Findings and Business Implications

Research Question 1:

Question: What is the relationship between the parental level of education and student's average scores across all subjects?

Findings: There's a very weak negative relationship between parental education and student scores. In simpler terms, this means that a parent's level of education doesn't have a significant impact on a student's average scores. The Pearson's correlation is -0.079 with a coefficient of determination of 0.006.

Business Implications: The negative Pearson's correlation coefficient suggests a weak inverse relationship between the parental level of education and the student's average scores, meaning students with parents who have a lower level of education tend to have slightly higher scores. However, the correlation is very weak and the coefficient of determination is close to 0, which indicates that only about 0.6% of the variation in student scores can be attributed to the level of parental education. Hence, parental education level doesn't seem to be a strong predictor of student performance. The management team could focus more on factors within the school's control to improve student outcomes.

Validation: The interpretation of Pearson's correlation and coefficient of determination is well-established in statistics, and we've adhered to those standards in our analysis. However, as with any statistical analysis, there are assumptions that need to be met for these measures to be valid, such as the linearity and homoscedasticity of the relationship. The validity of the answer also depends on the representativeness and quality of the data. An independent replication of the study with a new dataset would be a good way to confirm these findings.

Research Question 2:

Question: Do students who complete the test preparation course have better average scores than those who do not?

Findings: Yes, the data indicates that students who complete the test preparation course generally achieve better scores than those who don't. The T-statistic is 8.390944443482592 and the P-value is 1.633780203592351e-16.

Business Implications: The P-value is extremely small (much less than 0.05), indicating a statistically significant difference in the average scores between students who completed the test preparation course and those who did not. Specifically, students who completed the course

tend to have higher scores. The management team could use this information to encourage more students to complete the test preparation course and improve their academic performance.

Validation: The t-test is a common method for comparing means, and the assumptions for its validity include normality of the data and equal variances. The P-value gives us a measure of the strength of evidence in support of a significant difference. To verify the answer, we could again perform an independent replication of the analysis with a new dataset, or perform additional checks on the assumptions of the t-test.

Research Question 3:

Question: How does the type of lunch (standard vs free/reduced) affect the average student scores?

Findings: There's a significant difference in student scores based on lunch type. Students who receive standard lunch tend to perform better than those receiving free/reduced lunch. The T-statistic is 9.575113051511476 and the P-value is 7.736791812495384e-21.

Business Implications: The P-value is again extremely small, suggesting a significant difference in average scores between students receiving standard lunch and those receiving free/reduced lunch, with those receiving standard lunch typically scoring higher. This could be a proxy for socioeconomic status, implying that students from lower-income families may need additional support to achieve comparable academic outcomes. The school administration might consider implementing programs to mitigate these disparities.

Validation: The process is similar to that for Research Question 2. The interpretation of the t-test and the P-value are standardized, but it's crucial to ensure that the data meet the assumptions of the test. An independent replication of the study would help verify the findings.

Research Question 4:

Question: Are there differences in average student scores based on gender?

Findings: Yes, there is a significant difference in average student scores based on gender, but the data does not specify which gender performs better. The T-statistic is 4.169913529142 and the P-value is 3.311973638243664e-05.

Business Implications: The P-value is small, indicating a significant difference in the average scores between male and female students. Depending on the direction of this difference, the school might need to evaluate its teaching methods or support systems to ensure gender equity in educational outcomes.

Validation: The methodology is the same as in Research Question 2 and 3. Verification of the findings would involve checking the assumptions of the t-test, ensuring the quality of the data, and ideally, conducting an independent replication of the study.

Research Question 5:

Question: Does race/ethnicity have an impact on student scores across all subjects?

Findings: Yes, there is a significant difference in student scores based on race/ethnicity. The F-statistic is 9.096052313390889 and the P-value is 3.2258693161040875e-07.

Business Implications: The P-value is small, suggesting significant differences in average scores among students of different racial/ethnic groups. This highlights the importance of cultural competency in teaching and the potential need for targeted resources and supports to ensure equal opportunities for all students, regardless of their racial/ethnic background.

Validation: For ANOVA, assumptions include normality of the data, equal variances, and independent observations. The P-value from the F-test gives us a measure of the evidence for a significant difference among group means. Verification would involve checking these assumptions, ensuring the quality of the data, and conducting an independent replication of the study.