



Can You Count On Weather Forecasts?

TIME SERIES PLOT AND ONE-WAY ANOVA

What You'll Need
[Download the Minitab 17 Free Trial](#)
[Download the data set](#)

Can You Count On Weather Forecasts?

People consult weather forecasts to determine their activities, what to wear, or what to pack for a trip. But are these predictions trustworthy, or just a lot of hot air?

You can use the Assistant in Minitab Statistical Software to find out.

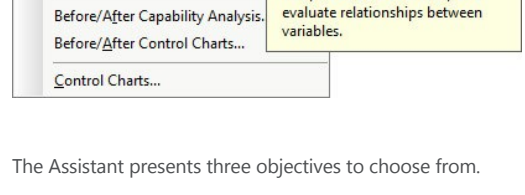
Try It Yourself

The [data set](#) collects 30 days' worth of next-day, 5-day, and 10-day high temperature forecasts (in degrees F) for State College, Pennsylvania—the home of Minitab's World Headquarters. The data sheet also includes the actual high temperature for each day, and the differences between the forecasted and actual high temperatures.

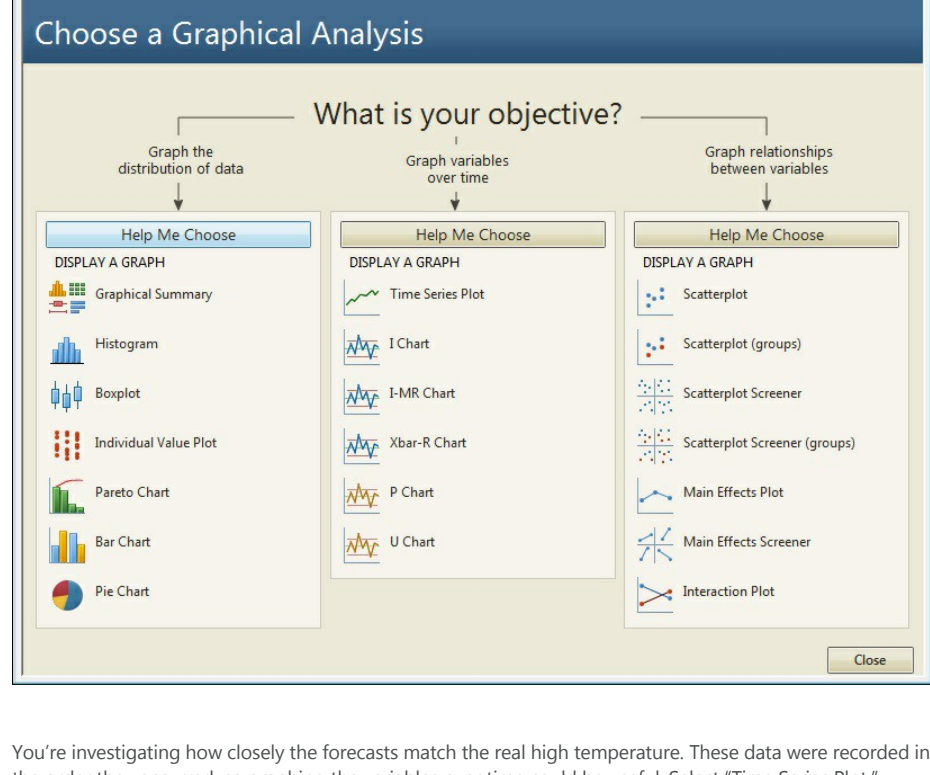
Date	10 day	5 day	next day	high	10-diff	5-diff	next-diff
May 12	69	77	77	77	-8	0	0
May 13	65	70	67	66	-1	4	1
May 14	64	64	63	63	1	1	0
May 15	65	66	66	68	-3	-2	-2
May 16	68	67	61	64	4	3	-3
May 17	69	59	62	61	8	-2	1
May 18	67	66	66	64	3	2	2
May 19	67	66	66	66	1	0	0
May 20	70	70	69	70	0	0	-1
May 21	70	78	77	75	-5	3	2
May 22	76	82	78	72	4	10	6

Step 1: See the Data

Graphing your data is always a good first step. Select **Assistant > Graphical Analysis...** for guidance on what graphs might offer insight about this data.

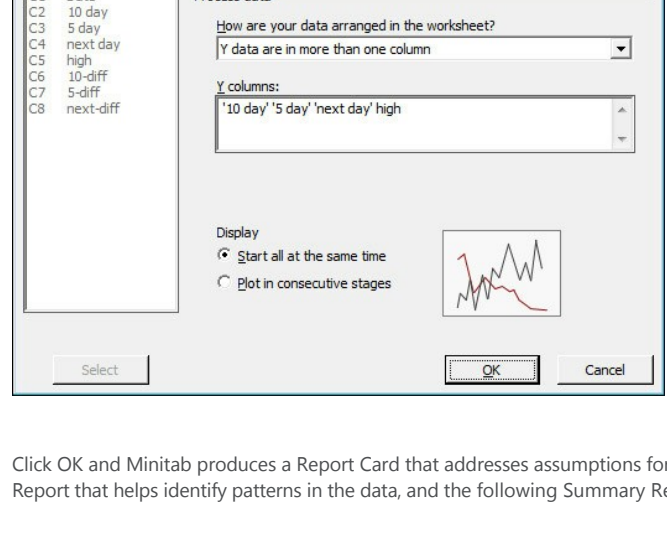


The Assistant presents three objectives to choose from.

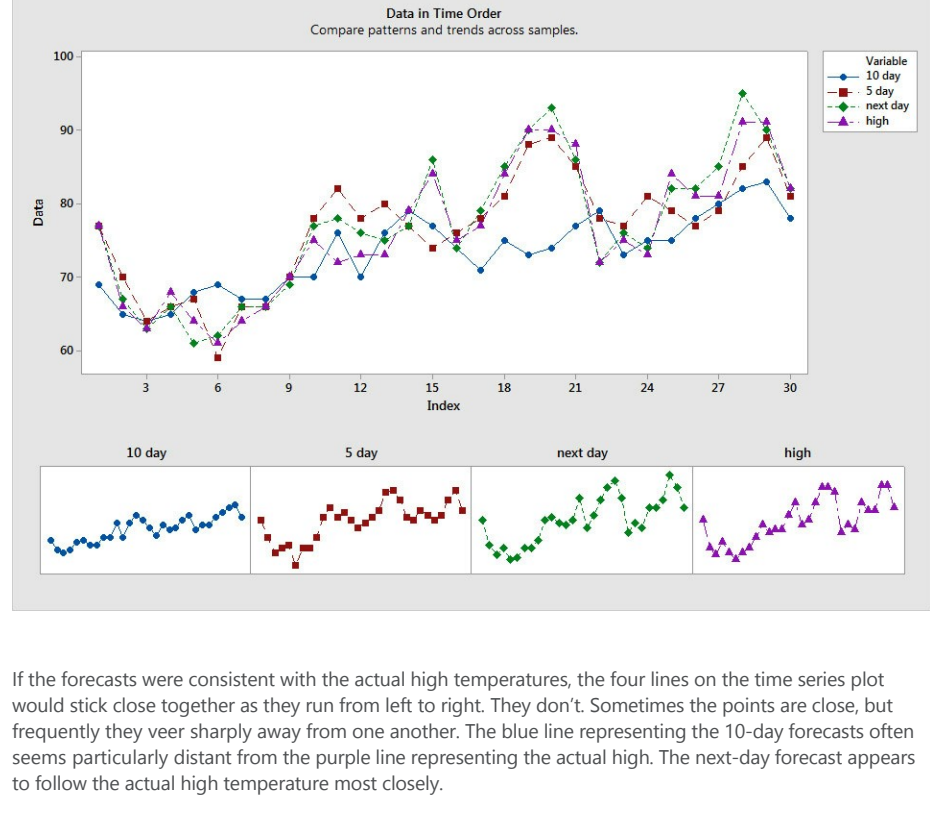


You're investigating how closely the forecasts match the real high temperature. These data were recorded in the order they occurred, so graphing the variables over time could be useful. Select "Time Series Plot."

Since each type of forecast is in its own column, select "Y data are in more than one column." Then enter the columns for each of the three forecasts as well as the one that lists the actual temperature.



Click OK and Minitab produces a Report Card that addresses assumptions for this analysis, a Diagnostic Report that helps identify patterns in the data, and the following Summary Report:

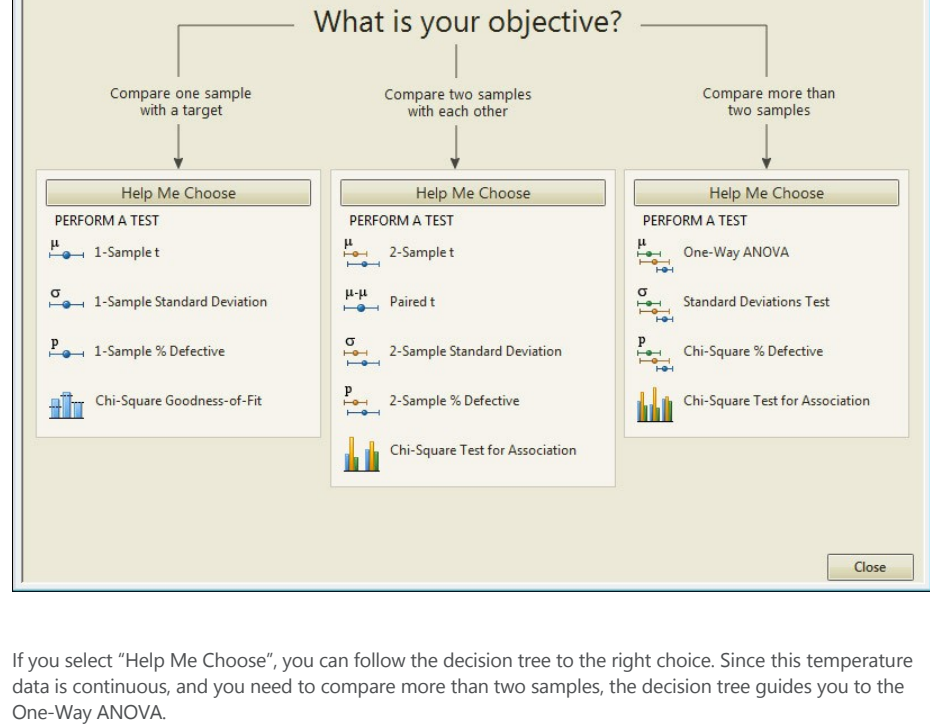


If the forecasts were consistent with the actual high temperatures, the four lines on the time series plot would stick close together as they run from left to right. They don't. Sometimes the points are close, but frequently they veer sharply away from one another. The blue line representing the 10-day forecasts often seems particularly distant from the purple line representing the actual high. The next-day forecast appears to follow the actual high temperature most closely.

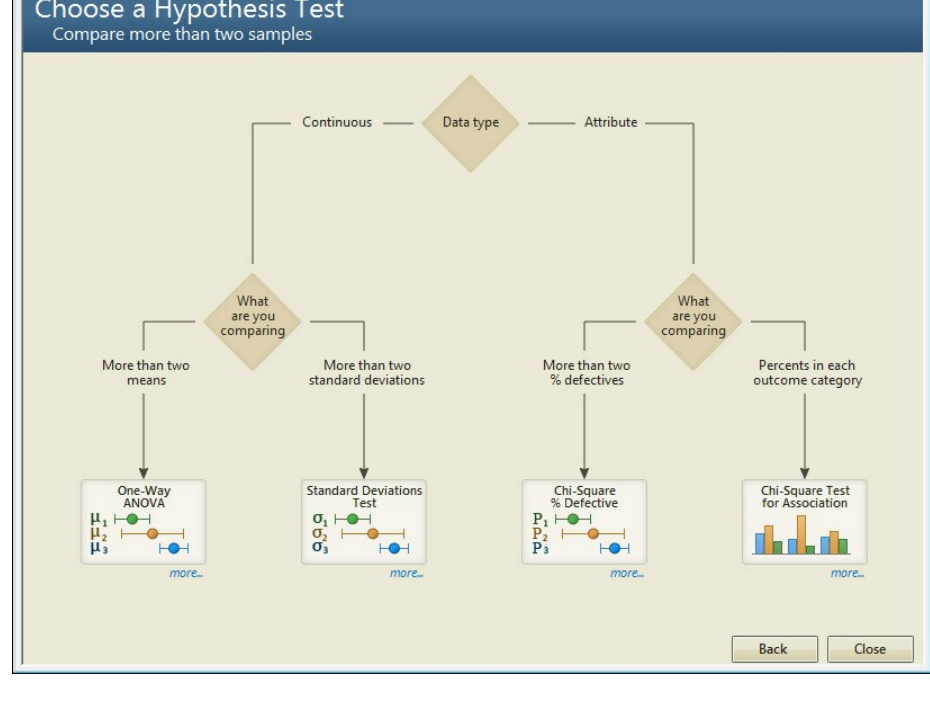
Step 2: Test the Hypothesis

The time series plot suggests that the next day, 5-day and 10-day forecasts are not equally reliable. To compare the three forecasts and see if the data support this hypothesis, use **Assistant > Hypothesis Tests...**

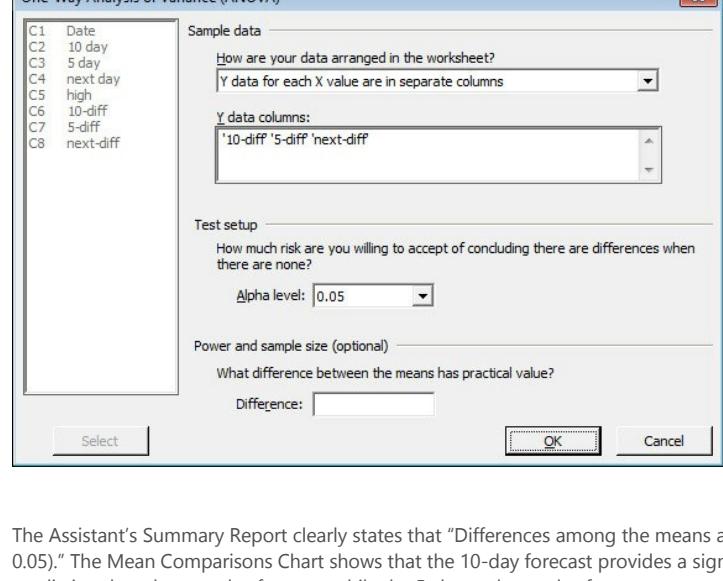
If you're not sure which test to use to analyze the data, let the Assistant guide you to the right choice. Since these data represent three types of forecasts, you want to compare more than two samples.



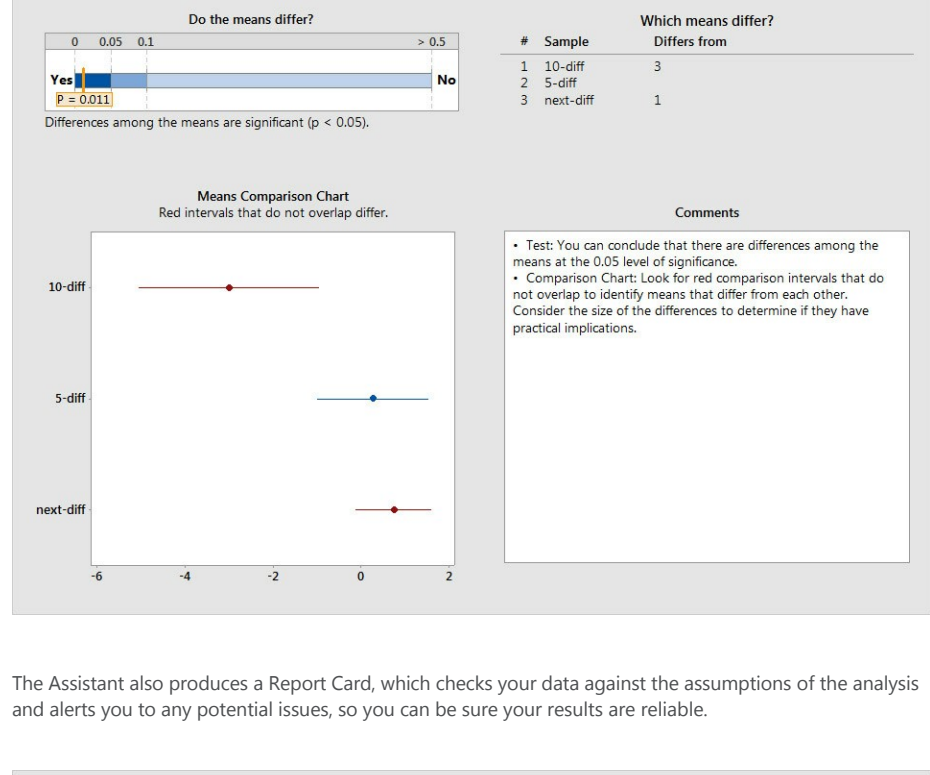
If you select "Help Me Choose", you can follow the decision tree to the right choice. Since this temperature data is continuous, and you need to compare more than two samples, the decision tree guides you to the One-Way ANOVA.



Fill out the dialog for One-Way ANOVA as shown, and click OK:



The Assistant's Summary Report clearly states that "Differences among the means are significant ($p < 0.05$). The Mean Comparisons Chart shows that the 10-day forecast provides a significantly worse weather prediction than the next day forecast, while the 5-day and next-day forecasts appear to be equally accurate.



The Assistant also produces a Report Card, which checks your data against the assumptions of the analysis and alerts you to any potential issues, so you can be sure your results are reliable.

Check	Status	Description
Unusual Data	Warning	One data point (row 15) is unusual compared to the others in 5-diff. Because unusual data can have a strong influence on the results, you should try to identify the cause of its unusual nature. Correct any data entry or measurement errors. Consider removing data that are associated with special causes and repeating the analysis.
Sample Size	Pass	The sample is sufficient to detect differences among the means.
Normality	Pass	Because all your sample sizes are at least 15, normality is not an issue. The test is accurate with nonnormal data when the sample sizes are large enough.
Equal Variance	Info	Minitab's Assistant uses Welch's method, which does not assume or require that the samples have equal variances. Research shows that the test performs well with unequal variances, even when the sample sizes are not equal.

The Report Card has flagged one data point as being unusual. However, a review of this data point shows that it is accurate, so it should remain in the analysis. The Report Card also confirms that the sample size is sufficiently large to detect differences and to satisfy assumptions of normality.

You've Weathered this Analysis!

Your analysis has demonstrated that if you're going to depend on weather reports to decide what to wear, the next-day forecast or even the 5-day forecast are the most reliable options.

Of course, weather is a highly variable, extremely complex process—and this data covers only 30 days in one location. Would the results be the same for 30 days of precipitation for your part of the world? Now you know how to use the Assistant to find out!

What else could you discover by using the Assistant to analyze your data?