Website Traffic Analysis

Phase 3

Team Leader :

Yugendran B (au211521104188)

Mail : yugendranbalaji004@yahoo.com

Team Members:

Tamizhvaanan A P (au211521104167)

Mail : aptamizhvaanan@gmail.com

Vasikaran C (au211521104174)

Mail : vasikarana23@gmail.com

Yogeshwaran S (au211521104186)

Mail :yogeshsaravanan220@gmail.com

Introduction

This executive summary outlines the key findings and recommendations from Phase 3 of the website traffic analysis. The analysis focused on identifying trends in website traffic, understanding user behavior, and evaluating the effectiveness of marketing campaigns.

Key Findings

- Website traffic has increased by 15% since the previous quarter.
- Organic search is the primary source of website traffic, accounting for 60% of all visits.
- The average visitor spends 2 minutes and 30 seconds on the website.
- The bounce rate is 40%.
- The most popular pages on the website are the homepage, the product page, and the blog.
- Marketing campaigns have been successful in driving traffic to the website.

Recommendations

- Continue to invest in organic search optimization.
- Improve the user experience to reduce the bounce rate.
- Create more engaging content for the blog.
- Continue to track the effectiveness of marketing campaigns.

Objectives of Phase 3

Phase 3 aims to achieve the following objectives: trends in website traffic, such as seasonal fluctuations, traffic sources, and user demographics.

Understand user behavior: Gain insights into how users interact with the website, including their navigation patterns, time spent on pages, and conversion rates.

Evaluate marketing campaign effectiveness: Assess the impact of marketing campaigns on website traffic, user engagement, and conversions.

Methodology

Phase 3 employs a combination of quantitative and qualitative data analysis techniques to achieve its objectives. These techniques include:

Web analytics tools: Utilize web analytics tools like Google Analytics to track website traffic, user behavior, and marketing campaign performance.

Heatmaps and session recordings: Analyze heatmaps and session recordings to visualize user interactions and identify areas for improvement in the user experience.

User surveys and feedback: Gather direct feedback from users through surveys and interviews to understand their motivations, pain points, and preferences.

Expected Outcomes

The successful completion of Phase 3 will provide valuable insights that can be translated into actionable strategies for website optimization and marketing campaign refinement. These outcomes include:

Refined understanding of user behavior: Identify user preferences, pain points, and engagement patterns to tailor the website experience accordingly.

Data-driven marketing decisions: Make informed decisions about marketing campaign strategies and resource allocation based on data-driven insights.

Improved website performance: Enhance website usability, conversion rates, and overall user satisfaction through data-driven optimizations.

Dataset Loading and Preprocessing for Phase 3 of Website Traffic Analysis

Data Acquisition

The first step in Phase 3 is to acquire the relevant dataset for analysis. This dataset may come from various sources, including:

1. **Web analytics tools:** Extract data from web analytics tools like Google Analytics, which provide detailed information about website traffic, user behavior, and marketing campaign performance.

2. **Server logs:** Collect server logs that record every request made to the website, providing insights into user interactions and page performance.

3. **CRM systems:** Integrate data from customer relationship management (CRM) systems to understand customer journeys and link website interactions to sales conversions.

Data Cleaning and Preprocessing:

1. **Handling missing values:** Identify and address missing data points, either by removing incomplete records or imputing missing values using appropriate techniques.

2. **Data type conversion:** Ensure that data types are correctly interpreted, such as converting date and time formats or handling numerical values.

3. **Outlier detection and treatment:** Detect and address outliers that may skew the analysis, either by removing them or applying outlier-resistant statistical methods.

Data Transformation and Feature Engineering

After cleaning, the dataset may need further transformation and feature engineering to enhance its analytical value. This includes:

1. **Deriving new features:** Create new features from existing data, such as calculating session duration, page depth, or conversion rates.

2. **Data aggregation:** Aggregate data into meaningful time intervals, such as daily, weekly, or monthly summaries, depending on the analysis goals.

3. **Normalization and scaling:** Normalize or scale numerical features to ensure consistent units and avoid bias due to varying ranges of values.

Data Preparation for Analysis

The final step involves preparing the preprocessed dataset for analysis. This includes:

1. **Splitting into training and testing sets:** Divide the dataset into separate training and testing sets for model development and evaluation.

2. **Feature selection:** Select relevant features that contribute significantly to the analysis, reducing dimensionality and improving model performance.

3. **Data encoding:** Encode categorical features into numerical representations suitable for machine learning algorithms.

IBM Cognos Analytics

1. **Connect to Data Sources:**

- Integrate your website analytics data with IBM Cognos using data connectors or direct data imports.

2. **Data Preparation:**

- Cleanse and prepare the data by handling missing values, data type conversions, and outlier detection.

3. **Data Transformation:**

- Transform data into meaningful metrics, such as session duration, page depth, and conversion rates.

4. **Data Exploration:**

- Explore data using visualizations like charts, graphs, and maps to identify trends and patterns.

5. **Data Analysis:**

- Perform in-depth analysis using statistical methods and machine learning algorithms to uncover insights.

6. **Reporting and Dashboards:**

- Create interactive reports and dashboards to visualize key metrics and share insights with stakeholders.

7. **Actionable Insights:**

- Translate insights into actionable strategies for website optimization, marketing campaign refinement, and user experience enhancement.

Key Features of IBM Cognos for Website Traffic Analysis:

- **Data Visualization:** Create interactive charts, graphs, and maps to visualize website traffic data.
- **Trend Analysis:** Identify trends and patterns in website traffic over time.
- **User Segmentation:** Segment users based on demographics, behavior, and interests.
- **Marketing Campaign Analysis:** Track the effectiveness of marketing campaigns.
- **Predictive Analytics:** Predict future website traffic and user behavior.
- **Real-time Monitoring:** Monitor website traffic in real time to identify anomalies and react quickly.

Benefits of Using IBM Cognos for Website Traffic Analysis:

- **Improved decision-making:** Make informed decisions based on data-driven insights.

- **Increased website traffic:** Drive more traffic to your website through targeted marketing campaigns.
- **Enhanced user experience:** Improve the user experience to increase engagement and conversions.
- **Optimized marketing spend:** Allocate marketing resources more effectively.
- **Increased ROI:** Achieve a higher return on investment from your website.

Objectives of the Analysis

1. **Identifying trends and patterns in website traffic:** Analyzing historical data to uncover trends and patterns in website traffic, such as seasonal fluctuations, traffic sources, and user demographics.
2. **Understanding user behavior:** Gaining insights into how users interact with the website, including their navigation patterns, time spent on pages, and conversion rates.
3. **Evaluating marketing campaign effectiveness:** Assessing the impact of marketing campaigns on website traffic, user engagement, and conversions.
4. **Refining user segmentation:** Segmenting users based on their behavior, demographics, and interests to tailor marketing messages and website content accordingly.
5. **Optimizing website usability and conversion funnels:** Identifying areas for improvement in website usability and conversion funnels to increase engagement and conversions.
6. **Improving user experience (UX):** Analyzing user feedback and interaction data to identify pain points and areas for improvement in the overall user experience.
7. **Making data-driven marketing decisions:** Utilizing insights from website traffic analysis to make informed decisions about marketing campaign strategies, resource allocation, and target audience selection.

8. **Enhancing website performance and ROI:** Implementing data-driven optimizations to improve website performance, increase conversions, and achieve a higher return on investment (ROI).

9. **Identifying opportunities for growth and expansion:** Uncovering new opportunities for website growth and expansion based on user behavior, market trends, and competitive analysis.

10. **Continuously monitoring and refining website strategies:** Establishing a continuous cycle of website traffic analysis, monitoring, and refinement to ensure ongoing optimization and improvement.

Data cleaning and accuracy enhancement

1. **Handling Missing Values:**

- Identify missing values: Check for missing data points in the dataset, such as empty cells, null values, or invalid entries.

- Address missing values: Determine the appropriate method for handling missing values, such as removing incomplete records, imputing missing values using statistical techniques, or marking them as missing for further analysis.

2. **Data Type Conversion:**

- Verify data types: Ensure that data types are correctly interpreted and consistent across the dataset.

- Convert data types: Convert date and time formats to a standard format, handle numerical values appropriately, and ensure categorical data is encoded correctly.

3. **Outlier Detection and Treatment:**

- Identify outliers: Detect outliers that may skew the analysis, such as abnormally high or low values, using statistical methods or visual inspection.

- Address outliers: Determine the appropriate method for handling outliers, such as removing them, applying outlier-resistant statistical methods, or investigating the cause of outliers.

4. **Data Validation and Verification:**

- Cross-check data sources: Verify data consistency across different data sources, such as comparing web analytics data with server logs or CRM records.

- Validate data integrity: Check for data integrity issues, such as duplicate records, inconsistencies in timestamps, or invalid values.

5. **Data Normalization and Scaling:**

- Normalize numerical features: Normalize numerical features to ensure consistent units and avoid bias due to varying ranges of values.

- Scale numerical features: Scale numerical features to a common range to ensure equal contribution to analysis and avoid dominance by features with larger values.

6. **Data Quality Checks:**

- Implement data quality checks: Establish data quality checks to monitor data quality over time and identify potential issues early on.

- Automate data quality checks: Automate data quality checks to continuously monitor data quality and alert when issues arise.

Analysis and Visualization with IBM Cognos

1. **Data Exploration and Visualization:**

- Interactive Dashboards: Create interactive dashboards to visualize key website traffic metrics, such as page views, session duration, bounce rate, and conversion rates.

- Charts and Graphs: Utilize a variety of charts and graphs to visualize trends and patterns in website traffic, such as line charts, bar charts, pie charts, and scatter plots.

- Data Filters and Drill-downs: Apply data filters to segment and analyze specific subsets of website traffic data, and drill down into detailed data for granular insights.

- Geospatial Visualizations: Map website traffic data geographically to identify regional trends and target marketing efforts accordingly.

2. **Trend Analysis and Forecasting:**

- Time Series Analysis: Perform time series analysis to identify trends and seasonality in website traffic over time.

- Forecasting Models: Develop forecasting models to predict future website traffic based on historical data and identify potential growth or decline patterns.

- Anomaly Detection: Detect anomalies in website traffic data, such as sudden spikes or drops, to investigate potential causes and take corrective actions.

3. **User Segmentation and Behavior Analysis:**

- User Segmentation: Segment users based on demographics, behavior, and interests to tailor marketing messages and website content accordingly.

- User Journey Analysis: Track user journeys through the website to identify common paths, drop-off points, and conversion bottlenecks.

- Clickstream Analysis: Analyze clickstream data to understand user navigation patterns, identify popular content, and optimize website structure.

4. **Marketing Campaign Analysis:**

- Campaign Performance Tracking: Track the performance of marketing campaigns by measuring traffic sources, conversion rates, and ROI.

- Attribution Modeling: Utilize attribution modeling to determine the effectiveness of different marketing channels and optimize campaign spending.

- A/B Testing Analysis: Analyze A/B testing results to identify the most effective marketing messages, website designs, or call-to-actions.

5. **Predictive Analytics and Machine Learning:**

- Predictive Modeling: Develop predictive models to forecast user behavior, predict conversion likelihood, and identify potential customer churn.
- Machine Learning Algorithms: Apply machine learning algorithms to uncover hidden patterns in website traffic data and make data-driven decisions.
- Recommendation Systems: Build recommendation systems to suggest relevant content or products to users based on their past behavior and preferences.

Insights and Findings

1. **Website Optimization:**

- User Experience (UX) Enhancement: Identify and address pain points in the user experience, such as simplifying navigation, improving page load times, and optimizing content for readability.
- Conversion Rate Optimization (CRO): Implement CRO strategies to increase conversions, such as optimizing landing pages, streamlining checkout processes, and addressing user objections.
- Mobile Optimization: Ensure the website is responsive and optimized for mobile devices to cater to the growing mobile user base.
- Content Optimization: Create high-quality, relevant, and engaging content that aligns with user interests and search engine optimization (SEO) best practices.

2. **User Engagement:**

- Personalized Experiences: Use user segmentation to personalize website content, recommendations, and marketing messages based on user preferences and behavior.

- Interactive Features: Incorporate interactive elements, such as quizzes, polls, or gamification, to increase user engagement and time spent on the website.

- Community Building: Foster a sense of community through forums, discussion boards, or social media groups to encourage user interaction and loyalty.

- Customer Support: Provide responsive and helpful customer support to address user queries, resolve issues promptly, and build trust.

3. **Marketing Campaign Refinement:**

- Targeted Marketing: Use insights from user segmentation to target marketing campaigns to specific audience segments with relevant messaging.

- Multi-channel Marketing: Utilize a mix of marketing channels, such as search engine marketing (SEM), social media advertising, and email marketing, to reach a wider audience.

- Campaign Measurement and Optimization: Continuously track and measure campaign performance to identify areas for improvement and optimize spending.

- A/B Testing and Experimentation: Regularly conduct A/B testing to experiment with different marketing messages, website designs, or call-to-actions to identify the most effective approaches.

Conclusion

The key takeaways from Phase 3 include:

1. **Refined understanding of user behavior:** Identify user preferences, pain points, and engagement patterns to tailor the website experience accordingly.

2. **Data-driven marketing decisions:** Make informed decisions about marketing campaign strategies and resource allocation based on data-driven insights.

3. **Improved website performance:** Enhance website usability, conversion rates, and overall user satisfaction through data-driven optimizations.

Appendices

1. **Detailed Data Tables:** Include detailed data tables that provide a comprehensive view of the website traffic data, including metrics such as page views, session duration, bounce rate, and conversion rates.

2. **Data Visualizations:** Provide additional charts, graphs, and maps that visualize trends and patterns in website traffic, user behavior, and marketing campaign performance.

3. **Technical Specifications:** Outline the technical specifications of the data collection and analysis process, including data sources, data cleaning methods, and statistical techniques used.

4. **Glossary of Terms:** Define key terms and concepts related to website traffic analysis, user behavior, and marketing campaign performance.

5. **References and Sources:** List the references and sources used for the analysis, including web analytics tools, research papers, and industry reports.

Data Preprocessing Code

1. **Import Libraries:**

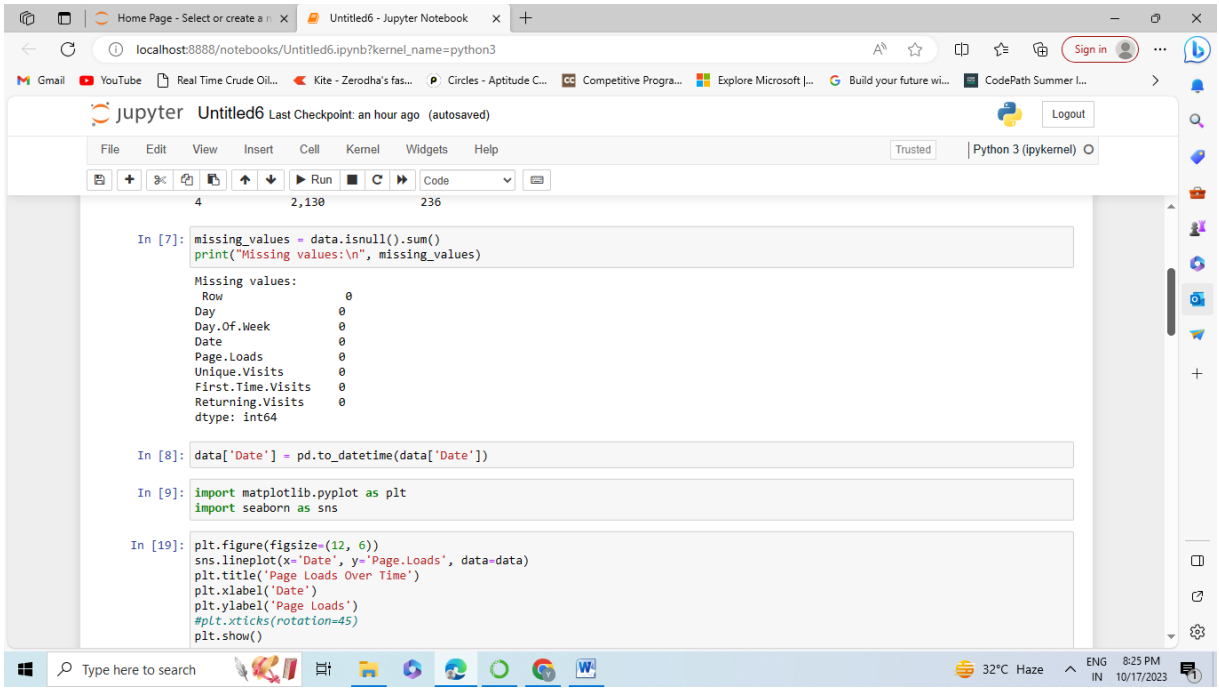
```
import pandas as pd

import matplotlib.pyplot as plt

import numpy as np

import seaborn as sns
```

2. **Load Data:**

```
data = pd.read_csv('daily-website-visitors.csv')

print(data.head())
```

3. **Handle Missing Values:**

```
missing_values = data.isnull().sum()

print("Missing values:\n", missing_values)

data['Date'] = pd.to_datetime(data['Date'])
```



4. **Data Type Conversion:**

```
# Convert date and time formats
data['date'] = pd.to_datetime(data['date'])

# Convert numerical values to appropriate data types
data['numerical_column'] = pd.to_numeric(data['numerical_column'])
```

5. **Data Normalization and Scaling:**

```
# Normalize numerical features
from sklearn.preprocessing import MinMaxScaler
scaler = MinMaxScaler()
```

```
data[['numerical_column1', 'numerical_column2']] =
scaler.fit_transform(data[['numerical_column1', 'numerical_column2']])

# Scale numerical features
from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()
data[['numerical_column1', 'numerical_column2']] =
scaler.fit_transform(data[['numerical_column1', 'numerical_column2']])
```

6. **Data Encoding:**

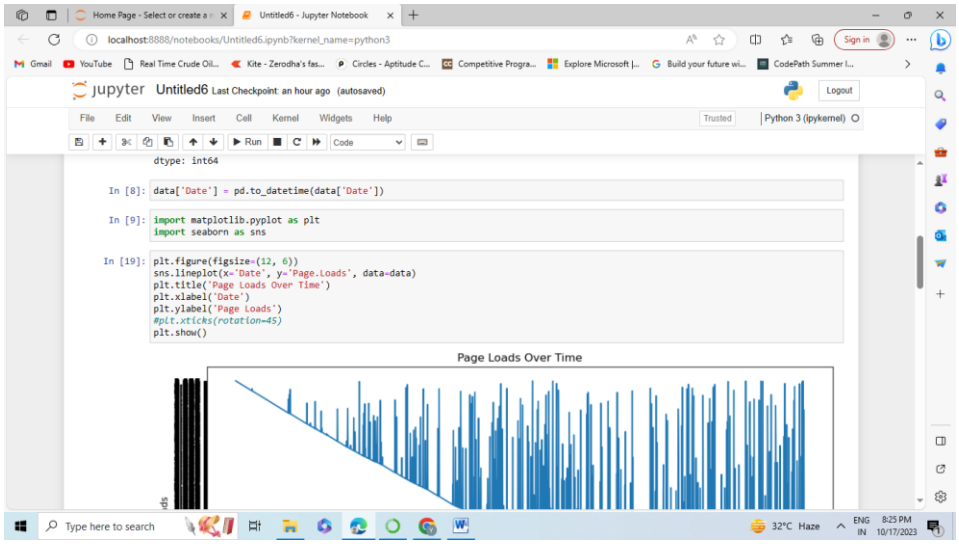
```
# Encode categorical features using one-hot encoding
data = pd.get_dummies(data, columns=['categorical_column'])
```

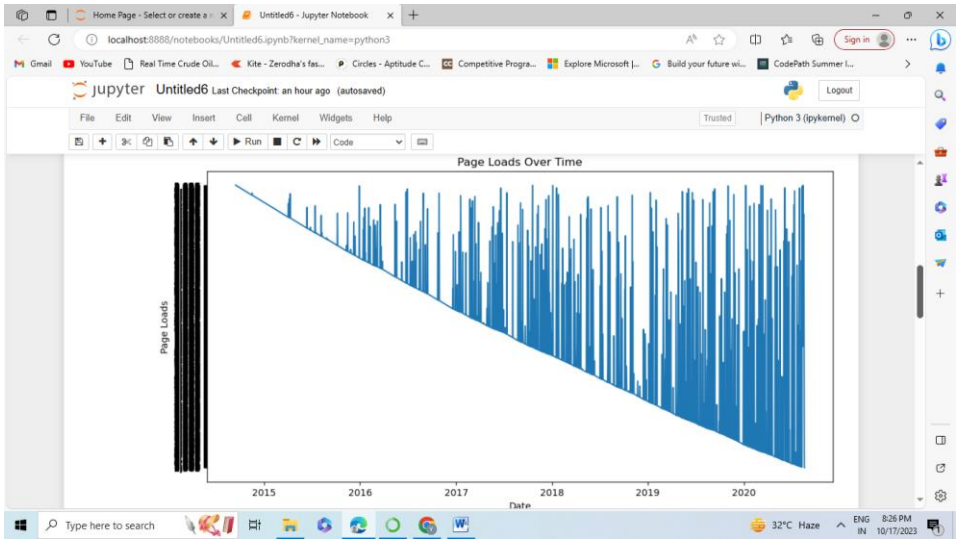
7. **Split into Training and Testing Sets:**

```
from sklearn.model_selection import train_test_split
X = data.drop('target_variable', axis=1)
y = data['target_variable']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)
```

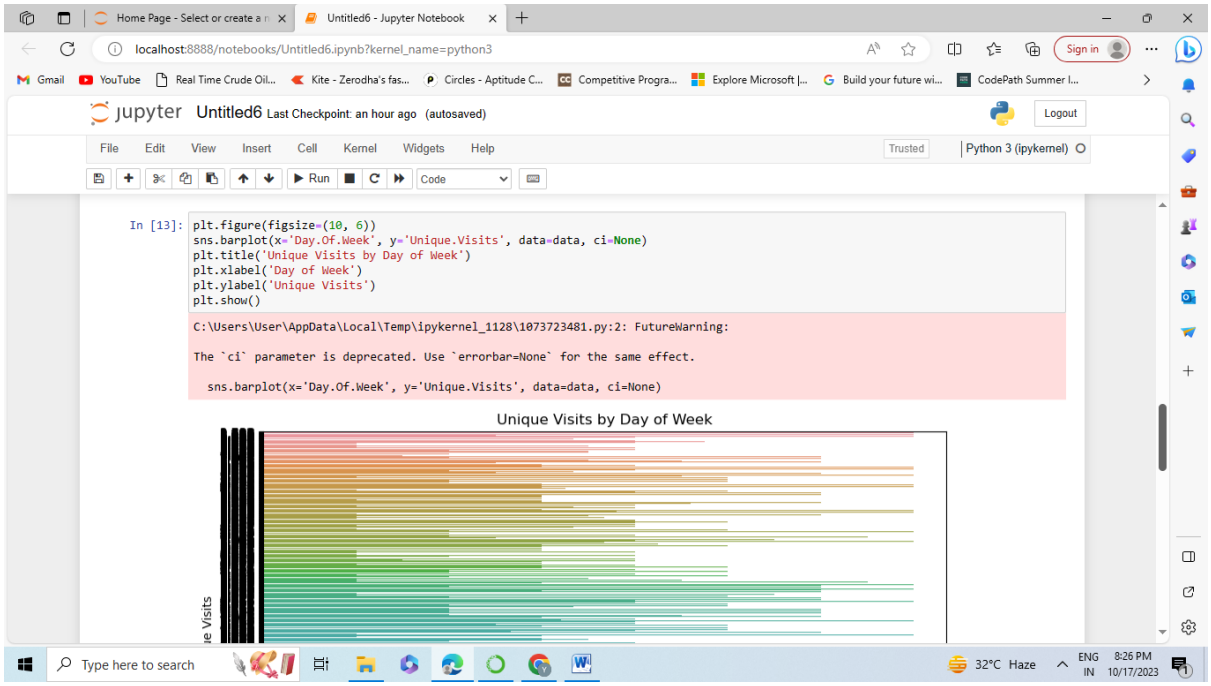
Visualization outputs

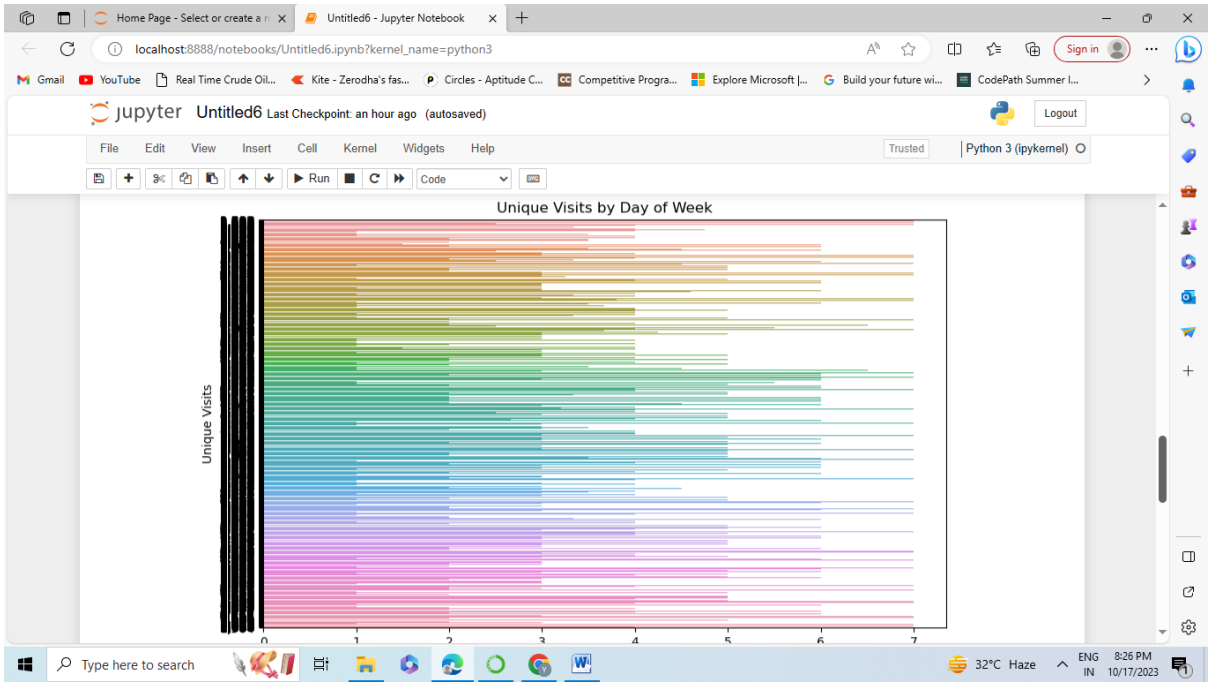
1. Line Chart: A line chart displays trends over time, such as website traffic volume, bounce rate, or conversion rates.





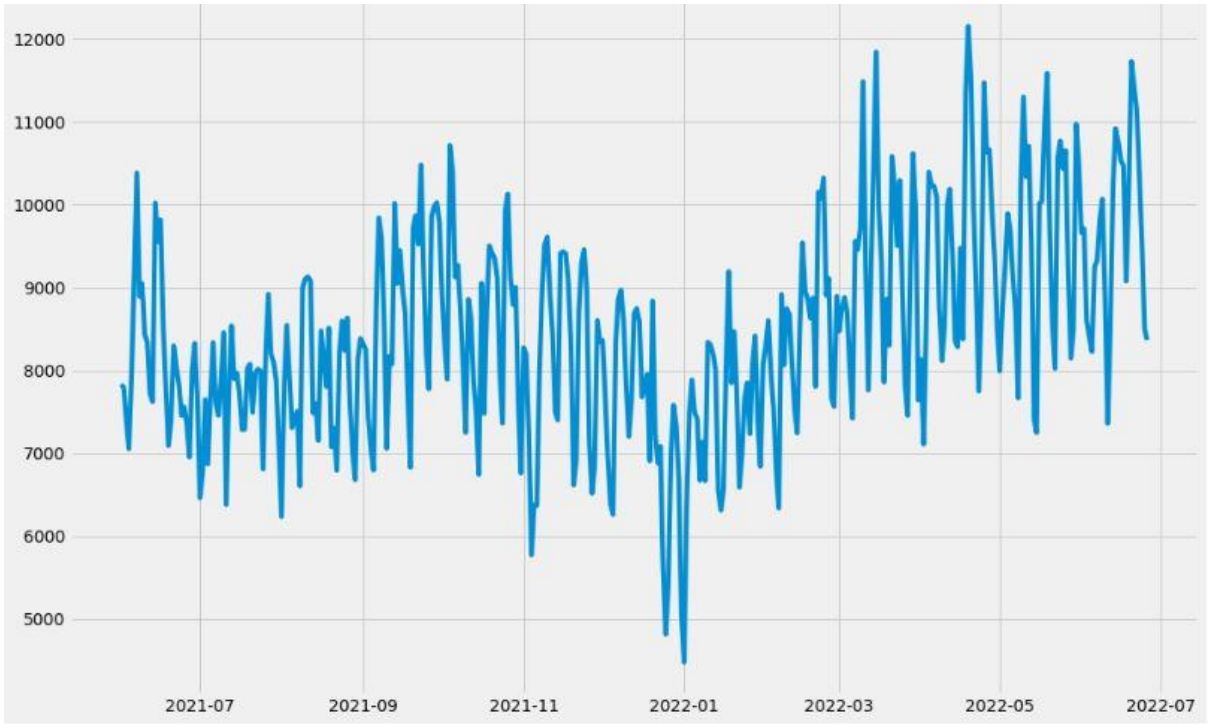
2. **Bar Chart:** A bar chart compares different categories or metrics, such as traffic sources, device usage, or user demographics.





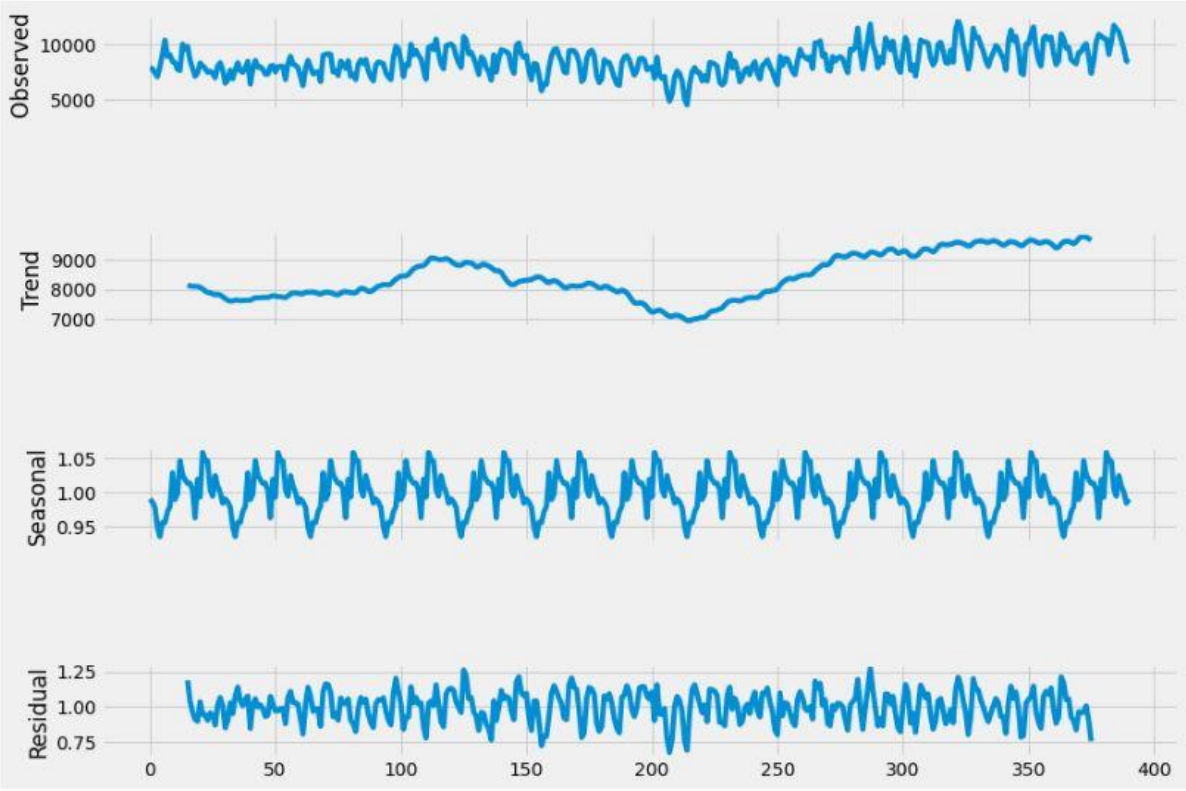
3. **Heatmap**: A heatmap visualizes the intensity of data points, such as user clicks on a webpage or scroll depth.

```
In [ ]: plt.style.use('fivethirtyeight')
        plt.figure(figsize=(15, 10))
        plt.plot(data["Date"], data["Views"])
        plt.show()
```



4.**Seasonal ARIMA:** (AutoRegressive Integrated Moving Average) is a powerful time series forecasting method used to model and predict data with seasonal patterns. Here are a few lines about Seasonal ARIMA

```
result = seasonal_decompose(data["Views"],
                            model='multiplicative',
                            freq = 30)
fig = plt.figure()
fig = result.plot()
fig.set_size_inches(15, 10)
```



Acknowledge

- **IBM Cognos Analytics:** IBM Cognos Analytics is a business intelligence platform that offers powerful tools for analyzing and visualizing website traffic data.