Codebook for Attrition Dataset Data Overview

## Credentials

The direct link to data is: <https://www.kaggle.com/joniarroba/noshowappointments>

## Business goal

This data was collected to predict if someone is to no-show an appointment?

## Data description

This data set is a data frame of 14 variables over more than a 110k rows. Each row represents an appointment of patient ID that is sold by the company.

## 

## Variables description

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **ID** | | **Variable Name** | **Description** | **Type** | **Possible values** |
| 1 | PatientId | | Patient ID - random UUID | Numeric | Random ID |
| 2 | | AppointmentID | Appointment ID - random UUID | Numeric | Random ID |
| 3 | | Gender | Gender of the patient | Textual | F=Female M=Male |
| 4 | | ScheduledDay | Scheduled day of appointment | DayTime |  |
| 5 | | AppointmentDay | Actual appointment day | DayTime |  |
| 6 | | Age | Age of the patient | Numeric | -1(NA) : 115 |
| 7 | | Neighbourhood | Neighbourhood name (Brazil) | Textual | Name |
| 8 | | Scholarship | Social welfare scholarship received? | Binary | 0=No  1=Yes |
| 9 | | Hipertension | Hypertension existence | Binary | 0=No  1=Yes |
| 10 | | Diabetes | Diabetes existence | Binary | 0=No  1=Yes |
| 11 | | Alcoholism | Alcoholism existence | Binary | 0=No  1=Yes |
| 12 | | Handcap | Amount of handicaps a person presents | Numeric | 0 : 4 |
| 13 | | SMS\_received | SMS reminder received? | Binary | 0=No  1=Yes |
| 14 | | No-show | Show/No show to the appointment | Textual | No/Yes |

## Business questions

1. Does the age and medical condition of the patient affect the no-show?
2. Is there a relationship between no-show to a specific doctor?
3. Are there more no-shows on a specific time of the year?
4. Is there a relationship between the duration from schedule to appointment to the probability of no-show?
5. Does the patient neighborhood can predict a no-show?

## Who needs to review the business question?

The medical center management

## Related links

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3962267/>

<https://rstudio-pubs-static.s3.amazonaws.com/260312_b0f48fe809e44515a168cf5a2ce5ac75.html>