**Report for Smile Guru data**

**Data Cleaning in SQL:**

1. **Loaded the dataset, ensuring all values were in text format.**

* Imported the dataset into SQL, ensuring all data was initially treated as text to prevent automatic conversions or misinterpretations.
* This step helped maintain data integrity and allowed for controlled type conversions later in the process.

1. **Converted event\_date to date type and replaced null values with a valid date.**

* Checked for null values in the event date column and replaced them with a standard date to maintain consistency.
* Converted the column from text format to a proper date type to enable accurate date-based analysis and filtering.

1. **Added a unique ID column for each record.**

* Introduced an ID column to uniquely identify each record.
* This helped in tracking data modifications and ensuring each row had a unique identifier.

1. **Removed rows that had null values in three key columns.**

* Identified and removed records where critical fields (such as event\_name, device\_category, and page\_location) contained null values.
* Ensured data completeness and avoided misleading or incomplete insights in further analysis.

1. **Changed the data type of all columns to match their actual values.**

* Converted numerical columns to appropriate data types (e.g., integer, float) for efficient calculations.
* Standardized categorical fields to VARCHAR format for better readability and processing.
* Applied constraints where necessary to maintain data consistency.

1. **Filled remaining null values with "Unknown" to retain as much data as possible.**

* Instead of removing records with missing values in non-critical fields, replaced nulls with "Unknown" or default values.
* This minimized data loss while still ensuring meaningful analysis.

1. **Exported the cleaned dataset as a CSV file for visualization in Power BI.**

* After all cleaning steps, saved the final version of the dataset in CSV format.
* This prepared the data for further visualization and analysis in Power BI.

**Data visualization in power bi**

1. **Bar Chart – Pages with the Most Views**

* Highlights the most visited pages, with the **Main** and **About** pages receiving the highest views.
* Less visited pages include **Orthodontics and Emergency Dental**, indicating areas for potential content improvement.

1. **Tree Map – Count of Event Names by Page**

* Shows the distribution of user interactions on different pages.
* The **Main** page has the highest engagement, while other service pages like **Teeth Whitening and Root Canal** also see notable activity.

1. **Pie Chart – Event Type Distribution**

* **Page Views (24.03%)** are the most recorded events, indicating users frequently navigate through pages.
* **Click (23.55%) and Scroll (22.6%) events** show user engagement with content, while **Form Submissions (8.45%)** highlight conversion opportunities.

1. **Pie Chart – Device OS Distribution**

* **Android (28.49%)** and **Windows (28.3%)** are the most used operating systems.
* A significant number of users also access the site via **iOS (28.21%)**, emphasizing the need for cross-platform optimization.

1. **Pie Chart – Device Category Distribution**

* **Mobile devices (29.25%)** and **Desktops (28.02%)** are the most commonly used platforms.
* **Tablet users (26.02%)** also form a large portion, suggesting the importance of tablet-friendly design.

1. **Filter Pane (Slicers) – Page, Device OS, Device Category, Event Name**

* These filters allow users to **interactively analyze** data by selecting specific pages, devices, categories, or events.
* Users can apply multiple filters to get **custom insights** based on selected criteria.
* Helps in narrowing down the data to understand trends for particular pages or devices.

1. **KPI Card – Total Visitors (1053)**

* Displays the **total number of unique visitors**, providing a high-level metric of website traffic.
* Acts as a **performance indicator** to track website engagement.
* Can be dynamically updated based on filters applied in the dashboard.