Course Project – Storing Output

Shaun Pritchard

Rasmussen College

QMB4400

Tina Harclerode

May 2, 2021

**Course Project – Storing Output**

This project I loaded my python libraries, bloated the no-show CSV file into a pandas dataframe, and then check the validity of the data. After this, I created a new data frame *df1* to organize the no-show data. I then use the shape method to get the total number of patients with no shows equal to yes. I then created another variable call *tb1* for the ***PatientId*** and *tb2* for ***Neighbourhood*** value counts. I then used the head method to validate the data for both tables. I then proceeded to create a new connection database called ***apt.db***. I then created 2-new tables in sqlite database for ***PatientId*** and ***Neighbourhood*** value counts of *yes*. I then validated the data in each of these new SQL tables. I then use the ***concat*** method to join the tables for easy visualization followed by testing the output with the shape and head method followed by outputting the Joint tables to a new CSV file called ***NoShowOutput\_mod4.csv***.

I have included all project files as required for this course project as well as a working URL to the Google code cloud script.

<https://colab.research.google.com/drive/1zq_7SEbJHDa8sp8odvecFzDDC6ktuVnw?usp=sharing>







