Data Analysis Reference Guide

If it's been a while since your last statistics course, or you simply don't consider yourself a "math" person, even the words "data analysis" can be intimidating.

Here at EvalFest, we want to make sure our partners (and the broader community) feel empowered to dig into their own data to find information that will be useful to them.

The EvalFest team has been busy creating videos and handouts to help support you in this process. This guide is meant to be a simple, comprehensive place to figure out the best steps to take in examining your data and where to find resources to guide you through each step of the process.

Step 1: What data do I have access to and what software do I want to use?

What data do I have if I am an EvalFest partner?

- Attendee Survey data (yes!!)
- Attendance data, festival description, festival mission statement (probably)
- Follow-Up Survey data (maybe)
- Mini-grant data (maybe)

At the most basic level you need to determine what data you have access to and the format your data are in. All EvalFest partners have access to their own Attendee Survey data. Some partners also have access to their Follow-Up Survey data or data from any mini-grants they contributed to or lead. The full EvalFest database of 2017 Attendee data and the full database of Follow-Up Survey data from 2018 are also available to the entire community. Your organization might also have data you want to analyze, so remember that the information you learn here can also be used to examine that information as well!

What software do I have access to?

EvalFest offers resources for you to be able to analyze your data in either Excel, SPSS, or R. All three options have their strengths and weaknesses. Deciding which to use will likely involve a balancing act between what you have access to and the kinds of analysis you need to do.

- Excel (almost everybody has it already; limited stats available)
- SPSS (has all the stats you will likely need and a point-click interface; can be cost-prohibitive)
- R (free and has all the stats you will need; steep learning curve)

Step 2: What kind of variables are included in my data and how are they organized?

The next step is to dig into data management by organizing and cleaning your data. This step allows you to (1) understand what data you have and (2) ensure the quality of the data you are analyzing.

Where can I find resources on data management and cleaning?

EvalFest team members have created a video series for partners on this topic. The video series is available on the EvalFest website at: https://evalfest.org/data-management/

If you would like to use your EvalFest data, you can find both the Attendee Survey database and the Follow-Up Survey database on the EvalFest Resources Drive.

Step 3: How can I start to analyze my data?

Now that you have your data organized and cleaned, a good place to begin data analysis is with basic descriptive statistics and exploratory data analysis (EDA).

- Descriptive statistics, in a nutshell, describe your data (imagine that!) through calculating things like mean, median, and standard deviation
- EDA is another way to get summary information about your data, typically done in visual format through charts and graphs

Both descriptive statistics and EDA should be approached with a certain level of caution if you are just getting to know your data. As you walk through these steps, always pause and ask, "Hey, do these numbers make sense to me given what else I already know about the study population/data set/etc.? Is this what I would have expected?" If the answer is "no," that does not necessarily mean your analysis is wrong, but it should send you back to look at your data for any issues.

Where can I find resources to get me started with descriptive statistics and EDA?

- Head over to the EvalFest website to the Data Analysis Resources section
- Open up the handout and the video on EDA for your selected software

We recommend these four steps for using the videos:

1. Read through the handout
2. Watch the training video all the way through
3. Watch the videos again – doing each step yourself with the Practice Data Set
4. If your practice analysis matches that in the video, move on to analyzing your own data

Step 4: How can I figure out what relationships exist in my data between different variables?

Once you have a handle on descriptive statistics, you can move on to other types of data analysis.

Conceptual and "how to" videos are available for the following:

- Chi Square (for exploring the relationship between two categorical variables)
- T-Test (for exploring the group differences between 2 groups on a rating or continuous variable)
- One-Way ANOVA (for exploring group differences between 2+ groups on a rating of continuous variable)
- P-Value (which is not a test, but will help you understand statistical significance)

In a nutshell, these tests can help you determine what type of relationships or associations exist within your data. For example, do men and women rate their festival experience the same way? What about different racial and/or ethnic groups? These tests can help you answer these questions… and more!

Where can I find resources to guide me through analyzing my data?

- Open up the data set of your choice
- Head over to the EvalFest website and access the Data Analysis Resources page
- Follow the links to access the resources for the type of test you want to conduct based upon your software