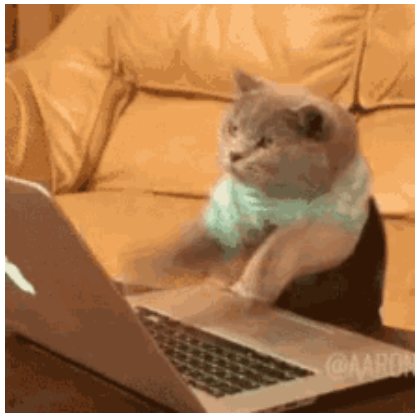
Introduction to PyROOT and RDataFrame

USCMS Undergraduate Summer Internship 2024

June 4, 2024



What to expect

- Lecture + exercises
- Lecture: 7 modules
- Exercises: Work together in the afternoon
- Run everything with Jupyter Notebook

https://github.com/Ari-mu-l/software-carpentry/tree/PURSUE2024

01-root-intro.ipynb	Make descriptions better for the presentation mode	6 months ago
02-histograms-and-graphs.ipynb	minor fixes	2 months ago
03-tfile-read-write-ttree.ipynb	Updated with interactive exercises	2 months ago
04-rdataframe-basics.ipynb	Updated with interactive exercises	2 months ago
05-rdataframe-collections.ipynb	Updated with interactive exercises	2 months ago
06-rdataframe-features.ipynb	minor fixes	2 months ago
07-rdataframe-advanced.ipynb	Updated with interactive exercises	2 months ago

What to expect

Code examples

In [1]:
```
import ROOT
h = ROOT.TH1D(name="h", title="My histo", nbinsx=100, xlow=-5, xup=5)


h.FillRandom("gaus", ntimes=5000)
```

```
Welcome to JupyROOT 6.24/07
```

In [2]:
```
# Try it!
# When you are not sure how ot use a function, check its documentation by running ?ROOT.<function_name>
?ROOT.TH1D
```

Execute

Directed small hands-on tasks that helps you to deepen your memory

In [ ]:
```
# Try it!
# To-do: import the ROOT module
import

# To-do: Create an empty histogram called "h1", with 50 bins, ranging from -1 to 1
h1 = ROOT.TH1D(name=, title=, nbinsx=, xlow=, xup=)

# To-do: Fill the histogram with 1000 entries randomly sampled from a gaussian distribution
h1.FillRandom("", ntimes=)
```

Fill-in blanks

What to expect

- Learn the basics of pyROOT and RDataFrame
- It’s normal to not remember all the details of each function or tool
- What’s more important than trying to memorize them is to establish the ability and habit of reading documentations and online resources, so that you know how and where to find answers when you work on your research projects!
- Get familiar with cmslpc clusters → You will very likely work on your future research at US CMS on the cmslpc clusters

Setting up

- CMS Offline Software (CMSSW) contains all the packages we need in this workshop. It is also what we use to perform physics analyses
- Access CMSSW using cmslpc clusters
- Follow the instructions in the github repository (https://github.com/Ari-mu-l/software-carpentry/tree/PURSUE2024)
- If you cannot access cmslpc, access the repository through this link https://mybinder.org/v2/gh/Ari-mu-l/software-carpentry.git/PURSUE2024