Exploratory Data Analysis

Aditya Pandey
Yash Kamble

Introduction

Exploratory Data Analysis (EDA) is an approach to analyzing data sets to summarize their main characteristics, often with visual methods. The objective is to discover patterns, spot anomalies, and test hypotheses, guiding the next steps in data analysis and decision making.

01

Overview

Define Exploratory Data Analysis (EDA)

EDA is a critical step in the data analysis process, involving summarizing the main characteristics of a dataset. It emphasizes graphical representations, helping analysts understand the data's underlying structure and relationships.

Objective of the analysis

The primary objective of this EDA is to extract meaningful insights from website traffic data, identify trends and patterns, and guide strategic decisions based on the analysis.

Importance of EDA in website traffic

Conducting EDA on website traffic data helps to understand user behavior, optimize user experience, and measure engagement metrics effectively, leading to informed improvements in web strategies.

02

Dataset

Description of the dataset

The dataset consists of records captured from website visits, including unique page views, user sessions, traffic sources, and user demographics. This information is crucial for comprehensive analysis.

Data source information

The data is sourced from web analytics
tools such as Google Analytics, which
tracks user interactions on the website,
providing essential data for analysis.

Key features to analyze

Key features in the dataset include visitor counts, session durations, bounce rates, traffic sources, and user demographics. These features will be analyzed to derive actionable insights.

03

Data Cleaning

Handling missing values

To handle missing values, various strategies were applied, such as imputation, where missing entries are replaced with the mean, median, or mode, and removal of records that had excessive missing data. It ensures that the analysis is based on complete information.

Data transformation methods

Data transformation methods applied include normalization and scaling of numerical features, which help in adjusting the data to fit within a specific range without distorting differences in the ranges of values. This aids in better analysis and interpretation.
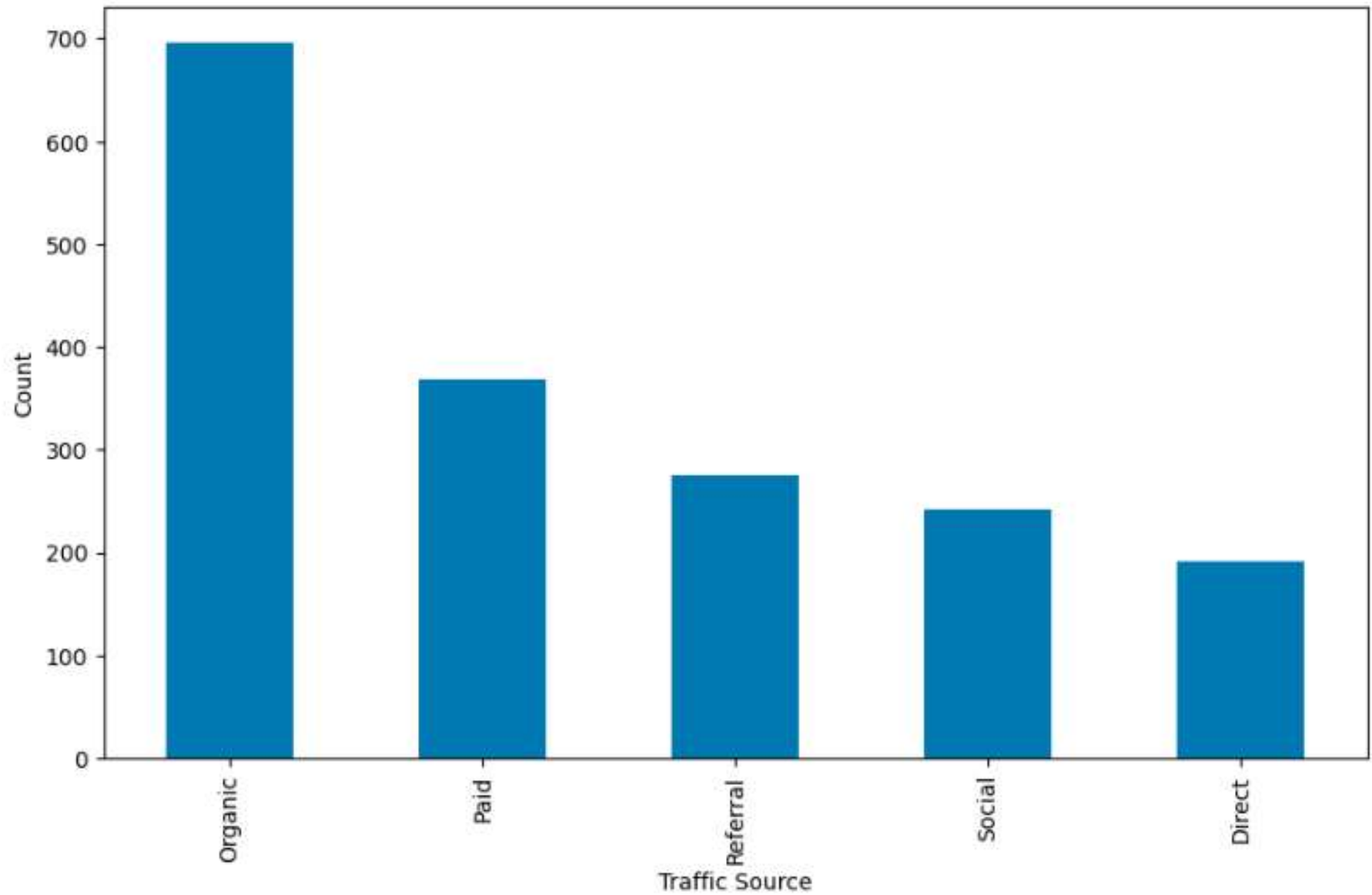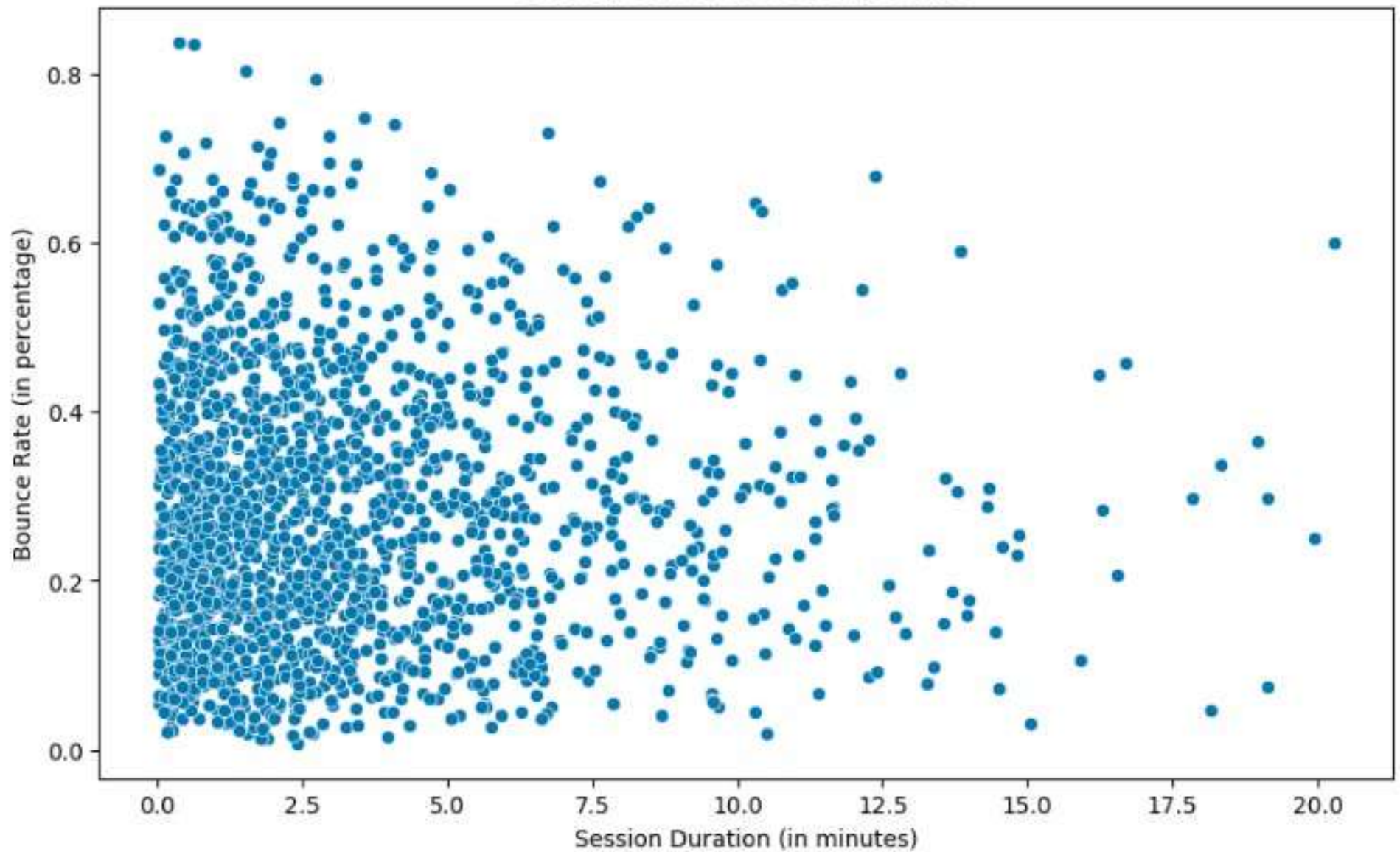
04

Visualization

Graphs and charts for data analysis

Graphs and charts, such as bar charts, pie charts, and histograms, are essential tools for visualizing data analysis results. They provide an easy-to-understand representation of complex data, enabling stakeholders to grasp insights at a glance.
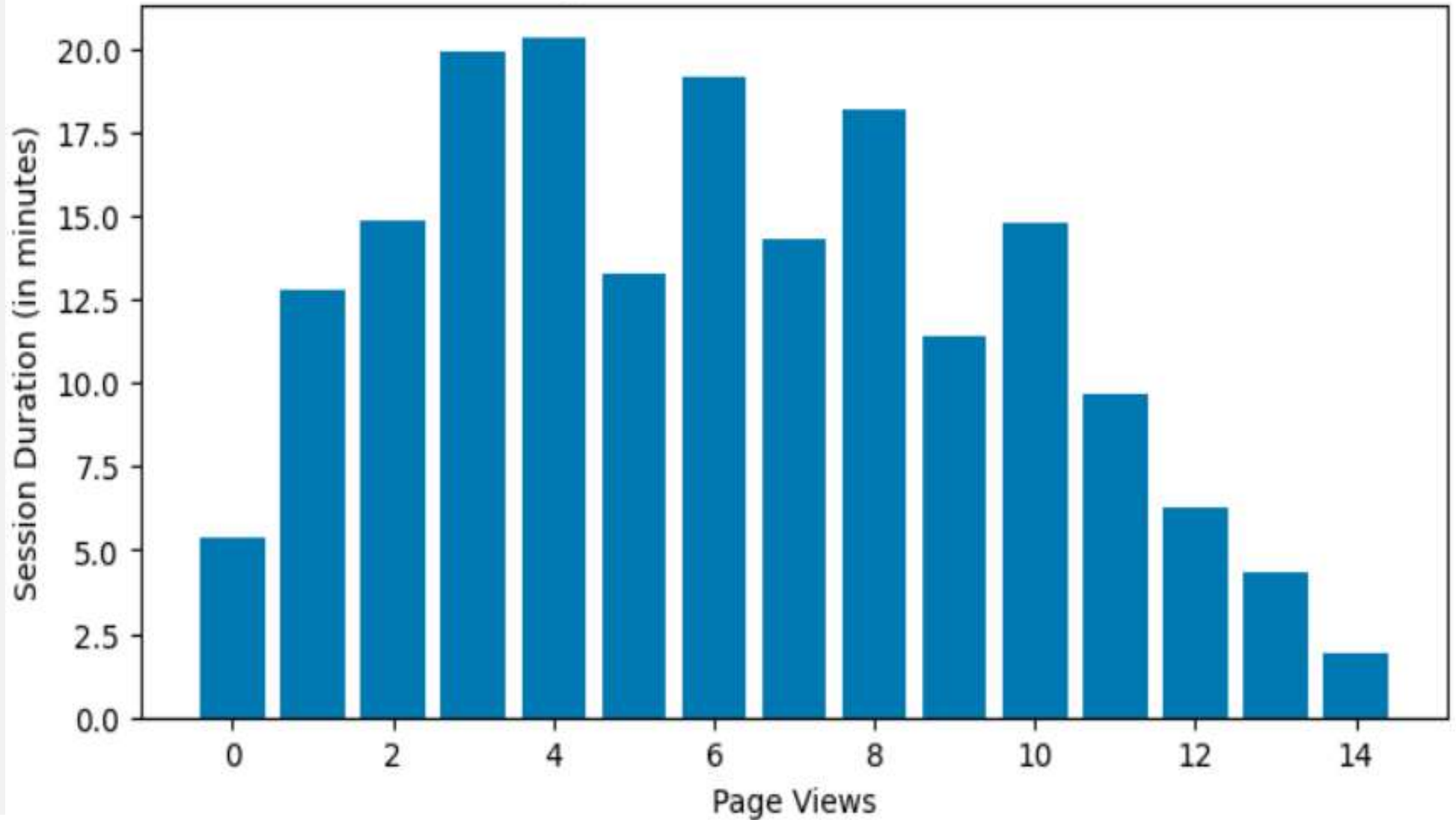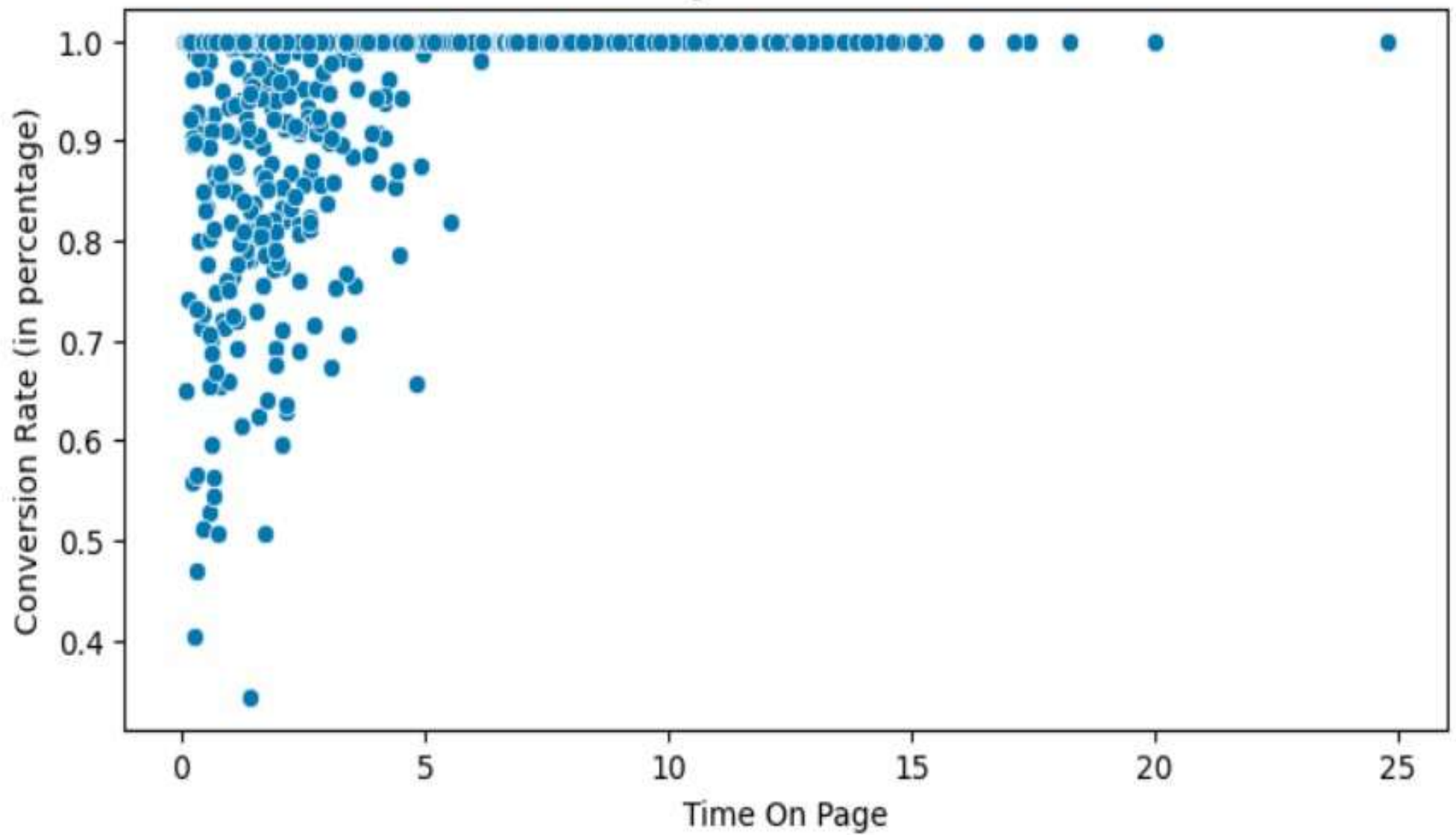
Traffic Source Distribution

Session Duration vs Bounce Rate
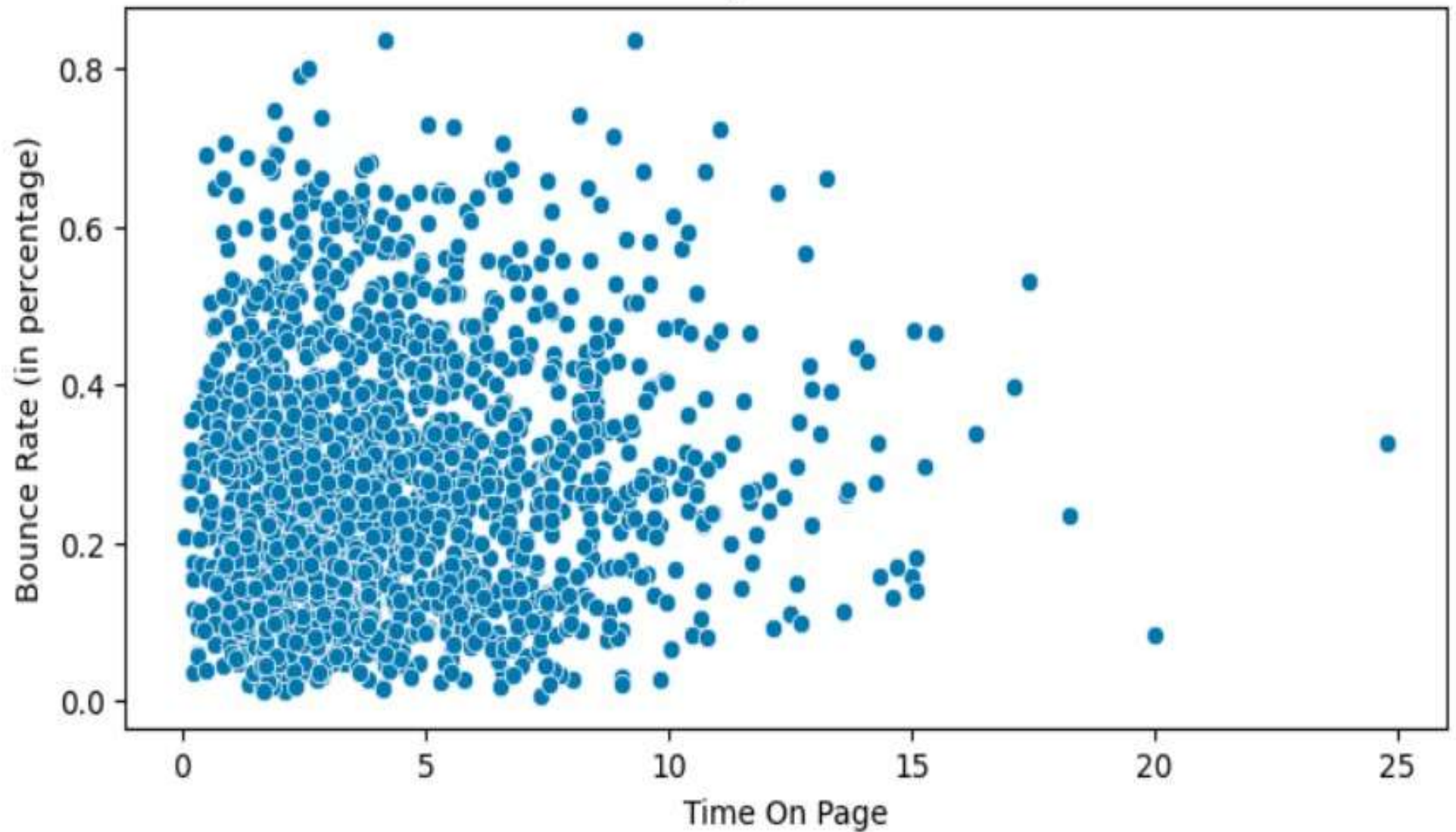
Page Views vs Session Duration

Time On Page vs Conversion Rate

Time On Page vs Bounce Rate

Conclusion

In conclusion, the exploratory data analysis on website traffic reveals significant insights into user behavior, helping identify key patterns and trends that can drive decision-making. Future strategies should focus on leveraging these insights for website optimization.

Thank you!

Do you have any questions?