Data Suppressions and Small Group Insights Report

1. How many data points are suppressed (s) across all tables, and what subgroups are most affected?

- **Total Suppressed Data Points**: 4,486,928

- **Subgroups Most Affected**:

    - **ELA**:

        - Multiracial: **16,843 suppressions**

        - Black or African American: **12,977 suppressions**

        - Hispanic or Latino: **11,462 suppressions**

    - **Math**:

        - Multiracial: **16,644 suppressions**

        - Black or African American: **12,892 suppressions**

        - Asian or Native Hawaiian/Other Pacific Islander: **12,635 suppressions**

    - **Science**:

        - Multiracial: **3,445 suppressions**

        - Black or African American: **2,712 suppressions**

        - Hispanic or Latino: **2,642 suppressions**

This indicates that race/ethnicity subgroups experience significant data suppression across all subjects.

2. How does combining small groups (e.g., Small Group Total) affect overall proficiency rates?

- **Small Group Proficiency Rate**: 47.07%

- **Overall Proficiency Rate**: 45.21%

- **Impact**: Combining small groups slightly improves the overall proficiency rate, suggesting that small groups might perform better on average compared to larger subgroups.

3. Is there a notable performance difference between suppressed and non-suppressed groups?

- **Suppressed Groups Proficiency Rate**: NaN (due to missing or invalid data)

- **Non-Suppressed Groups Proficiency Rate**: 45.21%

The missing data for suppressed groups highlights the need for better handling or imputation methods for suppressed values to enable performance comparisons.

4. Are suppressed subgroups more common in specific grades or subjects?

- **Suppressions by Grade/Subject**: No specific grades or subjects show significant suppressions in the provided data.

This suggests that suppressions are distributed evenly across grades and subjects or are more subgroup-focused.

5. How often is gender data suppressed in comparison to race/ethnicity data?

- **Gender Suppressions**: **92,136**
- **Race/Ethnicity Suppressions**: **1,744,358**

Race/ethnicity subgroups face substantially more data suppression compared to gender-related subgroups, reflecting broader challenges in reporting data for smaller racial/ethnic populations.

Key Insights

1. **Suppression Prevalence**:
    - A significant number of data points are suppressed, particularly for racial/ethnic subgroups.

2. **Small Group Benefits**:
    - Combining small groups slightly improves proficiency rates, indicating a potential benefit in aggregating data for smaller populations.

3. **Data Challenges**:
    - Suppressed data remains difficult to analyze, especially for certain subgroups and metrics.

Recommendations

1. **Improved Data Handling**:
    - Develop methods to impute or estimate suppressed data points to enable a more comprehensive analysis.

2. **Targeted Support**:
    - Focus on the subgroups with the highest suppression rates, such as racial/ethnic minorities, to address reporting challenges.

3. **Leverage Small Groups**:
    - Explore the performance of small groups further to identify effective strategies that could benefit larger populations.