

No-Show Appointments

Questions:

What factors impacted the attendance of appointments?

What kind of impact did pre-existing health issues have on appointment attendance?

What does sending SMS reminders help with appointment attendance?

Does the day of an appointment affect attendance?

Analysis:

First, I checked for duplicated rows. With the un-modified dataset of no-show appointments, there wasn't any duplicates. If there were duplicates, that would introduce false frequencies into the analysis. There were a couple of indexes that were incorrectly spelled or were in a format that would make analysis more cumbersome: 'Hipertension' to 'Hypertension', 'Handcap' to 'Handicap', 'No-show' to 'NoShow', 'Neighbourhood' to 'Neighborhood'. I then evaluated the dataset for factors that would have no direct impact to the patient: PatientId, AppointmentID, Age, Gender, ScheduledDay, Scholarship, Neighborhood. These factors that were deemed unimpactful were dropped from the dataset. Finally, before actual analysis, the data type in the AppointmentDay index was converted to datetime objects to make analysis easier later on.

For analysis, the data for whether or not a patient has health complications or any addictions was separated into a class of data where patients show or don't show, then from that, I separate once again by whether or not the patient had that complication or addiction. From there, I did a histogram of this info. For health complications and addictions, the dataset are binary, so the values are either true or false and made it simple to use a histogram as it shows us the frequency of either.

Plots:

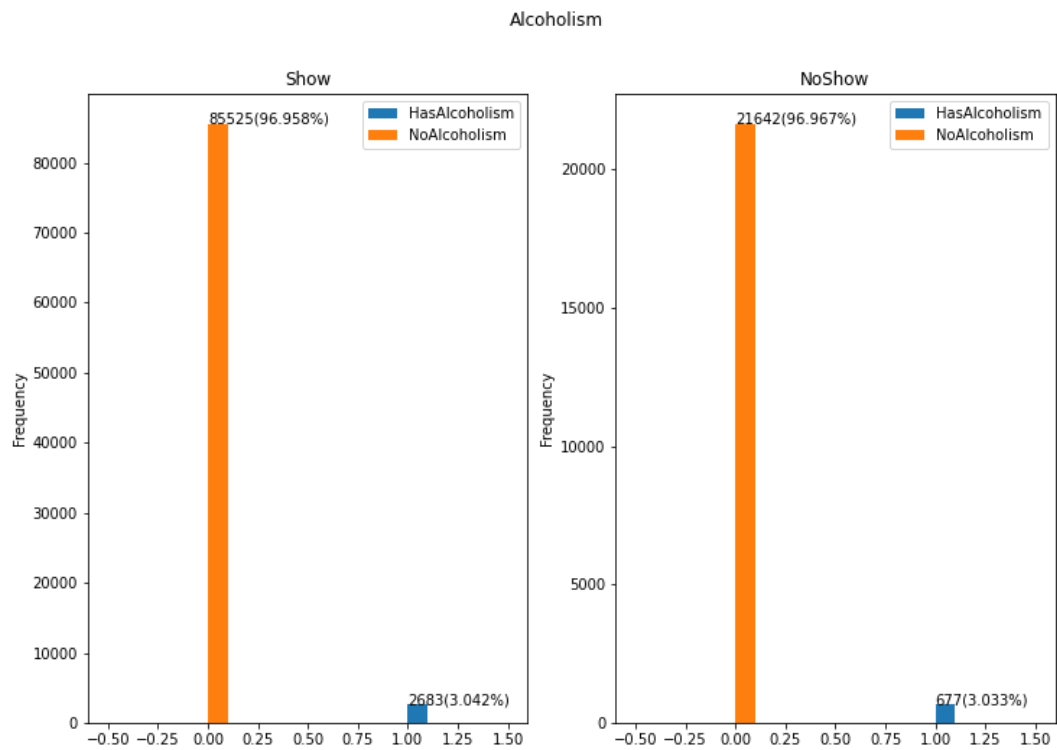


Figure 1: Alcoholism in patients who show vs noshow

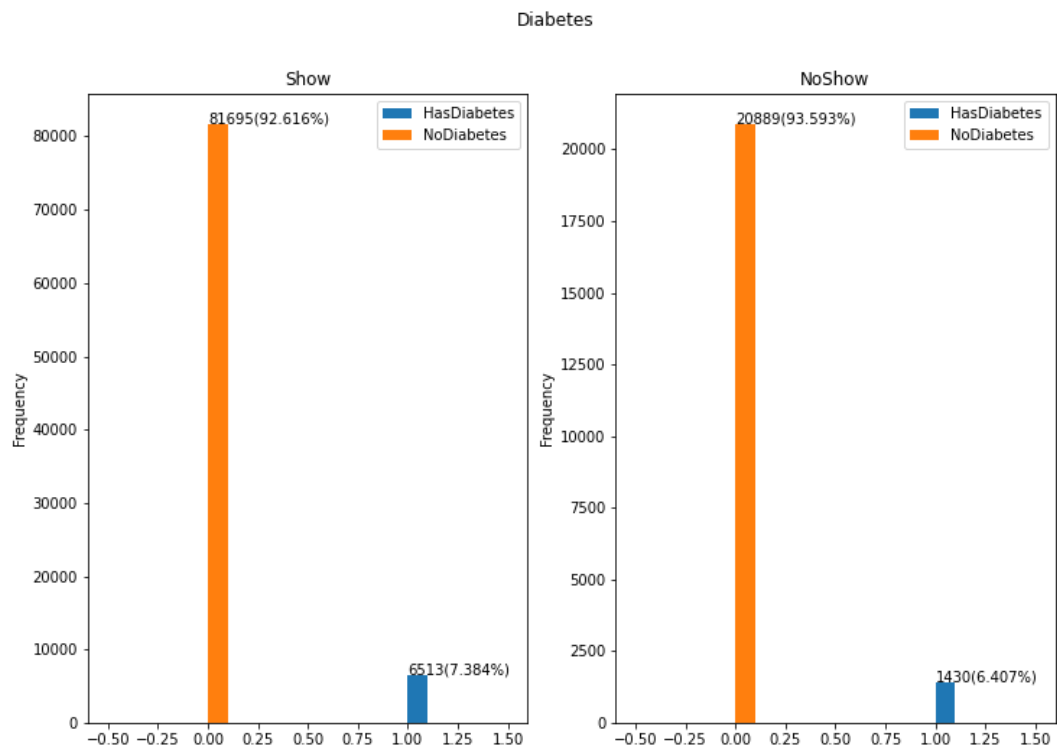


Figure 2: Diabetes in patients who show vs noshow

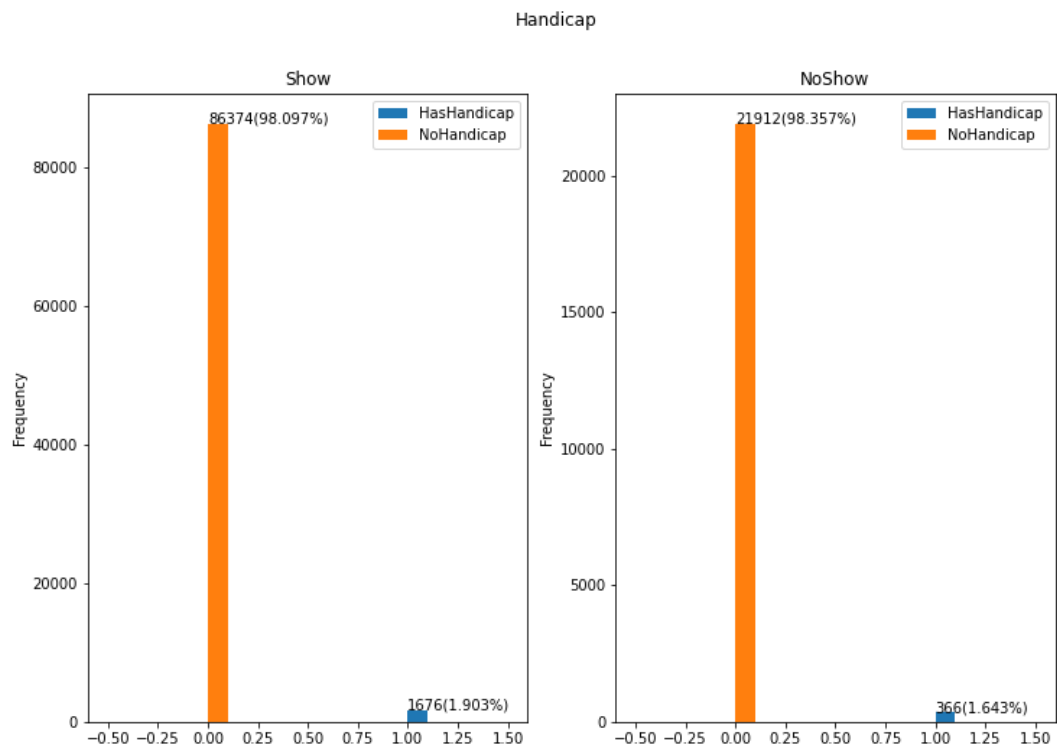


Figure 3: Patients who show vs noshow with a handicap

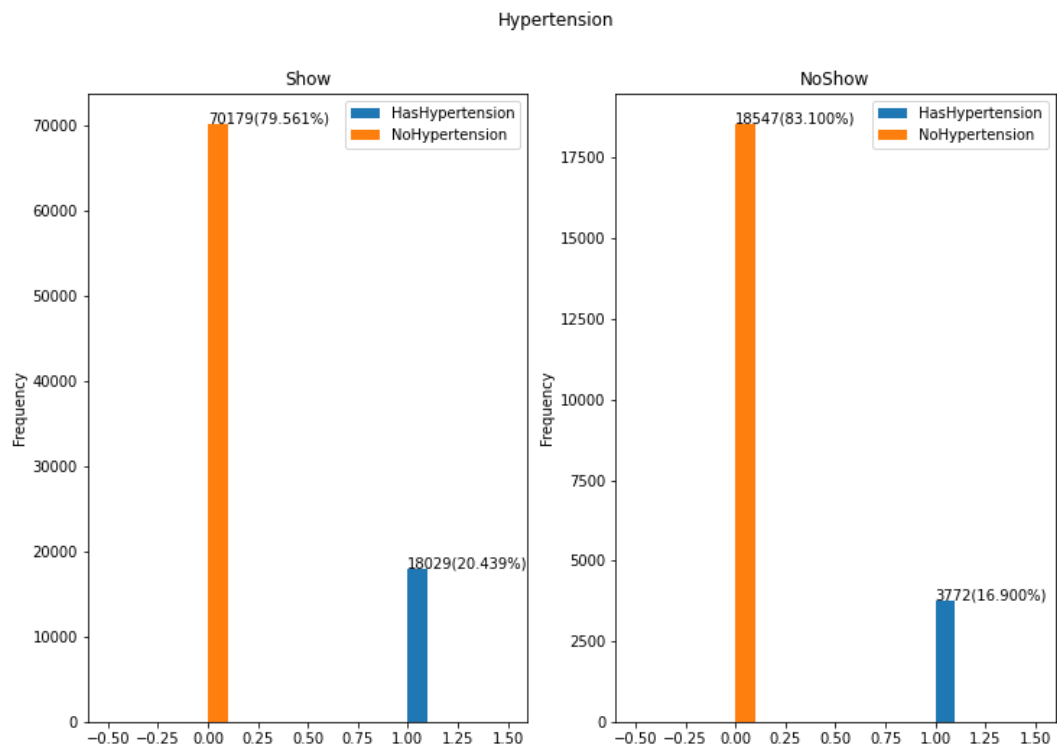


Figure 4: Hypertension in patients who show vs no show

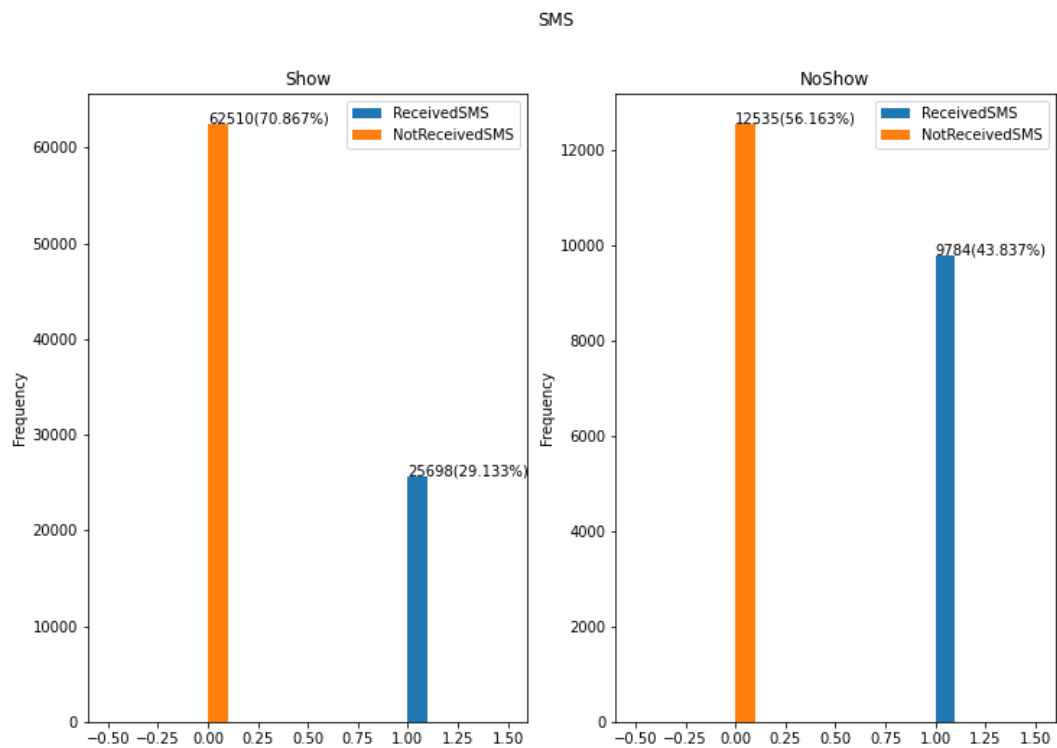


Figure 5: Patients who show vs noshow if receiving SMS reminders

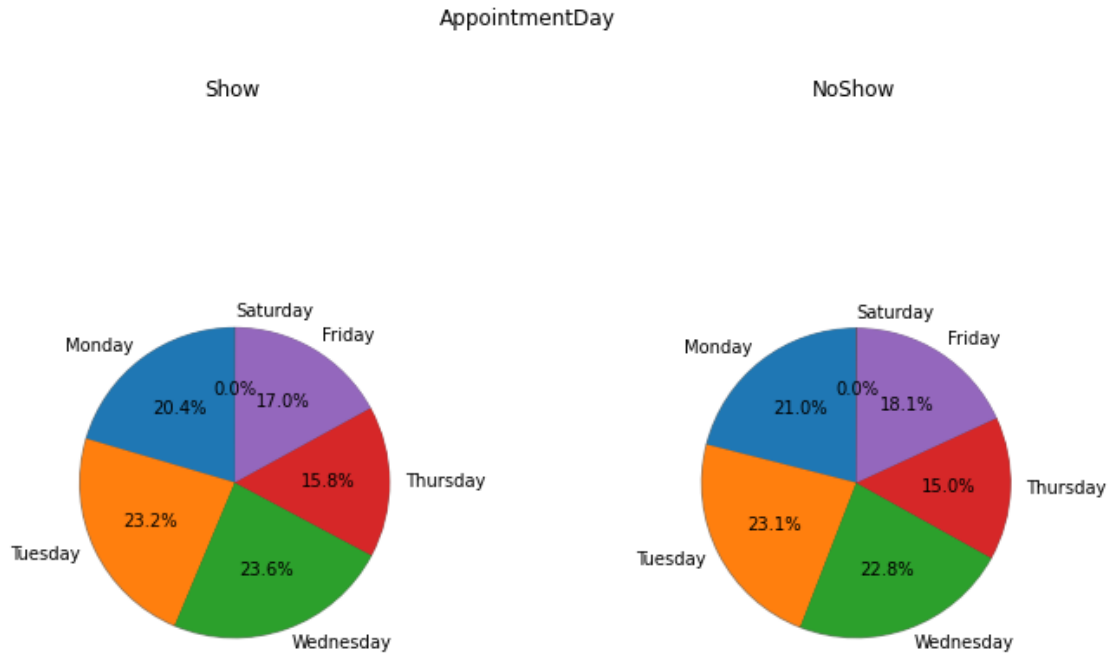


Figure 6: Patients who show vs no show on certain days

Answers:

What factors impacted the attendance of appointments?

Contrary to initial belief, having health complications did not negatively affect the patients' appointment attendance. I would have expected a higher percentage of no shows from patients that were positive with health complications, but that was not the case. Instead, there was an increased percentage of people that did not show up to appointments if they received SMS.

What kind of impact did pre-existing health issues have on appointment attendance?

There was no impact with a patient's appointment attendance if they had pre-existing health issues.

What does sending SMS reminders help with appointment attendance?

Contrary to initial belief, sending SMS to patients did not help lower the percentage of no shows occurring. Instead, there was an increased percentage (around 14%) of no shows among those that did receive a SMS.

Does the day of an appointment affect attendance?

The day of an appoint had very minimal effect to attendance. There was a 1.1% increase of a patient not attending an appointment if the appointment was made for Friday and a 0.6% increase of not showing up on Monday.

Sites used for reference:

<https://matplotlib.org/stable/api/index.html>

<https://pandas.pydata.org/pandas-docs/stable/reference/index.html>

<https://stackoverflow.com/questions/49199164/increasing-pie-chart-size-with-matplotlib-radius-parameter-appears-to-do-nothin>

<https://stackoverflow.com/questions/39698672/pandas-replace-boolean-value-with-string-or-integer>

<https://numpy.org/doc/stable/reference/index.html>

<https://www.doxygen.nl/manual/docblocks.html>