Assignment 2

Date 11/04/2023 AM624D Submission by: 18/04/2023

****

Submission Guidelines:

1. Create a python notebook file (.ipynb) with name **AssnNo_Name_RollNo**

   Eg: Assn2_Rahul_Kottah_12-12-01

2. Create a folder on your Github profile with your Name_RollNo. Upload the python notebook file in this folder.

3. No Handwritten assignments. Only ipynb notebooks uploaded in your github profile to be submitted.

4. Share the link of your Github uploaded notebook/folder in Google classroom assignment reply.

****

1 Read the housingdata.csv file into pandas DataFrame and display first six rows of the DataFrame

2 Display the column names and row index in separate cells.

3 a) How many numbers of "STATE" exist in the dataset. Hint: use shape
b) How many unique states exist in the dataset

4 Retrieve the list of all NaN/Null/Empty cells in the form of Boolean list

5 Drop all rows with N/A,NA,na values in Num_Bedrooms

6 Replace the NaN and String entries in Num_Bath with previous entry

7. Replace the empty values in "SQ_FT" with the mean of the all entries.

8. What is the sum of SQ_FT of all entries in the dataset.

9. Add a new column "NUM_STORE" and with all the values as 1 for all rows

10. Create a DataFrame containing all entries of TN state only

14. Create a DataFrame having SQ_FT area greater than 1000 and display the DataFrame

15. Create a DataFrame having only first Three columns and First Three rows

16. Display the state having highest average price per square foot of housing area. Hint: Average of (TotalPrice/Sq_Ft)