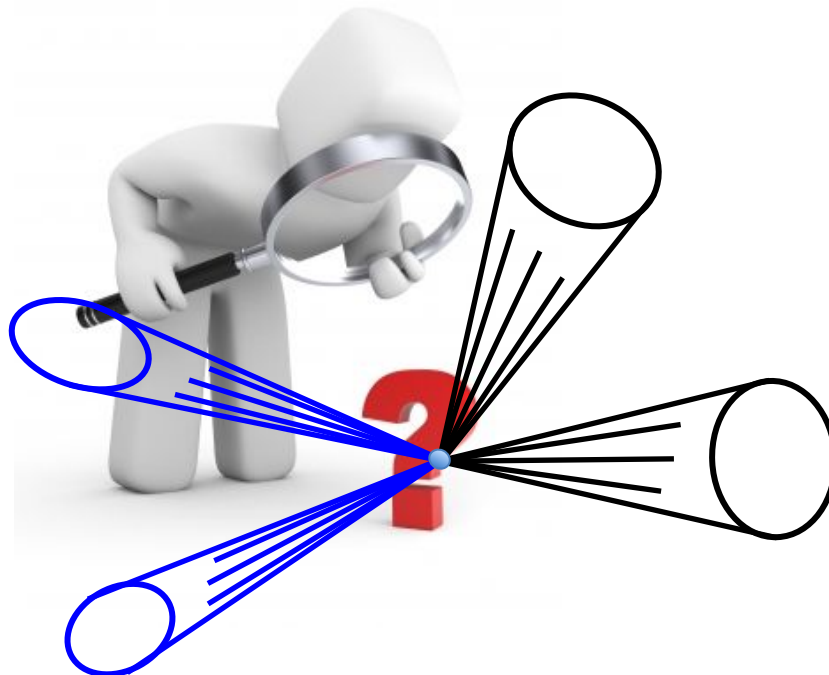# ATLAS Data and Tools Release for Outreach and Education
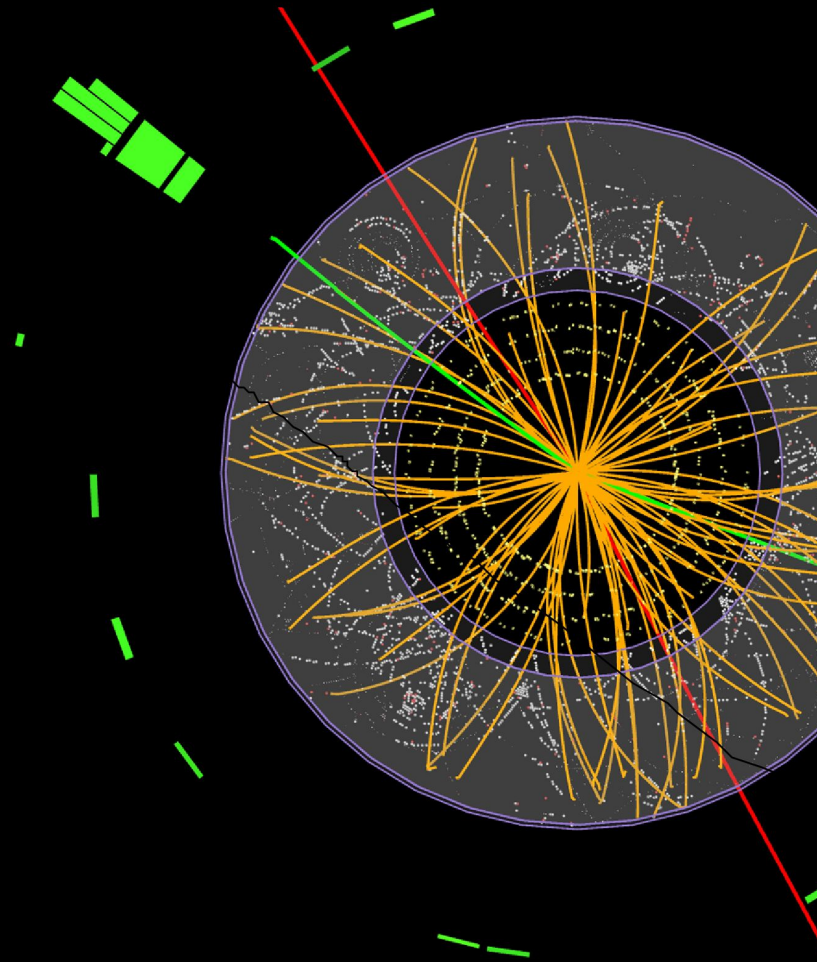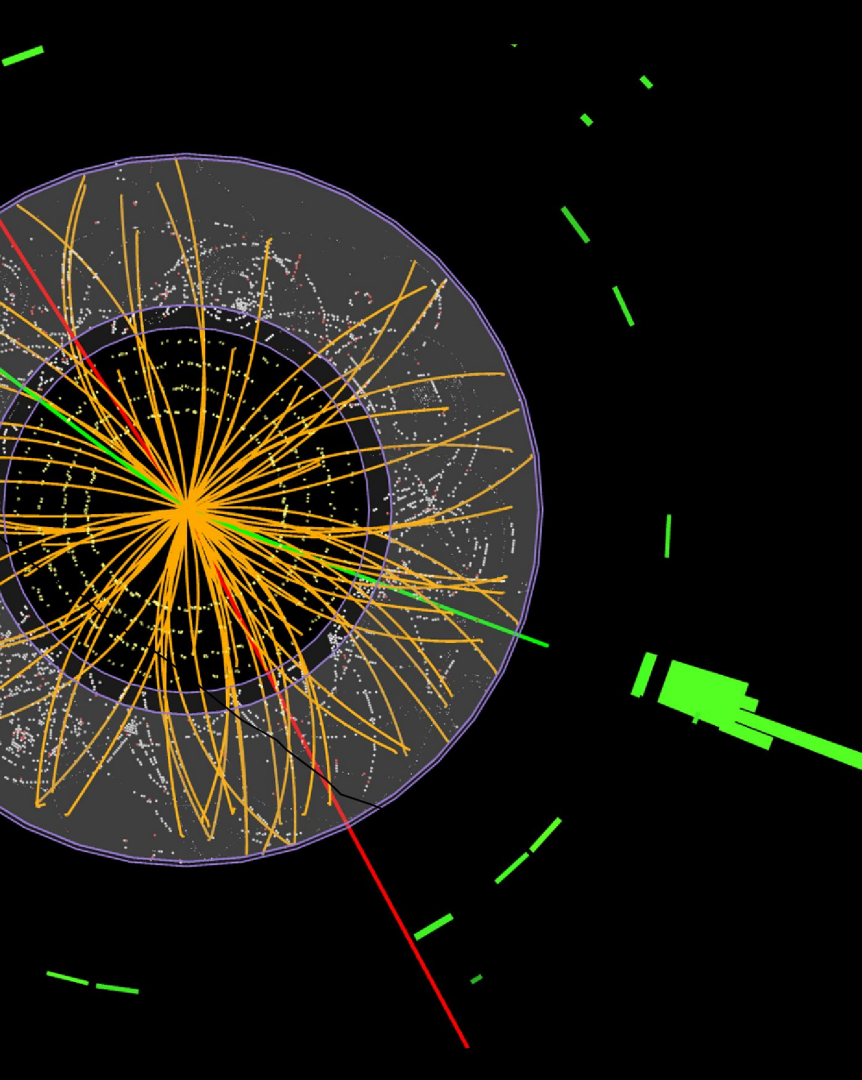
Arturo Sánchez, on behalf the ATLAS Data and Tools Group, ATLAS Outreach

University of Naples and INFN

ATLAS Weekly Meeting, **May 31st, 2016**

The ATLAS collaboration is releasing 1 fb⁻¹ of 8 TeV data and MC for educational use this Summer 2016

...today we'll show the main objectives, including the creation, use and deployment of analysis *tools*

3

ATLAS
Outreach group

1. Data Release
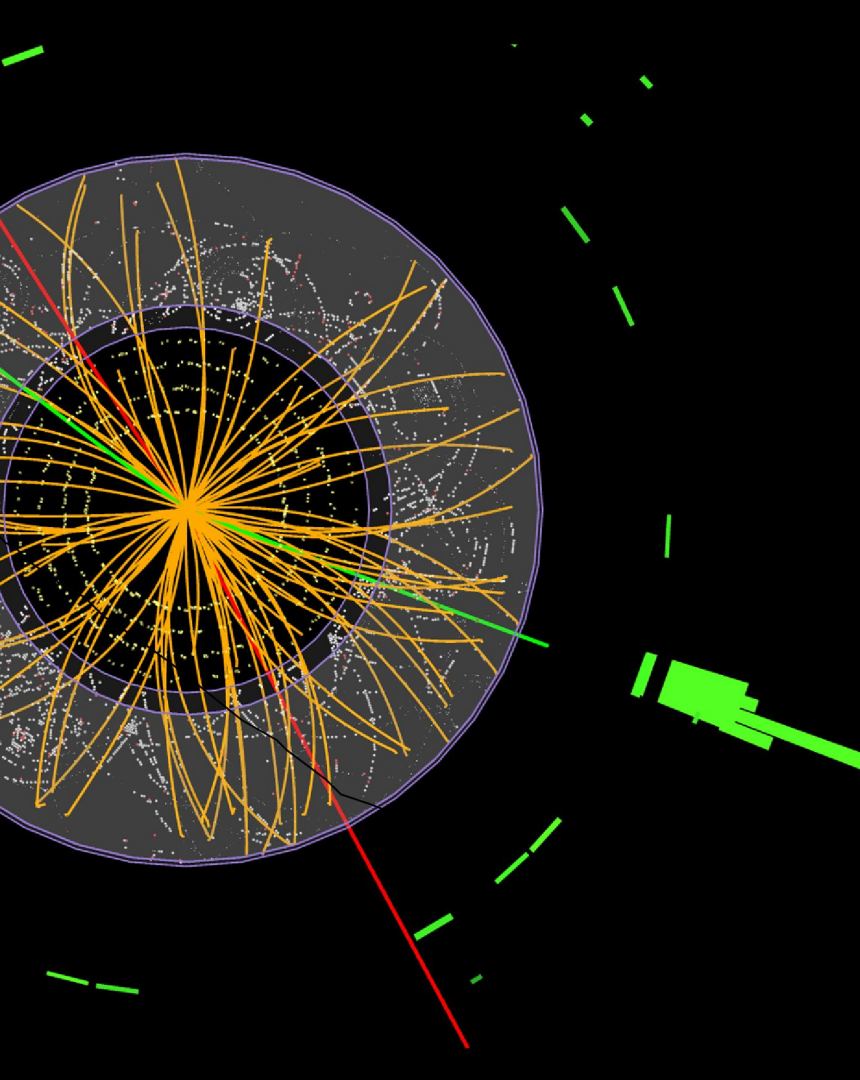
- **What do we want to release?**

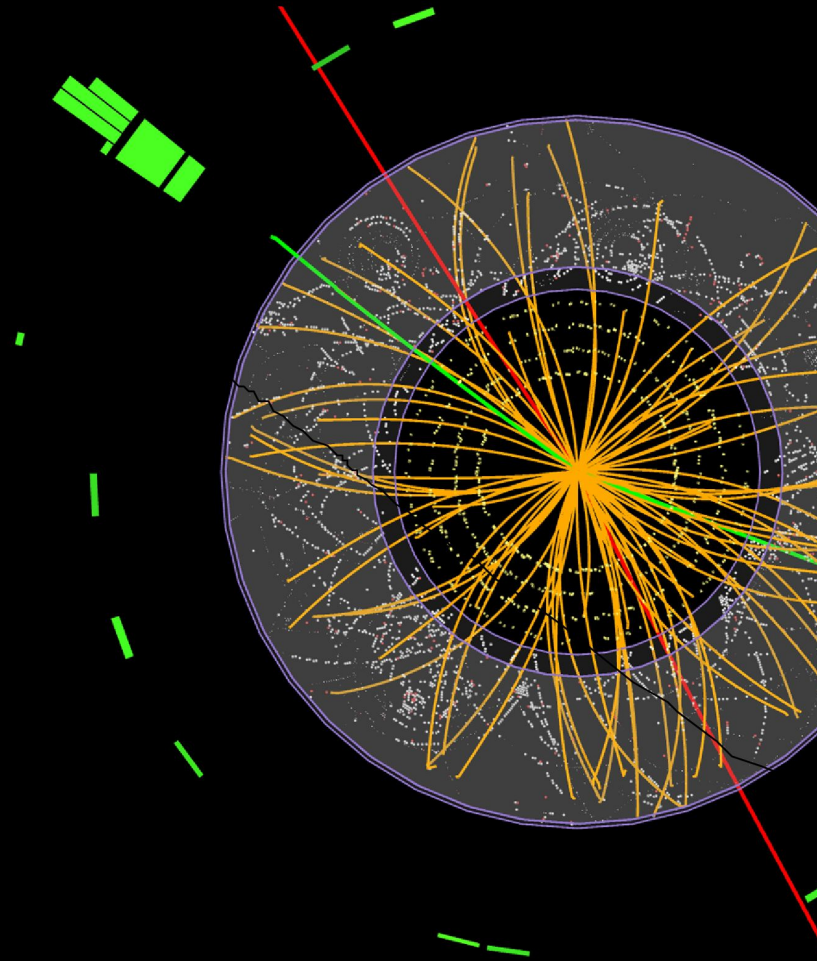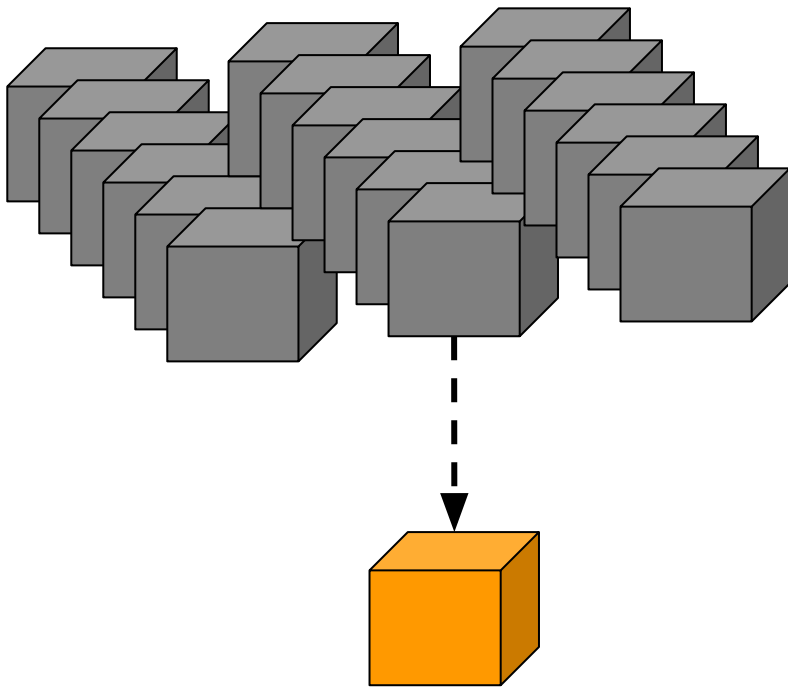**Data and analysis tools for educational use aimed at the universitary level* and beyond**

*Advanced High school students also may benefit

5

- **How we are doing it?**
Releasing data and MC samples in ROOT *tree* format; where the ATLAS collaboration decided 1 fb⁻¹ would be sufficient

The release...

...contains millions of events from 2012 data and MC that will be used by:

Target Audiences

- Students
  - Advance High School and University
  - PhD Students
- Researches
  - Training for new analyzers
  - Outside ATLAS scientific community
- General Public
  - Outreach
  - Curious people!

Type of sample-ntuples

A set of filtered AnalysisTop final analysis-kind (Run1) ntuples

- **Data Egamma string**
  - 1 root file from Period D
  - Total *final* #Events ~ 33.6MM
- **Data Muons string**
  - 1 root file from Period D
  - Total *final* #Events ~ 33.8MM
- **MC signals and backgrounds**
  - 38 samples, one per root file
  - 4 Higgs + 13 Z' signals
  - 3 VV + 6 (s)top + 12 W/Z+jet
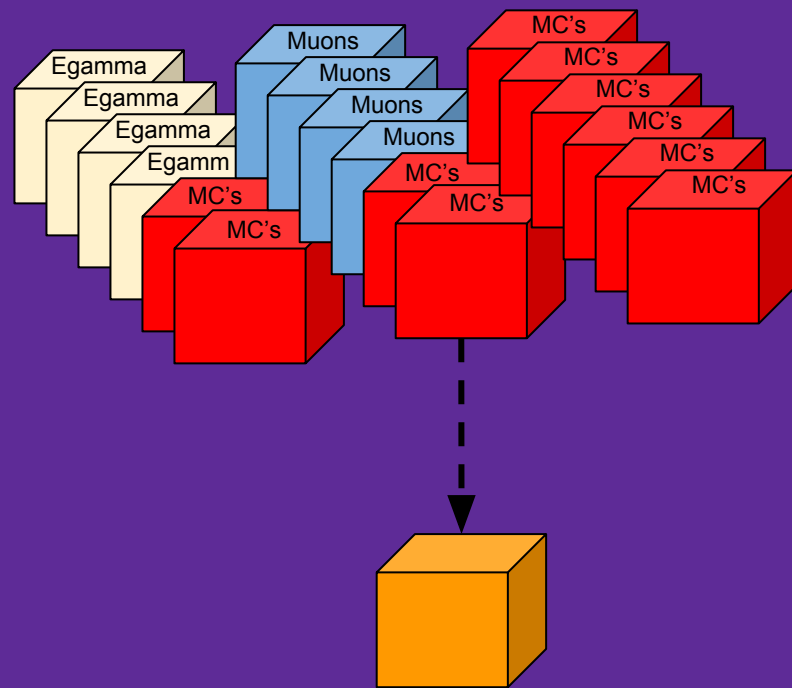
The release...

Approval Plots for the ATLAS Open Data published dataset

ATLAS Outreach Note
Plots from selected analyses using the published dataset to showcase the possible **abilities and limitations**

https://cds.cern.ch/record/2156957

ATLAS NOTE

May 31, 2016

Approval Plots for the ATLAS Published Dataset

S. Cheatham[f], A. Sanchez[i], K. Shaw[f], F. Socher[a]

[a]IKTP, TU Dresden (DE)
[f]INFN Udine & ICTP Trieste (IT)
[i]INFN Sezione di Napoli & Dipartimento di Fisica, Universit di Napoli (IT)
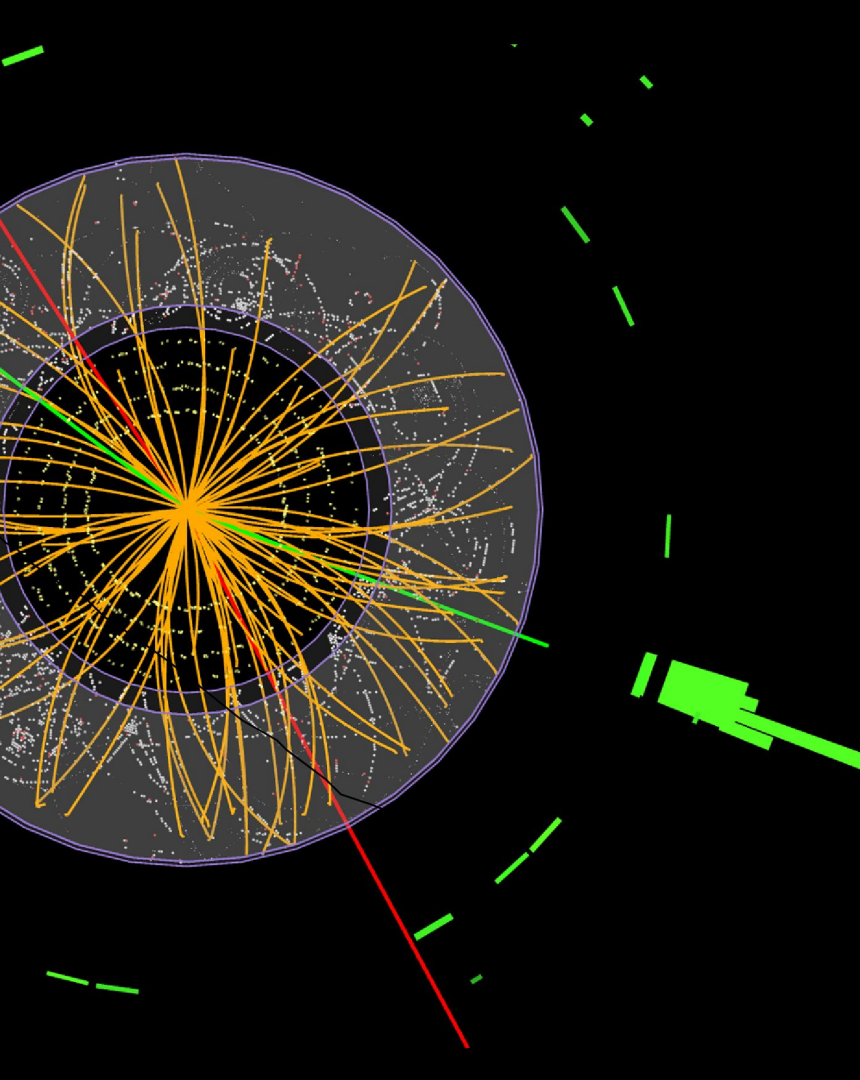
Abstract

This document presents approval plots from selected analyses using the ATLAS Open Data dataset. This dataset containing 1 fb^{-1} of data and matching Monte Carlo simulation,

period	N_{Events}^{pre}	N_{Events}^{sample}	$\mathcal{L}$ [pb^{-1}]
Egamma	7917590	33575219	1000.6
Muons	7028084	33815203	1000.6

Contents

- **And we are not alone in this effort**
  Several ATLAS members in universities around the world have been developing educational programs using the samples, testing them and give us feedback

Physics Lab course:
Technische Universitaet Dortmund (DE)
link

MasterClasses and HEP Education:
University of Oslo (NO), Dresden and Athens (GR)
link

Data & Tools

VRML Viewer, a web based event display:
Georgian Technical University (GE)
link

CEVALE2VE Outreach Group:
Several ATLAS and not ATLAS physics institutions/members
link

https://atlas-outreach.web.cern.ch/data-tools-group

Physics Lab course:
Technische

Ma...
h...
Uni...

Are you or your team part of an Education/Outreach project and you want to tell us more about it? Please join to atlas-outreach-data-and-tools@cern.ch and ask for spot in the Monthly meeting!

Several ATLAS
not ATLAS physics
institutions/members
link

12

- **The dataset has been prepared in harmony with the tools and educational use cases**

Looking to maximize their educational power we have created a platform where people can develop their own exercises and new ideas
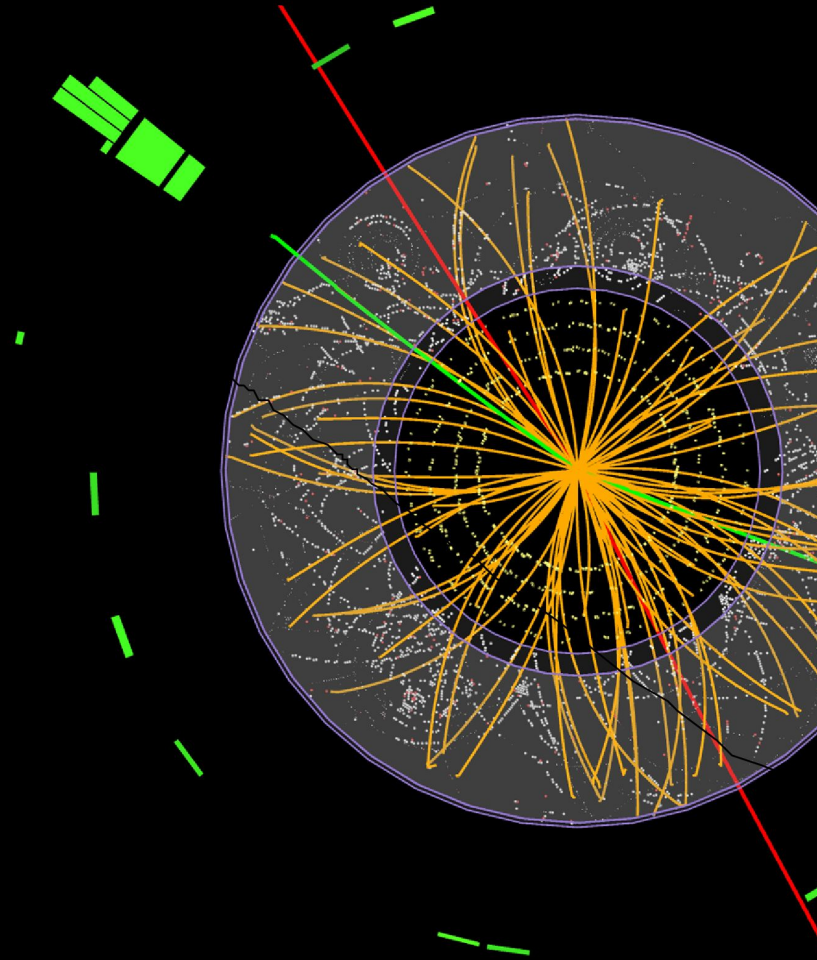
ATLAS
Outreach group

2. The Tools

Levels of interaction with the Datasets and the Tools

Level 1

Level 2

Level 3

Visualisation and qualitative "touch" of the datasets

Development and use of existing tools that allow the visualisation and the **interaction** with the ATLAS samples, events and detector.

Quantitative analysis on the Web

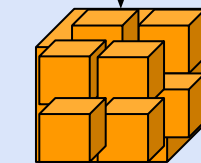Several efforts are ongoing in order to produce *meaningful* results out of the samples using web-based and cloud computing solutions.

Sophisticated but compact analysis toolbox

Tested tools like Virtual Machine (VM) that contains all the inputs, code, visualisations and analysis tools in a single space for more complex or long term educational and analysis tasks.

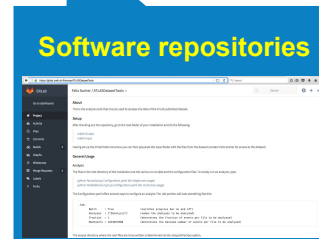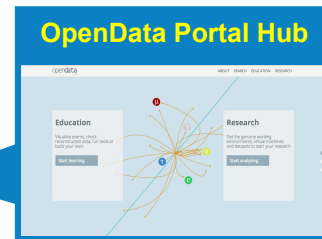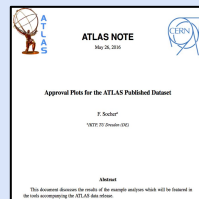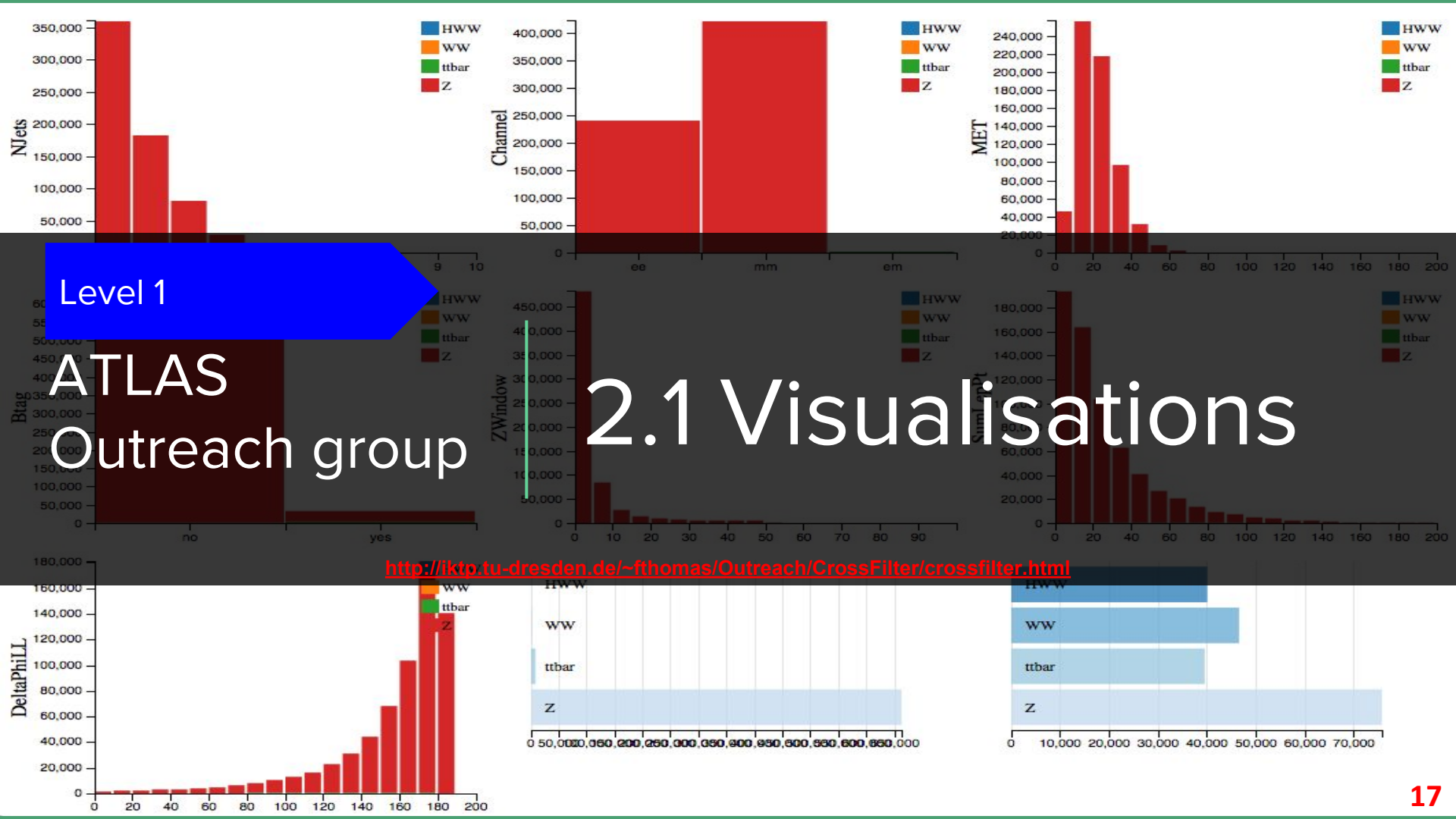# Surrounding the set of samples we have:

Website platform

Interactive Apps

Official Note

Externals collaborative repo

OpenData Portal Hub

Software repositories

Level 1

ATLAS
Outreach group

2.1 Visualisations

http://iktp.tu-dresden.de/~fthomas/Outreach/CrossFilter/crossfilter.html

Visualisations and Analysis

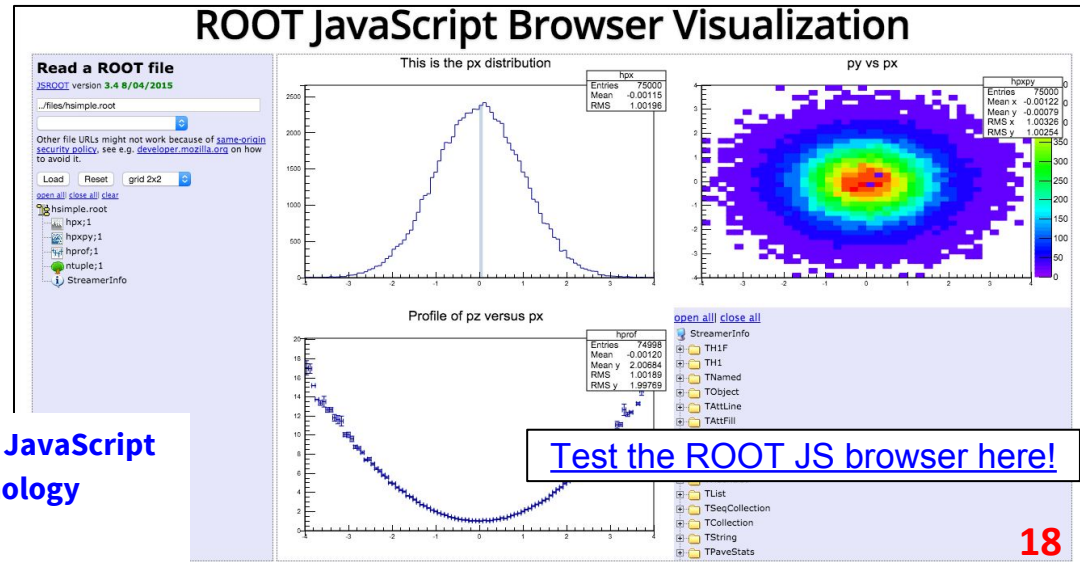There are several ways to visualise the dataset in form of histograms allowing to the user to interact with them!

Some examples here:

http://iktp.tu-dresden.de/~fthomas/Outreach/CrossFilter/crossfilter.html

http://atlasui03.na.infn.it:8088/~sanchez/ATLAS/crossfilter-gh-pages/crossfilter.html

Using *(mainly)* JavaScript and web technology

http://iopscience.iop.org/article/10.1088/1742-6596/513/5/052005

Test the ROOT JS browser here!

18

PAPER • OPEN ACCESS

Exploring JavaScript and ROOT technologies to create Web-based ATLAS analysis and monitoring tools

on behalf of the ATLAS Collaboration

Article PDF

22 Total downloads

Export citation and abstract

BibTeX   RIS

+ Share this article

Abstract

References

Related content

JavaScript ROOT

ROOT I/O in JavaScript

Geant4 application in a Web browser

ROOT I/O in JavaScript

Visual Physics Data Analysis in the Web Browser

brightrecruits.com jobs

Electronics Design Engineer
Vector Recruitment Limited

R&D Manager
Vector Recruitment Limited

Head of Tokamak and Neutral Beam Operations
United Kingdom Energy Authority, CCFE

More jobs   Post a job

Abstract

We explore the potential of current web applications to create online interfaces that allow the visualization, interaction and real cut-based physics analysis and monitoring of processes through a web browser. The project consists in the initial development of web-based and cloud computing services to allow students and researchers to perform fast and very useful cut-based analysis on a browser, reading and using real data and official Monte-Carlo simulations stored in ATLAS computing facilities. Several tools are considered: ROOT, JavaScript and HTML. Our study case is the current cut-based $H \rightarrow ZZ \rightarrow llqq$ analysis of the ATLAS experiment. Preliminary but satisfactory results have been obtained online.

Level 2

ATLAS
Outreach group

2.2 Web Analysis

19

Root + Jupyter NoteBooks

User's use RootBooks (NoteBooks + Root Kernel) to interact with the samples!

https://root.cern.ch/

Learning a lot thanks to the CERN ROOT Team!!

…**We are now writing the Notebooks** that read those samples and create example physics analysis under **ROOT**…

…**You can help us as well, subscribe and get in contact:**

atlas-outreach-data-and-tools@cern.ch

Level 3

ATLAS Outreach group

2.3 *Extended* Analysis

https://gitlab.cern.ch/fthomas/ATLASDatasetTools

ATLAS Outreach Tutorial VM 2016

22

Analysis Framework for Education

https://gitlab.cern.ch/fthomas/ATLASDatasetTools

GitLab

Felix Socher / ATLASDatasetTools ⌄

Search

Go to dashboard

🏠 Project

⚙️ Activity

🗂 Files

🕒 Commits

🔩 Builds   0

📊 Graphs

🕐 Milestones

📋 Merge Requests   0

🏷 Labels

⑂ Forks

About

This is the analysis code that may be used to analyse the data of the ATLAS published data

Setup

After checking out the repository, go to the root-folder of your installation and do the follo

```
mkdir Output
mkdir Input
```

Having set up the initial folder structure you can then populate the *Input* folder with the fi

General Usage

Analysis

The files in the root directory of the installation are the various run scripts and the configu

```
python Runscript.py Configuration.yaml (for single core usage)
python ParallelRunscrupt.py Configuration.yaml (for multi core usage)
```

The Configuration.yaml offers several ways to configure an analysis The Job portion will lo

```
Job:
    Batch     : True          (switches progress bar on a
    Analyses  : ["ZAnalysis"] (names the analyses to be ex
    Fraction  : 1             (determines the fraction of
    MaxEvents : 1234567890    (determines the maximum numb
```

The output directory where the root-files are to be written is determined via the *OutputDirectory* option.

The Analysis framework allows the interaction and analysis of the samples through of a series of analysis examples:

They produce histograms, cutflows and final plots.

This python framework allows the development of existing and new analysis.

And of course, any ATLAS member is invited to contribute!
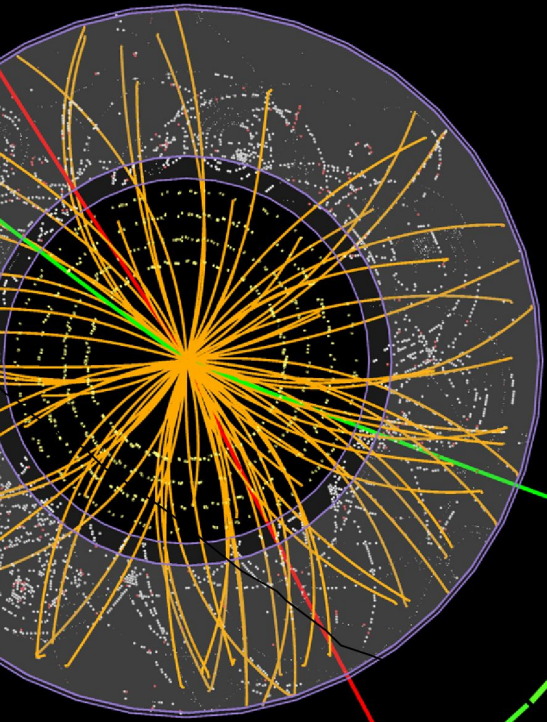
- **The analysis framework code can produce a large set of analysis plots. How do they look?**
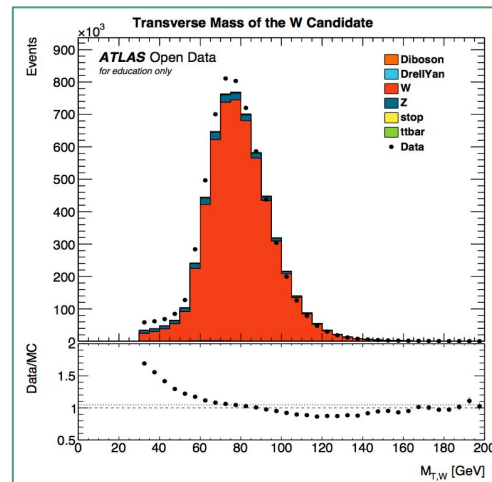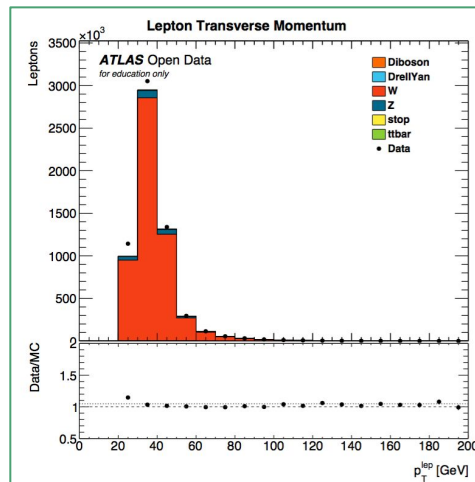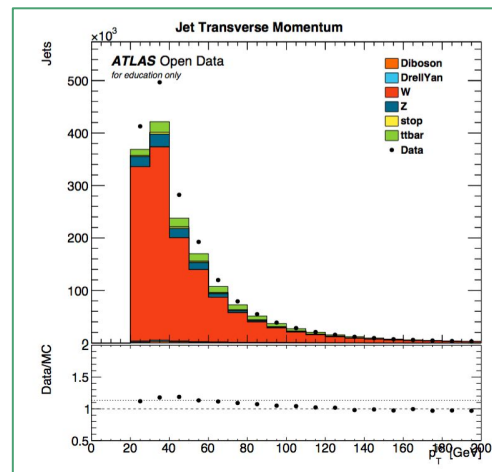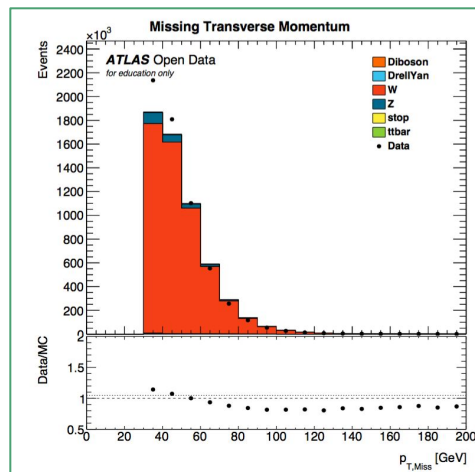
They look good! And when they have discrepancies, we took the time to understand where those come from:

24

W Analysis

Aim: reconstruction of the W boson in leptonic channel

Sources of discrepancy data/mc:

- dominant: due to the missing MC QCD contribution.
- *b*-tagging scale factors are not included.
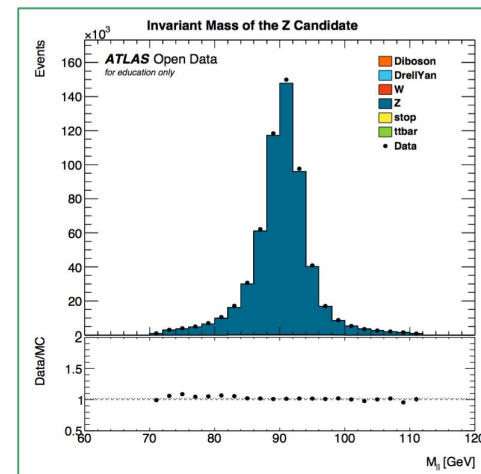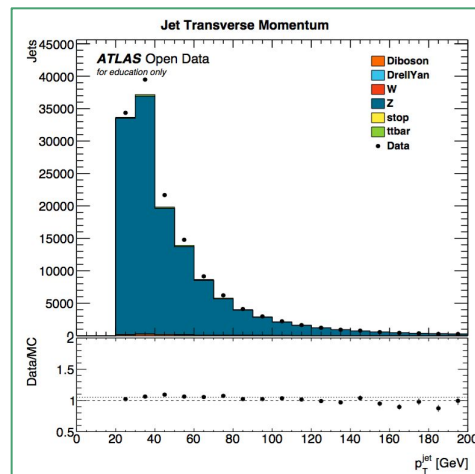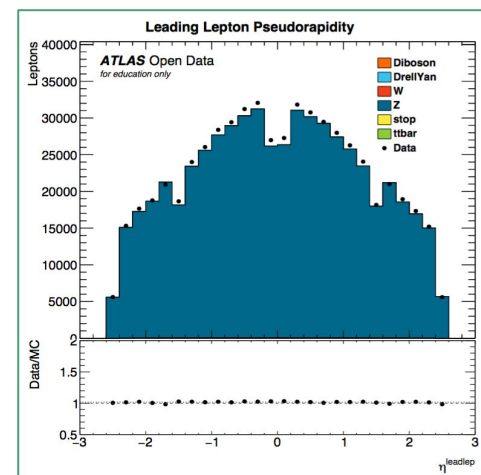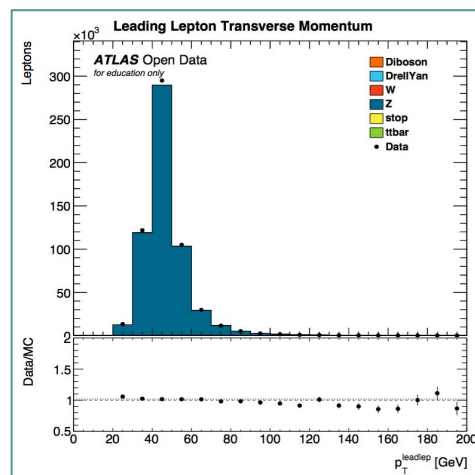- others including systematics, lumi uncertainties also play a role

Z Analysis

Aim: reconstruction of the Z boson in leptonic channel

Sources of discrepancy data/mc:

- dominant: missing of *b*-tagging scale factors
- uncertainties such as JER/JES were not included
- Others such as missing QCD MC, systematics and lumi uncertainties

HWW Analysis

Aim: search of the SM higgs boson in WW channel:

- Important point, it should not be possible to make a "discovery" with this set of data
- The higgs signal accounts for 3% of the contribution
- Educational purposes are always the unique objective of the samples, code and other tools
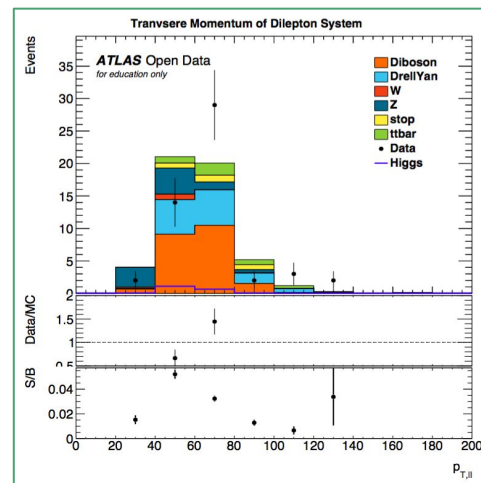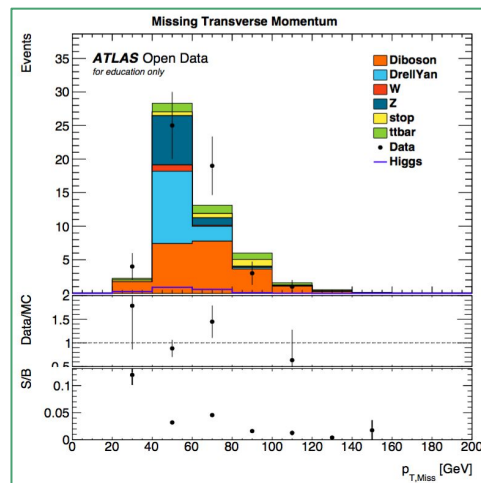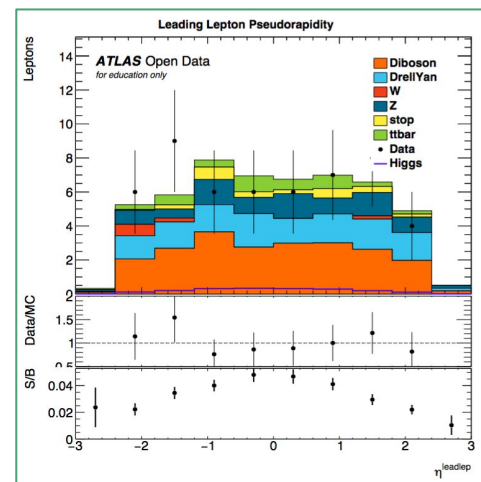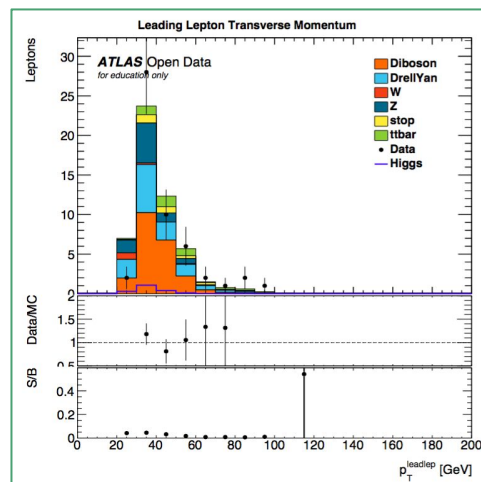
Virtual Machine(s)

We have a complete version of a Scientific Linux VM to perform the analysis and visualisations

- *Several VMs will be available for the public with different level of content, allowing the maximum of functionality for people with good and not-so-good internet services.*

python/C++ analysis framework software →

Samples in ROOT format Other auxiliary files →

Virtual Machine and space at Ixplus for ATLAS members

← ROOT 5&6 frameworks and other open source software

← AFS + CVMFS file systems e.g. A lot of software!

28

ATLAS
Outreach group

5. Platforms

5.1 ATLAS Open Data Website

ATLAS Outreach group

http://atlasopendata.web.cern.ch

An ATLAS-opendata Website I

The ATLAS-opendata Website allows the user to find the dataset samples, tools and documentation in order to start to work.

This is a place for questions, feedback and other possible interactions with the users
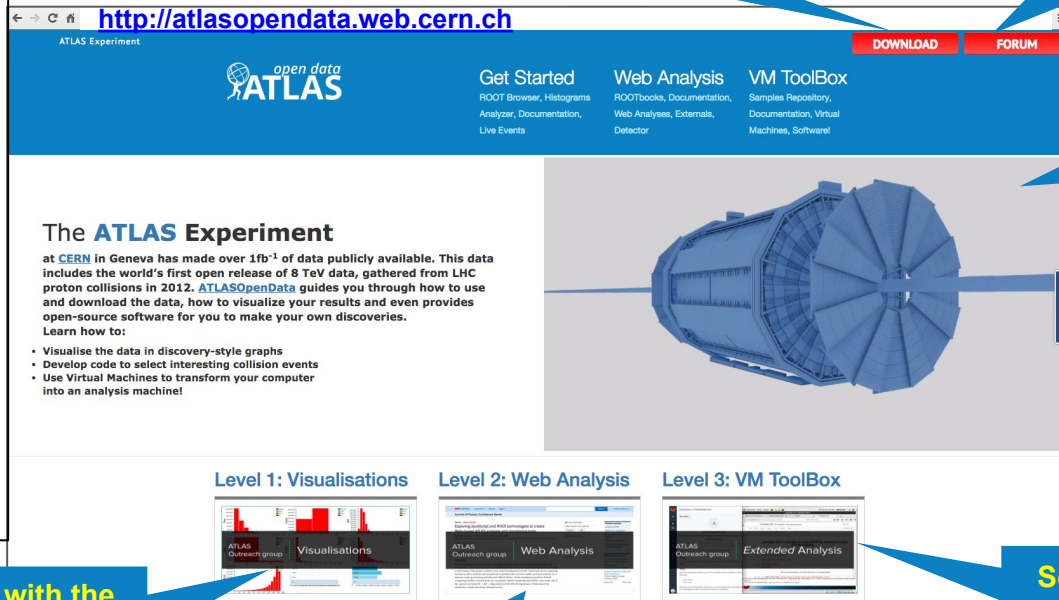
MC and data samples repository

Public Forum for Q&A, user's reports,...

http://atlasopendata.web.cern.ch

ATLAS Experiment

DOWNLOAD   FORUM

open data ATLAS

Get Started
ROOT Browser, Histograms
Analyzer, Documentation,
Live Events

Web Analysis
ROOTbooks, Documentation,
Web Analyses, Externals,
Detector

VM ToolBox
Samples Repository,
Documentation, Virtual
Machines, Software!

The ATLAS Experiment
at CERN in Geneva has made over 1fb⁻¹ of data publicly available. This data includes the world's first open release of 8 TeV data, gathered from LHC proton collisions in 2012. ATLASOpenData guides you through how to use and download the data, how to visualize your results and even provides open-source software for you to make your own discoveries.
Learn how to:
• Visualise the data in discovery-style graphs
• Develop code to select interesting collision events
• Use Virtual Machines to transform your computer into an analysis machine!

Provide Feedback

Video documentation!

Feedback button (JIRA ticket)

Level 1: Visualisations

Level 2: Web Analysis

Level 3: VM ToolBox

Section with the visualisations, introductory documentation

RootBooks examples and documentation

Software + VMs repository and documentation

31

An ATLAS-opendata Website II

Two sections are dedicated to the interaction Levels 1 and 2:

Web-based appls to visualise histograms, apply cuts and explore in *real time* the behavior of your selection!

An ATLAS-opendata Website III

The **Forum** is available for suggestions, questions or reports from external users alongside the feedback button

The Extended Analysis section consists in a serie of repositories and documentation where the user can get the software, VMs and dataset for further exploration

ATLAS Outreach group

5.1 CERN Open Data Portal

http://opendata.cern.ch/

At the same time… The release is uploaded to the
CERN OpenData portal

Through the CERN Open Data portal one can access the samples, code and VMs prepared for this release.

It will redirect the user to our dedicated website for more details and documentation

Education

http://opendata.cern.ch/

The CMS (Compact Muon Solenoid) experiment is one of two large general-purpose detectors built on the Large Hadron Collider (LHC). Its goal is to investigate a wide range of physics such as the characteristics of the Higgs boson, extra dimensions or dark matter.

Explore CMS >

ALICE (A Large Ion Collider Experiment) is a heavy-ion detector designed to study the physics of strongly interacting matter at extreme energy densities, where a phase of matter called quark-gluon plasma forms. More than 1000 scientists are part of the collaboration.

Explore ALICE >

The ATLAS (A Toroidal LHC ApparatuS) experiment is a general-purpose detector exploring topics like the properties of the Higgs-like particle, extra dimensions of space, unification of fundamental forces and evidence for dark matter candidates in the Universe.

Explore ATLAS >

The LHCb (Large Hadron Collider beauty) experiment aims to record the decay of particles containing b and anti-b quarks, known as B mesons. The detector is designed to gather information about the identity, trajectory, momentum and energy of each particle.

Explore LHCb >

For education purposes, the complex primary data need to be processed into a format (examples below) that is good for simple applications. Get in touch if you wish to build your own applications similar to those shown here
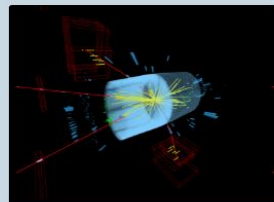
Visualise events >
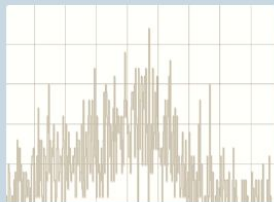
Visualise histograms >

Learning Resources >

35

ATLAS Outreach group | Summary

Exciting times for the ATLAS Outreach and the Collaboration: for first time we will release the first 8 TeV data and MC samples to the public!
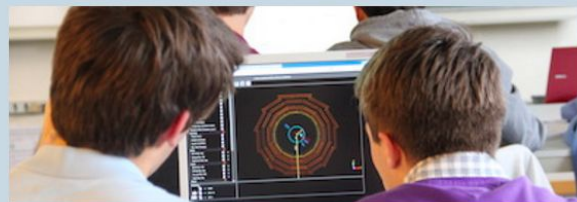
We have the ATLAS open data **dataset, tools, and a website** for documentation

The development of *useful* and accessible tools is a priority!

Next:

➔ The dataset approval plots will go for approval soon
➔ We are working on the website and the documentation (Send comments/suggestions please!)

http://atlasopendata.web.cern.ch

36

ATLAS
Outreach group | Looking Forward

This is just the beginning…

➔ We intend to extend the documentation & RootBooks
➔ We expect more groups to come forward wanting to develop their own lab course with the datasets
➔ And we will start to look toward releasing 13 TeV data this Winter with more advanced tools

ATLAS
Outreach group

Looking Forward

We will be putting the whole Virtual Machine, including tools, datasets and documentation onto a USB Stick for you, and anyone can buy one!

It is portable, and everyone in the world, even in countries with low bandwidth, can access the ATLAS Open Data!!

ATLAS
Outreach group

Backup