Title:
Publication-quality data visualizations using the R tidyverse

Instructors: Jacob Lahne, Virginia Tech, USA
Kyle Hamilton, Flavor Linguist LLC, USA
Elizabeth Clark, McCormick & Company, USA

Even medium-sized sensory studies often yield large amounts of noisy data with many variables. Good visualizations can condense these data so patterns are visible at a glance, but it's all too easy to make plots that mislead or confuse. The free and flexible R programming language allows researchers to quickly generate and iterate on highly-customizable visualizations suitable for presentations and publications, but the number of possibilities can quickly become overwhelming. In this tutorial, we aim to teach participants how to make useful visualizations of highly multivariate data in R. The R skills covered, including data import and visualization, are necessary for making most plots.

Participants will get hands-on experience coding in R, gaining familiarity with flexible techniques and packages. This tutorial will demonstrate the process and best practices for creating side-by-side plots for viewing many variables at once as well as "multivariate maps" from dimensionality-reducing analyses like Principal Components Analysis or Correspondence Analysis. It is not intended as a statistics course and will not dive deeply into specific analyses, but by the end participants will visualize Check-All-That-Apply data using multiple simultaneously-generated bar plots and Correspondence Analysis biplots. These methods allow participants to visually represent ordinal or categorical data about stimuli, and so could also be used on data from open comment questions, existing texts, or many other kinds of sensory surveys.

In this tutorial, we will introduce the audience to *ggplot2* with the aim of developing sufficient basic skills to visualize multivariate sensory and consumer data. We will provide a learning dataset for the analysis—a set of free response comments and overall liking scores from a central location test on berries. We will teach participants how to import and plot data using user-friendly, *tidyverse* R programming. All resources used in the tutorial are open-source and will remain available to attendees, including an R script covering the full workflow. This is **not** a beginner's introduction to R or to the *tidyverse*. Attendees will likely find the material most useful as a complement to existing data manipulation skills in R, but pre-workshop materials on the R *tidyverse* will be sent out in advance for those wishing to catch up.

At the end of the tutorial, attendees will be able to prepare raw sensory data for common multivariate visual representations in R.

Duration	3 hours
Audience	Sensory and consumer scientists with basic R/RStudio familiarity and who want to make better data visualizations
Background	Basic familiarity with data types, variables, functions, and installing/using packages in R/RStudio. Basic understanding of statistics is helpful but not required. We will email registered participants before the workshop with some basic setup requirements (R/RStudio software installation) and suggested "catch-up" reading for participants worried about their level of existing R knowledge.
Laptop	This is a coding workshop, and so we ask all participants to bring a laptop with access to RStudio. We will ask for minimal pre-work (installation of R/RStudio). You may need IT to install RStudio for you on a work laptop.