

Project-2: No-Show Appointments: Udacity

Contents

Introduction

Data Wrangling

Exploratory Data Analysis

Conclusions

Introduction

This dataset collects information from 100k medical appointments in Brazil and is focused on

the question of whether or not patients show up for their appointment. A number of characteristics

about the patient are included in each row

This dataset has 14 characteristics about the patient as PatientID, AppointmentID, Gender,

Scheduled Day, Appointment Day, Age, Neighbourhood, Scholarship, Hypertension,

Diabetes, Alcoholism, Handicap, SMS_Received and No_Show.

= important for us to know in order to predict if a patient will show up for their scheduled appointment?

Question : what is the relation between Count of alcoholism and no show ?

Question : How does Receiving an SMS have an effect on the patient's show up ?

Question : what is the relation between (gender type, the age) and the patient's show up ?

Question : How Scholarship affect the patient's show up ?

last Question : what is the relation between handicap and no show ?

Data Wrangling

General Properties

Data Cleaning

Dropping (PatientID, AppointmentID)

Deleting outliers and wrong data

Typo's in column's name should be corrected

Converting some columns' types to the correct ones

Exploratory Data Analysis

Data Cleaning

Deleting outliers and wrong data

Typo's in column's name should be corrected

Converting some columns' types to the correct ones

Dropping (PatientID, AppointmentID)