

Health Care EDA

29.08..2023

Satyajit Pattnaik


Pragra

Overview

Perform EDA and generate insights for the given dataset.

About the dataset

File name: Data.csv

- 
1. **PatientId**: This column represents the unique identifier for each patient. It is used to differentiate between different individuals in the dataset.
 2. **AppointmentID**: This column contains the unique identifier for each appointment. It is used to differentiate between different appointments made by the patients.
 3. **Gender**: This column indicates the gender of the patient. It can have two possible values: "Male" or "Female".
 4. **ScheduledDay**: This column represents the date and time when the appointment was scheduled. It provides information about when the patient made the appointment.
 5. **AppointmentDay**: This column represents the date when the appointment is scheduled to take place. It provides information about the date on which the patient is expected to visit the healthcare facility.
 6. **Age**: This column represents the age of the patient. It contains numerical values indicating the age in years.
 7. **Neighbourhood**: This column represents the neighborhood where the healthcare facility is located. It provides information about the geographical location of the facility.
 8. **Scholarship**: This column indicates whether the patient is enrolled in the Brazilian welfare program Bolsa Família. It can have two possible values: "1" if the patient is enrolled and "0" if not.
 9. **Hipertension**: This column indicates whether the patient has hypertension (high blood pressure). It can have two possible values: "1" if the patient has hypertension and "0" if not.
 10. **Diabetes**: This column indicates whether the patient has diabetes. It can have two possible values: "1" if the patient has diabetes and "0" if not.
 11. **Alcoholism**: This column indicates whether the patient is an alcoholic. It can have two possible values: "1" if the patient is an alcoholic and "0" if not.
 12. **Handicap**: This column represents the level of handicap/disability the patient has. It can have values ranging from 0 to 4, where 0 indicates no handicap and higher values indicate higher levels of handicap.
 13. **SMS_received**: This column indicates whether the patient received an SMS reminder for the appointment. It can have two possible values: "1" if the patient received an SMS reminder and "0" if not.
 14. **No-show**: This column indicates whether the patient showed up for the appointment or not. It can have two possible values: "Yes" if the patient did not show up and "No" if the patient showed up.

Deliverables

Perform the EDA in a python notebook, and prepare a presentation talking about the insights, and upload the same along with the data in a GitHub repository.