CMS Data Analysis

Daniela Gayoso
Colegio Latino Cordillera
Santiago, Chile

Patrick Mooney
Trinity School
South Bend, IN

Jill Ziegler
Hamilton High West
Hamilton, NJ

Neel Bhagat
Bethany Christian Schools
Goshen, IN

Patrick Evans
Trinity School
South Bend, IN

Logan Schwartzendruber
Bethany Christian Schools
Goshen, IN

Available Data Sets

- 2000 di-muon events around the J/ψ
- 2000 di-electron events around the J/ψ
- 2000 di-electron events around the Y
- 100,000 di-muon events in the invariant mass range 2-110 GeV
- 100,000 di-electron events in the invariant mass range 2-110 GeV
- 500 di-electron events around the Z boson
- 500 di-muon events around the Z boson
- 500 events of W to eν
- 500 events of W to μν
- 13 Higgs candidate events: H to γγ, 4e, 2e2μ, 4μ
- ~50 1/pb single muons for top quark analysis (508,561 events)

Available Data Sets

- 2000 di-muon events around the J/ψ
- 2000 di-electron events around the J/ψ
- 2000 di-electron events around the Y
- 100,000 di-muon events in the invariant mass range 2-110 GeV
- 100,000 di-electron events in the invariant mass range 2-110 GeV
- 500 di-electron events around the Z boson
- 500 di-muon events around the Z boson
- 500 events of W to eν
- 500 events of W to μν
- 13 Higgs candidate events: H to γγ, 4e, 2e2μ, 4μ
- ~50 1/pb single muons for top quark analysis (508,561 events)

Available Data Sets

- 2000 di-muon events around the J/ψ
- 2000 di-electron events around the J/ψ
- 2000 di-electron events around the Y
- 100,000 di-muon events in the invariant mass range 2-110 GeV
- 100,000 di-electron events in the invariant mass range 2-110 GeV
- 500 di-electron events around the Z boson
- 500 di-muon events around the Z boson
- 500 events of W to eν
- 500 events of W to μν
- 13 Higgs candidate events: H to γγ, 4e, 2e2μ, 4μ
- ~50 1/pb single muons for top quark analysis (508,561 events)

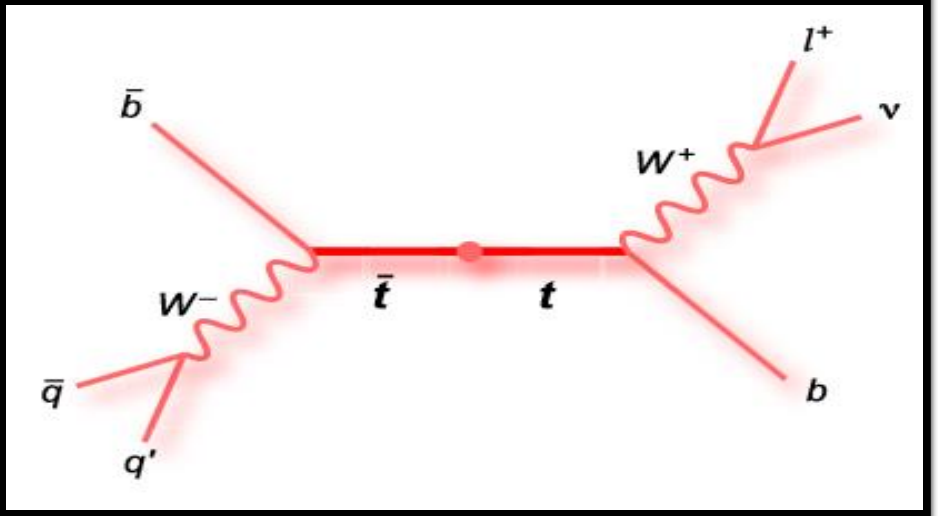
Why the $t\bar{t}$ Data Set?

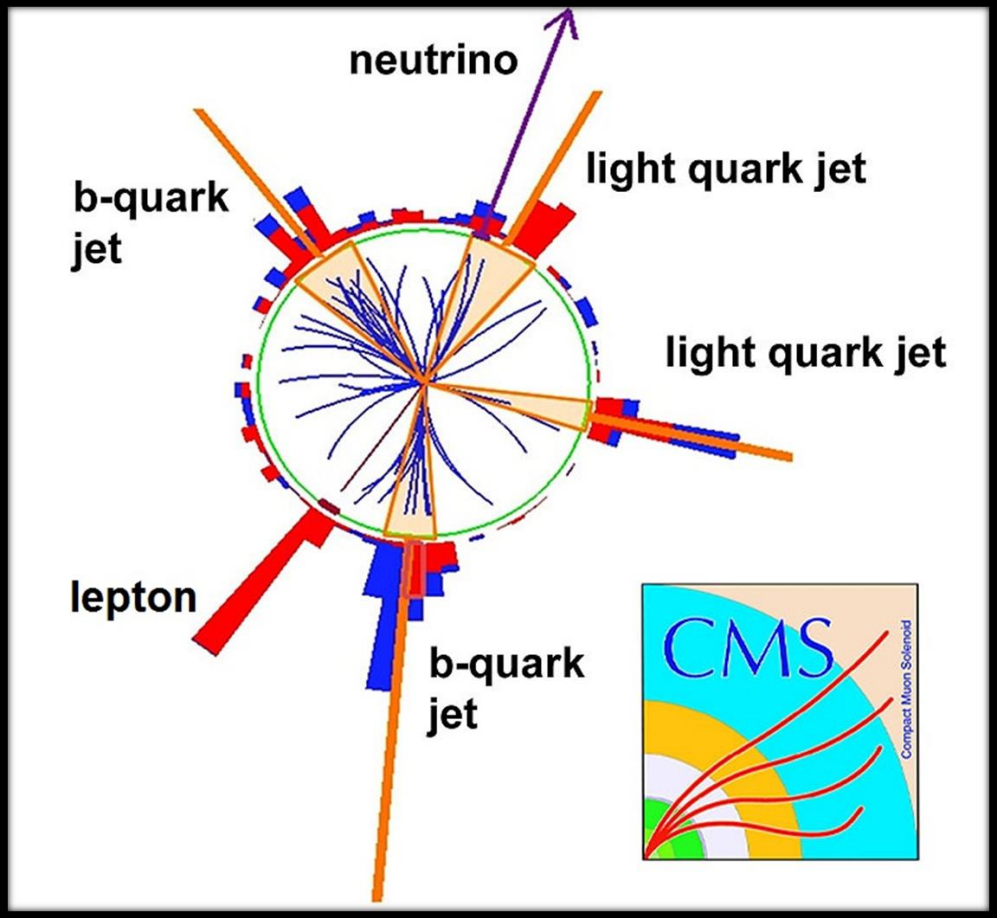
- Reasons for chosing the $t\bar{t}$ data set:
  - well-designed accompanying activity
  - activity requires use of simulated data to determine cuts and efficiencies
  - provides an opportunity to learn Linux, C++, and ROOT

Preparatory Activity

- We downloaded data and several example programs to our local Linux machine
  - created individual accounts for each member of the group
- We used MobaXterm to give us virtual terminals on that machine to access the data
- We studied C++ for two weeks before starting work with ROOT

$t\bar{t}$ Events

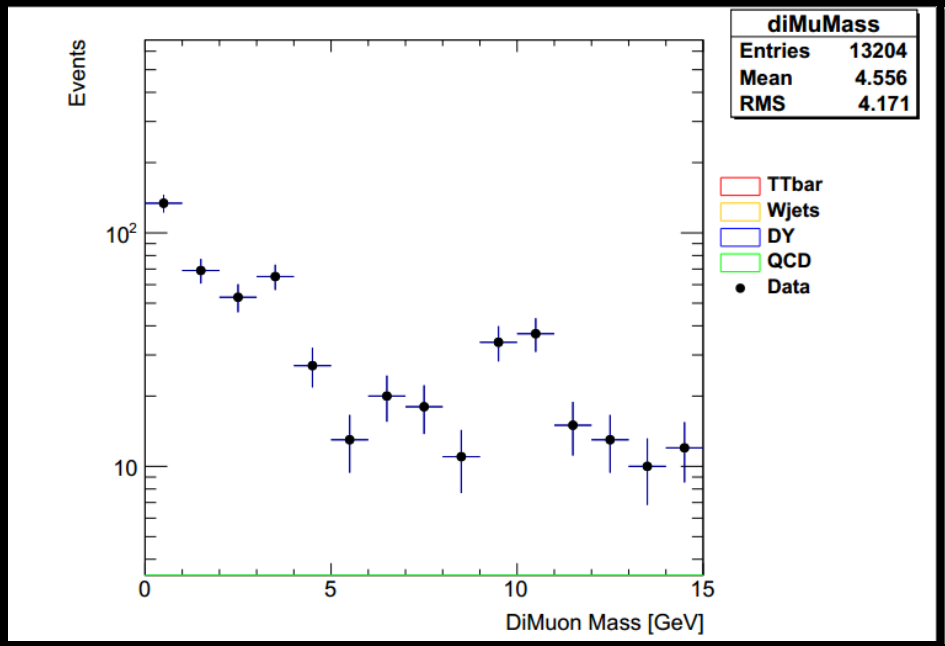


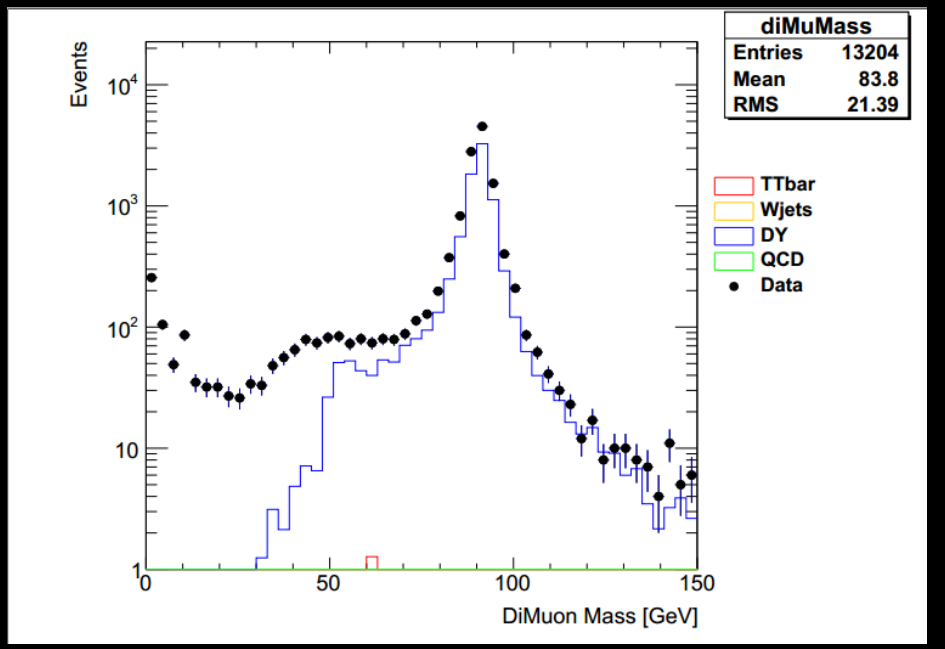


Preliminary Plots

1. Diagnostic Plots

2. DiMuon Mass Plots





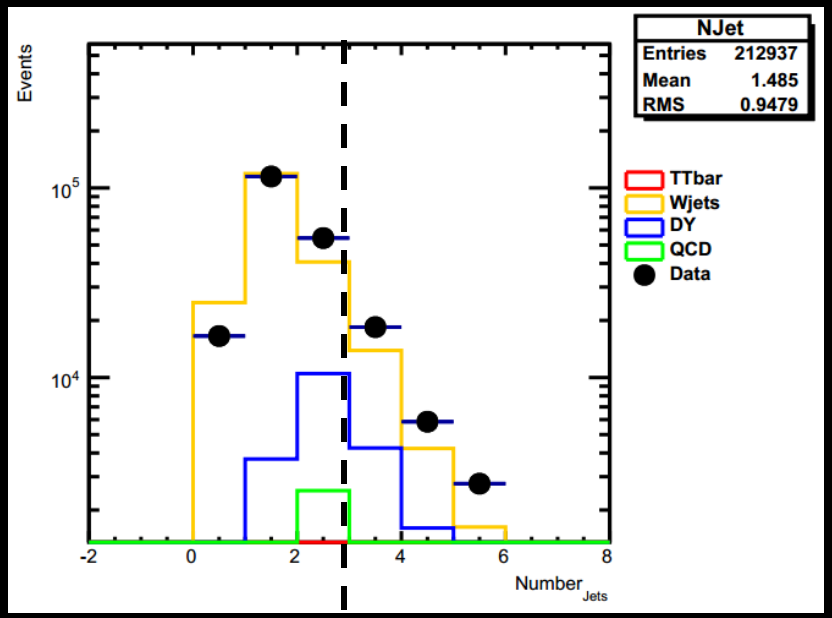
Enriching the Top Content of the Data Sample: Strategy

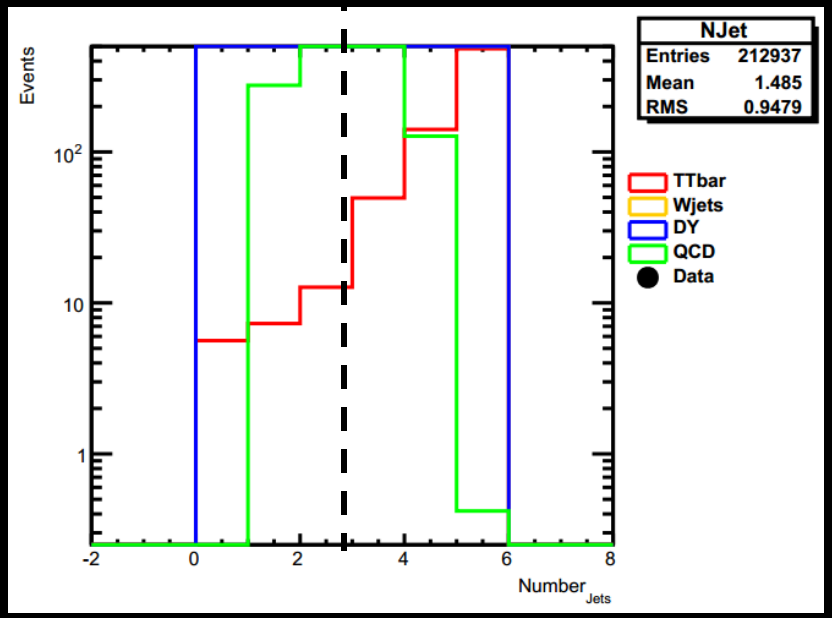
- Identify measured quantities that allow us to distinguish between $t\bar{t}$ events and expected background
- Number of simulated events
  - $t\bar{t}$: 383,167 (signal)
  - Drell-Yan: 109,656 (background)
  - W + Jets: 76,196 (background)
  - QCD: 137 (background)

  total: 569,156

Enriching the Top Content of the Data Sample

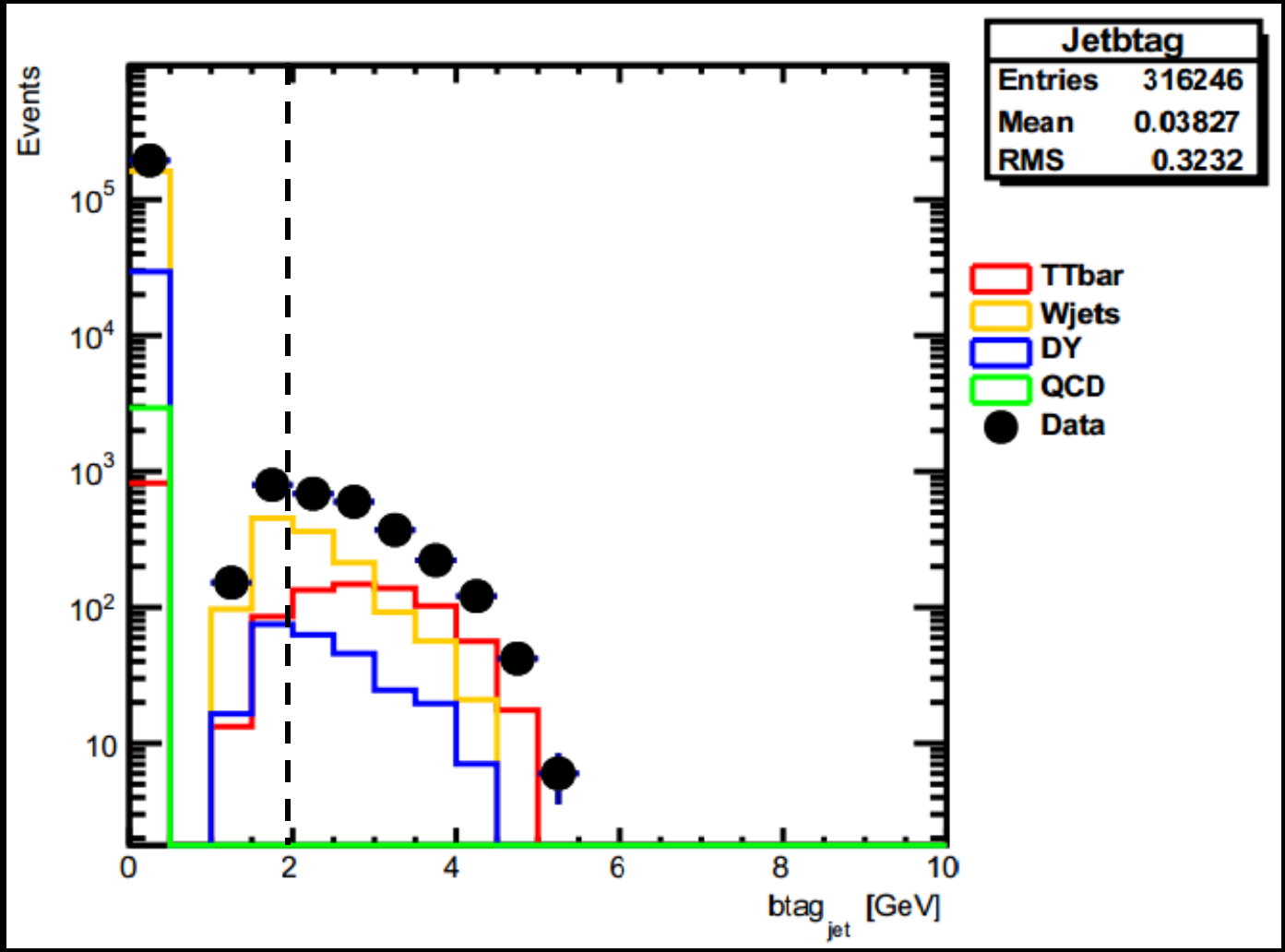
- NJets





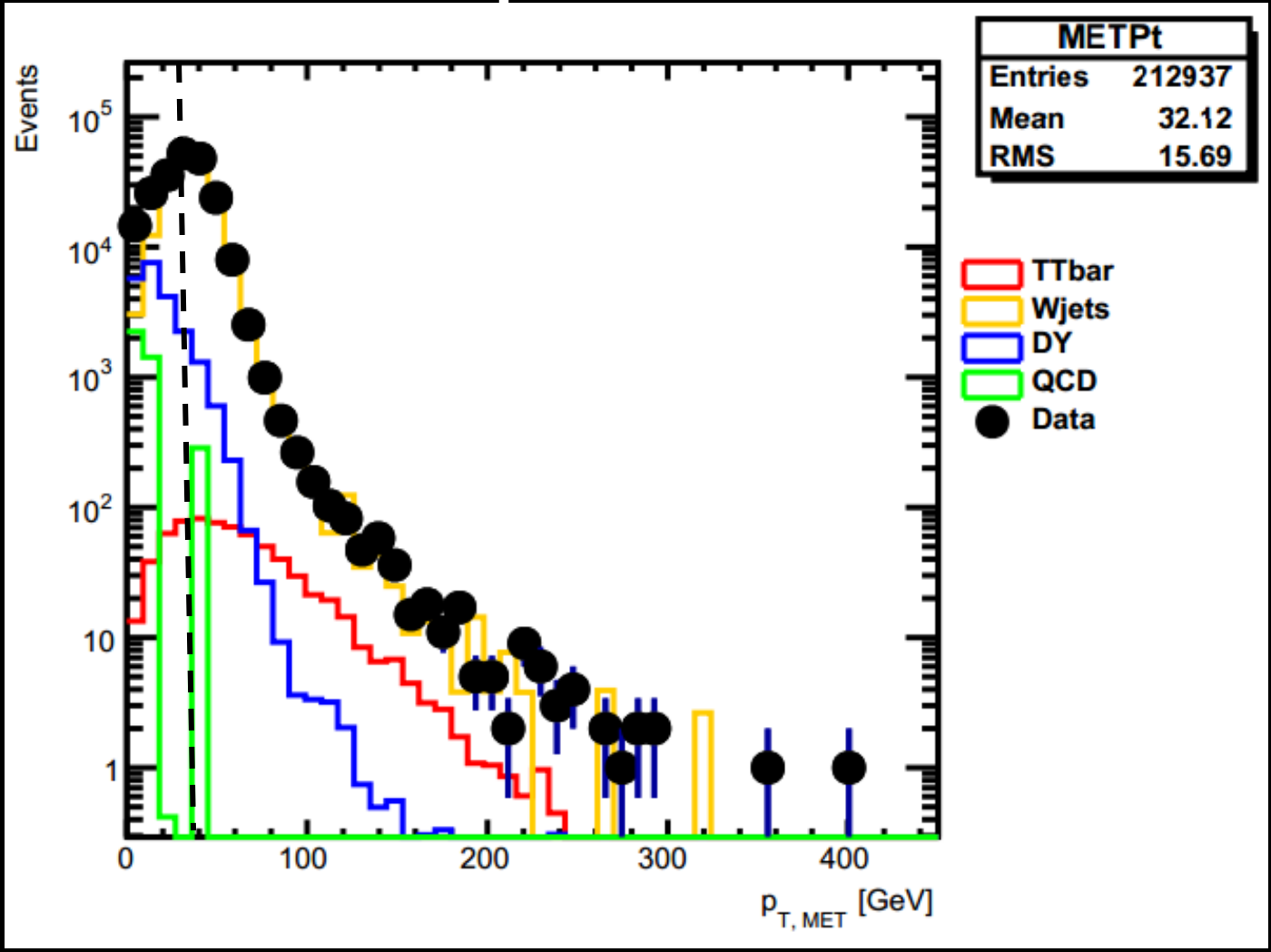
Enriching the Top Content of the Data Sample

- bTags



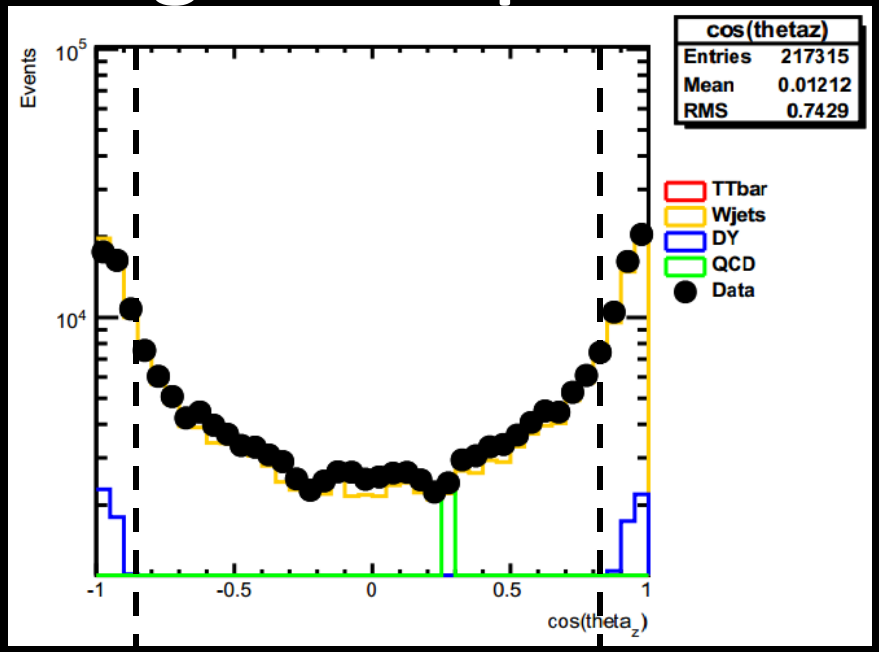
Enriching the Top Content of the Data Sample

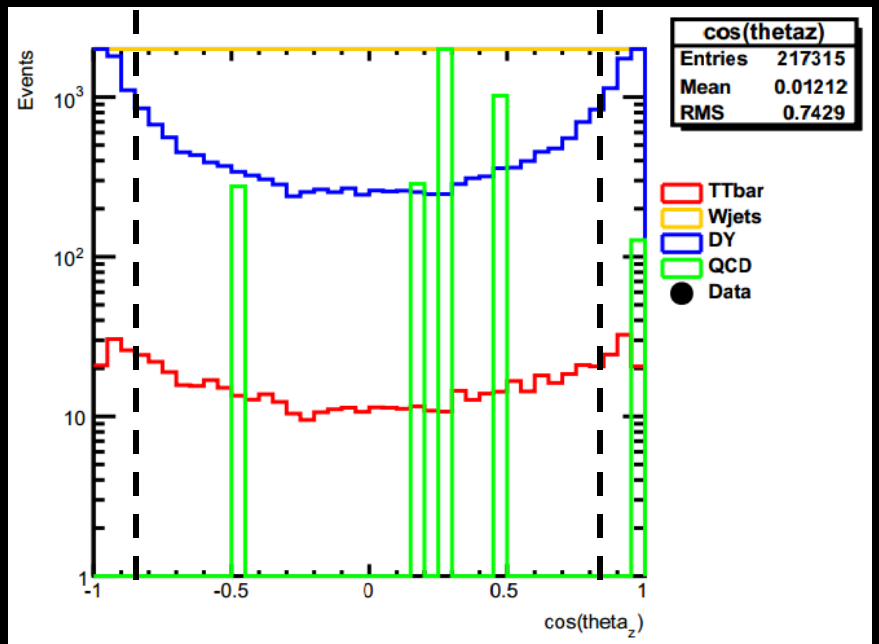
- MET



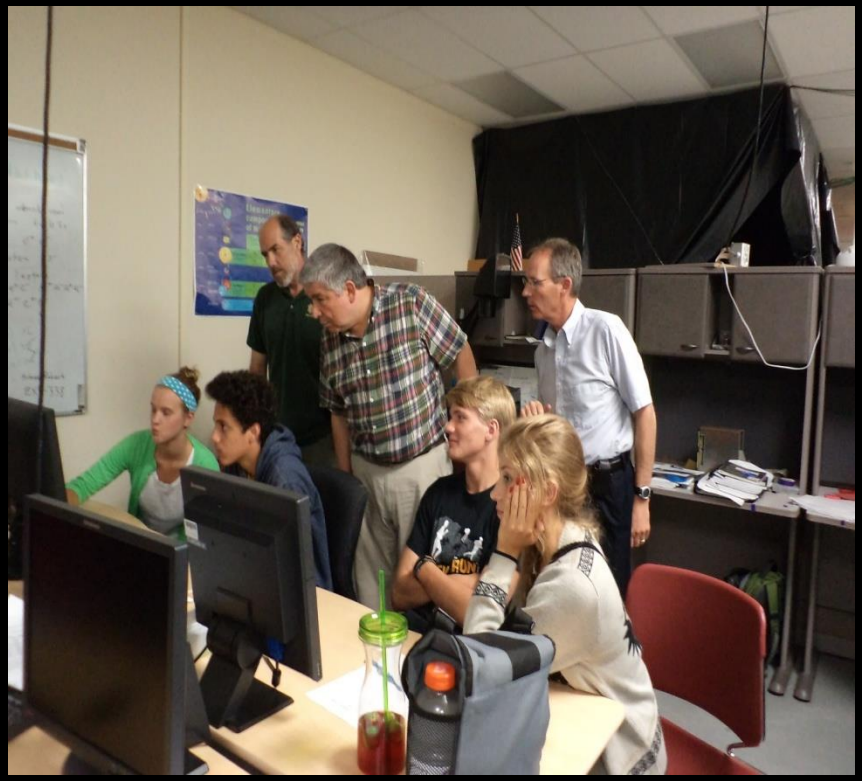
Enriching the Top Content

- Angular Distribution





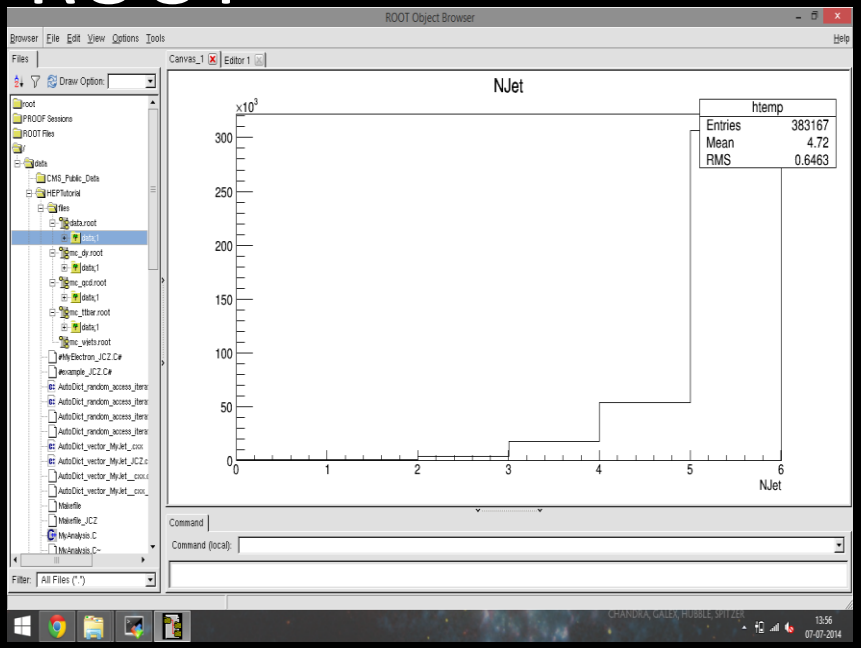
Quarknet



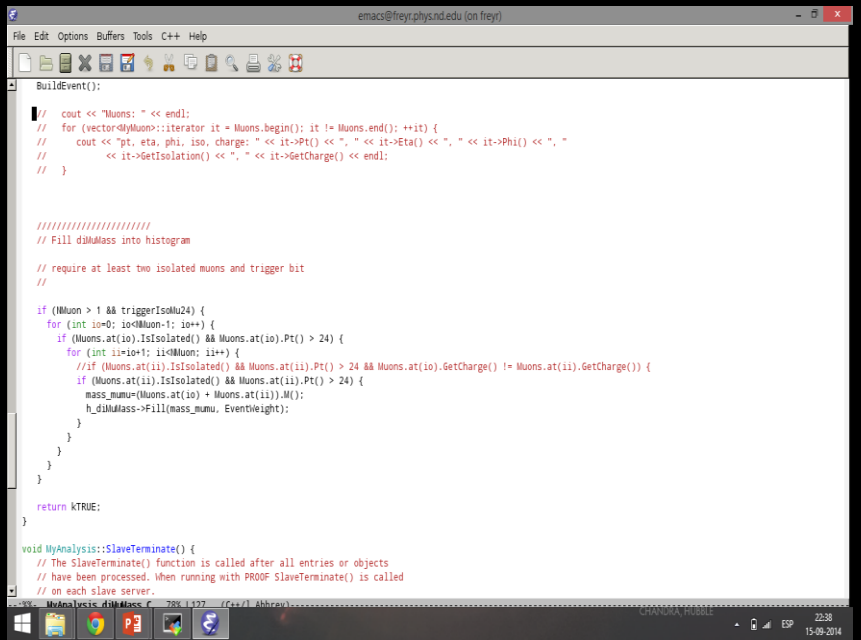
Quarknet

CMS Data Group

ROOT

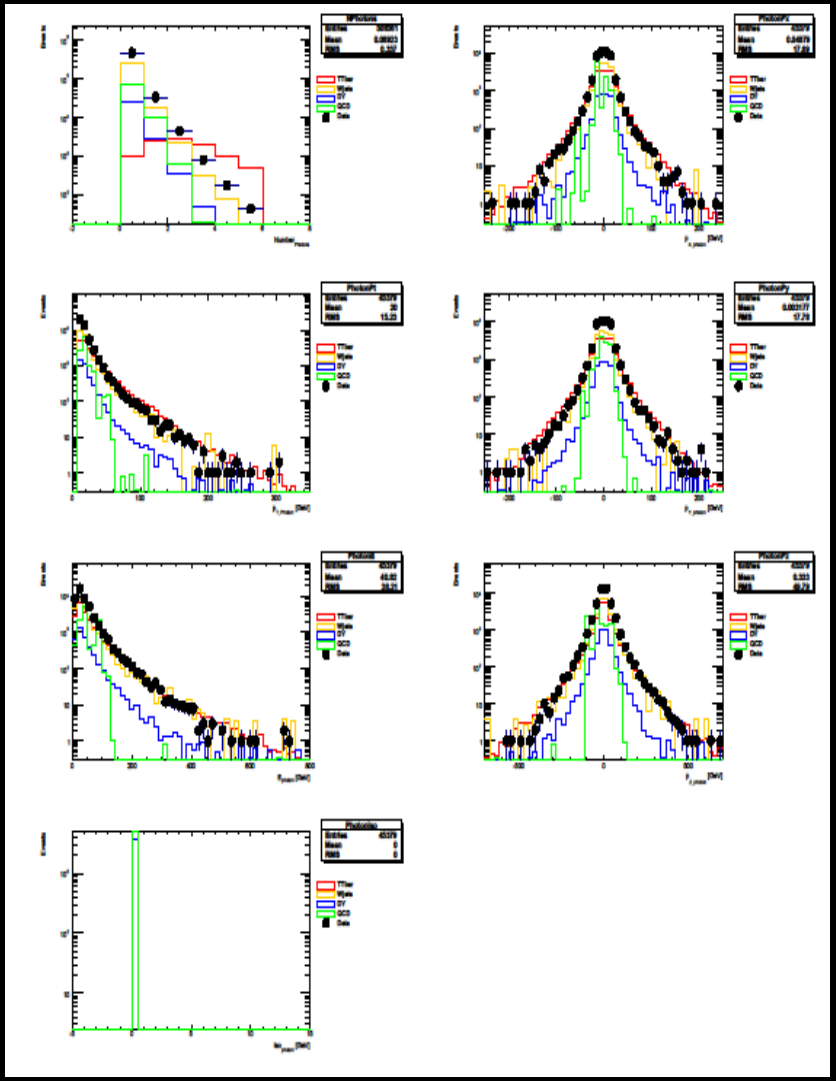


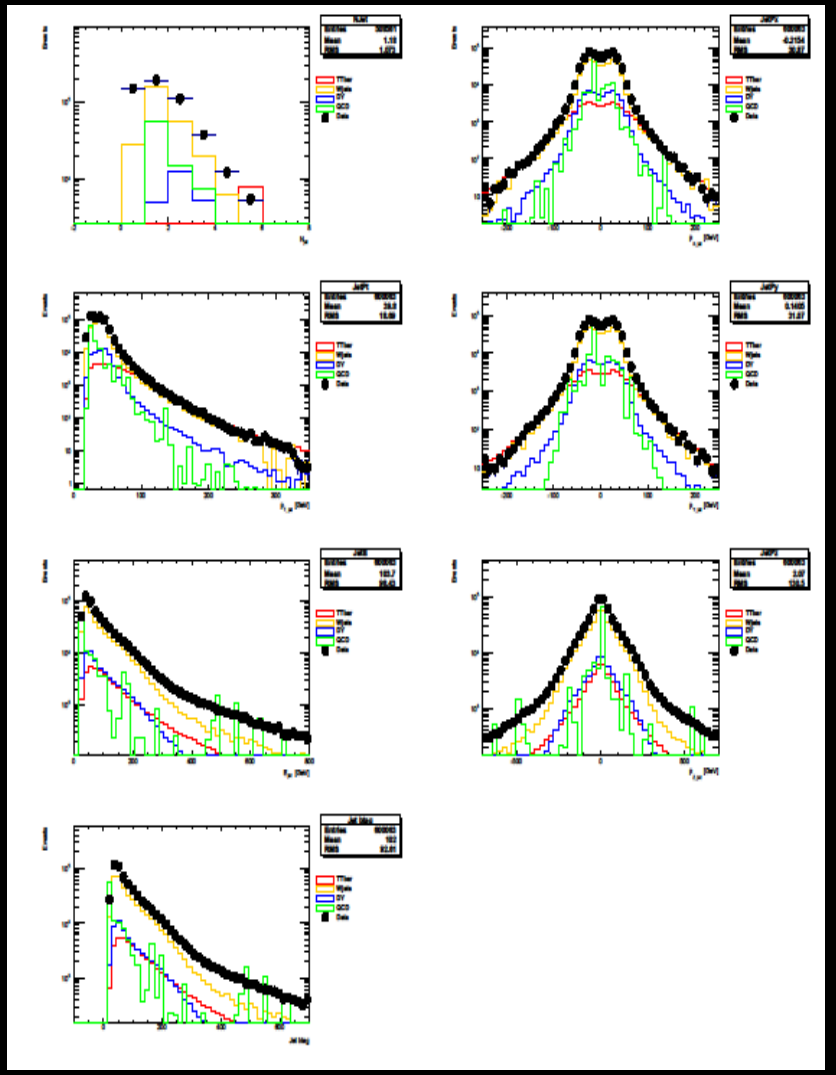
C++



Quarknet

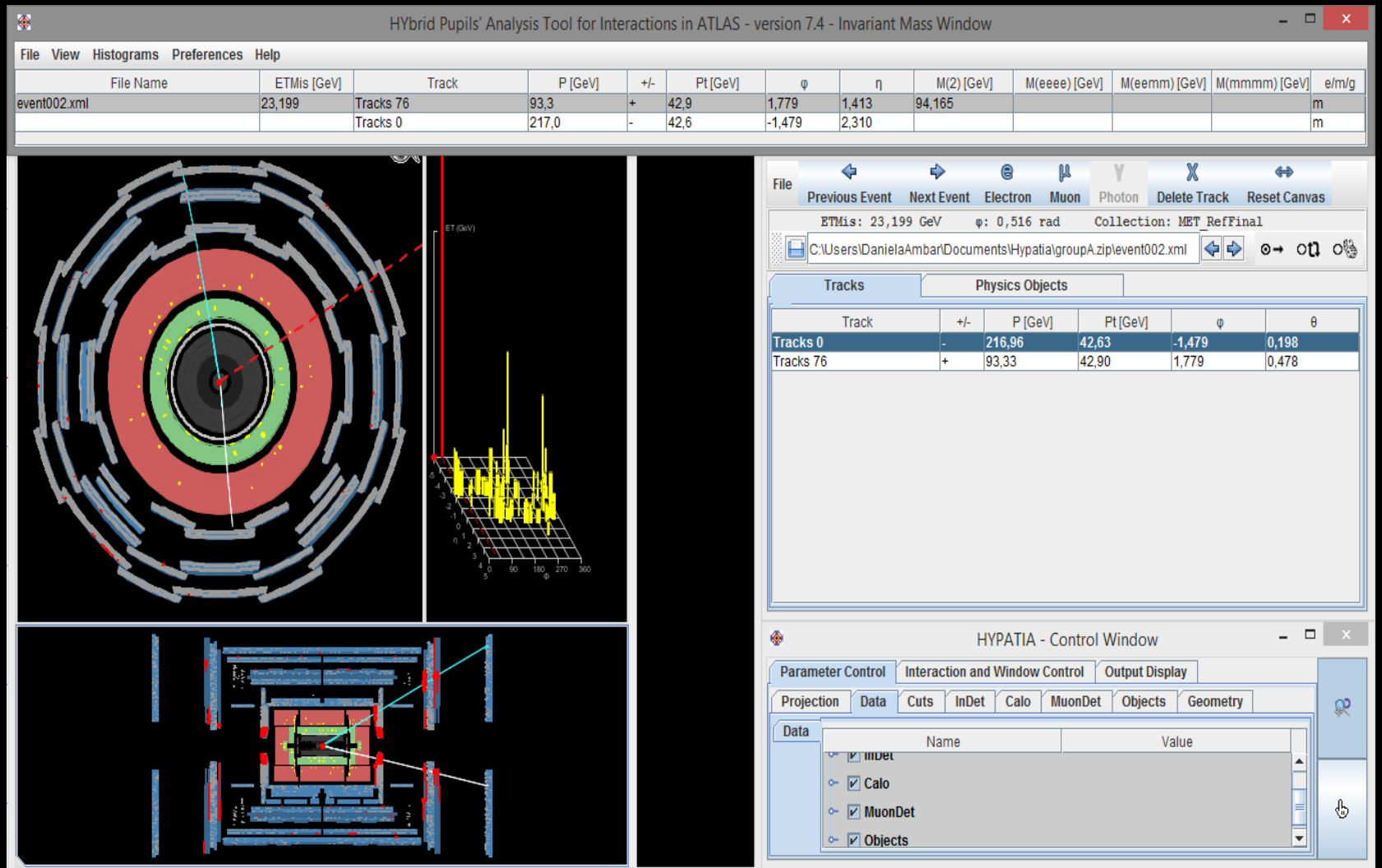
Photons





Quarknet

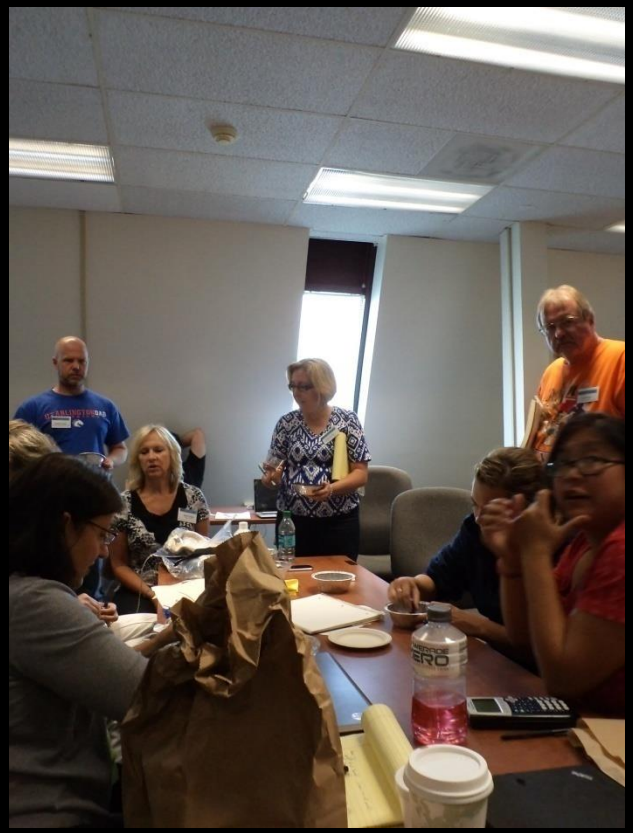
Masterclass Hypatia



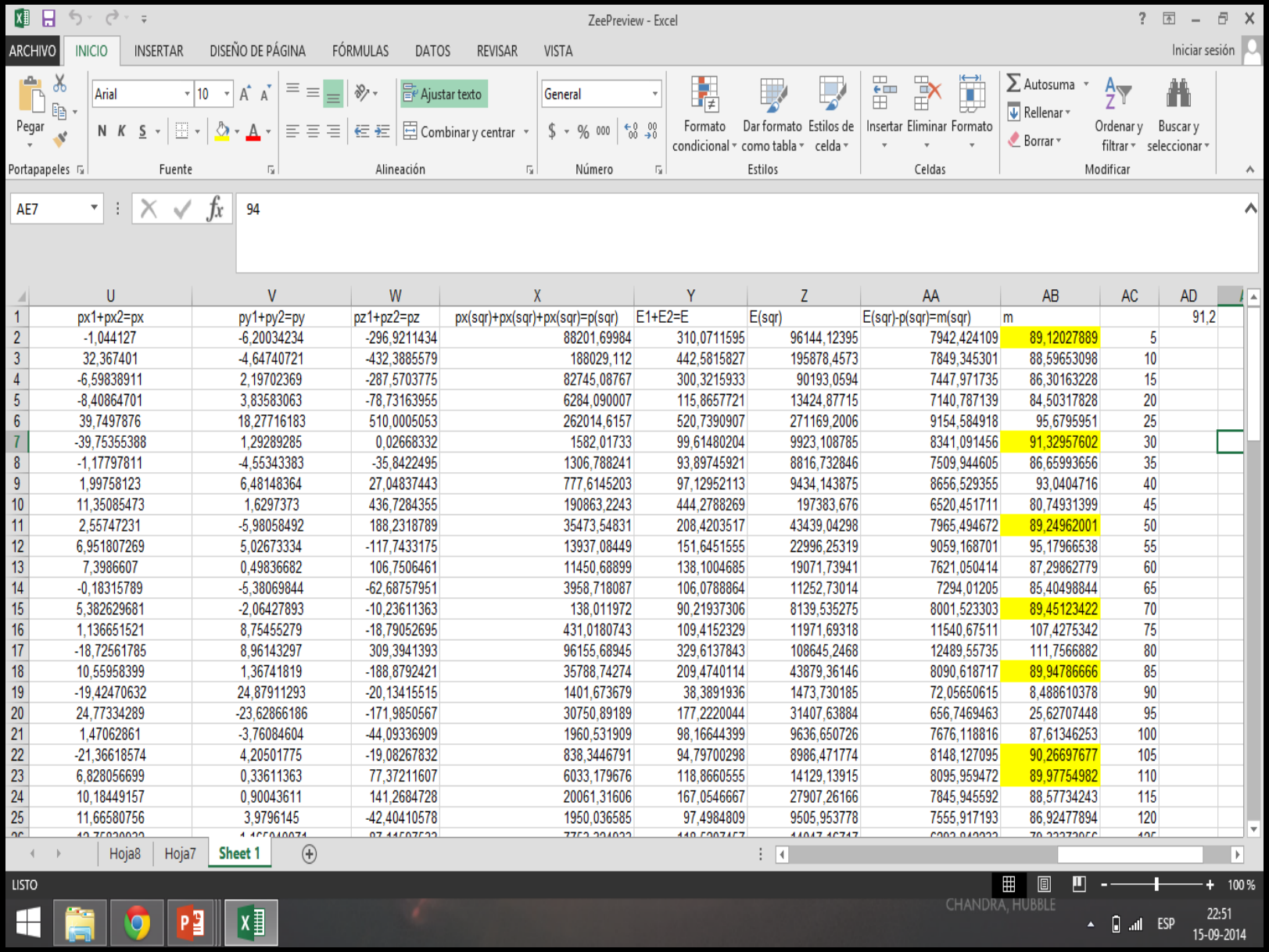
Fermilab

FERMILAB

Activities

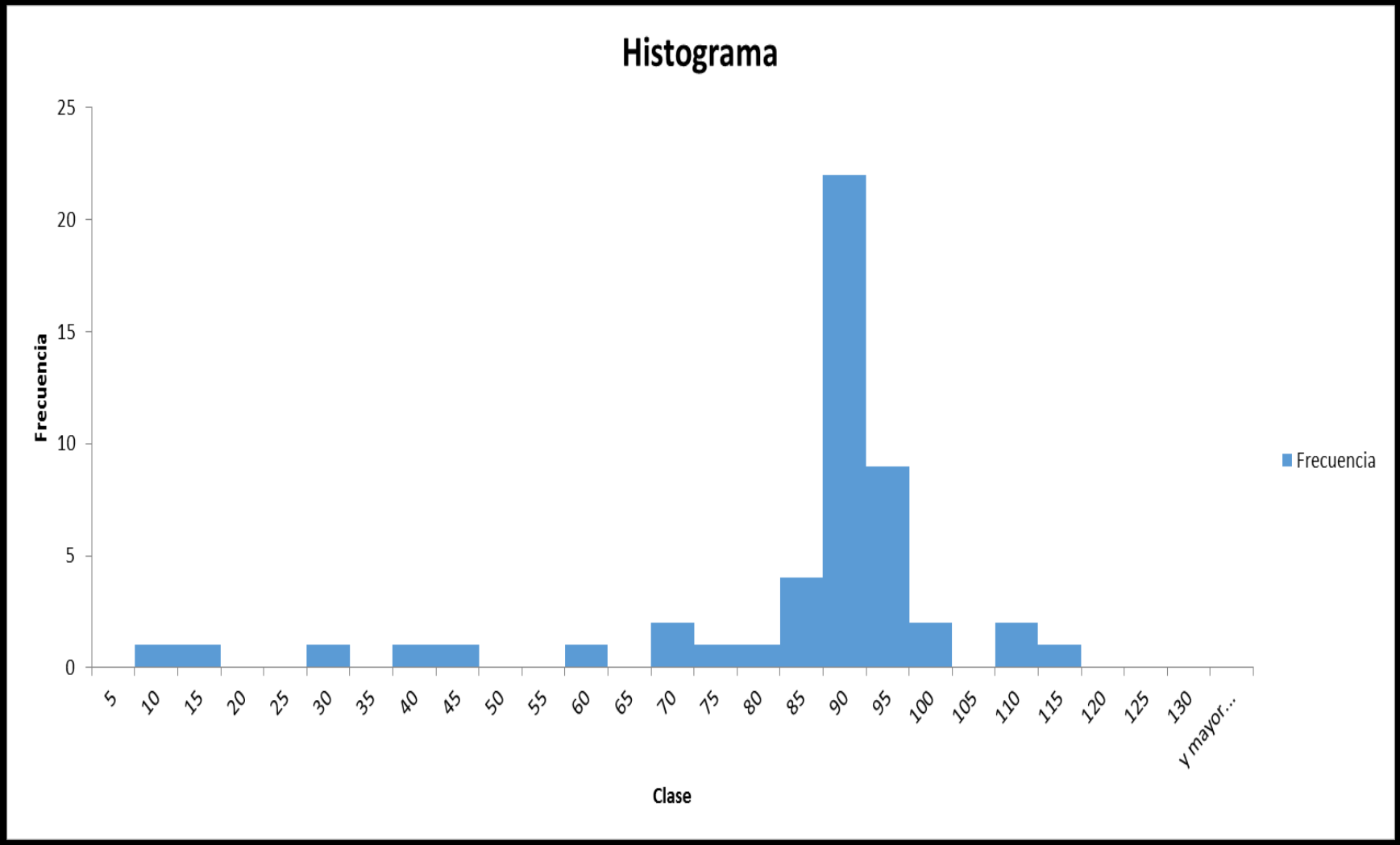


FERMILAB



	U	V	W	X	Y	Z	AA	AB	AC	AD
1	px1+px2=px	py1+py2=py	pz1+pz2=pz	px(sqr)+px(sqr)+px(sqr)=p(sqr)	E1+E2=E	E(sqr)	E(sqr)-p(sqr)=m(sqr)	m		91,2
2	-1,044127	-6,20034234	-296,9211434	88201,69984	310,0711595	96144,12395	7942,424109	89,12027889	5	
3	32,367401	-4,64740721	-432,3885579	188029,112	442,5815827	195878,4573	7849,345301	88,59653098	10	
4	-6,59838911	2,19702369	-287,5703775	82745,08767	300,3215933	90193,0594	7447,971735	86,30163228	15	
5	-8,40864701	3,83583063	-78,73163955	6284,090007	115,8657721	13424,87715	7140,787139	84,50317828	20	
6	39,7497876	18,27716183	510,0005053	262014,6157	520,7390907	271169,2006	9154,584918	95,6795951	25	
7	-39,75355388	1,29289285	0,02668332	1582,01733	99,61480204	9923,108785	8341,091456	91,32957602	30	
8	-1,17797811	-4,55343383	-35,8422495	1306,788241	93,89745921	8816,732846	7509,944605	86,65993656	35	
9	1,99758123	6,48148364	27,04837443	777,6145203	97,12952113	9434,143875	8656,529355	93,0404716	40	
10	11,35085473	1,6297373	436,7284355	190863,2243	444,2788269	197383,676	6520,451711	80,74931399	45	
11	2,55747231	-5,98058492	188,2318789	35473,54831	208,4203517	43439,04298	7965,494672	89,24962001	50	
12	6,951807269	5,02673334	-117,7433175	13937,08449	151,6451555	22996,25319	9059,168701	95,17966538	55	
13	7,3986607	0,49836682	106,7506461	11450,68899	138,1004685	19071,73941	7621,050414	87,29862779	60	
14	-0,18315789	-5,38069844	-62,68757951	3958,718087	106,0788864	11252,73014	7294,01205	85,40498844	65	
15	5,382629681	-2,06427893	-10,23611363	138,011972	90,21937306	8139,535275	8001,523303	89,45123422	70	
16	1,136651521	8,75455279	-18,79052695	431,0180743	109,4152329	11971,69318	11540,67511	107,4275342	75	
17	-18,72561785	8,96143297	309,3941393	96155,68945	329,6137843	108645,2468	12489,55735	111,7566882	80	
18	10,55958399	1,36741819	-188,8792421	35788,74274	209,4740114	43879,36146	8090,618717	89,94786666	85	
19	-19,42470632	24,87911293	-20,13415515	1401,673679	38,3891936	1473,730185	72,05650615	8,488610378	90	
20	24,77334289	-23,62866186	-171,9850567	30750,89189	177,2220044	31407,63884	656,7469463	25,62707448	95	
21	1,47062861	-3,76084604	-44,09336909	1960,531909	98,16644399	9636,650726	7676,118816	87,61346253	100	
22	-21,36618574	4,20501775	-19,08267832	838,3446791	94,79700298	8986,471774	8148,127095	90,26697677	105	
23	6,828056699	0,33611363	77,37211607	6033,179676	118,8660555	14129,13915	8095,959472	89,97754982	110	
24	10,18449157	0,90043611	141,2684728	20061,31606	167,0546667	27907,26166	7845,945592	88,57734243	115	
25	11,66580756	3,9796145	-42,40410578	1950,036585	97,4984809	9505,953778	7555,917193	86,92477894	120	

FERMILAB


Histograma
Frecuencia
25
20
15
10
5
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100
105
110
115
120
125
130
y mayor...
Clase
Frecuencia

How about in Chile?

- Classes to learn about Hadronic structure and Collider and particle physics.
- Classes using software hypatia.
- Connect relativity with particle physics through calculations of Z mass.
- Make videos about Quark Workbench.
- Watch and discussed movies about relativity and particle physics, for example Interstellar.

Plans for Next Summer

- Continue $t\bar{t}$ analysis
  - Develop a set of cuts
  - Determine efficiencies
  - Create $t\bar{t}$ mass plots
  - Calculate $t\bar{t}$ cross-section and mass

Conclusions

- We've made substantial progress toward analyzing CMS data in its native format
- We are in a good position to finish the $t\bar{t}$ analysis by the end of next summer
- We are better poised to access and analyze CMS data in general

- In October 2014 LHC Collaborations released portions of their datasets for education and research purposes at: http://opendata.cern.ch/