

# VP - Data Analysis toolbox

Mauro Donegà - Severian Gvasaliya  
ETH - Zürich

# Why these lectures

This is not a full fledged statistics course !

If interested there are several good courses at ETH.

The goal is to provide you with the analysis tools you need to properly extract information from the data you take in the experiments.

For this reason we will typically skip formal derivations and focus on the concepts. The goal is “understand what you are doing when you analyse your data!”

“The first principle is that you must not fool yourself and you are the easiest person to fool.”

*Richard P. Feynman*

The course is expected to last only the first half of the semester.

If any of you wants to go deeper in the subject, we are willing to address those question in the second part of the semester.

# Contents

- data representation: plots, histograms
- statistical and systematic uncertainties
- error estimation and propagation
- characterise datasets: mean, variance, etc..
- probability density functions
- covariance and correlation matrices
- parameters estimators and their properties (fits):
  - Least Squares Fits
  - Maximum Likelihood Fits
- goodness of fit
- how to generate simple Monte Carlo samples

# Organization

Lectures : 15:45 - 16:30 Monday (in English)

Lecturers: Mauro Donegà, Seva Gvasalyia

Room: HPV G 5

Exercises: 16:45 – 17:30 Mondays (in English)

Five Teaching Assistants = five groups :

Puneet Murthy, Michael Reichmann, Daniele Ruini, Diego Sanz, Matthew Singleton

Rooms: HIT F 11.1 » HIT J 51 » HIT J 53 » HIT K 51 » HIT K 52 »

Students sign up for a group on moodle (each room limited to ~24)

Please signup for HIT F11.1 if you don't have a laptop

*(first exercise session troubleshoot the anaconda installation)*

All slides / exercises are on moodle:

<https://moodle-app2.let.ethz.ch/course/view.php?id=11448>

# Organization

Exercise scheduling:

[New exercises](#) provided at the latest the Friday before class on Moodle (except first week)

[Solutions](#) provided the Friday after the class and Moodle;

[Students are invited to hand in the exercises.](#) Those will be checked/graded to give feedback to the student but will *not* contribute in any way to the final pass/fail which will be entirely determined by the experiments+reports

We will probably setup a [questionnaire to get feedback from you](#) on the lectures/exercises, but before that please talk to us ! We look forward to your comments/suggestions.

# Bibliography

Introduction to error analysis, J.R. Taylor

University Science books: ISBN-10: 093570275X

Statistics for nuclear and particle physics, L.Lyons,

Cambridge University Press; ISBN-10: 0521379342

Statistical data analysis, G. Cowan,

Oxford University Press; ISBN-10: 0198501552

*We will point to precise bibliography at the end of each lecture*

