**Healthcare No-Show Appointment Analysis**

**Introduction**

Missed medical appointments, or “no-shows,” can cause resource wastage and delays in healthcare services. This project analyzes patterns in patient appointment data to understand the reasons behind missed appointments and derive actionable insights. The focus is on gender-based behavior, age groups, and other critical factors contributing to patient no-shows.

**Abstract**

This project analyzes a real-world dataset collected from Kaggle titled *"Medical Appointment No-Shows."* The dataset contains information about over 100,000 appointments in Brazil, with various patient details such as gender, age, scholarship status, and whether they showed up for their appointment. Using Python for data cleaning and Power BI for visualization, we explore trends that may indicate which factors contribute most to patient no-shows. Two new columns — Male and Female — were also added using DAX formulas for targeted analysis.

**Tools Used**

* **Python Libraries**: Pandas, Sklearn
  + Data cleaning and preprocessing
  + Handling null values and type conversions
* **Power BI**
  + Visualizing insights using charts, filters, and dashboards
  + Created DAX-calculated columns for gender classification:
    - Male = IF(Gender = "M", "Male", BLANK())
    - Female = IF(Gender = "F", "Female", BLANK())

**Steps Involved in Building the Project**

1. **Data Collection**  
   The dataset was sourced from Kaggle, titled *Medical Appointment No-Shows*.
2. **Data Cleaning (Python)**
   * Removed irrelevant columns
   * Handled null and incorrect values
   * Converted date columns to datetime format
   * Verified consistency of "No-show" labels
3. **Data Processing (Python)**
   * Encoded categorical variables if necessary
   * Verified unique values for each attribute
   * Checked for outliers or unrealistic values (e.g., negative age)
4. **Data Import in Power BI**
   * Loaded the cleaned dataset into Power BI
   * Built interactive dashboards using bar charts, filters, and pie charts
5. **DAX Columns for Gender Classification**
   * Added two new columns:
     + **Male**: Displays "Male" if Gender is "M"
     + **Female**: Displays "Female" if Gender is "F"
6. **Visualization**
   * No-show rates by gender, age, and weekday
   * Number of patient as per gender and health problem
   * Gender-based attendance rates using the new columns

**Conclusion**

The analysis revealed that certain demographic groups (such as younger patients and females) showed distinct trends in missing appointments. Additional factors such as day of the week, scholarship status, and reminders also play significant roles. By integrating Python and Power BI, this project offers a scalable framework for healthcare data analysis. The added gender-specific columns improved the clarity and depth of gender-based insights, allowing for more precise visualization and planning.