Course Project – Values

Shaun Pritchard

Rasmussen College

QMB4400

Tina Harclerode

April 21, 2021

For this course project I first important my python libraries and then imported my data set ***noshowappointments.csv***. After importing the data, I pre-process the data by checking the data frame to verify that they that integrity was good. Then I processed the Null data variables by checking the null values determining which columns had no values and listing those no values in the dataset with descriptive statistics. Then I targeted the *Null* values and use several methods to replace the Null values with NaN. I then proceeded to verify the data replacements took place and then created a new dataframe ***df1*** and ***df2*** targeting the *no-show* column setting all of its variables removing any Null values. I've then verified the output of that data and proceeded to Ignore NaN values from the calculation and determine PatientID, Count(No-Show=Yes) and Ignore NaN values from the calculation and determine Calculate Neighbourhood, count(No-Shows=Yes)..

I have included a full jupyter notebook file of my python script and output as well as provided a link below to the Google colab file with my work as follows.

<https://colab.research.google.com/drive/1GkqjOVahCfsn69bBD8w9yVlK9FlE_ohQ?usp=sharing>











