

## Udacity – ALX Data Analysis Project 1

### Exploratory Analysis of the No Show Appointment Kaggle Data

**Introduction:** This analysis is based on a dataset from Brazil containing about 100k medical appointments. The goal is to find out how patients respond to their appointment and to also investigate on factors that influence the various responses.

Neighborhood indicates where the hospital is located

Scholarship indicates if the patient is enrolled in Brazilian welfare or not.

1 stands for something being present while

0 stands for it's absence

For the No-show column, No stands for a show up for the appointment while yes stands for absence at the appointment.

### Questions Posed:

1. How many patients were living with different disease conditions and which is the most common?
2. What's the relative number of male patients compared to females?
3. What role does the Scholarship and SMS notification play in the responses of patients to return for appointments? Does it improve responses?
4. Do Males tend to neglect appointments more than females?

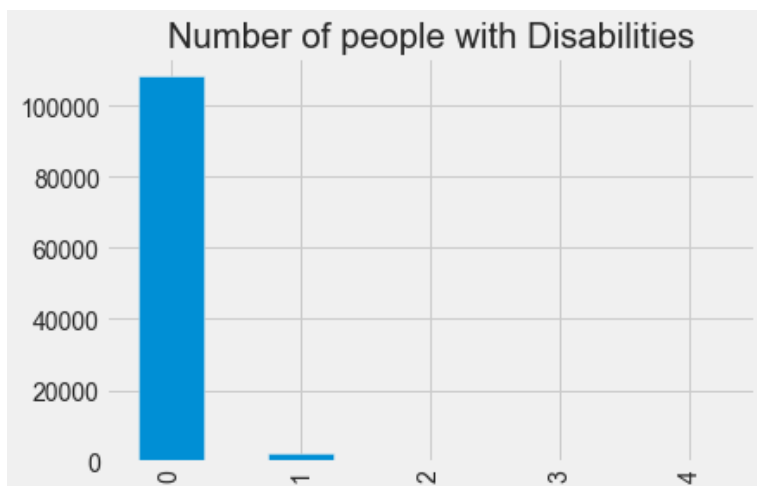
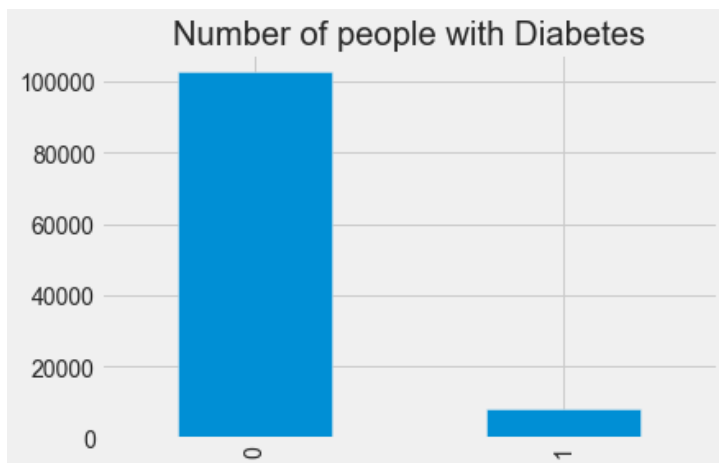
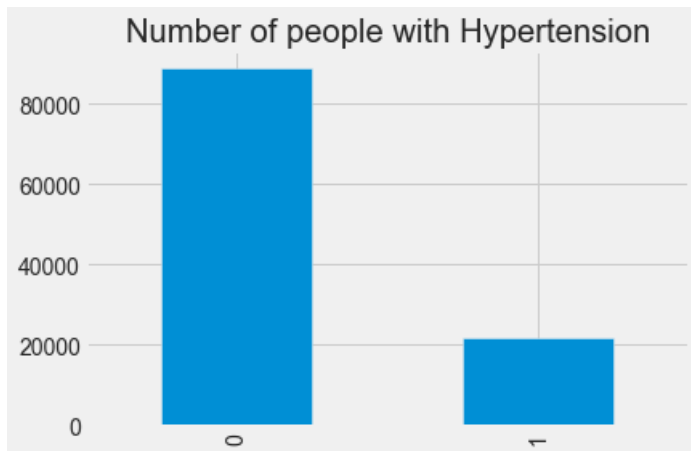
To investigate the questions, I mostly used charts to examine the relationships between all variables involved in giving insight to the questions posed.

### Data Wrangling

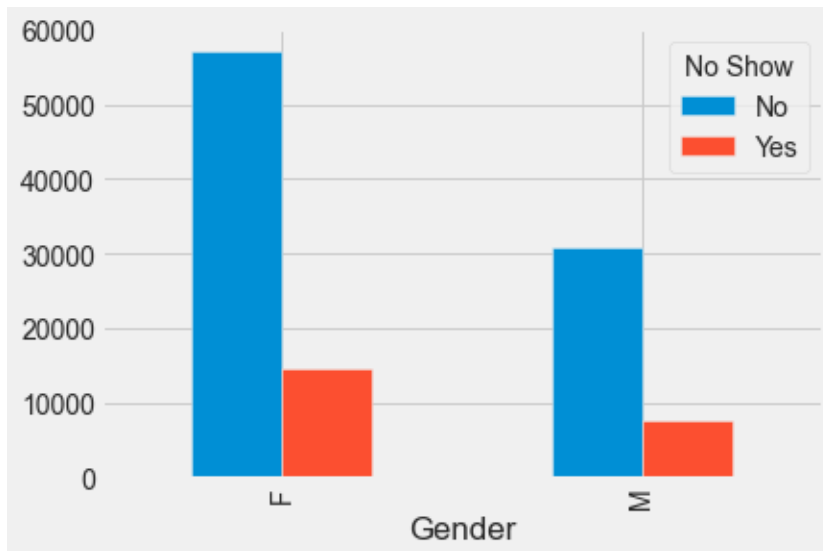
- Renaming of few columns for ease of use.
- Missing values were checked for
- Duplicates were checked for
- Outliers were checked for and fixed

## Conclusion

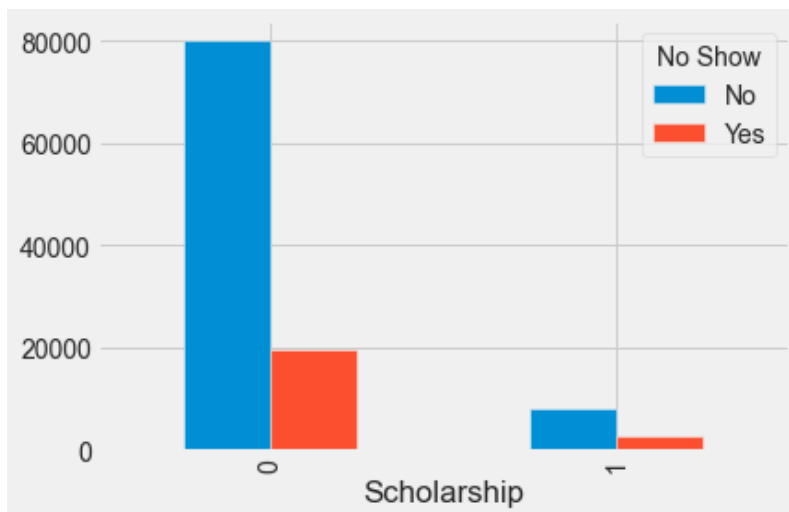
1. From the exploratory data analysis, we're able to answer some of the posed questions.  
Hypertension is the most common disease condition of the three mentioned.



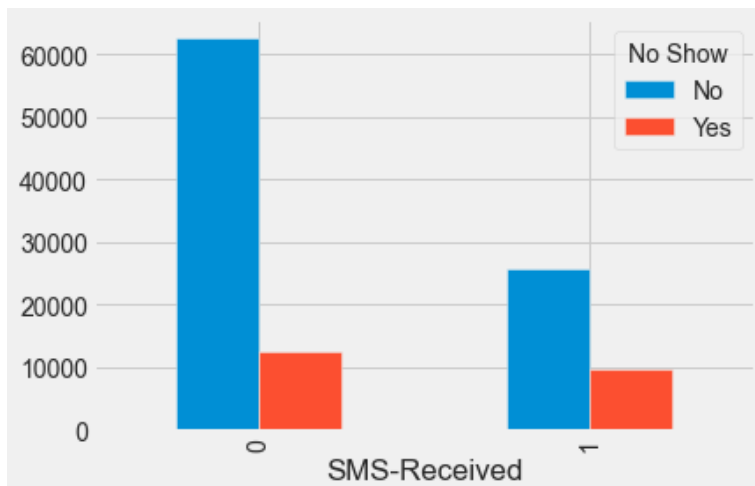
2. There are more of female patients compared to the male and more of the female percentage show up for their appointment compared to the male.



3. A higher percentage of those not on scholarship showed up for their appointment while a lesser ratio of those on scholarship showed up perhaps people value what they pay for, more than what is free.



4. From the analysis, male patient neglect their appointment compared to the female



**Limitation:** time spent at the various hospital locations were not recorded. This might have helped to know if time spent at the hospital can serve as a motivation as to whether or not to show up.