

# Table of Contents

<b>Extended JTerm Pre-Workshop Exercises - First Set - 15 Oct 2009.....</b>	<b>1</b>
Introduction.....	1
How to obtain an LPC computer account.....	1
Exercise 1 - Cut and Paste.....	1
Exercise 2 - Simple Edit Exercise.....	2
Exercise 3 - Setup a release area CMSSW_3_1_4.....	3
Exercise 4 - Find data in DBS ( Database Bookkeeping Service).....	3
Exercise 5 - EDM ( Event Data Model framework) standalone utilities - edmDumpEventContent, edmProvDump, edmEventSize.....	3
Exercise 6 - Make a PATtuple where PAT means Physics Analysis.....	5
Get Grid Certificate and CMS VO registration.....	6
Get CERN account.....	6
Exercise 7 - Fireworks - CMS Event Display.....	7
Exercise 8 - Grid/Crab exercises ( A valid Grid Certificate and CMS VORMS membership is a pre-requisite for this) - to be completed by ERIC VAANDERING.....	10
Exercise 9 - Run EDAnalyzer to make your PATtuple.....	10
Exercise 10 - FWLite exercises.....	11

# Extended JTerm Pre-Workshop Exercises - First Set - 15 Oct 2009

## Introduction

The purpose of the pre-workshop exercises is for prospective workshop attendees to become familiar with the basic software tools required to perform physics analysis at CMS. Please run and complete these exercises and post the results in the online form provided. The exercises are standalone. A large amount of additional information about these exercises is available in the twikis that we reference. Please remember that twikis evolve but aim to provide the best information at any time.

To perform the first set of exercises an LPC account is required.

## How to obtain an LPC computer account

- cmslpc account
  - ◆ Exercises will be performed on `cmslpc.fnal.gov` cluster.

**When logging in to `cmslpc` from a windows machine, the following links might be helpful:**

- Windows Kerberos at FNAL [↗](#)
- Hints for setting up Cygwin [↗](#)

AFTER the above you are ready for the exercises below We assume that you are familiar with basic unix commands like `ls`, `ls -altrh`, `cp` etc. and can edit a file using any of the editors like `pico`, `emacs`, `nedit` etc.

## Exercise 1 - Cut and Paste

Login to `cmslpc.fnal.gov`

To verify if cut and paste from one window to another works, cut and paste the following and then hit return

```
~cplager/runThisCommand.py "asdf;klasdjf;kakjsdf;akjf;aksdljf;a"\
"sldjfqewradsfafaw4efaefawefzdxffasdfw4ffawefawe4fawasdffadsfef"
```

The response should be your username followed by alphabets unique to your username, for example like this:

```
success: malik znyvx
```

QUESTION - Post this unique alphanumeric characters.

If you only run the command without any cut and paste, like the following:

```
somebody@cmslpc11> ~cplager/runThisCommand.py
```

you should get

```
Error: You must provide the secret key
```

If you paste incorrectly, you should get

```
Error: You didn't paste the correct input string
```

If you run it from a wrong computer ( not cmslpc, say from your laptop locally), you should get

```
bash: ~cplager/runThisCommand.py: No such file or directory
```

OR

```
Unknown user: cplager.
```

## Exercise 2 - Simple Edit Exercise

This is to test and make sure that the user can edit files on cmslpc.

Log into cmslpc, run this command:

```
cp ~cplager/editThisCommand.py .
```

Then open `editThisCommand.py` in your editor and edit the 11th line adding a # (hash character) to the front of the line. So the lines should start of looking like this:

```
# Please comment the line below out by adding a '#' to the front of  
# the line.  
raise RuntimeError, "You need to comment out this line with a #"
```

and be changed to:

```
# Please comment the line below out by adding a '#' to the front of  
# the line.  
#raise RuntimeError, "You need to comment out this line with a #"
```

Save the file and run the command:

```
user@cmslpc12> ./editThisCommand.py
```

If this is successful, you will see this

```
cplager@cmslpc12> ./editThisCommand.py  
success: cplager 0x1851DCBA
```

QUESTION - Paste the bottom line into the window.

If you did not successfully edit the file, you'll see an error message such as this:

```
cplager@cmslpc12> ./editThisCommand.py  
Traceback (most recent call last):  
  File "./editThisCommand.py", line 11, in ?  
    raise RuntimeError, "You need to comment out this line with a #"  
RuntimeError: You need to comment out this line with a #
```

## Exercise 3 - Setup a release area CMSSW\_3\_1\_4

```
source /uscmsst1/prod/sw/cms/cshrc uaf
cmscvsroot CMSSW
mkdir YOURWORKINGAREA
cd YOURWORKINGAREA
scram p CMSSW CMSSW_3_1_4
cd CMSSW_3_1_4/src
cmsenv
```

Run the following command:

```
echo $CMSSW_BASE
```

QUESTION - Paste the output of the above command

## Exercise 4 - Find data in DBS ( Database Bookkeeping Service)

Go to the url [DBS discovery](#) and in the menu driven interface, choose the info from the pull down menu, so that your choices look like this:

▣ Show result... ▣ Hide result...

### ADVANCED KEYWORD SEARCH

DBS instances  [HELP](#)

### MENU-DRIVEN INTERFACE

Physics groups

Data tier   
 composed tier, e.g. GEN-SIM:

Software releases

Data types

Primary dataset/  
MC generators

then click on "Find" and after few seconds it brings another page. On this page there are two data sets. Look for the one that says: /RelValTTbar/CMSSW\_3\_1\_4-STARTUP31X\_V2-v1/GEN-SIM-RECO.

QUESTIONS - What is the size of this data? Click on "plain" to see the number of files it contains. How many files does it have? Is this data at FNAL?

The files it contains should look like this

```
'/store/relval/CMSSW_3_1_4/RelValTTbar/GEN-SIM-RECO/STARTUP31X_V2-v1/0006/AC0641BB-73B1-DE
.....
```

More information about on accessing data in DBS can be found in [WorkBookDataSamples](#)

## Exercise 5 - EDM ( Event Data Model framework) standalone utilities - edmDumpEventContent, edmProvDump,

## edmEventSize

The overall collection of CMS software, referred to as CMSSW, is built around a Framework, an Event Data Model (EDM), and Services needed by the simulation, calibration and alignment, and reconstruction modules that process event data so that physicists can perform analysis. The primary goal of the Framework and EDM is to facilitate the development and deployment of reconstruction and analysis software. The CMS Event Data Model (EDM) is centered around the concept of an Event. An Event is a C++ object container for all RAW and reconstructed data related to a particular collision. To understand what is in a datafile and more, several EDM utilities are available. In this exercise, one will use three of the several EDM utilities available. They would very useful throughout one's physics analysis beyond this workshop. More on these EDM utilities can be found at [WorkBookEdmUtilities](#). These together with [CMSSW](#) and [CMS LXR Cross Referencer](#) can help you understand the CMS code and also write your own analysis specific code.

- Use `edmDumpEventContent` to see what class names etc. to use in order to access the objects in the RECO data file you located above

- ◆ To do this do

```
edmDumpEventContent --all --regex caloJet dcap:///pnfs/cms/WAX/11/store/relval/CMSSW_3_1_4/RelValTTbar/GEN-SIM
```

- ◆ Note how `/pnfs/cms/WAX/11/` has been prefixed to the file name. Adding this gives you the physical location of the file at Fermilab.
- ◆ Open and look at the file `EdmDumpEventContent.txt`. It has information divided into four columns (roughly). The first column is C++ class type of the data, second is module label, third is product instance label and fourth is process name. You can read more at [Identifying Data in the Event](#).
- ◆ QUESTION - How many types of CaloJet module labels are there? What are there names
- ◆ NOTE: Instead of the above you can also try above without options `--all --regex caloJet`. This will dump the entire event content as follows:

- ◇ To do this do

```
edmDumpEventContent dcap:///pnfs/cms/WAX/11/store/relval/CMSSW_3_1_4/RelValTTbar/GEN-SIM
```

- ◇ This is a file with many more lines than the one with option `--all --regex caloJet`

- To aid in understanding the full history of an analysis, the framework accumulates provenance for all data stored in the standard ROOT output files. Using `edmProvDump` one can print out all the tracked parameters which were used to create the data file. One can see what modules were run, CMSSW version etc. used when the RECO file was made. To use it do

- ◆ To do this do

```
edmProvDump dcap:///pnfs/cms/WAX/11/store/relval/CMSSW_3_1_4/RelValTTbar/GEN-SIM
```

- ◆ NOTE: `EdmProvDump.txt` is a huge file. Open and look at this file and locate `Processing History` ( about 20 lines from the top).
- ◆ QUESTION - Which version of `CMSSW_???` does the processing history say that was used for processing the data?

- You can use `edmEventSize` to know the size of different branches in your data file. The details about are here [SWGGuideEdmEventSize](#)

- ◆ To do this do

```
edmEventSize -v dcap:///pnfs/cms/WAX/11/store/relval/CMSSW_3_1_4/RelValTTbar/GEN-SIM
```

- ◆ Open and look at file `EdmEventSize.txt` and locate the line `recoCaloJets_antikt5CaloJets__RECO..`. There are two numbers next to it that measure plain and the compressed size of this branch.
- ◆ QUESTION - What are two numbers next to above line? Write them.

## Exercise 6 - Make a PATtuple where PAT means Physics Analysis

What is a PATuple and all about it can be found at [SWGuidePAT](#) and the latest PAT tutorial [PAT tutorial](#).

Using the `WorkBookPATSimpleExample` in Module 1 of PAT tutorial [PAT tutorial](#), we will plot the  $P_T$  distribution of the PAT muons. Note that we will use the datafile ( which is a RECO file) that you found above in the DBS instead of the default data file used in Module 1 ( you can dig out the default file later) but for now we will simply make a PAT tuple called `simplePAT.root`.

Make sure at this point your are in the directory `YOURWORKINGAREA/CMSSW_3_1_4/src` To complete this exercise follow these steps:

1. Open your favorite editor and cut paste the contents of the file `simplePAT_cfg.py`. Save this file as `simplePAT_cfg.py`.

2. Now run the following command

```
cmsRun simplePAT_cfg.py
```

While this command is executed you will see an output like this.

3. Now open the root file you just created. Note that this file is created in the `YOURWORKINGAREA/CMSSW_3_1_4/src` directory where you should always be for this tutorial. To open the root file do:

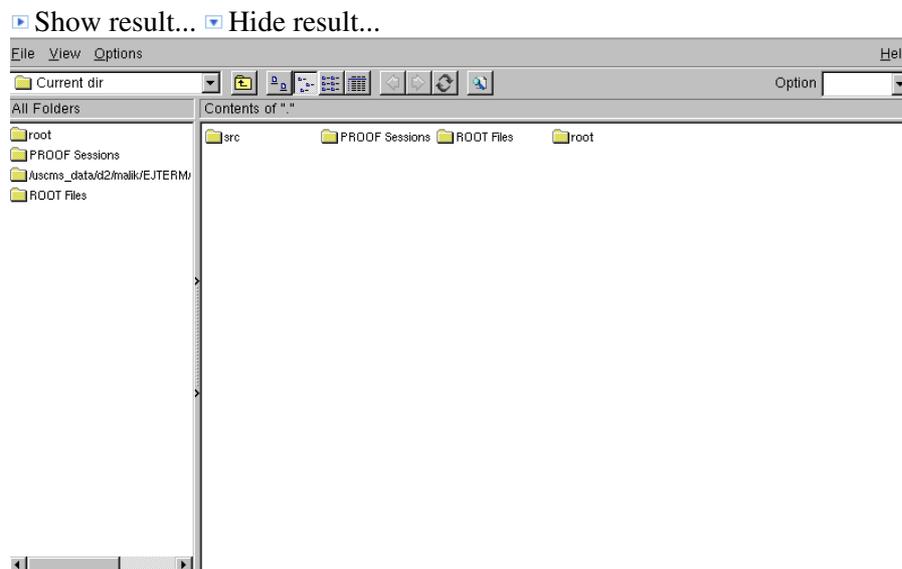
```
root -l simplePAT.root
```

On root prompt type `root [1] gStyle->SetOptStat(111111);` and then `TBrowser b;` like this

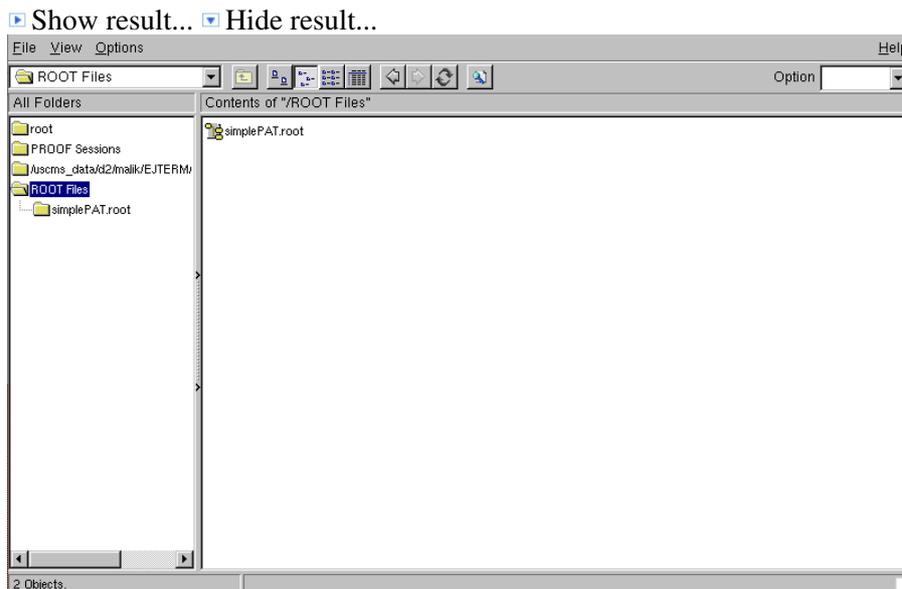
```
root [1] gStyle->SetOptStat(111111);
```

```
root [1] TBrowser b;
```

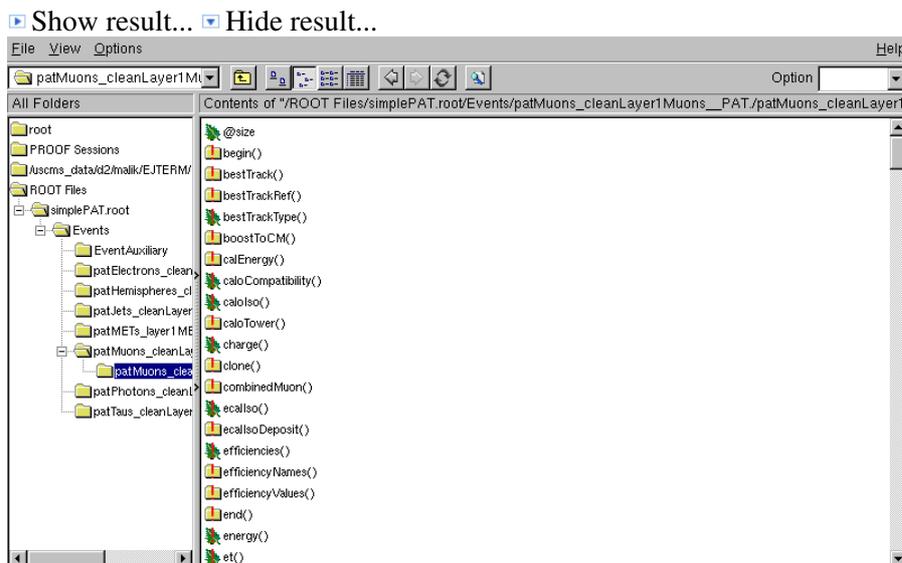
This opens a window that looks like this:



On this window click on `ROOT Files` on the left menu and now the window looks like this:



Click `simplePAT.root`, then on `Events`, then on `patMuons_cleanLayer1Muons__PAT` and then on `patMuons_cleanLayer1Muons__PAT.obj`. You should now see a window that looks like this:



Scroll way down ( not too fast) and click on `pt ()`. You should now see PAT Muon Pt distribution.  
QUESTION - What is the mean value of the muon `pt()`?

## Get Grid Certificate and CMS VO registration

Use the following link for this: [Get Your Grid Certificate and CMSVO](#)

## Get CERN account

\* Use the following link for CMS CERN account: [CMS CERN account](#)

- ♦ CERN account is needed to login in to any elearning web-site ( later used in the final tutorial), to get a file from afs area or any future need you may have beyond the tutorial.
- ♦ **To keep the information flow going for CERN account, please ask your team leader do the necessary "signing" after the online form has been treated by the secretariat.**

## Exercise 7 - Fireworks - CMS Event Display

Fireworks is CMS' graphical tool to display events for physics. We have renamed the file

/pnfs/cms/WAX/11/store/relval/CMSSW\_3\_1\_4/RelValTTbar/GEN-SIM-RECO/STARTUP31X\_V2-v1/0006/AC0641BB that you have been working with to EJTermDataForEventDisplay.root for simplicity.

There are the following two ways that one can star event display. Please be a little patient as few messages would pop up on the screen and it will take few seconds to open ( may be up to a minute).

\*\*\*\*\*

### 1. From cmslpc account ( slow but easy)

- To run it from your cmslpc account, run the following command from your  
YOURWORKINGAREA/CMSSW\_3\_1\_4/src working area

```
cmsShow /uscms_data/d2/malik/EJTERM/CMSSW_3_1_4/src/EJTermDataForEventDisplay.root
```

You can also run directly as follows by using the data file's url location as follows:

```
cmsShow http://www-d0.fnal.gov/~malik/EJTermDataForEventDisplay.root
```

\*\*\*\*\*

### 2. Running it locally on your desktop ( fast BUT one has to download the fireworks distribution locally on one's desktop)

- Using the recipe at WorkbookFireworks download the 31X distribution locally on your laptop
- Then change to the directory to cmsShow31 by doing

```
cd cmsShow31
```

- Then download the data file as follows
  - ◆ For LINUX laptop

```
wget http://www-d0.fnal.gov/~malik/EJTermDataForEventDisplay.root
```

- ◆ For Macintosh laptop

```
curl http://www-d0.fnal.gov/~malik/EJTermDataForEventDisplay.root > EJTermDataForEventDisplay.root
```

After you download the fireworks distribution run the following commands from your desktop's local command prompt to display fireworks ( you must be in the directory called cmsShow31.

```
./cmsShow EJTermDataForEventDisplay.root
```

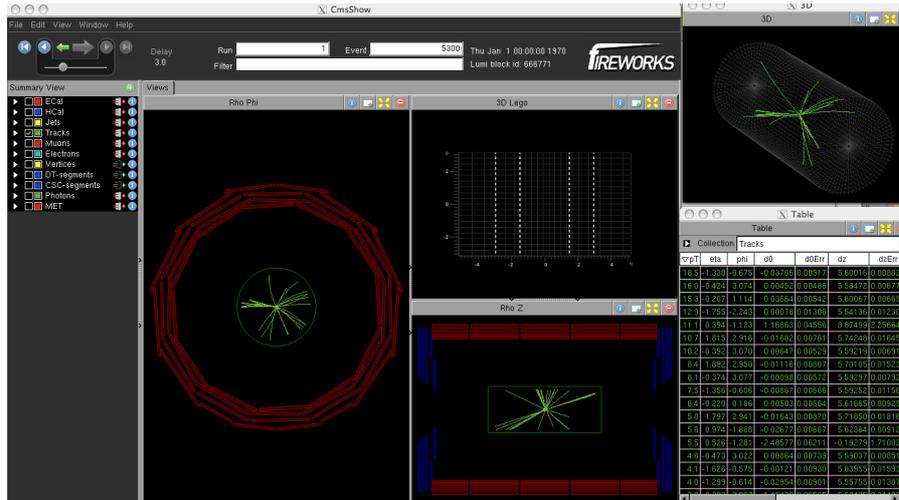
In case you do not want to download the file locally, you can run it directly from its url location as follows:

```
./cmsShow http://www-d0.fnal.gov/~malik/EJTermDataForEventDisplay.root
```

\*\*\*\*\*

Following either of the above method, the event display should start. You may UNCEHCK all the objects in the summary view on the left EXCEPT tracks. You should now be seeing the following type of graphics ( with just green tracks)

Show result... Hide result...



Once you have event display started three windows will open. We ignore the two small ones. The event display will display the first event. You can use the arrows on the top to go to next event or play/pause through all events.

To do the exercise do the following steps:

1. Write event number 8550 where you see event displayed as 8501, press enter.
2. Now you should see the green tracks displayed for event 8550. Not all the tracks are displayed though.
3. Click on the little arrow button on the left of the Tracks in the summary view
4. This will pull a display of all the tracks, showing pt, eta, phi etc., like this:

Show result... Hide result...

<input type="checkbox"/> <input type="checkbox"/> ECal <input type="checkbox"/> <input type="checkbox"/> HCal <input type="checkbox"/> <input type="checkbox"/> Jets <input checked="" type="checkbox"/> <input type="checkbox"/> Tracks				pt	eta	phi
<input type="checkbox"/> <input type="checkbox"/> Track 0	0.6	-0.5	-3.0			
<input type="checkbox"/> <input type="checkbox"/> Track 1	1.3	-0.3	-3.0			
<input type="checkbox"/> <input type="checkbox"/> Track 2	0.9	1.2	-2.8			
<input type="checkbox"/> <input type="checkbox"/> Track 3	0.8	-1.8	-2.7			
<input type="checkbox"/> <input type="checkbox"/> Track 4	2.0	-1.7	-2.5			
<input type="checkbox"/> <input type="checkbox"/> Track 5	1.2	-1.4	-2.6			
<input type="checkbox"/> <input type="checkbox"/> Track 6	0.7	-1.2	-2.4			
<input type="checkbox"/> <input type="checkbox"/> Track 7	0.7	0.8	-2.3			
<input checked="" type="checkbox"/> <input type="checkbox"/> Track 8	5.6	1.0	-1.9			
<input checked="" type="checkbox"/> <input type="checkbox"/> Track 9	2.3	1.0	-1.8			
<input type="checkbox"/> <input type="checkbox"/> Track 10	0.8	0.8	-1.6			
<input type="checkbox"/> <input type="checkbox"/> Track 11	2.0	0.6	-1.6			
<input type="checkbox"/> <input type="checkbox"/> Track 12	1.8	0.1	-1.5			
<input type="checkbox"/> <input type="checkbox"/> Track 13	1.2	0.3	-1.5			
<input type="checkbox"/> <input type="checkbox"/> Track 14	1.0	0.5	-1.2			
<input type="checkbox"/> <input type="checkbox"/> Track 15	0.7	0.0	-1.0			
<input type="checkbox"/> <input type="checkbox"/> Track 16	0.7	-1.4	-0.8			
<input type="checkbox"/> <input type="checkbox"/> Track 17	0.8	-1.3	-0.9			
<input type="checkbox"/> <input type="checkbox"/> Track 18	2.0	-1.4	-0.8			
<input checked="" type="checkbox"/> <input type="checkbox"/> Track 19	18.5	-1.3	-0.7			
<input type="checkbox"/> <input type="checkbox"/> Track 20	1.1	-1.5	-0.7			
<input checked="" type="checkbox"/> <input type="checkbox"/> Track 21	4.0	-1.3	-0.6			
<input type="checkbox"/> <input type="checkbox"/> Track 22	2.5	-1.4	-1.5			

5. Scroll down and see how many tracks this event has.
6. QUESTION - What are the number of tracks in this event 8550?
7. You can of course play with other buttons and menu.

## Exercise 8 - Grid/Crab exercises ( A valid Grid Certificate and CMS VORMS membership is a pre-requisite for this) - to be completed by ERIC VAANDERING

- a. Exercise on having a grid certificate
- b. Getting a /store/user area from your assigned Tier2
- c. Running a Grid Job

## Exercise 9 - Run EDAnalyzer to make your PATtuple

If you want to manipulate, analyze, draw histograms etc. from the data in the PATtuple, you need can do it using EDAnalyzer or in FWLite. More on these can be found on WorkbookWriteFrameworkModule and FWLiteExecutable. While EDAnalyzer needs full framework to run, FWLite can be run on your laptop by downloading the appropriate FWLite distribution.

In the `/YOURWORKINGAREA/CMSSW_3_1_4/src` directory cut and paste the following commands ( you can select all and paste all together and commands will be executed sequentially)

```
cvs co -r patTutorial_sept09_module1 PhysicsTools/PatExamples/src
cvs co -r patTutorial_sept09_module1 PhysicsTools/PatExamples/interface
cvs co -r patTutorial_sept09_module1 PhysicsTools/PatExamples/BuildFile
cvs co -r patTutorial_sept09_module1 PhysicsTools/PatExamples/plugins/BuildFile
cvs co -r patTutorial_sept09_module1 PhysicsTools/PatExamples/plugins/PatBasicAnalyzer.cc
cvs co -r patTutorial_sept09_module1 PhysicsTools/PatExamples/test/analyzePatBasics_cfg.py
```

To run, you need to compile the code by doing the following

```
scram b
```

Now you open the file (using your favourite editor), like this

```
pico PhysicsTools/PatExamples/test/analyzePatBasics_cfg.py
```

and change the input file to "simplePAT.root you created in Exercise 4. This is done like below in the config file `analyzePatBasics_cfg.py`

```
process.source = cms.Source("PoolSource",
    fileNames = cms.untracked.vstring(
        'file:simplePAT.root'
    )
)
```

To run the code do ( assuming you are still in the `src` directory)

```
cmsRun PhysicsTools/PatExamples/test/analyzePatBasics_cfg.py
```

After the config file runs, you should see a file called `analyzePatBasics.root` . Browse through different histograms.

QUESTIONS - How many histograms are there? Which histogram is empty?

## Exercise 10 - FWLite exercises

Go through the twiki - FWLiteExecutable

Note that FWLite twiki tells you to be in a working area by saying `cd $CMSSW_BASE/src`. You are already working in that area called `YOURWORKINGAREA/CMSSW_3_1_4/src`. Even if you execute the command `cd $CMSSW_BASE/src` you stay in the same working area where you already are.

AFTER completing the exercises on the twiki:

Open the root file `myZPeakModified.root` and draw the modified Z-peak as follows:

```
root -l myZPeakModified.root
root [0] Zmass->Draw();
```

QUESTION - Report the mean of the MODIFIED Z Mass.

Questions/Problems/Suggestion - [mailto: malik@fnalNOSPAMPLEASE.gov](mailto:malik@fnalNOSPAMPLEASE.gov) , Phone - 630-840-6441

-- SudhirMalik - 2009-10-06

---

This topic: [Sandbox > EJTermPreWorkshopOnlineWork](#)

Topic revision: r24 - 2009-10-18 - SudhirMalik



Copyright &© 2008-2021 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

or Ideas, requests, problems regarding TWiki? use [Discourse](#) or [Send feedback](#)