|  |
| --- |
| **Data Set:** Cleaned Grades1000 |
| **Question / Problem:**   1. Does completing a prep test affect the average math grade, reading grade and writing grade |
| **Variables:**  MathGrade:Dependent, Numerical  ReadingGrade:Dependent, Numerical  WritingGrade:Dependent, Numerical  PrepTest: Independent, Categorical |
| **Techniques used:**   1. **Data Preparation**    * The Cleaned Grades1000 dataset was loaded.    * Missing values were handled appropriately (e.g., excluding invalid data using na.rm = TRUE for numerical calculations). 2. **Grouping by PrepTest**    * The group\_by() function from the dplyr package was applied to classify data based on whether students completed the PrepTest. This separated the observations into "Yes" and "No" groups for comparison. 3. **Calculation of Averages**    * The summarise() function was used to calculate the mean for each subject (MathGrade, ReadingGrade, and WritingGrade) within each group (PrepTest completed vs. not completed).    * The mean() function, with the argument na.rm = TRUE, ensured missing values were ignored to prevent calculation errors. 4. **Data Reshaping for Visualization**    * The pivot\_longer() function transformed the dataset from a wide format (columns for MathGrade, ReadingGrade, WritingGrade) to a long format. This allowed each subject grade to be represented in a single column, making it compatible with the visualization process. 5. **Visualization Using** ggplot2    * A **bar chart** was created with ggplot() to visually compare the average grades for students who completed the PrepTest vs. those who did not.    * The aes() function was used to set the x-axis (subjects), y-axis (average grades), and colour (fill) based on PrepTest status.    * The bars were displayed side-by-side using the geom\_bar() function.    * Titles and axis labels were added with labs() to make the chart easy to interpret. |
| **Analysis and Visualisation:**            Students who completed the PrepTest generally had higher average grades across all three subjects suggesting a positive effect on academic performance. |
| **Considerations:**  Include additional independent variables (e.g., hours spent studying, attendance) to refine the analysis. |