Principles of Data Science

Assignment – 1

Kevin Suchetan Tupili Clement (16354142)

Bar Plot:

The chart reveals the distribution of math scores across different ranges, broken down by gender. Females consistently outperformed males in the higher score ranges (60-70, 70-80, 80-90). Conversely, males show a higher count in the lower score ranges (40-50, 50-60), indicating a possible trend of underperformance in math among male students.

Violin Plot:

The violin plot illustrates the impact of test preparation courses on student performance. Students completing a test preparation course achieved a noticeably higher median total score (220.5) than those who did not (196.0). This difference suggests a potential benefit of test preparation in improving overall scores.

Heat Map:

The heatmap shows parental educational attainment across different racial groups. Group E has a higher proportion of parents with college experience, while Group A shows a relatively lower percentage across most education levels, particularly in associate's and bachelor's degrees.

Box Plot:

The box plot compares the distribution of total student scores across different levels of parental education. Students whose parents hold a master's degree demonstrate the highest median total score, suggesting a potential positive correlation between advanced parental education and student performance. Conversely, students whose parents have only some high school education exhibit the lowest median total score, indicating a possible link between lower parental education levels and lower student achievement.

Histogram

Females show a slightly higher concentration of total scores in the upper range (around 225-275) compared to males, indicating a potential tendency for females to achieve higher overall scores. The male distribution appears to be more spread out with a slightly higher frequency in the lower to mid-range (around 150-225), suggesting a broader range of performance among male students.