**Introduction**

In this project, we will be analyzing the data that collected from 100k medical appointments in Brazil, by focusing on the main question of whether or not patients show up for their appointment. A number of characteristics about the patient are included in the row’s dataset such as (AppointmentID, Gender, Age, Neighbourhood, Scholarship, Hipertension, Diabetes, Alcoholism, Handcap, SMS\_received, No-show) we will be associate them in order to answer these questions:

* How many patients, whether they are male or female who show up their appointments?
* Did the patients who show up their appointments received a reminder such SMS message?
* Did the patients who show up their appointments have chronic illness such as hipertension and diabetes and how old are they?
* Did the patients who have alcoholism and they have scholarship show up their appointments?

‘ScheduledDay’ tells us on what day the patient set up their appointment.

‘Neighborhood’ indicates the location of the hospital.

‘Scholarship’ indicates whether or not the patient is enrolled in Brasilian welfare program Bolsa Família.

Be careful about the encoding of the last column: it says ‘No’ if the patient showed up to their appointment, and ‘Yes’ if they did not show up.

|  |  |  |  |
| --- | --- | --- | --- |
| ScheduledDay | AppointmentDay | SMS\_received | No-show |