When we printed the first five rows of the SAT data, the output looked like this:



DBN                                    SCHOOL NAME  \

0  01M292  HENRY STREET SCHOOL FOR INTERNATIONAL STUDIES

1  01M448            UNIVERSITY NEIGHBORHOOD HIGH SCHOOL

2  01M450                     EAST SIDE COMMUNITY SCHOOL

3  01M458                      FORSYTH SATELLITE ACADEMY

4  01M509                        MARTA VALLE HIGH SCHOOL

​

 Num of SAT Test Takers SAT Critical Reading Avg. Score SAT Math Avg. Score  \

0                     29                             355                 404

1                     91                             383                 423

2                     70                             377                 402

3                      7                             414                 401

4                     44                             390                 433

​

 SAT Writing Avg. Score

0                    363

1                    366

2                    370

3                    359

4                    384

We can make a few observations based on this output:

* The DBN appears to be a unique ID for each school.
* We can tell from the first few rows of names that we only have data about high schools.
* There's only a single row for each high school, so each DBN is unique in the SAT data.
* We may eventually want to combine the three columns that contain SAT scores -- SAT Critical Reading Avg., Score SAT Math Avg. Score, and SAT Writing Avg. Score -- into a single column to make the scores easier to analyze.

Given these observations, let's explore the other data sets to see if we can gain any insight into how to combine them.

Instructions

* Loop through each key in data. For each key:
  + Display the first five rows of the dataframe associated with the key.

Answer:

# for key in data

for k in data:

# print only 5 lines for all keys

print (data[k].head())