

# Advanced Methods in Data Analysis

## Outline of the course (labs):

1. Statistics and Data Analysis
2. Multivariate Techniques and Machine Learning
3. Physics Modeling, Simulation and Monte Carlo Methods
4. Regression, Classification, Clustering and Retrieval

First three parts will focus on applications in physics (mostly in High Energy Physics)

The last part will discuss more typical „Data Science“ problems and solutions.

# LABS: how we will operate it

- Labs is only 45min, so it will be not a time for you to write a code.**
- We will go through content of assignments, present results of analyses and observations, have short oral presentations.**
- It will be also occasion to share with everybody problems, exchange snippets of the code or interesting observations.**

# Getting your ETCs for labs

**This is not a course of programming, but you will be expected to write programs.**

- Basic choices are: C++/Root or Python + Anaconda libraries.
- You can use also R or other Data Science specific programming language/library

**I will not be teaching you programming or helping to debug your code, you are on your own ...**

**For labs you will be graded with:**

- completed assignments
- personalised project
- short topical presentations

**Graded will be not (necessarily) quality of a code, but maturity of how you analyse and interpret the data.**

**To pass the course you need to collect at least **60 scores**.**

# Assignments, Projects, Short presentations

## PEGAZ system:

**This system we will use to collect your assignments/projects/short presentations**

- I will be sending you back comments
- **You will see your grades there**

**Please don't use email to send me your scripts!**