**Speaker:** Dr. Mani Lakshminarayanan

**Affiliation:** Pfizer, Incorporated. Collegeville, Pennsylvania, USA

**Time:** 3pm - 4pm, 5th September, Friday.

**Venue:** SETS Auditorium, ISI-Chennai.

**Title:** Grammar of Graphics: A Visual Data Excursion in R

Abstract: Statisticians or any quantitative analysts are interested in terms such as “information,” “knowledge,” and “data”. These terms are critical in understanding the concept of “Big Data” that’s currently sweeping the scientific community. By definition, data refers to structured information in some schematic form such as a table or a list, and knowledge is derived from studying and synthesizing the information hidden in the data. Data is often but not always quantitative, and it is often derived by processing unstructured information. When we visualize data, we are interested in portraying abstract relationships among variables that contain the unstructured information. Visualization has been used for centuries to map our world (cartography) and describe the animal and plant kingdoms (scientific illustration). Data visualization, which is more abstract, emerged more recently. An early innovator was William Playfair, whose extensive charting of economic data in the 1800s (Wainer & Spence 2005a, Wainer & Spence 2005b) contributed to its emergence. The early history of visualization has been richly — and beautifully— documented (Friendly & Denis 2006, Tufte 1983, Tufte 1990, Ford 1992, Wainer 2000) (Ref: Dianne Cook and Deborah Swayne, 2007).

Grammar of Graphics, a term originally coined by Leland Wilkinson (2005), is a concise tool that enables us to describe the components of a graph that is the backbone of any data visualization. A recent R package, ggplot2, implements Wilkinson’s Grammar of Graphics. Written by Hadley Wickham, it was set out a set of general unifying principles for the data visualization. For this reason, ggplot2 offers a more elegant and arguably more natural approach than does the base R graphics package.

In this presentation, we will first introduce the grammar of graphics, followed by several basic and advanced examples, which will describe the features of ggplot2. We will also contrast the features in ggplot2 with other R tools such as ‘plot’ and ‘lattice’. Our primary goal in this presentation is to make the audience comfortable in using the freeware R and expose them to what it offers to the World of Graphics that is easy to use with abundant flexibility.

Key Words: Data Visualization; R Graphics; Statistical Summaries; Big Data.