

# Analysing City School Data using Pandas and Jupyter Notebooks

## Given Background

You are the new Chief Data Scientist for your local government area. In this capacity, you'll be helping the school board and mayor make strategic decisions regarding future school budgets and priorities. As a first task, you've been asked to analyse the area-wide standardised test results. You'll be given access to every student's maths and reading scores, as well as various information on the schools they attend. Your task is to aggregate the data to showcase obvious trends in school performance.

## Objectives

- A. To create a [summary report](#) that includes the data given below(1-9), prepared using Pandas and Jupyter Notebook.
- B. To analyse the data to showcase obvious trends in school performance and add the [conclusions](#) to the report.

## Required Output Data includes:

1. Local Government Area (LGA) Summary
2. School Summary
3. Highest-Performing Schools (by % Overall Passing)
4. Lowest-Performing Schools (by % Overall Passing)
5. Maths Scores by Year
6. Reading Scores by Year
7. Scores by School Spending
8. Scores by School Size
9. Scores by School Type

### A. SUMMARY REPORT

The objective is to analyse the trends in school performance based on various information derived from the given dataset. The dataset provides data including the details of students in different schools under the local self-government, which includes their student ID, some personal information including name and gender, math scores, reading scores, year/grade budget per student etc. The analysis was done to compare average math and reading scores, average math and reading pass per cent and overall pass per cent of students in different schools. The comparison was also made based on the type of schools (independent/ government), the size of the school and the average budget that was allocated per student.

#### 1. Local Government Area (LGA) Summary

To create a high-level snapshot of the local government area's critical metrics in a DataFrame including the following:

- Total number of unique schools
- Total Students
- Total budget
- Average maths score
- Average reading score

- % passing maths (the percentage of students who passed maths)
- % passing reading (the percentage of students who passed reading)
- % overall passing (the percentage of students who passed maths AND reading)

	Total no:of Schools	Total no:of Students	Total Budget	Average Math Score	Average Reading Score	% Passing Maths	% Passing Reading	% Overall Passing
0	15	39,170	\$24,649,428.00	70.338192	69.980138	86.078632	84.426857	72.808272

## 2. School Summary

To create a DataFrame that summarises key metrics about each school including the following:

- School name
- School type
- Total Students
- Total school budget
- Per student budget
- Average maths score
- Average reading score
- % passing maths (the percentage of students who passed maths)
- % passing reading (the percentage of students who passed reading)
- % overall passing (the percentage of students who passed maths AND reading)

	School Type	Total Students	Total School Budget	Per Student Budget	Average Math Score	Average Reading Score	% Passing Maths	% Passing Reading	% Overall Passing
school_name									
Bailey High School	Government	4976	3124928	628.0	72.352894	71.008842	91.639871	87.379421	80.084405
Cabrera High School	Independent	1858	1081356	582.0	71.657158	71.359526	90.850377	89.074273	80.785791
Figueroa High School	Government	2949	1884411	639.0	68.698542	69.077993	81.654798	82.807731	67.650051
Ford High School	Government	2739	1763916	644.0	69.091274	69.572472	82.438846	82.219788	67.469880
Griffin High School	Independent	1468	917500	625.0	71.788147	71.245232	91.212534	88.487738	81.335150
Hernandez High School	Government	4635	3022020	652.0	68.874865	69.186408	80.949299	81.877023	66.364617
Holden High School	Independent	427	248087	581.0	72.583138	71.660422	89.929742	88.524590	78.922717
Huang High School	Government	2917	1910635	655.0	68.935207	68.910525	81.693521	81.453548	66.712376
Johnson High School	Government	4761	3094650	650.0	68.843100	69.039277	82.062592	81.978576	67.191766
Pena High School	Independent	962	585858	609.0	72.088358	71.613306	91.683992	86.590437	79.209979
Rodriguez High School	Government	3999	2547363	637.0	72.047762	70.935984	90.797699	87.396849	79.419855
Shelton High School	Independent	1761	1056600	600.0	72.034072	70.257808	91.538898	86.712095	78.875639
Thomas High School	Independent	1635	1043130	638.0	69.581651	69.768807	83.853211	82.629969	69.480122
Wilson High School	Independent	2283	1319574	578.0	69.170828	68.876916	82.785808	81.296540	67.455103
Wright High School	Independent	1800	1049400	583.0	72.047222	70.969444	91.777778	86.666667	79.722222

Formatted and cleaned:

	School Type	Total Students	Total School Budget	Per Student Budget	Average Math Score	Average Reading Score	% Passing Maths	% Passing Reading	% Overall Passing
school_name									
Bailey High School	Government	4976	\$3,124,928.00	\$628.00	72.352894	71.008842	91.639871	87.379421	80.084405
Cabrera High School	Independent	1858	\$1,081,356.00	\$582.00	71.657158	71.359526	90.850377	89.074273	80.785791
Figueroa High School	Government	2949	\$1,884,411.00	\$639.00	68.698542	69.077993	81.654798	82.807731	67.650051
Ford High School	Government	2739	\$1,763,916.00	\$644.00	69.091274	69.572472	82.438846	82.219788	67.469880
Griffin High School	Independent	1468	\$917,500.00	\$625.00	71.788147	71.245232	91.212534	88.487738	81.335150
Hernandez High School	Government	4635	\$3,022,020.00	\$652.00	68.874865	69.186408	80.949299	81.877023	66.364617
Holden High School	Independent	427	\$248,087.00	\$581.00	72.583138	71.660422	89.929742	88.524590	78.922717
Huang High School	Government	2917	\$1,910,635.00	\$655.00	68.935207	68.910525	81.693521	81.453548	66.712376
Johnson High School	Government	4761	\$3,094,650.00	\$650.00	68.843100	69.039277	82.062592	81.978576	67.191766
Pena High School	Independent	962	\$585,858.00	\$609.00	72.088358	71.613306	91.683992	86.590437	79.209979
Rodriguez High School	Government	3999	\$2,547,363.00	\$637.00	72.047762	70.935984	90.797699	87.396849	79.419855
Shelton High School	Independent	1761	\$1,056,600.00	\$600.00	72.034072	70.257808	91.538898	86.712095	78.875639
Thomas High School	Independent	1635	\$1,043,130.00	\$638.00	69.581651	69.768807	83.853211	82.629969	69.480122
Wilson High School	Independent	2283	\$1,319,574.00	\$578.00	69.170828	68.876916	82.785808	81.296540	67.455103
Wright High School	Independent	1800	\$1,049,400.00	\$583.00	72.047222	70.969444	91.777778	86.666667	79.722222

### 3. Highest-Performing Schools (by % Overall Passing)

To sort the schools by % Overall Passing in descending order and display the top 5 rows and save the results in a DataFrame called "top\_schools".

	School Type	Total Students	Total School Budget	Per Student Budget	Average Math Score	Average Reading Score	% Passing Maths	% Passing Reading	% Overall Passing
school_name									
Griffin High School	Independent	1468	\$917,500.00	\$625.00	71.788147	71.245232	91.212534	88.487738	81.335150
Cabrera High School	Independent	1858	\$1,081,356.00	\$582.00	71.657158	71.359526	90.850377	89.074273	80.785791
Bailey High School	Government	4976	\$3,124,928.00	\$628.00	72.352894	71.008842	91.639871	87.379421	80.084405
Wright High School	Independent	1800	\$1,049,400.00	\$583.00	72.047222	70.969444	91.777778	86.666667	79.722222
Rodriguez High School	Government	3999	\$2,547,363.00	\$637.00	72.047762	70.935984	90.797699	87.396849	79.419855

### 4. Lowest-Performing Schools (by % Overall Passing)

To sort the schools by % Overall Passing in ascending order and display the top 5 rows and save the results in a DataFrame called "bottom\_schools".

	School Type	Total Students	Total School Budget	Per Student Budget	Average Math Score	Average Reading Score	% Passing Maths	% Passing Reading	% Overall Passing
school_name									
Hernandez High School	Government	4635	\$3,022,020.00	\$652.00	68.874865	69.186408	80.949299	81.877023	66.364617
Huang High School	Government	2917	\$1,910,635.00	\$655.00	68.935207	68.910525	81.693521	81.453548	66.712376
Johnson High School	Government	4761	\$3,094,650.00	\$650.00	68.843100	69.039277	82.062592	81.978576	67.191766
Wilson High School	Independent	2283	\$1,319,574.00	\$578.00	69.170828	68.876916	82.785808	81.296540	67.455103
Ford High School	Government	2739	\$1,763,916.00	\$644.00	69.091274	69.572472	82.438846	82.219788	67.469880

### 5. Maths Scores by Year

To create a dataframe that lists the average maths score for students of each year level (9, 10, 11, 12) at each school.

	Year 9	Year 10	Year 11	Year 12
school_name				
Bailey High School	72.493827	71.897498	72.374900	72.675097
Cabrera High School	72.321970	72.437768	71.008299	70.604712
Figueroa High School	68.477804	68.331586	68.811001	69.325282
Ford High School	69.021609	69.387006	69.248862	68.617811
Griffin High School	72.789731	71.093596	71.692521	71.469178
Hernandez High School	68.586831	68.867156	69.154412	68.985075
Holden High School	70.543307	75.105263	71.640777	73.409639
Huang High School	69.081754	68.533246	69.431345	68.639316
Johnson High School	69.469286	67.990220	68.637730	69.287393
Pena High School	71.996364	72.396000	72.523438	71.187845
Rodriguez High School	71.940722	71.779808	72.364811	72.154626
Shelton High School	72.932075	72.506696	70.097087	72.331536
Thomas High School	69.234273	70.057007	69.657831	69.369822
Wilson High School	69.212361	69.455446	68.378965	69.787472
Wright High School	71.741176	72.179012	73.275862	70.848238

Formatted and cleaned:

	Year 9	Year 10	Year 11	Year 12
school_name				
Bailey High School	72.494	71.897	72.375	72.675
Cabrera High School	72.322	72.438	71.008	70.605
Figueroa High School	68.478	68.332	68.811	69.325
Ford High School	69.022	69.387	69.249	68.618
Griffin High School	72.790	71.094	71.693	71.469
Hernandez High School	68.587	68.867	69.154	68.985
Holden High School	70.543	75.105	71.641	73.410
Huang High School	69.082	68.533	69.431	68.639
Johnson High School	69.469	67.990	68.638	69.287
Pena High School	71.996	72.396	72.523	71.188
Rodriguez High School	71.941	71.780	72.365	72.155
Shelton High School	72.932	72.507	70.097	72.332
Thomas High School	69.234	70.057	69.658	69.370
Wilson High School	69.212	69.455	68.379	69.787
Wright High School	71.741	72.179	73.276	70.848

## 6. Reading Scores by Year

To create a DataFrame that lists the average reading score for students of each year level (9, 10, 11, 12) at each school.

	Year 9	Year 10	Year 11	Year 12
school_name				
Bailey High School	70.901920	70.848265	70.317346	72.195525
Cabrera High School	71.172348	71.328326	71.201245	71.856021
Figueroa High School	70.261682	67.677588	69.152327	69.082126
Ford High School	69.615846	68.988701	70.735964	68.849722
Griffin High School	72.026895	70.746305	72.385042	69.434932
Hernandez High School	68.477569	70.621842	68.418199	69.244136
Holden High School	71.598425	71.096491	73.310680	70.481928
Huang High School	68.670616	69.516297	68.740638	68.671795
Johnson High School	68.719286	69.295029	69.969115	67.992521
Pena High School	70.949091	72.324000	71.703125	71.513812
Rodriguez High School	70.902921	70.137500	71.424453	71.414449
Shelton High School	70.715094	69.879464	71.150485	69.070081
Thomas High School	69.672451	69.741093	70.749398	68.730769
Wilson High School	68.683043	68.412541	68.796327	69.888143
Wright High School	71.823529	71.386831	70.475862	69.821138

Formatted and cleaned:

	Year 9	Year 10	Year 11	Year 12
school_name				
Bailey High School	70.902	70.848	70.317	72.196
Cabrera High School	71.172	71.328	71.201	71.856
Figueroa High School	70.262	67.678	69.152	69.082
Ford High School	69.616	68.989	70.736	68.850
Griffin High School	72.027	70.746	72.385	69.435
Hernandez High School	68.478	70.622	68.418	69.244
Holden High School	71.598	71.096	73.311	70.482
Huang High School	68.671	69.516	68.741	68.672
Johnson High School	68.719	69.295	69.969	67.993
Pena High School	70.949	72.324	71.703	71.514
Rodriguez High School	70.903	70.138	71.424	71.414
Shelton High School	70.715	69.879	71.150	69.070
Thomas High School	69.672	69.741	70.749	68.731
Wilson High School	68.683	68.413	68.796	69.888
Wright High School	71.824	71.387	70.476	69.821

## 7. Scores by School Spending

To create a table that breaks down school performance based on average spending ranges (per student) and includes the following metrics:

- Average maths score
- Average reading score
- % passing maths (the percentage of students who passed maths)
- % passing reading (the percentage of students who passed reading)
- % overall passing (the percentage of students who passed maths AND reading)

	Average Math Score	Average Reading Score	% Passing Math	% Passing Reading	% Overall Passing
Spending Ranges (Per Student)					
<\$585	71.364587	70.716577	88.835926	86.390517	76.721458
\$585-630	72.065868	71.031297	91.518824	87.292423	79.876293
\$630-645	69.854807	69.838814	84.686139	83.763585	71.004977
\$645-680	68.884391	69.045403	81.568470	81.769716	66.756253

Formatted and cleaned:

	Average Math Score	Average Reading Score	% Passing Math	% Passing Reading	% Overall Passing
Spending Ranges (Per Student)					
<\$585	71.36	70.72	88.84	86.39	76.72
\$585-630	72.07	71.03	91.52	87.29	79.88
\$630-645	69.85	69.84	84.69	83.76	71.00
\$645-680	68.88	69.05	81.57	81.77	66.76

## 8. Scores by School Size

To create a DataFrame called `size_summary` that breaks down school performance based on school size (small, medium, or large) and Includes the following metrics:

- Average maths score
- Average reading score
- % passing maths (the percentage of students who passed maths)
- % passing reading (the percentage of students who passed reading)
- % overall passing (the percentage of students who passed maths AND reading)

	Average Math Score	Average Reading Score	% Passing Math	% Passing Reading	% Overall Passing
School Size					
Small (<1000)	72.335748	71.636864	90.806867	87.557513	79.066348
Medium (1000-2000)	71.421650	70.720164	89.846560	86.714149	78.039785
Large (2000-5000)	69.751809	69.576052	84.252804	83.301185	70.293507

Formatted and cleaned:

	Average Math Score	Average Reading Score	% Passing Math	% Passing Reading	% Overall Passing
School Size					
Small (<1000)	72.34	71.64	90.81	87.56	79.07
Medium (1000-2000)	71.42	70.72	89.85	86.71	78.04
Large (2000-5000)	69.75	69.58	84.25	83.30	70.29

## 9. Scores by School Type

To create a new DataFrame called `type_summary` that shows school performance based on the "School Type" and Includes the following metrics:

- Average maths score
- Average reading score
- % passing maths (the percentage of students who passed maths)
- % passing reading (the percentage of students who passed reading)
- % overall passing (the percentage of students who passed maths AND reading)

	Average Math Score	Average Reading Score	% Passing Math	% Passing Reading	% Overall Passing
School Type					
Government	69.834806	69.675929	84.462375	83.587562	70.698993
Independent	71.368822	70.718933	89.204043	86.247789	76.973340

Formatted and cleaned:

	Average Math Score	Average Reading Score	% Passing Math	% Passing Reading	% Overall Passing
School Type					
Government	69.83	69.68	84.46	83.59	70.70
Independent	71.37	70.72	89.20	86.25	76.97



## B. CONCLUSIONS

1. The number of students passing either maths or science is high (80+) compared to the students passing both (70+) when looking at the data for the whole local self-government jurisdiction area.

	Total no:of Schools	Total no:of Students	Total Budget	Average Math Score	Average Reading Score	% Passing Maths	% Passing Reading	% Overall Passing
0	15	39,170	\$24,649,428.00	70.338192	69.980138	86.078632	84.426857	72.808272

2. A similar trend is observed while analysing school-wise data. Looking at the math and reading pass per cent and overall pass per cent in different schools, The number of students passing either maths or science is high (80+) compared to the students passing both (70+) irrespective of them being small, large, independent or government-run.

	School Type	Total Students	Total School Budget	Per Student Budget	Average Math Score	Average Reading Score	% Passing Maths	% Passing Reading	% Overall Passing
school_name									
Bailey High School	Government	4976	\$3,124,928.00	\$628.00	72.352894	71.008842	91.639871	87.379421	80.084405
Cabrera High School	Independent	1858	\$1,081,356.00	\$582.00	71.657158	71.359526	90.850377	89.074273	80.785791
Figueroa High School	Government	2949	\$1,884,411.00	\$639.00	68.698542	69.077993	81.654798	82.807731	67.650051
Ford High School	Government	2739	\$1,763,916.00	\$644.00	69.091274	69.572472	82.438846	82.219788	67.469880
Griffin High School	Independent	1468	\$917,500.00	\$625.00	71.788147	71.245232	91.212534	88.487738	81.335150
Hernandez High School	Government	4635	\$3,022,020.00	\$652.00	68.874865	69.186408	80.949299	81.877023	66.364617
Holden High School	Independent	427	\$248,087.00	\$581.00	72.583138	71.660422	89.929742	88.524590	78.922717
Huang High School	Government	2917	\$1,910,635.00	\$655.00	68.935207	68.910525	81.693521	81.453548	66.712376
Johnson High School	Government	4761	\$3,094,650.00	\$650.00	68.843100	69.039277	82.062592	81.978576	67.191766
Pena High School	Independent	962	\$585,858.00	\$609.00	72.088358	71.613306	91.683992	86.590437	79.209979
Rodriguez High School	Government	3999	\$2,547,363.00	\$637.00	72.047762	70.935984	90.797699	87.396849	79.419855
Shelton High School	Independent	1761	\$1,056,600.00	\$600.00	72.034072	70.257808	91.538898	86.712095	78.875639
Thomas High School	Independent	1635	\$1,043,130.00	\$638.00	69.581651	69.768807	83.853211	82.629969	69.480122
Wilson High School	Independent	2283	\$1,319,574.00	\$578.00	69.170828	68.876916	82.785808	81.296540	67.455103
Wright High School	Independent	1800	\$1,049,400.00	\$583.00	72.047222	70.969444	91.777778	86.666667	79.722222

3. There are more independent schools in the top 5 (best-performing schools) list than government schools and more government schools in the bottom 5 (worst-performing schools) when analysed based on overall pass per cent. And also government schools in both the top 5 and bottom 5 categories receive more funding than their independent counterparts.

### Top Performing Schools (By % Overall Passing)

	School Type	Total Students	Total School Budget	Per Student Budget	Average Math Score	Average Reading Score	% Passing Maths	% Passing Reading	% Overall Passing
school_name									
Griffin High School	Independent	1468	\$917,500.00	\$625.00	71.788147	71.245232	91.212534	88.487738	81.335150
Cabrera High School	Independent	1858	\$1,081,356.00	\$582.00	71.657158	71.359526	90.850377	89.074273	80.785791
Bailey High School	Government	4976	\$3,124,928.00	\$628.00	72.352894	71.008842	91.639871	87.379421	80.084405
Wright High School	Independent	1800	\$1,049,400.00	\$583.00	72.047222	70.969444	91.777778	86.666667	79.722222
Rodriguez High School	Government	3999	\$2,547,363.00	\$637.00	72.047762	70.935984	90.797699	87.396849	79.419855

### Bottom Performing Schools (By % Overall Passing)

	School Type	Total Students	Total School Budget	Per Student Budget	Average Math Score	Average Reading Score	% Passing Maths	% Passing Reading	% Overall Passing
school_name									
Hernandez High School	Government	4635	\$3,022,020.00	\$652.00	68.874865	69.186408	80.949299	81.877023	66.364617
Huang High School	Government	2917	\$1,910,635.00	\$655.00	68.935207	68.910525	81.693521	81.453548	66.712376
Johnson High School	Government	4761	\$3,094,650.00	\$650.00	68.843100	69.039277	82.062592	81.978576	67.191766
Wilson High School	Independent	2283	\$1,319,574.00	\$578.00	69.170828	68.876916	82.785808	81.296540	67.455103
Ford High School	Government	2739	\$1,763,916.00	\$644.00	69.091274	69.572472	82.438846	82.219788	67.469880

4. Reading and maths scores are almost similar among different class years.

Reading scores:

	Year 9	Year 10	Year 11	Year 12
school_name				
Bailey High School	70.902	70.848	70.317	72.196
Cabrera High School	71.172	71.328	71.201	71.856
Figueroa High School	70.262	67.678	69.152	69.082
Ford High School	69.616	68.989	70.736	68.850
Griffin High School	72.027	70.746	72.385	69.435
Hernandez High School	68.478	70.622	68.418	69.244
Holden High School	71.598	71.096	73.311	70.482
Huang High School	68.671	69.516	68.741	68.672
Johnson High School	68.719	69.295	69.969	67.993
Pena High School	70.949	72.324	71.703	71.514
Rodriguez High School	70.903	70.138	71.424	71.414
Shelton High School	70.715	69.879	71.150	69.070
Thomas High School	69.672	69.741	70.749	68.731
Wilson High School	68.683	68.413	68.796	69.888
Wright High School	71.824	71.387	70.476	69.821

Maths scores:

	Year 9	Year 10	Year 11	Year 12
school_name				
Bailey High School	72.494	71.897	72.375	72.675
Cabrera High School	72.322	72.438	71.008	70.605
Figueroa High School	68.478	68.332	68.811	69.325
Ford High School	69.022	69.387	69.249	68.618
Griffin High School	72.790	71.094	71.693	71.469
Hernandez High School	68.587	68.867	69.154	68.985
Holden High School	70.543	75.105	71.641	73.410
Huang High School	69.082	68.533	69.431	68.639
Johnson High School	69.469	67.990	68.638	69.287
Pena High School	71.996	72.396	72.523	71.188
Rodriguez High School	71.941	71.780	72.365	72.155
Shelton High School	72.932	72.507	70.097	72.332
Thomas High School	69.234	70.057	69.658	69.370
Wilson High School	69.212	69.455	68.379	69.787
Wright High School	71.741	72.179	73.276	70.848

5. Comparing the average overall passing score of independent schools and government schools to the area's average, Independent schools have a higher overall passing rate of 76.97 compared to 70.70 and 72.80 for government schools and the area average.
6. Similarly, the performance of independent schools is above the area average and that of government schools is below the area average when analysing various matrices like average math and reading scores, and pass per cent in math and reading.

	Total no:of Schools	Total no:of Students	Total Budget	Average Math Score	Average Reading Score	% Passing Maths	% Passing Reading	% Overall Passing
0	15	39,170	\$24,649,428.00	70.338192	69.980138	86.078632	84.426857	72.808272

	Average Math Score	Average Reading Score	% Passing Math	% Passing Reading	% Overall Passing
School Type					
Government	69.83	69.68	84.46	83.59	70.70
Independent	71.37	70.72	89.20	86.25	76.97



7. *There appears to be no direct correlation between the overall passing rate and the average budget allocated per student.*

	Average Math Score	Average Reading Score	% Passing Math	% Passing Reading	% Overall Passing
<b>Spending Ranges (Per Student)</b>					
<\$585	71.36	70.72	88.84	86.39	76.72
\$585-630	72.07	71.03	91.52	87.29	79.88
\$630-645	69.85	69.84	84.69	83.76	71.00
\$645-680	68.88	69.05	81.57	81.77	66.76

8. *The overall passing rate, the average math and reading scores and the average math and reading pass per cent decrease as the school size increase.*

	Average Math Score	Average Reading Score	% Passing Math	% Passing Reading	% Overall Passing
<b>School Size</b>					
Small (<1000)	72.34	71.64	90.81	87.56	79.07
Medium (1000-2000)	71.42	70.72	89.85	86.71	78.04
Large (2000-5000)	69.75	69.58	84.25	83.30	70.29

9. *Independent schools are seen to perform better than government schools when analysed based on the average math and reading scores, average math and reading pass per cent and overall pass per cent.*

	Average Math Score	Average Reading Score	% Passing Math	% Passing Reading	% Overall Passing
<b>School Type</b>					
Government	69.83	69.68	84.46	83.59	70.70
Independent	71.37	70.72	89.20	86.25	76.97