Data Analytics Capstone Topic Approval Form

Student Name: Justin Huynh

Student ID: 012229514

Capstone Project Name: Digital Marketing Analytics

Project Topic: Digital Marketing Campaign Analysis

☒ **This project does not involve human subjects research and is exempt from WGU IRB review.**

Research Question: How does the marketing channel used in digital campaigns impact customer conversion rates, and which channel is the most effective in driving conversions?

Hypothesis: **Null hypothesis**: There is no significant difference in the conversion rates across different marketing channels. **Alternate Hypothesis**: There is a significant difference in the conversion rates across different marketing channels.

Context: In today's competitive business landscape, marketing teams need to justify their ad spend and maximize ROI on their marketing investments. With the prominence of various digital marketing channels, marketers face the challenge of allocating their budgets to drive the most conversions. This dataset provides key insights into various campaigns to run across multiple channels with marketing metrics that make it ideal for data-driven analysis. Analyzing the effectiveness of each campaign channel can help marketers optimize their strategies. For instance, if a certain channel consistently underperforms in driving conversions compared to others, marketers can either improve their approach or reduce investment in that channel. On the other hand, identifying the most successful channels enables businesses to focus resources on what's working, thus increasing overall ROI. The analysis also helps in tailoring marketing strategies to target demographics more effectively. Understanding whether different age groups, income levels, or genders respond better to specific channels provides insights into customer preferences. This information can be leveraged to create more personalized and effective campaigns to improve both customer experience and profitability for a business. By using data analysis to identify the most effective marketing channels for maximizing conversions, businesses can make informed decisions that lead to better marketing efficiency and higher revenue.

Data: For this capstone project, I will be using an existing dataset that has been sourced from Kaggle. The dataset contains information about digital marketing campaigns metrics. The dataset is composed of 8,000 records with 20 columns, which offers a comprehensive view of customer behavior and campaign effectiveness. The relevant data being used for this analysis includes: **CustomerID**: A unique identifier for each customer. **Age**: The age of the customer. **Gender**: The gender of the customer (Male/Female). **Income**: The annual income of the customer. **CampaignChannel**: The channel used for the marketing campaign (i.e. social media, Email, PPC). **CampaignType**: The type of campaign (i.e. Awareness, Retention, Conversion). **AdSpend**: The amount spent on the advertising campaign. **ClickThroughRate**: The ratio of users who clicked on an ad compared to the total who viewed it. **ConversionRate**: The percentage of visitors who took the desired action (i.e. purchase). **Conversion**: A binary indicator showing whether a customer converted (1) or did not convert (0).

The dataset was pulled from Kaggle, which is a site that's typically open-source and made publicly available in the data science field for the purpose of learning and research. The data is freely accessible and can be used for projects without restrictions.

The data belongs to the original contributor(s) who uploaded it to Kaggle. By making the dataset publicly available, the contributor has granted users the right to download and use the dataset for various purposes like school projects. Since this dataset is open-source and available on Kaggle, anyone is permitted to use it without violating any usage rights or terms of service.

Data Gathering: For this project, we'll first source the dataset from Kaggle, which is a publicly available data science platform. Then we'll download from Kaggle and upload to the IDE Jupyter Notebook in a CSV format. Lastly, we'll perform initial data cleaning and transforming the data if necessary to ensure the data is in a usable format for analysis.

Data Analytics Tools and Techniques: For this project, we'll use Python with Jupyter Notebook for data analysis using the following libraries and packages:

- **Pandas**: For data manipulation and cleaning.
- **NumPy**: For numerical computations.
- **Matplotlib/Seaborn**: For data visualization and exploratory analysis.
- **SciPy and Statsmodels**: For statistical tests such as ANOVA to compare conversion rates across different campaign channels.
- **Scikit-learn**: If machine learning models are needed to predict customer conversion based on campaign channel and demographic data.

Justification of Tools/Techniques: Exploratory Data Analysis (EDA) will help uncover patterns, detect anomalies, and visualize relationships between different variables from the dataset. Python's libraries like Pandas and Matplotlib/Seaborn are powerful for performing EDA through data visualization so we can identify trends and relationships between variables like the marketing channel and customer conversion rate.

To test the research hypothesis that different marketing channels impact conversion rates, we will conduct a One-Way ANOVA (Analysis of Variance). ANOVA will allow us to test whether there are statistically significant differences in conversion rates across the different marketing channels.

If further analysis is required to predict customer conversion, we can use logistic regression. It is ideal for predicting binary outcome based on multiple independent variables since the outcome is either a conversion (1) or no conversion (0).

Project Outcomes: This project's outcome is to generate a model and visualizations that will help give recommendations for optimizing digital campaigns in order to determine whether we should reject or fail to reject the null hypothesis. The deliverables will be to first perform an EDA > testing the hypothesis through ANOVA analysis > create data visualizations to illustrate key findings > create a predictive model that predicts customer conversions based on available data > create a final report summarizing the research, methodology, analysis, and recommendations for improving a company's marketing strategy to improve ROI.

Projected Project End Date: 10/31/24

Sources: https://www.kaggle.com/datasets/arpit2712/digital-marketing-company

Course Instructor Signature/Date:

☒ The research is exempt from an IRB Review.

☐ An IRB approval is in place (provide proof in appendix B).

Course Instructor's Approval Status: Approved

Date: 9/22/24 Click here to enter a date.

Reviewed by: Daniel J. Smith, PhD, MBA

Comments: Click here to enter text.