Appointment No-Shows - Inferential Statistics

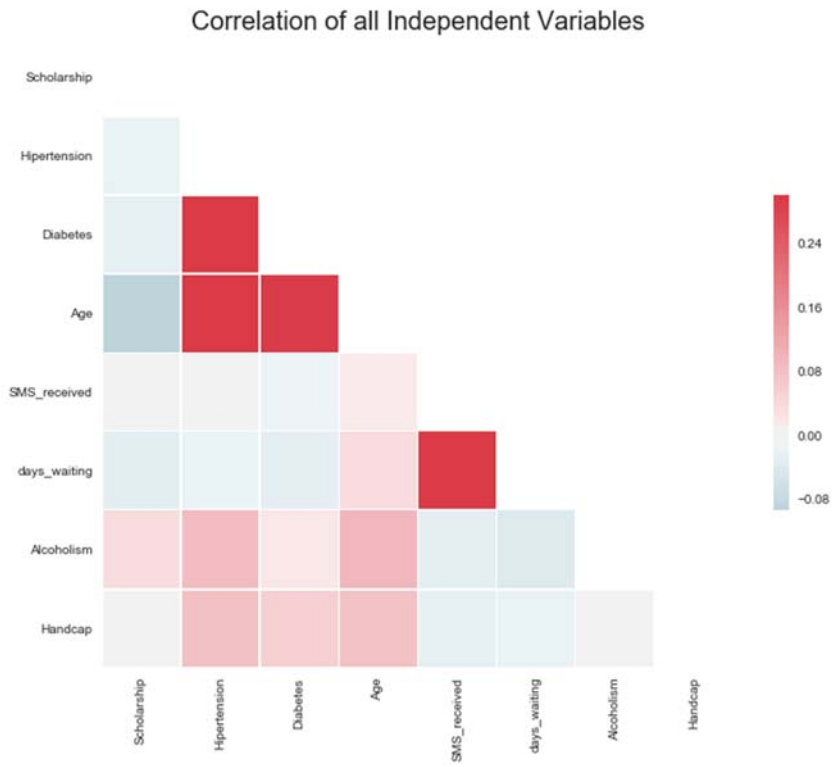
Dovid Burns

2/5/18

Springboard Capstone One

Looking at the data story telling indicated that the strongest indicators of correlation with missing an appointment seem to be appointment day, handicap, hypertension, diabetes, scholarship, received a text reminder, age, time between appointment scheduled to appointment day and neighborhood. Additionally, our initial survey indicated that there did not seem to be a correlation between missing an appointment and alcoholism and gender. To further statistically verify these claims we ran t-test for independence and chi-squared tests between these variables. To check for correlation between the independent variables above we created a correlation heat map of all the numerical variables.

Using a chi-squared contingency test built into python we found the following p-values for correlation with no-show rate: appointment day = 1.8e-5, handicap = 0.11, hypertension = 2.4e-32, diabetes = 4.9e-7, scholarship = 3.2e-22, received a text reminder =0.0, age binned = 6e-174, time between appointment scheduled to appointment day binned = 0.0 and neighborhood = 1.3e-60. The p-values that we expected to not show correlation was p-value for gender = 0.19 and alcoholism = 0.98. Holding these test to a 99% accuracy, we would only accept variables as being statistically significant in predicting no-shows with a p-value of less than 0.01. Therefore, appointment day, hypertension, diabetes, scholarship, received a text reminder, age binned, time between appointment scheduled to appointment day binned, and neighborhood are all good predictors of a patient not showing up for an appointment. The strongest indicators based on smallest p-value are scholarship, hypertension, received a text, age binned, time between appointment scheduled to appointment day binned, appointment day, and neighborhood.



Looking at the above heat map of correlation we see that there does seem to be some correlations between independent variables. We see a strong positive correlation between diabetes and hypertension. We also see a strong positive correlation between age and hypertension, age and diabetes and a slight negative correlation between age and scholarship. Lastly, we see a strong positive correlation between the time between scheduled and the appointment date with receiving a reminder text.