Ministry of Education

Analysis Report of Pakistan School's Dataset for Educational Policy Improvements

Prepared and reported by
Abdullah Akmal

Table of Contents

1. Executive Summary
    a. Key Findings
    b. Recommendations
2. Introduction
    a. Objectives
    b. Dataset Overview
3. Methodology
    a. Data Cleaning
4. Findings and Analysis
    a. Visualizations and Insights
5. Recommendations and Policy Implications
6. Conclusion
7. Appendices
    a. Heatmap of Average teachers per school categorized by district and school type
    b. Security Levels % by Districts Table
    c. Boundary Walls % by Districts Table
    d. Electricity % by Districts Table
    e. Drinking Water % by Districts Table

1. Executive Summary

This report analyzes a comprehensive dataset of schools across Pakistan to assess the state of educational infrastructure and outcomes. The objective is to identify gaps in infrastructure, regional disparities, and potential areas for policy intervention for the Ministry of Education.

Key Findings

In Pakistan, 48% of schools are male, and 52% are female, with an average enrollment distribution of 29.5% in rural areas and 70.5% in urban areas. The majority (69%) of schools operate in Urdu as the medium of instruction, while 19.5% are bilingual (Urdu and English), and only 11.5% are English medium. Most schools (99%) are functional and operate in the morning shift. However, 23.33% of schools lack security, and 6.39% have unsatisfactory security measures, leaving only 70.29% with adequate security. Infrastructure gaps include 2.68% of schools without boundary walls, 1.56% without electricity, and 0.48% without access to drinking water.

Recommendations:

- **Medium of Instruction**: Increase training and resources for teachers to deliver bilingual or English medium education to meet global standards.
- **Security**: Prioritize the deployment of security personnel and installation of surveillance systems in schools lacking security.
- **Boundary Walls:** Introduce partnerships with local communities and NGOs to fast-track the construction of boundary walls.
- **Electricity:** Provide solar energy solutions to schools without electricity, especially in remote areas.
- **Drinking Water:** Install water filtration systems and boreholes in schools without access to drinking water.

2. Introduction:

Education is a critical driver for national development. This project, initiated by the Ministry of Education, aims to evaluate the existing school infrastructure and propose data-driven recommendations. By identifying gaps and prioritizing regions for intervention, this report will contribute to a more equitable education system.

Objectives:

- **Analyze the KPIs:** Examine how schools and facilities are distributed among enrollment outcomes.
- **Evaluate Infrastructure:** Assess the availability and condition of school infrastructure.
- **Provide Actionable Insights:** Identify underperforming regions and propose targeted improvements.

Dataset Overview:

- Key Columns:
- *School Details:* ID, Name, District, Enrollment.
- *Infrastructure:* Toilets, Boundary Wall, Water Availability.
- *Staffing:* Teachers, Non-Teachers.
- Data Scope: Covers all provinces in Pakistan, capturing urban and rural school statistics.
- Limitations: The dataset has several missing or invalid values, including blanks, zeros, and extreme values. Variables like moza, street_name, uc_name, na_no, pp_no, and several others need data cleaning.
- Source: Ministry of Education school survey data.

---

3. Methodology:

I have used Microsoft excel for analyzing this dataset through pivot-tables and pivot-charts.

Data Cleaning:

The dataset has several missing or invalid values, including blanks, zeros, and extreme values, which have been replaced with NULL or zero. Variables like moza, street_name, uc_name, na_no, pp_no, and several others were cleaned by converting invalid entries, including empty cells and extreme values, into NULL or zero. Additionally, fields such as head_grade, school_ownership, and various upgrade_*_year columns were also cleaned by converting blank or zero values to NULL where appropriate.

4. Findings and Analysis:

All the key findings from the executive summary are explained here through charts and interpretations.

Visualizations and Insights:

Enrollment by School Gender:

Figure 1 represents Pink Bar are for females' schools which are 52% and Blue Bar are for males which are 48% which tells us that there more female schools than male schools.

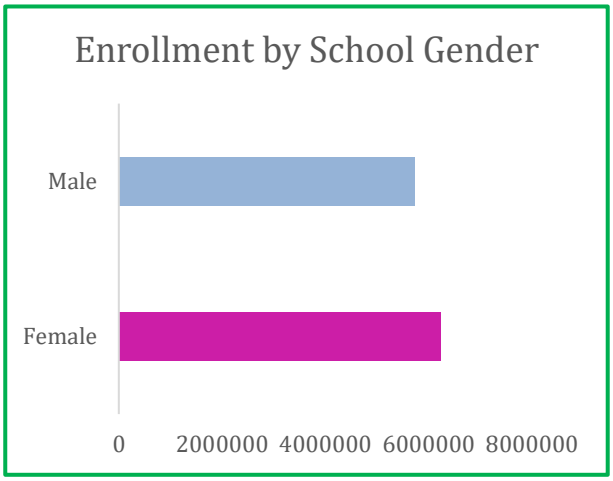

Figure 1

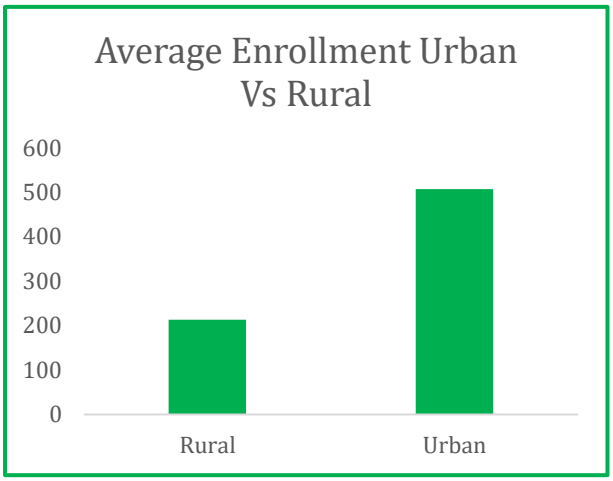

Figure 2

Average Enrollment Urban Vs Rural:

Figure 2 represents the average enrollment distribution is 29.5% in rural areas and 70.5% in urban areas of students. From this dataset 507 students are enrolled in urban areas and 213 students are enrolled in rural areas out of 720 students.

Distribution among Medium

Figure 3 represents the majority (69%) of schools operate in Urdu as the medium of instruction, while 19.5% are bilingual (Urdu and English), and only 11.5% are English medium. In Pakistan, there are more schools that use Urdu as the primary medium of instruction compared to English or a combination of both.

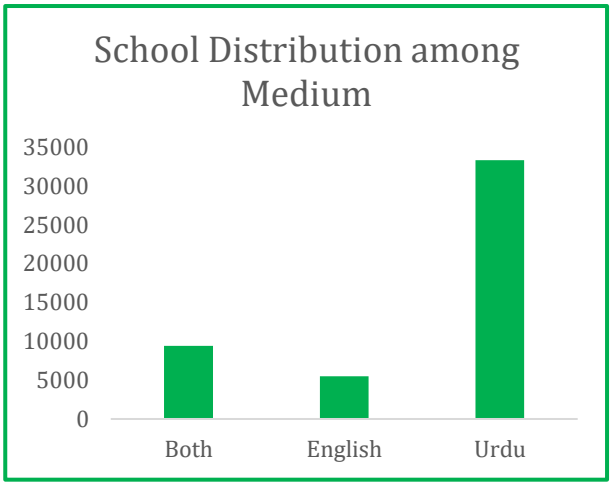

Figure 3

Functional Vs Non-Functional Schools and Schools by Shift:

Figures 4 and 5 represent that 99% of schools are functional and 99% of schools are in the morning shift. Most schools from the Ministry of education are functional and operate in the morning shift.

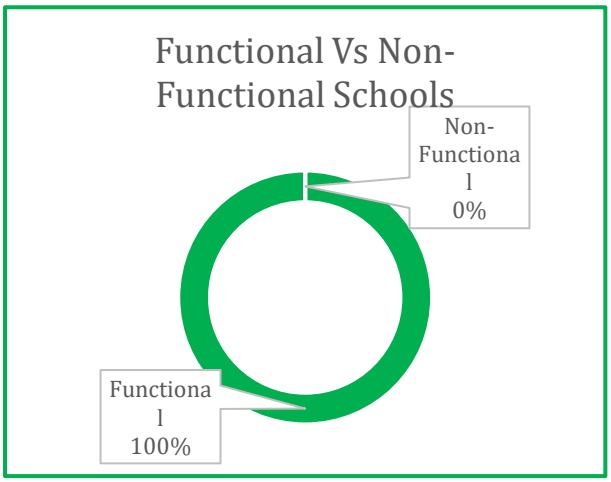

Figure 4

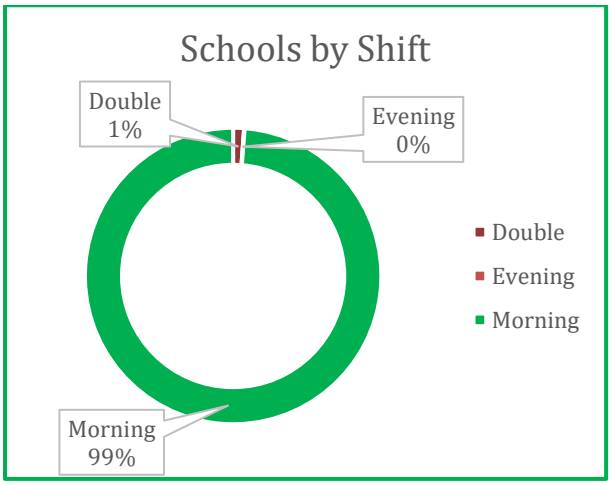

Figure 5

Security Levels % by Districts:

Figure 6 represents that most schools in the districts have satisfying security levels. However, there are some districts where a significant number of schools' security isn't available or didn't not satisfying security level. Rahim Yar Khan district have highest security and without security level among all the districts of Pakistan. In Overall numbers there are 23.33% of schools lacking security, and 6.39% have unsatisfactory security measures, leaving only 70.29% with adequate security.

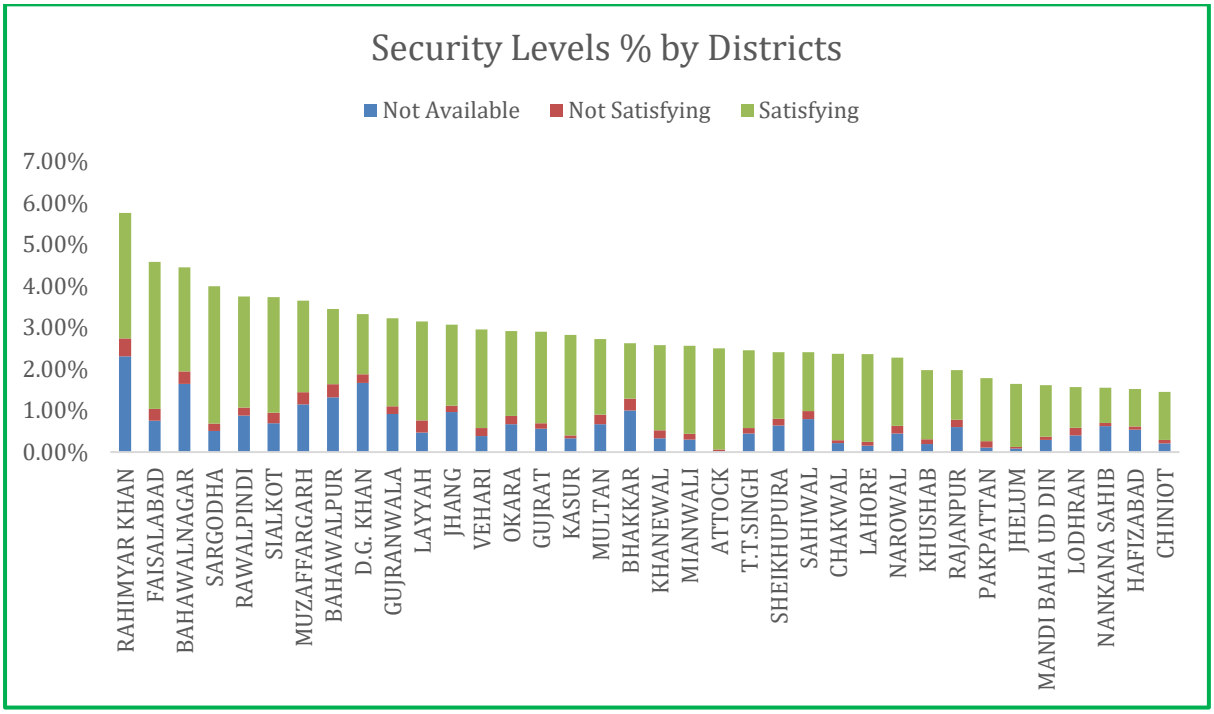

Figure 6

Boundary Walls % by Districts:

Figure 7 represents that most schools in districts have boundary walls. Rahim Yar Khan district have highest boundary wall % among schools whereas Bahawalnagar district has highest no boundary wall% among the schools. The infrastructure gaps include 2.68% of schools without boundary walls among all districts.

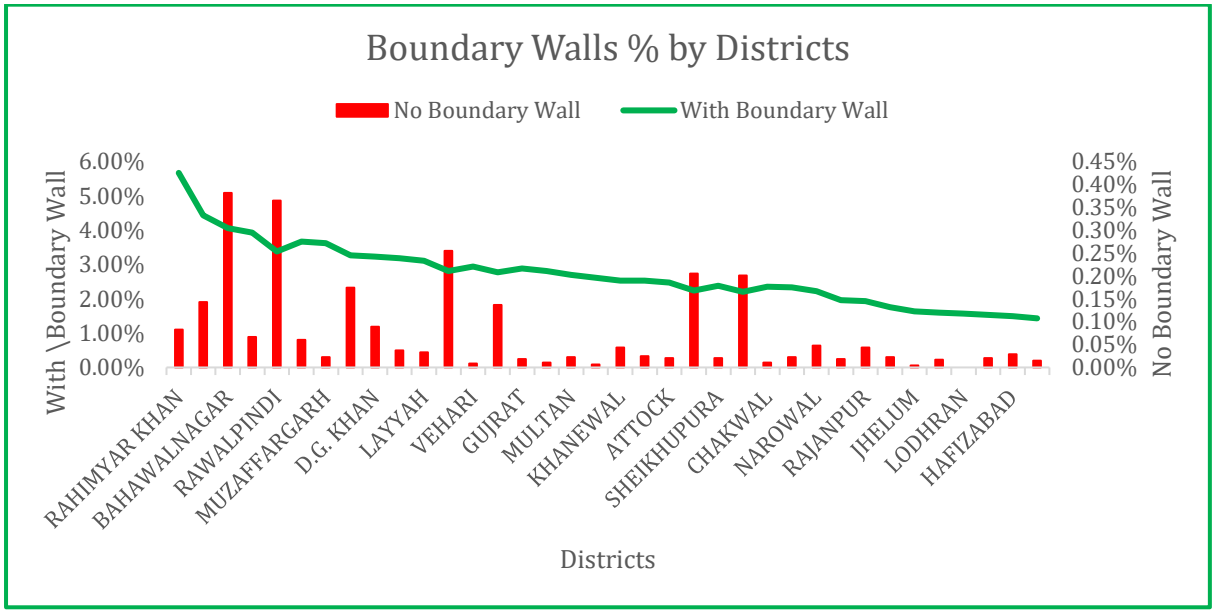

Figure 7

Electricity % by Districts:

Figure 8 represents that most schools in districts have electricity. Rahim Yar Khan district has the highest electricity % among schools as well as Rahim Yar Khan district has the highest electricity % among the schools. Overall, 1.56% of schools are without electricity among all districts.

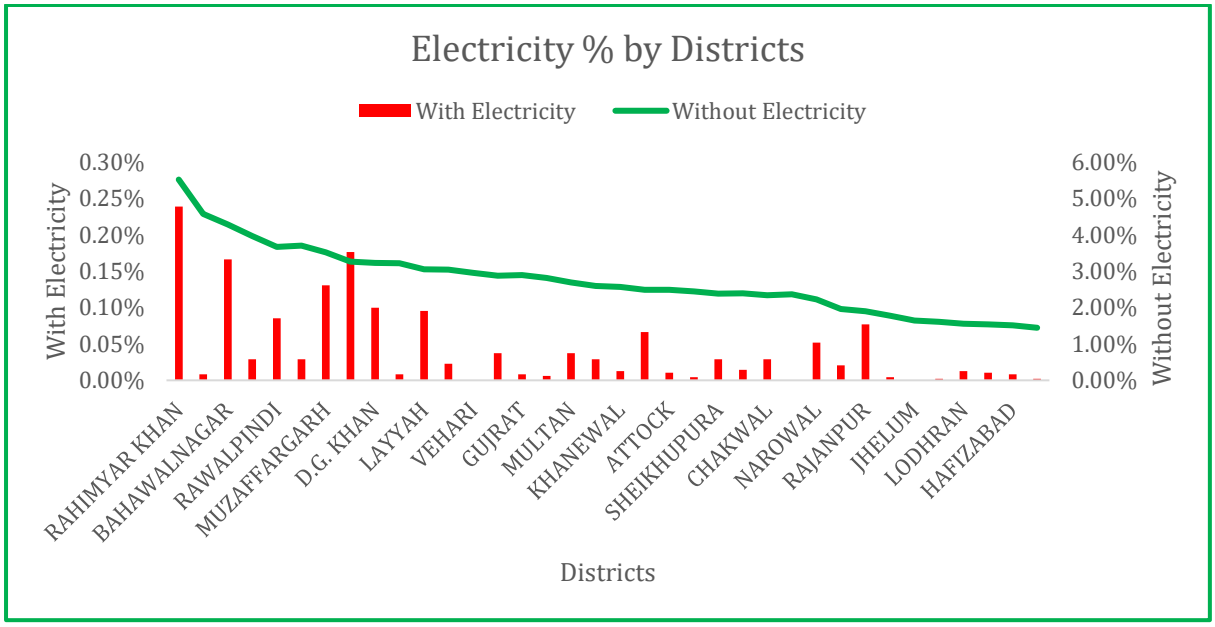

Figure 8

Drinking Water % by Districts:

Figure 9 represents that most schools in districts have drinking water available. Rahim Yar Khan district has the highest drinking water % among schools whereas D.G Khan district has the highest without drinking water % among the schools. Overall, 0.48% of schools are without access to drinking water among all districts.

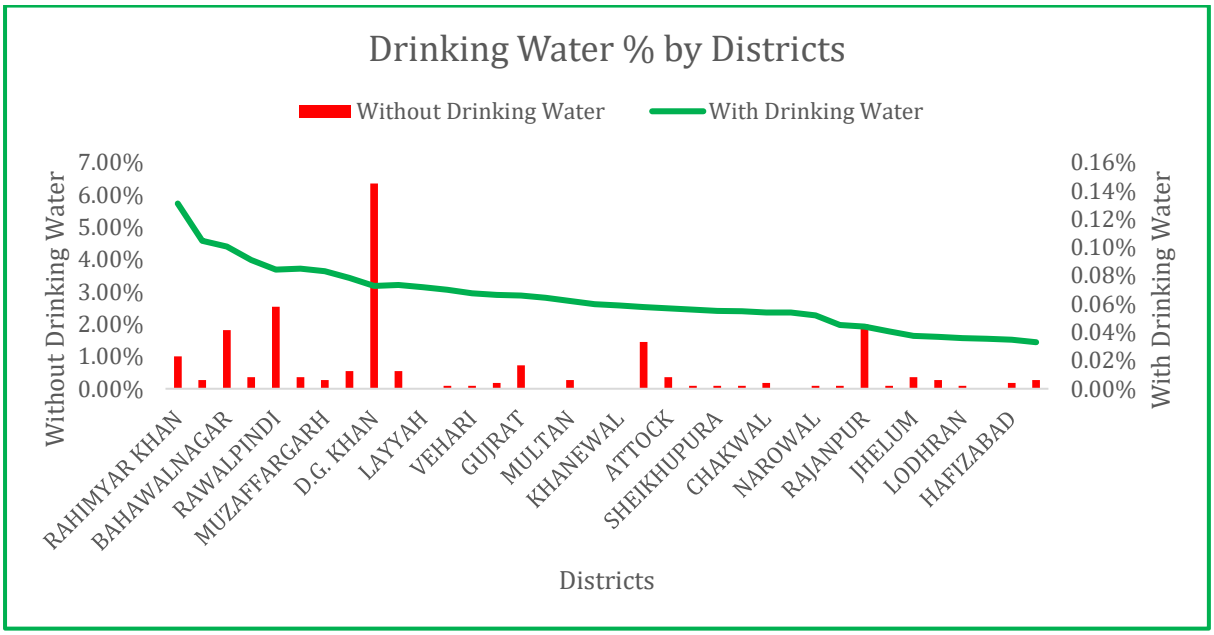

Figure 9

---

5. Recommendations and Policy Implications

To address these challenges and improve the quality of education in Pakistan, the following policy recommendations are proposed:

1. **Enhance Infrastructure:**
   a. **Prioritize Rural Areas:** Allocate additional resources to rural schools to improve infrastructure, particularly in terms of electricity and water supply.
   b. **Public-Private Partnerships:** Encourage public-private partnerships to accelerate infrastructure development.
   c. **Solar Energy Solutions:** Implement solar energy solutions in schools located in remote areas to ensure a reliable power supply.
2. **Strengthening Security:**
   a. **Mandatory Security Audits:** Conduct regular security audits of all schools to identify vulnerabilities and implement necessary measures.

b. **Community Involvement:** Engage with local communities to establish school safety committees.
c. **Training and Capacity Building:** Provide training to school staff on security protocols and emergency response procedures.

3. **Promote Bilingual and English-Medium Education:**
   a. **Teacher Training:** Invest in teacher training programs to equip educators with the skills to deliver effective bilingual and English-medium instruction.
   b. **Curriculum Development:** Develop curriculum materials that cater to the needs of diverse learners and promote critical thinking and problem-solving skills.
   c. **Resource Allocation:** Allocate sufficient resources to schools to support the implementation of bilingual and English-medium programs.
4. **Data-Driven Decision Making:**
   a. **Regular Data Collection:** Establish a robust data collection system to monitor school performance and identify areas for improvement.
   b. **Data Analysis and Visualization:** Utilize data analytics tools to generate actionable insights.
   c. **Evidence-Based Policymaking:** Use data to inform policy decisions and allocate resources effectively.

6: Conclusion:

The analysis of the Pakistani school dataset reveals a complex picture of the country's educational landscape. While most schools are functional and operate in the morning shift, significant disparities exist in terms of infrastructure, security, and medium of instruction.

Key findings include:

- **Gender Disparity:** A slight imbalance exists, with more female schools than male schools.
- **Urban-Rural Divide:** Urban schools have significantly higher enrollment rates compared to rural schools.
- **Medium of Instruction:** Urdu remains the dominant medium of instruction, with a smaller proportion of schools offering bilingual or English-medium education.
- **Infrastructure Gaps:** A small but concerning percentage of schools lack essential infrastructure like boundary walls, electricity, and drinking water.
- **Security Concerns:** A significant number of schools have inadequate security measures, posing potential risks.

7: Appendices:

Heatmap of Average teachers per school categorized by district and school_type.

Average of Teache	Column Label										
Row Labels	PSSP School	Govt. Schoo	Junior Mode	Community Scho	Model Schoo	M.C Local	Danish Adopted Scho	Centre of Excellenc	Comprehensiv	Technical Hig	Pilot Secondar
ATTOCK		7.649901381	9.25	8.704545455	6.602941176	11.33333333		4	14.44444444	18	25.75
BAHAWALNAGAR		6.178211587	6.814814815	6.674242424	9.27739726	8.76		11	12.625	11	15.33333333
BAHAWALPUR		6.979625102	8.7	6.397790055	10.66013072	9.423076923		15	13.375	23.25	22.66666667
BHAKKAR		6.080550098	4.333333333	6.081081081	7.944954128	8.277777778		4	9.818181818	19.5	13.5
CHAKWAL		7.562268804	4.8	8.618320611	7.122807018	8.5625	13	17.5	10.91891892	14.66666667	16.66666667
CHINIOT	3	6.50310559	4.333333333	8.449275362	5.950980392	7.071428571	10	4	0	12.5	37.5
D.G. KHAN	4	6.037865749	5	5.4	7.100775194	5.27027027		8	23.33333333	13.4	17.5
FAISALABAD		9.185801442	7.083333333	10.072	23.1509434	15.29357798	9.5	15	24.44444444	27.6	14.33333333
GUJRANWALA		9.313060818	5.833333333	7.990909091	6.036630037	8.896551724		14.66666667	22.53846154	17.125	22.83333333
GUJRAT		8.546906188	5.272727273	11.90410959	7	8.2		30	22	23.14285714	17.6
HAFIZABAD		6.763157895	3.125	6.968085106	5.421052632	8	13		10.09090909	31.66666667	29.66666667
JHANG		7.158318426	5.5625	7.186666667	6.903381643	10.8			30.66666667	16.5	23
JHELUM		8.507604563	11.8	8.969230769	7.261437908	8.083333333			19	15.5	12.5
KASUR		8.319806763	5.428571429	9.236842105	8.08411215	10.02777778			14.26666667	19.4	
KHANEWAL		9.026348808	4.777777778	8.784090909	9.545801527	10.48387097			12.6	22.33333333	18.5
KHUSHAB		6.387139108	4.384615385	7.952941176	8.5625	8.923076923			11.5	28	22
LAHORE		12.26587302	12.69565217	16.02150538	25.87837838	16.640625	9	42	63.66666667	26.25	38
LAYYAH		6.067307692	5.3	6.201834862	9.866666667	5.866666667		8	6.222222222	18.75	8.5
LODHRAN		7.554828151	4.75	8.909090909	9.375	6.75			11	27.5	
MANDI BAHA UD DIN		8.873747495	5.666666667	9.858823529	6.179104478	9.727272727			11.45454545	21.8	27.66666667
MIANWALI		6.241610738	6.25	6.887218045	8.450980392	5.028571429			26	6	
MULTAN		8.254677755	22	9.021052632	8.4	10.1369863	5	4	37.6	22.6	50
MUZAFFARGARH		6.70398482	10.5	8.265306122	11.42857143	8.666666667			23.33333333	23	33.5
NANKANA SAHIB		7.74829932	19.5	9.727272727	8.337078652	17	20		14.42857143	45.14285714	24
NAROWAL		9.046587216	5.555555556	7.897959184	8.880597015	8.416666667		11	14.18181818	9.5	
OKARA		7.64689781	22	7.878787879	10.05932203	17.42105263		6	10.5	18.25	20.5
PAKPATTAN		6.836411609		9.405405405	9.021276596	18.33333333			30		
RAHIMYAR KHAN		5.808303887	8.090909091	6.659793814	9.072289157	9.378378378			14.33333333	14.625	22.6
RAJANPUR		6.211347518	4.857142857	5.875	9.744680851	6		5	5	25	21
RAWALPINDI		9.230293663	7.769230769	9.034188034	6.125560538	15.06382979		29	21.41176471	25.25	16.2
SAHIWAL		8.790870488	17.125	8.409090909	14.94444444	8.64516129			28.75	15	58
SARGODHA		7.66	11.875	7.74789916	8.487272727	10.59090909			17.22222222	20.28571429	23.83333333
SHEIKHUPURA		7.795918367	7.285714286	8.303571429	10.02739726	9.921052632	3.333333333	8	12.30769231	19.9	26.66666667
SIALKOT		8.308219178	6	8.608695652	6.97752809	12.66666667	23.75		12.55555556	22	22
T.T.SINGH		8.522248244		9.333333333	9.040983607	8.565217391	5	9	9.818181818	20.57142857	13
VEHARI		7.330170778	6.777777778	6.957142857	12.03797468	6.285714286		4	8	16.66666667	15
Grand Total	**3.5**	**7.6020035**	**7.6184834**	**8.059011164**	**8.52174849**	**10.960185**	**12.4375**	**12.51851852**	**15.35179153**	**21.08379888**	**21.53271028**

Security Levels % by Districts Table:

Count of security	Column Labels		
Row Labels	Not Available	Not Satisfying	Satisfying
RAHIMYAR KHAN	2.34%	0.43%	3.02%
FAISALABAD	0.77%	0.30%	3.48%
BAHAWALNAGAR	1.66%	0.30%	2.50%
SARGODHA	0.50%	0.17%	3.22%
SIALKOT	0.70%	0.25%	2.80%
MUZAFFARGARH	1.17%	0.30%	2.23%
RAWALPINDI	0.88%	0.19%	2.62%
BAHAWALPUR	1.34%	0.33%	1.80%
D.G. KHAN	1.66%	0.21%	1.45%
GUJRANWALA	0.94%	0.18%	2.08%
LAYYAH	0.47%	0.29%	2.39%
JHANG	0.98%	0.15%	1.95%
VEHARI	0.39%	0.19%	2.39%
GUJRAT	0.58%	0.14%	2.22%
OKARA	0.69%	0.19%	2.02%
KASUR	0.33%	0.07%	2.43%
MULTAN	0.68%	0.23%	1.81%
BHAKKAR	1.03%	0.30%	1.35%
KHANEWAL	0.34%	0.19%	2.03%
MIANWALI	0.31%	0.14%	2.07%
ATTOCK	0.03%	0.03%	2.41%
T.T.SINGH	0.44%	0.13%	1.87%
SAHIWAL	0.80%	0.19%	1.40%
SHEIKHUPURA	0.65%	0.16%	1.58%
LAHORE	0.16%	0.09%	2.12%
CHAKWAL	0.21%	0.07%	2.03%
NAROWAL	0.45%	0.19%	1.67%
KHUSHAB	0.19%	0.12%	1.70%
RAJANPUR	0.61%	0.18%	1.18%
PAKPATTAN	0.11%	0.16%	1.54%
JHELUM	0.09%	0.04%	1.51%
MANDI BAHA UD DIN	0.30%	0.08%	1.23%
NANKANA SAHIB	0.64%	0.08%	0.85%
LODHRAN	0.41%	0.18%	0.97%
HAFIZABAD	0.54%	0.08%	0.86%
CHINIOT	0.22%	0.09%	1.15%
Grand Total	**23.61%**	**6.44%**	**69.96%**

Boundary Walls % by Districts Table:

Count of boundary_wall Row Labels	Column Labels No Boundary Wall	With Boundary Wall
RAHIMYAR KHAN	0.09%	5.71%
FAISALABAD	0.14%	4.41%
BAHAWALNAGAR	0.38%	4.08%
SARGODHA	0.07%	3.83%
SIALKOT	0.06%	3.70%
MUZAFFARGARH	0.02%	3.68%
RAWALPINDI	0.37%	3.33%
BAHAWALPUR	0.18%	3.28%
D.G. KHAN	0.09%	3.23%
GUJRANWALA	0.04%	3.16%
LAYYAH	0.03%	3.12%
JHANG	0.26%	2.82%
VEHARI	0.01%	2.96%
GUJRAT	0.02%	2.92%
OKARA	0.14%	2.77%
KASUR	0.01%	2.81%
MULTAN	0.02%	2.69%
BHAKKAR	0.01%	2.66%
KHANEWAL	0.04%	2.52%
MIANWALI	0.03%	2.49%
ATTOCK	0.02%	2.46%
T.T.SINGH	0.21%	2.23%
SAHIWAL	0.20%	2.19%
SHEIKHUPURA	0.02%	2.36%
LAHORE	0.02%	2.35%
CHAKWAL	0.01%	2.30%
NAROWAL	0.05%	2.26%
KHUSHAB	0.02%	2.00%
RAJANPUR	0.04%	1.93%
PAKPATTAN	0.02%	1.80%
JHELUM	0.00%	1.64%
MANDI BAHA UD DIN	0.02%	1.60%
NANKANA SAHIB	0.02%	1.55%
LODHRAN	0.00%	1.55%
HAFIZABAD	0.03%	1.45%
CHINIOT	0.01%	1.44%
Grand Total	**2.72%**	**97.28%**

Electricity % by Districts Table:

Count of electricity	Column Labels	
Row Labels	With Electricity	Without Electricity
RAHIMYAR KHAN	0.24%	5.56%
FAISALABAD	0.01%	4.54%
BAHAWALNAGAR	0.17%	4.28%
SARGODHA	0.03%	3.87%
SIALKOT	0.03%	3.73%
MUZAFFARGARH	0.13%	3.57%
RAWALPINDI	0.09%	3.61%
BAHAWALPUR	0.18%	3.28%
D.G. KHAN	0.10%	3.22%
GUJRANWALA	0.01%	3.19%
LAYYAH	0.10%	3.05%
JHANG	0.02%	3.06%
VEHARI	0.00%	2.97%
GUJRAT	0.01%	2.93%
OKARA	0.04%	2.87%
KASUR	0.01%	2.82%
MULTAN	0.04%	2.68%
BHAKKAR	0.03%	2.64%
KHANEWAL	0.01%	2.55%
MIANWALI	0.07%	2.45%
ATTOCK	0.01%	2.47%
T.T.SINGH	0.00%	2.44%
SAHIWAL	0.01%	2.38%
SHEIKHUPURA	0.03%	2.36%
LAHORE	0.00%	2.38%
CHAKWAL	0.03%	2.28%
NAROWAL	0.05%	2.25%
KHUSHAB	0.02%	2.00%
RAJANPUR	0.08%	1.90%
PAKPATTAN	0.00%	1.81%
JHELUM	0.00%	1.64%
MANDI BAHA UD DIN	0.00%	1.61%
NANKANA SAHIB	0.01%	1.56%
LODHRAN	0.01%	1.54%
HAFIZABAD	0.01%	1.47%
CHINIOT	0.00%	1.45%
Grand Total	**1.59%**	**98.41%**

Drinking Water % by Districts Table:

Count of drink_water Row Labels	Column Labels Without Drinking Water	With Drinking Water
RAHIMYAR KHAN	0.02%	5.77%
FAISALABAD	0.01%	4.55%
BAHAWALNAGAR	0.04%	4.42%
SARGODHA	0.01%	3.89%
SIALKOT	0.01%	3.75%
MUZAFFARGARH	0.01%	3.70%
RAWALPINDI	0.06%	3.64%
BAHAWALPUR	0.01%	3.45%
D.G. KHAN	0.15%	3.17%
GUJRANWALA	0.01%	3.18%
LAYYAH	0.00%	3.15%
JHANG	0.00%	3.08%
VEHARI	0.00%	2.97%
GUJRAT	0.02%	2.92%
OKARA	0.00%	2.90%
KASUR	0.00%	2.83%
MULTAN	0.01%	2.71%
BHAKKAR	0.00%	2.67%
KHANEWAL	0.00%	2.56%
MIANWALI	0.03%	2.48%
ATTOCK	0.01%	2.47%
T.T.SINGH	0.00%	2.44%
SAHIWAL	0.00%	2.39%
SHEIKHUPURA	0.00%	2.38%
LAHORE	0.00%	2.38%
CHAKWAL	0.00%	2.31%
NAROWAL	0.00%	2.30%
KHUSHAB	0.00%	2.02%
RAJANPUR	0.05%	1.93%
PAKPATTAN	0.00%	1.82%
JHELUM	0.01%	1.63%
MANDI BAHA UD DIN	0.00%	1.61%
NANKANA SAHIB	0.00%	1.57%
LODHRAN	0.00%	1.55%
HAFIZABAD	0.00%	1.48%
CHINIOT	0.01%	1.45%
Grand Total	**0.48%**	**99.52%**