PLATFORM USAGE-BASED SEGMENTATION USING PYTHON FOR A U.S. SAAS PRODUCT

1. Background

A fast-growing U.S.-based SaaS startup that offers productivity tools for remote teams was struggling with user retention during the first 30 days post-signup. Despite healthy traffic and sign-up volumes, nearly half of their users became inactive by week three.

The leadership team needed a data-driven segmentation model based on actual platform behavior, rather than demographics, to identify how different user types engage with the product and how onboarding could be adjusted for each group.

2. Objective

- To segment users based on in-app behavior during the first 30 days after account creation
- To apply unsupervised clustering techniques using Python for grouping users by behavioral patterns
- To deliver insights into usage frequency, feature adoption, and drop-off points
- To support onboarding redesign efforts that reduce early churn

3. Data Used

Source: Platform activity logs (for 60-day cohort)

Fields Included:

- User_ID, Signup_Date, Feature_Usage_Counts (per module), Session_Count, Avg_Session_Duration, Days_Active, Help_Center_Usage, Early_Upgrade_Flag

Sample Size: 13,200 users **Observation Period**: First 30 days post-signup **Churn Definition**: No login for 10+ consecutive days post day 15

4. Methodology

4.1 Feature Engineering

- Aggregated usage by module: Task Management, File Sharing, Video Calls, Integrations
- Created ratios:

- Avg_Feature_Use_Per_Day, Time_Between_Sessions, Support_Dependence_Index

- Scaled all numeric variables using MinMaxScaler

4.2 Clustering Approach

- Used **Hierarchical Clustering** with Ward linkage to explore natural groupings
- Verified results with **K-Means (k=4)** and **Silhouette Score (0.62)**
- Reduced dimensions for visualization using **PCA**

4.3 Tools Used

- Python Libraries: pandas, scikit-learn, scipy, matplotlib, seaborn

5. Segment Results

Segment Name	Size	Description
Power Users	12%	High usage across features, upgraded early, low support need
Feature Skimmers	29%	Used basic tools only (Tasks, Files), low engagement overall
Support Seekers	22%	Frequent logins but relied heavily on help center
Passive Registrants	37%	Very low engagement, never explored full feature set

6. Interpretation and Strategy

- **Power Users** received early upgrade offers and feedback surveys
- **Feature Skimmers** were shown feature walkthroughs inside the app to boost depth of usage
- **Support Seekers** were offered proactive chat support and onboarding calls
- **Passive Registrants** received personalized win-back email flows with product videos
- Suggested shifting onboarding from a linear tour to a **behavior-triggered path**
- Enabled Product team to tag friction points in low-retention segments

7. Reporting Output

- **Python Script (Jupyter Notebook)**:

    - Data prep, clustering, silhouette scoring, PCA visualizations
    - Segment assignment and export for downstream use
- **PDF Report (14 pages)**:
    - Cluster definitions and feature averages
    - Suggested onboarding modifications
    - PCA plots and retention overlay charts
- **Excel Output**:
    - Segment-wise counts
    - User_IDs with assigned segment
    - Session and churn overlay table

8. Business Impact

- Within 8 weeks:
    - Onboarding revamp reduced early churn by **21%**
    - Feature walkthrough CTR increased by **31%**
    - Customer Success team optimized effort on **Support Seekers**, reducing average query resolution time
    - Segment logic was integrated into their **Mixpanel dashboard** for continuous monitoring