Practical Data Visualization with R

When: **October 3, 2015**
Where: Constant Contact
1601 Trapelo Rd., Waltham, MA

Cost: $209 through September 10
$269 Sept. 11 - Sept. 22
$309 Sept. 23 - Oct. 2

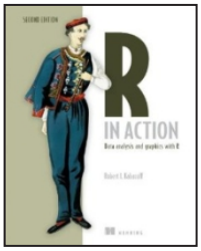
Speakers

Dr. Kabacoff is a seasoned researcher, with 30 years of experience in data analysis and data visualization.

As Vice President of Research for Management Research Group (1997-present), he consults widely with academic, government, and corporate organizations throughout North America, Western Europe, and the Pacific Rim.

As a Professor in the Center for Psychological Studies at Nova Southeastern University (1987-1997), he taught numerous graduate courses on multivariate statistics, statistical consulting, and research computing.

The second edition of his popular book R in Action: Data Analysis and Graphics with R, is now available



Back by Popular Demand!

In 2014 Robert Kabacoff delivered a GBC/ACM seminar on R. That seminar was a great introduction to R, focusing on learning R and applying it to Big Data Analytics. The attendees wanted more - specifically more on using R for data visualization, graphics, and presentations. Based on this demand, Dr. Kabacoff has agreed to deliver a follow-on workshop dedicated to visualization.

Visualization in R

Data visualization has become a central feature of modern data analysis. The R platform provides one of the most flexible and powerful platform for graphing data, understanding and evaluating statistical models, and communicating results to others.

This day-long workshop will provide practical review of R's major graphing capabilities, including its base functionality and exciting new capabilities provided by add on packages (including ggplot2, ggvis, and rCharts).

If you are looking at this workshop, you probably have some data that you need to collect, summarize, transform, explore, model, visualize, or present. If so, then R is for you! R has become the world-wide language for statistics, predictive analytics, and data visualization. It offers the widest range of methodologies for understanding data, from the most basic to the most complex and bleeding edge.

SEMINAR TOPICS

This workshop will provide a practical introduction to data visualization using R. The workshop will be highly practical and interactive. Participants are encouraged to bring a laptop and work through each of the examples as they are described.

Course Outline

I. Introduction – R basics; importing data; the base graphics system; creating univariate and multivariate plots; customizing plots (fonts, axes, grids, titles, labels, colors, symbols, legends); saving plots in various formats (e.g., pdf, svg, png, emf, jpeg).

II. Using ggplot2 – Creating sophisticated charts with the ggplot2 package; exploring complex data, customizing ggplot2 graphs.

III. Interactive Graphics – Visualizing data interactively; using the advanced capabilities of rCharts, googleVis and other packages. Creating visualizations for reports and websites.

Why you should attend this seminar

Robert Kabacoff notes: *"I think that there are two reasons why R can be challenging to learn quickly."*

First, while there are many introductory tutorials, none alone are comprehensive. In part, this is because much of the advanced functionality of R comes from hundreds of user contributed packages. Hunting for what you want can be time consuming, and it can be hard to get a clear overview of what procedures are available.

The second reason is more ephemeral. As users of statistical packages, we tend to run one prescribed procedure for each type of analysis. We carefully set up the run with all the parameters and options that we need. When we run the procedure, the resulting output may be a hundred pages long. We then sift through this output pulling out what we need and discarding the rest.

The paradigm in R is different. Rather than setting up a complete analysis at once, the process is highly interactive. You run a command, take the results and process it through another command (say a set of diagnostic plots), take those results and process it through another command (say cross-validation), etc. The cycle may include transforming the data, and looping back through the whole process again. You stop when you have fully analyzed the data.

presented by the Greater Boston Chapter of the ACM

Details and registration: www.gbcacm.org