Analysis on Appointment No Show

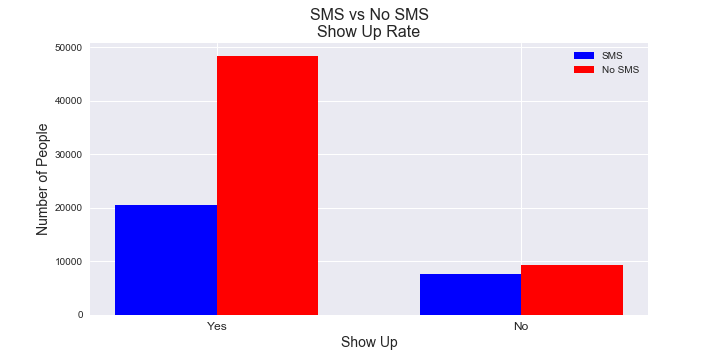
**What’s the Problem?**

There are many people who would make an appointment with the doctor’s office but never showed up at the appointed time. These individuals believe that if they miss this appointment, they can always reschedule a new one. They don’t realize that when they make an appointment with the doctors, they are making a commitment. If they do not show up to the appointment, it would be a waste of time for themselves, the doctors, and the other patients. When someone makes an appointment, it takes up a block of time which will prevent others from making an appointment at this time.

A **possible solution** to resolving this issue is to send out SMS messages to individuals who has made an appointment with the doctor’s office.

The **hypothesis** is that SMS messages will increase the show up rate of patients. A remainder for them to show up on appointed day will cause them to fulfill this commitment.

This **key metric** of analysis will focus on SMS messages. I want to calculate the proportion of show up after the doctor’s office send out SMS messages to remind patients of their appointment date. The five other variables that we will also be covering: age, scholarship, day difference, patients with conditions, and day of the week. **(Alpha value – 0.05)**

**SMS vs No SMS**

This is the results from the dataset:

* SMS Received sample size: 28248
* No SMS Received sample size: 57744
* Proportion of No Show (SMS): 0.27007221750212407 ~ 27.00%
* Proportion of No Show (No SMS): 0.16311651426988086 ~ 16.31%

|  |  |  |
| --- | --- | --- |
|  | No Show | Show |
| SMS | 7629 | 20619 |
| No SMS | 9419 | 48325 |

This bar graph above represents the Show Up Rate of patients who have received SMS messages and patients who did not receive SMS messages. The result is unexpected. It appears that patients who did not receive SMS messages are more likely to show up.

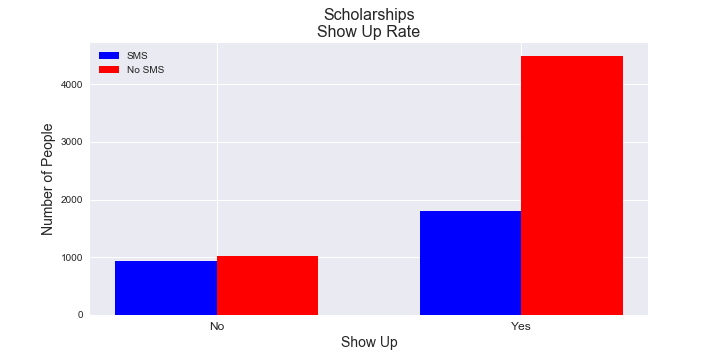
*T-test: Ttest\_indResult(statistic=37.244767666, pvalue=3.2893128454e-301)*

*Chi Value: 1364.50933234*

*P-Value: 1.08317178386e-298*

Using the table above, I performed a chi-square contingency test to test for the independence of the data. The t-test shows that SMS and No SMS are two vastly different samples of data. These values show that the data are dependent, meaning that SMS do affect the show up rate of appointments at the doctor’s office, but not a positive effect.

**Scholarship’s effect on Show Up Rate**

This variable is mainly to determine if people with scholarships are more likely to show up to appointments, considering these patients are model students who are more expected to fulfill their commitment.

*T-test: Ttest\_indResult(statistic=15.830644992, pvalue=1.2553049064e-55)*

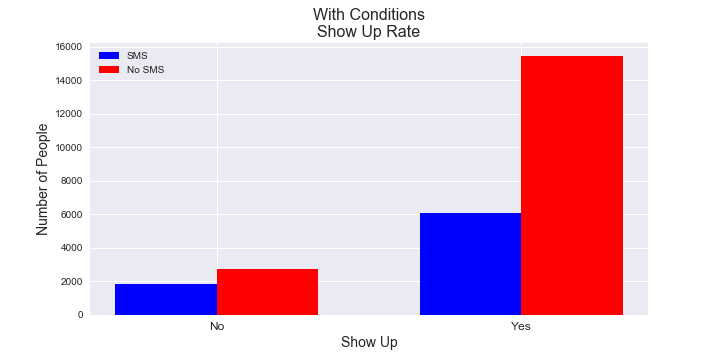
|  |  |  |
| --- | --- | --- |
|  | No Show | Show |
| SMS | 935 | 1800 |
| No SMS | 1028 | 4484 |

*Chi Value: 242.41981896*

*P-Value: 1.16704546016e-54*

The results are almost identical to SMS vs No SMS. The data are also dependent on each other. The graph also shows that patients are more likely to show up when they did not receive any SMS message.

**Patients’ with Conditions**

This category is concerning people who has some sort of illness or disabilities. I want to see if people who are ill will show up to their appointment after receiving a SMS message.

*T-test: Ttest\_indResult(statistic=16.7284468656, pvalue=1.720456767065e-62)*

|  |  |  |
| --- | --- | --- |
|  | No Show | Show |
| SMS | 1876 | 6080 |
| No SMS | 2736 | 15441 |

*Chi test stat: 276.31057299*

*P-Value: 4.78194911664e-62*

The result from running the t-test and chi-square test are also almost identical to the previous two variables. Patients who did not receive SMS message are still more likely to show up.

This contingency table will test for the independence of ‘no show’ rate to these three variables. From the looks of it, the three variables are still dependent of each other. Meaning that these variable in some ways do affect the ‘no show’ rate to the appointments for various reasons.

|  |  |  |  |
| --- | --- | --- | --- |
|  | With Condition | SMS | Scholarship |
| SMS | 1876 | 7629 | 935 |
| No SMS | 2736 | 9419 | 1028 |

*(Focus on No Show)*

*Chi test stat: 34.6784141286*

*P-Value: 2.94902226457e-08*

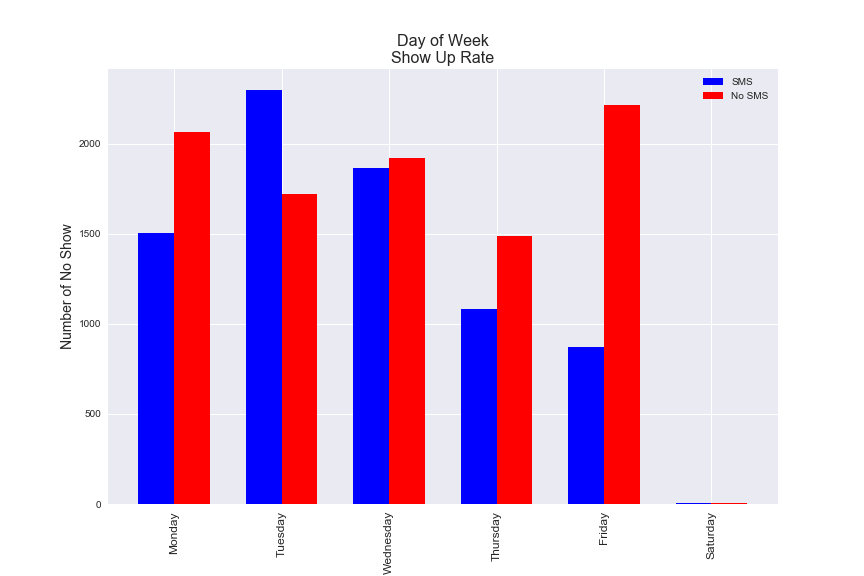
**Day of Week Effects on No Show**

Sometimes people might not be able to show up to appointments because of jobs or some other commitment which took priority. Taken that into consideration, I want to test for the independence of week days and no-show rate.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Mon | Tue | Wed | Thur | Fri | Sat |
| SMS | 1506 | 2300 | 1865 | 1081 | 874 | 3 |
| No SMS | 2065 | 1721 | 1919 | 1490 | 2218 | 6 |

*Chi test stat: 641.030789178*

*P-Value: 2.74853199958e-136*

The chi-square test shows that there is a relationship between days of the week and SMS messages on no-show rate. It appears that some days might cause people to not show up. In comparison to the other days, the appointments on Saturday are practically zero, we can exclude it since the data are insignificant. But from the bar-graph above, it seems that ‘no SMS message’ are more likely to show except for Tuesday and Wednesday.

**How Does Age Affect No-Show?**

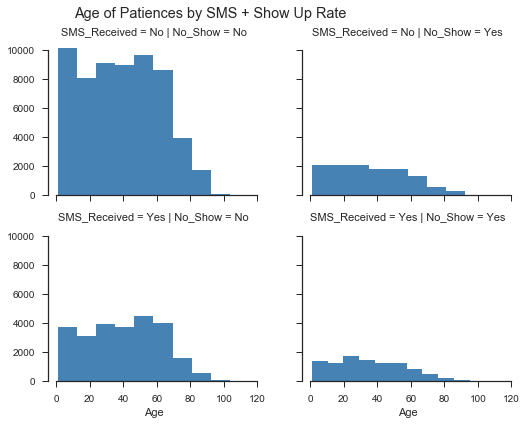
The four histograms show the show up rate according to age of patients. You can still see that no SMS received are still more likely to show up, however there isn’t an obvious relationship between Age and No-Show. Hence, I want to categorize the age into four groups for contingency independence test. (Focus on No-Show)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Kids | Young Adults | Old Adults | Seniors |
| SMS | 1523 | 3061 | 2182 | 863 |
| No SMS | 1693 | 3397 | 2970 | 1359 |

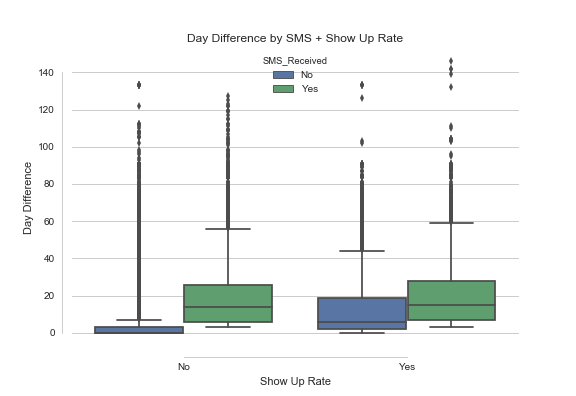
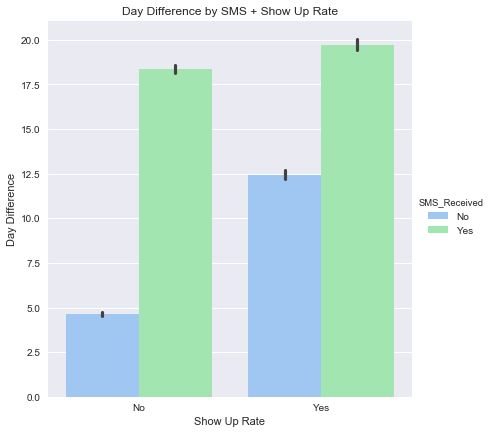
(kids: <= 24; young adults <= 44; old adults <=64; seniors 65+)

*Chi test stat: 70.5429089171 P-Value: 3.26587302995e-15*

The result displays that Age and SMS are indeed dependent on each other when it comes to No-Show Rates.



**Days of Difference between Scheduled and Appointment**

In theory, people tend to miss their appointment if they scheduled it months ahead of time because some issues might pop up during these times. People who make their appointment within a time interval of a week have a higher chance of showing up because they have an urgent problem. But is this true?

From the boxplot, most patients tend to make their appointment with a month of time. People who received SMS messages to remind them about the appointment also tend to make their appointments more in advance. However, the bar-graph shows people who do make appointments close to the appointment dates also do not show up as often as expected.

*Run t-test on: Day Difference SMS vs Day Difference No SMS*

T-test: Ttest\_indResult(statistic=-0.98096420032, pvalue=0.32788462585327183)

The outcome determines that we do not reject the assumption that the two samples have roughly the same average, which means they come from the same population.

**Conclusion**

The outcomes were unexpected and far from my hypothesis. Patients tend to not show up to appointments when they receive a SMS message to remind them about the appointment. I have covered the following variables: scholarships, conditions, day of week, age, and day differences. Although on average there are little to no difference on Show-Up Rate on SMS messages when it comes to day differences. The other four variables demonstrated that SMS messages and Show-Up Rates are dependent, meaning people do not show up because they received a SMS message. From this dataset, I can draw the conclusion SMS message is not the solution to resolving the issue of people not showing up to their appointments.

*Potential Bias and Future Improvement:*

*This is dataset is taken from one clinic in Brazil, there might be some other reasons to why patients do not show up to appointments after being sent a SMS message. Maybe some patients never received the message, or it was sent to a wrong number.*

*In the future, I will use SQL to join more data from other parts of the world to avoid bias.*