Apophenia

We'll now talk about Apophenia, which is used for statistical computations.

We'll cover the following

The Apophenia library is an open source library for working with data sets, statistics, and statistical models.

The Apophenia library provides functions for dealing with data (reading it in, storing it in convenient data structures, and writing it to files). It also provides functions for dealing with data using the SQL database language. This allows you to perform searches on data, extract subsets of data, etc.

Apophenia has lots of functions to fit data to statistical models such as OLS models (ordinary least squares), it provides for statistical tests like t-tests, F-tests, and it includes models like logit, probit, and multinomial models. It also includes maximum likelihood methods, Bayesian updating, and resampling methods like bootstrapping.

There is extensive online documentation and example code on the Apophenia website here. I suggest you dive deeper on your own if you want to explore it in more detail.

Modeling with Data (book)

Apophenia is maintained by Ben Klemens, who also wrote a book Modeling With Data that makes use of Apophenia. You can buy the book or you can download it for free, the author makes it available on the website here.

I highly recommend the book not only as a guide to using the Apophenia library, but more generally as a guide to how to use C on a daily basis for data analysis and modeling. Ben argues that really, C is not **that** much more of a pain in the neck than “higher-level” languages like Python and R, but C is **way** faster (in terms of execution speed). It’s worth a read.

execution speed). It's worth a read.

The last on our list is GNUplot, which as the name suggests, is used to make graphs and figures.