**Project Definition**

The project's main objective is to analyse website traffic data to extract valuable insights into user behaviour, identify popular pages, and understand traffic sources. By achieving these objectives, the project aims to assist website owners in enhancing the user experience on their site. This comprehensive project encompasses defining analysis objectives, collecting website traffic data, utilizing IBM Cognos for data visualization, and integrating Python code for advanced analysis.

**About Dataset:**

The variables are daily counts of page loads, unique visitors, first-time visitors, and returning visitors to an academic teaching notes website. There are 2167 rows of data spanning the date range from September 14, 2014, to August 19, 2020.

A visit is defined as a stream of hits on one or more pages on the site on a given day by the same user, as identified by IP address. Multiple individuals with a shared IP address (e.g., in a computer lab) are considered as a single user, so real users may be undercounted to some extent. A visit is classified as "unique" if a hit from the same IP address has not come within the last 6 hours. Returning visitors are identified by cookies if those are accepted.

All others are classified as first-time visitors, so the count of unique visitors is the sum of the counts of returning and first-time visitors by definition. The data was collected through a traffic monitoring service known as Stat Counter.

**Design Thinking:**

**Analysis Objectives**

To successfully analyse website traffic data, we need to define clear objectives:

**1. Identify Popular Pages:**

Determine which pages on the website are the most frequently visited by users. This will help in understanding which content is engaging and which may need improvement.

**2. Traffic Trends:**

Analyse trends in website traffic over time, including daily, weekly, and monthly patterns. This will provide insights into when the website experiences the most and least traffic.

**3. User Engagement Metrics:**

Calculate user engagement metrics such as average session duration, bounce rate, and conversion rates. These metrics will help gauge how effectively the website is engaging its visitors.

**Data Collection**

To collect website traffic data, we'll need to consider the following aspects:

**1. Data Sources:**

Identify the sources from which we will collect data. This may include web analytics tools like Google Analytics, server logs, or custom tracking scripts.

**2. Data Types:**

Determine the specific data points we want to collect, which could include page views, unique visitors, referral sources, user demographics, and more.

**3. Data Frequency:**

Decide how often data will be collected and updated. Real-time, daily, or weekly data collection frequencies should be considered based on project requirements.

**Visualization**

Visualization is a crucial step in presenting our insights effectively. We plan to use IBM Cognos for data visualization, and the following steps will be involved:

**1. Data Preparation:**

Clean and preprocess the collected data to ensure it's ready for visualization. This may involve handling missing data, aggregating data points, and formatting.

**2. Dashboard Creation:**

Design interactive dashboards in IBM Cognos that showcase the key insights. Dashboards should be user-friendly and intuitive, allowing website owners to explore the data easily.

**3. Report Generation:**

Create reports summarizing the findings, including charts, graphs, and tables. Reports should be concise and provide actionable insights.

**Python Integration**

To perform advanced analysis and potentially build machine learning models for prediction, Python will be integrated into the project:

**1. Data Processing:**

Python will be used to preprocess, clean, and transform data for more in-depth analysis. Libraries such as pandas and NumPy will be invaluable in this step.

**2. Advanced Analysis:**

Advanced analysis techniques, including machine learning algorithms, can be applied to uncover patterns and trends not easily discernible through traditional analysis methods.

**3. Predictive Modelling:**

Machine learning models can be developed to predict future traffic trends or user behaviour patterns, enabling website owners to make proactive improvements.

**Dataset Link:**[**https://www.kaggle.com/datasets/bobnau/daily-website-visitors**](about:blank)

**Conclusion**

In this phase, we have outlined the problem definition and design thinking for the project. We've identified clear analysis objectives, considered data collection methods, and discussed the visualization and Python integration aspects of the project. With a well-defined plan in place, we are ready to move on to the next phases of the project, including data collection, analysis, and reporting.