END OF FUNDING PERIOD REPORT
LAUC Statewide Grants, Research, & Professional Development Committee

Primary Applicant(s): Harrison Dekker
Campus: UC Berkeley
EmailL hdekker@berkeley.edu
Telephone: 510-250-7901

TITLE OF PROJECT: An Introduction to R

Time Period of Grant: 2013 Fiscal
Amount of Award Received: $500
Original Abstract as Submitted:

An Introduction to R
While it's likely that spreadsheets, with their relative ease of use and "universal" availability, will continue to be the preferred method of data analysis for the library community, compelling reasons exist for analysts to use more robust analytical tools. One such tool is R, an open-source statistical computing environment, created by and for data analysts, widely used across disciplines as disparate as bioinformatics and finance. In it's most basic form, R is a programming language designed and optimized for data manipulation, analysis, and visualization. A key feature of R platform is a "package" system that allows users to easily share code, data, and documentation, usually via established community-supported archives like CRAN, R-Forge, and Bioconductor. Thousands of high quality, well-documented packages are currently available and the number continues to grow.

The workshop will be of interest to those who routinely work with numeric data and who are interested in exploring new forms of analysis and visualization or who are interested in learning how a programmatic approach makes it easier to scale, reproduce, or re-use their analysis. The workshop should also appeal to those who are simply interested in becoming more computationally and data literate, skillsets that will be critical as libraries begin to explore new opportunities in the e-sciences and other data intensive areas.

This workshop will provide participants an opportunity to learn the basics of the R language and to learn more about the R "ecosystem". There will be hands on exercises using actual library data and all code samples and datasets used will be made available to participants. Prospective attendees are encouraged to download and install R, prior to the workshop and to work through some tutorials, such as those found on http://www.statmethods.net/ or http://www.r-bloggers.com/. For more background and to download R, visit http://cran.org.

I. ACCOMPLISHMENTS AND EVALUATION

My presentation was in the form of an all-day (6 hour) bootcamp at the 2012 ELAG (European Library Automation Group) Conference 2012 in Palma de Mallorca, Spain. ELAG is "Europe's premier conference on the application of information technology in libraries and documentation centres". It is a technical conference intended for professionals with appropriate technical experience. Accordingly, my workshop required previous experience with statistical programming software or a high level programming language. Sixteen participants attended the session. I am not aware of any formal evaluation conducted by the conference conveners.

II. IS YOUR PROJECT COMPLETED? Yes_X_   No__

III. FINANCIAL STATEMENT

Airfare and other travel expenses: $1462.07
Meals: $92.29
Lodging: $194.02
Registration: waived

Total: $1748.38

IV. SHARING YOUR PRODUCT/EVALUATION
Syllabus and code and datasets are available at the following url:
http://goo.gl/oWjN3e

V. NOTE
Information included in this report may be reprinted or posted on the web for dissemination to UCOP, other UC Libraries, and future potential LAUC grant applicants.