**Overview**

The objective of this project is to maximize profit by determining the optimal CTA combination that effectively encourages potential customers to click on the banner and schedule a meeting. Additionally, gaining insights into how different customer groups respond to various CTA combinations will be instrumental in enhancing overall website performance and user engagement.

**A diagram of a diagram

Description automatically generated**

*The flow guides customers from a webpage to a banner, form, and scheduled meeting, earning a bounty for the company.*

**Data Analysis**

**A graph with blue rectangular bars

Description automatically generatedA blue rectangular bar graph

Description automatically generated**

*Most users did not click the CTA, and an even smaller percentage proceeded to schedule an appointment.*

**Testing the split**

**Split Test Evaluation**A one-way and two-way ANOVA test was conducted to determine the most effective CTA combination.

* One-Way ANOVA: The p-value (<0.05) indicates significant performance differences between CTA combinations. Tukey's HSD test revealed that "Get Pre-Approved for a Mortgage in 5 Minutes Top" consistently outperformed other combinations.
* Two-Way ANOVA: Both CTA Copy and Placement significantly influenced performance, with placement having a stronger effect (lower p-value). However, no interaction effect was observed.
  + Top Placement: Outperformed Middle and Bottom in all scenarios.
  + Recommendations: Default to "Get Pre-Approved for a Mortgage in 5 Minutes" with Top Placement for maximum revenue if personalization isn't feasible. Use creative designs to enhance Middle Placement and avoid Bottom Placement due to weak performance.

A graph with black lines and white text

Description automatically generated with medium confidence

A graph with black lines and white text

Description automatically generated with medium confidence

**Customer Segmentation and User Behavior Analysis**

To understand user behavior, K-means clustering was applied after addressing data imbalance with Adaptive Synthetic Sampling and Chunking techniques.

* Optimal Clusters: Analyzing the SSE curve and Silhouette scores suggested 4-5 clusters, but overlap reduced the meaningful clusters to 3. A graph with a line

  Description automatically generatedA graph with red lines

  Description automatically generated
* Key findings:
  + Cluster 1: High visit counts but a lower click-through rate. Further analysis showed that **device type and scroll depth** significantly influenced this behavior.
    - Chi-Square Test: Device type was found to have a significant effect on click likelihood (p-value < 0.05).
  + Cluster 2: Primarily desktop users who access the site via browsers and favor Google search. They exhibit a higher click-through rate and are more likely to complete the CTA.
  + Cluster 3: Predominantly mobile users with lower engagement, highlighting the need for improved mobile interface design.

Recommendations:

1. Develop a responsive, user-friendly mobile interface to boost Cluster 3 engagement.
2. Tailor messaging for desktop users (Cluster 2) to capitalize on their higher conversion potential.
3. Experiment with interface features that encourage deeper scroll engagement for Cluster 1.

A graph of blue rectangular bars

Description automatically generatedA graph of blue rectangular bars

Description automatically generated

*Cluster 2 has the highest click-through rate (CTR) but the lowest average visit count, while Cluster 1 has the highest average visit count, but a lower CTR compared to Cluster 2.*

**Predictive Modeling for Optimal CTA Placement**

An ensemble model was built to recommend the optimal CTA placement and copy for individual users, maximizing revenue per decision.

* **Model Details**:
  + **Base Models**: Logistic Regression, K-Nearest Neighbors, and Decision Tree.
  + **Ensemble Approach**: A voting classifier combined model predictions, achieving high weighted metrics:
    - **Precision**: 0.893
    - **Recall**: 0.945
    - **F1 Score**: 0.918

**Key Outcomes**:

1. The model identifies the optimal combination of copy and placement for each user, driving personalized experiences.
2. Projected Impact:
   * **Engagement**: Higher click-through and conversion rates by aligning CTAs with user preferences.
   * **Revenue**: Exceeds the revenue generated by the top-performing CTA from Part 1 by leveraging personalized targeting.
3. Revenue Impact
   * Baseline Revenue per customer 12.03%
   * Optimized CTR 89.7%
   * Increased to **$107.93** from the baseline

**Business Recommendations**:

1. Implement the model as an A/B test alongside the current champion CTA to validate the predicted uplift in revenue.
2. Monitor model performance periodically to ensure consistent accuracy as user behavior evolves.
3. Incorporate additional features (e.g., time of visit, referrer source) to enhance the model's predictive power over time.