We would love to understand your data and machine learning technical skills, critical thinking on the business problem and insights at hand, and ability to derive actionable insights. Through this case, please provide a notebook of your choice (Jupyter, Colab, etc.), including your code, your written analysis and explanation for each question, and any visualizations you deem appropriate or relevant. Use appropriate modeling techniques, data analysis and visualization techniques to support your findings. There is an appendix with images and a data dictionary at the end of the doc. Thank you in advance for participating in the case study and enjoy!

## Case Study:

Financial Services stands out as the premier financial authority. When you visit Financialservices.com, the reviews, guides, and educational content have been developed by leading personal finance experts. Financial Services’s product comparison tools, calculators and educational content help over 100 million consumers make smarter financial decisions each year. No matter where you are in your financial journey, Financial Services can help you reach your goals.

One particular area in which Financial Services places a strong emphasis is mortgages, with the goal of advancing the visitor's decision-making journey and ultimately guiding them toward applying for and securing a mortgage. The mortgage team’s goal is to maximize revenue generated from users on our website by getting them to schedule an appointment with one of Financial Services’ mortgage partners. Once the appointment is scheduled, Financial Services gets paid a bounty (revenue). The amount depends on the variation of mortgage chosen for the appointment. For the purpose of this case, there are 4 variations (A, B, C, D), each with their own bounty associated with them.

The customer journey is straightforward. A user visits a page on the website. Once on the page, they have the option to click on a banner that takes them to a form that they can fill out to schedule an appointment with a mortgage lender. They have the option to choose from four mortgage types when they schedule the appointment.

To support this goal, the mortgage team conducted a split test, also called A/B testing, on various mortgage pages. This testing allows the team to compare the performance of website creative variations to see which one appeals more to visitors to maximize a targeted metric. The objective of this test was to learn which banner’s call-to-action (CTA) title copy and on-page placement combination will best entice visitors to click and enter the scheduling form in hope of increasing appointments and overall revenue. *See appendix for visual samples of each CTA and placement variation.*

The test has concluded and we have a data set of 100,000 decisions (rows) of the combination of CTA Copy and CTA placement being served randomly to users coming to the website. To help increase revenue per decision, we want to analyze the data for two purposes. One, we want to see if we have an outright winner in the split test to understand if we should have a new CTA combination that we show everyone. However, the team believes that certain groups of people may have performed differently based on the CTA copy and placement. If proven true and there are patterns in the data, we can use an algorithm to predict in real-time what CTA combination would be best to show to a future user to maximize the key metric.

Below, you will find a set of questions to help guide you in your process. Ensure you complete all 3 parts with a write up for each question. Be mindful of time, you should spend less time on part 1 than part 2 and 3.

### Part 1: Explore the split test. Tell us what happened

Possible things to consider: evaluate relevant metrics for the CTA combinations - which CTA Copy and CTA Placement did the best/worst? If we called one of these CTA combinations our champion (serve it 100% of the time), how much incrementally is that worth to us vs. the average of the rest of the split test?

### Part 2: Tell us about the site users and their behavior

Possible things to consider: Which groups of people tend to be more correlated or less correlated with our key metrics? Are there groups of people who drove higher/lower numbers when engaging with specific CTA copies and placements? What ways can you manipulate the columns/dataset to create features that increase predictive power towards our key metrics?

### Part 3: Predict the optimal CTA to maximize revenue

* Based on the findings from Part 2, build a predictive model that recommends the copy and on-page placement combination for the individual user so that the chances of that user’s revenue per decision is maximized, with the overall goal of creating more revenue per decision than what you found from the top performing CTA combination from part 1. Provide technical and business evaluation on the impact of your optimization, such as how model performance, engagement impact, and revenue impact.

## APPENDIX

### CTA Versions

| Version 1 |
| --- |
|  |
| Version 2 |
|  |
| Version 3 |
|  |

### 

### Page Placements

| Placement 1 - Top | Placement 2 - Middle | Placement 3 - Bottom |
| --- | --- | --- |
|  |  |  |

### Data Dictionary

| **Variables** | **Definition** |
| --- | --- |
| userId | Unique identifier for users visiting Financial Services |
| sessionReferrer | Source from which the user arrived (e.g., search engine) |
| browser | Browser used by the user (e.g., Chrome, Firefox) |
| deviceType | Category of device used by the user (e.g., mobile, desktop, tablet) |
| estimatedAnnualIncome | Estimated annual income of the user based on geographic location |
| estimatedPropertyType | Estimated property type (e.g., residential, commercial) |
| visitCount | Number of previous visits did the user have before the current visit |
| pageURL | URL of the current page visited |
| ctaCopy | Text of the call to action that was prompted to the user for potential engagement |
| ctaPlacement | Location of the call-to-action button on the page |
| editorialSnippet | Text surrounding the call-to-action |
| scrolledPage | If the user scrolled on the page or no AFTER the ctaCopy and ctaPlacement already loaded |
| scrollDepth | Percentage of the page length scrolled by the user. If scrolledPage is 0, than scrollDepth will be 0 |
| clickedCta | If the user clicked the call-to-action (yes/no) |
| scheduledAppointment | If the user scheduled an appointment regarding the mortgage application (yes/no) |
| mortgageVariation | Mortgage product variation the user applied for to discuss during the appointment. This only is present when someone scheduled an appointment |
| revenue | Revenue Financial Services received from the appointment scheduling and mortgage variation. If no scheduled appointment, this will be 0. |