An example we can discuss is comparing the performance of two search algorithms (Algorithm X and Algorithm Y) in finding the shortest path in a graph. Here are the steps in which a statistics hypothesis can be formed:

1. **Formulate the Null Hypothesis (H0) and the Alternative Hypothesis (H1):**

H0: There is no systematic difference in performance between Algorithm X and Algorithm Y in finding the shortest path.

H1: There is a systematic difference in performance between Algorithm X and Algorithm Y in finding the shortest path.

1. **Select a significance level (α):**

Let's choose a significance level of α = 0.05, indicating a 5% chance of observing the results due to random sampling error.

1. **Collect data:**

Gather a representative sample of test cases, such as different graphs with varying sizes and complexities. Run both Algorithm X and Algorithm Y on these test cases, recording the time taken by each algorithm to find the shortest path.

1. **Choose an appropriate statistical test:**

In this scenario, a suitable test could be the paired t-test since we have paired data (the performance of Algorithm X and Algorithm Y on the same set of test cases) and we are interested in comparing their mean performance. The t-test assumes that the data is approximately normally distributed.

1. **Compute the test statistic and p-value:**

Calculate the paired differences between the performance of Algorithm X and Algorithm Y on each test case. Then, use the paired t-test to calculate the test statistic and p-value based on these differences.Let's denote the performance difference between Algorithm X and Algorithm Y on each test case as d\_i, where i ranges from 1 to the total number of test cases. We calculate the mean difference (μ\_d) and the standard deviation of the differences (s\_d). Using these values, we can compute the test statistic (t):

t = (μ\_d - 0) / (s\_d / sqrt(n))

Here, n represents the number of paired samples (test cases). With the test statistic, we can then calculate the p-value associated with the observed test statistic using the t-distribution with (n-1) degrees of freedom.

1. **Compare the p-value with the significance level:**

If the p-value is less than the chosen significance level (α = 0.05), we reject the null hypothesis and conclude that there is a systematic difference in performance between Algorithm X and Algorithm Y in finding the shortest path.

If the p-value is greater than or equal to the significance level (α = 0.05), we fail to reject the null hypothesis and suggest that the observed performance differences could be due to sampling noise.

1. **Report the results:**

Summarize the findings, including the p-value and the test statistic. Additionally, provide any relevant effect size measures, such as the mean difference between Algorithm X and Algorithm Y, to quantify the magnitude of the observed performance differences.