**WEBSITE TRAFFIC ANALYSIS**

**PHASE - 5**

**COLLEGE OF ENGINEERING GUINDY - ANNA UNIVERSITY 0001**

TEAM MEMBERS:-

1)N.WAHEED AHMED

2)R A NITHISH KUMAR

3)B M KAVIN

4)ARAVIND KRISHNAN S

5)HEMANTH RAJ E

**Documentation:**

**Project Overview:**

In this website traffic analysis project, the objective is to gain valuable insights from website traffic data to improve website and marketing performance. Understanding visitor behaviour and preferences is essential for informed decision-making.

**Design Thinking Process:**

The design thinking process began with identifying the problem statement, which was framed as "How to drive more traffic to the website and convert more visitors into customers." This question guided the subsequent phases of development.

**Data Collection and Pre-processing:**

Data was collected from various sources, including Google Analytics, server logs, and marketing platforms. This data included information about user sessions, page views, traffic sources, demographics, and more.

Data pre-processing involved cleaning the data by removing duplicates, addressing missing values, and ensuring data consistency.

**Data Analysis:**

The analysis encompassed segmentation analysis, where visitors were grouped by demographics, interests, and behaviour.

Correlation analysis was conducted to identify relationships between traffic sources, page views, and conversion rates.

Regression analysis was used to build models predicting conversion likelihood based on various factors.

**Hypotheses and Testing:**

Hypotheses were formulated, including "Hypothesis 1: Visitors from social media sources are more likely to convert" and "Hypothesis 2: Demographics play a significant role in conversion rates."

Statistical tests were performed to evaluate these hypotheses and determine their significance.

**Insights and Recommendations:**

Analysis revealed insights such as "Insight 1: Social media marketing is the most effective channel for driving traffic and conversions" and "Insight 2: The 25-34 age group has the highest conversion rate."

Recommendations included increasing the budget for social media advertising and creating targeted content for the 25-34 age group.

**Documentation of Code:**

The code for data pre-processing, analysis, and modelling was documented and provided for review in the repository.

**ReadMe File:**

The README file explains how to run the code and lists necessary dependencies, ensuring easy replication of the analysis.

Submission:

**Code Files:**

The repository contains well-organized code files, including Python scripts for data pre-processing, analysis, and visualization.

**Dataset Source and Description:**

The dataset used for this project originated from Google Analytics and includes user interaction data, session information, and source details.

**GitHub Repository or Portfolio Link:**

The project has been shared on GitHub, allowing evaluators, stakeholders, and peers to access and review the analysis.

**Instructions for Replication:**

Detailed instructions are provided for replicating the analysis, setting up the project environment, and running the code. This covers data collection, pre-processing, and analysis steps.

**Example Outputs:**

Example outputs in the form of data visualizations, charts, and analysis summaries demonstrate the project's findings and insights.

By following this documentation and submission format, the website traffic analysis project can be effectively presented, reviewed, and replicated by others.