Forms Growth and Analysis

Agenda

I.   Data Analysis and Presentation of Findings
II. Business Intelligence & Dashboarding
III. Marketing

Data Analysis and Presentation of Findings

Process & Assumptions

1. **Getting Started**
   - Imported data and cleaned it (handled missing values, split multi-state entries)
   - Visualized basic stats to identify potential influential factors
2. **Assumptions Made**
   - **Habit formation** is **sending 3 or more envelopes in the both of the last 2 months**
   - Users with missing dates, companies, or multiple states are accurately represented
   - The current features in the dataset are relevant for predicting habit formation
3. **Pattern Identification**
   - Utilized logistic regression to assess feature importance
   - Checked correlations to discern relationships
   - Adjusted for class imbalance to refine model accuracy

Results

- Model **excels at predicting no habit formed**, but needs tuning to adjust for **imbalanced data**
- Balancing (Oversampling minority)
  - ➔ Better identifies habits formed, but with some false positives
- **Accuracy (91%)**
  - Decrease after oversampling due to precision, but still acceptable
- **ROC-AUC Score: 0.9635**
  - strong ability of the model to distinguish between the two classes, regardless of the threshold set for classification

- **Features**

  - **States** - significant impact on habit adoption

  - **User type** - moderate impact on habit adoption

Initial	precision	recall	f1-score	*support*
0 (no habit)	0.97	0.99	0.98	*24981*
1 (habit formed)	**0.7**	**0.39**	0.5	*1346*
accuracy			0.96	*26326*

Feature - State	Coefficient
South Carolina	1.35
Indiana	1.32
North Carolina	1.31
New Jersey	-1.41
Ontario	-1.51
New York	-1.76

Feature - User Type	Coefficient
Breeze User	1.07
Auditor	-0.91
Broker	-1.59
User Not Specified	-2.29

Actionable Insights

1. **Geographical Variation in Product Adoption**
   - States like South Carolina, Indiana, and North Carolina (US-SE) have the highest positive coefficients, indicating a strong positive correlation with habit formation
   - On the flip side, states like New York, Ontario, and New Jersey (US-NE) show strong negative correlations
     - ➔ ***Recommendation:*** *Investigate how marketing and sales approach different states, or how users differ across states (tech adoption, competitor usage rates, etc)*

2. **UserType Impact on Habit Formation**
   - Brokers and Auditors have strong negative correlations with habit formation
   - On the contrary, the Breeze users have a positive correlation
     - ➔ ***Recommendation:*** *Find out why brokers and auditors are less inclined to form habits. Figure out if they are using a different product, have less use cases, or are less likely to adopt technology*

3. **Interaction with Features and Habit Formation**
   - Certain product interactions, such as adding libraries, creating files, and adding forms, have a positive association with habit formation
     - ➔ ***Recommendation:*** *Explore encouraging user interaction with these features to accelerate product habit adoption. Ideas include tutorials, in app suggestions, or customer marketing campaigns*

Additional Recommendations
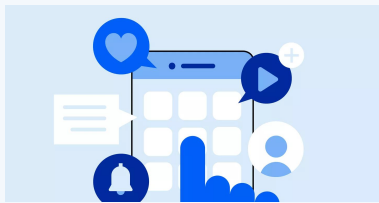
Dive into User Engagement Metrics

- Session duration
- Frequency of use
- Features used

Assess External Factors

- Market analysis: account for local real estate market dynamics
- Competitors: determine if users in specific regions, or specific user types prefer a competitor

User Profiling

- Demographics: age, years in real estate, education
- Tech proficiency
- User timeline

🔗Business Intelligence & Dashboarding