



## One-way ANOVA Test

1. Intro and Data Setup
  - a. If you've watched our video that explains the concept of an ANOVA test, you may recall that an ANOVA test helps you figure out when you really do have groups of people who behave or think differently along a particular phenomenon. That is, when people are acting like apples, oranges, pears, nectarines, etc., instead of like different shades of apples.
  - b. In this example, we are going to take a look at average Overall Rating of the festival by Time at Festival, using four groups for Time at Festival – 31-45 minutes, 46-60 minutes, 61-90 minutes, and More than 91 minutes. The Overall Rating question asked festival attendees, "How would you rate today's event overall?" Festival goers could choose an answer on a 5-point scale from Poor to Excellent.
  - c. Here, Time at Festival is the predictor variable – we want to know if the time spent at the festival predicts differences in overall festival rating. Overall Rating is our outcome variable – it is the variable that we want to learn more about.
  - d. The TimeatFestivalbyRating tab in the EvalFest Database for Excel.xlsx file contains the data we will use to conduct the one-way ANOVA test. Column A contains the Overall Rating data for festival-goers who had been at the festival for 31 to 45 minutes, Column B contains the Overall Rating data for festival-goers who had been at the festival for 46 to 60 minutes, and so forth.
2. Conducting the one-way ANOVA test
  - a. You will need to have the Data Analysis ToolPak loaded in Excel. Click on the Data tab, and if you don't see the Data Analysis icon on the top right, you will need to load this ToolPak.
    - i. To add it, go to Tools and click on Excel Add-ins...
    - ii. Check Analysis Toolpak and click OK.
    - iii. When you restart Excel, you should see the Data Analysis icon within the Data tab.
  - b. Click on Data Analysis.
  - c. Click on Anova: Single Factor and click OK.
  - d. Click in the Input Range box and highlight all of the columns containing Overall Rating data. In this example, that is columns A through D.
  - e. Make sure you indicate that the data are Grouped By Columns.
  - f. Check the box next to Labels in first row to indicate that row 1 contains labels.
  - g. Click the radio button next to Output Range and click in the Output Range box.
  - h. Click in the one cell on the worksheet where you would like Excel to begin displaying the output.

- i. Click OK.
3. Interpretation of the one-way ANOVA test
  - a. Excel will provide output – these are your results. Excel provides a Summary table of information as well as the results of the ANOVA test.
  - b. The Summary table shows the size of each group, as well as each group's mean and variance, which is related to the standard deviation. If you've forgotten what the standard deviation is, please visit the Data Analysis: Exploring Your Data section of the EvalFest website. If the group sizes are very different from one another, you should ask yourself if an ANOVA test is appropriate and whether or not to proceed. Here, our groups are fairly close in size, so we can feel confident to continue.
  - c. The ANOVA results table provides information about the 2 types of variance, the between groups variance and the within groups variance. We also see the degrees of freedom, the value of F, and the p-value. You're most interested in the p-value to determine if you have differences between one or more pairs of Time at Festival groups.
  - d. If you've watched the conceptual video, you may recall that the p-value is a measure of the risk you take in making a decision based on your data. If the p-value is less than .05, that means there is less than 5% chance that the results are based on something other than the data. Thinking about this another way, there is a 1 minus the p-value chance that you DO have differences in overall festival ratings among Time at Festival groups, and people are acting like apples, oranges, pears, etc., as opposed to acting more similarly, like different shades of apples. In more statistical terms, you use the p-value to determine if the mean of the outcome variable is statistically significantly different among the groups in the predictor variable.
  - e. For this scenario, we see that our p value is  $3.33 \times 10^{-15}$ , or .000000000000333. That number is less than .05, which means that we have one or more pairs of Time at Festival groups that differ statistically significantly in average overall rating of festival.
  - f. Note that the p-value simply tells you if there are pairs of groups for which the ratings are different – it doesn't tell you which groups differ statistically significantly from one another. To figure that out, we have to run what are called post hoc tests.
4. Bonferroni post hoc test
  - a. Excel doesn't have a way to conduct post hoc tests that compare all of your groups to each other at one time. You will have to conduct a series of separate post hoc independent samples t-tests for each pair, in this case, 6 tests. If you've forgotten how to run an independent samples t-test in Excel, please review the Independent Samples T-test video available on the EvalFest website.
  - b. Remember that the p-value is the risk that your results are actually from chance and not based on a real difference in your data. The value  $p < .05$  is the acceptable risk for one test. If you do more than one test, you have to adjust your risk level to help make sure that your estimates account for the extra level of risk. The more tests you conduct, the greater the possibility of making the wrong decision based on the p-value, so we adjust it for each post hoc test

we do.

- c. In this example, the number of tests we have to do is 6 and so we divide .05 by 6 – this makes it more difficult to get a statistically significant result, but if we do, then we know that we have re-assessed our risk level in a way that gives us confidence in our results. In this case, if the p-value for an independent samples t-test is less than  $0.05/6$ , or 0.008, then you know that the average overall rating between the two groups being tested is different – you have apples and oranges.
  - d. Because the average overall rating increases at each increase in Time at Festival, when we test each pair using an independent samples t-test, we can use the  $P(T \leq t)$  one-tail p-value for comparison to 0.008. The results of the 6 post hoc independent samples t-tests show that there are statistically significant differences in average overall rating between the following groups: 31-45 minutes and 61-90 minutes, 31-45 minutes and More than 91 minutes, 46-60 minutes and More than 91 minutes, and 61-90 minutes and More than 91 minutes. Keep in mind that, just because a result is statistically significant, you still have to decide whether or not the result has *practical implications* for your program. Please visit the Independent Samples T-Test in Excel video for more on this topic.
5. Final interpretation
- a. So how do you think about interpreting these results? Remember that these ratings were made on a five-point scale, where 1 was Poor and 5 was Excellent. The average ratings were above 4.0 for all groups, so somewhere between Very Good and Excellent on the scale. When interpreting these kinds of results, you want to think about the kinds of ratings you have overall and the group differences. So, for these results, we might say something like: Ratings for the Festival were quite positive. As festival-goers spent more time at the festival, average overall rating increased. Specifically, those who had been at the festival more than 91 minutes rated the festival statistically significantly more positively than all other groups.

