Website Statistics Analysis for E-commerce Optimization

Background:

In the dynamic landscape of e-commerce, understanding user behavior and website performance is crucial for optimizing conversions and maximizing revenue. The provided dataset offers a comprehensive view of user interactions on an e-commerce website, encompassing various dimensions such as page visits, session durations, traffic sources, and purchase behavior.

Objective:

The objective of this project is to leverage the dataset to analyze website statistics and derive actionable insights aimed at enhancing the e-commerce platform's performance and maximizing revenue generation. By examining user engagement metrics, traffic patterns, and transactional data, the goal is to identify areas of improvement and formulate strategic recommendations for optimization.

Dataset:
https://drive.google.com/drive/folders/1Iesv1Y7DVdv6Z4Oq8IwJ99_gMSEFecDk?usp=sharing

Expectations:

Through this analysis, the project aims to:

- Identify key drivers of user engagement and conversion on the e-commerce website.
- Provide actionable recommendations for optimizing website performance and enhancing user experience.
- Enhance revenue generation through targeted improvements in traffic acquisition, user engagement, and conversion optimization strategies.

Suggestions for approaching the case study:

Preprocessing of Data:

- Data Cleaning: Identify and handle missing values, correct data entry errors, and ensure consistency in data formats.
- Feature Engineering: Create new features if necessary, such as binary indicators for special days or whether all page categories were visited.
- Timestamp Conversion: Convert any date-related fields to a consistent format to facilitate analysis.

Univariate Analysis:

- Numerical Features: Plot histograms or box plots for each numerical feature (e.g., Administrative, Informational, ProductRelated, durations, BounceRates, ExitRates, PageValues) to identify outliers and understand their distribution.
- Categorical Features: Analyze the distribution of categorical variables (e.g., Month, OperatingSystems, Browser, Region, TrafficType, VisitorType, Weekend).

Correlation Analysis:

- Correlation Matrix: Calculate the correlation coefficients between numerical features to identify potential relationships.
- Visualization: Use heatmaps to visualize these correlations and detect multicollinearity or interesting relationships.

Class Distribution:

- Target Variable ('Revenue'): Check the distribution of the target variable to understand class balance and address any imbalances if necessary.

Summarize Key Metrics:

- Page Views and Durations: Summarize the number of pages visited and the time spent in each category (Administrative, Informational, ProductRelated).
- Bounce and Exit Rates: Calculate average bounce and exit rates for each page category to identify high drop-off points.

SpecialDay Analysis:

- Distribution Analysis: Explore the distribution of the SpecialDay feature and its values.

- Correlation with Revenue: Analyze how the proximity to special days correlates with the likelihood of generating revenue.

Feature Engineering:

- Binary Indicator for Page Categories: Generate a binary feature indicating whether the user visited all three page categories (Administrative, Informational, ProductRelated).

Exploratory Data Analysis (EDA):

- Visualize Relationships: Use scatter plots, pair plots, and heatmaps to visualize relationships between key numerical features.
- PageValues Analysis: Investigate the distribution of PageValues and its relationship with TrafficType, VisitorType, and Region.
- Session Lengths: Analyze user session lengths and their impact on conversion rates.

Behavior and Conversion Rate Analysis:

- Group Users: Group users based on various factors such as TrafficType, VisitorType, Region, and analyze differences in behavior and conversion rates.
- Conversion Funnel Analysis: Examine the steps leading to a conversion and identify potential drop-off points or barriers to completing a purchase.

Visualization and Reporting:

- Clear Visualizations: Create compelling visualizations (e.g., bar charts, line graphs, pie charts) to effectively communicate key findings.
- Comprehensive Reports: Write detailed reports summarizing the analysis, findings, and recommendations.
- Present Findings: Prepare presentations to share insights and actionable recommendations with stakeholders.