PATIENT CARE INSIGHTS

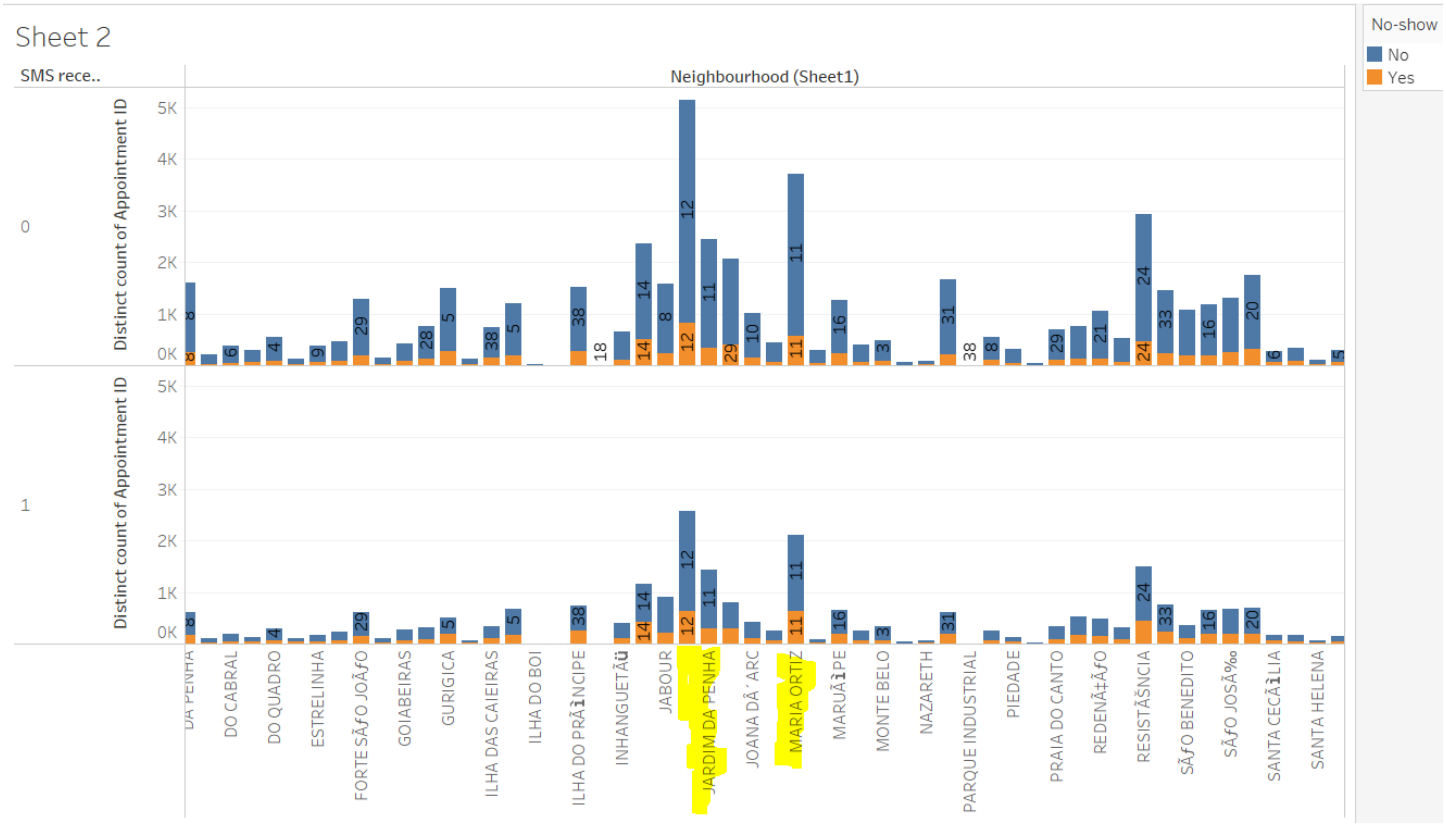
By:
Rakshit Soni
Texas A&M University
Management Information systems – Class of 2021

OBJECTIVE: The objective of this report is to explore and speculate the reasons why the patients might not be keeping their appointments and estimate if they will in the future.

1)This is a plot between neighborhoods and distinct count of Appointment IDs.

https://public.tableau.com/profile/rakshit.soni#!/vizhome/Book4_15915720600340/NeighbourhoodVSSMSs

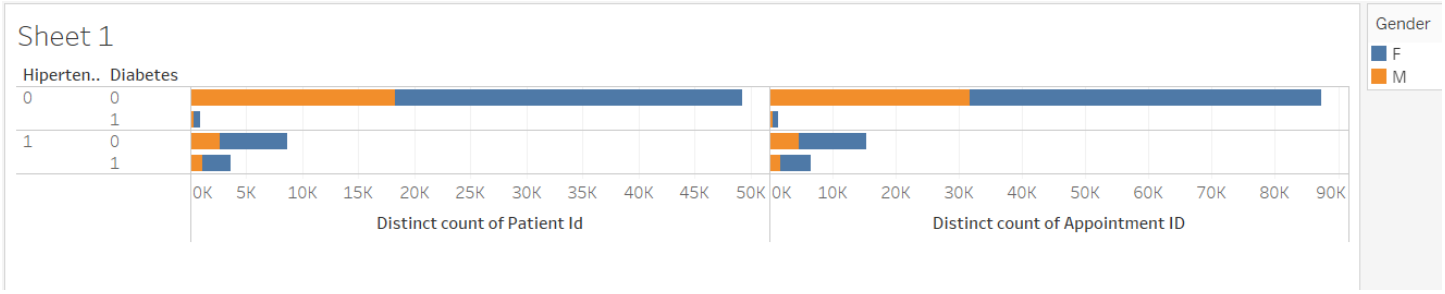
a) It seems as there are two or three places that have many patients which are not showing up for the appointments.
b) This might be because there might be a more competitive hospitals available in these neighborhoods. Moreover, we can see that the number of appointments has decreased for people who got SMSs.
c) This is probably because they were timely reminded and attended their meetings.
d) We can explore this further and even provide some sort of compensation for these patients to keep their appointments with our hospitals.



2) **Variation of Hyper-tension and Diabetes with patient ID(M/F)**

https://public.tableau.com/profile/rakshit.soni#!/vizhome/Viz1_15915721524670/PatientIDVsAppointmentID

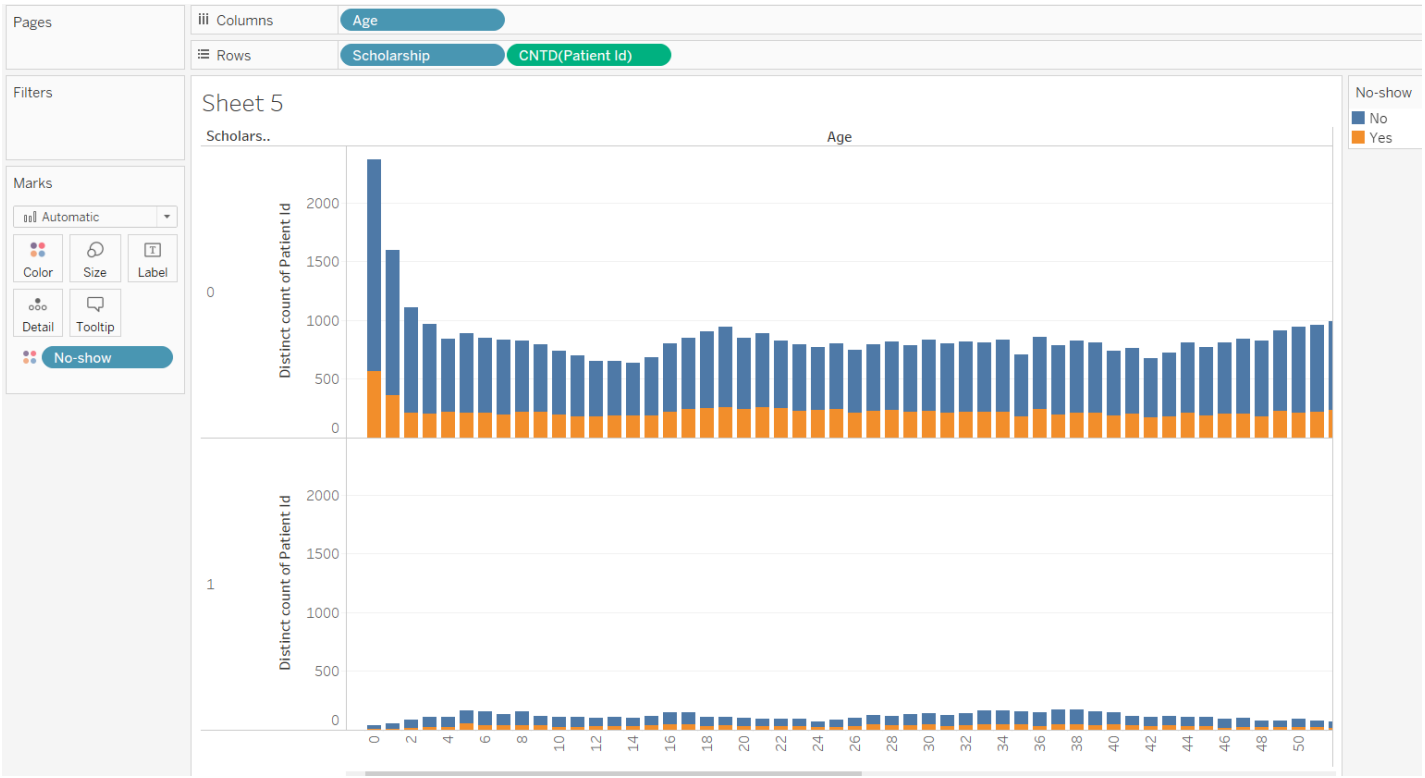
a) It seems that people who have diabetes and hypertension are booking equivalent number of appointments.
b) People who have none are booking more than double appointments
c) It seems that many people who are booking appointments and have neither of the conditions are a bit more liberal with attending of their appointments.



3)Co relation of Scholarship and No-show

https://public.tableau.com/profile/rakshit.soni#!/vizhome/Book3_15915720240270/AgeVSScholarships

a) This was an attempt to find out if there is any relationship between No-show and scholarship. Here I have assumed that scholarship is some sort of % discount on final appointment fee.
b) We can see a spike in the beginning i.e. a lot of the patients are young babies between 1 to 4 who were not given discount and consequently were a no show to the appointments probably because they found a better and cheaper alternative somewhere else.
c) Also, people who were given scholarship and showed up to the appointments were far less compared to people who were not given scholarships and still showed up to the appointments.
d) The distribution of scholarship seems very unbalanced with respect to age. A better distribution of scholarship can surly bump up the number of people who would attend the appointments.

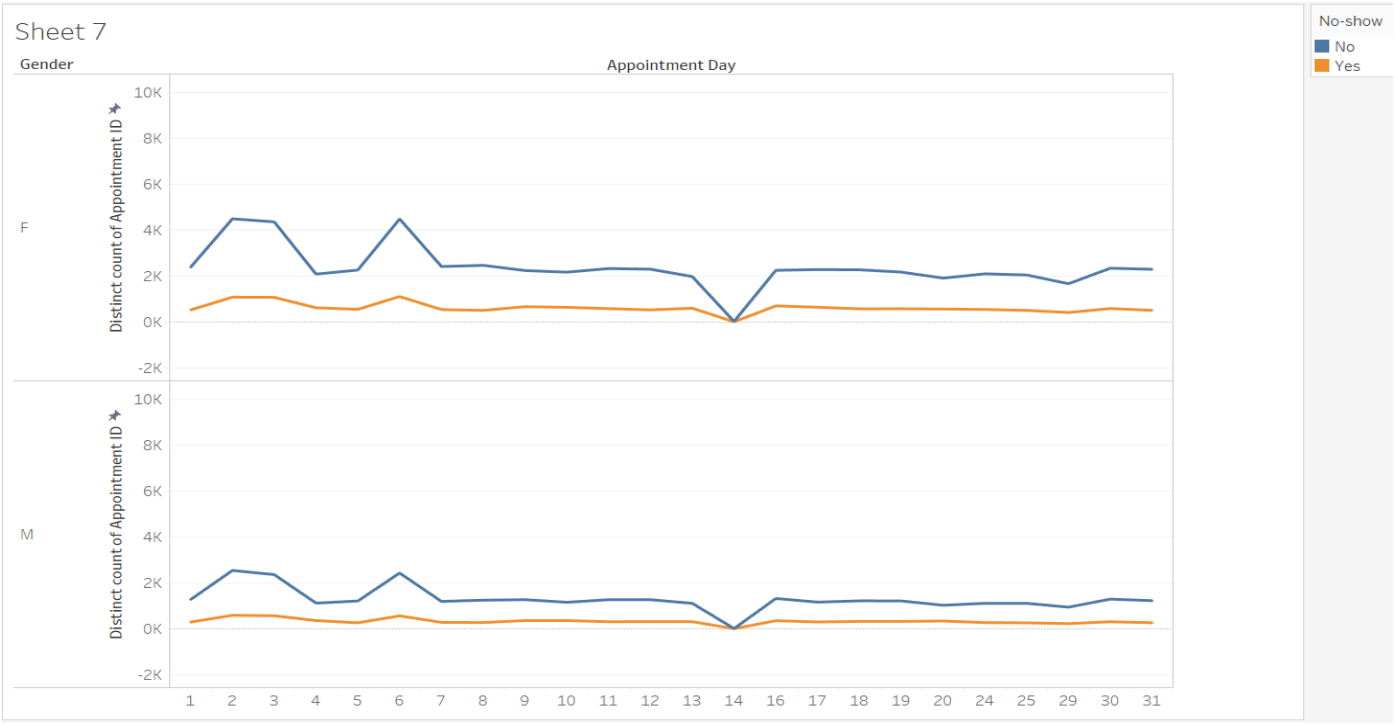


4) Trend for appointment made and if they showed up or not segregated by gender.

https://public.tableau.com/profile/rakshit.soni#!/vizhome/Book2_15915719489570/NoShowvariationbygender

a) The graph below shows that many appointments that were made were ultimately a no show.

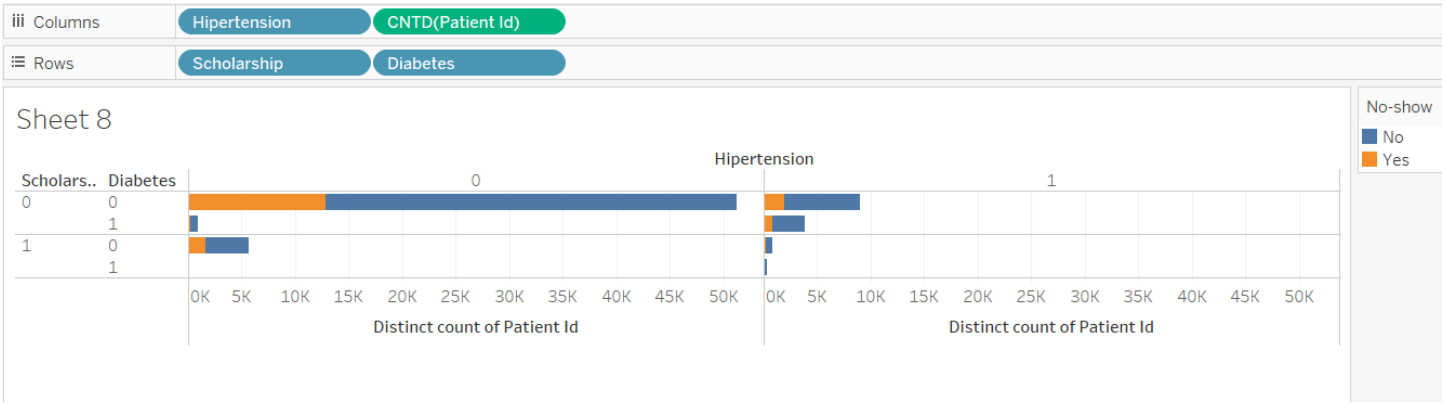
b) Majority of these appointments were by female patients.
c) Females who did not show up for the appointments were above 2K mark as compared to their male counter parts who were mostly below 2K.
d) The hospitals might be able to influence these numbers with a more favorable distribution of scholarships towards women.



5) Variation of people who have scholarship V/S Diabetes and Hypertension.

https://public.tableau.com/profile/rakshit.soni#!/vizhome/Book1_15915718736710/PatientIDVSDiabetesandHipertension

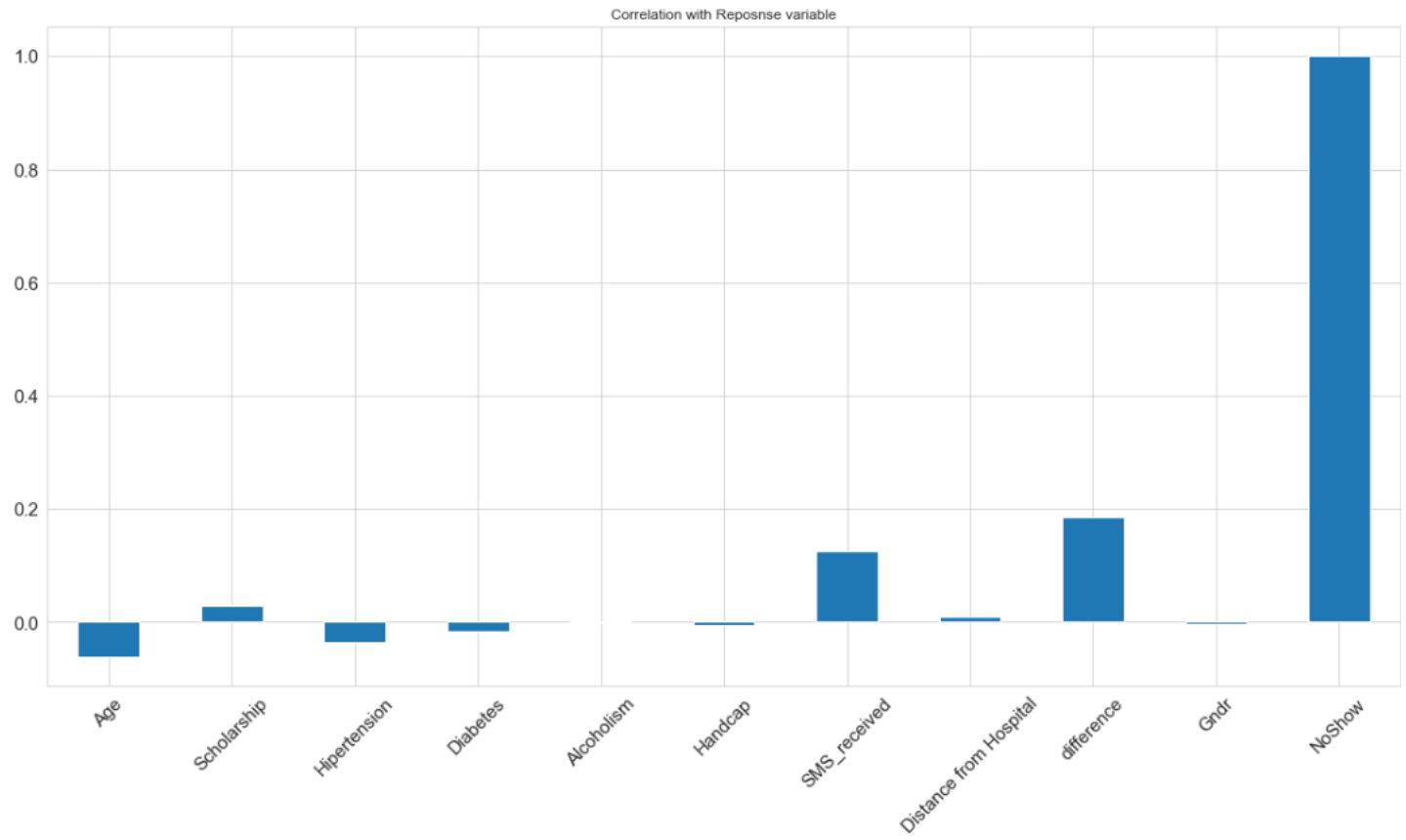
a) We see that 5000+ people who are neither diabetic nor have hyper tension get scholarship.
b) Out of which almost 4000 people did not show up.
c) This shows that business needs to give more scholarships to people who have both these conditions. That will also reduce the no show paramater.
d) Below is the graph for reference.



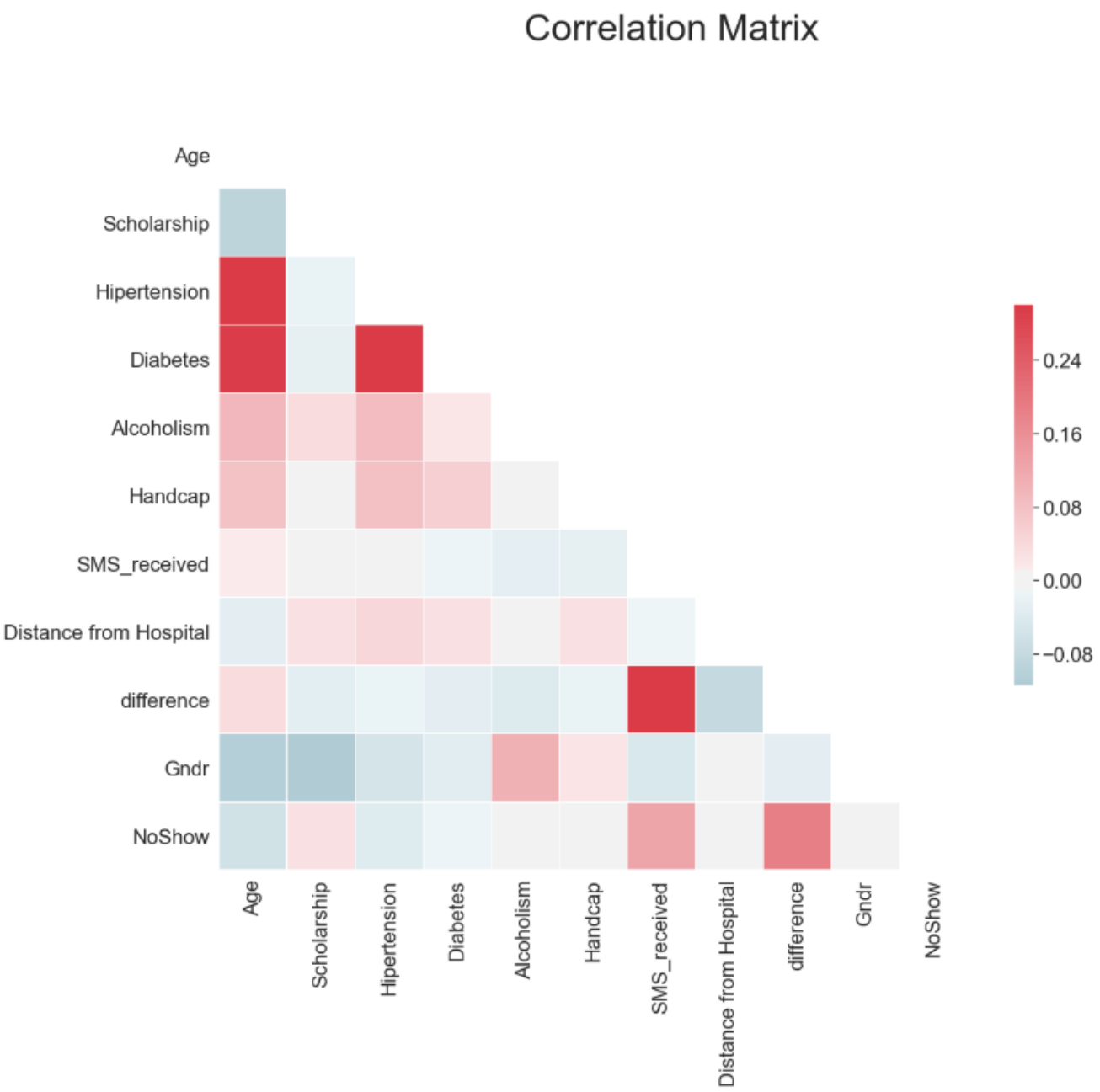
6) Predicting if the patient will show up (python code enclossed seperately)

The prediction includes exploratory data analysis. Conversion of binary values into 1s and 0s so it could be fed to the modal and application of Logistic regression and Naïve bayes classifier.

Exploratory data analysis



a) This correlation graph shows that SMS received and difference in the day is positively associated with people who showed up.
b) The more people received SMSs the more they were likely to show up. Similarly, the more time between appointment and schedule date the more likely patients will show up.
c) A policy that patients must have 2-3 days between schedule date and appointment date (unless an emergency) can help here.
d) Also, the IT system that the hospital is using must have automatic SMS triggering capability once an appointment is scheduled.
e) This will probably remind the patients that they have appointment and they won't forget and hence will show up.



We can see a strong co relation of hypertension and diabetes with age. This is common since with age these complications generally do happen.

Logistic Regression modal:

```
model F1-score in %: 70.98655453966853
model Accuracy-score in %: 79.2786272203625
model Precision-score in %: 70.36241996096452
```

This modal was chosen because of its simplicity. It predicts with 79% accuracy and has a descent F1 score

Naïve Bayes Classifier

```
Gaussian Naive Bayes model accuracy(in %): 77.35154859917368
Gaussian Naive Bayes model F1-score in %: 72.76119824480196
Gaussian Naive Bayes model Precision-score in %: 71.19185630091354
```

\

F1 score of Naïve Bayes is slightly better with a slightly decreased accuracy.

Thus, the predictions are 77% accurate.

REFERENCES

1. Python notebook

Patient care insights.pdf

2. Tableau Visualizations

https://public.tableau.com/profile/rakshit.soni#!/