



# PAKISTAN SCHOOL DATA ANALYSIS REPORT

Microsoft Excel Portfolio Project

## ABSTRACT

A comprehensive report on Pakistani school dataset analysis justifying the allocation of funds for specific educational initiatives and proposing data-driven marketing strategies to promote those initiatives effectively.

Amina Javaid

[aminajavid30@gmail.com](mailto:aminajavid30@gmail.com)

# Table of Contents

Pakistan School Data Analysis Report .....	3
Context .....	3
Objective .....	3
Problem Statement .....	3
Data Loading .....	4
Data Exploration.....	4
Data Cleaning .....	7
1. Missing Values.....	7
2. Duplicate Values.....	7
3. Individual Column Values.....	7
4. Rows with partially filled values.....	8
Data Analysis and Insights.....	8
1. Total number of schools in the dataset.....	8
2. Schools by Year of Establishment.....	8
3. Number of schools in Rural and Urban areas .....	9
4. Percentage of schools offering Urdu and English medium education or both in Rural and Urban areas.....	10
5. Percentage of schools by level in Rural and Urban areas .....	11
6. Distribution of schools according to school level.....	12
7. Number of schools categorized by gender (Male & Female).....	13
8. School level and medium with the lowest number of functional classrooms .....	14
9. Schools with respect to school ownership.....	16
10. Percentage of schools upgraded .....	16
11. Number of Male and Female students .....	17
12. Lowest number of Male & Female students by school level of schools established in the 2000s .....	
13. School level and medium by students enrolled .....	19
14. District with the highest enrollment in Primary schools since their establishment .....	20
15. District with the highest enrollment in Secondary schools since their establishment .....	22

16.	District with the highest enrollment in Higher Secondary schools since their establishment	24
17.	Top 5 Districts with the highest number of teachers and non-teachers in schools	26
18.	Distribution of teachers across different schools	28
a.	Teachers by School Location and Level	28
b.	Teachers by School Level and Medium	29
c.	Teachers by School Type	31
19.	Distribution of teachers based on their respective job posts	32
20.	Percentage of vacant and filled teaching and non-teaching posts	32
21.	Percentage of schools with satisfactory building conditions	33
22.	Percentage of schools based on their construct type in rural and urban areas	34
23.	Names of 5 Districts with the lowest availability of drinking water facilities in schools	35
24.	Names of 5 Districts with the lowest availability of electricity in schools	37
25.	Names of 5 Districts with the lowest presence of boundary walls in schools	39
26.	Names of 5 Districts with the lowest availability of toilets in schools	41
27.	Number of schools with satisfactory security measures	43
	Summary of Insights	43
	School Level & Upgradation	43
	Student Enrollment	44
	Staff (Teachers / Non-teachers)	45
	Infrastructure	46
	Facilities	46
	Recommendations	47
	Allocation of Funds	47
	Budget Plan	48
	Marketing Strategies	49
	References	50

# Pakistan School Data Analysis Report

## Context

An international ed-tech company has invested 3 Billion dollars in Punjab, Pakistan. The company aims to leverage technology to bridge educational gaps, ensure equal opportunities for all students and improve the infrastructure and facilities. A fund of 3 Million dollars is allocated to the initiatives that will contribute to improving education. The company requires a comprehensive report that justifies the allocation of funds and proposes data-driven marketing strategies. The report should outline specific initiatives and projects that will maximize the impact of the fund, ensuring that no money is left unattended and all resources are utilized optimally.

Following budget has been allocated for various improvements:

1. School level upgradation: Rs 10 Million
2. Infrastructure(toilets, boundary walls, classrooms, building condition): Rs 5 Million
3. Facilities such as drinking water, playgrounds, labs: Rs 2 Million
4. More middle and high schools in rural areas with separate schools for both genders: Rs 20 Million
5. New primary level school: Rs 10 Million

## Objective

Determine how the ed-tech company can have the most impact in improving the current education situation in Punjab.

## Problem Statement

Analyze a dataset on schools in Punjab and conduct a marketing feasibility analysis to generate a report justifying the allocation of funds and proposing data-driven marketing strategies.

## Data Loading

Data is present in the form of a CSV file called **Dataset\_schools – school.csv**.

Loaded the CSV file into Microsoft Excel using the following steps:

- Click on Data tab
- Get Data -> From File -> From Text/CSV
- Select the file and click 'Import'
- Click the 'Load' button

## Data Exploration

There are 48,274 rows and 56 columns.

1. **school\_id** – Unique identification number for each school.
2. **emiscode** – Education Management Information System code assigned to the school.
3. **school\_name** – The name of the school.
4. **district** - District of the school.
5. **tehsil** – Tehsil of the school.
6. **markaz** – Markaz of the school.
7. **moza** – Moza of the school.
8. **permanent\_address** – Permanent address of the school.
9. **street\_name** – Street where the school is located.
10. **uc\_name** – Union council related to the school.
11. **uc\_no** – Union council number.
12. **na\_no** – National Assembly number.
13. **pp\_no** – Provincial Assembly number.
14. **head\_name** – Name of the school head.
15. **head\_type** – Type of the school head (Permanent. Additional Charge, Lookafter Charge).
16. **head\_grade** – Pay scale or grade level of the head.
17. **school\_status** – Status of the school (Functional, Non-Functional).

18. **medium** – Medium of the school (English, Urdu, Both).
19. **school\_shift** – Shift of the school (Morning, Evening, Double).
20. **school\_location** – Location of the school (Rural, Urban).
21. **school\_gender** – The gender for which the school was initially built (Male, Female).
22. **gender\_studying** – The genders currently studying in the school (Male, Female, Both).
23. **school\_level** – The level of the school (Primary, Middle, Secondary, H. Sec, sMosque)
24. **school\_type** – The type of the school from one of the following:
- a. Centre of Excellence
  - b. Community School
  - c. Comprehensive
  - d. Danish Adopted School
  - e. Govt. School
  - f. Junior Model
  - g. M.C Local
  - h. Model School
  - i. Pilot Secondary
  - j. PSSP School
  - k. Technical High
25. **est\_year** – The year of establishment of the school.
26. **upgrade\_primary\_year** – The year in which the school level was upgraded to primary.
27. **upgrade\_middle\_year** – The year in which the school level was upgraded to middle.
28. **upgrade\_high\_year** – The year in which the school level was upgraded to high.
29. **upgrade\_high\_sec\_year** – The year in which the school level was upgraded to high secondary.
30. **bldg\_status** – The status of the school building (0, 1).
31. **school\_ownership** – Ownership of the school from one of the following:
- a. Building Provided By Local Residents
  - b. Education Department
  - c. Municipal Building

- d. On Rent
  - e. Property Of Any Other Institution Besides The Municipal Institution
  - f. Running In The Mosque
  - g. School Council provided building
  - h. Some Other Govt. School
32. **place\_status** – The status of the school place/location (0, 1).
33. **construct\_type** – Costruct type of the school (Completely Rough, Completely Solid, Partial Solid / Partial Rough).
34. **bldg\_condition** – The condition of the building from one of the following:
- a. Building Is Dangerous
  - b. Complete Building Needs Repairing
  - c. Needed Minor Repairing
  - d. Partial Building is Dangerous
  - e. Satisfying
35. **total\_area\_kanal** – Total area of the school in kanals.
36. **total\_area\_marla** – Total area of the school in marlas.
37. **covered\_area** – Covered area of the school.
38. **uncovered\_area\_kanal** – Uncovered area of the school in kanals.
39. **uncovered\_area\_marla** – Uncovered area of the school in marlas.
40. **functional\_classrooms** – Number of classrooms in the school that are functional.
41. **classes** – Total number of classes in the school.
42. **drink\_water** – Whether drinking water facility is available in the school or not (0, 1).
43. **drink\_water\_type** – Type of drinking water sources (Govt. Water, Hand Pump, Water Pump, Well, Other)
44. **drink\_water\_type\_other** – Other types of drinking water sources in the school.
45. **electricity** – Whether electricity is available or not available in the school (0, 1).
46. **electricity\_source** – Source of electricity in the school (Solar power, Wapda connection, Both)
47. **toilets** – Whether toilets are available or not available in the school (0, 1).

48. **total\_toilets** – Total number of toilets in the school.
49. **teachers\_toilets** – The number of teachers toilets in the school.
50. **boundary\_wall** – Whether a boundary wall is present or not around the school (0, 1).
51. **boundary\_wall\_state** – State of the boundary wall (Completed, Not Completed, Need Repairing)
52. **security** – Security provided to the school (Not Available, Satisfying, Not Satisfying)
53. **care\_giver** – Whether a caregiver is available or not available in the school (0, 1).
54. **enrollment** – The total number of students enrolled in the school.
55. **Teachers** – The number of teachers in the school.
56. **NonTeachers** – The number of non-teachers in the school.

## Data Cleaning

### 1. Missing Values

Checked and removed missing values in the data by filtering **school\_id** on 'Blanks'.

83 rows had missing **school\_id** and **emiscode**. They had wrong information **school\_name** (location information had been added in the 'school\_name' column. This seemed to be missing and erroneous data. Moreover, 83 out of 48,274 is just 0.17% of the total rows so it's appropriate to remove these rows.

### 2. Duplicate Values

Checked for duplicate values by using 'Remove Duplicates' in the Data tab.

There are NO duplicate values in the data.

### 3. Individual Column Values

- **moza** - Removed '-' from 8 entries and '-----' from 1 entry.
- **street\_name** - Removed '-' from 16 entries.
- **head\_name** - Replaced 'NULL' with Blank values in 172 entries.
- **head\_type** - Replaced 'NULL' with Blank values in 198 entries.
- **school\_type** - Replaced 'NA' with Blank values in 1507 entries.



- **upgrade\_primary\_year** – Replaced 'NULL' with Blank values in 11781 entries.
- **upgrade\_middle\_year** – Replaced 'NULL' with Blank values in 11780 entries.
- **upgrade\_high\_year** – Replaced 'NULL' with Blank values in 11780 entries.
- **upgrade\_high\_sec\_year** – Replaced 'NULL' with Blank values in 11781 entries.

#### 4. Rows with partially filled values

Removed 70 rows in data that have only school id and location information. They are only 0.15% of the data and cannot be used to do any analysis due to lack of information in them.

## Data Analysis and Insights

### 1. Total number of schools in the dataset

**PROCESS:** Each row in the dataset corresponds to one school. Counted the total number of rows in the dataset to find the total number of schools.

*Insight:*

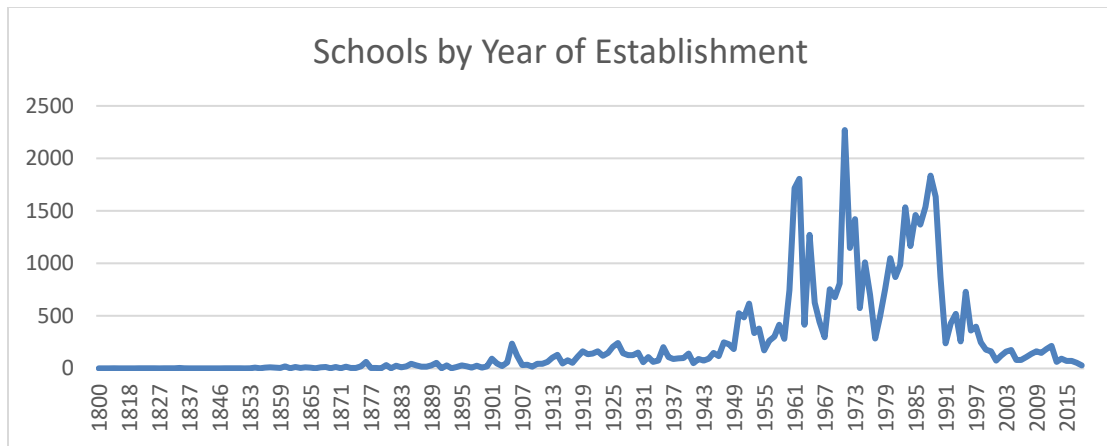
There are **48121 schools** in the dataset.

### 2. Schools by Year of Establishment

**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Rows:** est\_year

**Columns:** Count of school\_id



#### *Insight:*

- The highest number of schools have been established in the 60's, 70's and 80's with the top number being 2270 schools in 1971.

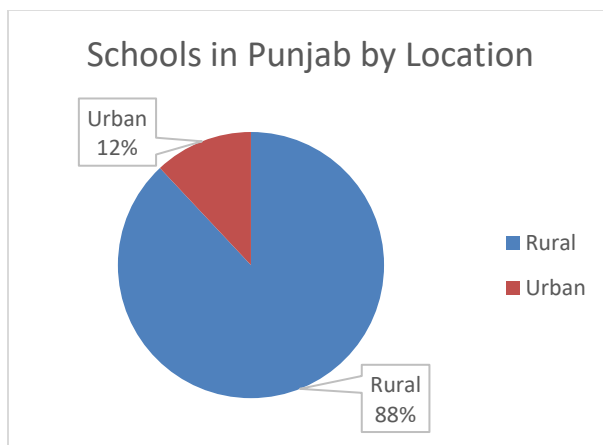
### 3. Number of schools in Rural and Urban areas

**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Rows:** school\_location

**Columns:** Count of school\_id

Row Labels	Count of school_id
Rural	42350
Urban	5771
<b>Grand Total</b>	<b>48121</b>



### Insights:

- 88% of the schools are located in Rural areas.
- 12% of the schools are located in Urban areas.

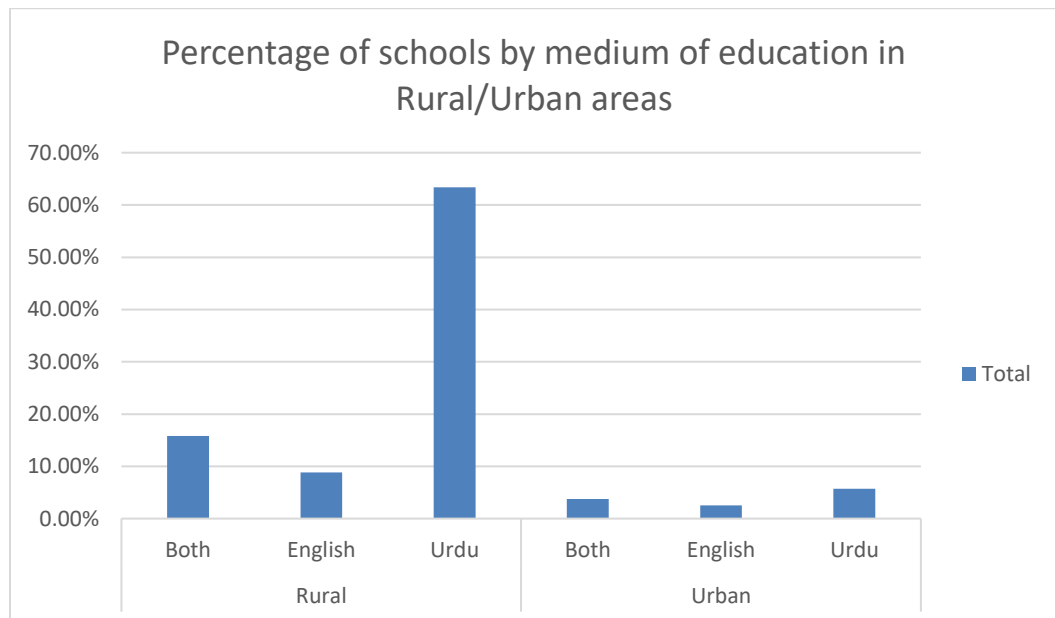
## 4. Percentage of schools offering Urdu and English medium education or both in Rural and Urban areas

**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Rows:** school\_level, medium

**Values:** Count of school\_id

Count of school_id		Total
school_location	medium	
Rural	Both	15.79%
	English	8.86%
	Urdu	63.36%
Urban	Both	3.76%
	English	2.52%
	Urdu	5.71%
Grand Total		100.00%



### Insights:

- The highest percentage (almost 63%) of schools provides Urdu medium education in rural areas.
- The lowest percentage (almost 2.5%) of schools provide English medium education in Urban areas.

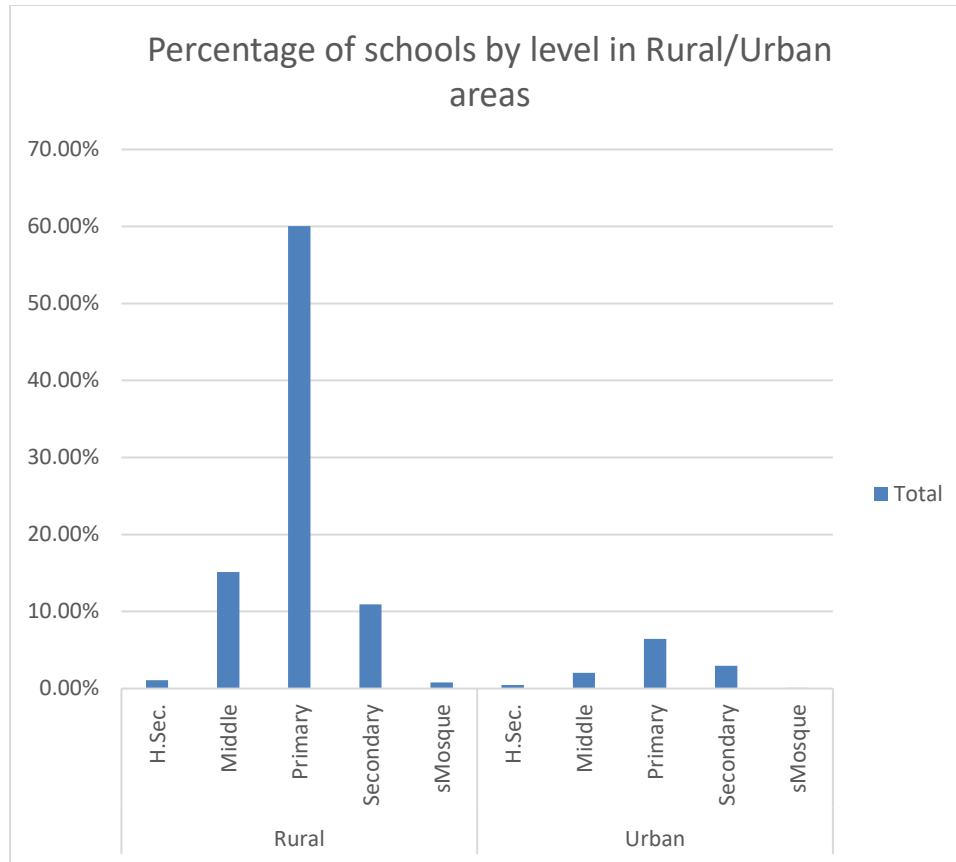
## 5. Percentage of schools by level in Rural and Urban areas

**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Rows:** school\_location, school\_level

**Values:** Count of school\_id

Row Labels	Count of school_id
<b>Rural</b>	<b>88.01%</b>
H.Sec.	1.10%
Middle	15.15%
Primary	60.03%
Secondary	10.92%
sMosque	0.81%
<b>Urban</b>	<b>11.99%</b>
H.Sec.	0.46%
Middle	2.06%
Primary	6.45%
Secondary	2.94%
sMosque	0.09%
<b>Grand Total</b>	<b>100.00%</b>



### Insights:

- Primary schools have the highest percentage in both rural and urban areas.
- sMosque and H.Sec. schools have the lowest percentage in both rural and urban areas.

## 6. Distribution of schools according to school level

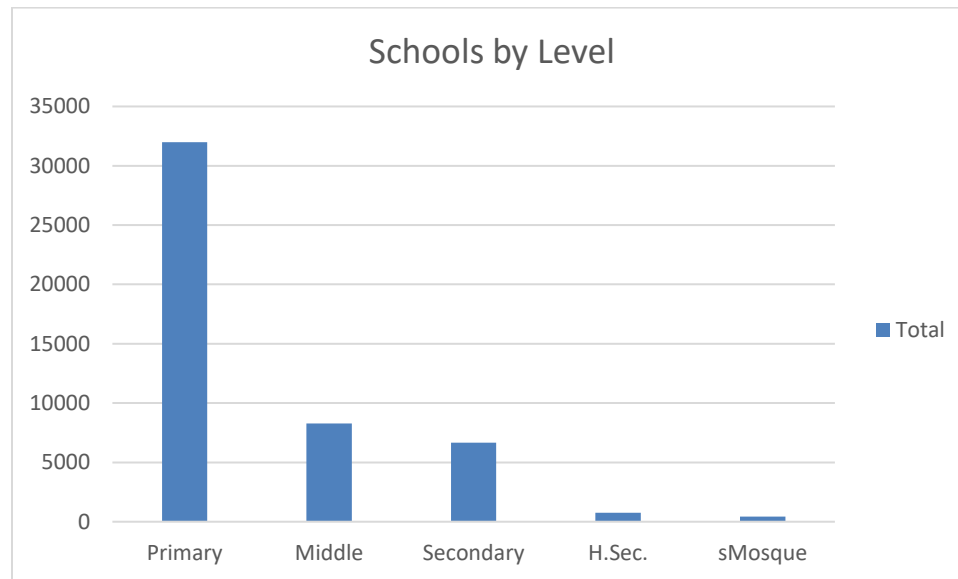
**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Rows:** school\_level

**Values:** Count of school\_id

**Sort** Values from Z to A by Count of school\_id

Row Labels	Count of school_id
Primary	31994
Middle	8277
Secondary	6668
H.Sec.	749
sMosque	433
<b>Grand Total</b>	<b>48121</b>



### Insights:

- The highest number of schools are Primary followed by Middle and then Secondary.
- The lowest number of schools are 'sMosque'.

## 7. Number of schools categorized by gender (Male & Female)

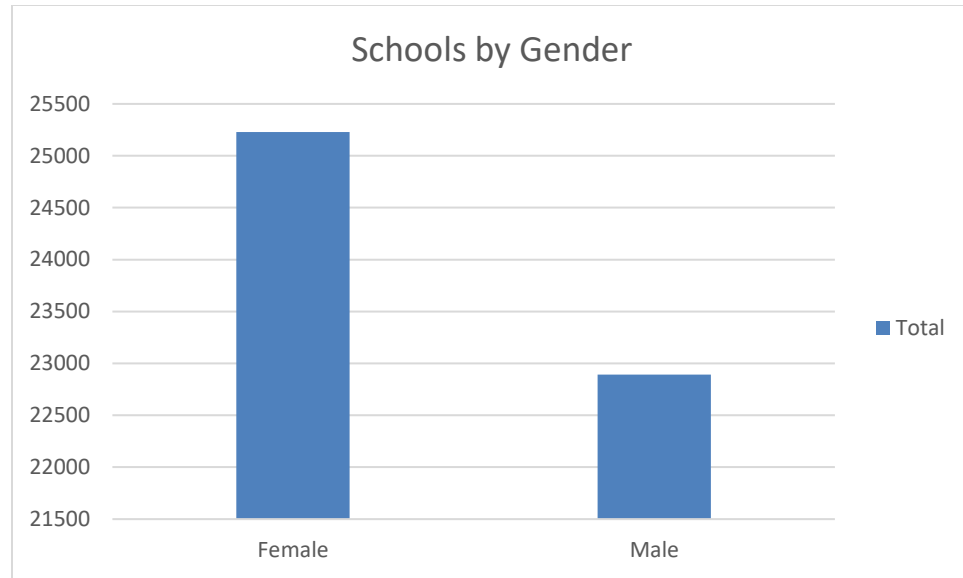
**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Rows:** school\_gender

**Values:** Count of school\_id

**Sort** Values from Z to A by Count of school\_id

Row Labels	Count of school_id
Female	25230
Male	22891
<b>Grand Total</b>	<b>48121</b>



*Insight:*

- There are 2339 more Female schools than Male schools in Punjab.

## 8. School level and medium with the lowest number of functional classrooms

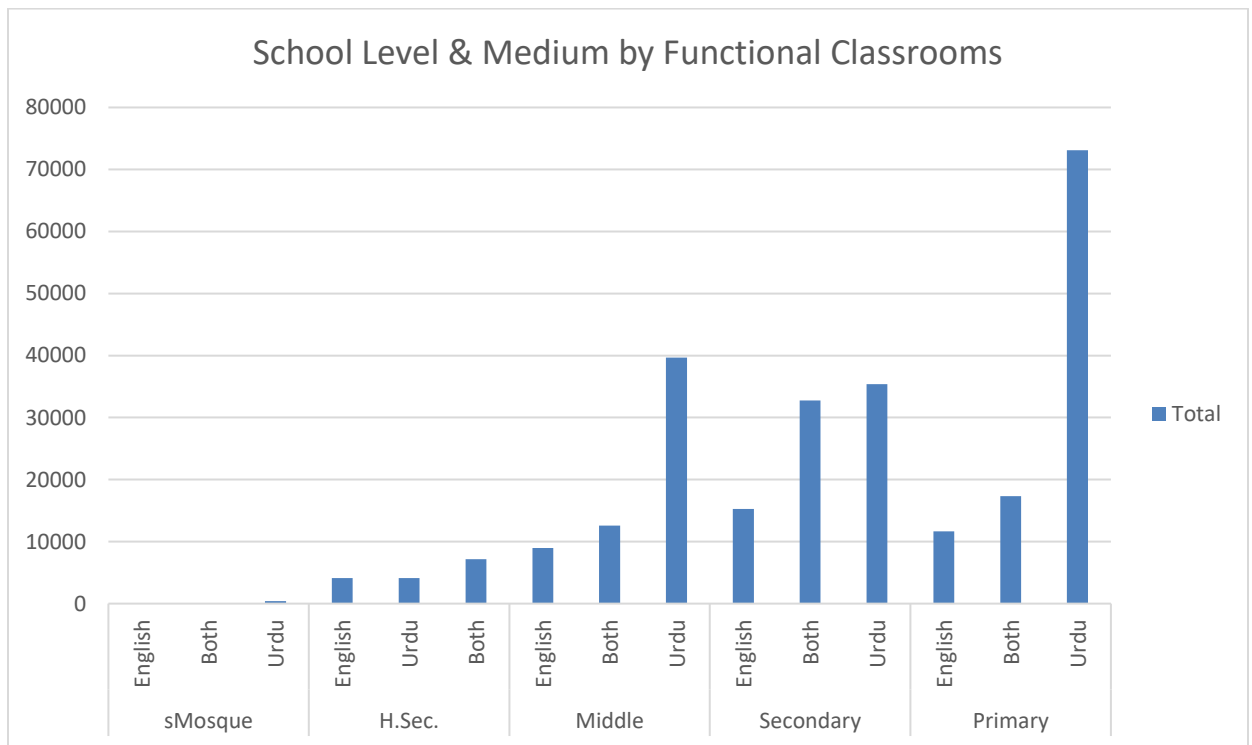
**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Rows:** school\_level, medium

**Values:** Sum of functional\_classrooms

**Sort** Values by school\_level and medium by Sum of functional\_classrooms

school_level	medium	Total
sMosque	English	29
sMosque	Both	80
sMosque	Urdu	403
H.Sec.	English	4114
H.Sec.	Urdu	4133
H.Sec.	Both	7149
Middle	English	8964
Middle	Both	12579
Middle	Urdu	39649
Secondary	English	15276
Secondary	Both	32748
Secondary	Urdu	35374
Primary	English	11671
Primary	Both	17343
Primary	Urdu	73111
<b>Grand Total</b>		<b>262623</b>



#### *Insights:*

- The school level with the lowest number of functional classrooms is 'sMosque' followed by 'H.Sec.'



- Within 'sMosque' and 'H.Sec.', English medium schools have the lowest number of functional classrooms.

## 9. Schools with respect to school ownership

Row Labels	Count of school_id
Education Department	45607
Building Provided By Local Residents	1028
Municipal Building	537
On Rent	275
Some Other Govt. School	183
Property Of Any Other Institution Besides The Municipal Institution	148
Running In The Mosque	132
School Council provided building	128
<b>Grand Total</b>	<b>48121</b>

### Insight:

- The Education Department owns the highest number of schools i.e. 45607.

## 10. Percentage of schools upgraded

**PROCESS:** Calculated each percentage by counting the number of upgraded schools in each column

Upgrade_Year	Percent_schools_upgraded
upgrade_primary_year	30.47
upgrade_middle_year	31.13
upgrade_high_year	15.35
upgrade_high_sec_year	1.60

### Insights:

- Almost 30% of the schools have been upgraded to Primary level.
- Almost 31% of the schools have been upgraded to Middle level.
- Almost 15% of the schools have been upgraded to High level.
- Almost 1.6% of the schools have been upgraded to Higher Secondary level.

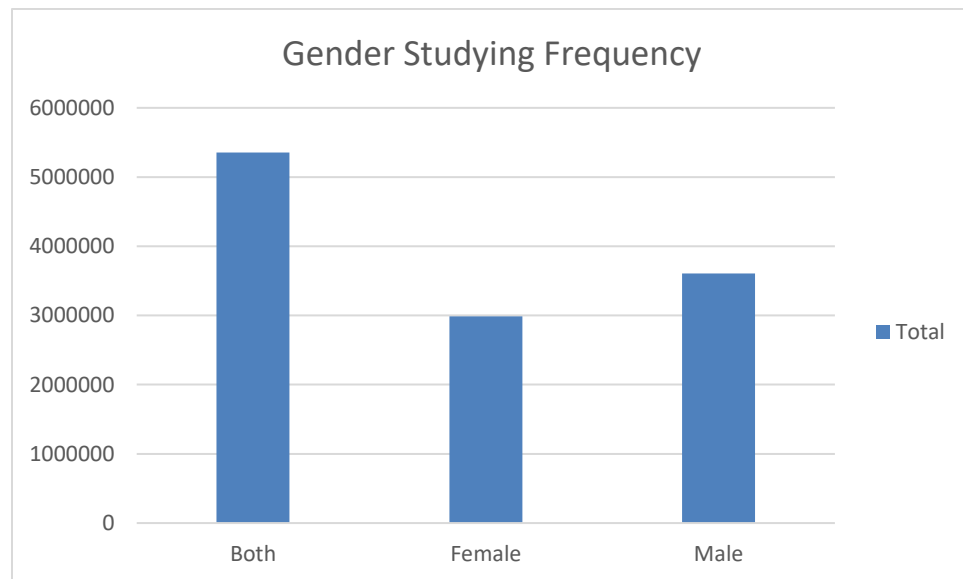
## 11. Number of Male and Female students

**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Rows:** gender\_studying

**Values:** Sum of enrollment

Row Labels	Sum of enrollment
Both	5354346
Female	2986045
Male	3607833
<b>Grand Total</b>	<b>11948224</b>



### *Insights:*

- The number of Male students enrolled are more than the number of Female students enrolled overall in schools where only one gender is studying.
- The number of enrollments in schools where both genders are studying is more than in gender-specific schools but we cannot tell about the number of males and females separately in those schools using this data.

## 12. Lowest number of Male & Female students by school level of schools established in the 2000s

**PROCESS:** Generate a PIVOT TABLE on the dataset.

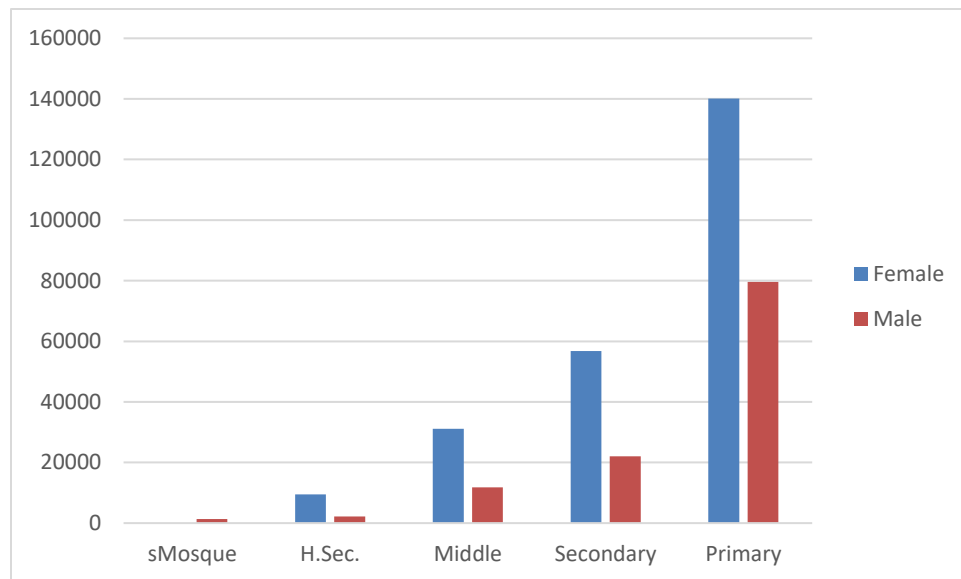
**Rows:** school\_level

**Columns:** school\_gender

**Values:** Sum of enrollment

**Sort** Values from A to Z by Sum of enrollment

Sum of enrollment Column Labels ▼			
Row Labels ▼	Female	Male	Grand Total
sMosque		1349	1349
H.Sec.	9467	2157	11624
Middle	31114	11789	42903
Secondary	56831	22086	78917
Primary	140064	79628	219692
<b>Grand Total</b>	<b>237476</b>	<b>117009</b>	<b>354485</b>



### *Insights:*

- The lowest students by gender have been enrolled in school level 'sMosque' which are Males (1349) followed by Males (2157) in 'H.Sec.' school for the schools established in 2000s.

- No females have been enrolled in 'sMosque' since 2000s.
- More females have been enrolled than males in all school levels other than 'sMosque' for the schools established in 2000s.
- The highest number of students (Males and Females) have been enrolled in Primary schools followed by Secondary, Middle and Higher Secondary schools since 2000s.

### 13. School level and medium by students enrolled

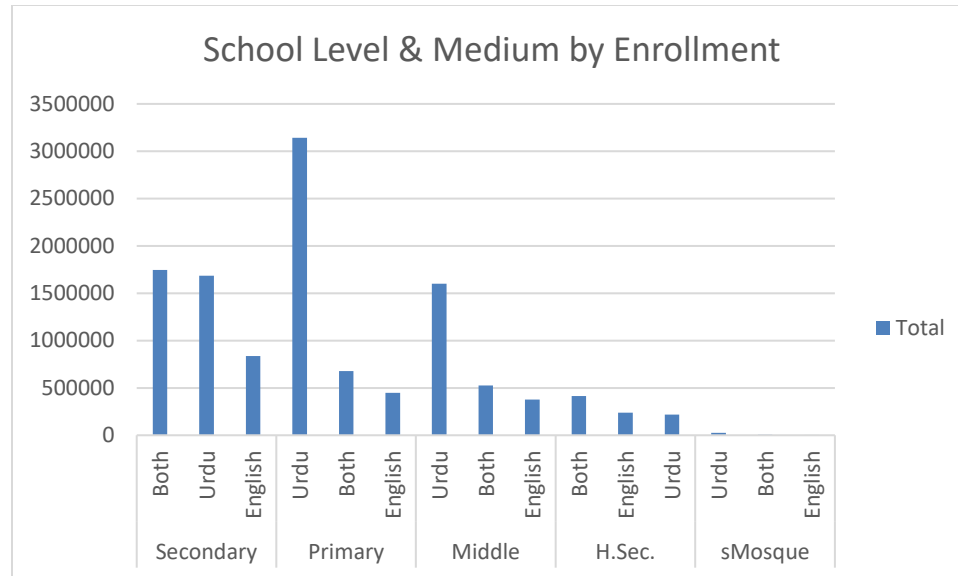
**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Rows:** school\_level, medium

**Values:** Sum of enrollment

**Sort** Values by school\_level and medium by Sum of enrollment

Sum of enrollment		Total
school_level	medium	
Secondary	Both	1746832
Secondary	Urdu	1686526
Secondary	English	838222
Primary	Urdu	3141024
Primary	Both	676559
Primary	English	449550
Middle	Urdu	1601941
Middle	Both	526540
Middle	English	378147
H.Sec.	Both	415456
H.Sec.	English	237170
H.Sec.	Urdu	217379
sMosque	Urdu	25780
sMosque	Both	5347
sMosque	English	1751
Grand Total		11948224



### Insights:

- The highest number of enrollments is in Secondary schools followed by Primary schools.
- The lowest number of enrollments is in sMosque level schools followed by H.Sec. schools.
- The highest number of enrollments is in Primary Urdu medium schools.
- Enrollments in English medium schools are the lowest in all school levels.

## 14. District with the highest enrollment in Primary schools since their establishment


**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Rows:** district

**Values:** Sum of enrollment

**Filter** school\_level by Primary

**Sort** Values from Z to A by Sum of enrollment

school_level	Primary	
<b>Row Labels</b>		<b>Sum of enrollment</b>
RAHIMYAR KHAN		257799
FAISALABAD		234888
BAHAWALNAGAR		188760
MUZAFFARGARH		184840
D.G. KHAN		176390
SARGODHA		160621
JHANG		153896
KASUR		151602
BAHAWALPUR		136240
MULTAN		136061
SIALKOT		134986
VEHARI		132077
OKARA		131889
KHANEWAL		130888
GUJRANWALA		129561
LAHORE		120464
LAYYAH		120390
T.T.SINGH		115485
GUJRAT		108364
PAKPATTAN		108198
BHAKKAR		102846
SHEIKHUPURA		98575
RAJANPUR		98489
RAWALPINDI		96860
NAROWAL		94505
SAHIWAL		93598
MIANWALI		89499
CHINIOT		87974
NANKANA SAHIB		73190
ATTOCK		72057
MANDI BAHA UD DIN		69993
LODHRAN		64909
KHUSHAB		64148
HAFIZABAD		60965
JHELUM		44058
CHAKWAL		42068
<b>Grand Total</b>		<b>4267133</b>

### *Insights:*

- District with the highest enrollment in Primary schools since their establishment is Rahimyar Khan with 257799 students enrolled.

## 15. District with the highest enrollment in Secondary schools since their establishment

**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Rows:** district

**Values:** Sum of enrollment

**Filter** school\_level by Secondary

**Sort** Values from Z to A by Sum of enrollment

school_level	Secondary	
<b>Row Labels</b>		<b>Sum of enrollment</b>
FAISALABAD		362868
LAHORE		353382
GUJRANWALA		186595
RAWALPINDI		180642
SARGODHA		173756
SIALKOT		165435
GUJRAT		157207
RAHIMYAR KHAN		153560
OKARA		143249
KASUR		141328
KHANEWAL		135106
SHEIKHUPURA		125310
MULTAN		124820
T.T.SINGH		122685
SAHIWAL		119178
BAHAWALNAGAR		110309
JHANG		109509
ATTOCK		101752
VEHARI		99163
BAHAWALPUR		99065
MANDI BAHU UD DIN		96709
CHAKWAL		94396
NAROWAL		90675
LAYYAH		86581
MUZAFFARGARH		82843
MIANWALI		76161
D.G. KHAN		75440
JHELUM		74328
BHAKKAR		66668
PAKPATTAN		65486
NANKANA SAHIB		57784
CHINIOT		56857
KHUSHAB		53342
LODHRAN		47296
HAFIZABAD		42826
RAJANPUR		39269
<b>Grand Total</b>		<b>4271580</b>



### *Insights:*

- District with the highest enrollment in Secondary schools since their establishment is Faisalabad with 362868 students enrolled.

## 16. District with the highest enrollment in Higher Secondary schools since their establishment


**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Rows:** district

**Values:** Sum of enrollment

**Filter** school\_level by H.Sec.

**Sort** Values from Z to A by Sum of enrollment

school_level	H.Sec.	
<b>Row Labels</b>		<b>Sum of enrollment</b>
FAISALABAD		99227
LAHORE		50918
SIALKOT		40423
MULTAN		38429
SARGODHA		36088
SAHIWAL		35339
RAWALPINDI		33499
RAHIMYAR KHAN		31133
SHEIKHUPURA		30830
BAHAWALPUR		28322
KHANEWAL		28009
VEHARI		27048
GUJRANWALA		25971
BAHAWALNAGAR		22815
GUJRAT		22618
JHANG		22285
D.G. KHAN		21770
KASUR		20300
NAROWAL		19752
HAFIZABAD		19179
T.T.SINGH		18219
ATTOCK		17879
MUZAFFARGARH		16695
RAJANPUR		16034
PAKPATTAN		15474
CHAKWAL		15162
MANDI BAHU UD DIN		14492
NANKANA SAHIB		14021
JHELMUM		13932
OKARA		12891
LODHRAN		12786
KHUSHAB		11998
MIANWALI		11274
LAYYAH		10050
BHAKKAR		8093
CHINIOT		7050
<b>Grand Total</b>		<b>870005</b>

### *Insights:*

- District with the highest enrollment in Higher Secondary schools since their establishment is Faisalabad with 99227 students enrolled.

## 17. Top 5 Districts with the highest number of teachers and non-teachers in schools

**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Rows:** district

**Values:** Sum of Teachers | Sum of NonTeachers

**Sort** Values from Z to A by Sum of Teachers | Sum of NonTeachers

Row Labels	Sum of Teachers
FAISALABAD	22996
RAHIMYAR KHAN	17601
RAWALPINDI	17112
LAHORE	16654
SARGODHA	15798
SIALKOT	15267
BAHAWALNAGAR	14560
GUJRANWALA	13853
BAHAWALPUR	12612
GUJRAT	12324
MUZAFFARGARH	12320
MULTAN	11882
OKARA	11734
KASUR	11637
KHANEWAL	11519
JHANG	11037
VEHARI	10843
SAHIWAL	10699
T.T.SINGH	10407
SHEIKHUPURA	10164
D.G. KHAN	9914
NAROWAL	9878
ATTOCK	9705
LAYYAH	9618
CHAKWAL	9208
MIANWALI	8510
BHAKKAR	8085
MANDI BAHA UD DIN	6814
JHELUM	6785
KHUSHAB	6463
NANKANA SAHIB	6397
RAJANPUR	6282
PAKPATTAN	6208
LODHRAN	5958
HAFIZABAD	5093
CHINIOT	4826
<b>Grand Total</b>	<b>390763</b>

Row Labels	Sum of NonTeachers
FAISALABAD	4918
RAWALPINDI	3438
SARGODHA	3418
LAHORE	3128
BAHAWALNAGAR	3015
RAHIMYAR KHAN	3006
SIALKOT	2713
ATTOCK	2591
KHANEWAL	2584
MULTAN	2416
GUJRAT	2413
KASUR	2313
OKARA	2313
JHANG	2288
SAHIWAL	2216
BAHAWALPUR	2182
GUJRANWALA	2168
T.T.SINGH	1983
SHEIKHUPURA	1975
MUZAFFARGARH	1973
VEHARI	1783
LAYYAH	1694
NAROWAL	1656
D.G. KHAN	1642
MANDI BAHA UD DIN	1406
BHAKKAR	1340
CHAKWAL	1316
JHELUM	1313
MIANWALI	1300
LODHRAN	1207
NANKANA SAHIB	1141
KHUSHAB	1093
RAJANPUR	910
PAKPATTAN	869
CHINIOT	836
HAFIZABAD	809
<b>Grand Total</b>	<b>73366</b>

### Insights:

- Top 5 districts with the highest number of teachers are:

1. FAISALABAD

2. RAHIMYAR KHAN
  3. RAWALPINDI
  4. LAHORE
  5. SARGODHA
- Top 5 districts with the highest number of non-teachers are:
    1. FAISALABAD
    2. RAWALPINDI
    3. SARGODHA
    4. LAHORE
    5. BAHAWALNAGAR
  - Districts with the lowest number of teachers are CHINIOT, HAFIZABAD, LODHRAN, PAKPATTAN and RAJANPUR.
  - Districts with the lowest number of non-teachers are HAFIZABAD, CHINIOT, PAKPATTAN, RAJANPUR and KHUSHAB.

## 18. Distribution of teachers across different schools

### a. Teachers by School Location and Level

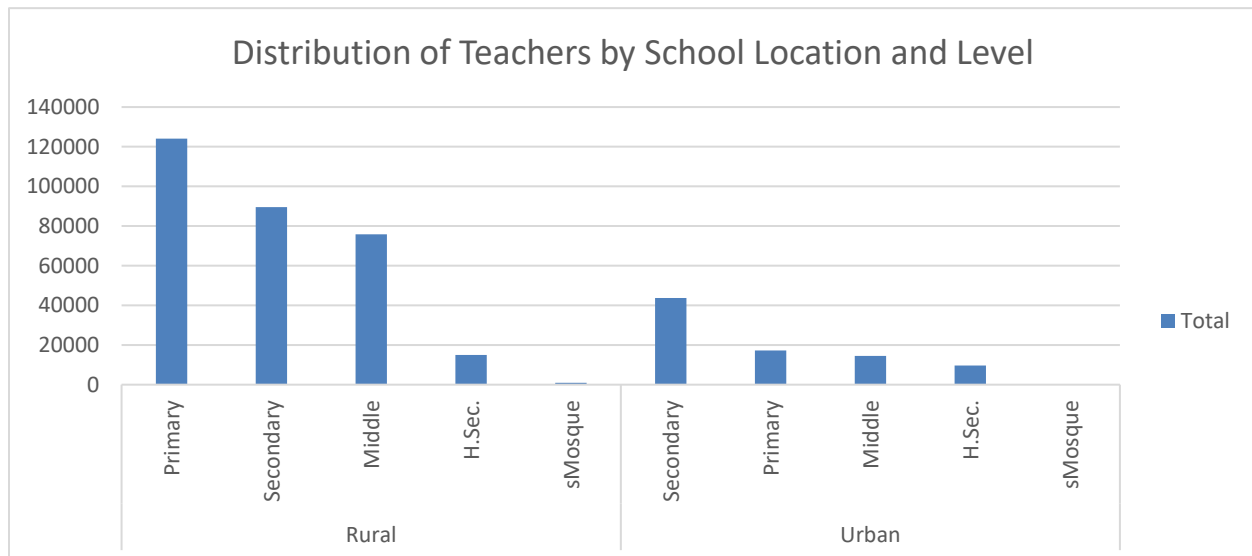
**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Rows:** school\_location, school\_level

**Values:** Sum of Teachers

**Sort** school\_location and school\_level from Z to A by Sum of Teachers

Sum of Teachers:		
school_locati	school_level	Total
Rural	Primary	124121
Rural	Secondary	89536
Rural	Middle	75887
Rural	H.Sec.	15035
Rural	sMosque	902
Rural Total		305481
Urban	Secondary	43665
Urban	Primary	17257
Urban	Middle	14552
Urban	H.Sec.	9707
Urban	sMosque	101
Urban Total		85282
Grand Total		390763



### Insights:

- There are more teachers in rural schools than urban schools – rural schools have almost 3.5 times more teachers than urban schools overall.
- Rural schools have highest number of teachers at the Primary level.
- Urban schools have highest number of teachers at the secondary level.

### b. Teachers by School Level and Medium

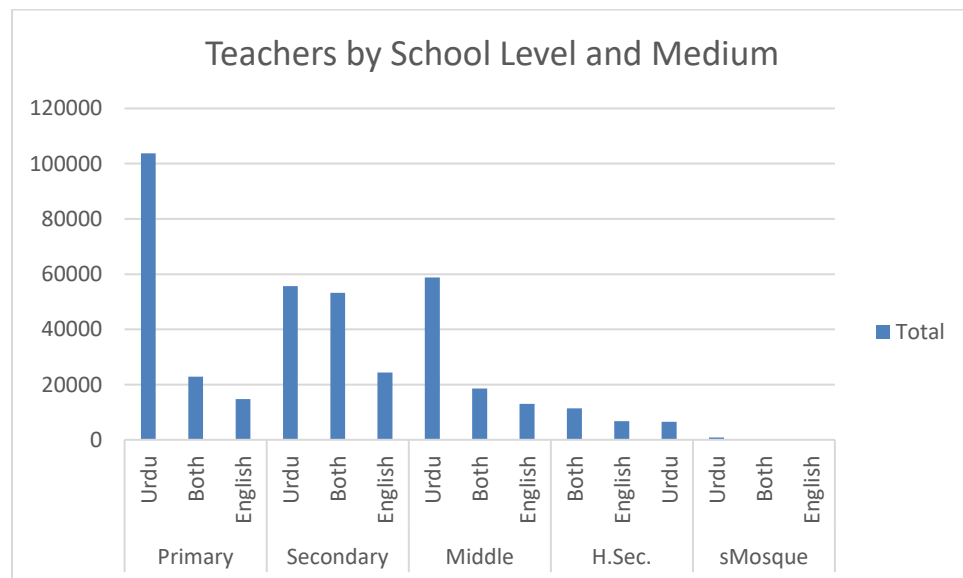
**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Rows:** school\_level, medium

**Values:** Sum of Teachers

**Sort** school\_level and medium from Z to A by Sum of Teachers

Sum of Teachers		
school_level	medium	Total
Primary	Urdu	103759
Primary	Both	22867
Primary	English	14752
Secondary	Urdu	55652
Secondary	Both	53181
Secondary	English	24368
Middle	Urdu	58765
Middle	Both	18610
Middle	English	13064
H.Sec.	Both	11440
H.Sec.	English	6759
H.Sec.	Urdu	6543
sMosque	Urdu	787
sMosque	Both	165
sMosque	English	51
Grand Total		390763



### Insights:

- Urdu medium Primary schools have the highest number of teachers.

- Primary schools have overall highest number of teachers followed by Secondary and then Middle schools.
- sMosque has the lowest number of teachers followed by H.Sec. schools.

### c. Teachers by School Type

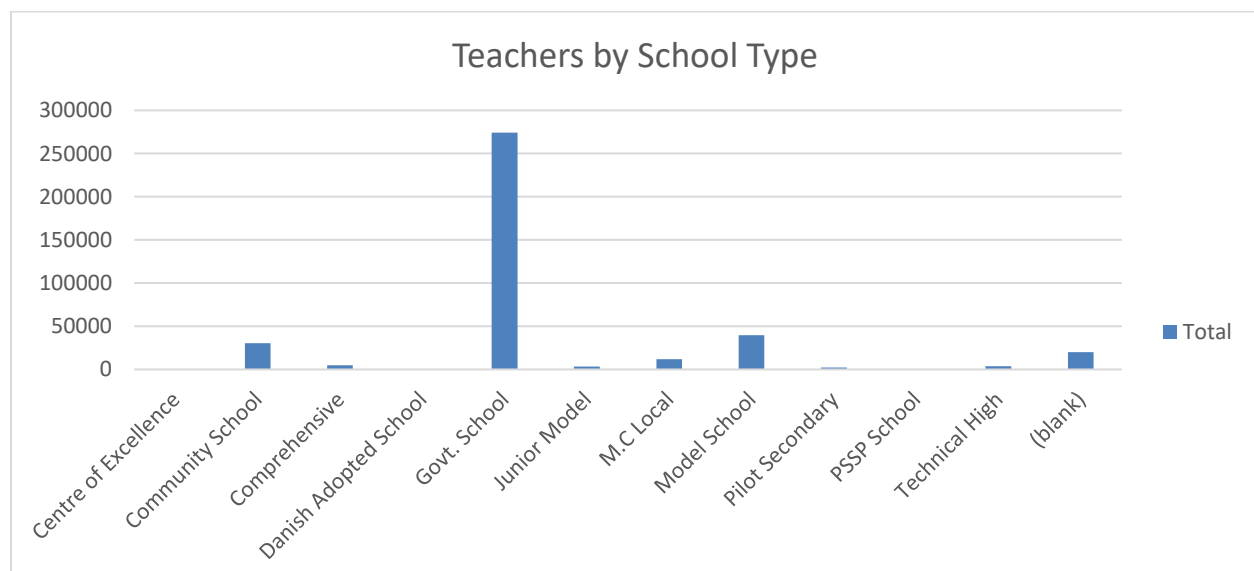
**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Rows:** school\_level, medium

**Values:** Sum of Teachers

**Sort** school\_level and medium from Z to A by Sum of Teachers

Row Labels	Sum of Teachers
Centre of Excellence	338
Community School	30285
Comprehensive	4704
Danish Adopted School	191
Govt. School	274379
Junior Model	3211
M.C Local	11837
Model School	39546
Pilot Secondary	2304
PSSP School	7
Technical High	3774
(blank)	20187
<b>Grand Total</b>	<b>390763</b>





*Insight:*

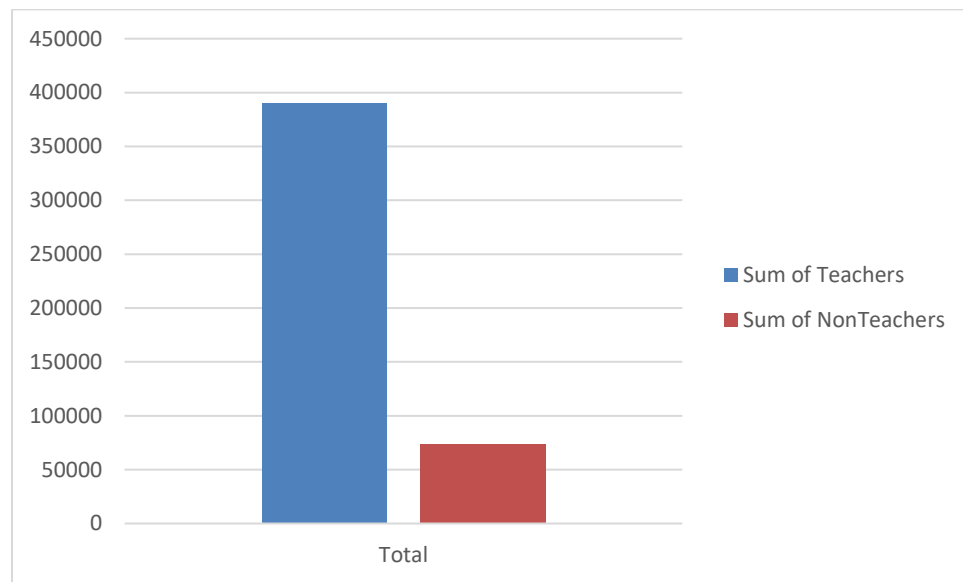
- Government schools have the highest number of teachers.

## 19. Distribution of teachers based on their respective job posts

**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Values:** Sum of Teachers, Sum of NonTeachers

Sum of Teachers	Sum of NonTeachers
390763	73366



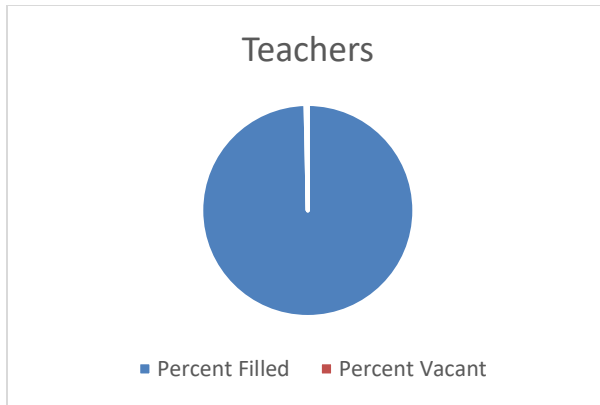
*Insight:*

- There are 390763 teachers and 73366 non-teachers in schools in Punjab.

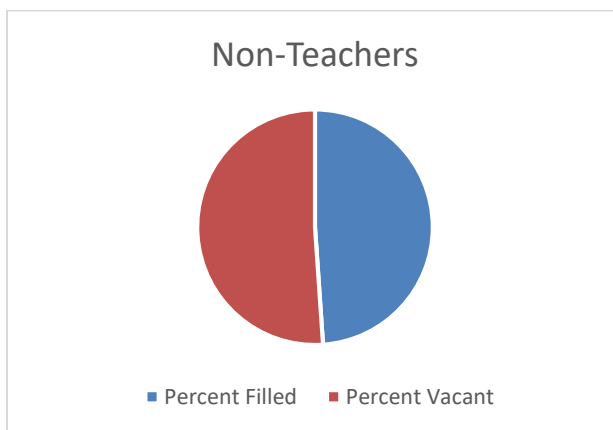
## 20. Percentage of vacant and filled teaching and non-teaching posts

**PROCESS:** Blank values in data have been considered as vacant posts and filled values have been considered as filled posts and their count has been used to calculate the percentages.

Post	Percent Filled	Percent Vacant
Teachers	99.67	0.33



Post	Percent Filled	Percent Vacant
NonTeachers	48.91	51.09



### *Insights:*

- 99.67% schools have teaching posts filled and 0.33% schools have teaching posts vacant.
- 48.91% schools have non-teaching posts filled and 51.09% schools have non-teaching posts vacant.

## 21. Percentage of schools with satisfactory building conditions

**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Rows:** bldg\_condition

**Values:** Count of school\_id -> Percentage of Column Total

Row Labels	Count of school_id
	0.50%
Building Is Dangerous	0.95%
Complete Building Needs Repairing	3.08%
Needed Minor Repairing	16.61%
Partial Building is Dangerous	3.92%
Satisfying	74.95%
<b>Grand Total</b>	<b>100.00%</b>

### Insights:

- Almost 75% of the schools have satisfactory building conditions.
- Almost 1% of the school buildings have been declared dangerous.

## 22. Percentage of schools based on their construct type in rural and urban areas

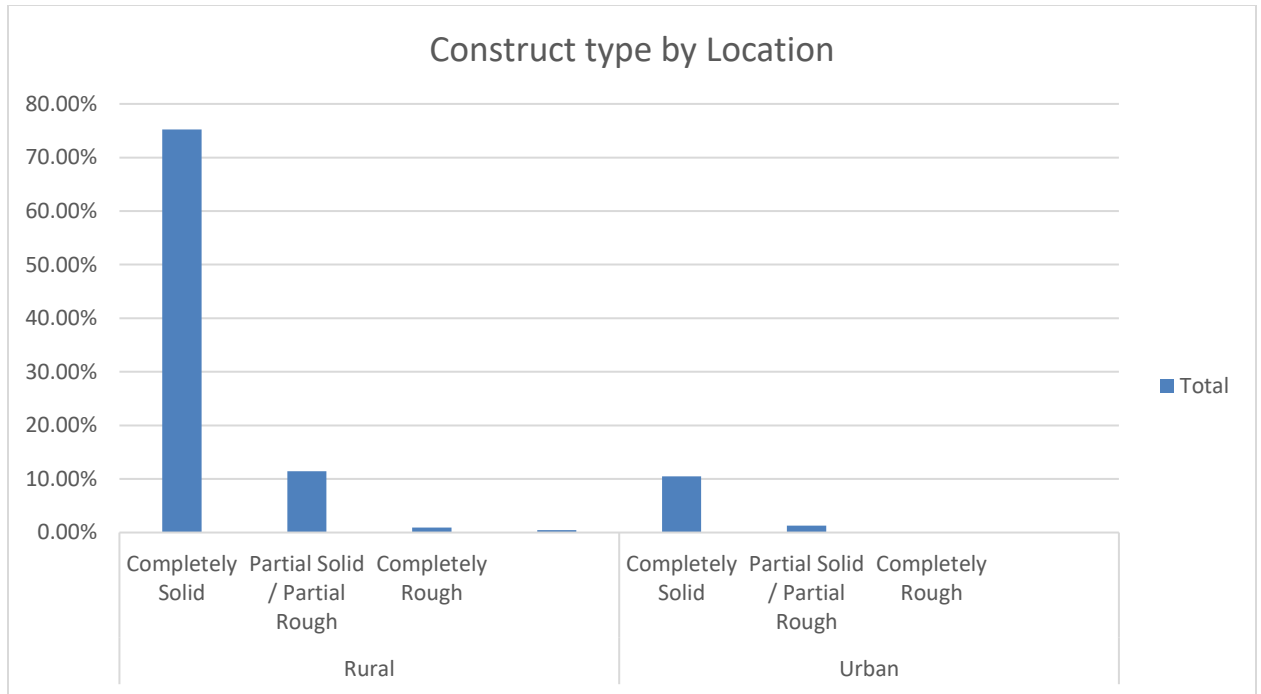
**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Rows:** school\_location, construct\_type

**Values:** Count of school\_id -> Percentage of Column Total

**Sort** Values from A to Z by Count of school\_id

Count of school_id		
school_location	construct_type	Total
[-] Rural	Completely Solid	75.21%
Rural	Partial Solid / Partial Rough	11.44%
Rural	Completely Rough	0.91%
Rural		0.44%
<b>Rural Total</b>		<b>88.01%</b>
[-] Urban	Completely Solid	10.51%
Urban	Partial Solid / Partial Rough	1.32%
Urban	Completely Rough	0.11%
Urban		0.06%
<b>Urban Total</b>		<b>11.99%</b>
<b>Grand Total</b>		<b>100.00%</b>



### *Insights:*

- Almost 13% of school buildings in rural areas have been constructed roughly or in a partially solid way.
- Almost 1.5% of school buildings in urban areas have been constructed roughly or in a partially solid way.

## 23. Names of 5 Districts with the lowest availability of drinking water facilities in schools

**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Rows:** district

**Values:** Sum of drink\_water

**Sort** Values from A to Z by Sum of drink\_water

Row Labels	Sum of drink_water
CHINIOT	694
HAFIZABAD	730
NANKANA SAHIB	748
LODHRAN	754
MANDI BAHA UD DIN	770
JHELUM	784
PAKPATTAN	857
RAJANPUR	928
KHUSHAB	950
NAROWAL	1094
CHAKWAL	1135
LAHORE	1137
SAHIWAL	1155
SHEIKHUPURA	1159
T.T.SINGH	1179
ATTOCK	1199
MIANWALI	1212
KHANEWAL	1237
BHAKKAR	1260
MULTAN	1309
KASUR	1358
GUJRAT	1385
OKARA	1400
VEHARI	1419
JHANG	1475
LAYYAH	1513
D.G. KHAN	1532
GUJRANWALA	1543
BAHAWALPUR	1651
MUZAFFARGARH	1751
RAWALPINDI	1775
SIALKOT	1790
SARGODHA	1921
BAHAWALNAGAR	2120
FAISALABAD	2205
RAHIMYAR KHAN	2759
<b>Grand Total</b>	<b>47888</b>

### *Insights:*

- Districts with the lowest availability of drinking water facilities in schools are:

- CHINIOT

2. HAFIZABAD
3. NANKANA SAHIB
4. LODHRAN
5. MANDI BAHA UD DIN

24. Names of 5 Districts with the lowest availability of electricity in schools.

**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Rows:** district

**Values:** Count of electricity

**Sort** Values from A to Z by Count of electricity

Row Labels	Count of electricity
CHINIOT	697
HAFIZABAD	732
NANKANA SAHIB	748
LODHRAN	755
MANDI BAHA UD DIN	773
JHELUM	788
PAKPATTAN	858
RAJANPUR	950
KHUSHAB	951
NAROWAL	1095
CHAKWAL	1137
LAHORE	1137
SAHIWAL	1156
SHEIKHUPURA	1160
T.T.SINGH	1180
ATTOCK	1203
MIANWALI	1228
KHANEWAL	1237
BHAKKAR	1260
MULTAN	1312
KASUR	1358
GUJRAT	1393
OKARA	1402
VEHARI	1420
JHANG	1476
LAYYAH	1513
GUJRANWALA	1549
D.G. KHAN	1602
BAHAWALPUR	1657
MUZAFFARGARH	1754
SIALKOT	1794
RAWALPINDI	1803
SARGODHA	1925
BAHAWALNAGAR	2140
FAISALABAD	2208
RAHIMYAR KHAN	2770
<b>Grand Total</b>	<b>48121</b>

### *Insights:*

- Districts with the lowest availability of electricity in schools are:

- CHINIOT

2. HAFIZABAD
3. NANKANA SAHIB
4. LODHRAN
5. MANDI BAHA UD DIN

25. Names of 5 Districts with the lowest presence of boundary walls in schools

**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Rows:** district

**Values:** Sum of boundary\_wall

**Sort** Values from A to Z by Sum of boundary\_wall



Row Labels	Sum of boundary_wall
CHINIOT	690
HAFIZABAD	718
NANKANA SAHIB	738
LODHRAN	755
MANDI BAHA UD DIN	765
JHELMUM	786
PAKPATTAN	847
RAJANPUR	929
KHUSHAB	942
SAHIWAL	1059
NAROWAL	1072
T.T.SINGH	1081
LAHORE	1126
CHAKWAL	1132
SHEIKHUPURA	1150
ATTOCK	1193
MIANWALI	1216
KHANEWAL	1216
BHAKKAR	1257
MULTAN	1301
OKARA	1336
JHANG	1353
KASUR	1353
GUJRAT	1384
VEHARI	1416
LAYYAH	1497
GUJRANWALA	1531
D.G. KHAN	1559
BAHAWALPUR	1573
RAWALPINDI	1628
MUZAFFARGARH	1743
SIALKOT	1765
SARGODHA	1893
BAHAWALNAGAR	1957
FAISALABAD	2139
RAHIMYAR KHAN	2730
<b>Grand Total</b>	<b>46830</b>

### *Insights:*

- Districts with the lowest presence of boundary walls in schools are:

- CHINIOT

2. HAFIZABAD
3. NANKANA SAHIB
4. LODHRAN
5. MANDI BAHA UD DIN

26. Names of 5 Districts with the lowest availability of toilets in schools

**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Rows:** district

**Values:** Sum of toilets

**Sort** Values from A to Z by Sum of toilets

Row Labels	Sum of toilets
CHINIOT	696
HAFIZABAD	729
NANKANA SAHIB	745
LODHRAN	755
MANDI BAHA UD DIN	772
JHELUM	788
PAKPATTAN	858
RAJANPUR	934
KHUSHAB	951
NAROWAL	1095
CHAKWAL	1136
LAHORE	1137
SAHIWAL	1155
SHEIKHUPURA	1158
T.T.SINGH	1180
ATTOCK	1201
MIANWALI	1227
KHANEWAL	1236
BHAKKAR	1260
MULTAN	1308
KASUR	1358
GUJRAT	1391
OKARA	1396
VEHARI	1420
JHANG	1476
LAYYAH	1511
GUJRANWALA	1548
D.G. KHAN	1585
BAHAWALPUR	1653
MUZAFFARGARH	1746
SIALKOT	1792
RAWALPINDI	1796
SARGODHA	1921
BAHAWALNAGAR	2096
FAISALABAD	2206
RAHIMYAR KHAN	2753
<b>Grand Total</b>	<b>47969</b>

### *Insights:*

- Districts with the lowest availability of toilets in schools are:

- CHINIOT

2. HAFIZABAD
3. NANKANA SAHIB
4. LODHRAN
5. MANDI BAHA UD DIN

## 27. Number of schools with satisfactory security measures

**PROCESS:** Generate a PIVOT TABLE on the dataset.

**Rows:** security

**Values:** Count of school\_id -> Percentage of Column Total

Row Labels	Count of school_id
Available	0.00%
Not Available	23.32%
Not Satisfying	6.39%
Satisfying	70.29%
<b>Grand Total</b>	<b>100.00%</b>

*Insight:*

- Almost 70% of the schools have satisfactory security measures.

## Summary of Insights

This dataset provides data about schools in Punjab, Pakistan. There are 48121 schools present in the dataset. After cleaning and analyzing the data in Excel and generating pivot tables and visualizations, it is observed that schools in Pakistan face a variety of problems related to upgradation, infrastructure, facilities, staff and security.

Following insights have been drawn after analyzing the data in detail:

### School Level & Upgradation

1. There are **48121 schools** in the dataset.
2. The highest number of schools have been established in the 60's, 70's and 80's with the top number being 2270 schools in 1971.
3. 88% of the schools are located in Rural areas.

4. 12% of the schools are located in Urban areas.
5. The highest percentage (almost 63%) of schools provides Urdu medium education in rural areas.
6. The lowest percentage (almost 2.5%) of schools provide English medium education in Urban areas.
7. Primary schools have the highest percentage in both rural and urban areas.
8. sMosque and H.Sec. schools have the lowest percentage in both rural and urban areas.
9. The highest number of schools are Primary followed by Middle and then Secondary.
10. The lowest number of schools are 'sMosque'.
11. There are 2339 more Female schools than Male schools in Punjab.
12. The school level with the lowest number of functional classrooms is 'sMosque' followed by 'H.Sec.'
13. Within 'sMosque' and 'H.Sec.', English medium schools have the lowest number of functional classrooms.
14. The Education Department owns the highest number of schools i.e. 45607.
15. Almost 30% of the schools have been upgraded to Primary level.
16. Almost 31% of the schools have been upgraded to Middle level.
17. Almost 15% of the schools have been upgraded to High level.
18. Almost 1.6% of the schools have been upgraded to Higher Secondary level.

## Student Enrollment

1. The number of Male students enrolled are more than the number of Female students enrolled overall in schools where only one gender is studying.
2. The number of enrollments in schools where both genders are studying is more than in gender-specific schools but we cannot tell about the number of males and females separately in those schools using this data.
3. The lowest students by gender have been enrolled in school level 'sMosque' which are Males (1349) followed by Males (2157) in 'H.Sec.' school for the schools established in 2000s.
4. No females have been enrolled in 'sMosque' since 2000s.

5. More females have been enrolled than males in all school levels other than 'sMosque' for the schools established in 2000s.
6. The highest number of enrollments by gender (Males and Females) is in Primary schools followed by Secondary, Middle and Higher Secondary schools since 2000s.
7. The overall highest number of enrollments is in Secondary schools followed by Primary schools.
8. The overall lowest number of enrollments is in sMosque level schools followed by H.Sec. schools.
9. The highest number of enrollments by level and medium of education is in Primary Urdu medium schools.
10. Enrollments in English medium schools are the lowest in all school levels.
11. District with the highest enrollment in Primary schools since their establishment is Rahimyar Khan with 257799 students enrolled.
12. District with the highest enrollment in Secondary schools since their establishment is Faisalabad with 362868 students enrolled.
13. District with the highest enrollment in Higher Secondary schools since their establishment is Faisalabad with 99227 students enrolled.

## Staff (Teachers / Non-teachers)

1. Top 5 districts with the highest number of teachers are:
  1. FAISALABAD
  2. RAHIMYAR KHAN
  3. RAWALPINDI
  4. LAHORE
  5. SARGODHA
2. Top 5 districts with the highest number of non-teachers are:
  - i. FAISALABAD
  - ii. RAWALPINDI
  - iii. SARGODHA
  - iv. LAHORE

#### v. BAHAWALNAGAR

3. Districts with the lowest number of teachers are CHINIOT, HAFIZABAD, LODHRAN, PAKPATTAN and RAJANPUR.
4. Districts with the lowest number of non-teachers are HAFIZABAD, CHINIOT, PAKPATTAN, RAJANPUR and KHUSHAB.
5. There are more teachers in rural schools than urban schools – rural schools have almost 3.5 times more teachers than urban schools overall.
6. Rural schools have highest number of teachers at the Primary level.
7. Urban schools have highest number of teachers at the secondary level.
8. Urdu medium Primary schools have the highest number of teachers.
9. Primary schools have overall highest number of teachers followed by Secondary and then Middle schools.
10. sMosque has the lowest number of teachers followed by H.Sec. schools.
11. Government schools have the highest number of teachers.
12. There are 390763 teachers and 73366 non-teachers in schools in Punjab.
13. 99.67% schools have teaching posts filled and 0.33% schools have teaching posts vacant.
14. 48.91% schools have non-teaching posts filled and 51.09% schools have non-teaching posts vacant.

### Infrastructure

1. Almost 75% of the schools have satisfactory building conditions.
2. Almost 1% of the school buildings have been declared dangerous.
3. Almost 13% of school buildings in rural areas have been constructed roughly or in a partially solid way.
4. Almost 1.5% of school buildings in urban areas have been constructed roughly or in a partially solid way.

### Facilities

1. Districts with the lowest availability of drinking water, electricity, boundary walls, and toilet facilities in schools are:

- a. CHINIOT
  - b. HAFIZABAD
  - c. NANKANA SAHIB
  - d. LODHRAN
  - e. MANDI BAHA UD DIN
2. Almost 70% of the schools have satisfactory security measures.

## Recommendations

### Allocation of Funds

1. Budget allocated to school level upgradation is justified as very less percentage of schools (almost 30% or less) have been upgraded to their higher levels.
2. Higher Secondary level schools and sMosque level schools are significantly neglected in both the rural and urban areas. More budget should be allocated for the construction of these schools alongwith Primary and Middle schools.
3. There are NO sMosque level schools for females so a portion of budget needs to be allocated to them as well.
4. New teachers would need to be appointed for new schools and existing teachers need to be trained with the latest trends and techniques in teaching so a fair amount of budget should be kept aside for them.
5. Some budget needs to be allocated for appointment of non-teaching staff as they are crucial for proper management of schools.
6. Allocation of budget for facilities like water, toilets, boundary walls and electricity is truly justified as the insights from data analysis truly depict its necessity.
7. Basic first aid facilities, scientific labs, computer labs, libraries and sports facilities must also be taken care of in budget allocation.
8. A portion of the budget must be allocated for digital marketing campaigns to maintain online presence of the schools and promoting high quality education in general.



## Budget Plan

3 Million dollars is approximately Rs 835 Million. The recommended budget plan to address different problems based on data analysis to improve the overall education situation in Pakistan could be as follows:

Area of improvement	Recommended Budget (PKR)
More middle and high schools in rural areas with separate schools for both genders	100 Million
More English medium schools in rural areas	100 Million
New Higher Secondary schools and sMosque schools in rural and urban areas	100 Million
School level upgradation	Rs 100 Million
Infrastructure: <ul style="list-style-type: none"> <li>• Building Condition</li> <li>• Classrooms</li> <li>• Boundary wall</li> <li>• Toilets</li> </ul>	Rs 200 Million
Facilities: <ul style="list-style-type: none"> <li>• Drinking water</li> <li>• Electricity</li> <li>• Security</li> </ul>	100 Million
Teachers (induction and training)	50 Million
Non-teaching staff	20 Million
Digital Marketing	20 Million
Other facilities: Playgrounds, Labs, Basic First Aid, Libraries, Sports	20 Million
Scholarships & Stipends for students	25 Million

## Marketing Strategies

The following marketing strategies could be adopted to promote educational initiatives:

1. Partnerships can be done with local businesses in the school areas to provide resources for the students.
2. Use digital marketing to attract companies who would like to help building new schools and improving the current ones.
3. Build an online presence of the schools to keep track of their achievements and upcoming events.
4. Use data analytics combined with digital marketing to identify the areas of improvement and focus on the areas lagging behind and to create awareness among people using data-driven insights.
5. Take advantage of government programs of improvement and ensure that schools are actively participating in those programs.
6. Work on the training and development of teachers to improve the overall quality of education and promote these activities through collaborations.

## References

- Pakistan Education Statistics 2020-21  
<http://library.aepam.edu.pk/Books/Pakistan%20Education%20Statistics%202020-21.pdf>
- Pakistan Punjab Education Fact Sheets 2022  
[https://data.unicef.org/wpcontent/uploads/2022/12/2022Pakistan\\_Factsheet\\_2022p11.28.22-1.pdf](https://data.unicef.org/wpcontent/uploads/2022/12/2022Pakistan_Factsheet_2022p11.28.22-1.pdf)
- School Education in Pakistan – A Sector Assessment  
<https://www.adb.org/sites/default/files/institutional-document/518461/pakistan-school-education-sector-assessment.pdf>
- School Information System – Government of the Punjab  
<https://sis.punjab.gov.pk/>
- Annual School Census Report for Settled Districts 2020-2021, Govt. of Khyber Pakhtunkhwa  
[http://175.107.63.45/newimwebsite/images/reports/ASC\\_Report\\_2020-21\\_Final.pdf](http://175.107.63.45/newimwebsite/images/reports/ASC_Report_2020-21_Final.pdf)
- Balochistan Education Statistics 2021-2022  
[http://www.emis.gob.pk/Uploads/BalochistanEducationStatistics/Balochistan\\_Education\\_Statistics\\_2021-22.pdf](http://www.emis.gob.pk/Uploads/BalochistanEducationStatistics/Balochistan_Education_Statistics_2021-22.pdf)
- Analysis Of Age Specific Data of Students upto Higher Secondary School Level in Pakistan  
<http://www.neas.gov.pk/SiteImage/Publication/AASDSHSL19.pdf>