



Seminario Aleatorio

Sesión 299

Data science with the tidyverse

Hadley Wickham
Chief Scientist, RStudio.

Resumen

My goal is to create an environment for data science where you can spend your precious mental energy on the problem at hand, rather than fighting with a programming language. To this end, I've developed a number of packages that make individual parts of the process easier (ggplot2 for visualisation, dplyr for data manipulation, tidyr for data tidying, ...). Recently I've been thinking more about how the pieces fit together. In the words of Hal Abelson, "No matter how complex and polished the individual operations are, it is often the quality of the glue that most directly determines the power of the system." In this talk, I'll discuss the idea of the tidyverse, a set of conventions that ties together disparate packages to provide a uniform interface for doing data science. Along the way you'll learn about tidy data, tibbles, list columns, pure functions, referential integrity, piping, and why ggplot2 should never have existed.

Viernes 4 de noviembre de 2016, 13:00 hrs.
Salón por Confirmar, Plantel Río Hondo

El Seminario Aleatorio está destinado tanto a profesores como a estudiantes, por lo que el Departamento de Estadística agradece a los profesores que colaboren invitando a sus alumnos a estas sesiones.

En la red: <http://estadistica.itam.mx/es/51/contenido/seminario-aleatorio-de-estadistica>