

These are instructions for creating an invariant mass histogram using EXCEL and CERN's open data. The data is from dimuon decay and will peak at around 90 GeV for the Z boson. This is actual data from the Compact Muon Solenoid (CMS) which has been acquired, analyzed, filtered and identified as collisions in the Large Hadron Collider (LHC) as those that exhibit two muons.

### Instructions for acquiring the open data:

1. Download the comma separated(.csv) file at: <http://opendata.cern.ch/record/700>. Or go to the opendata homepage and
  - ➔ click education (start learning)
  - ➔ click explore CMS
  - ➔ type dimuon event information into the search field
  - ➔ click the first search result
2. It is the largest file, MuRun2010B.csv, around 15 MB in size.

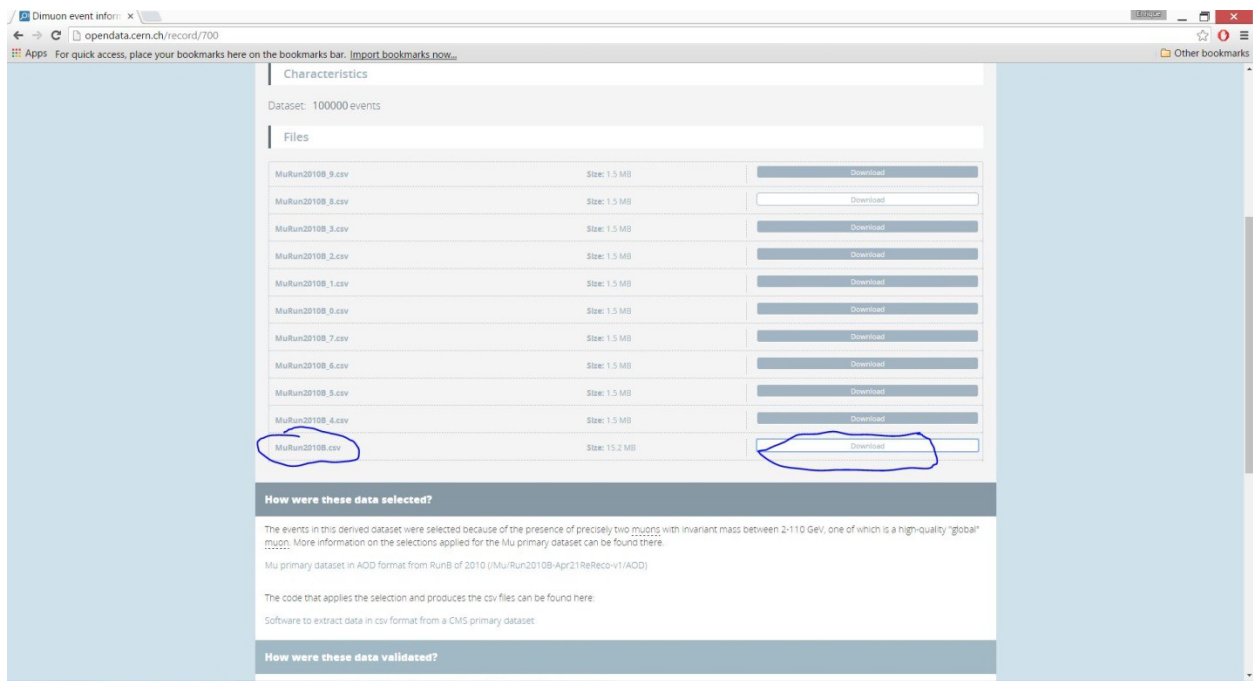


Figure 1: download MuRun2010B.csv (This is data from the first LHC run, taken in 2010)

Note: you can have students download the other files too. The largest dataset contains 100,000 collisions. The more data you have in the histogram, the more statistically accurate it will be.

### The physics involved:

It is known that Z bosons can transform(decay) into two muons. When analyzing the dimuons invariant masses from the data, it is possible to find out the masses of the particles which they are transformed(decayed) from.

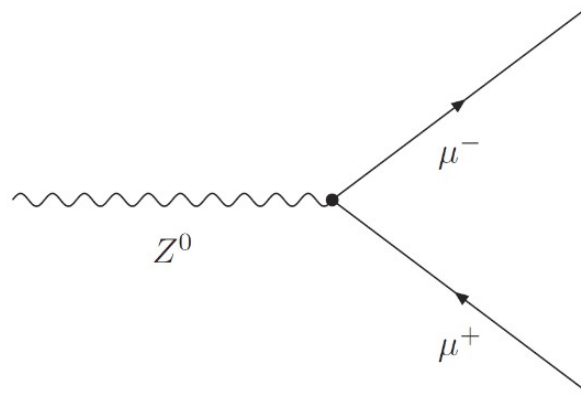
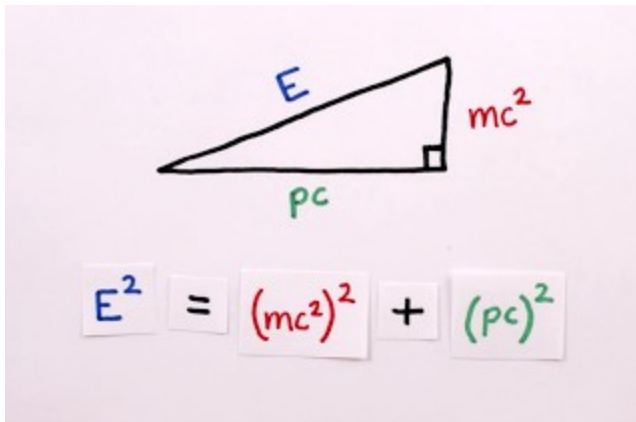


Figure 2: Feynman Diagram of Z-boson decay.



We can calculate the invariant mass using a right triangle analogy between mass, energy and momentum. This equation simplifies if we set the speed of light,  $c=1$ .

We get

Solving for mass  $m$ ,

When we have 2 muons, in order to get the invariant mass of the particle which produced the muons, we have to add the energies and momenta, so our equation becomes:

Momentum is a vector quantity and must be summed in accordance with vector rules. To get a resultant vector, we add the components first, and then get the magnitude of the resultant vector by using the Pythagorean theorem in 3-dimensions.

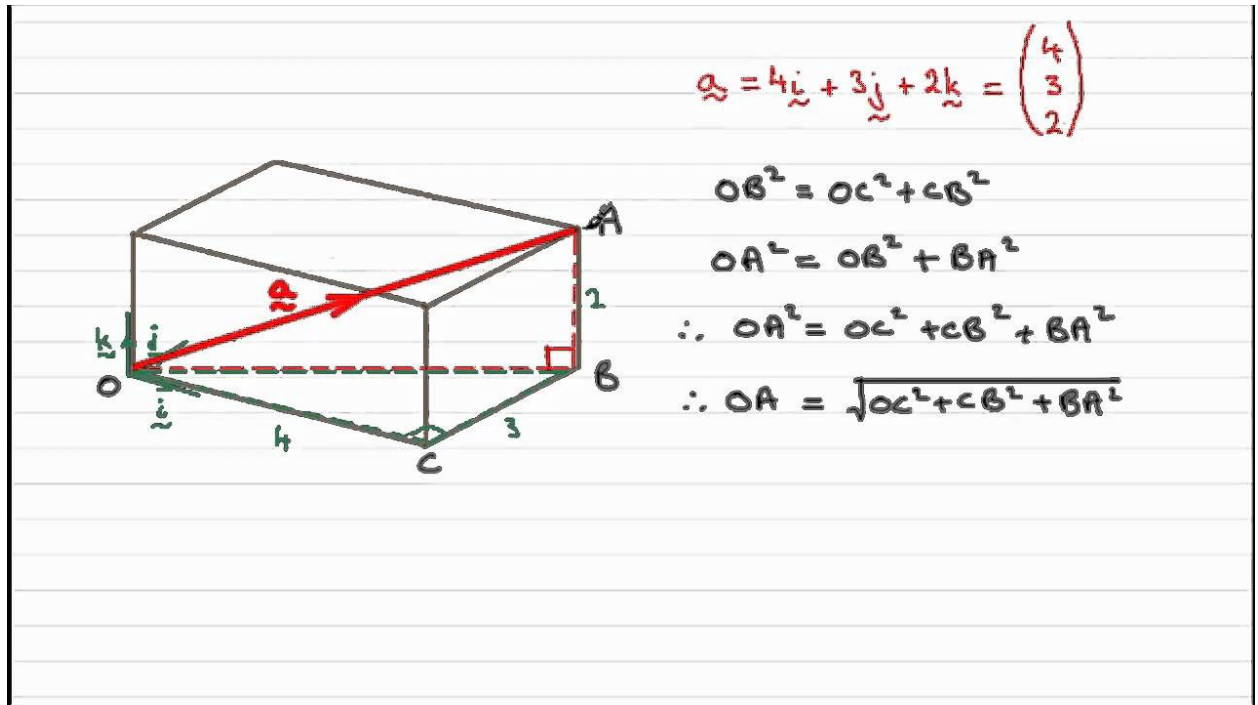


Figure 3: The magnitude of a 3d vector is just the sum of the squares, rooted.

Using open data:

1. Open the downloaded file with EXCEL or other spreadsheet program

B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	
1	Event	Type1	E1	px1	py1	pz1	pt1	eta1	phi1	Q1	Type2	E2	px2	py2	pz2	pt2	eta2
2	90830792	G	19.1712	3.81713	9.04323	-16.4673	9.81583	-1.28942	1.17139	1T	5.43984	-0.362592	2.62699	-4.74849	2.65189	-1.34587	
3	90862225	G	12.9435	5.12579	-3.98369	-11.1973	6.4918	-1.31335	-0.660674	-1G	11.8636	4.78984	-6.26222	-8.86434	7.88403	-0.966622	
4	90644850	G	12.3999	-0.849742	9.4011	8.04015	9.43943	0.77258	1.66094	1G	8.55532	-4.85155	6.97696	-0.983229	8.49797	-0.115445	
5	90678594	G	17.8132	-1.95959	2.80531	17.4811	3.42195	2.3335	2.18053	1G	9.42174	4.36523	0.168017	8.34713	4.36846	1.403	
6	90686690	G	7.95664	7.097	-1.31646	3.34613	7.21806	0.448399	-0.183411	1T	5.44467	-1.34176	1.38647	5.09025	1.92941	1.69738	
7	90692451	G	11.3415	0.816445	4.98596	10.1534	5.05236	1.44794	1.40849	1G	7.86621	0.602823	5.05381	5.99681	5.08964	1.00197	
8	90725500	G	18.1293	9.77963	7.17648	13.4728	12.1303	0.957512	0.633065	-1T	4.48788	-1.39991	0.374027	4.24621	1.44902	1.79621	
9	90749620	G	15.8762	1.5928	3.01336	15.5057	3.40842	2.21997	1.08455	-1G	7.74866	-1.91338	2.42923	7.10411	3.09228	1.56923	
10	90808356	G	8.99267	-1.6444	-4.20646	7.77551	4.51646	1.31172	-1.94345	1G	6.45454	0.860784	-4.41161	4.63105	4.4948	0.902647	
11	90851884	G	10.4195	-3.61903	-2.36646	-9.47933	4.32406	-1.52644	-2.56249	1G	25.6473	11.7951	-17.8219	14.1785	21.3716	0.622447	
12	90653677	G	10.679	2.05438	0.499999	10.4671	2.11435	2.30268	0.23874	-1G	9.74014	-6.3448	-7.15544	1.84454	9.56331	0.1917	
13	90733245	G	5.94088	1.89526	0.200522	5.6259	1.90584	1.80313	0.105409	1T	5.51003	-4.81074	-2.4305	1.13959	5.38986	0.209888	
14	90779622	G	18.3797	2.24914	3.19848	17.9586	3.9101	2.2293	0.957933	-1G	8.11221	-3.1851	2.77883	6.92316	4.22691	1.26889	
15	90788478	G	13.2996	1.33269	13.0342	-2.28086	13.1021	-0.173215	1.4689	1T	6.65812	0.323117	1.82718	-6.39346	1.85553	-1.95067	
16	90916664	G	10.6149	5.48515	8.64685	-2.79476	10.2399	-0.26965	1.0055	1G	24.3322	-14.8133	-9.85434	-16.5982	17.7916	-0.83314	
17	90969200	G	6.21065	0.402732	-3.07414	-5.38038	3.10041	-1.31862	-1.44053	-1G	12.0373	-8.90423	-7.12958	3.84303	11.4068	0.330837	
18	91067376	G	14.0765	1.64084	3.77976	13.4594	4.12055	1.89949	1.16123	1G	7.65029	-6.28799	-4.12087	1.41259	7.51801	0.186805	
19	91027753	G	9.2808	-1.94966	-7.25532	5.44805	7.51271	0.673171	-1.83332	1T	4.6726	1.1791	1.28327	4.33416	1.74271	1.6424	
20	90916331	G	8.4722	1.97601	7.82587	2.57256	8.07149	0.313559	1.32347	-1G	4.12802	-3.06666	-2.55515	-1.04698	3.99164	-0.259375	
21	91079051	G	11.7437	-10.8203	2.95501	3.47723	11.2166	0.305245	2.875	-1T	3.67172	0.830242	-0.966505	-3.44194	1.27414	-1.71953	
22	90942036	G	41.4068	4.09647	10.9996	39.7081	11.7376	1.93306	1.21429	-1T	2.88176	0.936893	0.121513	7.22045	0.94474	1.77966	
23	90985100	G	10.618	-4.3387	-0.754251	-9.66113	4.40377	-1.52711	-2.96947	1G	8.10394	2.39198	3.40745	-6.95201	1.16352	-1.28544	
24	91168213	G	167.782	39.0091	-30.5509	160.299	49.5486	1.89031	-0.664396	1T	3.14872	-0.426839	-1.08415	2.92331	1.16515	1.65055	
25	90928390	G	14.9175	-3.07583	2.70282	14.3442	4.09462	1.96659	2.42065	-1G	11.8474	5.45981	-0.919517	10.4736	5.5367	1.39411	
26	90951446	G	22.482	21.6862	5.29054	-2.67353	22.3222	-0.119485	0.239285	-1T	5.10747	-1.46378	-0.116118	-4.8907	1.46838	-1.91813	
27	90990646	G	9.40165	-1.53812	-1.01688	-9.21846	1.84388	-2.31234	-2.55744	1G	33.0052	5.01759	-9.95044	-31.0668	11.1439	-1.74911	
28	91129167	G	45.8122	5.66643	-14.2482	-43.1698	15.3336	-1.75839	-1.19228	-1T	8.07748	0.120785	-1.55773	-7.92423	1.56241	-2.32643	
29	91255034	G	13.0397	0.876393	5.15305	-11.9457	5.22705	-1.56443	1.40234	-1G	13.6012	-2.63238	7.81677	-10.8144	8.24811	-1.08522	

Figure 4: .csv file opened in LibreOffice

## Glossary for the columns:

- Run: The data collected by CMS in a given year are divided into sets called "Runs". For example, the data from 2010 were divided into "RunA" and "RunB", the latter being from the second part of the year. Run is a data collection period with a variable length of time
  - Event: One proton - proton collision
  - Type1: G = global muon, T = tracker muon (observed only in tracker)
  - E1: energy of muon number 1
  - px1: momentum in the direction of the x-axis of muon number 1
  - py1: momentum in the direction of the y-axis of muon number 1
  - pz1: momentum in the direction of the z-axis of muon number 1
  - pt1: transverse momentum of muon number 1
  - eta1 = in experimental particle physics, pseudorapidity  $\eta$ , is a commonly used spatial coordinate describing the angle of a particle relative to the beam axis. It is defined as
  
  - phi1 = is the azimuthal angle,  $\varphi$ , and is measured from the positive x-axis in the x-y-plane.
  - Q1: charge of muon number 1
  - Type2: G = global muon, T = tracker muon (observed only in tracker)
  - E2: energy of muon number 2
  - px2: momentum in the direction of the x-axis of muon number 2
  - py2: momentum in the direction of the y-axis of muon number 2
  - pz2: momentum in the direction of the z-axis of muon number 2
  - pt2: transverse momentum of muon number 2
  - eta2 = in experimental particle physics, pseudorapidity  $\eta$ , is a commonly used spatial coordinate describing the angle of a particle relative to the beam axis. It is defined as
  
  - phi2 = is the azimuthal angle,  $\varphi$ , and is measured from the positive x-axis in the x-y-plane.
  - Q2: charge of muon number 2
  - M: Invariant mass
2. To calculate the invariant mass, we need to first calculate the total momenta of both muons:
- a. First we calculate the sum of momenta of both muons in x, y & z-directions.
  - b. To do this, create a new column (named for example vector sum px) and type  $=(E2+N2)$ 
    - b.i. Then copy the equation for each of the cells in the column (to do this select cell V2, scroll all the way down to cell V100001, press down shift and select cell V100001. Then press ctrl + d.

V2:V100001  $\sum$  = =(E2+N2)

	M	N	O	P	Q	R	S	T	U	V
99974	8,26491	-3,03954	0,138833	7,68371	3,04271	1,65658	3,09595	1	6,51209	
99975	4,11088	1,12754	-0,427857	3,92859	1,20599	1,89689	-0,362676	-1	2,44222	
99976	8,71322	6,16271	-4,06202	4,62927	7,38099	0,591998	-0,582766	1	10,2813	
99977	6,99733	5,92772	0,216154	-3,71039	5,93166	-0,590587	0,0364487	1	8,84741	
99978	4,36982	0,294436	1,46183	-4,10615	1,49118	-1,73751	1,37204	-1	18,069	
99979	5,85586	-0,473105	1,09943	-5,73127	1,1969	-2,27007	1,97716	1	2,61924	
99980	8,23405	-1,07099	4,83625	6,57663	4,95342	1,09522	1,78873	-1	6,53904	
99981	10,2788	-0,537263	-4,06371	9,42551	4,09907	1,57005	-1,70224	-1	7,74001	
99982	3,91133	1,99077	2,86115	1,77144	3,48559	0,48855	0,962899	1	6,93768	
99983	3,63427	-2,75676	-1,8428	-1,48361	3,31597	-0,433691	-2,55235	1	7,49873	
99984	5,33429	5,10356	1,54172	-0,142348	5,33134	-0,0266971	0,29337	1	12,9462	
99985	4,54869	-4,47685	0,207371	0,770823	4,48165	0,171158	3,0953	-1	11,876	
99986	9,67424	-1,21325	8,83232	3,75473	8,91526	0,409608	1,70731	-1	12,8433	
99987	23,8796	6,8046	-4,02762	22,5322	7,90723	1,76977	-0,534439	1	2,33736	
99988	7,31658	-3,82168	-1,47399	-6,06162	4,09608	-1,18354	-2,77348	-1	2,93104	
99989	12,2514	11,2216	-3,29678	-3,64575	11,6959	-0,306873	-0,285748	-1	14,2309	
99990	6,20147	-5,38144	3,07747	-0,127977	6,19925	-0,0206425	2,62212	1	11,2614	
99991	5,20524	-5,15548	-0,516811	-0,487193	5,18131	-0,0938909	-3,04168	1	7,81819	
99992	3,68546	0,811564	0,65352	3,53352	1,04198	1,93538	0,67794	-1	2,69561	
99993	2,8846	1,12062	-0,340286	-2,63405	1,17114	-1,5498	-0,294811	-1	2,18911	
99994	5,04227	-4,61148	-1,90852	0,710773	4,99081	0,141939	-2,74919	1	12,7389	
99995	11,5498	0,68567	3,16948	11,0847	3,2428	1,94302	1,35774	1	5,21763	
99996	12,073	4,86687	-9,3491	-5,88675	10,54	-0,532927	-1,09083	-1	13,1144	
99997	5,48304	4,49503	3,04466	0,759731	5,42912	0,139484	0,595355	-1	6,61359	
99998	4,2913	-2,87423	-3,08156	0,804277	4,21393	0,189721	-2,3214	-1	9,37972	
99999	11,3934	6,3344	5,27644	-7,86341	8,24412	-0,84834	0,694531	1	22,7881	
100000	8,27498	-6,1363	1,89389	-5,21758	6,42191	-0,742373	2,84223	-1	12,8546	
100001	6,31165	0,629169	-1,9335	5,97423	2,0333	1,79872	-1,2562	1	2,7981	

Figure 5: Before copying the function to column V

b.ii. Repeat for y and z axes

$\sum$  = =(G2+P2)

	O	P	Q	R	S	T	U	V	W	X	Y
	py2	pz2	pt2	eta2	phi2	Q2	M	Vector sum px	Vector sum py	Vector sum pz	
2	2,62699	-4,74849	2,65189	-1,34587	1,70796	1	2,73205	3,454538	11,67022	-21,21579	
4	-6,26222	-8,86434	7,88403	-0,966622	-0,917841	1	3,10256	9,91563	-10,24591	-20,06164	
5	6,97696	-0,983229	8,49797	-0,115445	2,17841	-1	9,41149	-5,701292	16,37806	7,056921	
3	0,168017	8,34713	4,36846	1,403	0,0384708	1	7,74702	2,40564	2,973327	25,82823	
6	1,38647	5,09025	1,92941	1,69738	2,33981	-1	8,67727	5,75524	0,07001	8,43638	
3	5,05381	5,99681	5,08964	1,00197	1,45208	-1	2,30104	1,419268	10,03977	16,15021	
1	0,374027	4,24621	1,44902	1,79621	2,88051	1	8,38705	8,37972	7,550507	17,71901	
8	2,42923	7,10411	3,09228	1,56923	2,23796	1	4,14828	-0,32058	5,44259	22,60981	
4	-4,41161	4,63105	4,4948	0,902647	-1,3781	-1	3,1318	-0,783616	-8,61807	12,40656	
1	-17,8219	14,1785	21,3716	0,622447	-0,986147	-1	28,3605	8,17607	-20,18836	4,69917	
8	-7,15544	1,84454	9,56331	0,1917	-2,29622	1	14,236	-4,29042	-6,655441	12,31164	
4	-2,4305	1,13959	5,38986	0,209888	-2,67378	1	8,47813	-2,91548	-2,229978	6,76549	
1	2,77883	6,92316	4,22691	1,26889	2,42421	-1	6,79063	-0,93596	5,97731	24,88176	
7	1,82718	-6,39346	1,85553	-1,95067	1,39577	1	9,97321	1,655807	14,86138	-8,67432	
3	-9,85434	-16,5982	17,7916	-0,83314	-2,55458	1	27,5089	-9,32815	-1,20749	-19,39296	
3	-7,12958	3,84303	11,4068	0,330837	-2,46643	1	12,419	-8,501498	-10,20372	-1,53735	
9	-4,12087	1,41259	7,51801	0,186805	-2,56146	-1	15,1381	-4,64715	-0,34111	14,87199	
1	1,28327	4,33416	1,74271	1,6424	0,827677	1	7,92127	-0,77056	-5,97205	9,78221	
6	-2,55515	-1,04698	3,99164	-0,259375	-2,44693	1	11,2902	-1,09065	5,27072	1,52558	
2	-0,966505	-3,44194	1,27414	-1,71953	-0,861092	-1	11,5705	-9,990058	1,988505	0,03529	
3	0,121513	2,72045	0,94474	1,77966	0,128978	-1	3,50342	5,033363	11,121113	42,42855	
8	3,40745	-6,95201	4,1632	-1,28544	0,95874	1	7,98031	-1,94672	2,653199	-16,61314	
9	-1,08415	2,92331	1,16515	1,65055	-1,94587	-1	9,29907	38,582261	-31,63505	163,22231	
1	-0,919517	10,4736	5,5367	1,39411	-0,16685	1	9,56971	2,38398	1,783303	24,8178	
8	-0,116118	-4,8907	1,46838	-1,91813	-3,06243	-1	16,378	20,22242	5,174422	-7,56423	
9	-9,95044	-31,0668	11,1439	-1,74911	-1,10375	-1	6,56139	3,47947	-10,96732	-40,28526	
5	-1,55773	-7,92423	1,56241	-2,32643	-1,49341	-1	3,19125	5,787215	-15,80593	-51,09403	

Figure 6: Your spreadsheet should look like this after calculating the x, y & z momenta

- c. Next we need to calculate the sum of the total momenta of both muons. Name the column resultant momentum sum. Type  $=\sqrt{V^2+W^2+X^2}$  into cell Y2.
  - c.i. Copy the equation to all the cells in the column (up until Y100001)

V	W	X	Y
Vector sum px	Vector sum py	Vector sum pz	Resultant momentum sum
3,454538	11,67022	-21,21579	24,4588963971
9,91563	-10,24591	-20,06164	24,6123503452
-5,701292	16,37806	7,056921	18,7228660689
2,40564	2,973327	25,82823	26,1098686743
5,75524	0,07001	8,43638	10,2127467589
1,419268	10,03977	16,15021	19,0693625052
8,37972	7,550507	17,71901	21,0045989873
-0,32058	5,44259	22,60981	23,2578602988
-0,783616	-8,61807	12,40656	15,1263979716
8,17607	-20,18836	4,69917	22,2822844166
-4,29042	-6,655441	12,31164	14,638274426
-2,91548	-2,229978	6,76549	7,6970566096
-0,93596	5,97731	24,88176	25,6067615417
1,655807	14,86138	-8,67432	17,2871669104
-9,32815	-1,20749	-19,39296	21,5536380243
-8,501498	-10,20372	-1,53735	13,3699220306
-4,64715	-0,34111	14,87199	15,584878752
-0,77056	-5,97205	9,78221	11,4869829111
-1,09065	5,27072	1,52558	5,5944080185
-9,990058	1,988505	0,03529	10,1861011365
5,033363	11,121113	42,42855	44,1496970833

Figure 7: Our accomplishments so far

- d. Now we can calculate the invariant masses related to each of the dimuon events. To do that name the next column Invariant mass. Type  $=\sqrt{(D^2+M^2)-Y^2}$  into cell Z2.
  - d.i. Copy the equation to all the cells in the column (up until Z100001)

U	V	W	X	Y	Z
M	Vector sum px	Vector sum py	Vector sum pz	Resultant momentum sum	Invariant mass
2,73205	3,454538	11,67022	-21,21579	24,4588963971	2,7323390924
3,10256	9,91563	-10,24591	-20,06164	24,6123503452	3,1023250789
9,41149	-5,701292	16,37806	7,056921	18,7228660689	9,4114574544
7,74702	2,40564	2,973327	25,82823	26,1098686743	7,7470455409
8,67727	5,75524	0,07001	8,43638	10,2127467589	8,677264163
2,30104	1,419268	10,03977	16,15021	19,0693625052	2,3012034007
8,38705	8,37972	7,550507	17,71901	21,0045989873	8,3871122883
4,14828	-0,32058	5,44259	22,60981	23,2578602988	4,1480048625
3,1318	-0,783616	-8,61807	12,40656	15,1263979716	3,1318335189
28,3605	8,17607	-20,18836	4,69917	22,2822844166	28,3604277721
14,236	-4,29042	-6,655441	12,31164	14,638274426	14,235947463
8,47813	-2,91548	-2,229978	6,76549	7,6970566096	8,4781282944
6,79063	-0,93596	5,97731	24,88176	25,6067615417	6,7908069324
9,97321	1,655807	14,86138	-8,67432	17,2871669104	9,9731864422
27,5089	-9,32815	-1,20749	-19,39296	21,5536380243	27,5089164877
12,419	-8,501498	-10,20372	-1,53735	13,3699220306	12,4190524638
15,1381	-4,64715	-0,34111	14,87199	15,584878752	15,1381953346
7,92127	-0,77056	-5,97205	9,78221	11,4869829111	7,9212748444
11,2902	-1,09065	5,27072	1,52558	5,5944080185	11,2901790496
11,5705	-9,990058	1,988505	0,03529	10,1861011365	11,5705884645

Figure 8: Calculated invariant masses. Comparing the results with the values in column U would suggest that we have gotten our physics right!

3. Next we want to present our data in a histogram. The phases for doing that are:
  - a. Create a new column named Bin. We need this column to use as our vertical axis. It will allow us to count the number of events in specific energy ranges (called bins). Type 0 into cell AA2 and 1 into cell AA3. Then select cells AA2 and AA3 and drag down from the highlighted point until you reach cell AA117 and release.

		10,3121303184	104
		6,3587998313	105
		3,3709870971	106
		9,58560375	107
		5,8231624685	108
		7,9693418712	109
		13,6507525208	110
		9,8923555495	111
		5,5445469762	112
		3,7791050341	113
		3,8814100833	114
		12,2937811114	115

Z	AA
Invariant mass	Bin
2,7323390924	0
3,1023250789	1
9,4114574544	
7,7470455409	
8,677264163	
2,3012034007	

Figure 9: On the left – drag from the highlighted point. On the right – column AA filled with values from 0 to 115

- b. To create the histogram, we need to acquire the number of events in each of the previously created energy ranges (bins). To do that we need to use the frequency function. Create a new column called Number of events. Use the function wizard to get the frequency of the events on each bin.

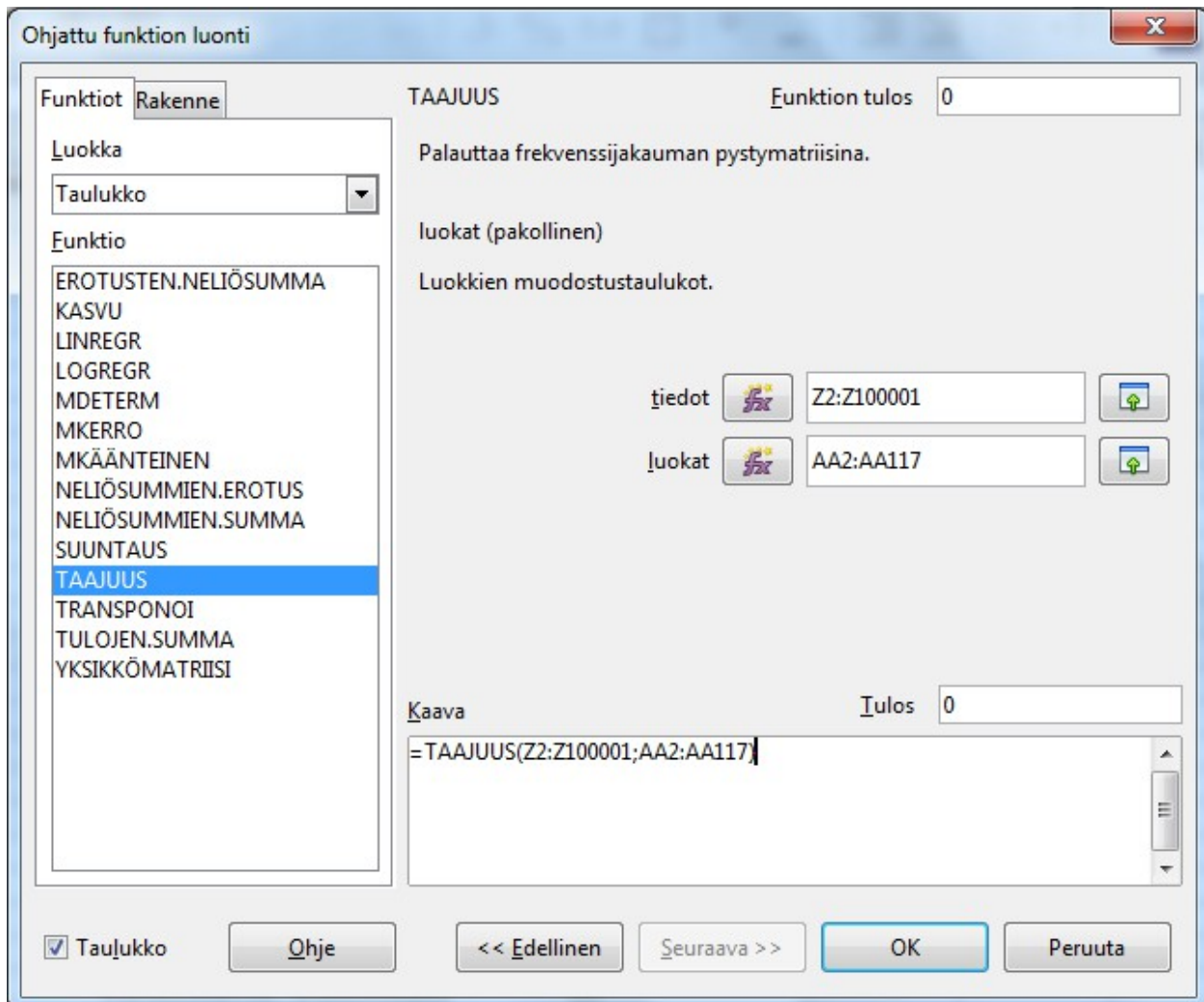


Figure 10: Finnish function wizard and frequency function

- c. The last thing left to do is the drawing of the histogram itself. Select all the cells in the Bin and Number of events column and plot a bar graph with the graphing tool
- Name the axes
  - Change the y-axis as logarithmic



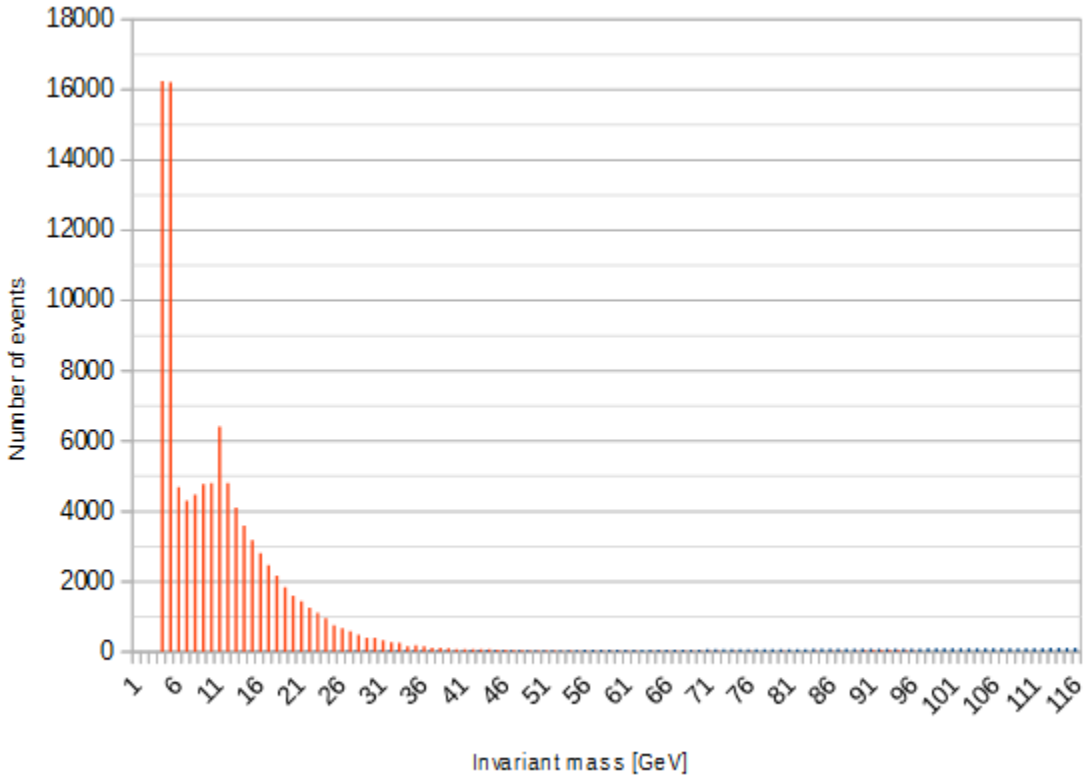


Figure 11: Amount of events in different energy ranges

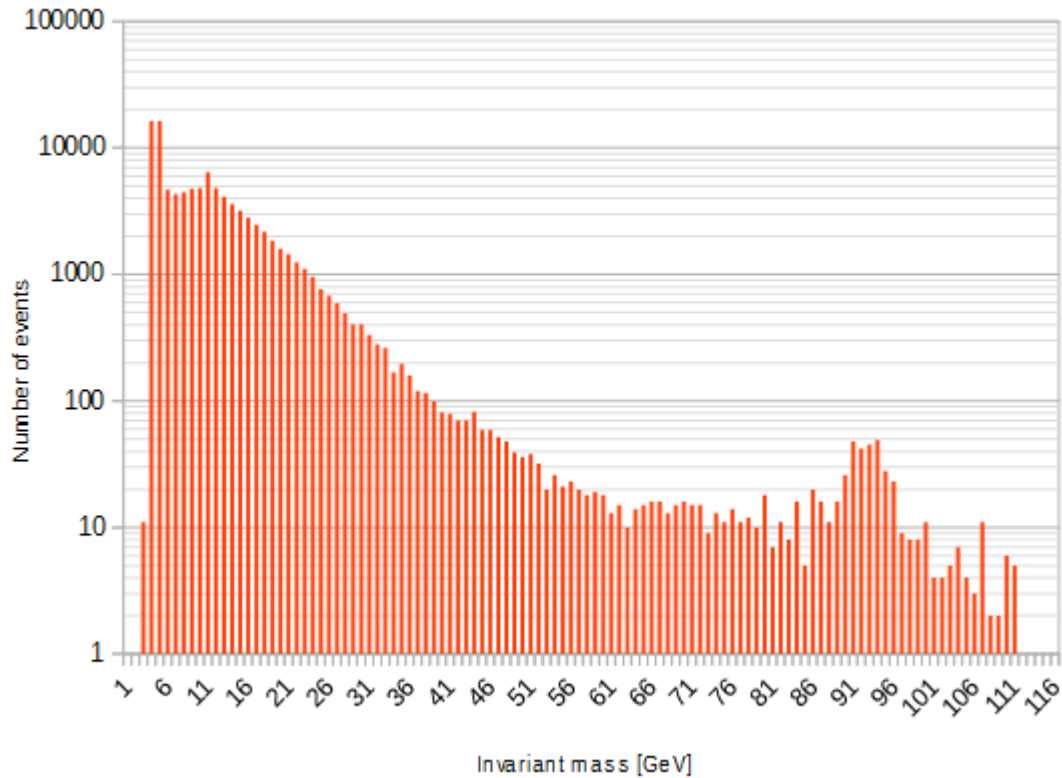


Figure 12: Amount of events in different energy ranges with a logarithmic scale

4. Analyzing the data
  - a. Why is there a bump at the 92 GeV range in the logarithmic histogram?
  - b. To what particle is the bump linked to?
  - c. Are there hints about other particles in the histogram?

## APPENDIX:

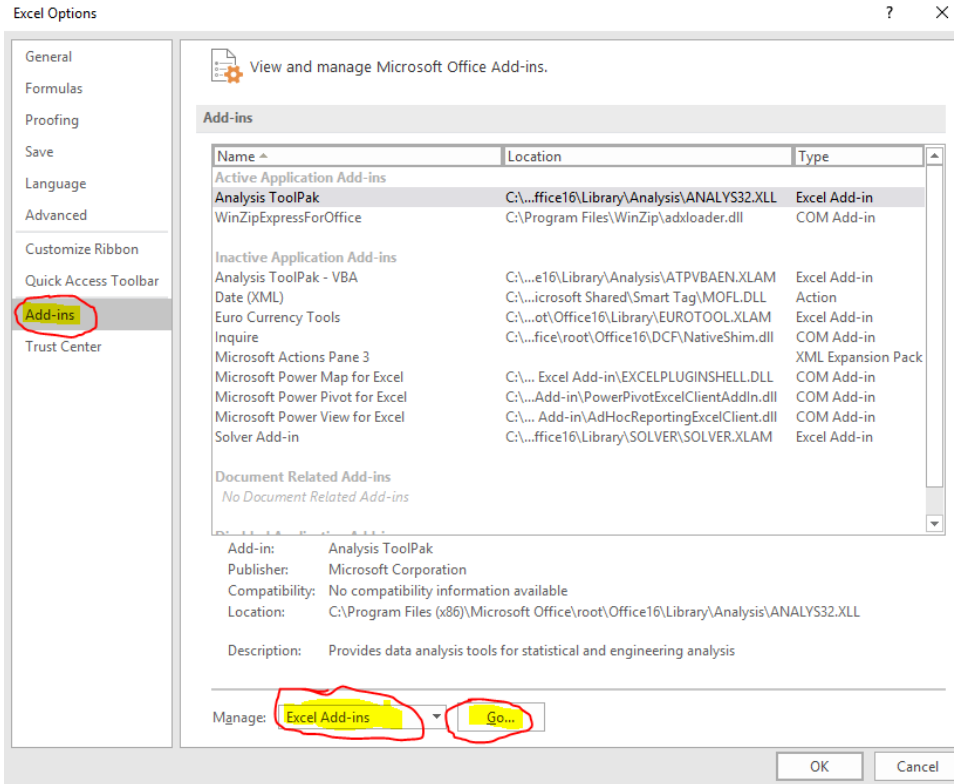
### Instructions on how to create a histogram with EXCEL:

#### 1) Active Excel Histogram package

Click "File" on the top left-hand side of the screen.

Click "Options" tab on the lower left hand side of screen.

Click "Add-ins" tab on the lower left hand side of the window. Then select "Excel Add-ins" and click "Go".

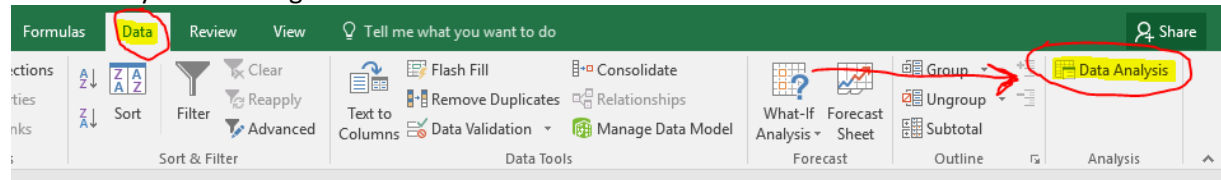


Tick the Analysis ToolPak box and click OK.

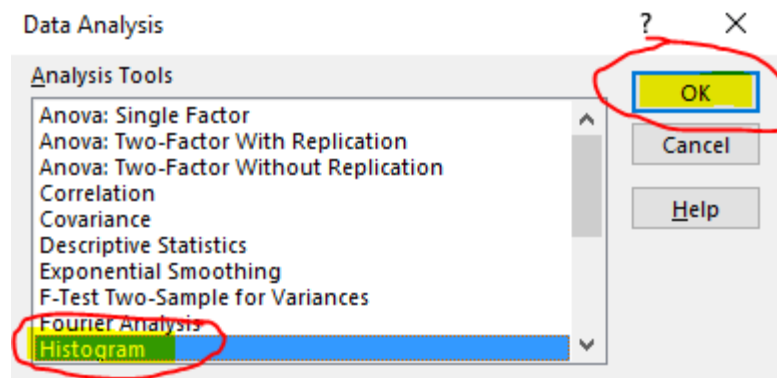
**2) Prepare Data for Histogram**

Click on the data tab.

“Data Analysis on the right hand side”.



Select “Histogram” and click “OK”.



In the “Input Range” type “\$A\$2:\$A\$2001”. In “Bin Range” type “\$B\$2:\$B\$31”. Click OK.

