

Investigate no show appointments

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. '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 Brazilian 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.

Data Wrangling

I loaded the dataset and explored the counts of many features and figured out the trend between SMS_received and no-show appointments and extract the days of week to explore if there is trend with the no-show and extract the hours to exclude the extreme values .

cleaned data

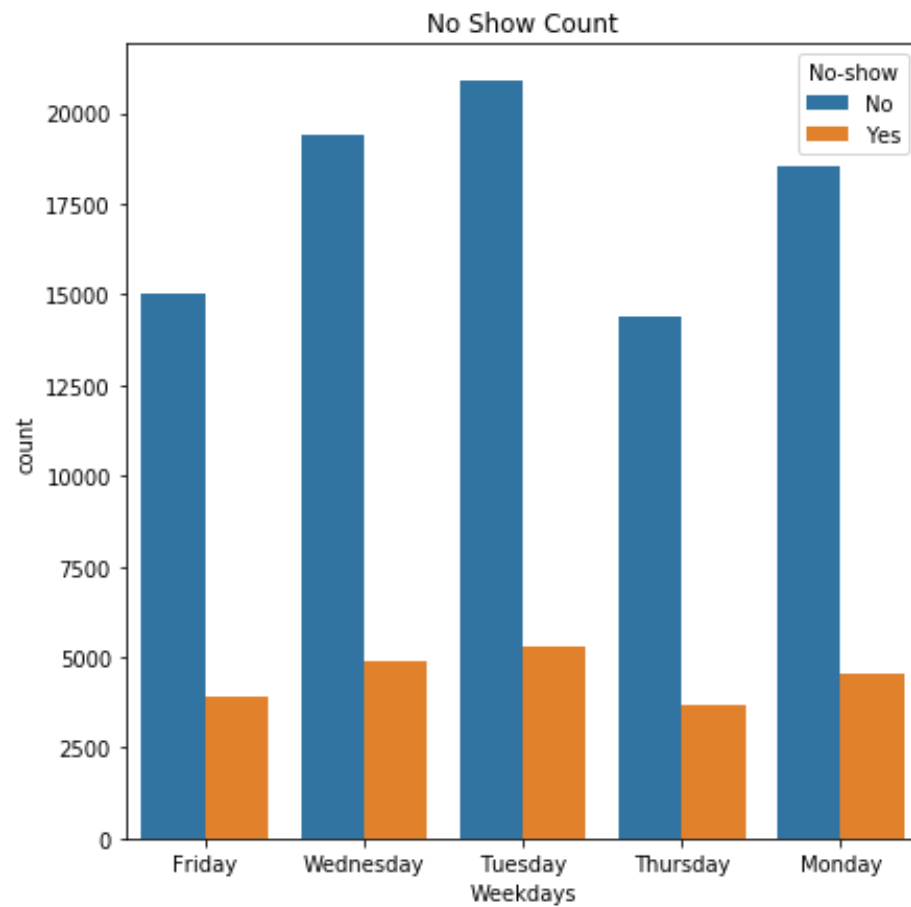
The Age less than 0 and more than 97 have been excluded .

The saturday has been excluded and it is only 24 day .

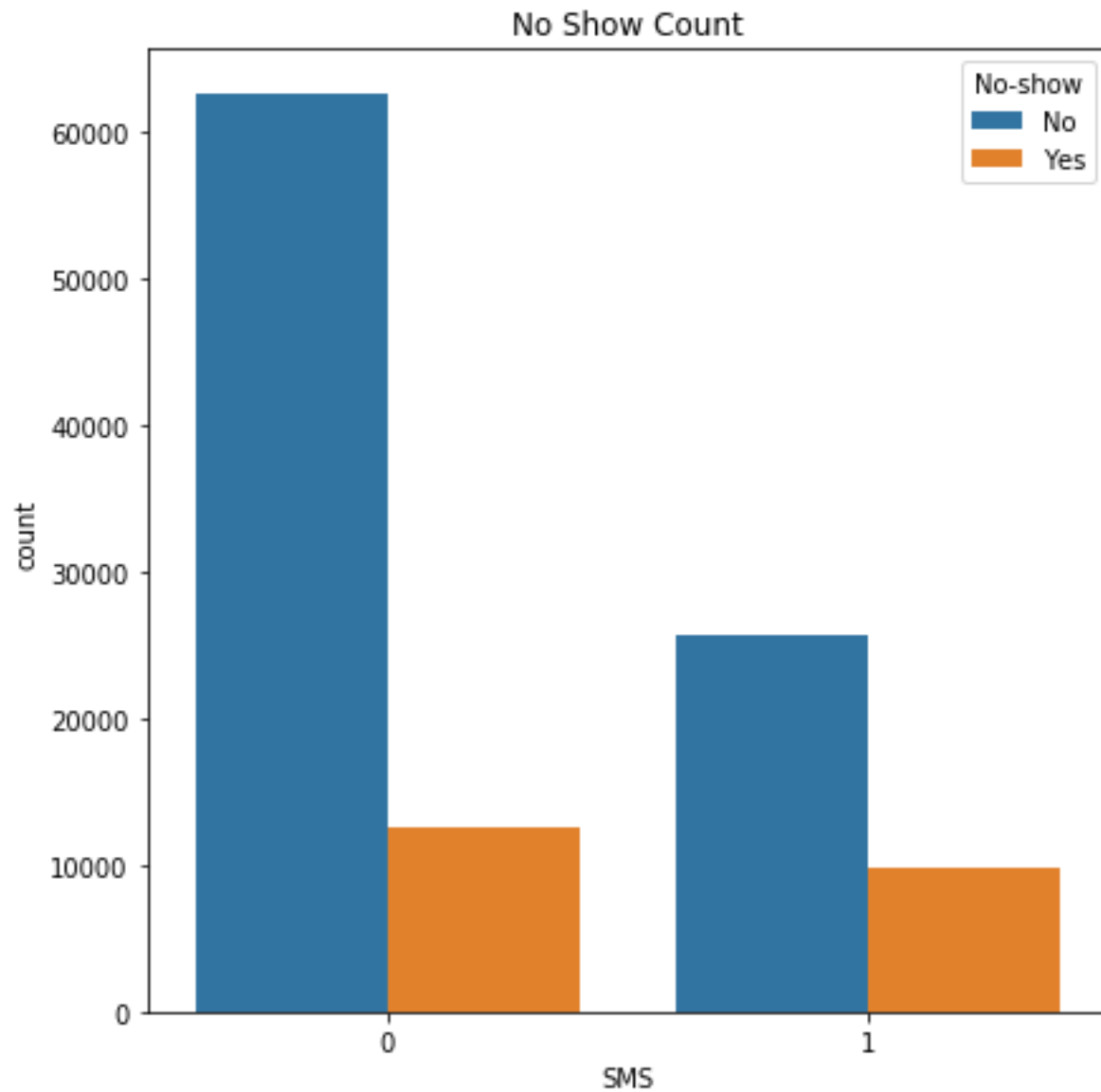
The extreme values of hours have been excluded and it's only 3 rows .

Research Questions

- Are there any trends between day of the week of the scheduled-appointment date and the No-Show rates?



- Are there any trends between SMS_received and the no-show rates?



Conclusions

Obviously there is no trend between days of the week and no-shows but its very clear between SMS and no-shows .