

Software & Tools

IPA and J-PAL staff members and researchers use a number of software programs to design randomized evaluations and analyze the data from them. This section contains resources – both internally-developed and external ones – to help users use some of this software.

1. Optimal Design
2. Stata
3. R

Optimal Design

Optimal Design is a free program for conducting power calculations, compatible only with Windows operating systems (for advice on running Optimal Design on Macs, see [here](#).) The program and documentation are available for download [here](#).

- [Guide to Doing Power Calculations](#) with examples from Optimal Design Software. This guide explains the trade-offs in designing a well powered randomized trial. It contains exercises and a step-by-step guide that demonstrates how varying parameters can affect the statistical power of a study.

Stata

Stata is a data analysis and statistical software program, compatible with both Windows and Mac operating systems. It

is widely used by economists and a number of technical resources developed at IPA and J-PAL have been designed with Stata users in mind. For more information, see [here](#).

- All versions of Stata come with a user guide and manuals—working through the user guide is a very efficient way to learn the basics of Stata
- [Writing Randomization Code in Stata: A Guide](#). This step-by-step guide uses data and an annotated Stata do-file to illustrate how a simple randomization can be carried out using Stata.
- [Power calculations in Stata: A Guide](#). This step-by-step guide uses data and an annotated Stata do-file to illustrate how power calculations can be carried out using Stata. It provides an example of both a conventional parametric and a non-parametric simulation method of calculating power.
- The UCLA Institute for Digital Research and Education [provides a number of valuable resources on learning and using Stata](#), organized by topic. The search functionality allows for looking up resources on specific commands e.g. [this page](#) on using the “egen” command to recode continuous variables into groups.
- Ray Kluender and Ben Marx have a number of hands-on [tips and suggestions](#) for working with data in Stata.
- Germán Rodríguez has a short introduction to Stata for new users, with an emphasis on data management and graphics, available [here](#).
- Data Science Central Data Science Central has created a few "[cheat sheets](#)" on using Stata for data science tasks and analysis. These will be of interest to both novice and advanced Stata users.
- The [randtreat command](#) performs random treatment assignment. It can handle an arbitrary number of treatments and uneven treatment fractions, which are

common in real-world randomized control trials. It also provides several methods to deal with 'misfits', a practical issue that arises in treatment assignment whenever observations can't be neatly distributed among treatments.

- [orth_out](#), a Stata program written by an IPA staff member for producing summary stats tables and orthogonality tables.
- [cfout](#), a Stata program written by an IPA staff member for reconciling multiple rounds of data entry.
- [bcstats](#), a Stata program written by an IPA staff member for conducting back checks on survey data.
- [odkmeta](#), a Stata program written by an IPA staff member to facilitate the easy import of survey data collected using the ODK platform, into Stata.

IPA has developed a number of self-guided Stata learning modules designed for different levels, from beginners to more advanced users:

- [The Stata 101 module](#) is designed for users with little or no knowledge of Stata
- [The Stata 102 module](#) is designed for users with some Stata experience, but who are not especially comfortable with the program.
- [The Stata 103 module](#) is designed for users who are familiar with the basic Stata commands and are comfortable working with the program.
- [The Stata 104 module](#) is designed for advanced users, with a focus on data cleaning.

R

R is a free software environment for statistical computing and graphics, compatible with both Windows and Mac operating systems. For more information, see [here](#).

These are many great resources that you can find online for learning R. We provide the following list of resources along with their descriptions to help you in your path of learning this open-source and very powerful programming language and statistical software:

- [Base R Cheat Sheet](#) - A cheat sheet of basic R commands.
- [R for Statistics 571 by Bret Larget](#) - A good compliment to this course in that it is fairly short, it teaches the basics of R, and allow you to learn of the statistical applications of R such as statistical tests, bootstrapping methods, and data visualization.
- [R-bloggers](#) - This website compiles the blogs of many data analyst who share the cool work they do with R! This source is good for seeing the potentials that R has for data analysis and visualization. You will see a diverse set of interesting projects that are possible with many of the techniques that you learn in this course
- [UCLA Statistical Consulting Group](#) - This website provides excellent references and a library of resources for R. The collection of resources and tutorials can be somewhat advanced, but it is a good reference to keep for when doing projects.
- [Randomization Inference \(RI\)](#) is an R package for performing randomization-based inference for experiments.
- Jorge Cimentada has written [an R version](#) of the J-PAL guide to writing randomization code in Stata.

Finally, we recommend the following courses, which are a part of the [MIT Economics MicroMasters program](#):

- [14.310x: Data Analysis for Social Scientists](#) - The course uses R to discuss methods for harnessing data to answer questions of economic and policy interest
- [JPAL 102x: Designing and Running Randomized](#)

Evaluations – The course provides templates for calculating power, monitoring data collection and data management on R. These resources will be updated on our website soon.

Please note that the practical research resources referenced here were curated for specific research and training needs and are made available for informational purposes only. Please [email us](#) for more information.

Published on *The Abdul Latif Jameel Poverty Action Lab*
(<https://www.povertyactionlab.org>)

<https://www.povertyactionlab.org/research-resources/software-and-tools>