**Medical Appointment- Updated Dataset**

**Description**

* This dataset contains information about medical appointments and whether patients showed up for their scheduled appointments. It is a cleaned and preprocessed version of the original Kaggle dataset. This updated version is ready for data analysis, visualization, and predictive modeling.

**Note**:  
"All data cleaning and preprocessing were performed using Microsoft Excel".

**Original Source**

The original dataset is from Kaggle: [Medical Appointment No Shows](https://www.kaggle.com/datasets/joniarroba/noshowappointments)

**Dataset Columns:**

* "PatientId": Unique identifier for each patient
* "AppointmentID": Unique appointment identifier
* "Gender": Gender of the patient (“Male” / “Female”)
* "ScheduledDay": The day and time the appointment was scheduled (formatted as “DD-MON- YYYY”)
* "AppointmentDay": Date of the actual appointment (date only)
* "Age”: Age of the patient
* “Neighbourhood”: Name of the hospital
* “Scholarship”: Whether the patient is enrolled in the Brazilian welfare program (“Yes” / “No”)
* “Hipertension”: Whether the patient has hypertension (“Yes” / “No”)
* “Diabetes”: Whether the patient has diabetes (“Yes” / “No”)
* “Alcoholism”: Whether the patient is an alcoholic (“Yes” / “No”)
* “Handcap”: Whether the patient has a handicap (“Yes” / “No”)
* “SMS\_received”: Whether the patient received an SMS reminder (“Yes” / “No”)
* “No-show”: Appointment status (“Attended” / “Missed\_Appointment”)

**Changes Made:**

* + Updated “Gender” values from “F” / “M” to “Female” / “Male”
  + Reformatted “ScheduledDay” to include day, month, year; time adjusted
  + Reformatted “AppointmentDay” to include date only
  + Changed “Neighbourhood” values to hospital names
  + Converted binary columns (“Scholarship”, “Hipertension”, “Diabetes”, “Alcoholism”, “Handcap”, “SMS\_received”) from “0/1” to “No/Yes”
  + Converted “No-show” values from “0/1” to “Missed\_Appointment/Attended”
  + Removed all duplicate rows
  + Ensured no null values remain in the dataset

**Usage:**

This cleaned dataset can be used for:

* + Data analysis
  + Predictive modeling (classification: “Attended” / “Missed\_Appointment”)
  + Visualizations of appointment patterns
  + Healthcare analytics projects

**File Details**

* + Medical\_appointments\_updated.csv”: Cleaned and preprocessed dataset (Excel used for all changes)