

Source: <https://tinyurl.com/y3a2zcy>



Source: <https://tinyurl.com/y5efbcow>

## Pilot project to improve the English Language Arts (ELA) assessment results

April 2019

Francihelena Uzcategui

## ❖ Introduction

For the 2014-2015 school year, the superintendent had proposed and conducted a pilot project: a new iPad-based learning system in five schools selected 610534, 609880, 610066, 610082, 610182.

“All 5th graders in the five schools received a school-issued iPad. During the school day, students spent 1 hour per day completing an ELA program on the iPad. The ELA program had levels of difficulty ranging from 1 - 10, with each level having activities on reading comprehension, writing, and vocabulary. Students needed to demonstrate their mastery of the ELA standards on each level to proceed to a higher and more difficult level.”

After applied the pilot test, the superintendent wanted a report to evaluate the feasibility and possible expansion of the project along all-district.

The Education Pioneers dataset has univariate and bivariate data to analyze - relations between Education and Demographic indexes, and how they affect the growth and development of students.

## ❖ **Data wrangling**

The data used for the analysis is from Education Pioneers provided for a fellowship challenge. It was a csv file with three tables, each of them with information about ELA results, Math results, and Demographic information (4 MB).

## ❖ **Data cleaning and Exploratory Data Analysis**

We used the Python libraries: Pandas, Numpy, and Matplotlib, to develop the queries, access to the tables, and compile the target information.

In general, to prepare the dataset, we imported the csv file as a data frame, verified data types and changed its, and selected only target columns.

There were several null values, unnamed columns, and duplicates values. We deal with them by `filter()`, `dropna()`, and `duplicated().sum()`.

## ❖ Questions

***1- How did the pilot school 5th graders' ELA mean scores compare to the non-pilot 5th graders in 2017 (2016 - 2017 school year)?***

**1.1- What was the average score of all the pilot schools combined?**

School_ID	Year	Test Name	# Students Tested	Overall_ELA_Mean_Scale_Score_numbers	Reading Mean Scale Score	Writing Mean Scale Score
1540	609880	2017 ELA/Literacy Grade 5	29	792.0	64	40
1669	610066	2017 ELA/Literacy Grade 5	27	786.0	63	43
1679	610082	2017 ELA/Literacy Grade 5	102	774.0	55	38
1753	610182	2017 ELA/Literacy Grade 5	24	768.0	56	39
1854	610534	2017 ELA/Literacy Grade 5	62	795.0	68	44
4915	609880	2016 ELA/Literacy Grade 5	30	791.0	68	42
5045	610066	2016 ELA/Literacy Grade 5	29	772.0	59	39
5055	610082	2016 ELA/Literacy Grade 5	112	768.0	54	38
5131	610182	2016 ELA/Literacy Grade 5	28	764.0	54	38
5232	610534	2016 ELA/Literacy Grade 5	58	793.0	64	45

The average score of the five pilot schools combined was 780.3

## 1.2- What was the average score of all the non-pilot schools combined?

The below table doesn't includes all non-pilot schools, only the first five rows.

School_ID	Year	Test Name	# Students Tested	Overall_ELA_Mean_Scale_Score_numbers	Reading Mean Scale Score	Writing Mean Scale Score
1408	400009	2017 ELA/Literacy Grade 5	30	709.0	36	18
1409	400011	2017 ELA/Literacy Grade 5	49	744.0	47	33
1410	400021	2017 ELA/Literacy Grade 5	55	718.0	38	24
1411	400023	2017 ELA/Literacy Grade 5	55	733.0	43	30
1412	400024	2017 ELA/Literacy Grade 5	65	734.0	44	30

The average score of all the non-pilot schools combined was 725.2

### ***Method used to prepare this first dataset:***

To prepare the dataset, we imported the csv file as a data frame, verified the data type, changed the Overall\_ELA\_Mean\_Scale\_Score data type (string to float), and selected only the specifics columns.

Filtering-"School\_ID", "Year", "Test Name", "# Students Tested", "Overall\_ELA\_Mean\_Scale\_Score\_numbers", "Reading Mean Scale Score", "Writing Mean Scale Score".

### ***Method used to obtain 1.1 and 1.2 answers:***

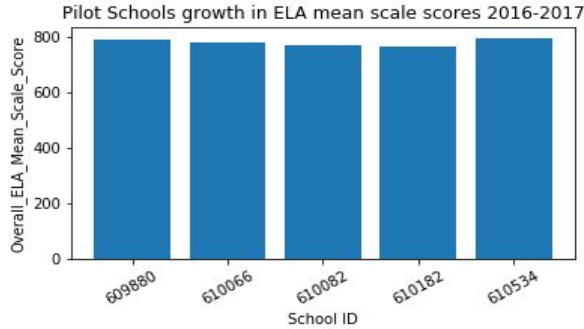
We filtered by the five pilot schools, nonpilot schools, and ELA/Literacy Grade 5.

For the five pilot schools we used .isin() and .mean() method.

For the nonpilot schools, we used a lambda expression: lambda x:x not in and the .mean() method. We applied the formula mean of the means to obtain the total\_mean\_pilot\_schools\_5thgrade.

## 2- Which pilot school exhibited the most growth in ELA mean scale scores?

### 2.1- Which pilot school exhibited the most growth in ELA mean scale scores from the 2016 test through the 2017 test



The school ID 610534 had the best performance ELA mean scale score 794.0

### 2.2- What is the percentage difference in this school's 5th grade growth from 2016 through 2017 compared to the other pilot schools' 5th grade average growth during the same time period?

Pilot_school_ID	Percentage difference 2016 & 2017
610534	0.251889
609880	0.126342
610066	1.797176
610082	0.77821
610182	0.522193

The school ID 610534 had a percentage difference of 0.251. Concerning the other pilot schools, they didn't show relevant variances.

***Method used to obtain 2.1 and 2.2 answers:***

Mainly, we grouped by pilot schools-School ID and ELA scores.`groupby('School_ID').Overall_ELA_Mean_Scale_Score_numbers.mean()`.

Then we plotted a bar graph with this information.

To calculate the percentage difference we used this formula:  $(\text{First Value} - \text{Second Value} / (\text{First Value} + \text{Second Value}) / 2) \times 100\%$



### ***3- What was the average percentage of students who received free/reduced lunch in the 5 pilot schools in 2017?***

	School_Year	School_ID	Total	Free_Reduced_Lunch_N	Free_Reduced_Lunch_%
13	2017	610182	360	199	0.5528
68	2017	609880	292	40	0.1370
380	2017	610066	196	64	0.3265
501	2017	610082	1107	190	0.1716
507	2017	610534	488	46	0.0943

Only 22% average percentage of students who received free/reduced lunch in the five pilot schools in 2017

### ***4- What was the average ELA mean test score of the non-pilot schools that have an equal or greater percentage of students who qualify for free/reduced lunch as the pilot schools? Only look at the 5th-grade class in the 2017 school year.***

In 2017, 79.72 % average percentage of students of non pilot schools received free/reduced lunch, and the average ELA mean test score was 723.96.

***Method used to prepare this second dataset:***

We imported the Demographic csv file as a data frame, verified the data type, and selected only the specific columns. Filtering-"School\_Year","School\_ID", "Total", "Free\_Reduced\_Lunch\_N", "Free\_Reduced\_Lunch\_%. Then, we merged the two data sets into one (first and second).

***Method used to obtain 3rd and 4th answer:***

We filtered nonpilot and pilot schools, ELA scores, Year, Free\_Reduced\_Lunch\_N, and Free\_Reduced\_Lunch\_%.

Finally, to determine the average percentages, we applied the simple rule of three.

## ❖ Limitations

Missing values-numerical values: Overall\_ELA\_Mean\_Scale\_Score column has several zeros values to handle. First, we replaced the zeros to NaN, later calculated the mean and reassigned this value to NaN, then we created new columns and dropped the initial columns with several zeros.

Object type: we verified the object type to select the method to use; otherwise, we got confused.

## ❖ Sources

Education Pioneers-Challenge exercise

Shivam, K. *Python -pandas.to\_numeric method*. Retrieve from [https://www.geeksforgeeks.org/python-pandas-to\\_numeric-method/](https://www.geeksforgeeks.org/python-pandas-to_numeric-method/)

Pythontic. *Subtract Method Of Counter*. Retrieve from <https://pythontic.com/containers/counter/subtract>

Python Software Foundation. *Container datatypes*. Retrieve from <https://docs.python.org/3/library/collections.html>