Title:

WebPulse: Intelligent Website Traffic Insights Using Power BI

Abstract:

This project presents an interactive Website Traffic Analysis Dashboard built using Power BI, designed to track and visualize key website performance metrics. The dashboard leverages data from SQL queries, providing insights into session duration, bounce rates, page views, and traffic sources. It allows businesses to make informed decisions regarding marketing strategies, user engagement, and website optimization, focusing on device, region, and browser segmentation. The visualization empowers users to understand trends, improve user retention, and optimize their digital strategies.

Dataset Details:

1. Size

- **Approximate Size:** 10MB (depending on the volume of traffic data)

2. Attributes/Columns

- **Device Lookup Dataset:**
  - **Columns:** Device_Key, Device_Name, Device_Type, Content_Segment, Device_Browser
- **Geo Lookup Dataset:**
  - **Columns:** Location_Key, Location_Name, Location_Country, Location_Region, Location_City
- **Website Traffic Data:**
  - **Columns:** Session_Id, Date_Key, Source_Key, Device_Key, Location_Key, Session_Duration(Seconds), Page_Views_Per_Session
- **Source Lookup Dataset:**
  - **Columns:** Source_Key, Source_Name, Source_Type, Source_Campaign

3. Rows/Entries:

- **Device Lookup:** 100+ rows (varies based on the number of devices tracked)
- **Geo Lookup:** 500+ rows (depending on the region/city coverage)
- **Website Traffic Data:** 50,000+ rows (depending on the website traffic collected)
- **Source Lookup:** 50+ rows (traffic source entries)

Tools Used:

1. **Power BI** - For data visualization and dashboard creation.
2. **SQL** - For querying and retrieving data from relational databases.
3. **Excel** - For organizing and cleaning the datasets before importing into Power BI.
4. **DAX (Data Analysis Expressions)** - For creating calculated columns and measures in Power BI.

Operations Used:

1. **Data Import & Integration:**
    - Import datasets from Excel and SQL.
    - Join datasets using common keys (Device_Key, Location_Key, Source_Key).
2. **Data Transformation:**
    - Clean and preprocess the data, ensuring accurate relationships.
    - Convert session duration from seconds to more readable units (e.g., minutes).
3. **Data Aggregation:**
    - Aggregating session duration, bounce rate, and page views over specific periods (e.g., by month, quarter).
4. **Segmentation:**
    - Segmenting data by device type, traffic source, and region to create meaningful insights.
5. **Data Visualization:**
    - Creating interactive charts and graphs (line charts, bar charts, maps, and pie charts).
    - Implementing time-series trends for metrics like session duration and bounce rates.
6. **DAX Calculations:**
    - Using DAX functions to compute key metrics (average session duration, bounce rate, etc.).

Procedure:

The project follows a systematic data analysis lifecycle, combining data integration, transformation, visualization, and insight generation using SQL and Power BI:

1. Data Collection & Preparation

- **Excel Files Used:** Which consist of Device lookup, Geo lookup, Website Traffic Data, Source lookup columns
- **SQL Queries:**
    - SQL was used to **join**, **clean**, and **transform** these datasets.

- Keys such as Device_Key, Location_Key, and Source_Key were used to relate different tables.

Complex queries were crafted to calculate aggregations such as:

Average session duration, Bounce rate , Total page views ,Segmentation by device, source, and region

2. Data Modelling in Power BI

- The cleaned and joined data was imported into Power BI.
- **Relationships** were created between the tables to form a star schema, enhancing performance and flexibility.
- Calculated columns and measures were built using **DAX** to compute key performance indicators (KPIs).

3. Dashboard Design & Visualization

The dashboard is designed with multiple visual layers and charts to provide an in-depth traffic analysis:

Session Duration Analysis

- Time-series charts show trends by month and quarter.
- Segment-based visuals break down session duration by:
  - **Device Type** (e.g., desktop, mobile)
  - **Region/City**

Bounce Rate Trends

- Line and bar charts display bounce rates over time.
- Comparative visuals highlight bounce rates across:
  - Devices (e.g., higher bounce rates on mobile)
  - Content segments (e.g., business, gaming)

Page Views

- Visuals track total page views and compare them by:
  - **Traffic Source** (email, organic, social, etc.)
  - **Device Type**

Traffic Source Breakdown

- Donut and bar charts represent metrics by source:
  - Total sessions, session duration, and bounce rate
  - Referral and organic traffic performance

Geographical Insights

- Maps and bar charts visualize:
  - Traffic performance by **region** (North, South, East, West)
  - City-specific data (e.g., Chennai, Hyderabad)

Device and Browser Usage

- Charts show:
  - Device-wise performance and engagement
  - Browser-based metrics (e.g., Chrome, Firefox)

4. Insights & Outcomes

- **User Behavior Trends**: Identified which devices and locations yield better session durations and lower bounce rates.
- **Marketing Optimization**: Helped in understanding the effectiveness of different traffic sources and campaigns.
- **Technical Improvements**: Detected browsers or platforms causing higher bounce rates, indicating potential UI/UX issues.

Power Query's:

Bounce Rate by devices

```
with cte as (
select
d. Device_Type as device1,
count(Session_Id) as Total_Sessions
from device_lookup d
join website_traffic_data w
on d.Device_Key=w.Device_Key
group by d.Device_Type
)
select
d. Device_Type as device_type,
(count(Session_Id)/cte.Total_Sessions)*100 as Bounce_Rate
from device_lookup d
join website_traffic_data w
on d.Device_Key=w.Device_Key
join cte on d.device_type=cte.device1
where Page_Views_Per_Session <2
group by device_type
```

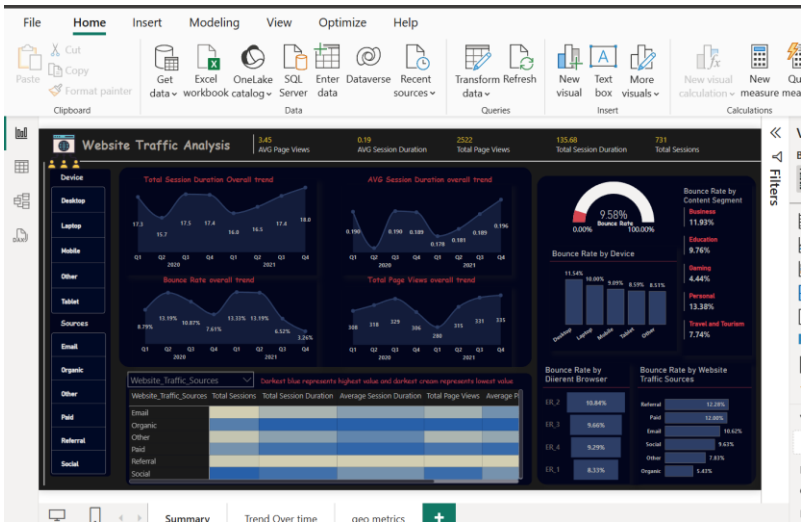
device_type	Bounce_Rate
Tablet	8.5890
Desktop	11.5385
Other	8.5106
Mobile	9.0909
Laptop	10.0000

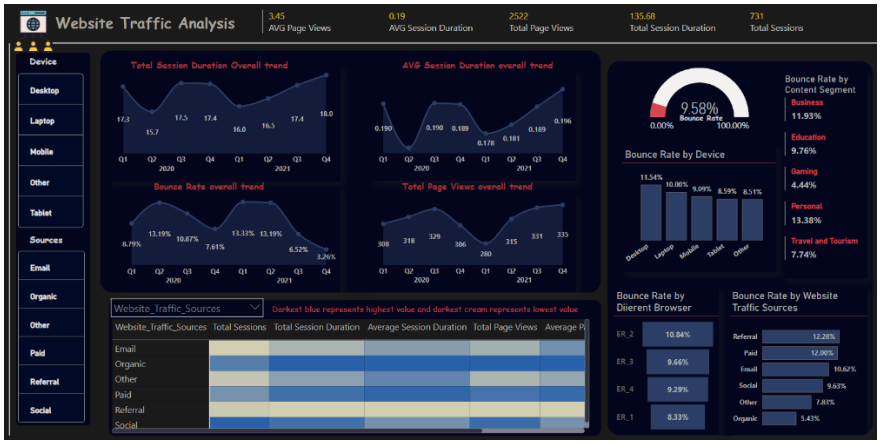
Bounce Rate by content segment

```
with cte as (
select
d.Content_Segment as segment1,
count(Session_Id) as Total_Sessions
from device_lookup d
join website_traffic_data w
on d.Device_Key=w.Device_Key
group by d.Content_Segment
)
Select
d.Content_Segment as Content_Segment,
(count(Session_Id)/cte.Total_Sessions)*100 as Bounce_Rate
from device_lookup d
join website_traffic_data w
on d.Device_Key=w.Device_Key
join cte on d.Content_Segment=cte.segment1
where Page_Views_Per_Session <2
group by d.Content_Segment
```

Content_Segment	Bounce_Rate
Personal	13.3803
Education	9.7561
Travel and Tourism	7.7419
Business	11.9318
Gaming	4.4444

Output's:





Inference:

1. **User Behavior Insights:**
   Session durations vary across devices, with mobile users showing shorter engagement. Certain regions exhibit higher traffic, presenting potential growth opportunities.

2. **Traffic Source Effectiveness:**
   Organic traffic and paid ads are performing well, but social media traffic needs optimization due to higher bounce rates.

3. **Optimization Recommendations:**
   Focus on improving mobile site performance and targeting high-performing regions with tailored content. Increased investment in organic and referral traffic can yield better engagement.

Conclusion:

The Website Traffic Analysis Dashboard enables clear visualization of user behavior and traffic trends, empowering data-driven decisions.
With Power BI's interactive features, businesses can efficiently monitor, analyse, and optimize website performance.