

Summary

In this project i investigate a dataset called 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.
- 'No_show' it says 'No' if the patient showed up to their appointment, and 'Yes' if they did not show up.

Progress outline:

using data analysis proccess to explore a data set answer questions we were interested on.

1-Clean data

2-explore data

3-answer questions by visualize the data

4-make conclusions

- you can navigate the data with IPythonpeople show more than people not show

Conclusions

female who will show is more than male

people not have Scholarship more than people have Scholarship in show

people not have Hypertension more than people have Hypertension in show

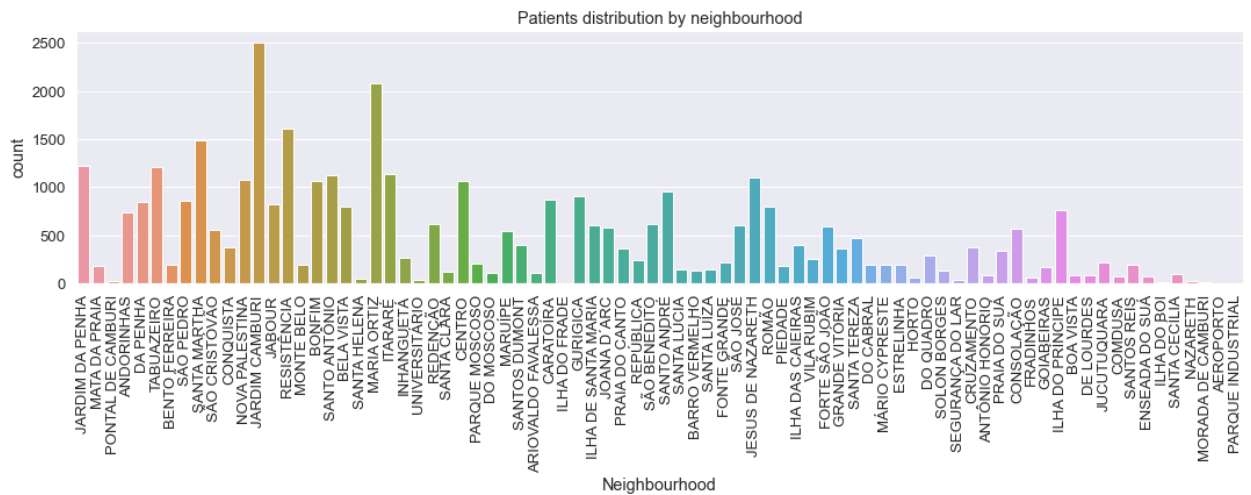
people not have Diabetes more than people have Diabetes in show

people not have Alcoholism more than people have Alcoholism in show

people not have Handicap more than people have Handicap in show

people not received him sms more than people received him sms in show

Number of patients varies from a neighbourhood to another, so lets investigate the ratio of not showing up The largest number found in (jardim camburi,mariaortiz,resistencia,santamartha)



age on 0 to 10 show more than any group ,,,,

people have [Scholarship Hypertension Diabetes Alcoholism Handicap] not show more than have not ,,,,,

people not received him sms show more than received him ,,,,