The CMS Data Analysis School experience
Nicola De Filippis
Dipartimento Interateneo di Fisica “M. Merlin”, Politecnico and INFN Bari, Italy
on behalf of the CMS Collaboration

The CMSDAS are the official schools for learning about CMS Data Analysis; they are coordinated by the CMS Schools Committee. Originated in 2010 at the LHC Physics Center (LPC) in Fermilab, CMS typically has three schools every year in Europe, Asia, and the US. So far it has been conducted at Fermilab, Pisa, CERN, DESY, Bari, Taipei, Kolkata, and Daegu.

Goals of the CMSDAS
- teach master students, Ph.D. and young postdocs for the data analysis with the CMS software via hands-on tutorials
- train them about timing and competition in doing their work
- expose to them about how to write a paper and the steps towards a publication in a scientific journal
- Provide venues for discussions: educational, professional and social

Plenary lectures
- The "big picture" in physics:
  - theory
  - CMS experimental physics results and challenges
  - LHC status, results and plans
  - CMS detector status, results and plans
  - CMS Software and computing tools

Operations
- Half of the first day devoted to plenary lectures on physics, detector and software tools
- two days of 9-12 "Short" exercises about objects reconstruction/identification/trigger/generators/statistics
- few hours for the writer publication training
- two and a half days of 8-10 "Long" exercises about physics analysis from official CMS physics groups (HIG, EXO, SUSY, SMP, TOP …)
- team of 6-8 students formed
- bunch of "facilitators" between the CMS experts selected to train the students
- last day "mini-symposium": competition between the analysis teams for the "Best Analysis Team" Prize

Preparation steps
- coordination between the host institute and the international CMS advisory committee for the school foreseen
- the time of the school is chosen carefully to avoid overlap with or closely preceding major physics conferences
- host institute must ensure a computing infrastructure be able to deal with hundreds of analysis jobs running in parallel and to host few TB of data
- simultaneous availability of multiple conference rooms equipped with wi-fi / cable connection / blackboard
- preparatory exercises on computing and software to be executed and answered by the students before the beginning of the school
- wikis for short and long exercises to be prepared well before the beginning of the school to be tested by the people from the host institute
- each student can choose only one "long" exercise and then they are mapped to at least 3 "short exercises" pointed by the facilitators for the assigned long exercise.

Results from surveys
- Participation level (mostly good students)
- Time duration of the school. Is that right?
  - A. too long: 1 (3.0%)
  - B. too short: 12 (37.0%)
  - C. just right: 20 (60.0%)

Results and conclusions:
- The analyses teams at the school were able to:
  - reproduce the latest public results by the collaboration
  - extend the state of the art on CMS either by using a larger dataset or by modifying selection to improve sensitivity
  - About 1000 users trained so far at 15 schools and more to follow worldwide
- After the CMSDAS training the students jumped immediately to start on their own physics analysis which otherwise it would take months of preparation

The CMSDAS training program is a successful model, proven to be a key for the preparation of new and young physicists