**FRPM – Student Poverty Data Case Study**

The California Department of Education (CDE) oversees the state's public school system, which is responsible for the education of more than six million children and young adults in more than 10,000 schools with 295,000 teachers.

The Director of Education sent you an email earlier today:

From: Director of Education

To: You (Analytics Engineer)

Subject: Impact of free school meals and school type on education

Hi - we’re glad to have assigned you to this project. I will soon be discussing the **impact** of **free school meals** and **school typ**e on **exam scores** to the Board of Education. I need answers to a few questions so I can bring the relevant data and insights to the meeting. Can you help me?

Here’s how you can get the data:

* Use this link to fetch the database <http://2016.padjo.org/files/data/starterpack/cde-schools/cdeschools.sqlite>
* It’s in SQLite format so it should be very familiar for you
* You can use any tool that can help you navigate the data

I need to know:

* What are the 10 best high schools for maths scores?
* Which counties don’t have any SAT scores?
* Are locally funded charter schools more likely than directly funded charter schools to enrol students on the FRPM (Free or Reduced-Price meals) program?
* Does being on the FRPM program impact math scores?
* Which school types are the best at achieving good reading and writing scores for students on the FRPM program?

We don’t have a lot of time, so I want you to pull the data together and run a quick analysis on each of these and report back tomorrow. Some of the queries are open to interpretation and **I’d like for you to send me your answers and opinions** to the questions above and **let me know your reasoning.**

**I’d like a clear (and brief) word document, rather than a powerpoint, accompanied by any data visualisations you consider necessary.**

Thanks, and we look forward to hearing from you.

* Director of Education

**Q1: What are the 10 best high schools for maths scores?**

The top 10 high schools by average math scores are as follows:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| AvgScrMath | School | District | County | City | EdOpsName | FundingType |
| 699.0 | Mission San Jose High | Fremont Unified | Alameda | Fremont | Traditional | None |
| 698.0 | Lynbrook High | Fremont Union High | Santa Cliara | San Jose | Traditional | None |
| 691.0 | Monta Vista High | Fremont Union High | Santa Clara | Cupertino | Traditional | None |
| 686.0 | Henry M. Gunn High | Palo Alto Unified | Santa Clara | Palo Alto | Traditional | None |
| 674.0 | Saratoga High | Los Gatos-Saratoga Joint Union High | Santa Clara | Saratoga | Traditional | None |
| 666.0 | University High | Irvine Unified | Orange | Irvine | Traditional | None |
| 660.0 | Cupertino High | Fremont Union High | Santa Clara | Cupertino | Traditional | None |
| 657.0 | San Marino High | San Marino Unified | Los Angeles | San Marino | Traditional | None |
| 657.0 | Palo Alto High | Palo Alto Unified | Santa Clara | Palo Alto | Traditional | None |
| 656.0 | Northwood High | Irvine Unified | Orange | Irvine | Traditional | None |

* 4 out of 10 are from the Fremont Union High District
* 6 out of 10 are from Santa Clara County
* all of these counties are classified as having schools managed in a traditional manner

**Q2: Which counties don’t have any SAT scores?**

Based on our current data, 57 counties have no information about SAT scores. These counties are:

|  |  |  |  |
| --- | --- | --- | --- |
| Alameda  Alpine  Amador  Butte  Calaveras  Colusa  Contra Costa  Del Norte  El Dorado  Fresno  Glenn  Humboldt  Imperial  Inyo  Kern | Kings  Lake  Lassen  Los Angeles  Madera  Marin  Mariposa  Mendocino  Merced  Modoc  Mono  Monterey  Napa  Nevada | Orange  Placer  Plumas  Riverside  Sacramento  San Benito  San Bernardino  San Diego  San Francisco  San Joaquin  San Luis Obisp  San Mateo  Santa Barbara  Santa Clara | Santa Cruz  Shasta  Sierra  Siskiyou  Solano  Sonoma  Stanislaus  Sutter  Tehama  Trinity  Tulare  Tuolumne  Ventura  Yolo  Yuba |

**Q3: Are locally funded charter schools more likely than directly funded charter schools to enrol students on the FRPM (Free or Reduced-Price meals) program?**

If we examine the global student counts for the academic year 2014-2015, we get the following distribution of charter schools by funding type:

|  |  |  |  |
| --- | --- | --- | --- |
| Charter Funding Type | Enrollment (K-12) | FRPM Count (K-12) | FRPM/Enrollment Ratio |
|  |  |  |  |
| Directly funded | 392 445 | 239 173 | 61 % |
| Locally funded | 151 713 | 683 69 | 45 % |

If we can trust the numbers we received lately, then we can assume that directly funded charter schools are 1.35 times more likely to offer the FRPM program than locally funded charter schools.

However, we have to keep in mind that in our sample of schools, two out of three are directly funded. Specifically:

* 846 beneficiaries were reported as being in directly funded charter schools
* 330 beneficiaries were reported as being in locally funded charter schools

|  |  |  |  |
| --- | --- | --- | --- |
|  | Charter Funding Type | Directly funded | Locally funded |
|  | **count** | 846 | 330 |
| Enrollment (K-12) |  |  |  |
| **mean** | 464 | 460 |
| **std** | 502 | 487 |
| **min** | 6 | 2 |
| **25%** | 194 | 191 |
| **50%** | 348 | 375 |
| **75%** | 528 | 558 |
| **max** | 5 333 | 5 051 |

Regardless of the total counts for each funding type, both directly funded and locally funded charter schools enroll, on average, around 460 students. By examining the descriptive statistics of both groups, we can safely assume that their distributions of student enrollments are quite similar. This assumption is further supported by a quick look at the histograms  
and their density estimations shown below.

Obraz zawierający tekst, diagram, zrzut ekranu, Wykres

Opis wygenerowany automatycznieObraz zawierający tekst, diagram, zrzut ekranu, Wykres

Opis wygenerowany automatycznie

The assumption about the similarity of distributions was formally confirmed by the Kolmogorov-Smirnov test:

**Kolmogorov-Smirnov Test:**

Statistic: 0.076

P-value: 0.120

Interpretation: Fail to reject the null hypothesis (distributions are similar).

The perspective changes when looking at the distributions of FRPM student counts for each group. On average, locally funded charter schools enroll fewer students in the FRPM program.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Charter Funding Type | Directly funded | Locally funded |
|  | **count** | 846 | 330 |
| FRPM Count (K-12) |  |  |  |
| **mean** | 283 | 207 |
| **std** | 316 | 273 |
| **min** | 1 | 1 |
| **25%** | 83 | 57 |
| **50%** | 198 | 130 |
| **75%** | 379 | 251 |
| **max** | 2 662 | 2 476 |

Obraz zawierający tekst, diagram, zrzut ekranu, Wykres

Opis wygenerowany automatycznie

Obraz zawierający tekst, diagram, zrzut ekranu, Wykres

Opis wygenerowany automatycznie

**Kolmogorov-Smirnov Test:**

Statistic: 0.172, P-value: 0.00001

Interpretation: Reject the null hypothesis (distributions are different)

**Q4: Does being on the FRPM program impact math scores?**

Based on a simple correlation analysis, it seems that enrollment in the FRPM program does not positively impact students' math scores. The correlation matrix indicates a negative relationship between the number of students involved in FRPM and their math scores:

* -0.64 for the correlation between math scores and the percentage of students enrolled in FRPM
* -0.26 for the correlation between math scores and the count of students enrolled  
  in FRPM

Obraz zawierający tekst, zrzut ekranu, diagram, kwadrat

Opis wygenerowany automatycznie

**Q5: Which school types are the best at achieving good reading and writing scores for students on the FRPM program?**

Considering the unbalanced dataset we are working with, the schools that likely achieve the best reading and writing scores are:

* Locally funded schools
* Alternative schools of choice or traditional schools

**Funding Type**

Obraz zawierający tekst, diagram, Plan, linia

Opis wygenerowany automatycznie

Obraz zawierający tekst, diagram, zrzut ekranu, Plan

Opis wygenerowany automatycznie

**School category 1**

Obraz zawierający diagram, tekst

Opis wygenerowany automatycznie

Obraz zawierający diagram, tekst, linia, design

Opis wygenerowany automatycznie

**Category 2 of School**

Obraz zawierający tekst, diagram, Plan, zrzut ekranu

Opis wygenerowany automatycznie

Obraz zawierający tekst, diagram, zrzut ekranu, Plan

Opis wygenerowany automatycznie

**Category 3 of School**

Obraz zawierający tekst, diagram, zrzut ekranu, Równolegle

Opis wygenerowany automatycznie

Obraz zawierający tekst, diagram, zrzut ekranu, Równolegle

Opis wygenerowany automatycznie