



LOS RETOS EN:

ATLAS OPEN DATA

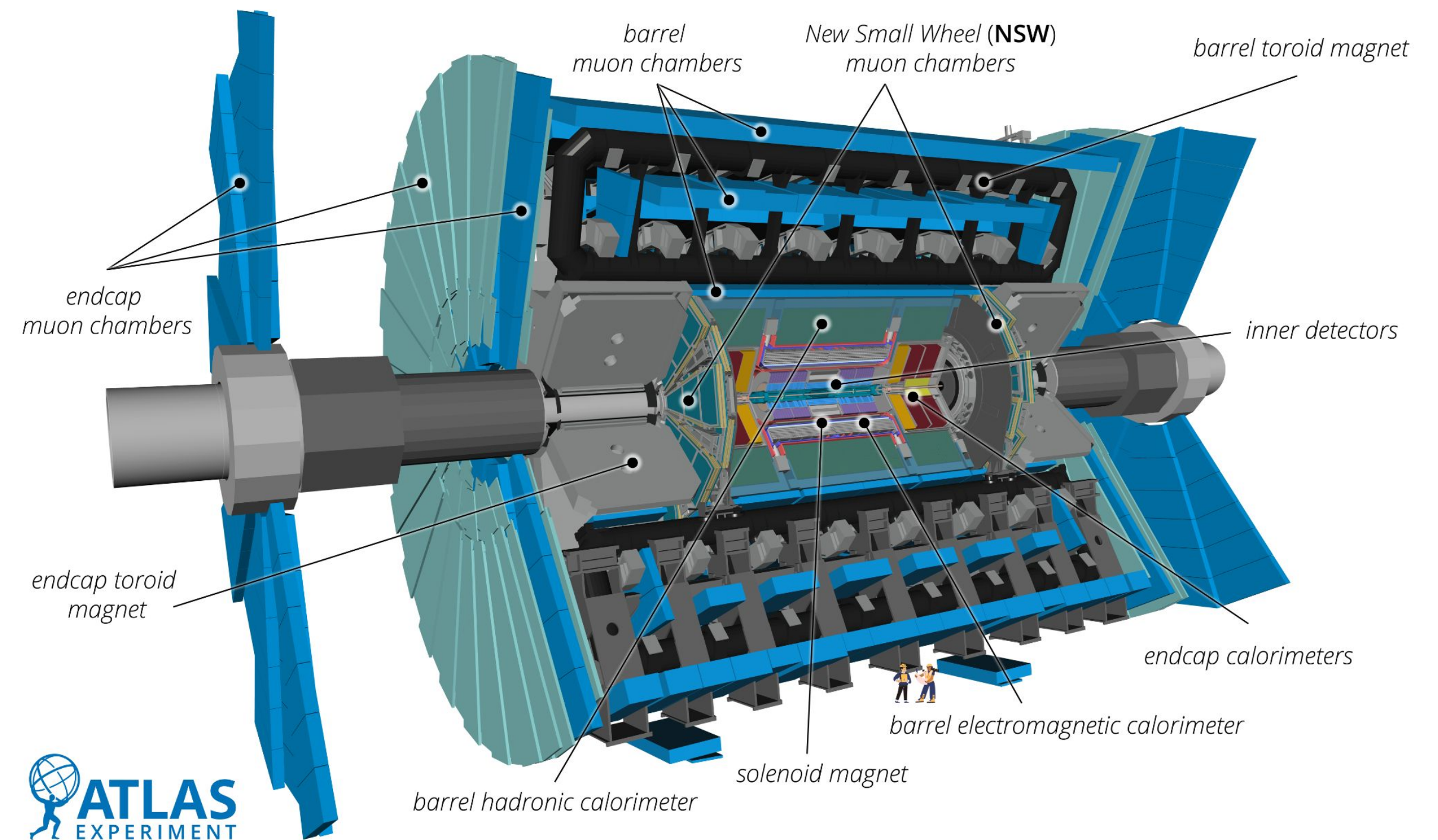
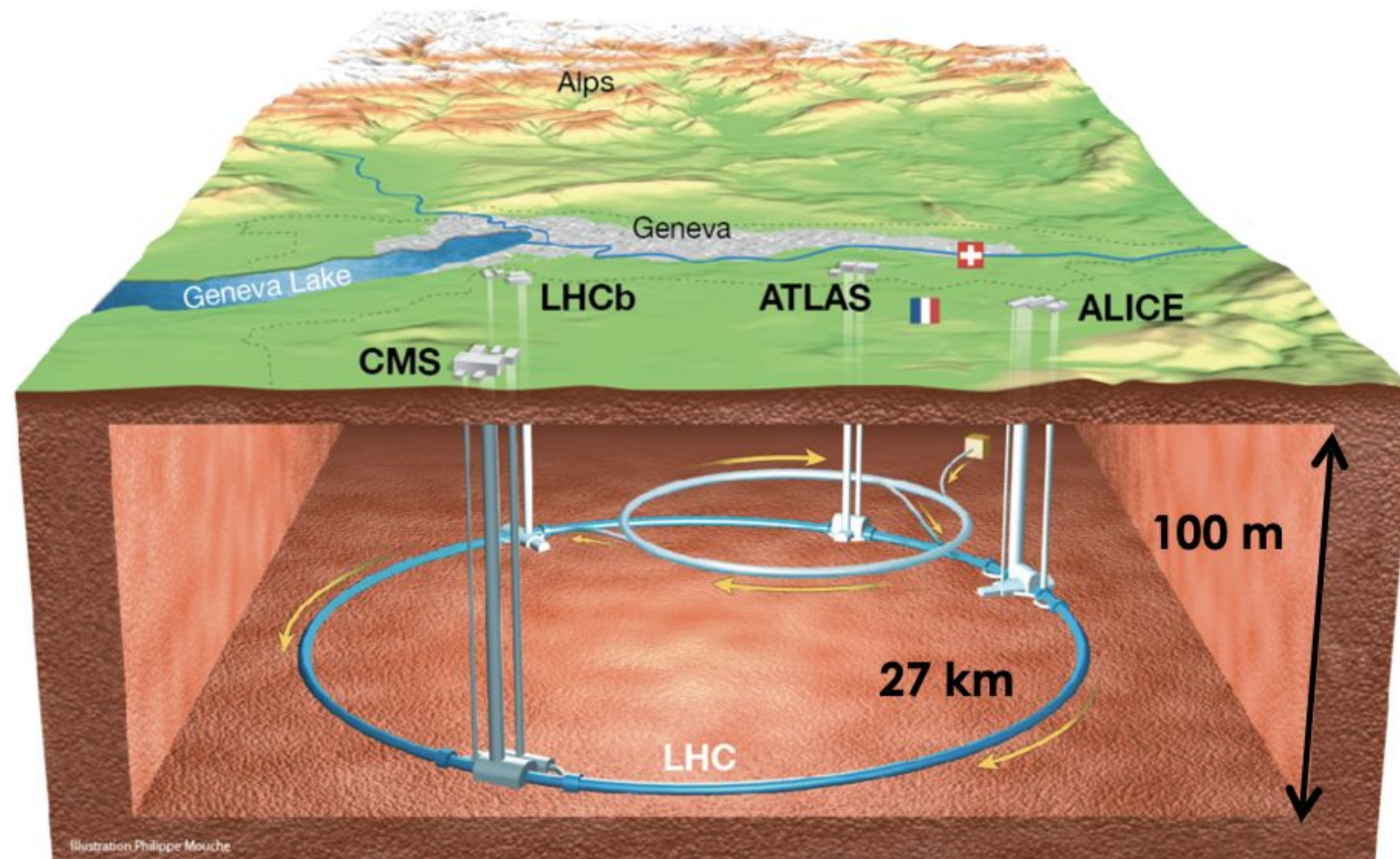
JULIO 19 - 2024

DILIA PORTILLO / CARLOS SANDOVAL

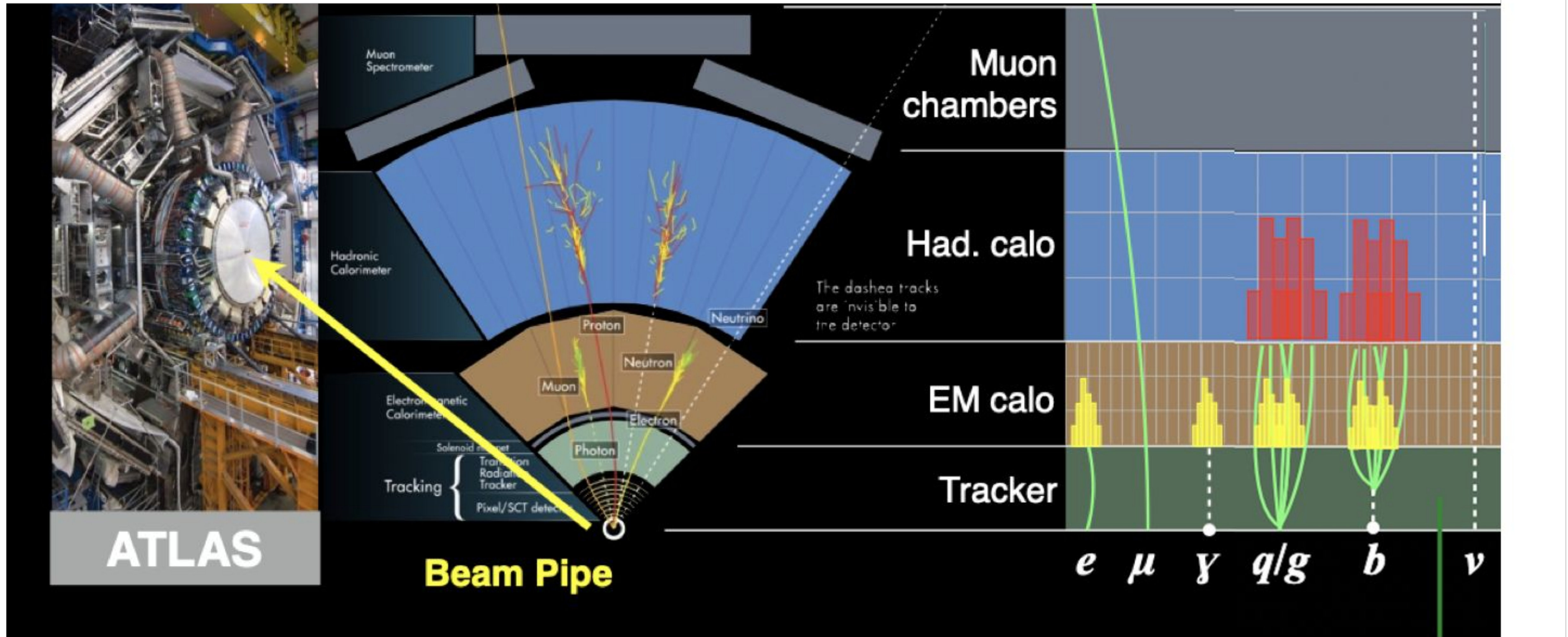
El experimento ATLAS en el LHC

- ▶ [ATLAS](#) es uno de los dos experimentos de propósito general Gran Colisionador de Hadrones ([LHC](#))

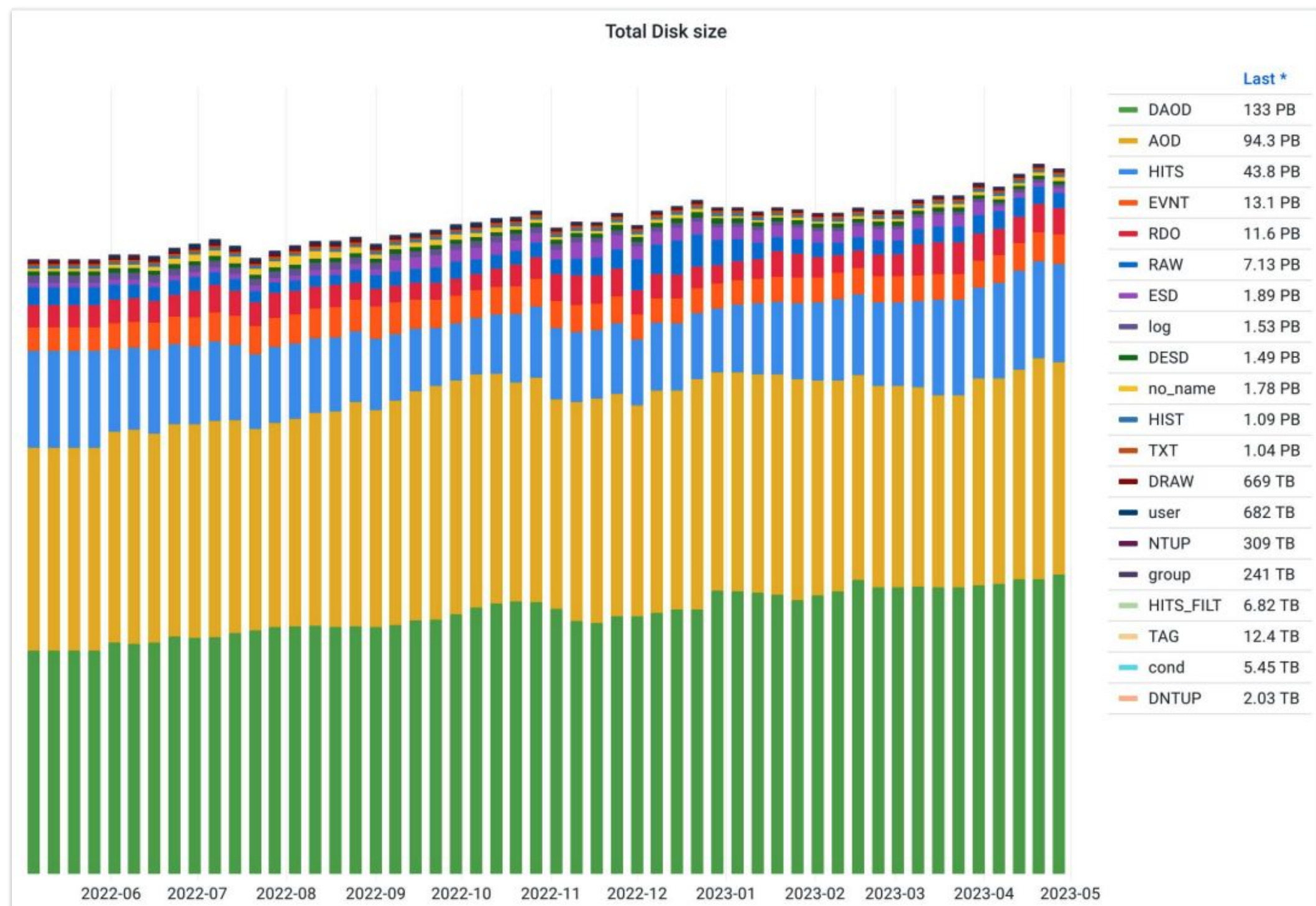
Buscando todo tipo de física, con especial énfasis en Higgs y más allá del Modelo Estándar de física de partículas



Reconstruyendo las partículas



Datos y Software de ATLAS en números



- ▶ ATLAS ha recogido ~4 PB de datos RAW hasta ahora
 - ▶ La gráfica incluye la replicación de datos
- ▶ Los datos de análisis reconstruidos ocupan >100 PB de espacio espacio
- ▶ La reconstrucción / simulación ATLAS / análisis común es de ~4 millones de líneas de C++
 - ▶ El tamaño instalado del software, con todas sus "dependencias", es O(10) GB



athena

Project ID: 53790



★ Unstar

174

Forks

2206

121,033 Commits 39 Branches 2,537 Tags 826.9 MiB Project Storage 415 Releases

The ATLAS Experiment's main offline software repository

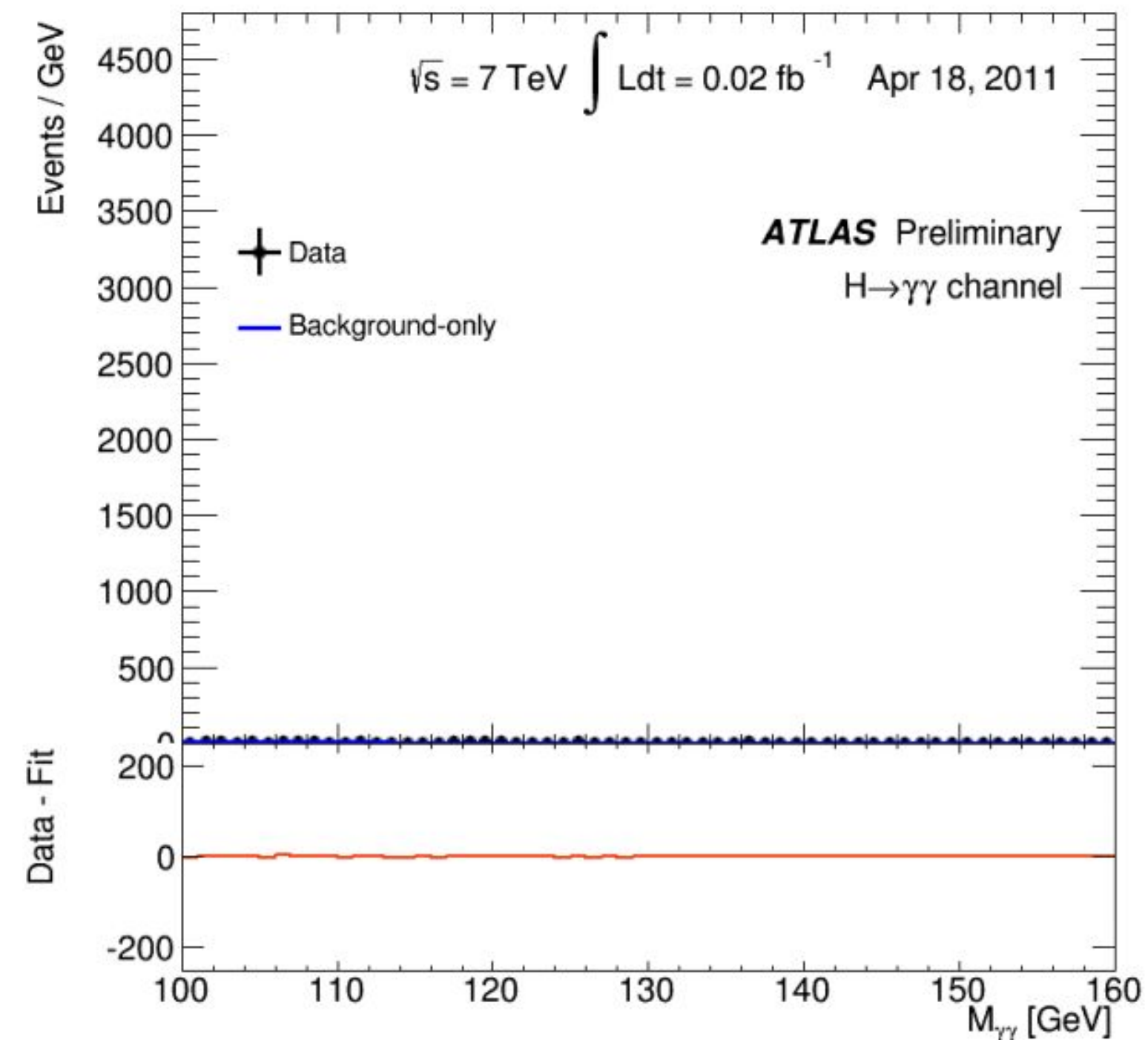
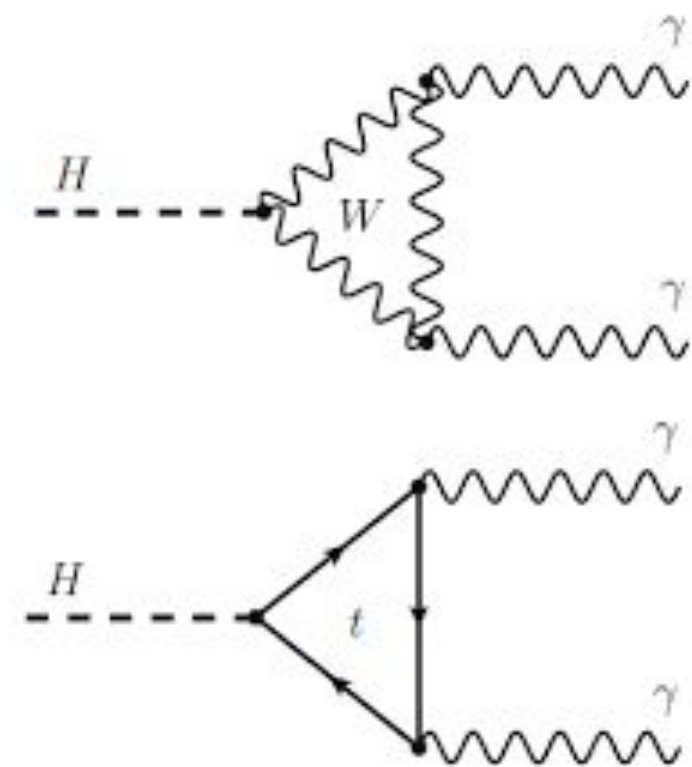
DOI 10.5281/zenodo.2641997

Doxygen main



Los análisis de datos en ATLAS

- ▶ Los análisis de datos requieren cientos de petabytes de almacenamiento de datos y cientos de miles de CPU para ejecutar millones de líneas de código C++.



- ▶ El proyecto ATLAS Open Data intenta ofrecer una muestra de estos análisis a estudiantes y al público en general

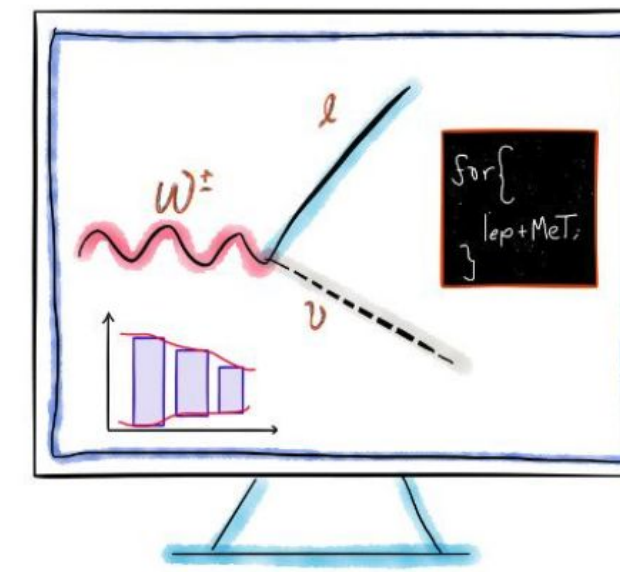
ATLAS Open Data Project

- ▶ Hacer públicos (algunos de) los datos de ATLAS tiene múltiples beneficios
 - ▶ Exponer a los estudiantes a estos datos, y al método de análisis de datos utilizado en el experimento, ayuda a despertar el interés por nuestro campo
 - ▶ Proporciona recursos para enseñar habilidades transferibles en técnicas de programación y análisis.
 - ▶ Los datos y el software pueden ayudar a los profesores a hablar de la física moderna de altas energías.
 - ▶ A largo plazo, debería mejorar la cultura científica del público.

En qué consiste la Open Data?

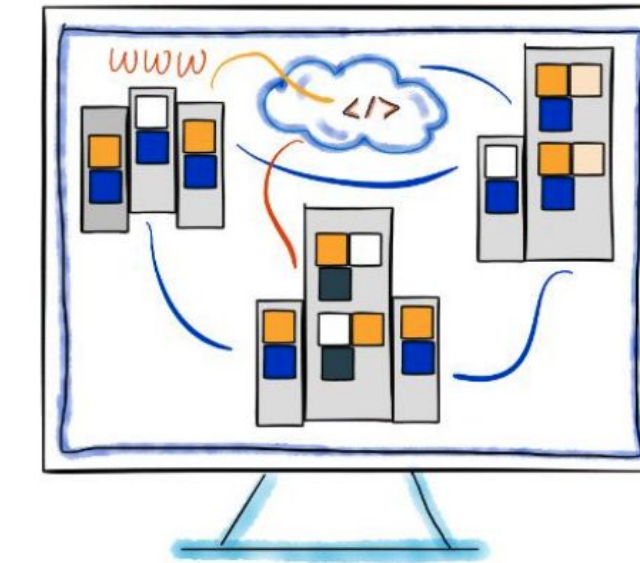
- ▶ Consiste en 2 "Campañas"
 - ▶ 1 fb-1 de datos de 8 TeV
 - ▶ 10 fb-1 de datos de 13 TeV
 - ▶ <https://cds.cern.ch/record/2707171/files/ANA-OTRC-2019-01-PUB-updated.pdf>
- ▶ Ambas acompañadas de las respectivas simulaciones de Monte Carlo
- ▶ Datos y simulación de MC preseleccionados con diferentes filtros para diferentes casos de estudio
- ▶ Permite el uso de un conjunto de datos mucho más pequeño en determinadas situaciones, para análisis / demostraciones más sencillos.

The 8 TeV samples



Learn more about the 2016 datasets

The 13 TeV samples

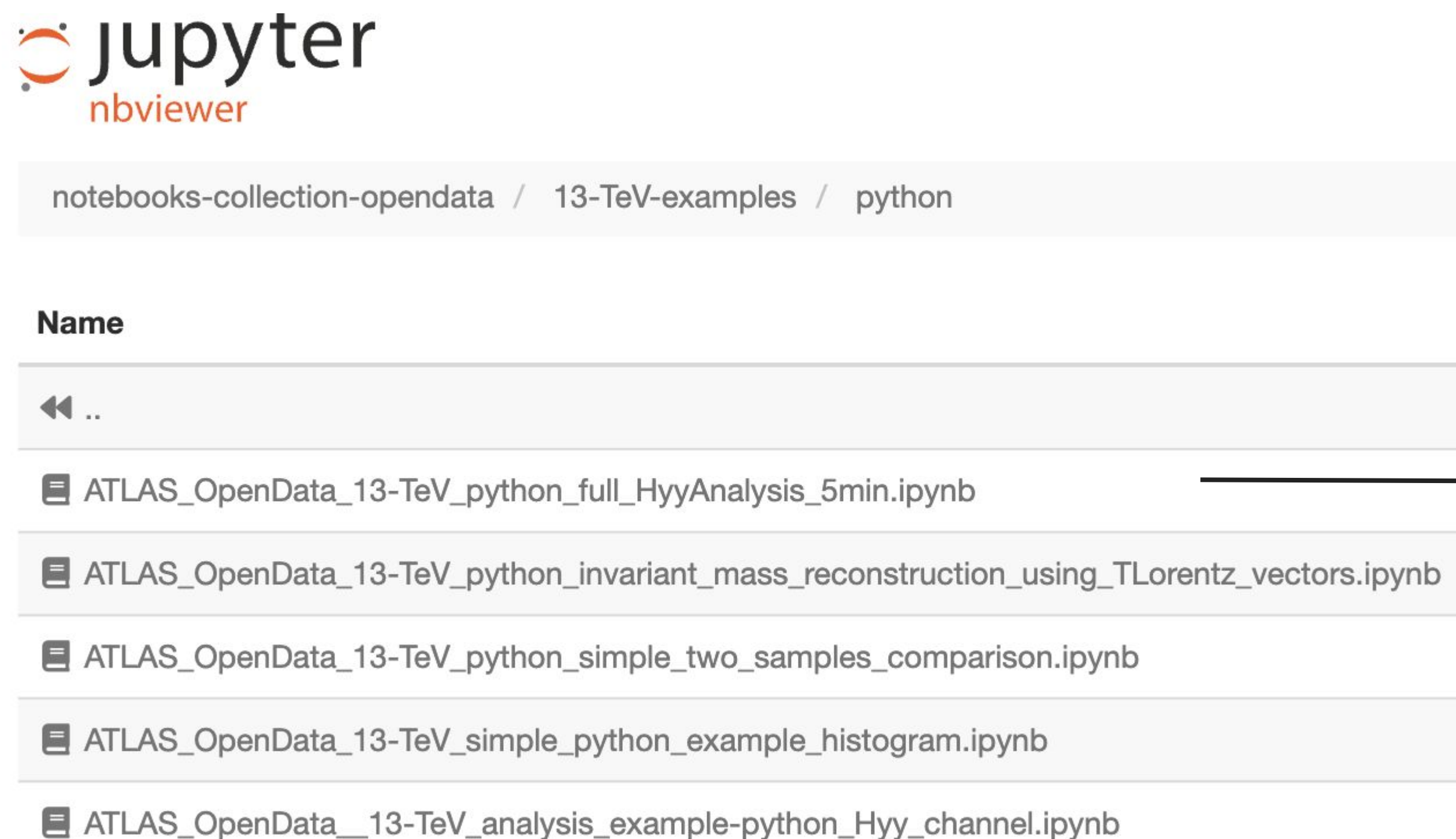


Explore the 10x more data in 2020 datasets

Description	Name	link to ZIP file
events selected with at least one lepton (electron or muon) and exactly one large-Radius jet (R = 1.0)	1largeRjet1lep	5.5 Gb
events selected with exactly one lepton (electron or muon). This is a very large collection, so, it was divided into three ZIP files	1lep	17 Gb, 20 Gb, 21 Gb
events selected with exactly one lepton (electron or muon) and exactly one hadronic-reconstructed tau	1lep1tau	1.3 Gb
events selected with at least two leptons (electron or muon)	2lep	24 Gb
events selected with exactly three leptons (electron or muon)	3lep	1.0 Gb
events selected with at least four leptons (electron or muon)	4lep	427 Mb
events selected with at least two photons	GamGam	1.5 Gb

Reto 9: ATLAS Open Data

- ▶ Elige uno (o más) de los 13 TeV Open data python examples listados aquí
- ▶ <https://nbviewer.org/github/atlas-outreach-data-tools/notebooks-collection-opendata/tree/master/13-TeV-examples/python/>



Buen candidato para comenzar!

Redescubrir el Higgs :)

Inspiración:

<https://www.kaggle.com/code/meirinevans/how-to-rediscover-the-higgs>

Reto 9: ATLAS Open Data, los pasos

1. Sigue los pasos y asegúrate de que el estudiante puede ejecutar el ejemplo tal cual e identifica las partes que técnica y científicamente son más complejas.
2. Mejora la explicación de la física para hacerla más accesible pedagógicamente. Considera también añadir ecuaciones, referencias y diagramas para mejorar la documentación.
3. Propón gráficos intermedios u otros gráficos finales que ayuden a comprender el estado final o el resultado y a ampliar la investigación.
4. Proporciona también la documentación en español para ampliar el alcance de estos recursos.

BONUS!

Si quieren ir más allá, y tienen buenos conocimientos de python y nociones de análisis ATLAS, también pueden proponer otro estado final y crear su propio ejemplo.