSCRIPTION CHURN DIAGNOSIS FOR AN ONLINE LANGUAGE LEARNING PLATFORM

1. Background and Problem Statement:

A leading online language learning platform offering English, Spanish, and French courses via monthly and annual subscription plans faced declining user retention, especially within the first three months of onboarding. Despite regular content updates and promotional discounts, monthly churn remained above 29%. The company lacked a structured churn definition, did not track behavior beyond login, and had no visibility into why high-intent users dropped off. This project aimed to quantify churn, analyze behavioral drivers, and design corrective strategies.

2. Objectives:

- Define and measure churn across monthly and annual plans
- Segment customers by churn risk using behavioral and demographic data
- Identify key features predicting early cancellation or drop-off
- Recommend retention improvements in onboarding and subscription engagement

3. Methodology:

3.1 Data Sources and Scope

- **Sample Size:** 67,400 users over 2 quarters
- **Platforms:** Web + Mobile App
- **Tools Used:** R (survival, tidyverse), PostgreSQL, Google Analytics
- Variables collected:
    - Course progress (% completed)
    - Login frequency
    - First course selection
    - Video lesson completion
    - Device type
    - Subscription type

- Last payment date

3.2 Churn Definition

- **Monthly Plan:** Considered churned if subscription not renewed within 5 days of expiry
- **Annual Plan:** Considered churned if no login or lesson activity for 60+ days

3.3 Survival Analysis

- Kaplan-Meier survival curves used to estimate time to churn by cohort
- Log-rank tests run between cohorts (by course, device, subscription type)
- Built a **Cox Proportional Hazards model** to analyze hazard rates across users
- Additional decision tree model used to identify high-risk behavior patterns

4. Results:

4.1 Churn Rate Breakdown

Subscription Type	30-Day Churn	90-Day Churn	Avg. Time to Churn
Monthly	31.2%	49.5%	46 days
Annual	14.8%	27.1%	78 days

- Highest churn observed in **English learners from mobile** using Android
- Students starting with grammar-focused content had **37% higher early churn** than those who began with conversation modules
- Users who completed 3+ lessons in the first week had **4x longer retention span**

4.2 Predictors of Churn (Cox Model)

Variable	Hazard Ratio	Interpretation
Incomplete onboarding (tutorial skipped)	2.47	147% higher risk of churn
No login for 7 consecutive days	2.19	Strong early churn signal
Using Android device	1.42	Slightly higher churn, likely UX-related

Started course but no video completion	2.10	Passive usage correlates with churn

5. Recommendations:

5.1 Onboarding Redesign

- Require 3 interactive actions during onboarding (lesson demo, streak setup, goal selection)
- Auto-trigger an NPS survey and email intervention if no video lesson is completed in the first 3 days

5.2 Early Engagement Campaign

- Launch 7-day progressive challenge to complete 3 lessons and unlock bonus content
- Dynamic emails with course progress and encouragement sent on Day 2, 5, and 7

5.3 Subscription Tier Adjustment

- Introduce **flexible pause option** for monthly users to reduce hard churn
- Add email prompts to annual users after 30-day inactivity with re-engagement offers

6. Future Work:

- Build real-time churn risk scoring integrated with campaign platform
- Incorporate user support ticket sentiment and ratings into churn model
- Launch A/B testing on course structure: grammar-first vs. conversation-first

7. Stakeholder Relevance:

Academic:

- Real-world demonstration of churn survival analysis using Kaplan-Meier and Cox regression
- Useful for case studies in SaaS analytics, customer behavior modeling, and EdTech retention

Corporate:

- Provides a replicable churn segmentation and prediction framework for subscription-based learning platforms
- Aligns user journey design with risk behavior patterns for proactive intervention
- Enhances profitability by reducing preventable early-stage cancellations