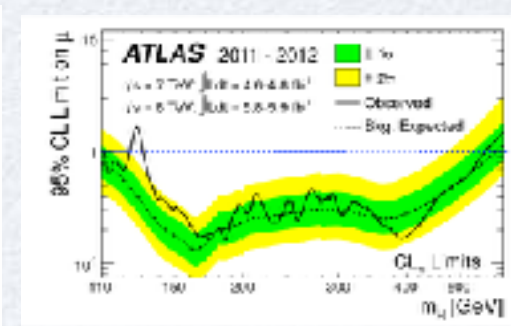
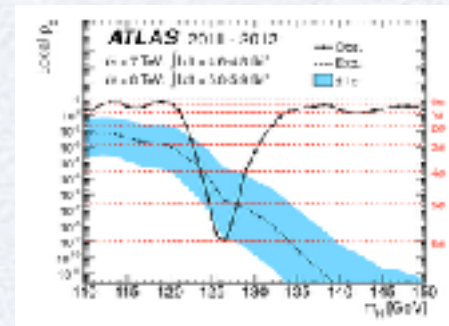
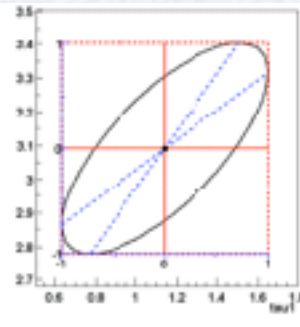
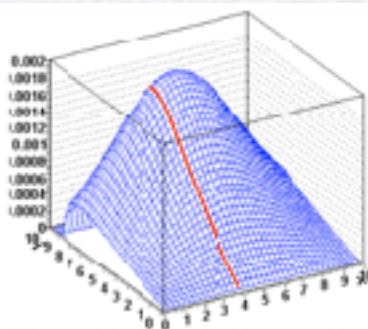
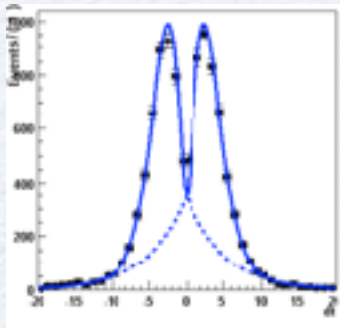


Practical Statistics for Particle Physics Analyses: Introduction to Computing Examples

Louis Lyons (Imperial College), Lorenzo Moneta (CERN)

IPMU, 27-29 March 2017



Introduction

- Hands-on session based on computing examples
- Examples uses **ROOT**
 - including its advanced statistical packages, RooFit and RooStats
- Examples are **interactive demos** made as **Jupyter** notebooks and written in C++
 - Jupyter supports also Python

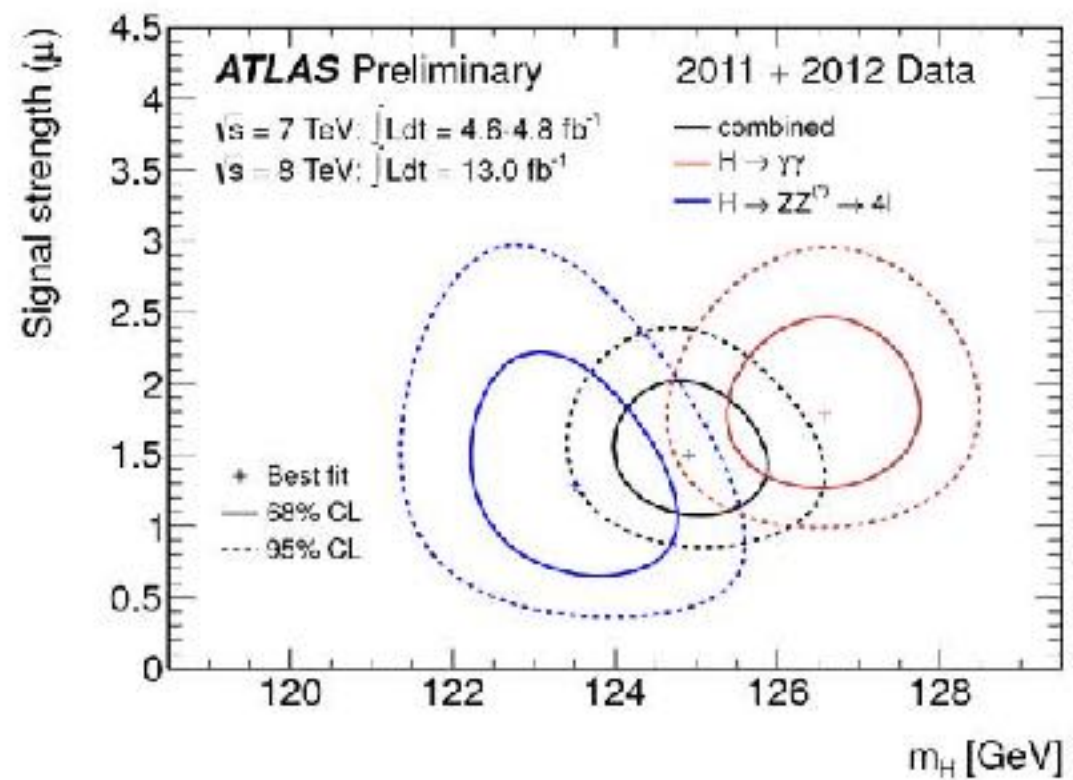
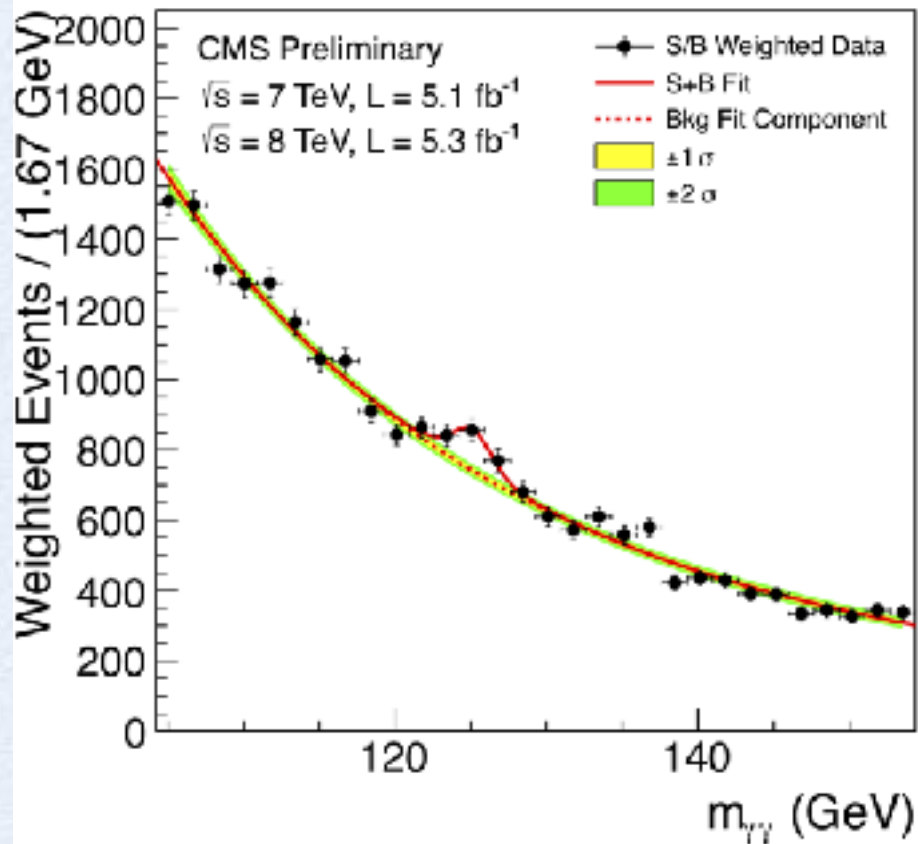
Computing Examples

- Interactive demos made of Jupyter notebooks
- **How to run them ?**
- We have 3 possible options:
 - 1. Running in the CERN notebook service**
 - SWAN (swan.cern.ch)
 - no installation required but needs a CERN account
 - 2. Install ROOT in your computer and run a local notebook**
 - [root —notebook](#)
 - 3. Use a Virtual Machine containing the installed software**
 - <https://cernbox.cern.ch/index.php/s/oKBM14VDqGFT7Wl>

Computing Examples

- We recommend running the notebooks at CERN in
 - **SWAN** (swan.cern.ch)
 - files will be stored in **CERNBOX** (shared file system at CERN)
- If you have a CERN account you can log-in in SWAN using your CERN credential
- Don't worry if you don't have a CERN computing account, some special accounts are provided for the course
 - Please fill in this case the form:
<https://goo.gl/forms/ZhoVzBjiQVEiTjsn2>
- If you have not used before **CERNBOX** you need to login first there:
 - cernbox.cern.ch

What can you do with ROOT?



Higgs analyses plots done with ROOT!

ROOT in a Nutshell

- ROOT is a software toolkit which provides building blocks for:

- Data processing
- Data analysis
- Data visualisation
- Data storage

An Open Source Project

*All contributions are warmly
welcome!*



- ROOT is written mainly in C++ (C++11 standard)
 - Bindings for Python are provided.
- Adopted in High Energy Physics and other sciences (but also industry)
- ~250 PetaBytes of data in ROOT format on the LHC Computing Grid
- Fits and parameters' estimations for discoveries (e.g. the Higgs)
- Thousands of ROOT plots in scientific publications

ROOT in a Nutshell

ROOT can be seen as a collection of building blocks for various activities, like:

- Data analysis: histograms, graphs, trees
- I/O: row-wise, column-wise storage of **any** C++ object
- Statistical tools (RooFit/RooStats): rich modeling and statistical inference
- Math: non trivial functions (e.g. Erf, Bessel), optimised math functions (VDT)
- C++ interpretation: fully C++11 compliant
- Multivariate Analysis (TMVA): e.g. Boosted decision trees, neural networks
- Advanced graphics (2D, 3D, event display).
- And more: HTTP servering, JavaScript visualisation.

Jupyter Notebooks

- A web-based interactive computing interface and platform that combines code, equations, text and visualisations.
 - Many supported languages: Python, Haskell, Julia, R ... One generally speaks about a “kernel” for a specific language
- No need to download any software
 - can execute code step by step
 - can display text and formulae
- In a nutshell: an “interactive shell opened within the browser”



<http://www.jupyter.org>

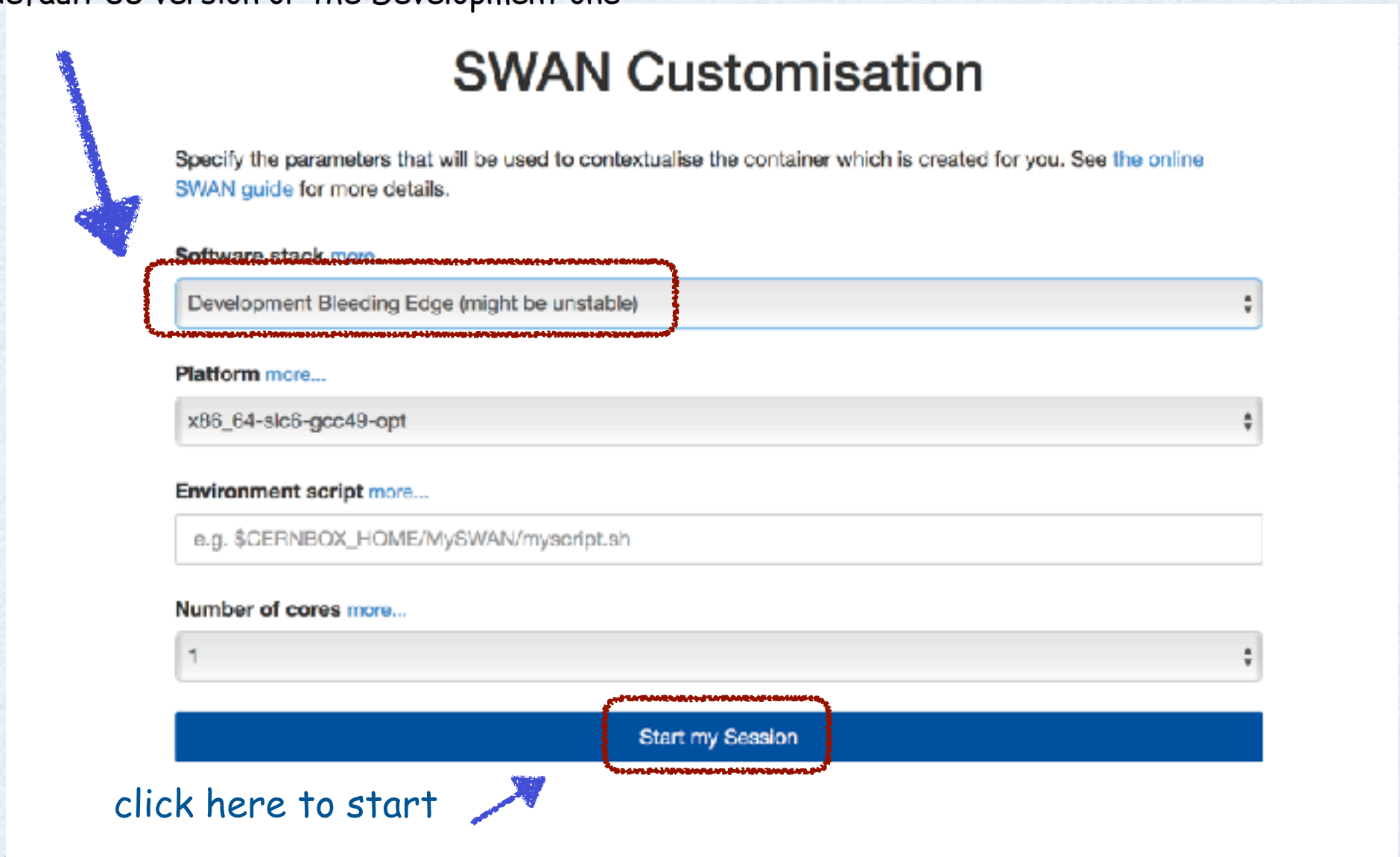
SWAN

- go to swan.cern.ch



Starting SWAN

make sure to have selected
the default 88 version or the Development one



The screenshot shows the 'SWAN Customisation' web interface. A blue arrow points to the 'Software stack' dropdown menu, which is highlighted with a red dashed box and contains the text 'Development Bleeding Edge (might be unstable)'. Below this, the 'Platform' dropdown menu is set to 'x86_64-slc6-gcc49-opt'. The 'Environment script' text input field contains 'e.g. \$CERNBOX_HOME/MySWAN/myscript.sh'. The 'Number of cores' dropdown menu is set to '1'. At the bottom, a blue button labeled 'Start my Session' is highlighted with a red dashed box. A blue arrow points to this button with the text 'click here to start'.

SWAN Customisation

Specify the parameters that will be used to contextualise the container which is created for you. See [the online SWAN guide](#) for more details.

Software stack [more...](#)

Development Bleeding Edge (might be unstable)

Platform [more...](#)

x86_64-slc6-gcc49-opt

Environment script [more...](#)

e.g. \$CERNBOX_HOME/MySWAN/myscript.sh

Number of cores [more...](#)

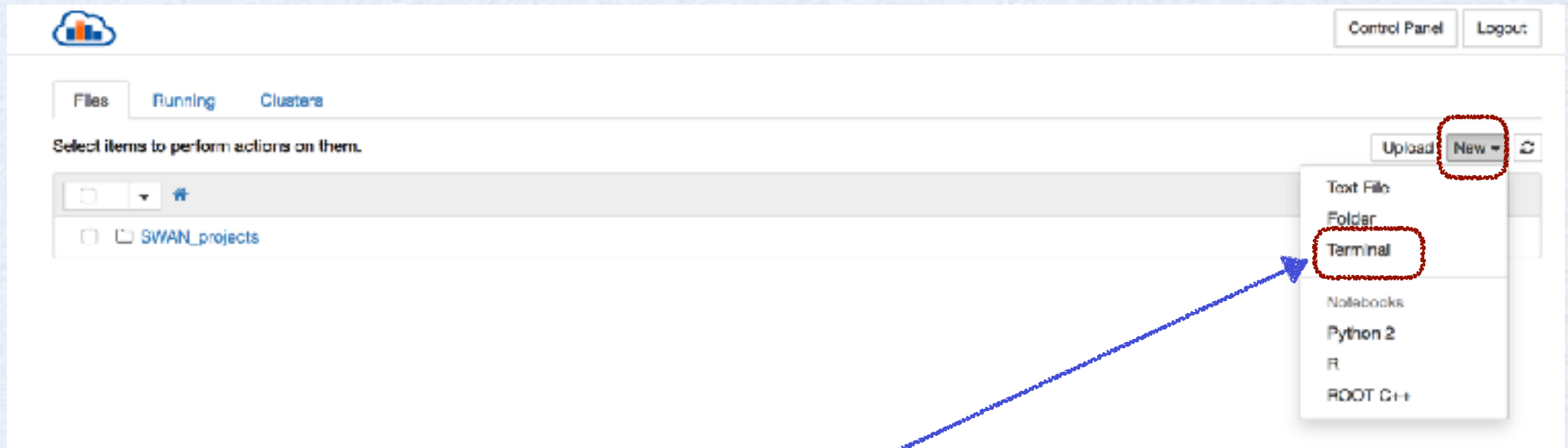
1

Start my Session

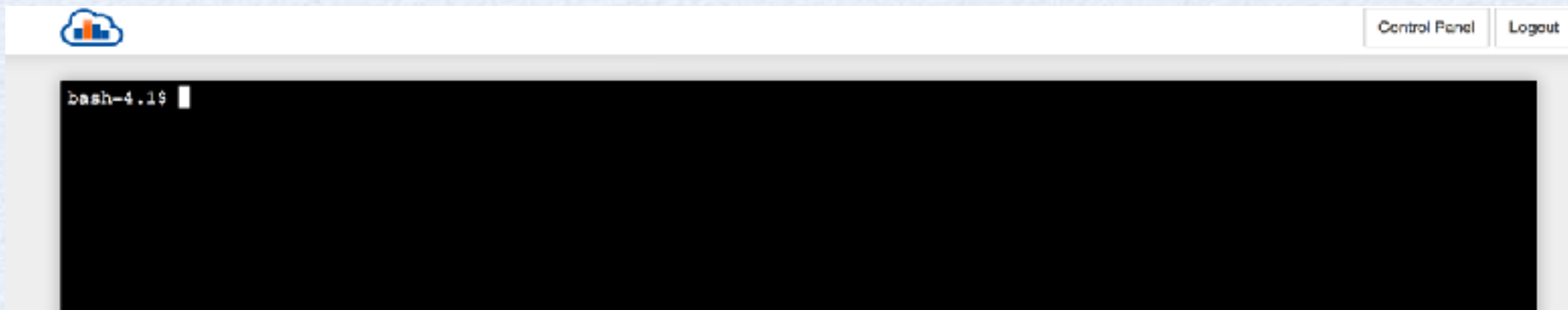
click here to start

Starting a Terminal in SWAN

After login cernbox home directory will be visible



Start a terminal window



Downloading the Notebooks

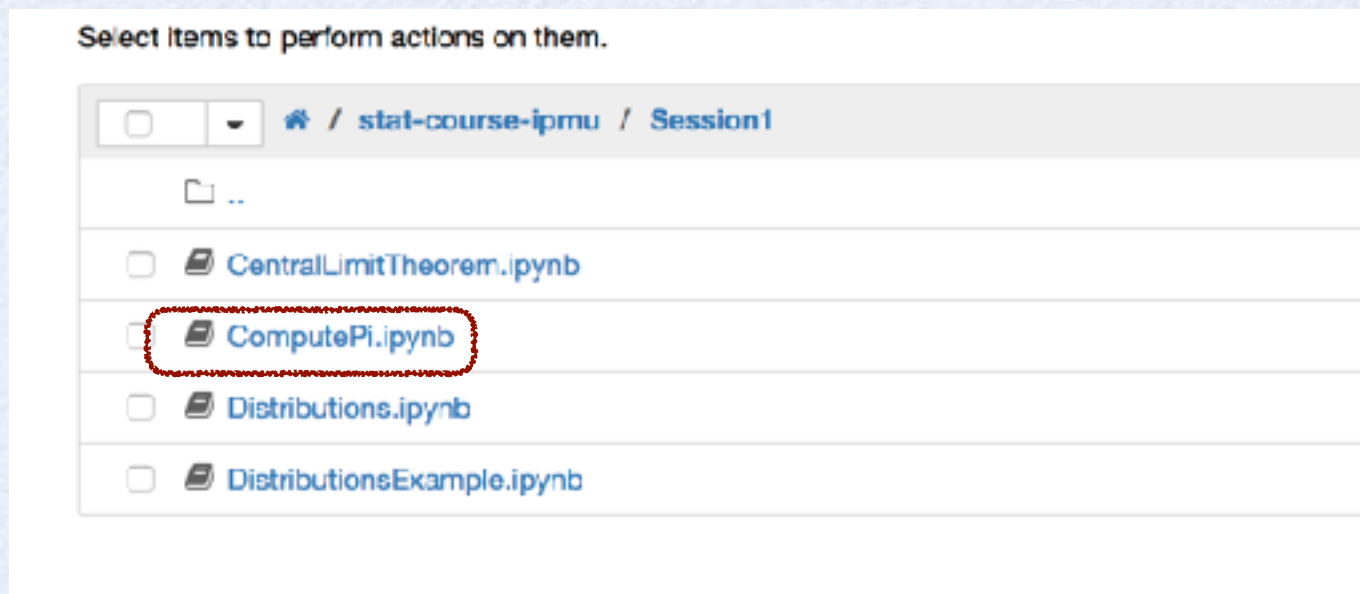
- After starting the terminal, get the notebooks for today's session.
- Download them from github

git clone <https://github.com/lmoneta/stat-course-ipmu.git>

```
bash-3.2$ git clone https://github.com/lmoneta/stat-course-ipmu.git
Cloning into 'stat-course-ipmu'...
remote: Counting objects: 53, done.
remote: Compressing objects: 100% (30/30), done.
remote: Total 53 (delta 23), reused 52 (delta 22), pack-reused 0
Unpacking objects: 100% (53/53), done.
bash-3.2$ cd stat-course-ipmu/Session1
bash-3.2$
```


Starting a Notebook

- After having downloaded the repository, go back to SWAN home page and go to the notebook folder
- ```
bash-3.2$ cd stat-course-ipmu/Session1
bash-3.2$
```
- select on of the notebook or create a new one



# Starting a Notebook



## Computation of the number $\pi$ using Monte Carlo

```
In [1]: n = 1000000000
(int) 1000000000
```

### Define the random number generator to use

```
In [2]: TRandom * r = new TRandom3();
```

```
In [3]: t = new TStopwatch()
(TStopwatch *) 0x7f7fab37cc40
```

### Computation of $\pi$

We compute  $\pi$  by calculating the area of a quarter of a circle with radius 1 using a Monte Carlo method. We generate 2 numbers  $(x,y)$  between 0 and 1 and we accept the numbers which are inside the circle:  $x^2 + y^2 \leq 1$



# References

- Presentation about Jupiter notebooks and SWAN
  - <https://cernbox.cern.ch/index.php/s/cpqdIOaNIwrlzJu>
- Basics of ROOT for fitting
  - <https://cernbox.cern.ch/index.php/s/SjJnYSkJrY9Kv8y>
- RooFit tutorial slides
  - <https://cernbox.cern.ch/index.php/s/kzGv4406IrpFpVM>
- RooStats tutorial slides
  - <https://cernbox.cern.ch/index.php/s/9DWJCObingTWB3r>