|  |
| --- |
| **Data Set:** Cleaned Grades1000 |
| **Question / Problem:**   1. Find the mean and median of the math grade where the student has completed the prep test |
| **Variables:**   * PrepTest: A categorical variable indicating if the student completed the prep test * MathGrade, ReadingGrade, WritingGrade: Numeric variables that contain the grades for each student in different subjects. * average\_grades: A new data frame created after summarizing the data with the mean grades. * average\_grades\_long: A reshaped version of average\_grades for easier plotting. |
| **Techniques used:**  **Filtering Data**:   * Filtered to include only students who completed the prep test (PrepTest == "Yes"). * Ensures that the analysis focuses solely on the target group.   **Statistical Analysis**:   * **Mean Calculation**: Used mean() to compute the average math grade. * **Median Calculation**: Used median() to find the central value of the math grade distribution. * Handled missing data using na.rm = TRUE to avoid inaccuracies due to null values.   **Data Summarization**:   * Grouped the data using group\_by() based on the prep test status. * Calculated summary statistics for each subject (math, reading, writing) using summarise().   **Visualization with ggplot2**:   * Created a bar chart to display the average grades for each subject. * Used aes() to map the x-axis (subjects) and y-axis (average grades), and set bar colors. * Used geom\_bar() to represent actual values with separated bars. * Added meaningful titles and axis labels using labs(). |
| **Analysis and Visualisation:**            **Findings:**   * Students who completed the prep test have a higher average math grade compared to those who did not. * This suggests a positive correlation between prep test completion and math performance.   **Bar Chart:**   * Depicts the average grades (math, reading, writing) for students who completed the prep test. * Bars are color-coded by subject, with heights representing mean grades. * Highlights the math grade as the highest average among the three subjects. |
| **Considerations:**  If space or simplicity is a concern, a table summarizing the averages could replace the chart, but it might reduce visual impact**.** |