Student Engagement Strategies & Drop-Off Analysis

Insights and Recommendations for Excelerate

Team 10: AI Data Powered Data Insights
Date: 10 March 2025

Agenda Overview

- Executive Summary
- Data Analysis Approach
- Data Cleaning Process
- Feature Engineering
- Exploratory Data Analysis (EDA)
- Predictive Modelling
- Model Evaluation
- Churn Analysis
- Drop-off Analysis
- Conclusion and Recommendations

Executive Summary

Project Objectives:

- Enhance student engagement & retention with data insights.
- Analyze impact of engagement time, scores & opportunity duration on churn.
- Provide proactive recommendations for a supportive learning environment.
- Identify key differences between retained & churned students.

DATA ANALYSIS APPROACH

Step 1

Data Cleaning &
Feature Engineering

EDA &
Visualization

Step 2

Step 3

Churn Analysis &
Predictive Modeling

Drop-off Analysis

Step 4

Data Cleaning Process

- **Column Renaming:** Standardized to lowercase with underscores.
- **Data Type Correction:** Consistent date formats.
- **Missing Value Imputation:** Median date for date columns, mode for categorical columns.
- **Country Name Standardization:** Mapping inconsistencies.
- **Institution Name Standardization:** Consolidated variations of institution names.

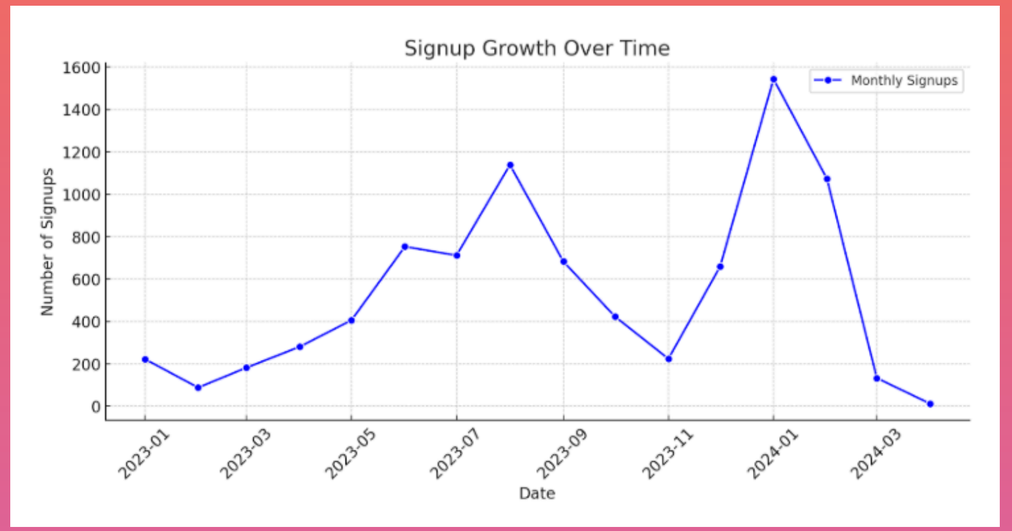
Feature Engineering

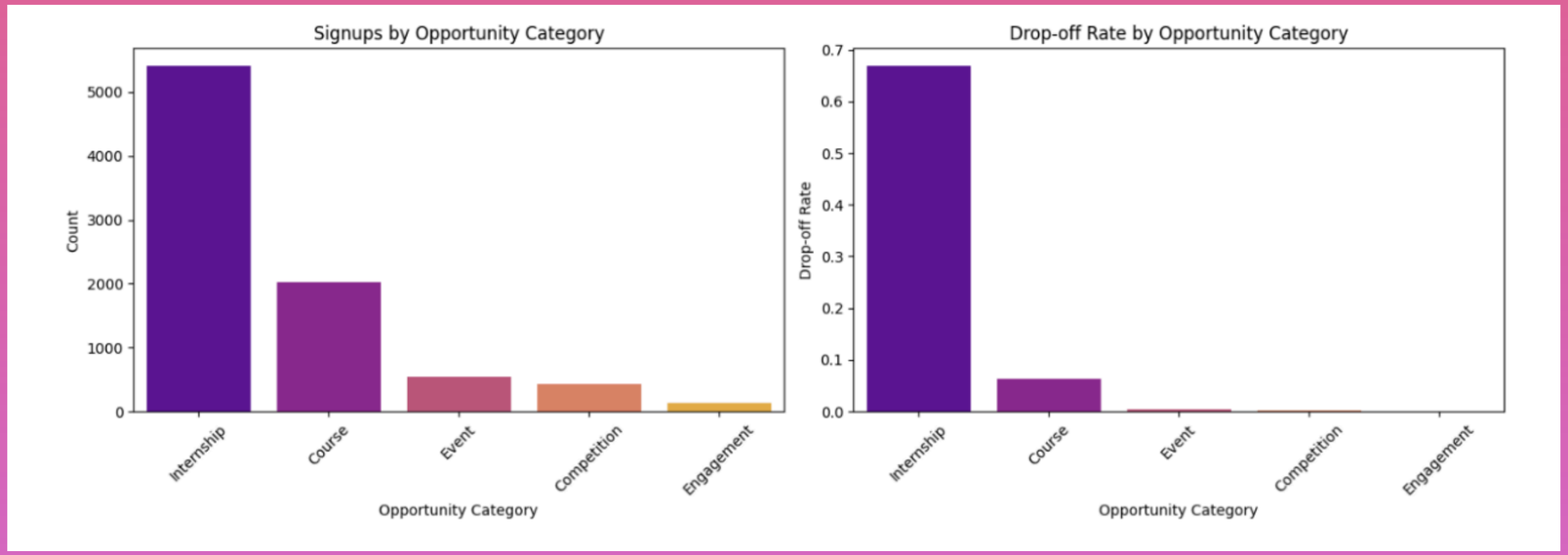
- **Age Calculation:** Age = current year - year of birth
- **Opportunity Duration:** Opportunity duration = opportunity end date - opportunity start date
- **Engagement Time:** Engagement time = opportunity start date - apply date
- **Engagement Score:** Weighted average of opportunity duration, age, and engagement time

Exploratory Data Analysis (EDA)

Signup & Engagement Trends

✅ Internships: Highest signups (5,414) but 66.9% drop-off
✅ Courses: 2nd highest signups (2,023) with 6.4% drop-off
✅ Competitions & Events: Low signups, high completion rates
✅ Peak Signups: January & August ⬆ | Drop-offs: March-April ⬇
✅ Higher Engagement = Higher Completion





Exploratory Data Analysis (EDA)

Demographic Insights

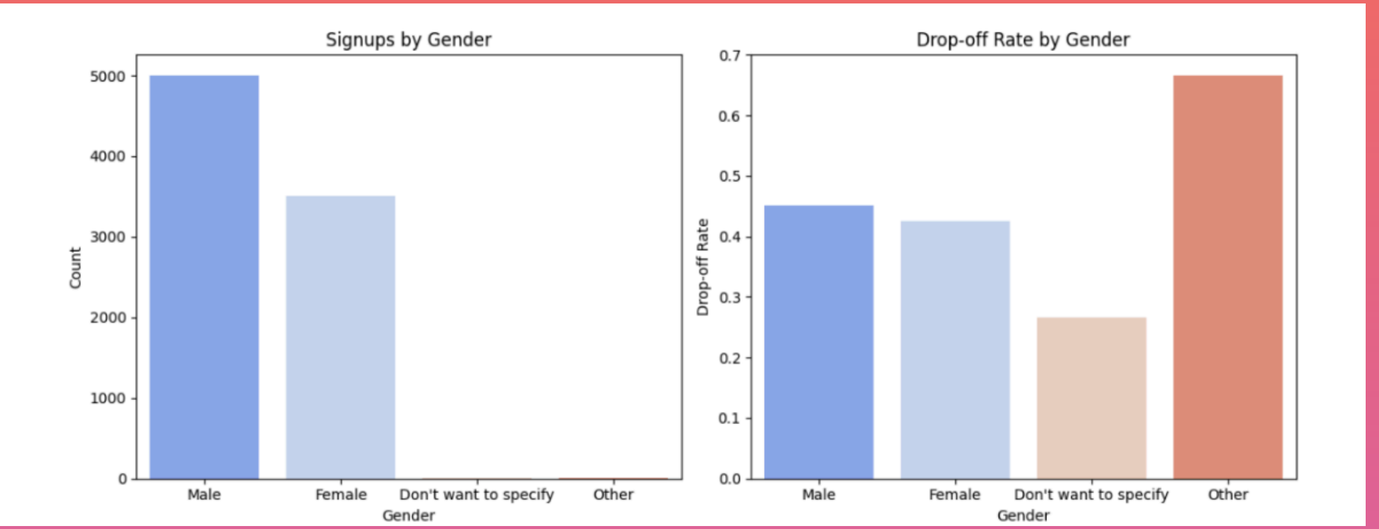
✅ Teenagers (≤19) & 50+ age group: Highest completion

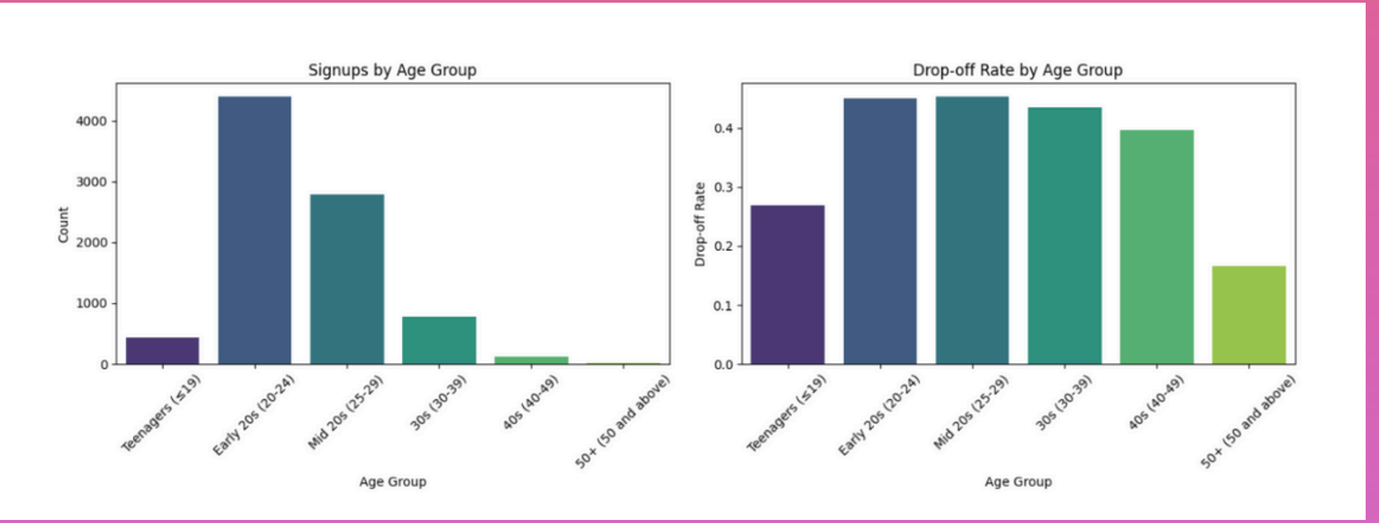
✅ Young Adults (20-29): Highest drop-offs (job/academic pressure)

✅ Females: Higher completion rates than males

✅ Geographic Trends:

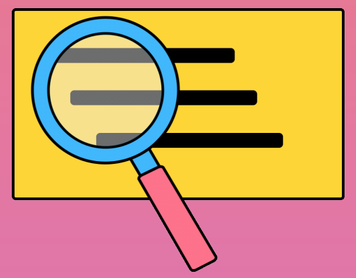
- US: Most signups, highest drop-offs
- Pakistan & Ghana: Strong retention





Predictive Modelling Process

Importing Libraries & Loading the dataset

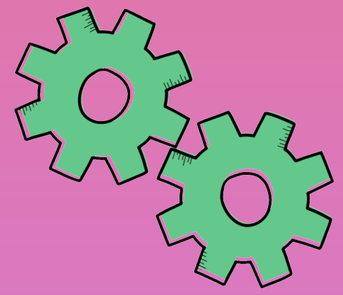


Feature Selection and Data Splitting

Model Selection & Training

Model Evaluation

Hyper-parameter Tuning



Model Comparison & Selection

Model Evaluation

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	99.7%	97.75%	99.67%	99.87%
Decision Tree	100.0%	100.0%	100.0%	100.0%
SVM	95.4%	82.1%	97.5%	87.8%
Logistic Regression	95.6%	82.7%	97.6%	88.3%

Churn Analysis

Objective: Identify key factors contributing to student churn and improve retention.

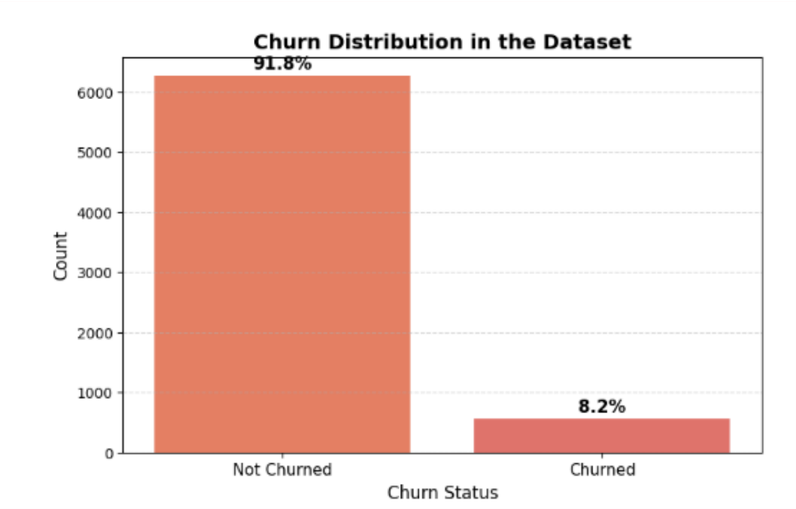
Key Findings:

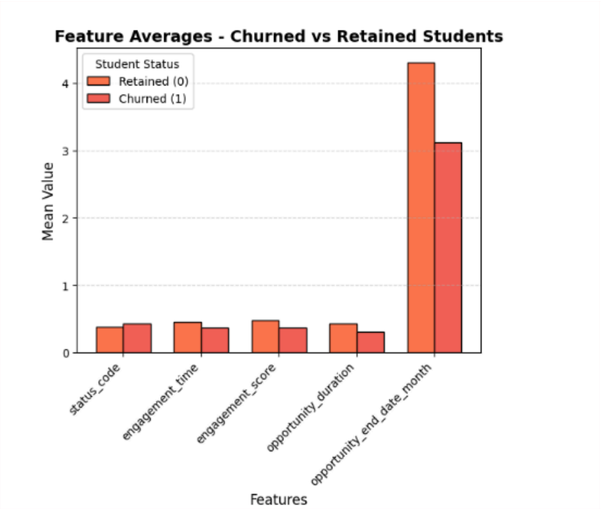
- High engagement reduces churn risk.
- Internship & career-related opportunities see higher drop-offs.
- Personalized learning & proactive interventions can enhance retention.

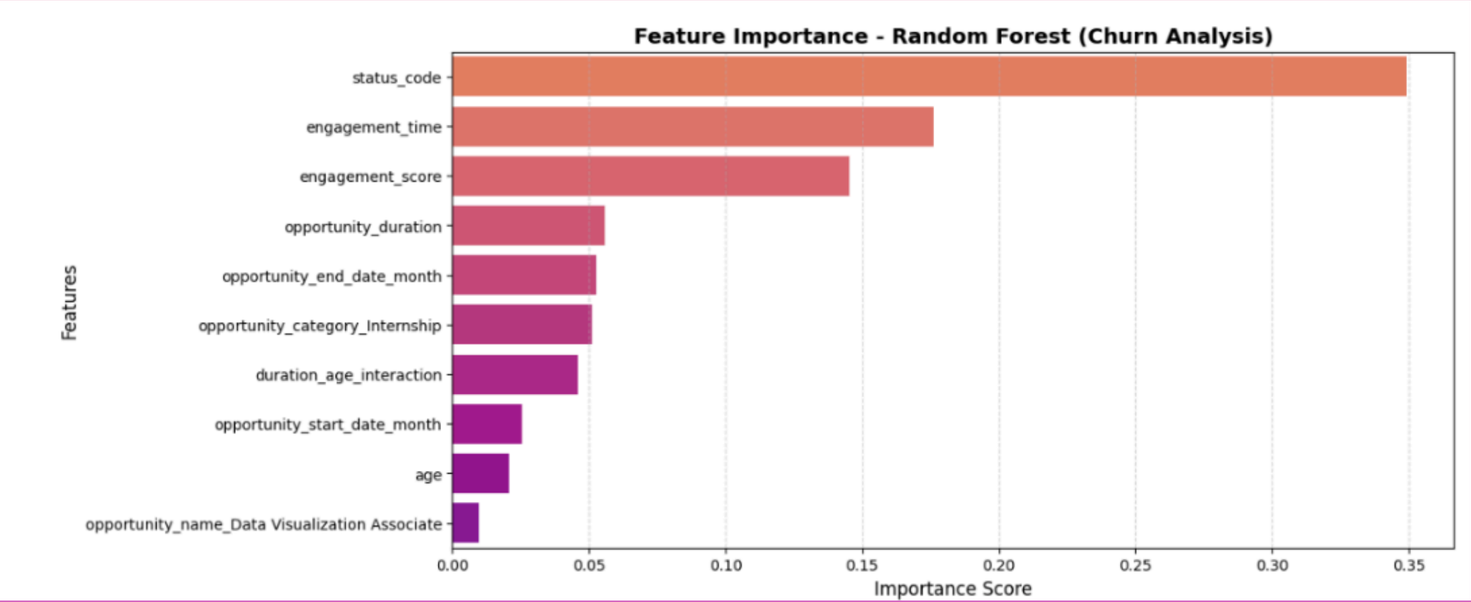
Recommendations:

- Implement data-driven engagement strategies & a basic recommendation system.

Churn Analysis Visualizations

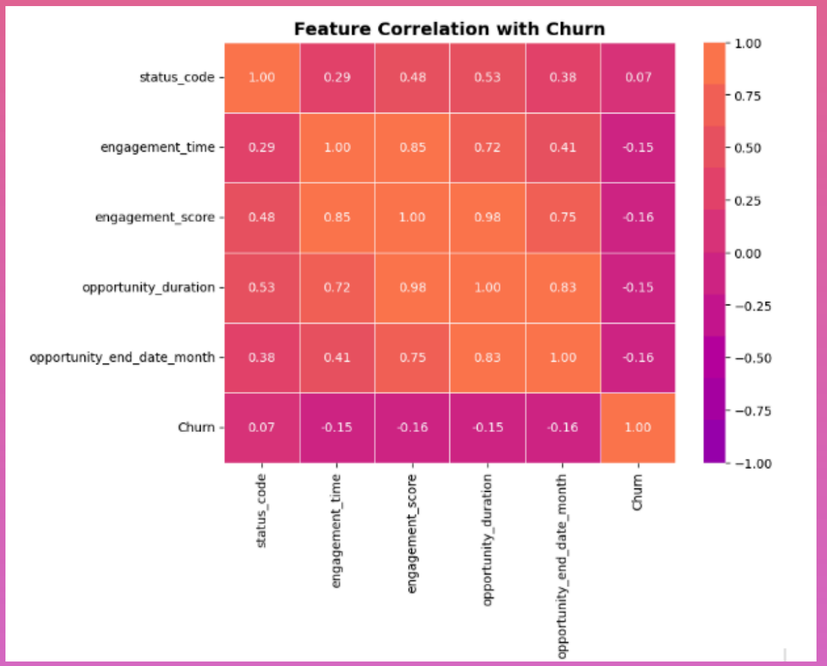






TOP FEATURES IMPACTING CHURN

- **Status Code (34.9%)** – Inactive students have the highest churn.
- **Engagement Time & Scores (32.1%)** – More engagement = lower churn.
- **Opportunity Duration (5-6%)** – Longer opportunities improve retention.



Drop-Off Analysis

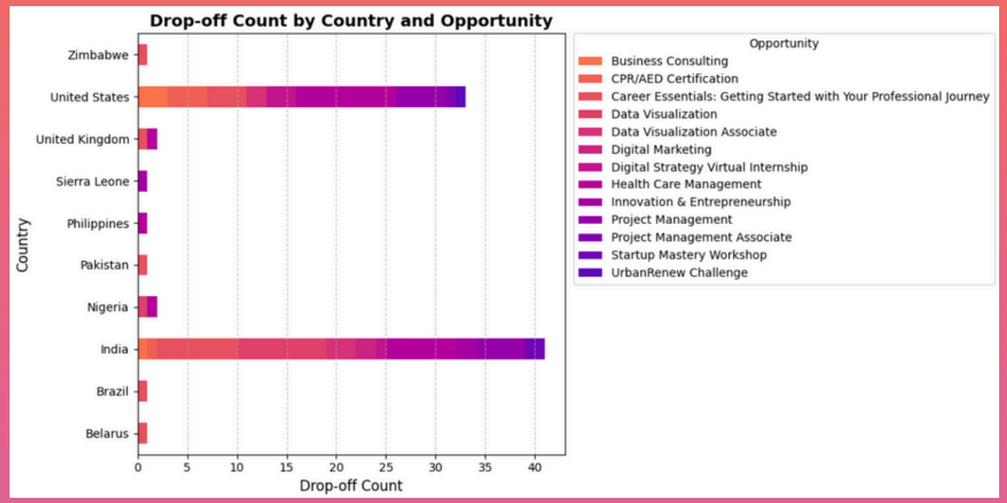
Key Findings:

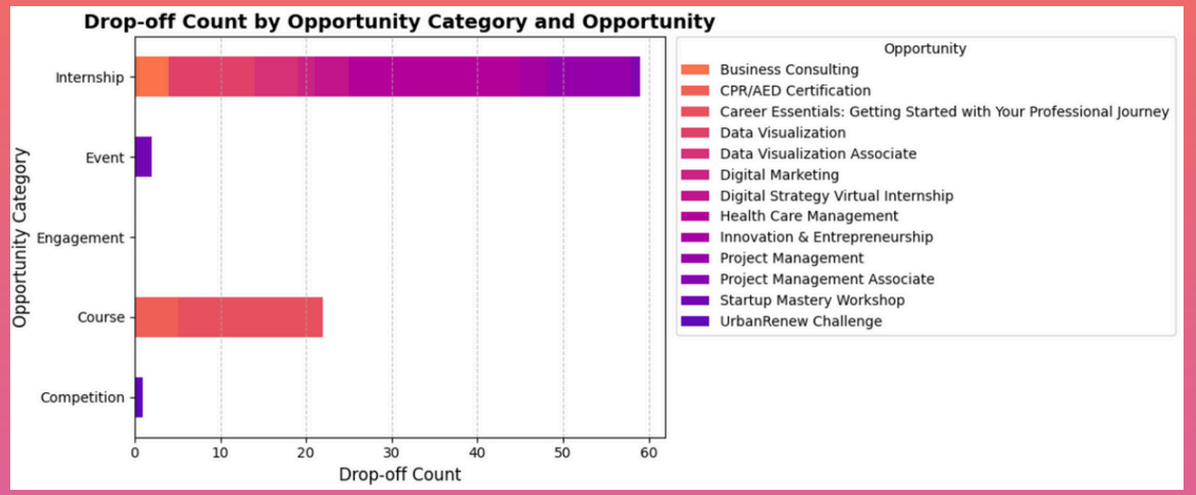
- **Highest Drop-Off:** Internships & Courses.
  - Set clear expectations before enrollment.
  - Provide mentorship & peer support.
  - Introduce progress tracking & rewards.
- **Lowest Drop-Off:** Engagement Opportunities (0%).
  - Leverage engagement strategies.
  - Add gamification, interactive sessions, & community engagement.
- **Age Group:** Teenagers & mid-20s have the highest drop-off rates.
  - Provide structured onboarding, interactive learning, and mentorship.
- **Opportunity Start Date:** Highest drop-offs in March 2023, November 2022, and April 2024.
  - Boost engagement for high drop-off periods through pre-start reminders, orientation webinars, mentorship, and expectation-setting.

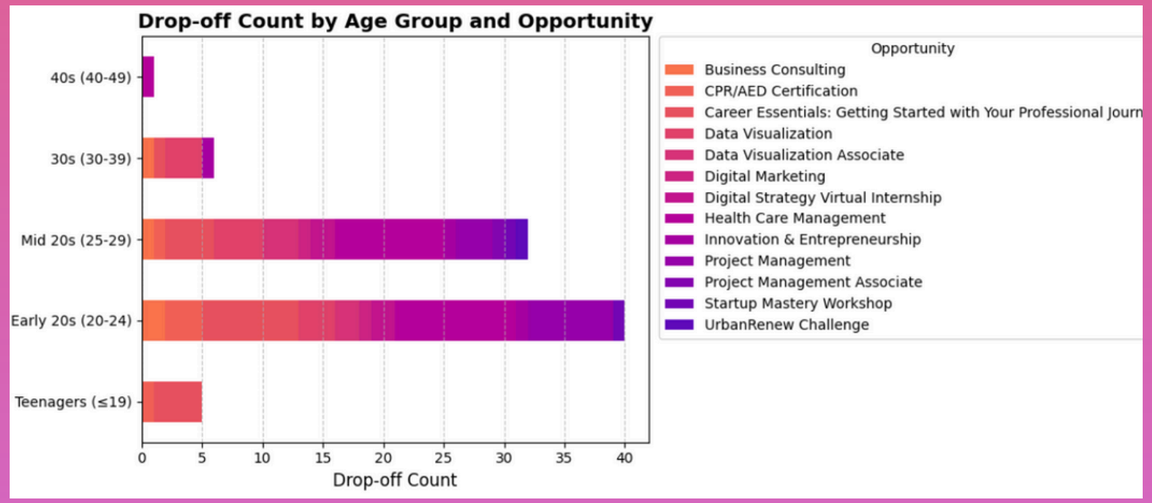
Actionable Recommendations:

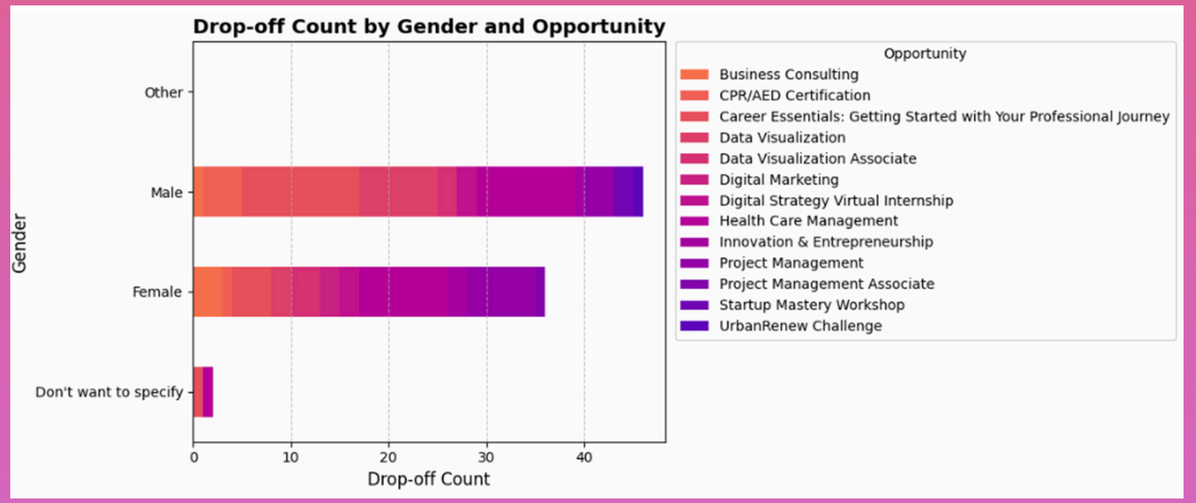
- Improve support during high drop-off periods.
- Use engagement metrics for retention.
- Provide mentorship & structured onboarding for younger learners.
- Optimize start dates & durations.
- Enhance inclusivity & support for underrepresented groups.

Drop-Off Analysis Visualization









Conclusion & Future Impact

Key Impact of this Project:

- Enhanced churn prediction
- Improved student retention
- Personalized learning recommendations

Future Scope & Next Steps:

- Expand Predictive Modeling
- Automate Interventions
- Refine & Scale Recommendations

Call to Action:

Integrate findings into Excelerate's platform to drive long-term student success & engagement!

IMPORTANT LINKS

- **Data Preprocessing & Cleaning**
  - Data Cleaning & Preprocessing Report
  - Cleaned & Preprocessed Data

- **Exploratory Data Analysis (EDA)**
  - EDA & Data Visualization (Colab)
  - Bivariate Analysis - Python Files Folder
  - Univariate Analysis - Insights Report
  - EDA Report

- **Predictive Modeling & Final Model**
  - Student Drop-Off Analysis
  - Consolidated Predictive Models
  - Final Model File

Data isn’t just numbers; it tells a story. And when we listen to that story, we unlock **insights** that drive **real impact.**

THANK YOU!