

Predicting High School AP Class Performance Using the Elementary School Performance in the Same District

Members, Team 14: Chenghao Gong, Thinh Le, Dapeng Zhang, Huadong Zhang, Abhimanyu Patel, Yifan Li

The Problem addressed and Research Question:

- Parents are constantly worried about the future of their children
- Our research is shaped to help the parents of san diego county and project it towards the whole country.
- The question we address is whether or not the elementary school performance affects the AP class passing rates of the students for the San diego county?





Hierarchy

- Elementary School Data
- High School Data
- High School/Elementary School Mapping
- Examining the Correlation
- Regression Model



Hierarchy

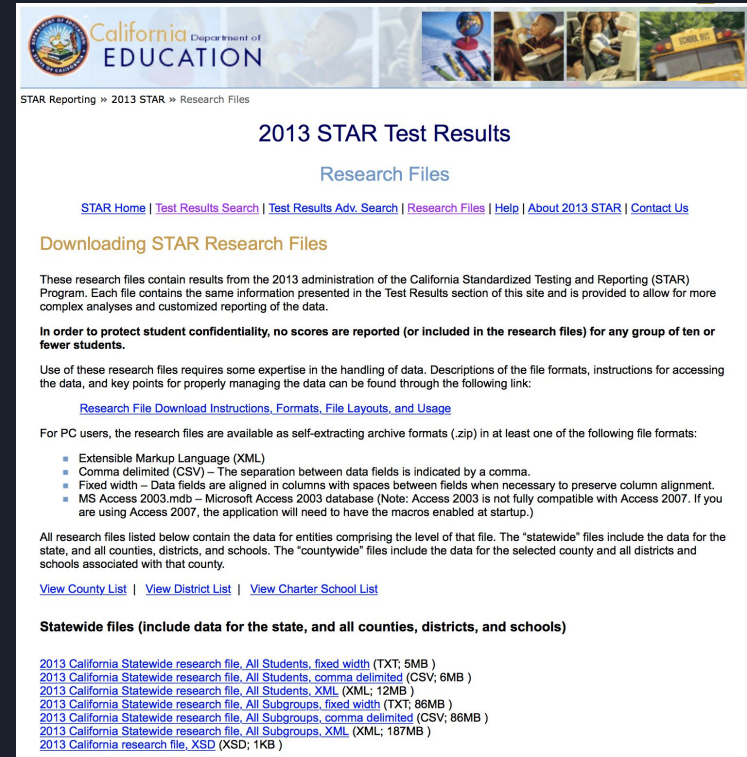
- Elementary School Data
- High School Data
- High School/Elementary School Mapping
- Examining the Correlation
- Regression Model

Elementary School Data

Data range: 2009-2014

<https://star.cde.ca.gov/star2013/ResearchFileList.aspx?rf=True&ps=True>

California Standardized Testing and Reporting (STAR)



The screenshot shows the "2013 STAR Test Results" page under the "Research Files" section. The header includes the California Department of Education logo and a navigation bar with links like "STAR Home", "Test Results Search", and "Research Files". The main content area is titled "2013 STAR Test Results" and "Research Files". It provides information about downloading research files, including a warning about student confidentiality and instructions for handling the data. A list of file formats (XML, CSV, XSD) is provided, along with a list of available files for download, including statewide and county-level data for various years and formats.

STAR Reporting » 2013 STAR » Research Files

2013 STAR Test Results

Research Files

[STAR Home](#) | [Test Results Search](#) | [Test Results Adv. Search](#) | [Research Files](#) | [Help](#) | [About 2013 STAR](#) | [Contact Us](#)

Downloading STAR Research Files

These research files contain results from the 2013 administration of the California Standardized Testing and Reporting (STAR) Program. Each file contains the same information presented in the Test Results section of this site and is provided to allow for more complex analyses and customized reporting of the data.

In order to protect student confidentiality, no scores are reported (or included in the research files) for any group of ten or fewer students.

Use of these research files requires some expertise in the handling of data. Descriptions of the file formats, instructions for accessing the data, and key points for properly managing the data can be found through the following link:

[Research File Download Instructions, Formats, File Layouts, and Usage](#)

For PC users, the research files are available as self-extracting archive formats (.zip) in at least one of the following file formats:

- Extensible Markup Language (XML)
- Comma delimited (CSV) – The separation between data fields is indicated by a comma.
- Fixed width – Data fields are aligned in columns with spaces between fields when necessary to preserve column alignment.
- MS Access 2003.mdb – Microsoft Access 2003 database (Note: Access 2003 is not fully compatible with Access 2007. If you are using Access 2007, the application will need to have the macros enabled at startup.)

All research files listed below contain the data for entities comprising the level of that file. The "statewide" files include the data for the state, and all counties, districts, and schools. The "countywide" files include the data for the selected county and all districts and schools associated with that county.

[View County List](#) | [View District List](#) | [View Charter School List](#)

Statewide files (include data for the state, and all counties, districts, and schools)

[2013 California Statewide research file, All Students, fixed width \(TXT; 5MB\)](#)
[2013 California Statewide research file, All Students, comma delimited \(CSV; 6MB\)](#)
[2013 California Statewide research file, All Students, XML \(XML; 12MB\)](#)
[2013 California Statewide research file, All Subgroups, fixed width \(TXT; 86MB\)](#)
[2013 California Statewide research file, All Subgroups, comma delimited \(CSV; 86MB\)](#)
[2013 California Statewide research file, All Subgroups, XML \(XML; 187MB\)](#)
[2013 California research file, XSD \(XSD; 1KB\)](#)

Elementary School Data

Important information:

- Mean Scale Score
- Percentage of Advanced
- Percentage Basic
- Percentage Below Basic
- Percentage Far Below
- Percentage Proficient

H	I		J		K	L	M	N	O
Total CAASPP Enrollment	Total Tested At Entity	Level	Total Tested with Scores	Grade	Test Id	CAASPP Reported Enrollment	Students Tested	Mean Scale Score	
3154	2744	2744	2744	3	3	445	390	338.6	
3154	2745	2745	2745	3	4	445	392	330.3	
3154	2745	2745	2745	4	4	428	385	431.7	
3154	2744	2744	2744	4	3	428	384	437	
3154	2745	2745	2745	5	4	444	376	531.1	
3154	2744	2744	2744	5	3	444	377	536.7	
3154	2744	2744	2744	6	3	449	406	639	
3154	2745	2745	2745	6	4	449	406	631.9	
3154	2745	2745	2745	7	4	471	417	732.9	
3154	2744	2744	2744	7	3	471	421	737.6	
3154	2744	2744	2744	8	3	462	415	841.7	
3154	2745	2745	2745	8	4	462	416	834	
3154	2745	2745	2745	11	4	455	353	937.1	
3154	2744	2744	2744	11	3	455	351	942.8	
3154	2745	2745	2745	13	4	3154	2745		
3154	2744	2744	2744	13	3	3154	2744		
2148	1870	1870	1870	3	4	291	255	330.3	
2148	1871	1871	1871	3	3	291	253	338.6	
2148	1871	1871	1871	4	3	299	269	437.3	
2148	1870	1870	1870	4	4	299	271	432.5	
2148	1871	1871	1871	5	3	289	245	538.2	
2148	1870	1870	1870	5	4	289	243	532.8	
2148	1870	1870	1870	6	4	314	287	631.5	
2148	1871	1871	1871	6	3	314	288	638.4	
2148	1871	1871	1871	7	3	316	286	737.2	
2148	1870	1870	1870	7	4	316	283	733	
2148	1870	1870	1870	8	4	315	283	834	
2148	1871	1871	1871	8	3	315	282	841.2	
2148	1870	1870	1870	11	4	324	248	936.6	
2148	1871	1871	1871	11	3	324	248	942.3	
2148	1870	1870	1870	13	4	2148	1870		
2148	1871	1871	1871	13	3	2148	1871		
1006	873	873	873	3	3	154	137	338.5	
1006	875	875	875	3	4	154	137	330.3	
1006	873	873	873	4	3	129	115	436.3	
1006	875	875	875	4	4	129	114	429.7	
1006	873	873	873	5	3	155	132	533.9	
1006	875	875	875	5	4	155	133	528.2	
1006	875	875	875	6	4	135	119	633	
1006	873	873	873	6	3	135	118	640.4	
1006	875	875	875	7	4	155	134	732.7	
1006	873	873	873	7	3	155	135	738.5	
1006	873	873	873	8	3	147	133	842.7	
1006	875	875	875	8	4	147	133	834	
1006	875	875	875	11	4	131	105	938.1	
1006	873	873	873	11	3	131	109	944	
1006	875	875	875	13	4	1006	875		



Hierarchy

- Elementary School Data
- High School Data
- High School/Elementary School Mapping
- Exam the Correlation
- Regression Model

High School Data

Advanced Placement(AP) pass rate data of high schools (in San Diego) from:

https://www.sandiegounified.org/advanced_placement

San Diego Unified SCHOOL DISTRICT

About Academics Schools Enrollment

Advanced Placement

Home

Overview

Contact Us

Course of Study

Course Information

My School

District Reports

Mandated Reporting

Schools in Session

Advanced Placement (AP) Program 2016

The Advanced Placement (AP) program consists of college-level courses and corresponding end-of-course examinations. It offers high school students the opportunity to earn credit or advanced standing at participating colleges and universities across the country.

AP courses cover a range of subject areas including art and music, language and literature (English and other languages), mathematics, science, and history and social science. In 2015-16, district schools' collective AP course offerings enabled students to take 33 different AP examinations.

AP examinations are developed by the College Board and administered each May. Tests consist of an open-response section (essay or problem-solving) and multiple choice questions. Exceptions to this test format include: (1) AP Studio Art, which is a portfolio assessment; (2) world language tests, which have a speaking component; and (3) AP Music Theory, which includes a sight-singing task. For more information, please see the [College Board AP website](#).

[Click for the Course Enrollment Summary.](#)

San Diego Unified SCHOOL DISTRICT

Advanced Placement (AP) Report for 2015-2016

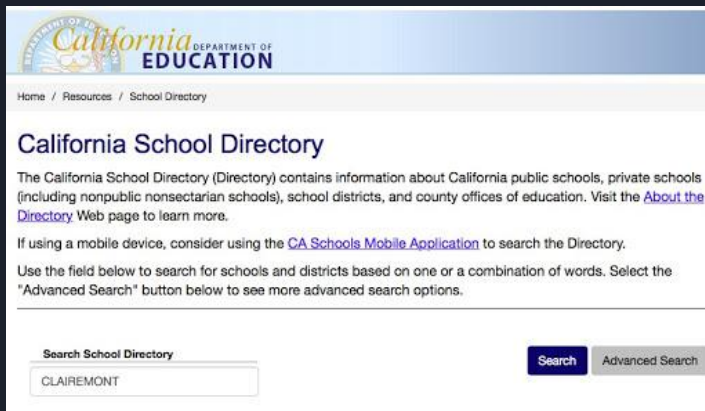
Grade 10 to 12 Student Summary By School

School	Total Gr 10-12 Students	Enrld in AP Course	AP Enrld and Took an AP Test	AP Enrld and Passed an AP Test
		N	N	N
		Pct	Pct of Enrld	Pct of Enrld/Passed
331 ALBA	15	0	0.0	0
332 CLAIREMONT	680	264	38.8	174
334 CRAWFORD	724	193	26.7	157
361 GARFIELD HIGH	235	2	0.9	0
336 HENRY	1,630	765	46.9	637
382 HOME & HOSP/TRANSITION SUPPORT	28	0	0.0	0
338 HOOVER	1,255	255	20.3	241
504 IHIGH VIRTUAL ACADEMY	51	9	17.6	3
733 KEARNY DIGITAL MEDIA & DESIGN	205	12	5.9	4
736 KEARNY ENG, INNOV & DESIGN	186	3	1.6	0
735 KEARNY INTERNATIONAL BUSINESS	226	11	4.9	9
734 KEARNY SCT	214	37	17.3	33
342 LA JOLLA HIGH	1,138	749	65.8	574
637 LINCOLN	848	259	30.5	189
346 MADISON	767	249	32.5	194
349 MIRA MESA	1,669	730	43.7	577
350 MISSION BAY	707	154	21.8	72
352 MORSE	1,158	322	27.8	290
395 MT. EVEREST	92	54	58.7	41
369 MUIR	23	0	0.0	0

High School Data

To make the combination, we got their CDS code from:

<https://www.cde.ca.gov/SchoolDirectory/districtschool?allSearch=ALBA&simpleSearch=Y&page=0&tab=1>



Home / Resources / School Directory

California School Directory

The California School Directory (Directory) contains information about California public schools, private schools (including nonpublic nonsectarian schools), school districts, and county offices of education. Visit the [About the Directory](#) Web page to learn more.

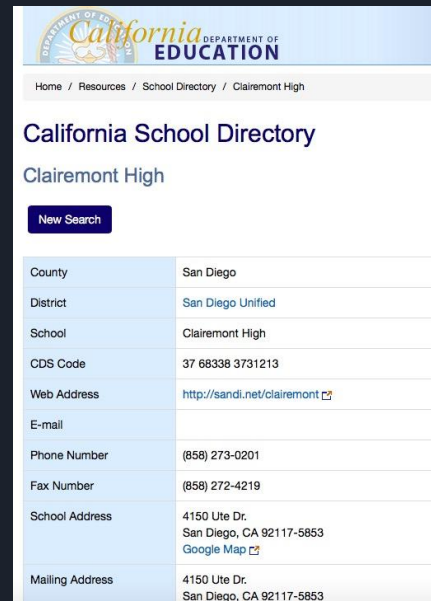
If using a mobile device, consider using the [CA Schools Mobile Application](#) to search the Directory.

Use the field below to search for schools and districts based on one or a combination of words. Select the "Advanced Search" button below to see more advanced search options.

Search School Directory

CLAIREMONT

Search Advanced Search



County	San Diego
District	San Diego Unified
School	Clairemont High
CDS Code	37 68338 3731213
Web Address	http://sandi.net/clairemont
E-mail	
Phone Number	(858) 273-0201
Fax Number	(858) 272-4219
School Address	4150 Ute Dr. San Diego, CA 92117-5853 Google Map
Mailing Address	4150 Ute Dr. San Diego, CA 92117-5853

High School Data

Including :

CDS Code

Area Code

School Name

AP Pass Rate

From 2009-2014

	A	B	C	D
1	CDS Code	Area Code	School	AP Examinations Passed
2	37683383731130	331	ALBA	NA
3	37683383731213	332	CLAIREMONT	70
4	37683380107177	334	CRAWFORD	29.8
5	37683383737962	361	GARFIELD HIGH	NA
6	37683383732781	336	HENRY	69.1
7	37683380126938	382	HOME & HOSP/TRANSITION SUPPORT	NA
8	37683383732997	338	HOOVER	56.1
9	37683380120410	504	IHIGH VIRTUAL ACADEMY	NA
10	37683380107078	733	KEARNY DIGITAL MEDIA & DESIGN	0
11	37683380107102	736	KEARNY ENG, INNOV & DESIGN	45.7
12	37683380107086	734	KEARNY SCT	38.7
13	37683383733508	342	LA JOLLA HIGH	78.7
14	37683380114025	637	LINCOLN	29.6
15	37683383733698	346	MADISON	51.5
16	37683383730181	349	MIRA MESA	66.4
17	37683383734431	350	MISSION BAY	67.6
18	37683383734654	352	MORSE	58.9
19	37683383735750	354	POINT LOMA	60.1
20	37683386040083	364	RILEY/NEW DAWN	NA
21	37683380107052	749	SAN DIEGO BUSINESS/LEADERSHIP	27.7
22	37683380107029	744	SAN DIEGO INT'L STUDIES	67.5
23	37683380107482	500	SAN DIEGO METRO CAREER & TECH	NA
24	37683380107060	750	SAN DIEGO MVP ARTS	NA
25	37683380107219	753	SAN DIEGO SCI TECH	25.3
26	37683383730371	368	SCPA	62.5
27	37683383730884	359	SCRIPPS RANCH	85.7
28	37683383730173	357	SERRA	53.9
29	37683383730231	362	TWAIN	NA
30	37683383730314	355	UNIVERSITY CITY	73
31				
32				
33				
34				
35				

	A	B	C	D
1	CDS Code	Area Code	School	AP Examinations Passed
2	37683383731130	331	ALBA	NA
3	37683383731213	332	CLAIREMONT	46.6
4	37683380107185	704	CRAWFORD CHAMPS	14.8
5	37683380107201	702	CRAWFORD IDEA	22
6	37683380107193	705	CRAWFORD LAW & BUSINESS	15
7	37683383737962	361	GARFIELD HIGH	NA
8	37683383732781	336	HENRY	60.7
9	37683380126938	382	HOME & HOSP/TRANSITION SUPPORT	NA
10	37683383732997	338	HOOVER	24.7
11	37683380120410	504	IHIGH VIRTUAL ACADEMY	NA
12	37683380107102	736	KEARNY ENG, INNOV & DESIGN	4.8
13	37683380107078	733	KEARNY DIGITAL MEDIA & DESIGN	24.6
14	37683380107086	734	KEARNY SCT	45.1
15	37683383733508	342	LA JOLLA HIGH	70.6
16	37683380114025	637	LINCOLN	22
17	37683383733698	346	MADISON	28.9
18	37683383730181	349	MIRA MESA	55.4
19	37683383734431	350	MISSION BAY	71.6
20	37683383734654	352	MORSE	38
21	37683383735750	354	POINT LOMA	41.1
22	37683386040083	364	RILEY/NEW DAWN	NA
23	37683380107052	749	SAN DIEGO BUSINESS/LEADERSHIP	36.4
24	37683380107045	746	SAN DIEGO COMMUNICATION	NA
25	37683380107029	744	SAN DIEGO INT'L STUDIES	55.2
26	37683380107037	745	SAN DIEGO LEADS	NA
27	37683380107482	500	SAN DIEGO METRO CAREER & TECH	NA
28	37683380107060	750	SAN DIEGO MVP ARTS	NA
29	37683380107219	753	SAN DIEGO SCI TECH	NA
30	37683383730371	368	SCPA	42.4
31	37683383730884	359	SCRIPPS RANCH	80.5
32	37683383730173	357	SERRA	39.5
33	37683383730231	362	TWAIN	NA
34	37683383730314	355	UNIVERSITY CITY	57.1



Hierarchy

- Elementary School Data
- High School Data
- High School/Elementary School Mapping
- Exam the Correlation
- Regression Model



Elementary/High School Mapping

Elementary School Data

District Codes	% Scored Advanced	%Scored Proficient
----------------	-------------------	--------------------

0
---	-----	-----

1
---	-----	-----

2
---	-----	-----

3
---	-----	-----

.
---	-----	-----

.
---	-----	-----

n
---	-----	-----

High School Data

District Codes	AP Pass Rates
----------------	---------------

0	...
---	-----

1	...
---	-----

2	...
---	-----

3	...
---	-----

.	...
---	-----

.	...
---	-----

n	...
---	-----

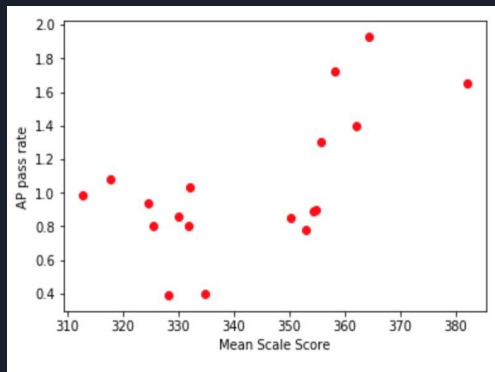


Hierarchy

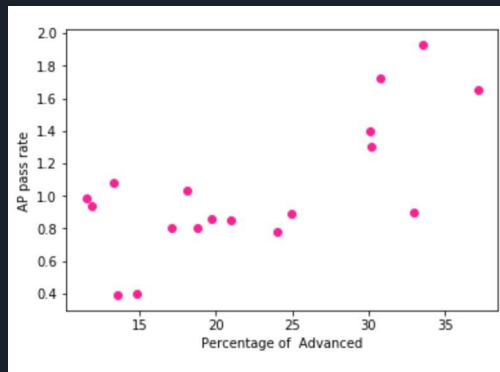
- Elementary School Data
- High School Data
- High School/Elementary School Mapping
- Examining the Correlation
- Regression Model

Correlation of Variables in 2011

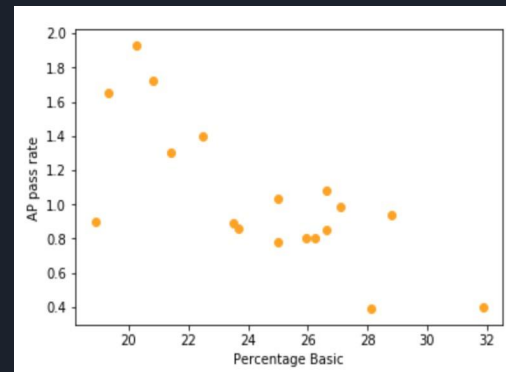
Mean Scale Score



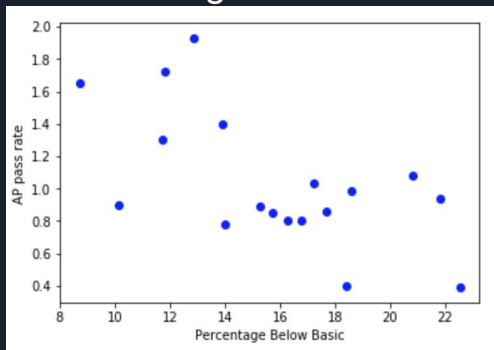
Percentage of Advanced



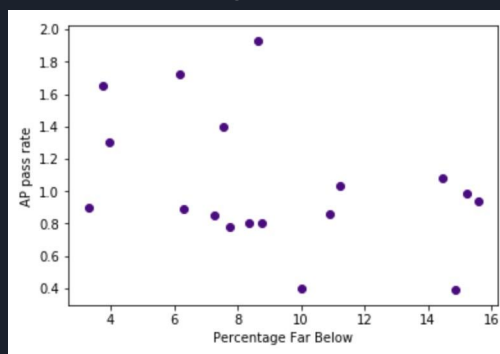
Percentage Basic



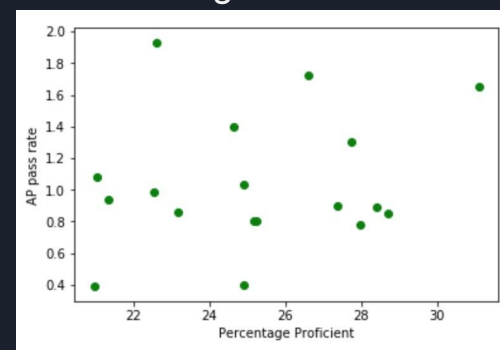
Percentage Below Basic



Percentage Far Below



Percentage Proficient



Correlation of Variables in 2009

'2009 Mean Scale Score':

0.174753168, 4.49354999

4.49354999, 505.046350

'2009 Percentage Basic':

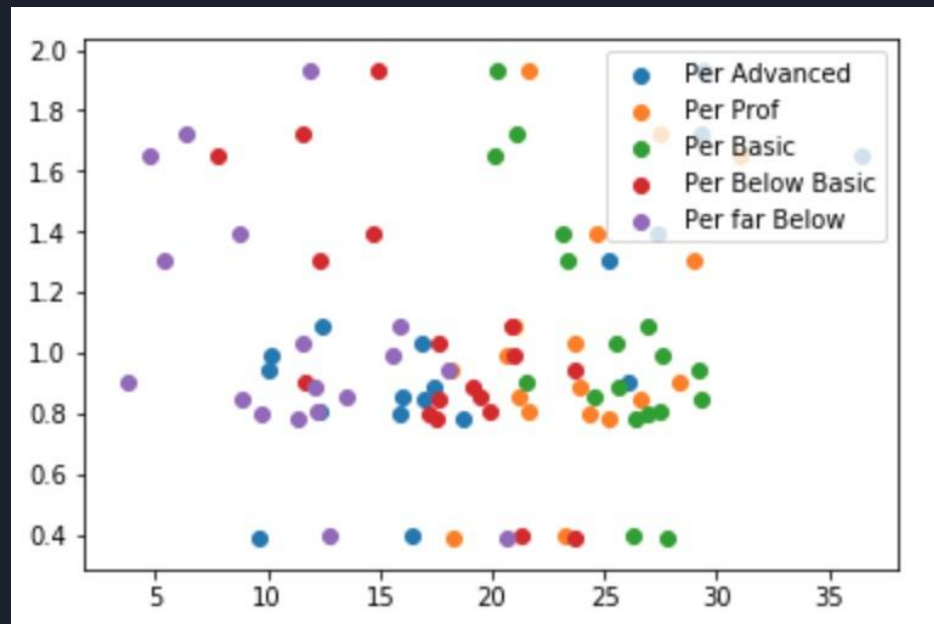
0.17475317, -0.69289134

-0.69289134, 22.12322183

'2009 Percentage Proficient':

0.17475317, -0.01977771

-0.01977771, 15.86612157





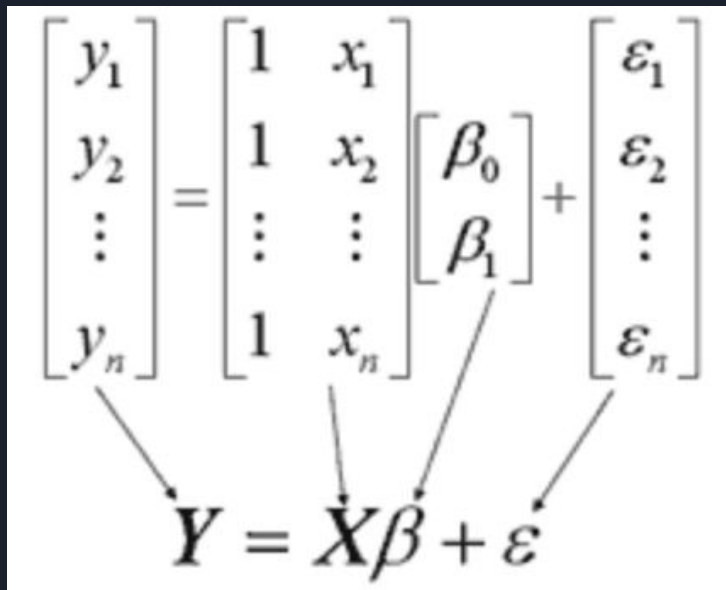
Hierarchy

- Elementary School Data
- High School Data
- High School/Elementary School Mapping
- Examining the Correlation
- Regression Model

Linear Regression

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$Y = X\beta + \varepsilon$



Objective: Find Beta that makes the error as small as possible

Y: High school AP performance

X: Elementary school performance

Beta: Parameters

e: Error



Equations to find Beta

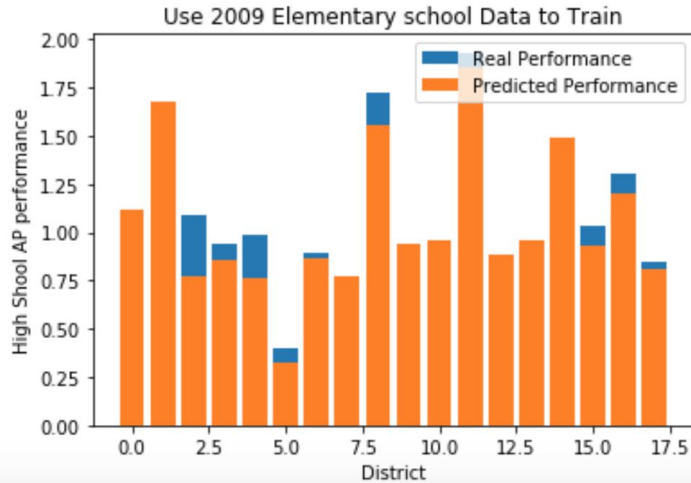
$$\text{Beta} = (X^T X)^{-1} X^T Y$$

Then we plug Beta back and find $G = \text{Beta} * X$ and comparing G with Y

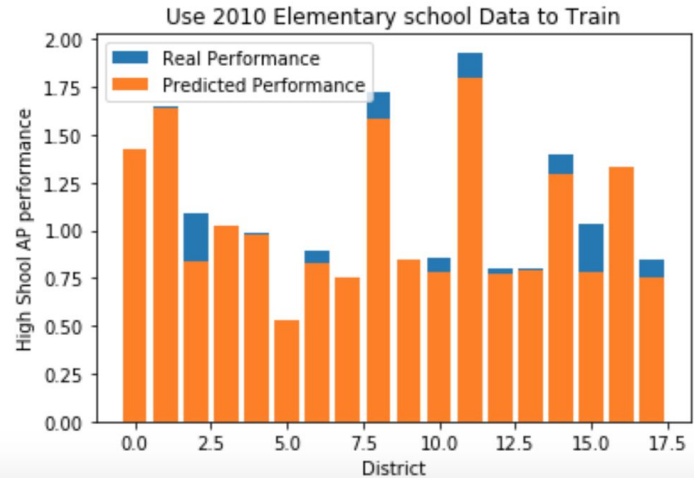
G: Predicted Y

Results of Linear Regression:

Average Error is:
0.135759303129

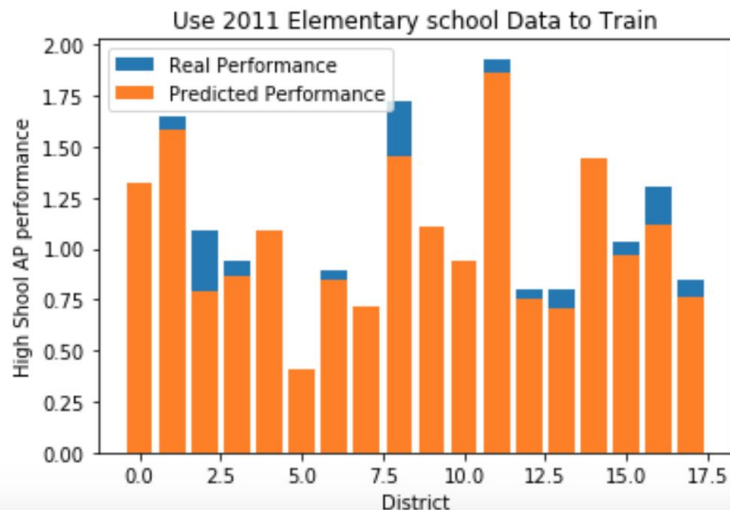


Average Error is:
0.13204368343

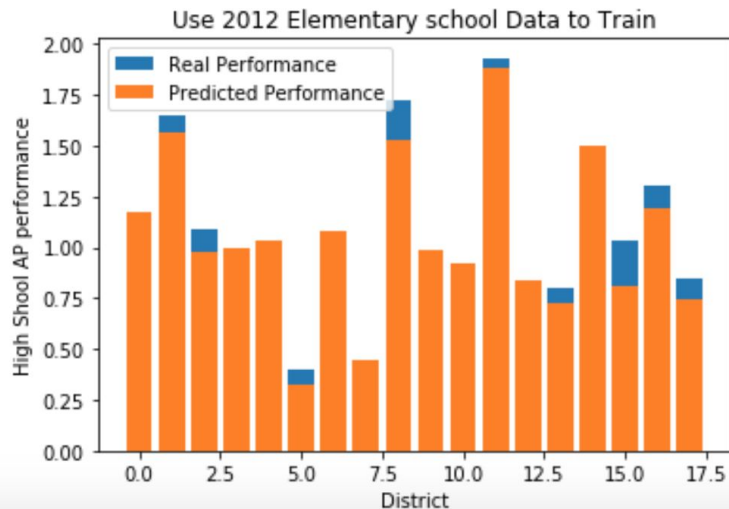


Results of Linear Regression (Cont):

Average Error is:
0.145673661587

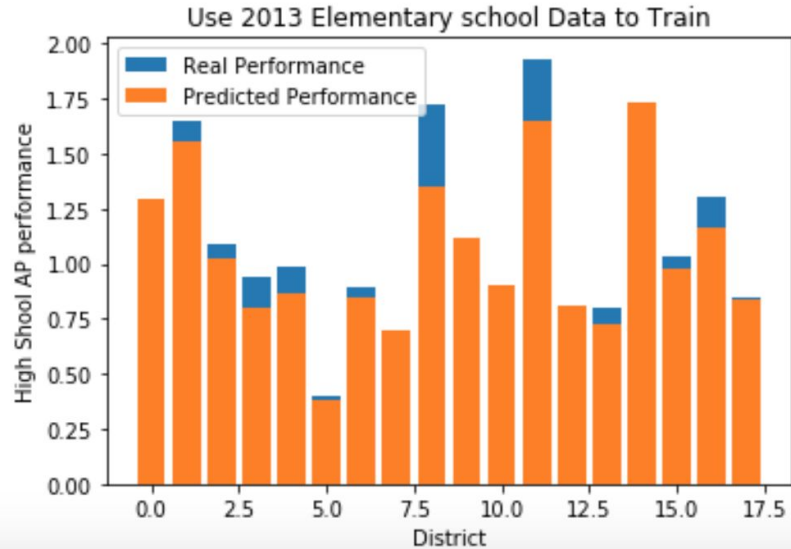


Average Error is:
0.11398609033

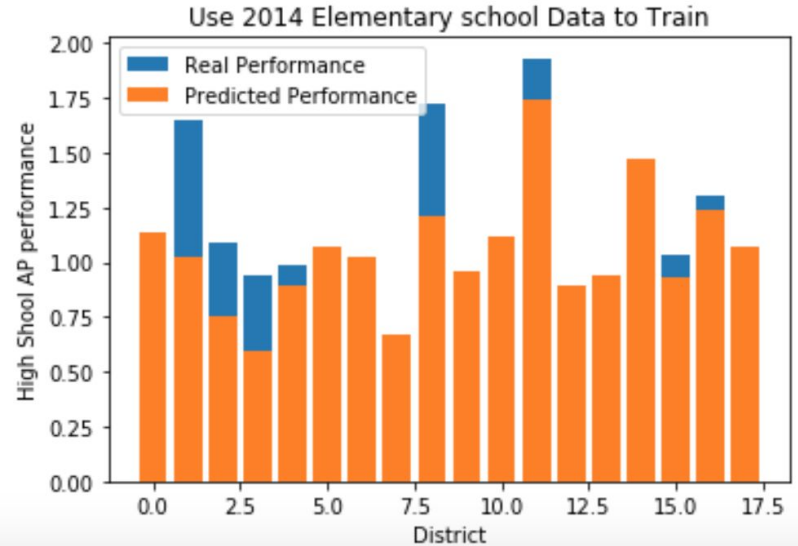


Results of Linear Regression (Cont):

Average Error is:
0.158841697762



Average Error is:
0.252576691706





Thank You!

Questions?