Okay, Rishi's project presents a classic scenario for comparing the means of two related groups. Here's my advice, explaining the rationale behind the choice of statistical test, as I would during a graduate-level statistics exam:

\*\*1. Identifying the Data Type and Research Question:\*\*

\* \*\*Dependent Variable:\*\* The selling price of textbooks (assumed to be a continuous, ratio-level variable).

\* \*\*Independent Variable:\*\* Bookstore (with two levels: Bookstore 1 and Bookstore 2).

\* \*\*Research Question:\*\* Is there a statistically significant \*difference\* in the average selling prices of textbooks between the two internet bookstores? Crucially, we're not asking if one is \*greater\* than the other, just if they're simply \*different\*

\*\*2. Recognizing the Data Structure: Paired/Dependent Samples:\*\*

The key observation here is that Rishi is collecting prices for the \*same\* textbooks from both bookstores. Each textbook's price at Bookstore 1 is directly related to its price at Bookstore 2. This means the data are \*\*paired\*\* or \*\*dependent\*\*. A single observation will have a price from bookstore one and a price from bookstore two. This pairing is crucial because it accounts for textbook-specific effects (e.g., how popular the textbook is, its condition, etc.) that might influence price, thereby increasing the precision of our test.

\*\*3. Choosing the Appropriate Statistical Test: The Paired Samples t-test\*\*

Given the paired structure of the data and the research question of comparing means, the most appropriate statistical test is the \*\*Paired Samples t-test\*\* (also known as the \*dependent samples t-test\* or the \*matched pairs t-test\*).

\* \*\*Why the Paired Samples t-test?\*\*

\* \*\*Addresses Dependence:\*\* This test is specifically designed to handle paired data. It calculates the \*difference\* in price for each textbook (Price at Bookstore 1 - Price at Bookstore 2). It then tests whether the \*average\* of these difference scores is statistically different from zero. This focuses the analysis on the \*within-pair\* differences, minimizing the impact of textbook-specific variability.

\* \*\*Efficiency and Power:\*\* By analyzing the differences, the paired t-test reduces the within-group variance (the variation within the bookstores), which leads to a more powerful test. This increases the chances of detecting a real difference between the average selling prices if one exists.

\* \*\*Testable Assumptions:\*\* the assumptions that need to be considered before the use of a paired t-test are that the differences follow a normal distribution. This can be assessed through the creation of a histogram of the difference scores, or through a Shapiro-Wilk test.

\* \*\*Why NOT an Independent Samples t-test?\*\*

\* \*\*Incorrect Treatment of Dependence:\*\* An independent samples t-test would treat the prices from the two bookstores as if they were \*independent\* samples. This would ignore the pairing, which would inflate the error and reduce the power of the test. It would also violate the assumption of independence of observations, which is a fundamental requirement for the independent t-test.

\*\*4. Steps for Rishi's Project:\*\*

1. \*\*Data Collection:\*\* Ensure Rishi collects prices for \*the same\* ten textbooks from both bookstores.

2. \*\*Calculate Difference Scores:\*\* For each textbook, subtract the price at Bookstore 2 from the price at Bookstore 1 (or vice versa, consistently).

3. \*\*Check Assumptions:\*\* Visually inspect a histogram or Q-Q plot of the difference scores. If the distribution appears approximately normal (or if the sample size is large enough), the assumptions are likely met. If the sample size is small, a Shapiro-Wilk test can be performed to assess normality, and if assumptions are violated, you could use a non-parametric test like a Wilcoxon signed-rank test.

4. \*\*Conduct the Paired Samples t-test:\*\* Use statistical software (like SPSS, R, or even Excel) to perform the test.

5. \*\*Interpret Results:\*\*

\* Look at the t-statistic and the p-value.

\* If the p-value is less than the significance level (e.g., 0.05), reject the null hypothesis (that there is no difference in average prices). Conclude that there is a statistically significant difference in the average selling prices between the two bookstores.

\* Calculate and report the mean difference and its standard error. Consider calculating a confidence interval for the mean difference to give a range of plausible values.

6. \*\*Report the results in a clear and concise manner.\*\*

\*\*In summary, the Paired Samples t-test is the most appropriate and powerful statistical test for Rishi's project because it properly accounts for the paired nature of the data, allowing for a valid comparison of the average selling prices of textbooks between the two internet bookstores.\*\*