pandas-challenge-\_ScreenShotDSBC\_29APR2023\_\_Raj

Week 4th - as of 4/29/2023

Module 4 Challenge

=============================================

Student – Raj Agrawal / DS bootcamp

===========================================

terminal output screen shot

===

**# based on below details, here are few observations -**

#1. \*\*\*\*\*\*\*\*\*\*\*\*District Summary

# Total number of unique schools 15

# Total students 39,170

# Total budget 24,649,428

# Average math score 78.98

# Average reading score 81.87

# % passing math (the percentage of students who passed math) 74.98

# % passing reading (the percentage of students who passed reading) 85.80

# % overall passing (the percentage of students who passed math AND reading) 65.17

# 2. maximum per student budget is = 83.98 Wilson High School

# 3. highest count of students -- Bailey High School 4976

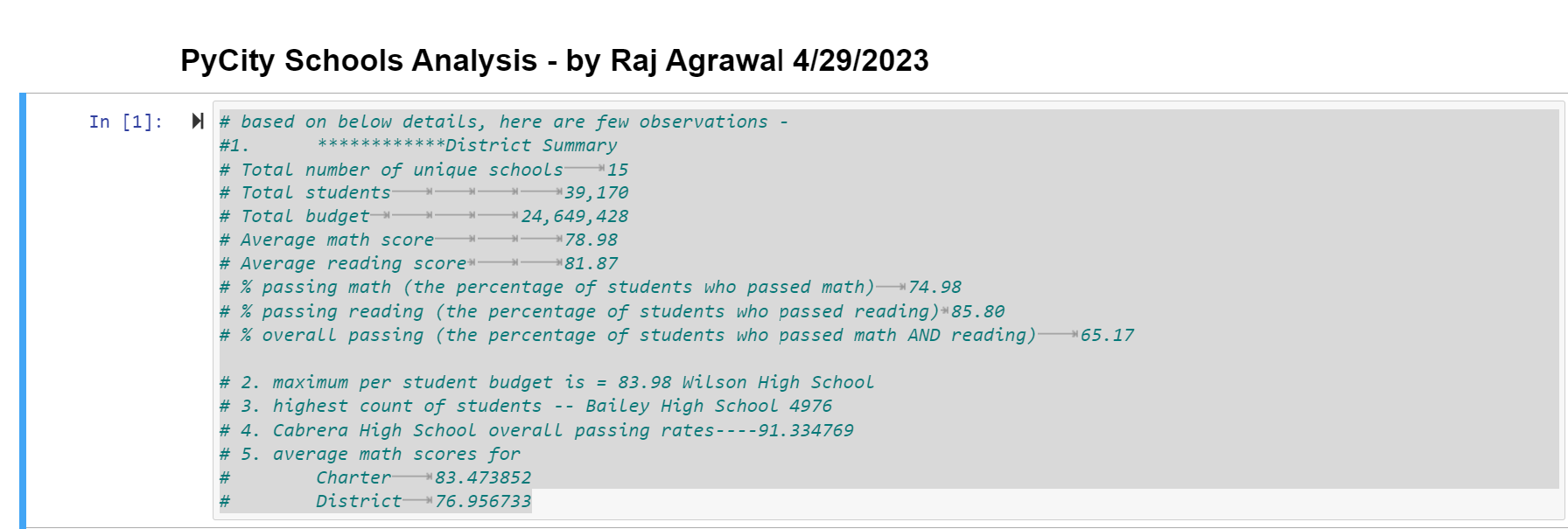
# 4. Cabrera High School overall passing rates----91.334769

# 5. average math scores for

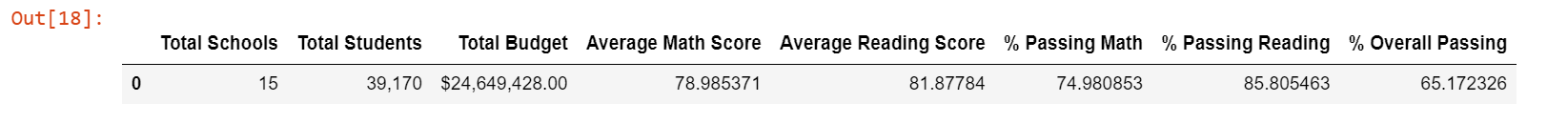
# Charter 83.473852

# District 76.956733

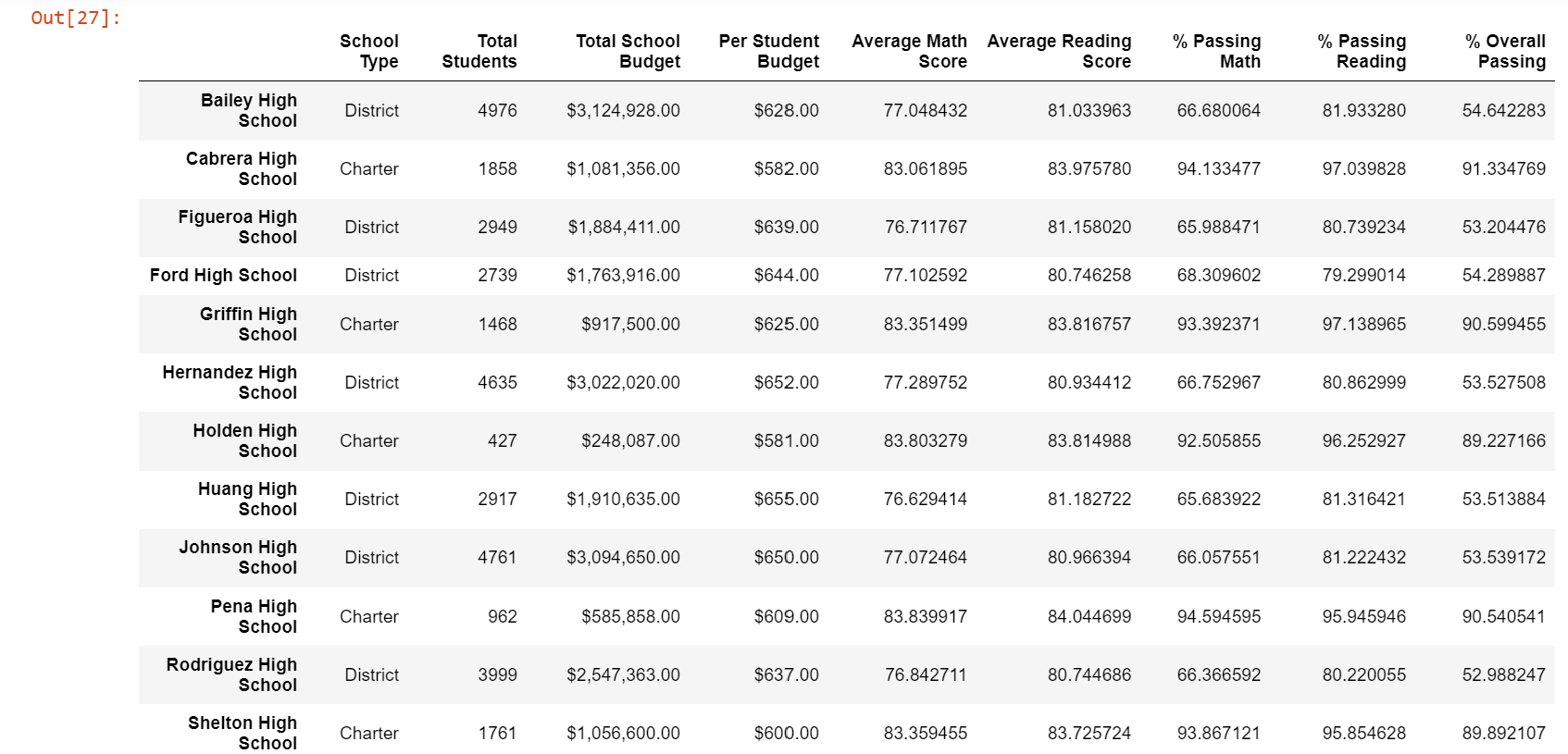
=======



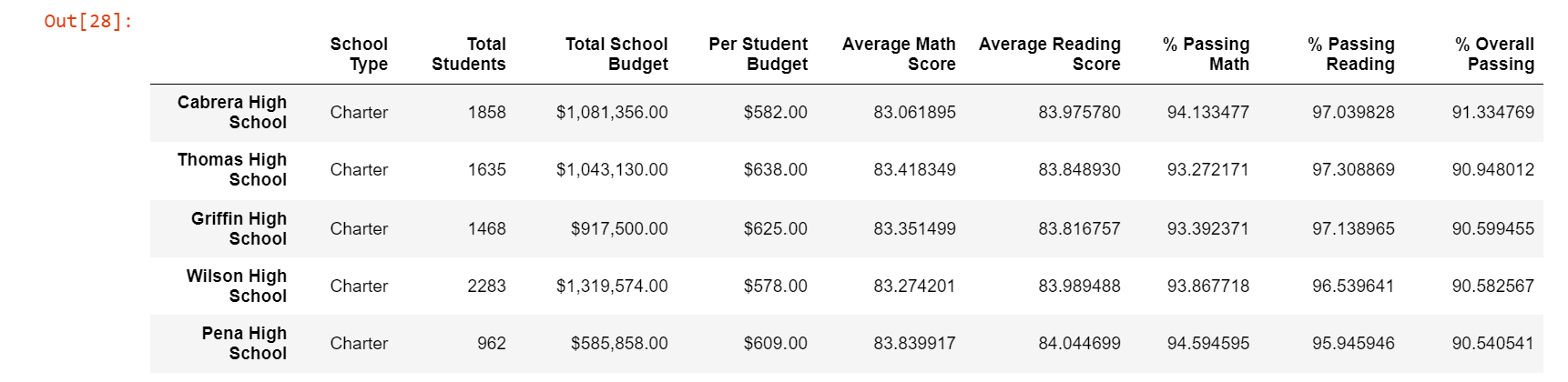
**District Summary**

****

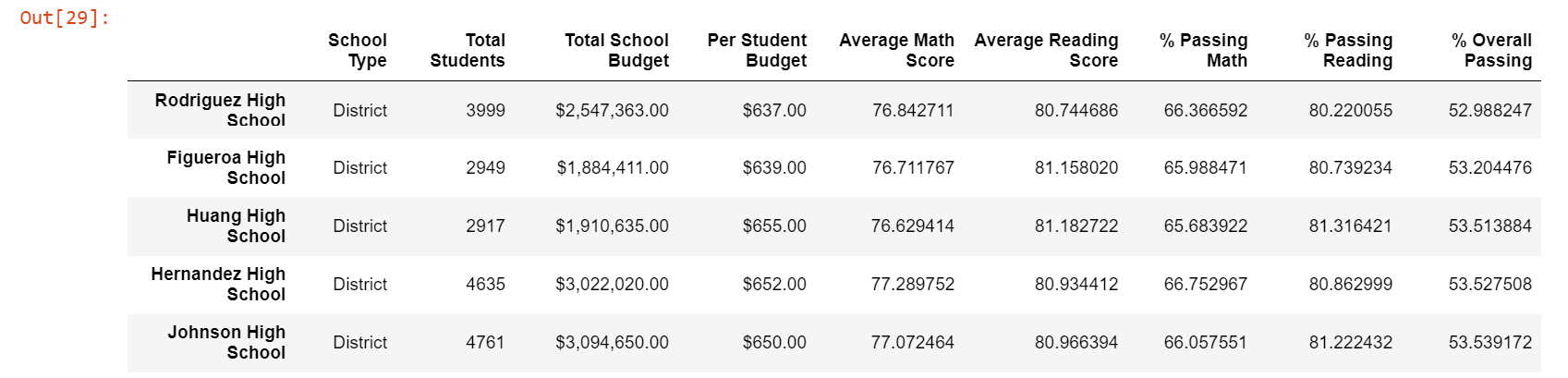
**School Summary**

****

**Highest-Performing Schools (by % Overall Passing)**

****

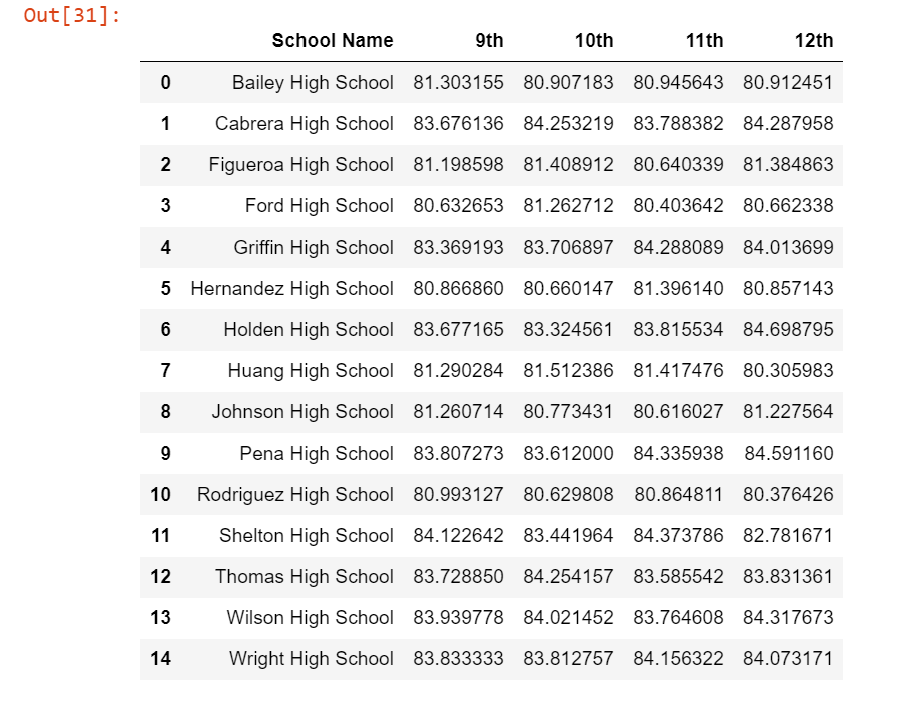
**Lowest-Performing Schools (by % Overall Passing)**

****

**Math Scores by Grade**

****

**Reading Scores by Grade**

****

**Scores by School Spending**

spending\_bins = [0, 585, 630, 645, 680]

labels = ["<$585", "$585-630", "$630-645", "$645-680"]

Use pd.cut to categorize spending based on the bins.

Use the following code to then calculate mean scores per spending range.

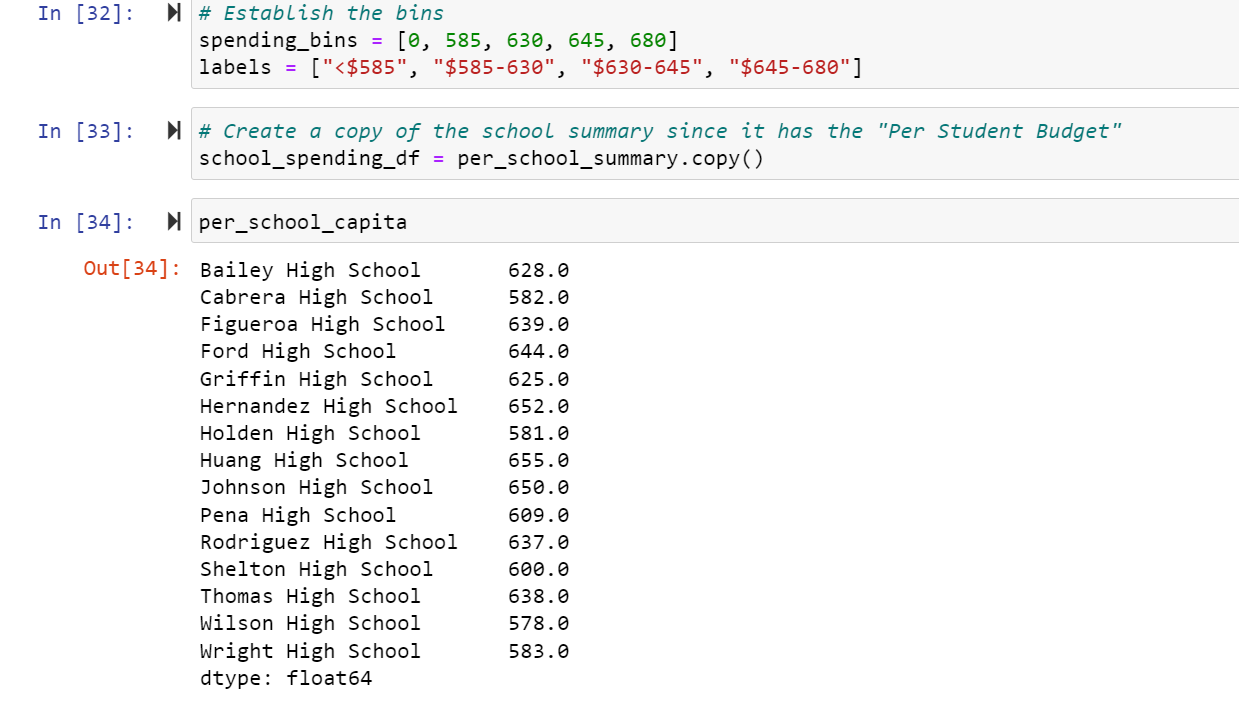
spending\_math\_scores = school\_spending\_df.groupby(["Spending Ranges (Per Student)"])["Average Math Score"].mean()

spending\_reading\_scores = school\_spending\_df.groupby(["Spending Ranges (Per Student)"])["Average Reading Score"].mean()

spending\_passing\_math = school\_spending\_df.groupby(["Spending Ranges (Per Student)"])["% Passing Math"].mean()

spending\_passing\_reading = school\_spending\_df.groupby(["Spending Ranges (Per Student)"])["% Passing Reading"].mean()

overall\_passing\_spending = school\_spending\_df.groupby(["Spending Ranges (Per Student)"])["% Overall Passing"].mean()

****

**Scores by School Size**

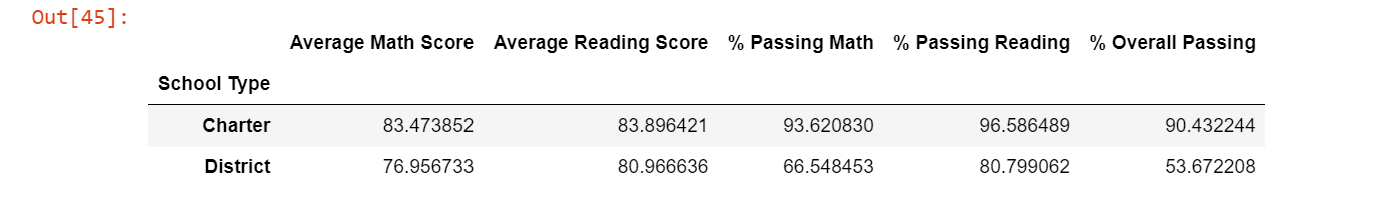
Use the following code to bin the per\_school\_summary.

size\_bins = [0, 1000, 2000, 5000]

labels = ["Small (<1000)", "Medium (1000-2000)", "Large (2000-5000)"]

****

**Scores by School Type**

****