Teaching ROOT to DESY Summer Students

Changes over the years

David Brunner and Dirk Krücker
ROOT Users Workshop, May 12, 2022

HELMHOLTZ

DESY.

The DESY Summer Student Program

A short advertisment:

https://summerstudents.desy.de

In 2022 back to a
in person program!

DESY has a longstanding tradition in particle physics and more - Summer student program for

- ATLAS, Belle II, CMS, ..

- Astrophysics

- HEP Theory

- Photonscience

Expected experience

- At least three years of full-time studies at university level but not yet a Master degree

- All HEP summer students (still?) need an introduction to ROOT

    - One of us (DK) involved for ~10y

Our course is attended by students from the underlined groups

Our Approach

Overview

- A two-day crash course of 4 hours each in the first week of the program

- Mixed lectures and tutorials, e.g. last pre-pandemic

 - 1 lecturer + 2 tutors

Preparation

- Prior to the program, students are asked to acquire basic Linux and programming skills (Python)

  - Many online courses available

- A simple email questionnaire for self-assessment

The basic idea: **get everyone on board**, despite the big differences in experience

```
What do you know about ROOT?        [ ] never heard about

                                    [ ] a little

                                    [ ] good basics

                                    [ ] fluently

What do you know about C++?         [ ] never heard about

                                    [ ] a little

                                    [ ] good basics

                                    [ ] fluently

What do you know about Python?      [ ] never heard about

                                    [ ] a little

                                    [ ] good basics

                                    [ ] fluently

What do you know about matplotlib?  [ ] never heard about

                                    [ ] a little

                                    [ ] good basics

                                    [ ] fluently

What kind of Notebook/operation system do you bring with
you? Linux/Windows/Mac
```

Questionaire

Examples
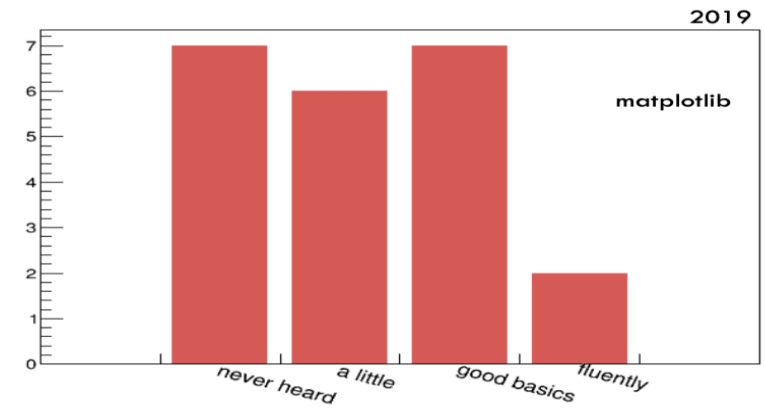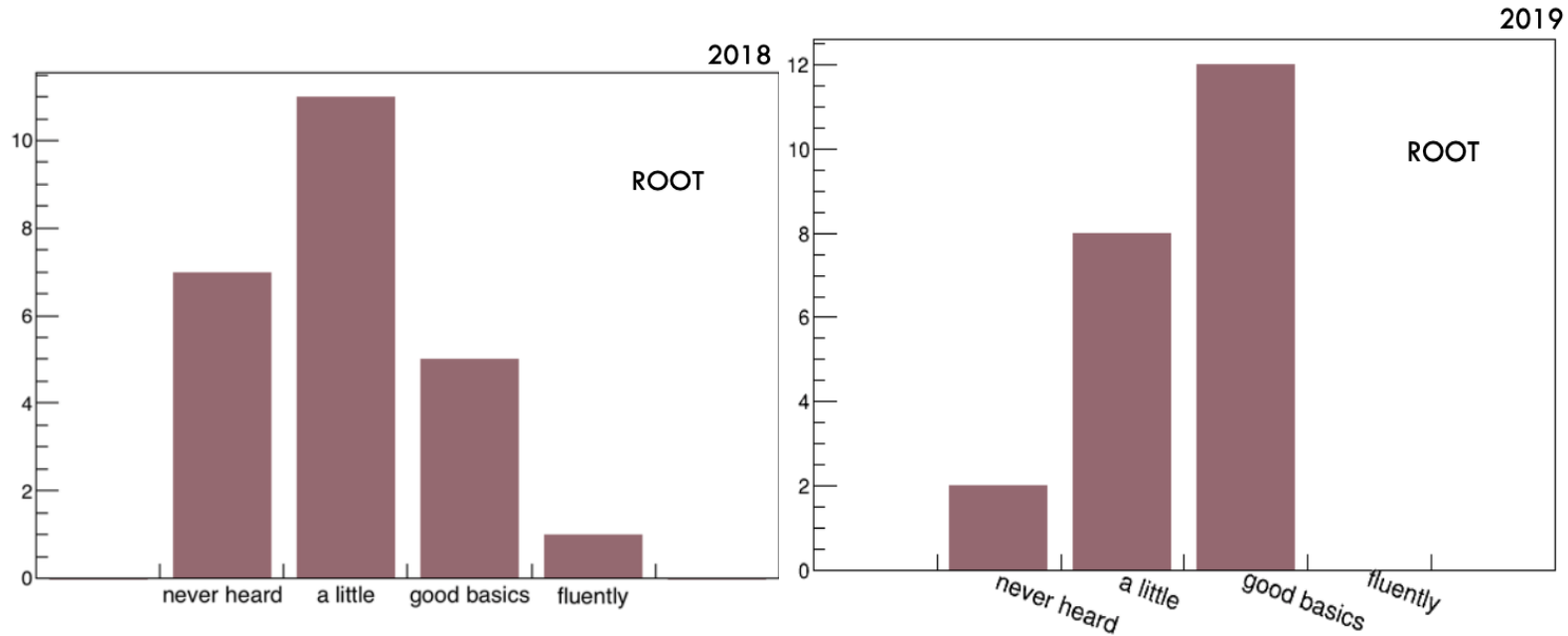
- Strong fluctuations in knowledge between students and years

- Increasing experience with matplotlib

Observations over the last decade
The introduction in ROOT has become more challenging

- In general, the experience in computing basics among HEP students seems to decline

- The student's experience is more scattered

  - Python

  - Javascript, web related stuff

  - C/C++ rare

  - Knowledge of OOP concepts cannot be expected

- For doing a HEP analysis Python has probably become the standard approach despite of performance problems

  - But C++ is essential: experiment frameworks, GEANT4, legacy code, etc.
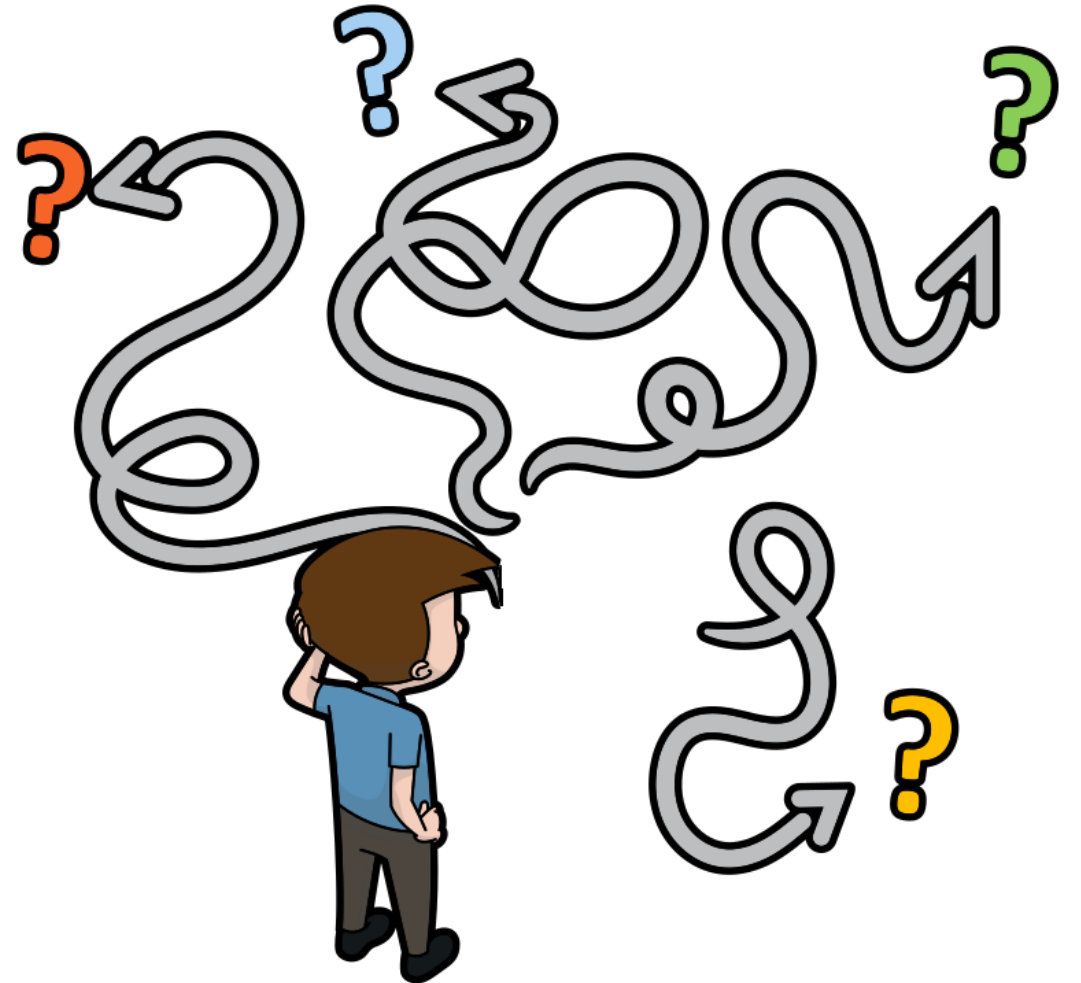
Observations over the last decade

## The introduction in ROOT has become more challenging

- The complexity of ROOT has exploded

  - Analysis tool & data format at the same time

- **First time contact is confusing**

  - Multiple implementation e.g.

    - TLorentzVector (all you need for HEP beginners)

    - GenVector class (probably optimal as data format)

  - Let alone

    - TMVA

    - RooStats

    - R/Javascript interface …

  - Different concepts how to read a tree over the years

    - Now RDataFrame …

Needs on our Side

Moving target

Summer students work in different DESY groups with multiple projects/needs

- Typically Python

- But also C++ scripts in form of ROOT macros are used

- Often within large frameworks,
  e.g. CMSSW

- No use to teach concepts that have not arrived in the analysis groups (e.g. RDataFrame)
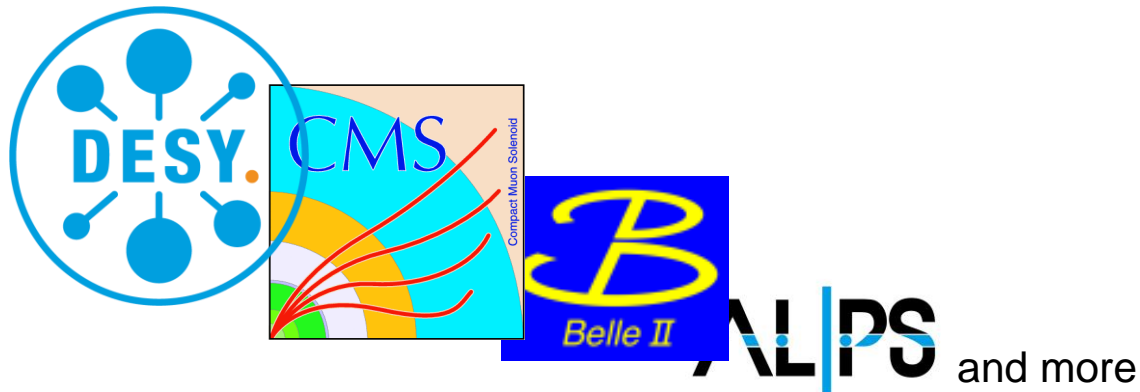
Basic knowhow to teach: **How to do nice plots**

- Read trees from rootfiles

  - Plain and custom files from experiments, e.g. nanoAOD

- Do selection

- Create plots/histograms

- Do fits

➔ Understand the idea of a class

  - OOP design of ROOT

  - Basic translation C/C++ <-> Python

- Learn how to access and digest online documentation

and more

How to teach

CMD line vs Jupyter etc.

Our approach for now is CMD line

- CMD line on WGS in cluster

- Mostly used in real HEP analysis

- Interactive use of ROOT

  - TBrowser, FitPanel etc
    GUIs look old-fashioned but are very useful

Jupyter notebooks are a great tool for tutorials but

- Click and forget problem

- Do not prepare for the work within a HEP analysis group

- Teach in Python or C++?

Explanining ROOT's Python and C++ API in parallel??

- Restrict to the basics

- Explain concepts

What we do is basically similar to CERN summerstudents introduction
e.g. (S. Hageboeck, E. Tejedor) 2019

Outlook

uproot has been a game changer

- Tools from the ML community may dominate soon

- ROOT incorporated experience from decades of doing data analysis in HEP

  - A simple weighted histograms with proper error bars does not exit in matplotlib

  - 4-vectors

- Despite of conda, ROOT installation in an python environment still feels heavy

  - LCG releases on CVMFS helps a lot

- Are DataFrames with Pandas-style really the way to go?

- Our group (SUSY search in CMS) moves to coffea/awkward/uproot

- The options for conducting an HEP analysis are becoming more diverse and it is becoming more difficult to decide what to teach our summer students

Thank you

Contact

Deutsches Elektronen-
Synchrotron DESY

D. Brunner and D. Krücker
Department FH/CMS

www.desy.de