STAT 1500 FINAL PROJECT REPORT (GROUP 2)

KAY-SALAMI MOTOLANI KARIMA 202165668

The project advanced smoothly without significant obstacles, thanks to the clarity of the provided materials and the proactive engagement of all team members. Our scheduled meetings were highly productive, with each person contributing their expertise and completing their responsibilities on time. A collaborative environment and effective communication minimized potential challenges, ensuring an overall seamless process.

That said, we did encounter a few minor challenges:

- One issue was deciding as a group whether to focus solely on students' Math scores, Portuguese scores, or both datasets. Although the Portuguese dataset included more students, we anticipated that cleaning it would be stressful and time-intensive, so we chose to combine both datasets.
- I also faced some difficulties in creating a new variable to analyze its correlation with the existing ones. For example, I needed to use variables like *fedu*, *medu*, *mjob*, and *fjob* to approximate family income.
- Finding a convenient time for everyone to meet was another challenge, as we had varying schedules and academic commitments. Despite these hurdles, we ultimately resolved them and successfully moved forward.

Contributions

Our general group was divided into subgroups, with each subgroup assigned specific tasks. Kwesi and Raheem (Group 1) were responsible for analyzing parents' education levels and students' grades, efficiently dividing the work: Raheem developed the regression model, while Kwesi focused on data representation. Chidera and I (Group 3) handled family support and grades. I primarily worked on designing and implementing visualizations for data exploration, including creating box plots, histograms, and other graphical tools to uncover trends and patterns. Additionally, I contributed to the analysis by validating statistical models and reviewing the assumptions of linear regression. I also helped compile the results into a cohesive summary and actively participated in group discussions by sharing key insights. Meanwhile, Chidera focused on data representation and its interpretation. Group 2, consisting of Emilie and Wellington, analyzed the relationship between family quality and grades. Emilie created visualizations and assisted with the analysis, while I worked on the regression model and findings interpretation. Kwesi formatted Group 2's work into a Word document, while Raheem did the same for Group 3's contributions.

As a team, we collaborated on writing the introduction, proofreading the document, and brainstorming recommendations, with each member contributing at least one suggestion Overall, this project was both educational and enjoyable, offering me a deeper understanding of my classmates' diverse perspectives and the challenges of analyzing real-world data.