# School Performance Clustering Report

## Objective

The purpose of this analysis is to group schools based on their academic performance, as measured by the **Pass Rate**, **Bachelor Rate**, **Learner Count**, **Teacher Attendance**, and **Learner Attendance**. This report identifies clusters of schools with similar characteristics, providing insights into performance trends and areas requiring intervention.

## Data Overview

The dataset contains information for 10 schools across different provinces. The key performance metrics are:

- **Pass Rate (%)**: Percentage of learners passing.
- **Bachelor Rate (%)**: Percentage of learners achieving bachelor's passes.
- **Learner Count**: The number of learners enrolled.
- **Teacher Attendance (%)**: The attendance rate of teachers.
- **Learner Attendance (%)**: The attendance rate of learners.

## Clustering Analysis Methodology

Hierarchical clustering was performed on the numerical data, including Pass Rate, Bachelor Rate, Learner Count, Teacher Attendance, and Learner Attendance. The following steps were followed:

1. **Data Standardization**: The numerical features were standardized to ensure that all variables contributed equally to the clustering process.
2. **Clustering Method**: The **Ward's Method** was used for hierarchical clustering, which minimizes the variance within clusters.
3. **Optimal Cluster Determination**: The **Silhouette Score** method was used to determine the optimal number of clusters, resulting in **3 clusters**.

## Dendrogram

A dendrogram was generated to visualize the hierarchical clustering. The branches of the dendrogram illustrate how schools are grouped based on their similarity in terms of the selected features.

## Cluster Results

The clustering analysis divided the schools into 3 distinct groups:

### Cluster 1: High Performance Schools

- **Characteristics**: High Pass Rate, High Bachelor Rate, High Teacher and Learner Attendance.
- **Schools**: School A, School E, School I, School J.

### Cluster 2: Moderate Performance Schools

- **Characteristics**: Moderate Pass Rate, Moderate Bachelor Rate, Moderate Teacher and Learner Attendance.
- **Schools**: School B, School F, School G.

### Cluster 3: Low Performance Schools

- **Characteristics**: Low Pass Rate, Low Bachelor Rate, Lower Teacher and Learner Attendance.
- **Schools**: School C, School D, School H.

## Insights and Observations

1. **Cluster 1 (High Performance Schools)**:
    - These schools have a high academic performance (Pass and Bachelor rates) as well as excellent attendance from both teachers and learners.
    - **Actionable Insight**: These schools can serve as benchmarks for the other clusters, providing best practices that can be shared with underperforming schools.
2. **Cluster 2 (Moderate Performance Schools)**:
    - These schools show moderate academic performance but could benefit from improved attendance rates, especially among teachers and learners.
    - **Actionable Insight**: Focus on improving attendance and providing additional academic resources to boost performance.
3. **Cluster 3 (Low Performance Schools)**:
    - These schools exhibit low academic performance and lower attendance rates, both for teachers and learners.
    - **Actionable Insight**: Targeted interventions are needed to improve both performance and attendance. This may include teacher training, learner support programs, and resources to address attendance issues.

## Next Steps

- **Investigate Underperforming Schools**: Delve deeper into the factors causing low performance and low attendance in Cluster 3 schools. Possible areas of focus include infrastructure, teacher training, learner engagement, and local socioeconomic factors.
- **Tailored Interventions**: Develop specific intervention plans for each cluster. For Cluster 1, consider mentorship and collaboration with other schools. For Cluster 2 and 3, introduce targeted support such as after-school programs, increased teacher training, and attendance monitoring systems.
- **Monitoring**: Track the progress of schools in each cluster and reassess their performance over time. Re-cluster the schools periodically to see the impact of the interventions.

## Conclusion

This clustering analysis has identified clear groups of schools based on performance and attendance data. Schools in Cluster 1 exhibit high academic achievement and strong attendance, while Cluster 3 schools require immediate attention to improve both academic outcomes and attendance. By targeting specific needs within each cluster, interventions can be developed to enhance overall school performance.

The dynamic use of **Silhouette Analysis** for optimal cluster determination ensures that the results are robust and reliable, allowing for more effective decision-making in educational interventions.