



Contribution ID: 6

Type: **not specified**

## Exploring EDA

*Monday 23 February 2015 13:30 (1 hour)*

An introduction to thinking in plots and searching for patterns. Understanding distributions, the interplay between variables of a single distribution and how to look at multiple variables. Principles explained with example data set and visualisation in R programming language.

Targeted audience: Anyone wishing to learn more about how to look at data. Data scientists in this context can be either physicists or computer scientists or people involved in marketing or recruitment.

Benefits of attending the lecture: Demonstration of the importance of understanding the nature of the data that were working with before beginning to analyse it. Examples of various software for data processing and visualisation will also be given.

Prerequisites: This lecture will be self contained and easy to follow for anyone with some experience in basic computing and statistics.

**Presenter:** CROFT, Vincent Alexander (NIKHEF (NL))

**Session Classification:** Exploring EDA, Clustering and Data Preprocessing