



# Estudo da Seção de Choque de Pares de Quarks Top usando Dados do Run II do CMS Open Data

Matheus Costa Reis  
email: mataiascost@gmail.com

# Sumário

- Objetivo e processo físico que estamos estudando;
- Preparação do ambiente para a análise;
- Conjunto de dados utilizados;
- Corte de qualidade nos objetos;
- Corte de seleção do sinal;
- Estimativa da massa do quark top!



# Objetivo



Regular Article - Theoretical Physics | [Open Access](#) | Published: 13 September 2017

# Measurement of the $t\bar{t}$ production cross section using events with one lepton and at least one jet in pp collisions at $\sqrt{s} = 13$ TeV

[The CMS collaboration](#), [A. M. Sirunyan](#), [A. Tumasyan](#), [W. Adam](#), [E. Asilar](#), [T. Bergauer](#), [J. Brandstetter](#), [E. Brondolin](#), [M. Dragicevic](#), [J. Erö](#), [M. Flechl](#), [M. Friedl](#), [R. Frühwirth](#), [V. M. Ghete](#), [C. Hartl](#), [N. Hörmann](#), [J. Hrubec](#), [M. Jeitler](#), [A. König](#), [I. Krätschmer](#), [D. Liko](#), [T. Matsushita](#), [I. Mikulec](#), [D. Rabady](#), ... [N. Woods](#)

+ Show authors

[Journal of High Energy Physics](#) 2017, Article number: 51 (2017) | [Cite this article](#)

Reproduzir parte dos resultados da análise feita pelo CMS!

Regular Article - Theoretical Physics | [Open Access](#) | Published: 13 September 2017

# Measurement of the $t\bar{t}$ production cross section using events with one lepton and at least one jet in $p\bar{p}$ collisions at $\sqrt{s} = 13$ TeV

[The CMS collaboration](#), [A. M. Sirunyan](#), [A. Tumasyan](#), [W. Adam](#), [E. Asilar](#), [T. Bergauer](#), [J. Brandstetter](#), [E. Brondolin](#), [M. Dragicevic](#), [J. Erö](#), [M. Flechl](#), [M. Friedl](#), [R. Frühwirth](#), [V. M. Ghete](#), [C. Hartl](#), [N. Hörmann](#), [J. Hrubec](#), [M. Jeitler](#), [A. König](#), [I. Krätschmer](#), [D. Liko](#), [T. Matsushita](#), [I. Mikulec](#), [D. Rabady](#), ... [N. Woods](#)

+ Show authors

[Journal of High Energy Physics](#) 2017, Article number: 51 (2017) | [Cite this article](#)

O que estamos medindo/procurando?

Regular Article - Theoretical Physics | [Open Access](#) | Published: 13 September 2017

# Measurement of the $t\bar{t}$ production cross section using events with one lepton and at least one jet in $p\bar{p}$ collisions at $\sqrt{s} = 13$ TeV

[The CMS collaboration](#), [A. M. Sirunyan](#), [A. Tumasyan](#), [W. Adam](#), [E. Asilar](#), [T. Bergauer](#), [J. Brandstetter](#), [E. Brondolin](#), [M. Dragicevic](#), [J. Erö](#), [M. Flechl](#), [M. Friedl](#), [R. Frühwirth](#), [V. M. Ghete](#), [C. Hartl](#), [N. Hörmann](#), [J. Hrubec](#), [M. Jeitler](#), [A. König](#), [I. Krätschmer](#), [D. Liko](#), [T. Matsushita](#), [I. Mikulec](#), [D. Rabady](#), ... [N. Woods](#)

+ Show authors

[Journal of High Energy Physics](#) 2017, Article number: 51 (2017) | [Cite this article](#)

O que estamos medindo/procurando?

Regular Article - Theoretical Physics | [Open Access](#) | Published: 13 September 2017

# Measurement of the $t\bar{t}$ production cross section using events with one lepton and at least one jet in pp collisions at $\sqrt{s} = 13$ TeV

[The CMS collaboration](#), [A. M. Sirunyan](#), [A. Tumasyan](#), [W. Adam](#), [E. Asilar](#), [T. Bergauer](#), [J. Brandstetter](#), [E. Brondolin](#), [M. Dragicevic](#), [J. Erö](#), [M. Flechl](#), [M. Friedl](#), [R. Frühwirth](#), [V. M. Ghete](#), [C. Hartl](#), [N. Hörmann](#), [J. Hrubec](#), [M. Jeitler](#), [A. König](#), [I. Krätschmer](#), [D. Liko](#), [T. Matsushita](#), [I. Mikulec](#), [D. Rabady](#), ... [N. Woods](#)

+ Show authors

[Journal of High Energy Physics](#) 2017, Article number: 51 (2017) | [Cite this article](#)

Quais as condições/informações do experimento?

Regular Article - Theoretical Physics | [Open Access](#) | Published: 13 September 2017

# Measurement of the $t\bar{t}$ production cross section using events with one lepton and at least one jet in $pp$ collisions at $\sqrt{s} = 13$ TeV

The CMS collaboration, [A. M. Sirunyan](#), [A. Tumasyan](#), [W. Adam](#), [E. Asilar](#), [T. Bergauer](#), [J. Brandstetter](#), [E. Brondolin](#), [M. Dragicevic](#), [J. Erö](#), [M. Flechl](#), [M. Friedl](#), [R. Frühwirth](#), [V. M. Ghete](#), [C. Hartl](#), [N. Hörmann](#), [J. Hrubec](#), [M. Jeitler](#), [A. König](#), [I. Krätschmer](#), [D. Liko](#), [T. Matsushita](#), [I. Mikulec](#), [D. Rabady](#), ... [N. Woods](#)

+ Show authors

[Journal of High Energy Physics](#) 2017, Article number: 51 (2017) | [Cite this article](#)

Quais as condições/informações do experimento?



Regular Article - Theoretical Physics | [Open Access](#) | Published: 13 September 2017

# Measurement of the $t\bar{t}$ production cross section using events with one lepton and at least one jet in pp collisions at $\sqrt{s} = 13$ TeV

[The CMS collaboration](#), [A. M. Sirunyan](#), [A. Tumasyan](#), [W. Adam](#), [E. Asilar](#), [T. Bergauer](#), [J. Brandstetter](#), [E. Brondolin](#), [M. Dragicevic](#), [J. Erö](#), [M. Flechl](#), [M. Friedl](#), [R. Frühwirth](#), [V. M. Ghete](#), [C. Hartl](#), [N. Hörmann](#), [J. Hrubec](#), [M. Jeitler](#), [A. König](#), [I. Krätschmer](#), [D. Liko](#), [T. Matsushita](#), [I. Mikulec](#), [D. Rabady](#), ... [N. Woods](#)

+ Show authors

[Journal of High Energy Physics](#) 2017, Article number: 51 (2017) | [Cite this article](#)

Informações sobre nosso processo físico?



Regular Article - Theoretical Physics | [Open Access](#) | Published: 13 September 2017

# Measurement of the $t\bar{t}$ production cross section using events with one lepton and at least one jet in pp collisions at $\sqrt{s} = 13$ TeV

[The CMS collaboration](#), [A. M. Sirunyan](#), [A. Tumasyan](#), [W. Adam](#), [E. Asilar](#), [T. Bergauer](#), [J. Brandstetter](#), [E. Brondolin](#), [M. Dragicevic](#), [J. Erö](#), [M. Flechl](#), [M. Friedl](#), [R. Frühwirth](#), [V. M. Ghete](#), [C. Hartl](#), [N. Hörmann](#), [J. Hrubec](#), [M. Jeitler](#), [A. König](#), [I. Krätschmer](#), [D. Liko](#), [T. Matsushita](#), [I. Mikulec](#), [D. Rabady](#), ... [N. Woods](#)

+ Show authors

[Journal of High Energy Physics](#) 2017, Article number: 51 (2017) | [Cite this article](#)

Informações sobre nosso processo físico?

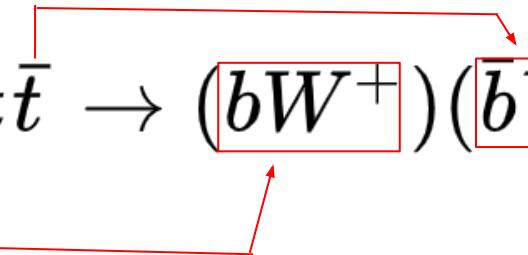
# Mais informações sobre o nosso processo

- O processo físico que queremos estudar é o

$$t\bar{t} \rightarrow (bW^+)(\bar{b}W^-) \rightarrow (b\ell^+\nu_\ell)(\bar{b}q\bar{q})$$

# Mais informações sobre o nosso processo

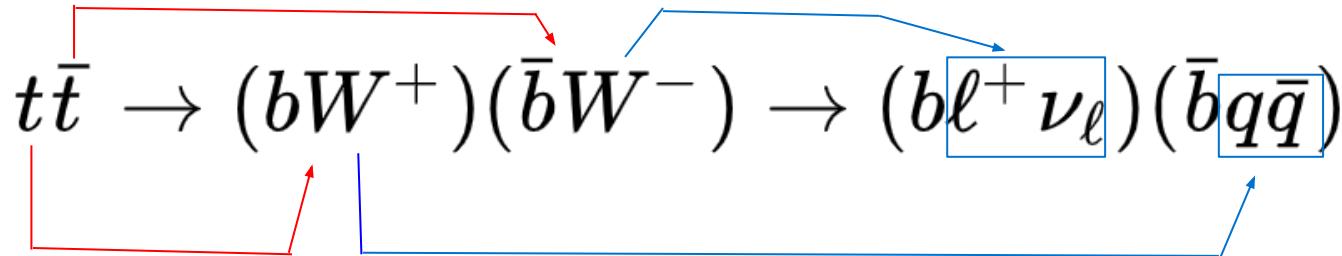
- O processo físico que queremos estudar é o

$$t\bar{t} \rightarrow (bW^+) (\bar{b}W^-) \rightarrow (b\ell^+ \nu_\ell) (\bar{b}q\bar{q})$$


The diagram shows the decay chain of a top-antitop quark pair. A red bracket groups the first two terms,  $(bW^+)$  and  $(\bar{b}W^-)$ . Another red bracket groups the last two terms,  $(b\ell^+ \nu_\ell)$  and  $(\bar{b}q\bar{q})$ . Red arrows point from the right side of the first bracket to the left side of the second bracket, highlighting the  $W$  bosons.

# Mais informações sobre o nosso processo

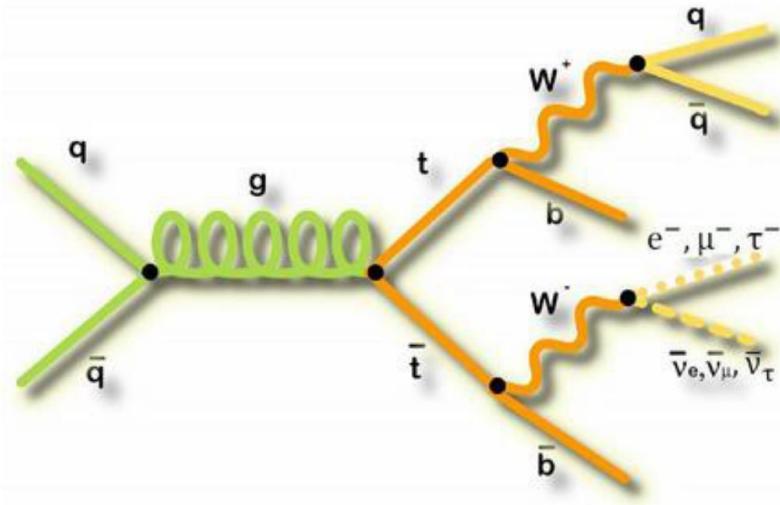
- O processo físico que queremos estudar é o

$$t\bar{t} \rightarrow (bW^+)(\bar{b}W^-) \rightarrow (b\ell^+\nu_\ell)(\bar{b}q\bar{q})$$


# Mais informações sobre o nosso processo

- O processo físico que queremos estudar é o

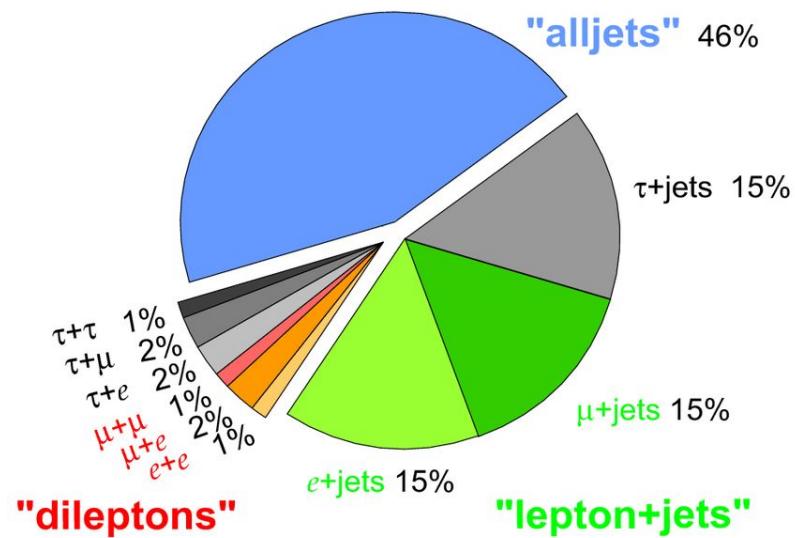
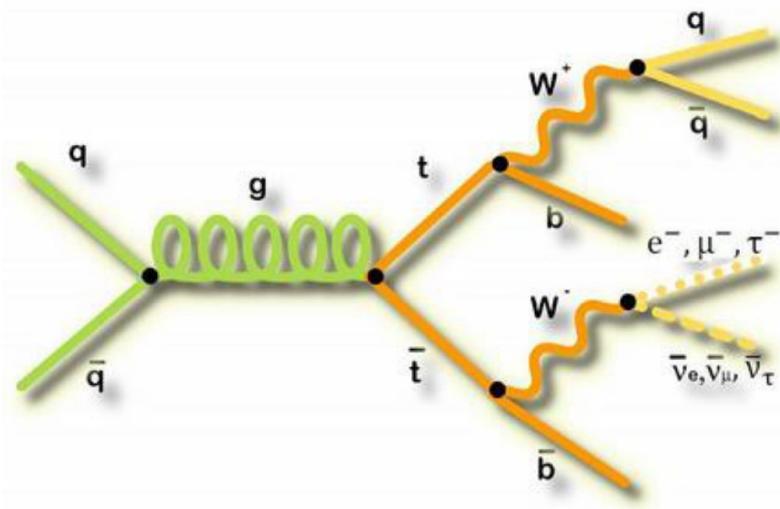
$$t\bar{t} \rightarrow (bW^+)(\bar{b}W^-) \rightarrow (b\ell^+\nu_\ell)(\bar{b}q\bar{q})$$



# Mais informações sobre o nosso processo

- O processo físico que queremos estudar é o

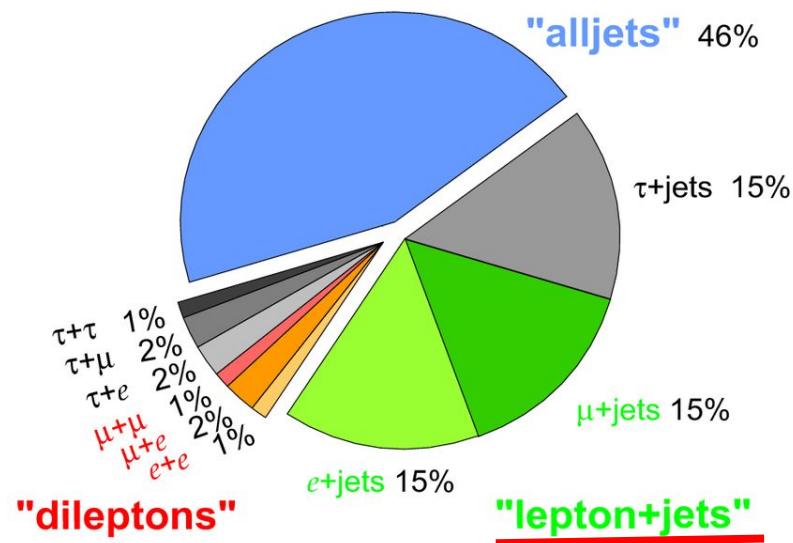
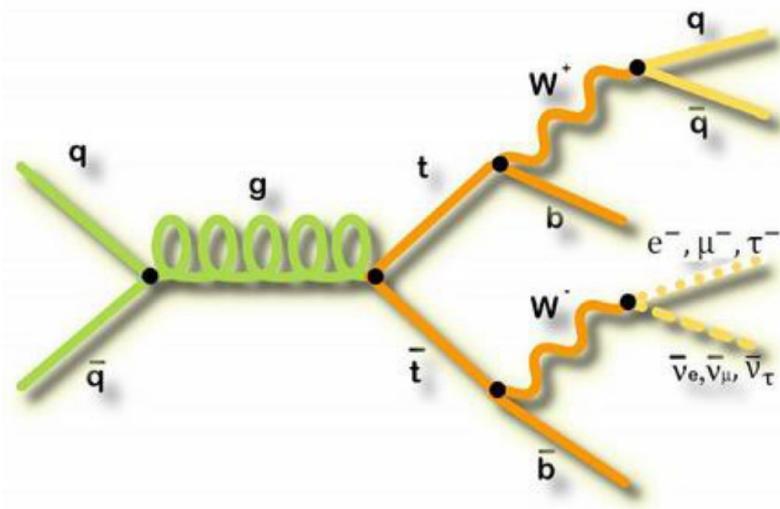
$$t\bar{t} \rightarrow (bW^+)(\bar{b}W^-) \rightarrow (b\ell^+\nu_\ell)(\bar{b}q\bar{q})$$



# Mais informações sobre o nosso processo

- O processo físico que queremos estudar é o

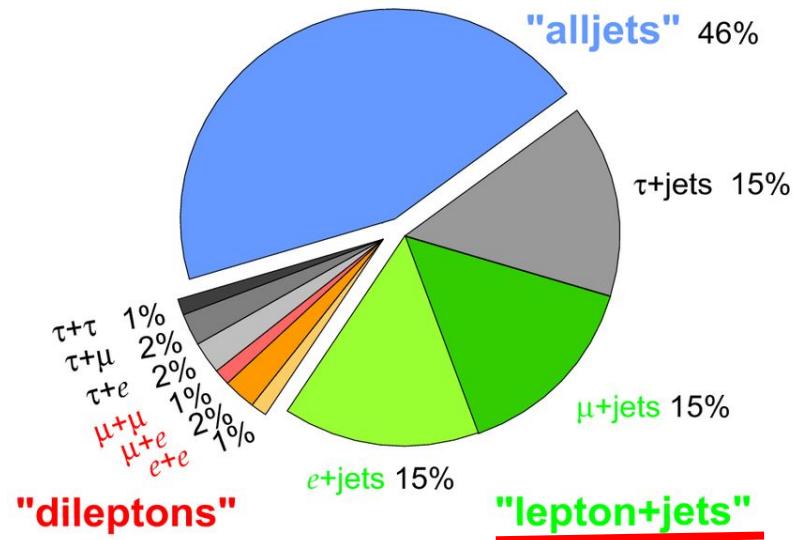
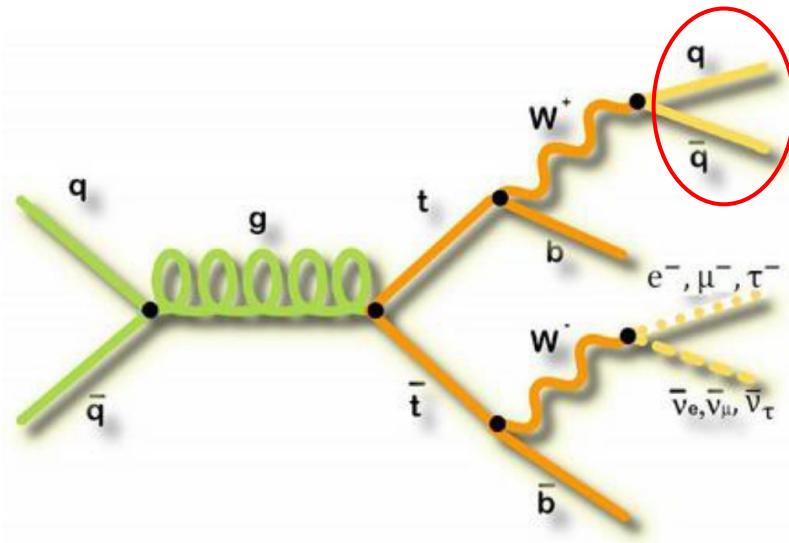
$$t\bar{t} \rightarrow (bW^+)(\bar{b}W^-) \rightarrow (b\ell^+\nu_\ell)(\bar{b}q\bar{q})$$



# Mais informações sobre o nosso processo

- O processo físico que queremos estudar é o

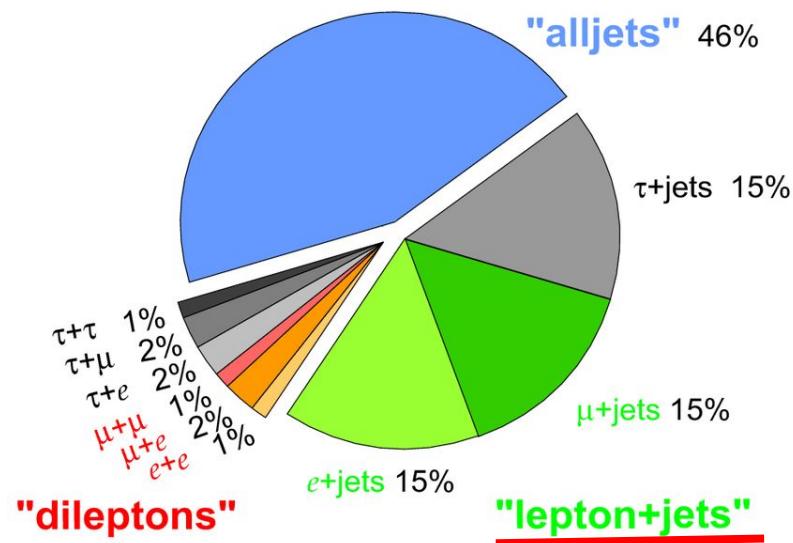
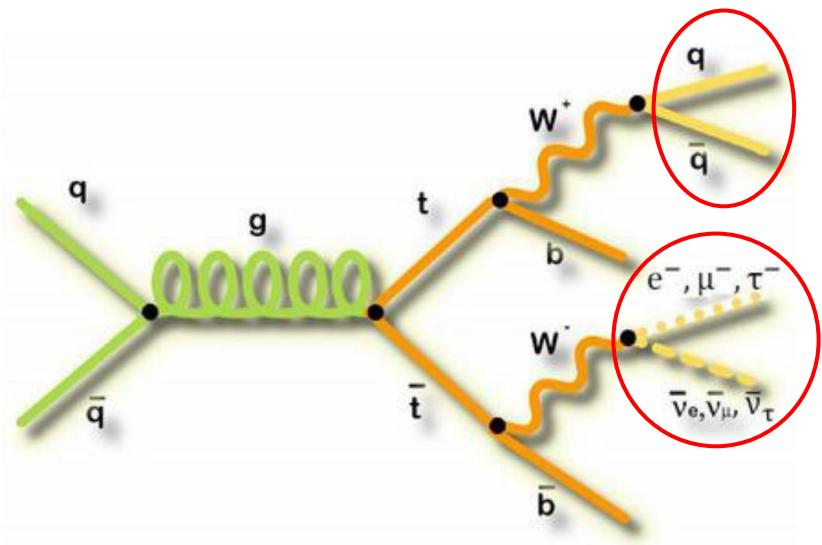
$$t\bar{t} \rightarrow (bW^+)(\bar{b}W^-) \rightarrow (b\ell^+\nu_\ell)(\bar{b}q\bar{q})$$



# Mais informações sobre o nosso processo

- O processo físico que queremos estudar é o

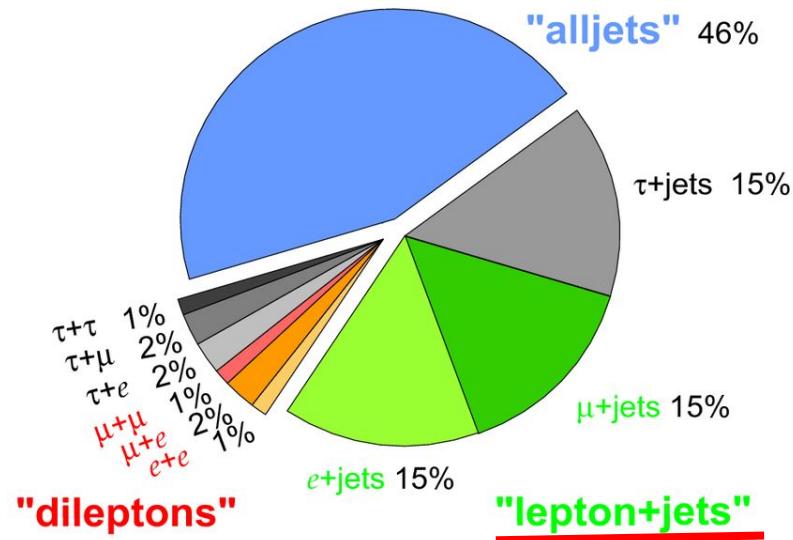
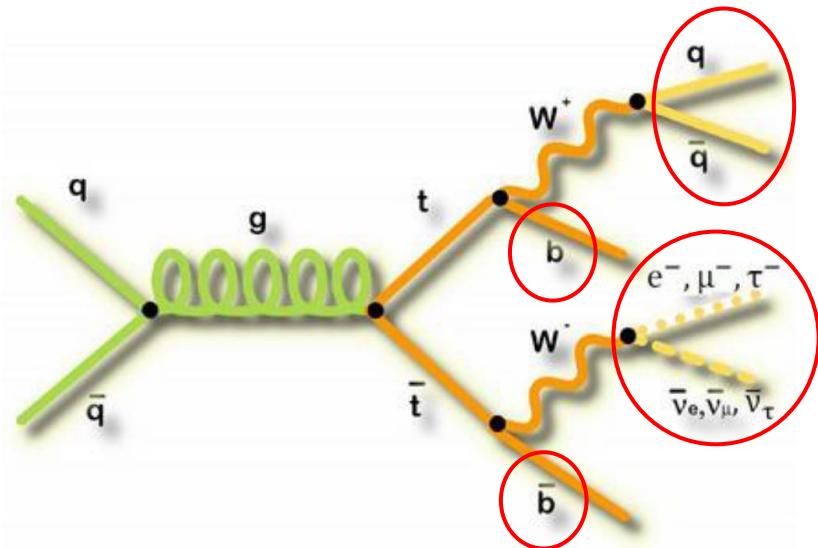
$$t\bar{t} \rightarrow (bW^+)(\bar{b}W^-) \rightarrow (b\ell^+\nu_\ell)(\bar{b}q\bar{q})$$



# Mais informações sobre o nosso processo

- O processo físico que queremos estudar é o

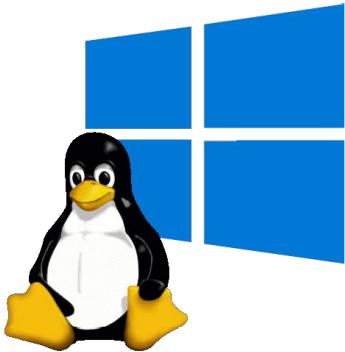
$$t\bar{t} \rightarrow (bW^+)(\bar{b}W^-) \rightarrow (b\ell^+\nu_\ell)(\bar{b}q\bar{q})$$





# Preparando o ambiente

# Windows Subsystem for Linux



Caso esteja utilizando o Windows, será aconselhado utilizar o Windows Subsystem for Linux (WLS), que é uma ferramenta que te permite executar binários e scripts em Linux diretamente no Windows.

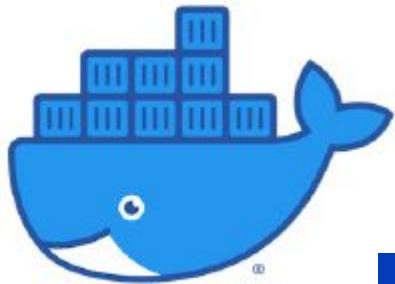
Para instalá-lo basta ir no PowerShell do windows e digitar:

```
wsl --install
```

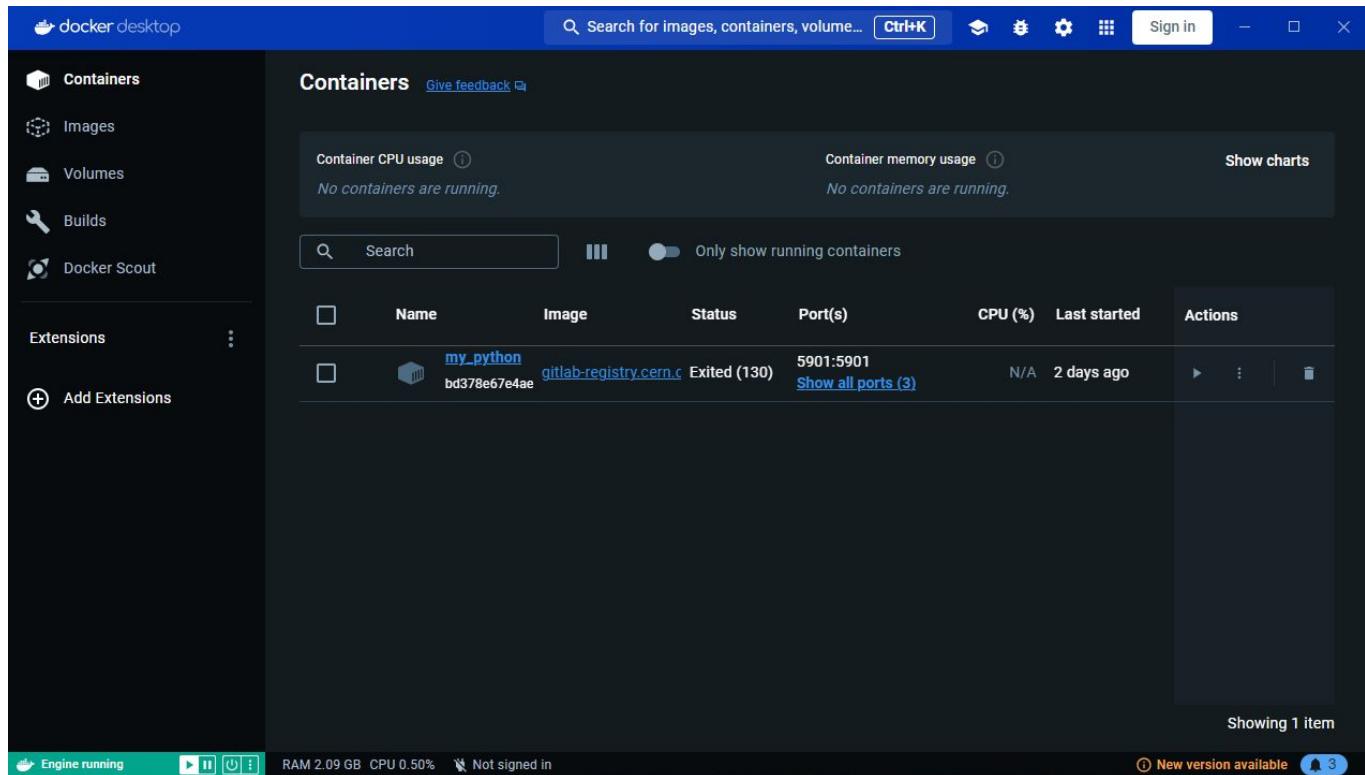
Fica assim no PowerShell

```
C:\Windows\system32> wsl --install
```

# Baixando e executando o Docker



Para baixar o docker desktop clique no [link](#), execute e abra o programa.



The screenshot shows the Docker Desktop interface. On the left, a sidebar includes 'Containers', 'Images', 'Volumes', 'Builds', 'Docker Scout', 'Extensions' (with an 'Add Extensions' button), and a 'Search' bar. The main area is titled 'Containers' and displays a table with one item:

Name	Image	Status	Port(s)	CPU (%)	Last started	Actions
my_python	gitlab-registry.cern.c	Exited (130)	5901:5901 <a href="#">Show all ports (3)</a>	N/A	2 days ago	 

At the bottom, status indicators show 'Engine running', system resources ('RAM 2.09 GB CPU 0.50%'), and a note 'Not signed in'. A message at the bottom right says 'Showing 1 item'. The top bar includes a search bar, keyboard shortcut 'Ctrl+K', and various icons for account, notifications, and settings.



# Baixando o contêiner que usaremos

Dentro do WSL, crie uma pasta onde vão ficar os arquivos do seu projeto:

```
cd  
mkdir cms_open_data_python  
chmod -R 777 cms_open_data_python
```

Caso esteja no Linux:

```
docker run -it --name my_python -P -p 8888:8888 --net=host --env="DISPLAY" -v  
$HOME/.Xauthority:/home/cmsusr/.Xauthority:rw -v ${HOME}/cms_open_data_python:/code  
gitlab-registry.cern.ch/cms-cloud/python-vnc:python3.10.5
```

Caso esteja no MacOS ou Windows WLS2:

```
docker run -it --name my_python -P -p 5901:5901 -p 6080:6080 -p 8888:8888 -v  
${HOME}/cms_open_data_python:/code gitlab-registry.cern.ch/cms-cloud/python-vnc:python3.10.5
```

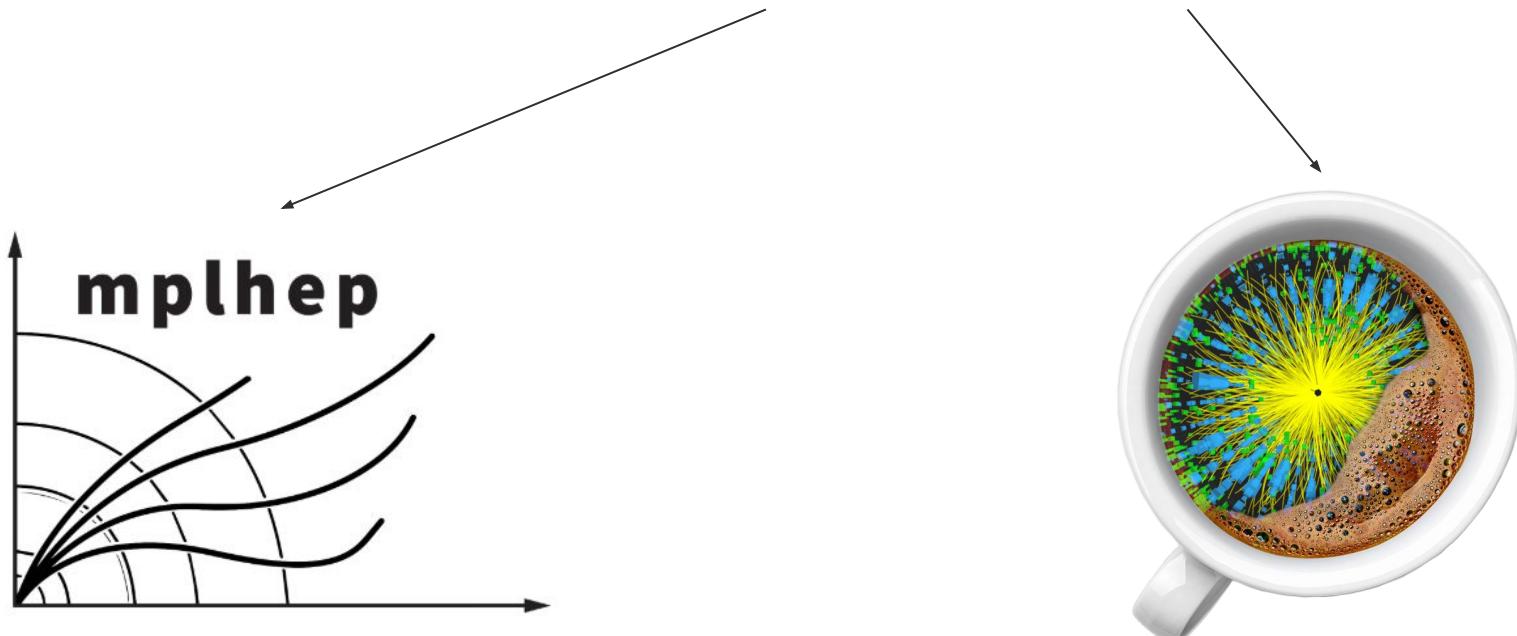
# Acessando o container

Para acessar o container use o comando:

```
docker start -i my_python
```

Baixe as seguintes versões dos pacotes:

```
pip install vector==0.8 hist mplhep==0.3.26 cabinetry coffea==0.7.21
```



# Acessando o container

Para acessar o container use o comando:

```
docker start -i my_python
```

Baixe as seguintes versões dos pacotes:

```
pip install vector==0.8 hist mplhep==0.3.26 cabinetry coffea==0.7.21
```

Execute o Jupyter notebook

```
jupyter-lab --ip=0.0.0.0 --no-browser
```





# Como acessar o jupyter notebook

```
(base) bellis@titan:~$ docker run -it --name my_python -P -p 8888:8888 --net=host --env="DISPLAY" -v $HOME/.Xauthority:/home/cmsusr/.Xauthority:rw -v ${HOME}/cms_open_data_python:/code gitlab-registry.cern.ch/cms-cloud/python-vnc:latest
WARNING: Published ports are discarded when using host network mode
cmsusr@titan:/code$ jupyter-lab --ip=0.0.0.0 --no-browser
[I 2023-07-08 20:22:45.113 ServerApp] jupyterlab | extension was successfully linked.
[I 2023-07-08 20:22:45.121 ServerApp] nbclassic | extension was successfully linked.
[I 2023-07-08 20:22:45.122 ServerApp] Writing Jupyter server cookie secret to /home/cmsusr/.local/share/jupyter/runtime/jupyter_cookie_secret
[I 2023-07-08 20:22:45.336 ServerApp] notebook_shim | extension was successfully linked.
[I 2023-07-08 20:22:45.351 ServerApp] notebook_shim | extension was successfully loaded.
[I 2023-07-08 20:22:45.352 LabApp] JupyterLab extension loaded from /usr/local/venv/lib/python3.10/site-packages/jupyterlab
[I 2023-07-08 20:22:45.352 LabApp] JupyterLab application directory is /usr/local/venv/share/jupyter/lab
[I 2023-07-08 20:22:45.354 ServerApp] jupyterlab | extension was successfully loaded.
[I 2023-07-08 20:22:45.357 ServerApp] nbclassic | extension was successfully loaded.
[I 2023-07-08 20:22:45.357 ServerApp] Serving notebooks from local directory: /code
[I 2023-07-08 20:22:45.358 ServerApp] Jupyter Server 1.18.1 is running at:
[I 2023-07-08 20:22:45.358 ServerApp] http://titan:8888/lab?token=6ce000cacd64984166f8d5462bb9404fcffcc1fc6123777e
[I 2023-07-08 20:22:45.358 ServerApp] or http://127.0.0.1:8888/lab?token=6ce000cacd64984166f8d5462bb9404fcffcc1fc6123777e
[I 2023-07-08 20:22:45.358 ServerApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation).
[IC 2023-07-08 20:22:45.360 ServerApp]
```

To access the server, open this file in a browser:  
file:///home/cmsusr/.local/share/jupyter/runtime/jpserver-9-open.html  
Or copy and paste one of these URLs:  
<http://titan:8888/lab?token=6ce000cacd64984166f8d5462bb9404fcffcc1fc6123777e>  
or <http://127.0.0.1:8888/lab?token=6ce000cacd64984166f8d5462bb9404fcffcc1fc6123777e>

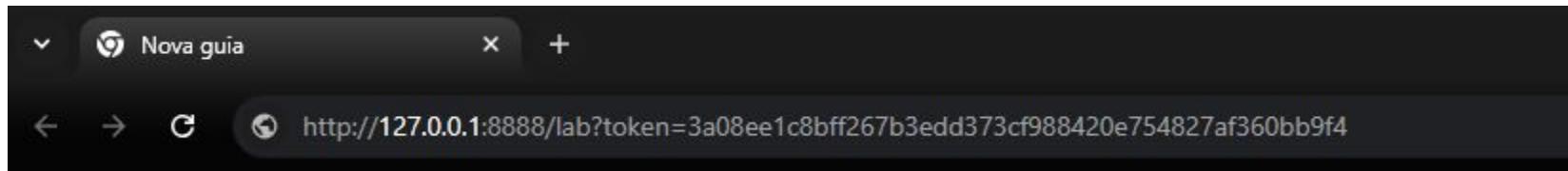


# Como acessar o jupyter notebook

```
(base) bellis@titan:~$ docker run -it --name my_python -P -p 8888:8888 --net=host --env="DISPLAY" -v $HOME/.Xauthority:/home/cmsusr/.Xauthority:rw -v ${HOME}/cms_open_data_python:/code gitlab-registry.cern.ch/cms-cloud/python-vnc:latest
WARNING: Published ports are discarded when using host network mode
cmsusr@titan:/code$ jupyter-lab --ip=0.0.0.0 --no-browser
[I 2023-07-08 20:22:45.113 ServerApp] jupyterlab | extension was successfully linked.
[I 2023-07-08 20:22:45.121 ServerApp] nbclassic | extension was successfully linked.
[I 2023-07-08 20:22:45.122 ServerApp] Writing Jupyter server cookie secret to /home/cmsusr/.local/share/jupyter/runtime/jupyter_cookie_secret
[I 2023-07-08 20:22:45.336 ServerApp] notebook_shim | extension was successfully linked.
[I 2023-07-08 20:22:45.351 ServerApp] notebook_shim | extension was successfully loaded.
[I 2023-07-08 20:22:45.352 LabApp] JupyterLab extension loaded from /usr/local/venv/lib/python3.10/site-packages/jupyterlab
[I 2023-07-08 20:22:45.352 LabApp] JupyterLab application directory is /usr/local/venv/share/jupyter/lab
[I 2023-07-08 20:22:45.354 ServerApp] jupyterlab | extension was successfully loaded.
[I 2023-07-08 20:22:45.357 ServerApp] nbclassic | extension was successfully loaded.
[I 2023-07-08 20:22:45.357 ServerApp] Serving notebooks from local directory: /code
[I 2023-07-08 20:22:45.358 ServerApp] Jupyter Server 1.18.1 is running at:
[I 2023-07-08 20:22:45.358 ServerApp] http://titan:8888/lab?token=6ce000cacd64984166f8d5462bb9404fcffcc1fc6123777e
[I 2023-07-08 20:22:45.358 ServerApp] or http://127.0.0.1:8888/lab?token=6ce000cacd64984166f8d5462bb9404fcffcc1fc6123777e
[I 2023-07-08 20:22:45.358 ServerApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation).
[IC 2023-07-08 20:22:45.360 ServerApp]
```

To access the server, open this file in a browser:  
file:///home/cmsusr/.local/share/jupyter/runtime/jpserver-9-open.html  
Or copy and paste one of these URLs:  
<http://titan:8888/lab?token=6ce000cacd64984166f8d5462bb9404fcffcc1fc6123777e>  
or <http://127.0.0.1:8888/lab?token=6ce000cacd64984166f8d5462bb9404fcffcc1fc6123777e>

Vá no seu browser e digite um dos URLs, isso te levará para o seu jupyter notebook





# Jupyter notebook

The screenshot shows the JupyterLab interface running at `127.0.0.1:8888/lab`. On the left, a file browser displays a list of files:

Name	Last Modified
workshop2...	15 days ago
agc_sche...	15 days ago
analyze.C	8 days ago
analyze.py	8 days ago
MYANALY...	15 days ago
myoutput.r...	17 days ago
output.zip	8 days ago

The main area is a "Launcher" pane with sections for "Notebook", "Console", and "Other".

- Notebook:** Two entries for "Python 3 (ipykernel)".
- Console:** One entry for "Python 3 (ipykernel)".
- Other:** Icons for "Terminal", "Text File", "Markdown File", "Python File", and "Show Contextual Help".



# Jupyter notebook

A screenshot of a web browser displaying the JupyterLab interface at the URL 127.0.0.1:8888/lab. The interface includes a file explorer on the left showing local files like 'workshop2...', 'agc\_sche...', 'analyze.C', 'analyze.py', 'MYANALY...', 'myoutput.r...', and 'output.zip'. The main area is a 'Launcher' pane containing a grid of icons. One icon, 'Notebook' (represented by a yellow square icon), is highlighted with a red box and a red arrow points from it to a callout box. The callout box contains the text: 'Clique nesse ícone, isso vai te levar ao seu jupyter notebook!' (Click this icon, it will take you to your Jupyter notebook!). Other icons in the launcher include 'Console' (Python 3 (ipykernel)), 'Terminal', 'Text File', 'Markdown File', 'Python File', and 'Show Contextual Help'.

Clique nesse ícone, isso vai te levar ao seu jupyter notebook!



# Baixando o repositório do projeto

Baixe o repositório do projeto que usaremos:

```
git clone https://github.com/cms-opendata-workshop/workshop2022-lesson-ttbarljetsanalysis-payload.git  
cd workshop2022-lesson-ttbarljetsanalysis-payload
```

FIM

Com isso você estará apto a usar todas as funções e acessar os dados desse projeto através da sua própria máquina!



# Baixando o repositório do projeto

Baixe o repositório do projeto que usaremos:

```
git clone https://github.com/cms-opendata-workshop/workshop2022-lesson-ttbarljetsanalysis-payload.git  
cd workshop2022-lesson-ttbarljetsanalysis-payload
```

FIM

Com isso você estará apto a usar todas as funções e acessar os dados desse projeto através da sua própria máquina!

Você já consegue acessar os dados do Open Data Portal através de comandos no jupyter notebook como esses:

```
import uproot  
import awkward as ak  
  
events =  
uproot.open('root://eospublic.cern.ch//eos/opendata/cms/upload/od-workshop/ws2021/myoutput_odws2022-ttbaj  
ets-prodV2.0_merged.root')['events']  
events
```



# Começando a análise!

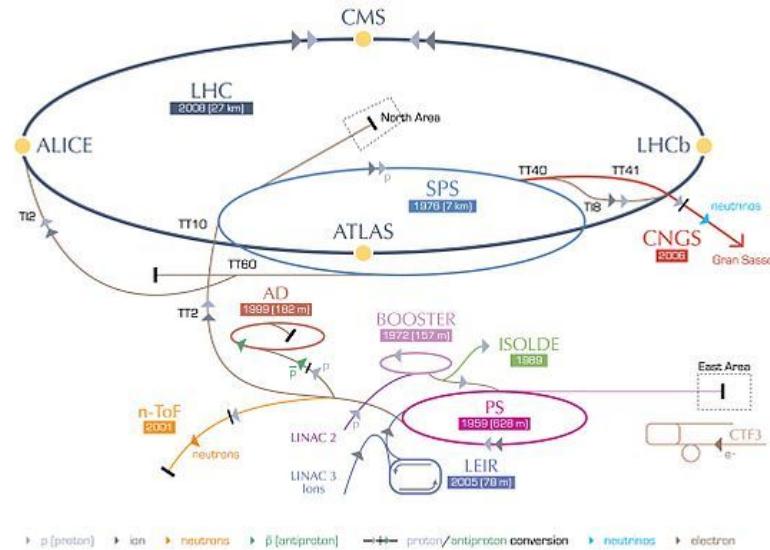
# Contextualizando



Aceleramos os prótons  
através do complexos de  
aceleradores do CERN



CERN's accelerator complex



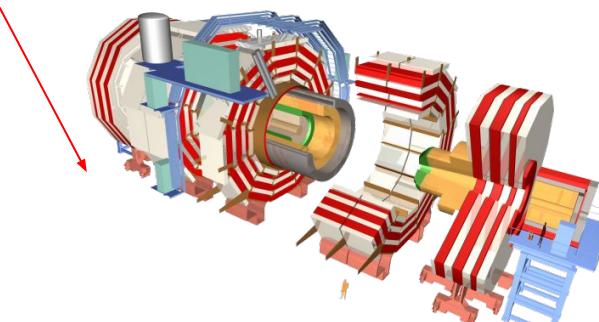
LHC Large Hadron Collider   SPS Super Proton Synchrotron   PS Proton Synchrotron

AD Antiproton Decelerator   CTF3 Clic Test Facility   CNGS Cern Neutrinos to Gran Sasso   ISOLDE Isotope Separator OnLine DDevice  
LEIR Low Energy Ion Ring   LINAC LINear ACcelerator   n-ToF Neutrons Time Of Flight

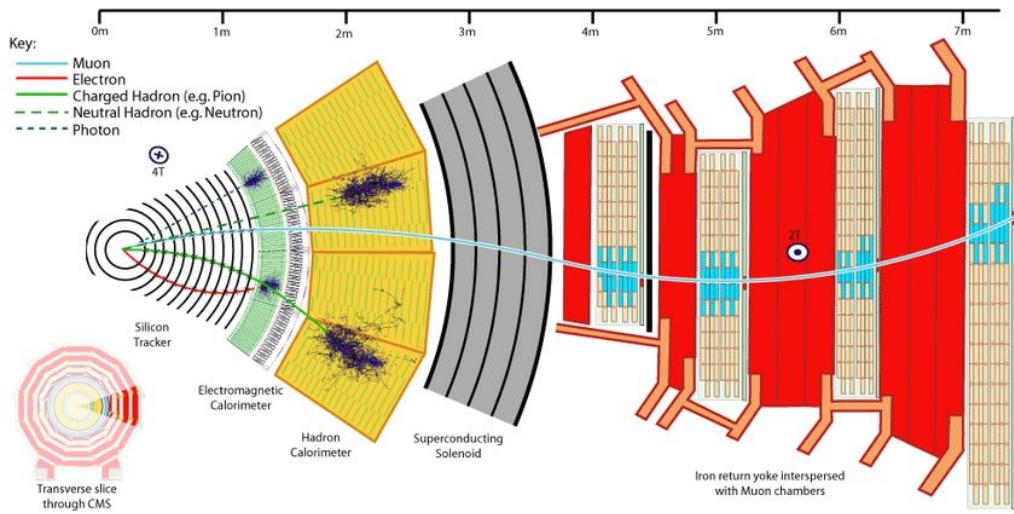
European Organization for Nuclear Research | Organisation européenne pour la recherche nucléaire

© CERN 2008

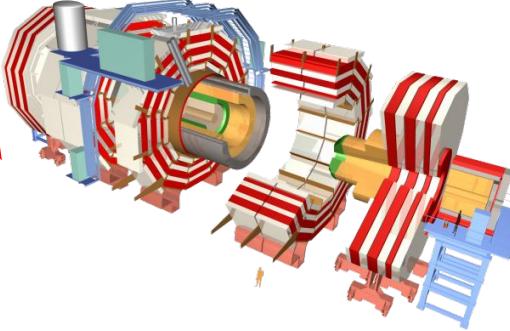
# Contextualizando



Colidimos os prótons em pontos específicos, onde estão localizados nossos detectores

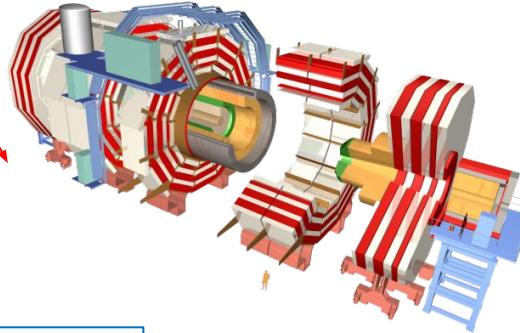


# Contextualizando

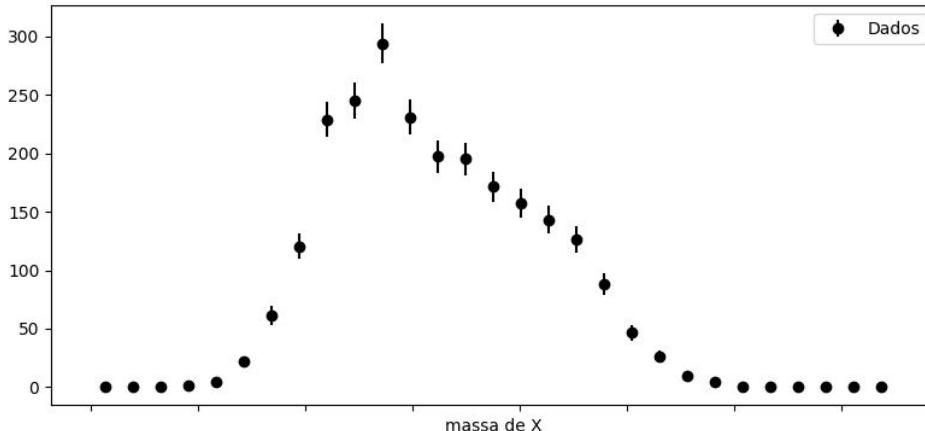


Utilizamos uma gama de softwares de reconstrução para identificar e armazenar informações das colisões

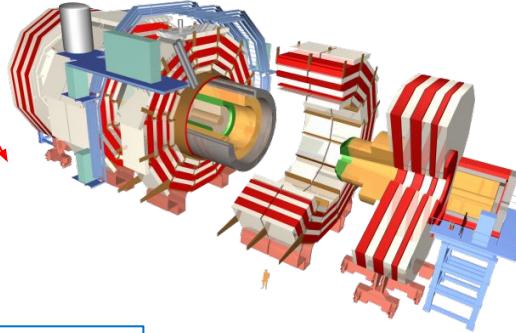
# Contextualizando



Estudamos diferentes distribuições, referentes às partículas previamente reconstruídas pelos softwares!



# Contextualizando



Estudamos diferentes distribuições, referentes às partículas previamente reconstruídas pelos softwares!

Essa distribuição é composta por quantos processos físicos diferentes?

```

    static void Main(string[] args)
    {
        string str = "Hello, World!";
        string result = ReverseString(str);
        Console.WriteLine(result);
    }

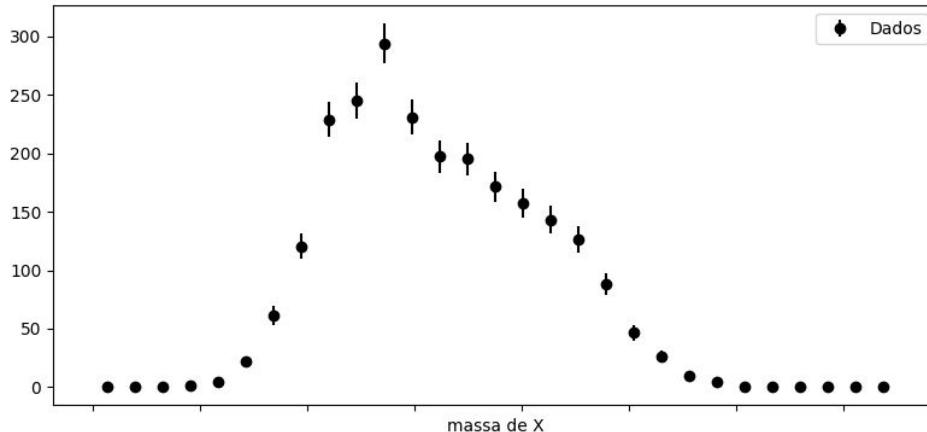
    static string ReverseString(string str)
    {
        if (str == null || str.Length == 0)
            return str;

        int start = 0;
        int end = str.Length - 1;

        while (start < end)
        {
            char temp = str[start];
            str[start] = str[end];
            str[end] = temp;
            start++;
            end--;
        }

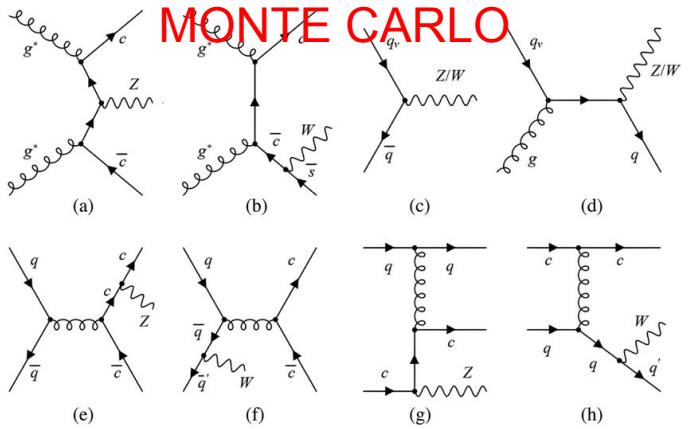
        return str;
    }
}

```



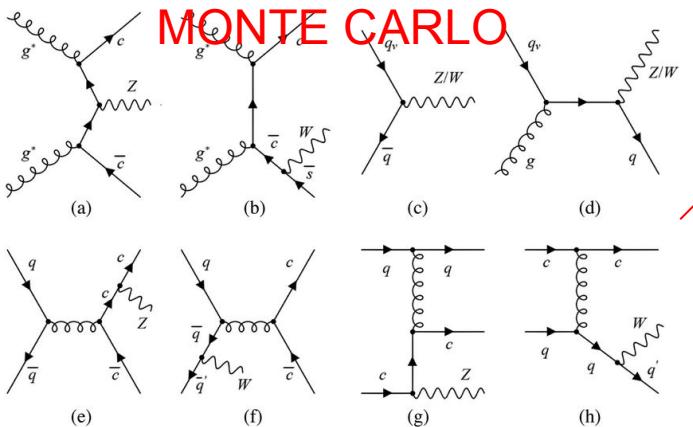
# Contextualizando

MONTE CARLO



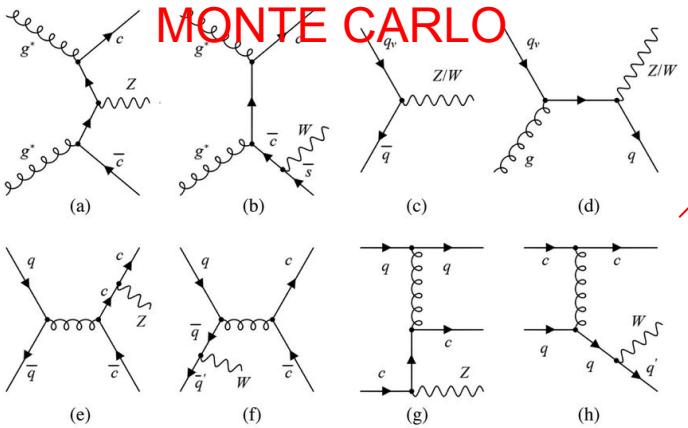
Monte Carlo é o método estatístico utilizado para criarmos dados que representam os diferentes processos físicos (diferentes diagramas de Feynman!)

# Contextualizando



Submetemos os dados gerados, pelo método de Monte Carlo, a uma simulação do nosso experimento, no nosso caso o CMS. Com isso são gerados dados similares aos *outputs* do detector!

# Contextualizando



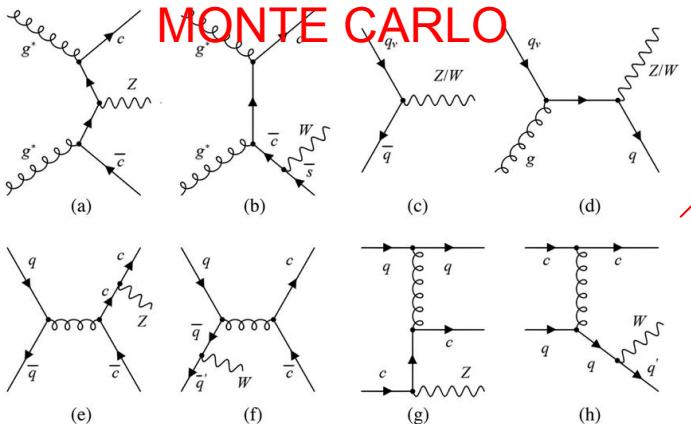
Aplicamos esses dados aos mesmos softwares de reconstrução utilizados anteriormente (para os dados!).

```
    void check(Condition c) {
        if (c == null) return;
        if (c instanceof Condition) {
            Condition condition = (Condition)c;
            if (condition.getCondition() != null) {
                check(condition.getCondition());
            }
            if (condition.getFunction() != null) {
                Function function = condition.getFunction();
                if (function.getFunction() != null) {
                    check(function.getFunction());
                }
                if (function.getCondition() != null) {
                    check(function.getCondition());
                }
            }
        }
    }
}

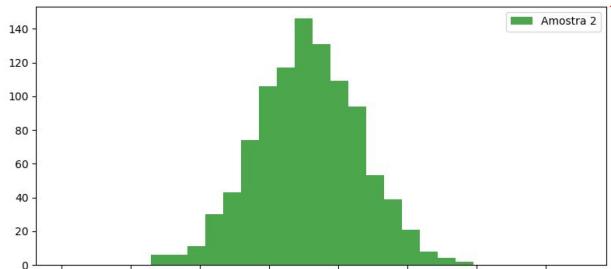
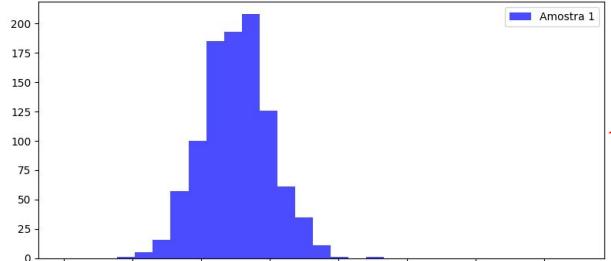
class Condition {
    Condition getCondition() { return null; }
    Function getFunction() { return null; }
}

class Function {
    Condition getCondition() { return null; }
    String getFunction() { return null; }
}
```

# Contextualizando

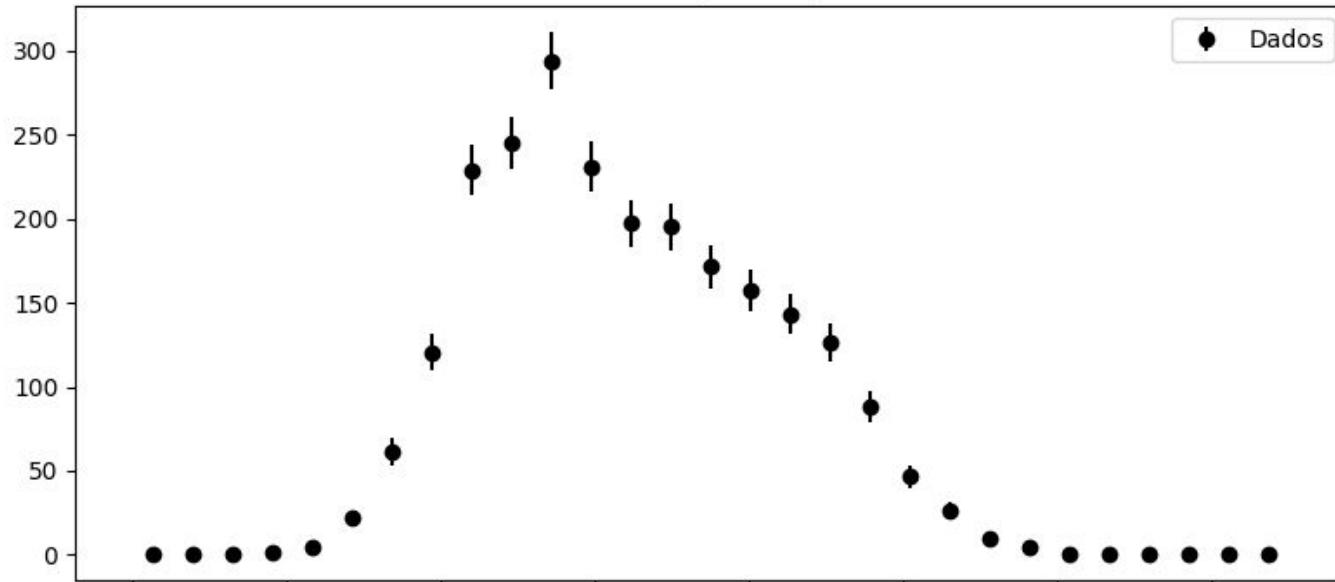


# MONTE CARLO



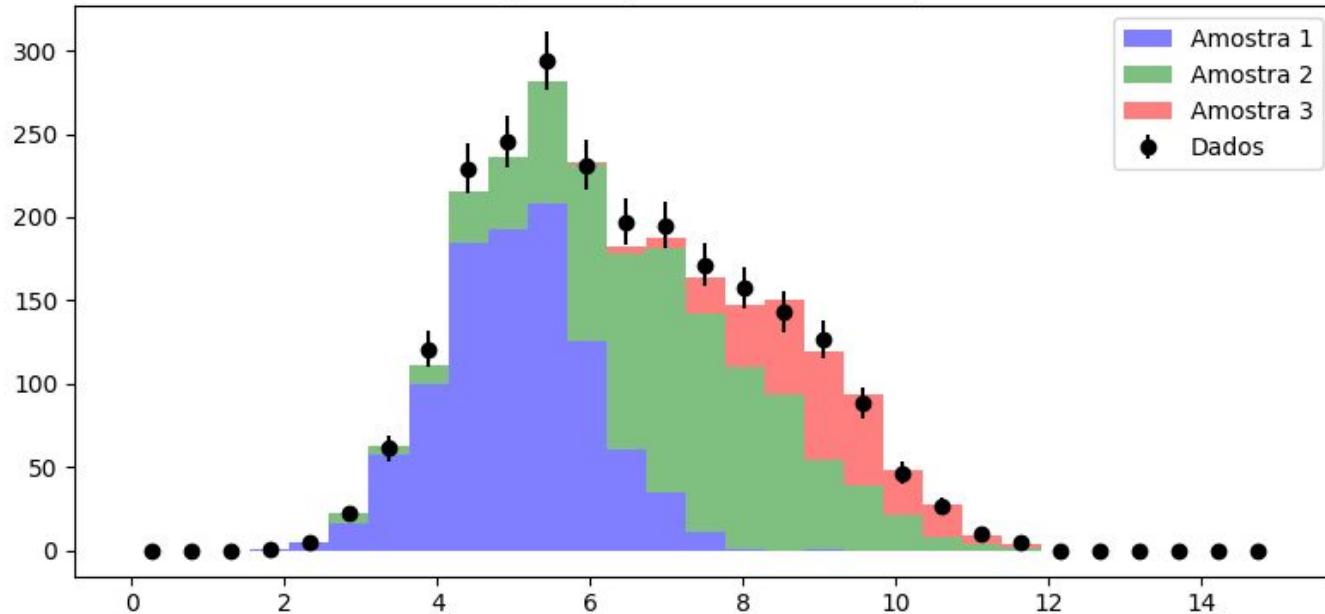
E com isso somos capazes de identificar qual a contribuição de diferentes tipos de processos físicos!

# Contextualizando



A distribuição geradas pelas colisões é, inicialmente uma **caixa-preta**, não sendo possível mensurar a contribuição de diferentes processos físicos!

# Contextualizando



Mas utilizando a **simulação** somos capazes de identificar a contribuição de diferentes processos! Podendo assim, identificar se as nossas distribuições advém somente do processo que queremos estudar!

# Resumindo

- Os dados que vamos usar para mensurar nossa medida é o que vem das colisões!
- Não temos certeza se os dados que estamos trabalhando são do processo físico que queremos.
- Para que tenhamos certeza utilizamos a **simulação**.
- Através dela vamos conseguir guiar melhor a tomada de decisão com relação a filtragem (cortes de qualidade/seleção de sinal) e tratamento dos dados (fatores de escala e correções no geral).

# Resumindo

- Os dados que vamos usar para mensurar nossa medida é o que vem das colisões!
- Não temos certeza se os dados que estamos trabalhando são do processo físico que queremos.
- Para que tenhamos certeza utilizamos a **simulação**.
- Através dela vamos conseguir guiar melhor a tomada de decisão com relação a filtragem (cortes de qualidade/seleção de sinal) e tratamento dos dados (fatores de escala e correções no geral).

Podemos agora avançar para a primeira etapa da análise:  
**Lista de dados/amostras que vamos utilizar**

# Lista de dados e amostras

type	dataset	variation	CODP name	xsec [pb]	size (TB)	Rec ID
colisions data	data	nominal	/SingleMuon/Run2015D-16Dec2015-v1/MINIAOD		1.4	24119
			/SingleElectron/Run2015D-08Jun2016-v1/MINIAOD		2.6	24120
signal	ttbar	nominal	TT_TuneCUETP8M1_13TeV-powheg-pythia8	831.76	3.4	19980
		scaledown	TT_TuneCUETP8M1_13TeV-powheg-scaledown-pythia8		1.4	19983
		scaleup	TT_TuneCUETP8M1_13TeV-powheg-scaleup-pythia8		1.3	19985
		ME_var	TTJets_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8		1.3	19949
		PS_var	TT_TuneEE5C_13TeV-powheg-herwigpp		0.8	19999
backgrounds	single_top_t_chan	nominal	ST_t-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1	44.33	0.52	19397
	single_atop_t_chan	nominal	ST_t-channel_antitop_4f_leptonDecays_13TeV-powheg-pythia8_TuneCUETP8M1		26.38	0.042
	single_top_tW	nominal	ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.6	0.03	19419
			ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1		35.6	0.03
wjets	nominal		WJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	61526.7	3.8	20548

# Lista de dados e amostras

type	dataset	variation	CODP name	xsec [pb]	size (TB)	Rec ID
colisions data	data	nominal	/SingleMuon/Run2015D-16Dec2015-v1/MINIAOD		1.4	24119
			/SingleElectron/Run2015D-08Jun2016-v1/MINIAOD			
signal	ttbar	nominal	TT_TuneCUETP8M1_13TeV-powheg-pythia8	831.76	3.4	19980
			TT_TuneCUETP8M1_13TeV-powheg-scaledown-pythia8			
			TT_TuneCUETP8M1_13TeV-powheg-scaleup-pythia8			
		ME_var	TTJets_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	1.3	1.3	19949
		PS_var	TT_TuneEE5C_13TeV-powheg-herwigpp			
backgrounds	single_top_t_chan	nominal	ST_t-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1	44.33	0.52	19397
	single_atop_t_chan	nominal	ST_t-channel_antitop_4f_leptonDecays_13TeV-powheg-pythia8_TuneCUETP8M1			
	single_top_tW	nominal	ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.6	0.03	19419
		nominal	ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1			
wjets	wjets	nominal	WJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	61526.7	3.8	20548

# Lista de dados e amostras

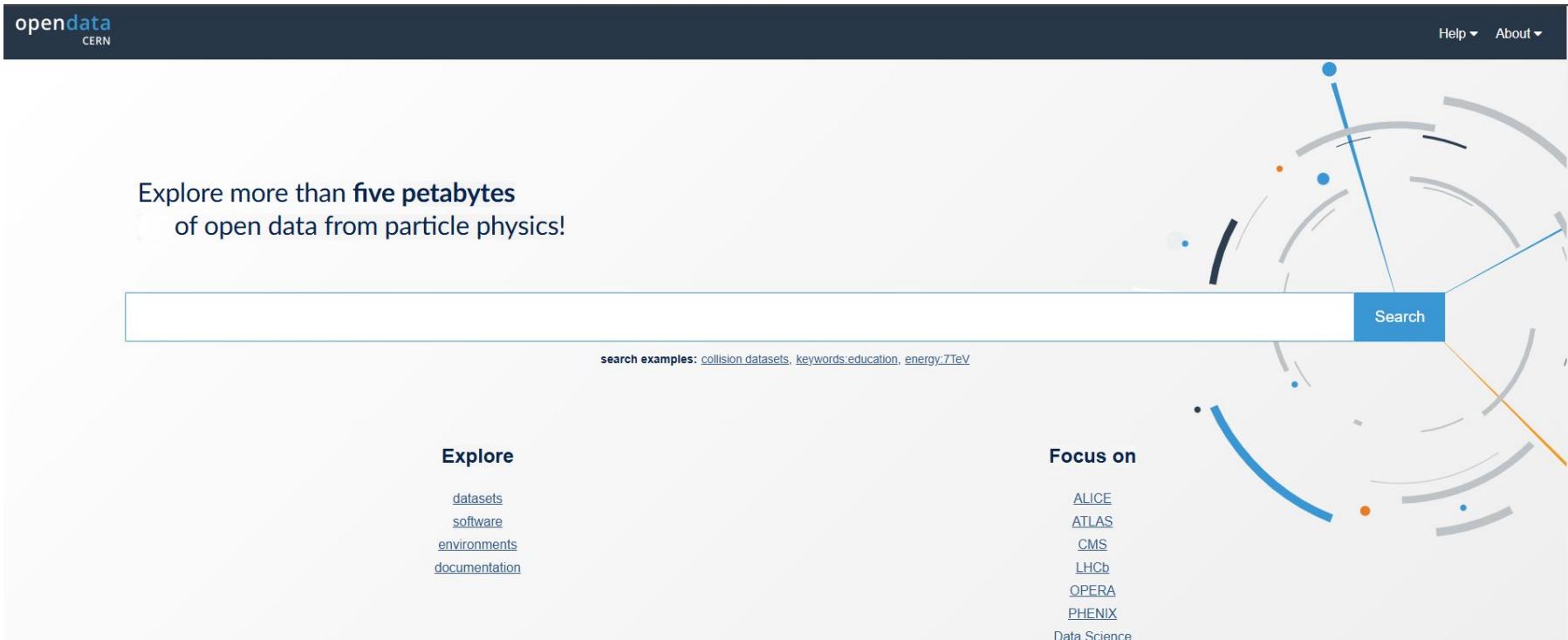
type	dataset	variation	CODP name	xsec [pb]	size (TB)	Rec ID
colisions data	data		nominal /SingleMuon/Run2015D-16Dec2015-v1/MINIAOD		1.4	24119
			/SingleElectron/Run2015D-08Jun2016-v1/MINIAOD		2.6	24120
signal	ttbar	nominal	TT_TuneCUETP8M1_13TeV-powheg-pythia8	831.76	3.4	19980
		scaledown	TT_TuneCUETP8M1_13TeV-powheg-scaledown-pythia8		1.4	19983
		scaleup	TT_TuneCUETP8M1_13TeV-powheg-scaleup-pythia8		1.3	19985
		ME_var	TTJets_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8		1.3	19949
		PS_var	TT_TuneEE5C_13TeV-powheg-herwigpp		0.8	19999
backgrounds	single_top_t_chan	nominal	ST_t-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1	44.33	0.52	19397
	single_atop_t_chan	nominal	ST_t-channel_antitop_4f_leptonDecays_13TeV-powheg-pythia8_TuneCUETP8M1	26.38	0.042	19407
	single_top_tW	nominal	ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.6	0.03	19419
			ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.6	0.03	19412
wjets	nominal		WJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	61526.7	3.8	20548

# Lista de dados e amostras

type	dataset	variation	CODP name	xsec [pb]	size (TB)	Rec ID
colisions data	data	nominal	/SingleMuon/Run2015D-16Dec2015-v1/MINIAOD		1.4	24119
			/SingleElectron/Run2015D-08Jun2016-v1/MINIAOD		2.6	24120
signal	ttbar	nominal	TT_TuneCUETP8M1_13TeV-powheg-pythia8	831.76	3.4	19980
		scaledown	TT_TuneCUETP8M1_13TeV-powheg-scaledown-pythia8		1.4	19983
		scaleup	TT_TuneCUETP8M1_13TeV-powheg-scaleup-pythia8		1.3	19985
		ME_var	TTJets_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8		1.3	19949
		PS_var	TT_TuneEE5C_13TeV-powheg-herwigpp		0.8	19999
backgrounds	single_top_t_chan	nominal	ST_t-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1	44.33	0.52	19397
	single_atop_t_chan	nominal	ST_t-channel_antitop_4f_leptonDecays_13TeV-powheg-pythia8_TuneCUETP8M1	26.38	0.042	19407
	single_top_tW	nominal	ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.6	0.03	19419
			ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.6	0.03	19412
	wjets	nominal	WJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	61526.7	3.8	20548

# Utilizando o CERN Open Data Portal

- Para achar os datasets que se quer trabalhar, usaremos o [CERN Open Data Portal](#)



opendata  
CERN

Help ▾ About ▾

Explore more than **five petabytes** of open data from particle physics!

search examples: [collision datasets](#), [keywords:education](#), [energy:7TeV](#)

**Explore**

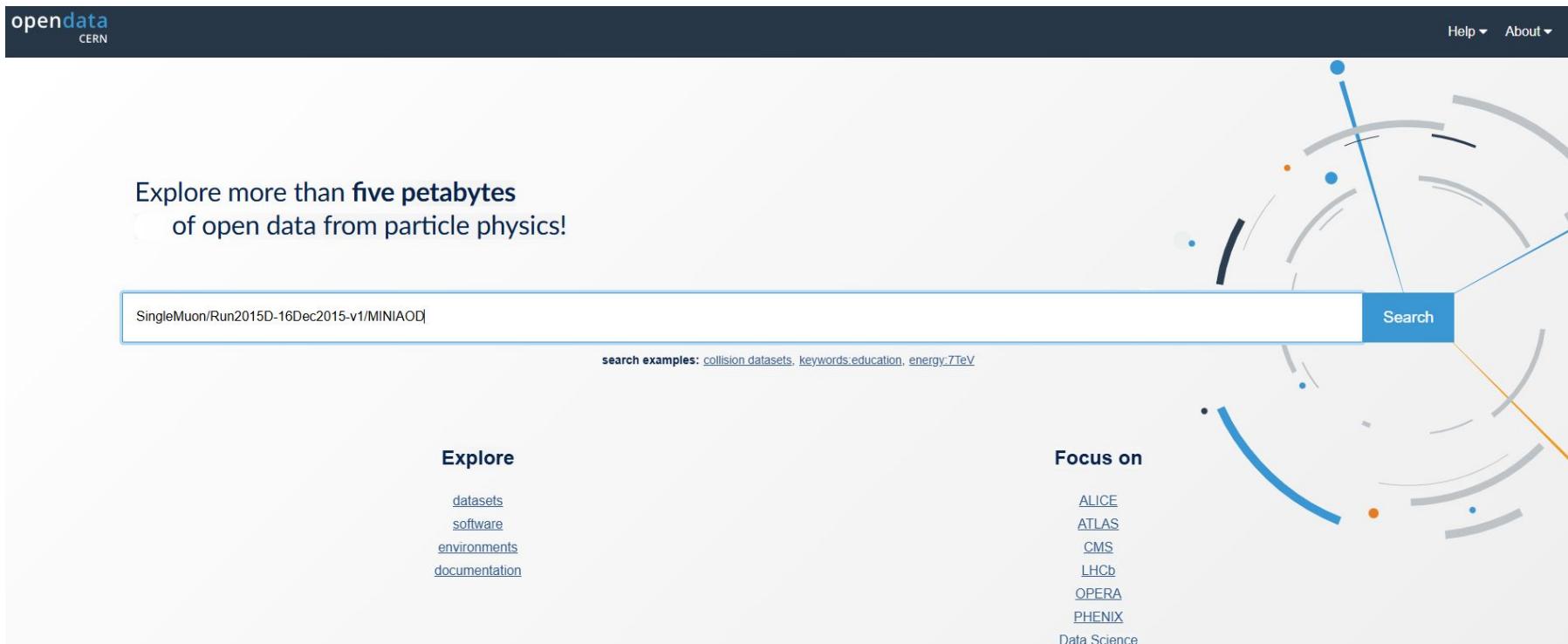
[datasets](#)  
[software](#)  
[environments](#)  
[documentation](#)

**Focus on**

[ALICE](#)  
[ATLAS](#)  
[CMS](#)  
[LHCb](#)  
[OPERA](#)  
[PHENIX](#)  
[Data Science](#)

# Utilizando o CERN Open Data Portal

- Para achar os datasets que se quer trabalhar, usaremos o [CERN Open Data Portal](#)



Explore more than **five petabytes** of open data from particle physics!

SingleMuon/Run2015D-16Dec2015-v1/MINIAOD

search examples: collision datasets, keywords:education, energy:7TeV

**Explore**

[datasets](#)  
[software](#)  
[environments](#)  
[documentation](#)

**Focus on**

[ALICE](#)  
[ATLAS](#)  
[CMS](#)  
[LHCb](#)  
[OPERA](#)  
[PHENIX](#)  
[Data Science](#)



# Escolhendo o dataset correto

open data CERN Help

SingleMuon/Run2015D-16Dec2015-v1/MINIAOD Search

3 result(s) found Sort by Best match

Current parameters Clear all

Availability include on-demand datasets

Type

- Dataset (3)
- Collision (2)
- Derived (1)

Experiment CMS (3)

Year 2015 (3)

**/SingleMuon/Run2015D-16Dec2015-v1/MINIAOD**

SingleMuon primary dataset in MINIAOD format from RunD of 2015. Run period from run number 256630 to 260627. The list of validated runs, which must be applied to all analyses, either with the full v...

Dataset Collision CMS

**/SingleMuon/Run2015D-16Dec2015-v1/AOD**

SingleMuon primary dataset in AOD format from RunD of 2015. Run period from run number 256630 to 260627. The list of validated runs, which must be applied to all analyses, either with the full valid...

Dataset Collision CMS

**Run2015D\_SingleMuon dataset in reduced NanoAOD-like format**

This dataset contains information extracted from different physics objects from the 2015 /SingleMuon/Run2015D-16Dec2015-v1/MINIAOD dataset, readable with bare ROOT or other ROOT-compatible software...

( 1 ) 10 results per page

# Escolhendo o dataset correto

opendata  
CERN

Help

SingleMuon/Run2015D-16Dec2015-v1/MINIAOD

Search

3 result(s) found

Sort by Best match

Current parameters Clear all

Availability

include on-demand datasets

Type

Dataset (3)

Collision (2)

Derived (1)

Experiment

CMS (3)

Year

2015 (3)

**/SingleMuon/Run2015D-16Dec2015-v1/MINIAOD**

SingleMuon primary dataset in MINIAOD format from RunD of 2015. Run period from run number 256630 to 260627. The list of validated runs, which must be applied to all analyses, either with the full v...

**Dataset** **Collision** **CMS**

**/SingleMuon/Run2015D-16Dec2015-v1/AOD**

SingleMuon primary dataset in AOD format from RunD of 2015. Run period from run number 256630 to 260627. The list of validated runs, which must be applied to all analyses, either with the full valid...

**Dataset** **Collision** **CMS**

**Run2015D\_SingleMuon dataset in reduced NanoAOD-like format**

This dataset contains information extracted from different physics objects from the 2015 /SingleMuon/Run2015D-16Dec2015-v1/MINIAOD dataset, readable with bare ROOT or other ROOT-compatible software...

( 1 )

10▼ results per page

# Dataset SingleMuon/Run2015D-16Dec2015-v1/MINIAOD

- No final da página, vão estar disponíveis as listas de endereços dos datasets

## File Indexes

Filename	Size		
CMS_Run2015D_SingleMuon_MINIAOD_16Dec2015-v1_10000_file_index.txt	128.8 KiB	<a href="#">List files</a>	<a href="#">Download Index</a>
CMS_Run2015D_SingleMuon_MINIAOD_16Dec2015-v1_10001_file_index.txt	101.6 KiB	<a href="#">List files</a>	<a href="#">Download Index</a>
CMS_Run2015D_SingleMuon_MINIAOD_16Dec2015-v1_20000_file_index.txt	792.0 bytes	<a href="#">List files</a>	<a href="#">Download Index</a>
CMS_Run2015D_SingleMuon_MINIAOD_16Dec2015-v1_60000_file_index.txt	15.9 KiB	<a href="#">List files</a>	<a href="#">Download Index</a>

# Dataset SingleMuon/Run2015D-16Dec2015-v1/MINIAOD

- No final da página, vão estar disponíveis as listas de endereços dos datasets

## File Indexes

Filename	Size	
CMS_Run2015D_SingleMuon_MINIAOD_16Dec2015-v1_10000_file_index.txt	128.8 KiB	<a href="#">List files</a> <a href="#">Download index</a>
CMS_Run2015D_SingleMuon_MINIAOD_16Dec2015-v1_10001_file_index.txt	101.6 KiB	<a href="#">List files</a> <a href="#">Download index</a>
CMS_Run2015D_SingleMuon_MINIAOD_16Dec2015-v1_20000_file_index.txt	792.0 bytes	<a href="#">List files</a> <a href="#">Download index</a>
CMS_Run2015D_SingleMuon_MINIAOD_16Dec2015-v1_60000_file_index.txt	15.9 KiB	<a href="#">List files</a> <a href="#">Download index</a>

# Dataset SingleMuon/Run2015D-16Dec2015-v1/MINIAOD

- No final da página, vão estar disponíveis as listas de endereços dos datasets

## File Indexes

Filename	Size	
CMS_Run2015D_SingleMuon_MINIAOD_16Dec2015-v1_10000_file_index.txt	128.8 KiB	<a href="#">List files</a> <a href="#">Download index</a>
CMS_Run2015D_SingleMuon_MINIAOD_16Dec2015-v1_10001_file_index.txt	101.6 KiB	<a href="#">List files</a> <a href="#">Download index</a>
CMS_Run2015D_SingleMuon_MINIAOD_16Dec2015-v1_20000_file_index.txt	792.0 bytes	<a href="#">List files</a> <a href="#">Download index</a>
CMS_Run2015D_SingleMuon_MINIAOD_16Dec2015-v1_60000_file_index.txt	15.9 KiB	<a href="#">List files</a> <a href="#">Download index</a>

```
root://eospublic.cern.ch//eos/opendata/cms/Run2015D/SingleMuon/MINIAOD/16Dec2015-v1/10000/00006301-CAA8-E511-AD39-549F35AD8BC9.root
root://eospublic.cern.ch//eos/opendata/cms/Run2015D/SingleMuon/MINIAOD/16Dec2015-v1/10000/0034202D-A3A8-E511-B9C-00259073E3DA.root
root://eospublic.cern.ch//eos/opendata/cms/Run2015D/SingleMuon/MINIAOD/16Dec2015-v1/10000/0043758E-ECA8-E511-B849-002618FDA287.root
root://eospublic.cern.ch//eos/opendata/cms/Run2015D/SingleMuon/MINIAOD/16Dec2015-v1/10000/004C08BC-C8A8-E511-943C-00266CFAEE0.root
root://eospublic.cern.ch//eos/opendata/cms/Run2015D/SingleMuon/MINIAOD/16Dec2015-v1/10000/005416D9-E0A8-E511-8AA1-0CC47A4C8E46.root
root://eospublic.cern.ch//eos/opendata/cms/Run2015D/SingleMuon/MINIAOD/16Dec2015-v1/10000/00649463-BE8-E511-9266-0CC47A4D769C.root
root://eospublic.cern.ch//eos/opendata/cms/Run2015D/SingleMuon/MINIAOD/16Dec2015-v1/10000/00AB7295-C6A8-E511-AEAF-0CC47A4C8E8A.root
root://eospublic.cern.ch//eos/opendata/cms/Run2015D/SingleMuon/MINIAOD/16Dec2015-v1/10000/00E42A7E-BAA8-E511-B23B-0CC47A4C8E86.root
root://eospublic.cern.ch//eos/opendata/cms/Run2015D/SingleMuon/MINIAOD/16Dec2015-v1/10000/00F3E56E-BE8-E511-A6A2-3417EBE722FA.root
root://eospublic.cern.ch//eos/opendata/cms/Run2015D/SingleMuon/MINIAOD/16Dec2015-v1/10000/02004059-D9A8-E511-8790-0CC47A78A440.root
root://eospublic.cern.ch//eos/opendata/cms/Run2015D/SingleMuon/MINIAOD/16Dec2015-v1/10000/0205A0BE-C8A8-E511-8F5B-00266CFADEC0.root
root://eospublic.cern.ch//eos/opendata/cms/Run2015D/SingleMuon/MINIAOD/16Dec2015-v1/10000/02120061-B9A8-E511-AFA7-00266CFFCD50.root
root://eospublic.cern.ch//eos/opendata/cms/Run2015D/SingleMuon/MINIAOD/16Dec2015-v1/10000/0213C018-EBA8-E511-8E2A-0025905B858E.root
root://eospublic.cern.ch//eos/opendata/cms/Run2015D/SingleMuon/MINIAOD/16Dec2015-v1/10000/0217776C-E1A8-E511-B061-0025904C5DDA.root
root://eospublic.cern.ch//eos/opendata/cms/Run2015D/SingleMuon/MINIAOD/16Dec2015-v1/10000/02720220-A2A8-E511-911C-003048FFCB6A.root
```

# Problema do armazenamento

type	dataset	variation	CODP name	xsec [pb]	size (TB)	Rec ID
colisions data	data	nominal	/SingleMuon/Run2015D-16Dec2015-v1/MINIAOD		1.4	24119
			/SingleElectron/Run2015D-08Jun2016-v1/MINIAOD			
signal	ttbar	nominal	TT_TuneCUETP8M1_13TeV-powheg-pythia8	831.76	3.4	19980
			TT_TuneCUETP8M1_13TeV-powheg-scaledown-pythia8			
			TT_TuneCUETP8M1_13TeV-powheg-scaleup-pythia8			
		ME_var	TTJets_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	1.3	1.3	19949
			TT_TuneEE5C_13TeV-powheg-herwigpp			
backgrounds	single_top_t_chan	nominal	ST_t-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1	44.33	0.52	19397
	single_atop_t_chan	nominal	ST_t-channel_antitop_4f_leptonDecays_13TeV-powheg-pythia8_TuneCUETP8M1			
	single_top_tW	nominal	ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.6	0.03	19419
		ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1				
wjets	nominal		WJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	61526.7	3.8	20548

# Problema do armazenamento

type	dataset	variation	CODP name	xsec [pb]	size (TB)	Rec ID
colisions data	data	nominal	/SingleMuon/Run2015D-16Dec2015-v1/MINIAOD		1.4	24119
			/SingleElectron/Run2015D-08Jun2016-v1/MINIAOD		2.6	24120
signal	ttbar	nominal	TT_TuneCUETP8M1_13TeV-powheg-pythia8	831.76	3.4	19980
		scaledown	TT_TuneCUETP8M1_13TeV-powheg-scaledown-pythia8		1.4	19983
		scaleup	TT_TuneCUETP8M1_13TeV-powheg-scaleup-pythia8		1.3	19985
		ME_var	TTJets_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8		1.3	19949
		PS_var	TT_TuneEE5C_13TeV-powheg-herwigpp		0.8	19999
backgrounds	single_top_t_chan	nominal	ST_t-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1	44.33	0.52	19397
	single_atop_t_chan	nominal	ST_t-channel_antitop_4f_leptonDecays_13TeV-powheg-pythia8_TuneCUETP8M1		26.38	0.042
	single_top_tW	nominal	ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.6	0.03	19419
			ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1		35.6	0.03
wjets	nominal		WJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	61526.7	3.8	20548

Mais de 16.5 TB!

# Dados filtrados

- A partir daqui vamos trabalhar com dados filtrados;
- Esse dados são provenientes das listas apresentadas, mas foi aplicado a eles filtros de *triggers*:
  - **HLT\_Ele22\_eta2p1\_WPLoose\_Gsf**: Eventos com pelo menos um elétron identificado de forma, com momento maior que 22 GeV e  $|\eta| < 2,1$ ;
  - **HLT\_IsoMu20\_v**: Eventos com pelo menos um muon considerado isolado com momento maior que 20 GeV;
  - **HLT\_IsoTkMu20\_v**: Eventos com pelo menos um muon, identificado com informação do detector de traços, que tenha momento maior que 20 GeV
- Com esses filtros foram reduzidos de 16,5 TB para 60 GB!
- Esse dados filtrados já estão no HepGrid Brazil! Então caso queira acessar eles, entre na pasta:

[/opendata/cms/derived-data/POET/23-Jul-22\\*](https://opendata.cmsderived-data.cern.ch/POET/23-Jul-22)

# Como acessamos os dados?

- Utilizando o *uproot*, fica fácil acessar os dados:

```
import uproot
import awkward as ak
import hist
from hist import Hist
from coffea.nanoevents import NanoEventsFactory, BaseSchema
```

Importando os pacotes que vamos usar

# Como acessamos os dados?

- Utilizando o *uproot*, fica fácil acessar os dados:

```
import uproot
import awkward as ak
import hist
from hist import Hist
from coffea.nanoevents import NanoEventsFactory, BaseSchema
```

Importando os pacotes que vamos usar

```
events = uproot.open(
    'root://eospublic.cern.ch//eos/opendata/cms/derived-data/POET/')
```

Utilizamos uproot para acessar/abrir o dataset

# Como acessamos os dados?

- Utilizando o *uproot*, fica fácil acessar os dados:

```
import uproot
import awkward as ak
import hist
from hist import Hist
from coffea.nanoevents import NanoEventsFactory, BaseSchema
```

Importando os pacotes que vamos usar

```
events = uproot.open(
    'root://eospublic.cern.ch//eos/opendata/cms/derived-data/POET/')
```

Utilizamos uproot para acessar/abrir o dataset

```
events.keys()
['events;1']
```

Podemos usar `.keys()` para ver os índices do dataset. Ele nos retorna `['events;1']`, indicando que podemos acessar mais informações através desse índice.

# Como acessamos os dados?

```
events['events'].keys()
```

Podemos visualizar o que há dentro do *events*, através do *.keys()*

# Como acessamos os dados?

```
events['events'].keys()
```

Podemos visualizar o que há dentro do *events*, através do *.keys()*

```
['numberelectron', 'nelectron_e', 'electron_e', 'nelectron_pt', 'electron_pt', 'nelectron_px', 'electron_px', 'nelectron_py', 'electron_py', 'nelectron_pz', 'electron_pz', 'nelectron_eta', 'electron_eta', 'nelectron_phi', 'electron_phi', 'nelectron_ch', 'electron_ch', 'nelectron_iso', 'electron_iso', 'nelectron_veto', 'electron_veto', 'nelectron_isLoose', 'electron_isLoose', 'nelectron_isMedium', 'electron_isMedium', 'nelectron_isTight', 'electron_isTight', 'nelectron_dx', 'electron_dx', 'nelectron_dz', 'electron_dz', 'nelectron_dxError', 'electron_dxError', 'nelectron_dzError', 'electron_dzError', 'nelectron_ismvaLoose', 'electron_ismvaLoose', 'nelectron_ismvaTight', 'electron_ismvaTight', 'nelectron_ip3d', 'electron_ip3d', 'nelectron_sip3d', 'electron_sip3d', 'numberfatjet', 'nfatjet_e', 'fatjet_e', 'nfatjet_pt', 'fatjet_pt', 'nfatjet_eta', 'fatjet_eta', 'nfatjet_phi', 'fatjet_phi', 'nfatjet_ch', 'fatjet_ch', 'nfatjet_mass', 'fatjet_mass', 'nfatjet_corrpt', 'fatjet_corrpt', 'nfatjet_corrptUp', 'fatjet_corrptUp', 'nfatjet_corrptDown', 'fatjet_corrptDown', 'nfatjet_corrptSmearUp', 'fatjet_corrptSmearUp', 'nfatjet_corrptSmearDown', 'fatjet_corrptSmearDown', 'nfatjet_cormass', 'fatjet_cormass', 'nfatjet corre', 'fatjet corre', 'nfatjet corrpx', 'fatjet corrpx', 'nfatjet corrpy', 'fatjet corrpy', 'nfatjet corrpz', 'fatjet corrpz', 'nfatjet_prunedmass', 'fatjet_prunedmass', 'nfatjet_softdropmass', 'fatjet_softdropmass', 'nfatjet_tau1', 'fatjet_tau1', 'nfatjet_tau2', 'fatjet_tau2', 'nfatjet_tau3', 'fatjet_tau3', 'nfatjet_subjet1btag', 'fatjet_subjet1btag', 'nfatjet_subjet2btag', 'fatjet_subjet2btag', 'nfatjet_subjet1hflav', 'fatjet_subjet1hflav', 'nfatjet_subjet2hflav', 'fatjet_subjet2hflav', 'numberjet', 'njet_e', 'jet_e', 'njet_pt', 'jet_pt', 'njet_eta', 'jet_eta', 'njet_phi', 'jet_phi', 'njet_ch', 'jet_ch', 'njet_mass', 'jet_mass', 'njet_btag', 'jet_btag', 'njet_hflav', 'jet_hflav', 'njet_corrpt', 'jet_corrpt', 'njet_corrptUp', 'jet_corrptUp', 'njet_corrptDown', 'jet_corrptDown', 'njet_corrptSmearUp', 'jet_corrptSmearUp', 'njet_corrptSmearDown', 'jet_corrptSmearDown', 'njet_cormass', 'jet_cormass', 'njet corre', 'jet corre', 'njet_corrpx', 'jet_corrpx', 'njet_corrpy', 'jet_corrpy', 'njet_corrpz', 'jet_corrpz', 'btag_Weight', 'btag_WeightUp', 'btag_WeightDn', 'met_e', 'met_pt', 'met_px', 'met_py', 'met_phi', 'met_significance', 'met_rawpt', 'met_rawphi', 'met rawe', 'numbermuon', 'nmuon_e', 'muon_e', 'nmuon_pt', 'muon_pt', 'nmuon_px', 'muon_px', 'nmuon_py', 'muon_py', 'nmuon_pz', 'muon_pz', 'nmuon_eta', 'muon_eta', 'nmuon_phi', 'muon_phi', 'nmuon_ch', 'muon_ch', 'nmuon_isLoose', 'muon_isLoose', 'nmuon_isMedium', 'muon_isMedium', 'nmuon_isTight', 'muon_isTight', 'nmuon_isSoft', 'muon_isSoft', 'nmuon_isHighPt', 'muon_isHighPt', 'nmuon_dx', 'muon_dx', 'nmuon_dz', 'muon_dz', 'nmuon_dxError', 'muon_dxError', 'nmuon_dzError', 'muon_dzError', 'nmuon_pfrelico03all', 'muon_pfrelico03all', 'nmuon_pfrelico04all', 'muon_pfrelico04all', 'nmuon_pfrelico04DBCorr', 'muon_pfrelico04DBCorr', 'nmuon_TkIso03', 'muon_TkIso03', 'nmuon_jetidx', 'muon_jetidx', 'nmuon_genpartidx', 'muon_genpartidx', 'nmuon_ip3d', 'muon_ip3d', 'nmuon_sip3d', 'muon_sip3d', 'numberphoton', 'nphoton_e', 'photon_e', 'nphoton_pt', 'photon_pt', 'nphoton_px', 'photon_px', 'nphoton_py', 'photon_py', 'nphoton_pz', 'photon_pz', 'nphoton_eta', 'photon_eta', 'nphoton_phi', 'photon_phi', 'nphoton_ch', 'photon_ch', 'nphoton_chIso', 'photon_chIso', 'nphoton_nhIso', 'photon_nhIso', 'nphoton_phIso', 'photon_phIso', 'nphoton_isLoose', 'photon_isLoose', 'nphoton_isMedium', 'photon_isMedium', 'nphoton_isTight', 'photon_isTight', 'nPV_chi2', 'PV_chi2', 'nPV_ndof', 'PV_ndof', 'PV_npvs', 'PV_npvsGood', 'nP_V_x', 'PV_x', 'nP_V_y', 'PV_y', 'nP_V_z', 'PV_z', 'trig_Ele22_eta2p1_WPloose_Gsf', 'trig_IsoMu20', 'trig_IsoTkMu20', 'numbertau', 'ntau_e', 'tau_e', 'ntau_pt', 'tau_pt', 'ntau_px', 'tau_px', 'ntau_py', 'tau_py', 'ntau_pz', 'tau_pz', 'ntau_eta', 'tau_eta', 'ntau_phi', 'tau_phi', 'ntau_ch', 'tau_ch', 'ntau_mass', 'tau_mass', 'ntau_decaymode', 'tau_decaymode', 'ntau_iddecaymode', 'tau_iddecaymode', 'ntau_idisoraw', 'tau_idisoraw', 'ntau_idisovloose', 'tau_idisovloose', 'ntau_idisoloose', 'tau_idisoloose', 'ntau_idisomedium', 'tau_idisomedium', 'ntau_idisotight', 'tau_idisotight', 'ntau_idantieletight', 'tau_idantieletight', 'ntau_idantimutight', 'tau_idantimutight']
```

# Utilizando o AGCSchema

```
from agc_schema import AGCSchema
agc_events = NanoEventsFactory.from_root('root://eospUBLIC.cern.ch//eos
                                         schemaClass=AGCSchema,
                                         treePath='events').events()
```

Importando e  
aplicando a função  
do pacote

# Utilizando o AGCSchema

```
from agc_schema import AGCSchema
agc_events = NanoEventsFactory.from_root('root://eospublic.cern.ch//eos
                                         schemaclass=AGCSchema,
                                         treepath='events').events()
```

Importando e  
aplicando a função  
do pacote

```
agc_events.fields
```

```
['jet',
 'photon',
 'muon',
 'fatjet',
 'tau',
 'electron',
 'met',
 'trig',
 'btag',
 'PV']
```

Utilizando a função fields, podemos acessar de forma organizada os diferentes objetos físicos que são pertinentes a análise.

# Acessando informações dos múons

Para acessar informações dos múons

```
agc_events.muon.fields|
```

# Acessando informações dos múons

Para acessar informações dos múons

```
agc_events.muon.fields|
```

```
['pt', 'px', 'py', 'pz', 'eta', 'phi', 'ch', 'isLoose', 'isMedium', 'isTight', 'isSoft', 'isHighPt', 'dxy', 'dz', 'dxyError', 'dzError', 'pfreliso03all',  
'pfreliso04all', 'pfreliso04DBCorr', 'TkIso03', 'jetidx', 'genpartidx', 'ip3d', 'sip3d', 'energy']
```

# Acessando informações dos múons

Para acessar informações dos múons

```
agc_events.muon.fields|
```

```
['pt', 'px', 'py', 'pz', 'eta', 'phi', 'ch', 'isLoose', 'isMedium', 'isTight', 'isSoft', 'isHighPt', 'dxy', 'dz', 'dxyError', 'dzError', 'pfreliso03all',  
'pfreliso04all', 'pfreliso04DBCorr', 'TkIso03', 'jetidx', 'genpartidx', 'ip3d', 'sip3d', 'energy']
```

```
import matplotlib.pyplot as plt  
# Flatten the jagged array before we histogram it  
fig = plt.figure(figsize=(6,4))  
  
values = ak.num(agc_events.muon, axis=1)  
  
plt.hist(values,bins=10,range=(0,10), label='Todos os múons')  
  
plt.xlabel(f'Número de múons', fontsize=14)  
plt.ylabel(f'Eventos', fontsize=14)  
  
plt.legend()|
```

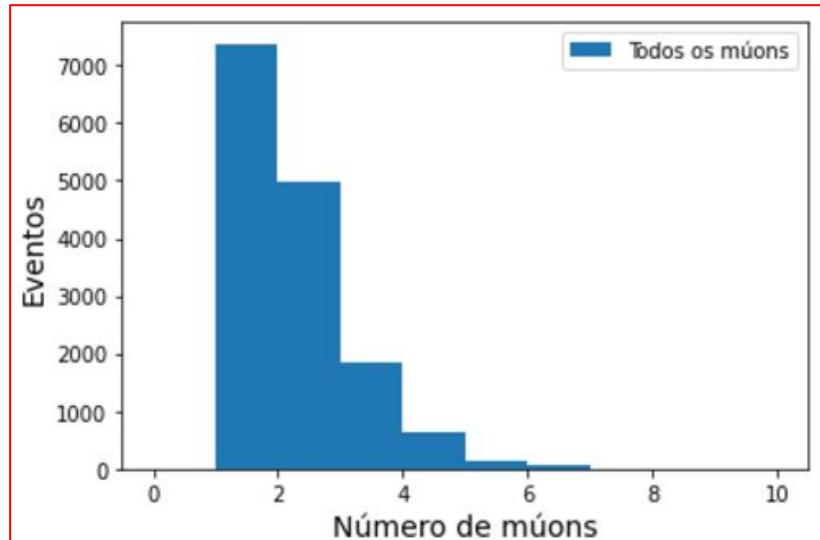
# Acessando informações dos múons

Para acessar informações dos múons

```
agc_events.muon.fields|
```

```
['pt', 'px', 'py', 'pz', 'eta', 'phi', 'ch', 'isLoose', 'isMedium', 'isTight', 'isSoft', 'isHighPt', 'dxy', 'dz', 'dxyError', 'dzError', 'pfreliso03all',  
'pfreliso04all', 'pfreliso04DBCorr', 'TkIso03', 'jetidx', 'genpartidx', 'ip3d', 'sip3d', 'energy']
```

```
import matplotlib.pyplot as plt  
# Flatten the jagged array before we histogram it  
fig = plt.figure(figsize=(6,4))  
  
values = ak.num(agc_events.muon, axis=1)  
  
plt.hist(values,bins=10,range=(0,10), label='Todos os múons')  
  
plt.xlabel(f'Número de múons', fontsize=14)  
plt.ylabel(f'Eventos', fontsize=14)  
  
plt.legend()|
```



# Outros exemplos de distribuições

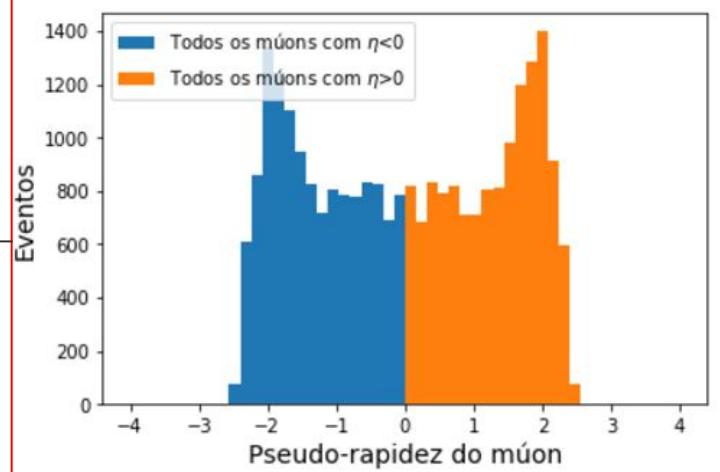
```
import matplotlib.pyplot as plt
# Flatten the jagged array before we histogram it
fig = plt.figure(figsize=(6,4))

values = ak.flatten(agc_events.muon.eta)

plt.hist(values[values<0],bins=50,range=(-4,4), label='Todos os m\xf3tons com $\eta$<0')
plt.hist(values[values>0],bins=50,range=(-4,4), label=r'Todos os m\xf3tons com $\eta$>0')

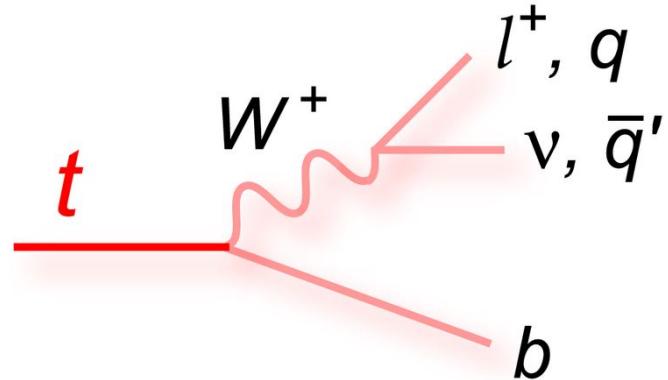
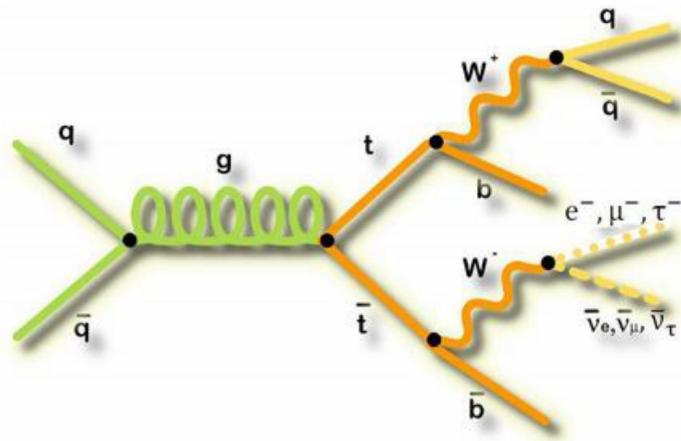
plt.xlabel(f'Pseudo-rapidez do m\xf3ton',fontsize=14)
plt.ylabel(f'Eventos',fontsize=14)

plt.legend()
```



# Amostra de Sinal X Background

- O **sinal** representa o processo físico que queremos estudar, nesse estudo é o ttbar (slide 19). Identificamos o processo através das partículas presentes no seu estado final.
- Diferentes processos físicos podem apresentar estados finais similares! Comprometendo a **pureza** dos nossos dados!
- Esse processos similares são chamados de **background** e nosso objetivo é reduzir sua contribuição o máximo possível.



# Amostra de Sinal X Background

- O **sinal** representa o processo físico que queremos estudar, nesse estudo é o ttbar (slide 19). Identificamos o processo através das partículas presentes no seu estado final.
- Diferentes processos físicos podem apresentar estados finais similares! Comprometendo a **pureza** dos nossos dados!
- Esse processos similares são chamados de **background** e nosso objetivo é reduzir sua contribuição o máximo possível.



Uma analogia é que estamos jogando **cara-a-cara**, mas com processos físicos! Então agora precisamos aplicar seleções visando manter somente o sinal!

# Corte de qualidade dos objetos

- Antes que sejam aplicados cortes visando manter o **sinal** e eliminar o máximo de **background**, faz-se necessário aplicar cortes de qualidades nos objetos físicos presentes no estado final do nosso sinal (elétrons, múon e jatos).

# Corte de qualidade dos objetos

- Antes que sejam aplicados cortes visando manter o **sinal** e eliminar o máximo de **background**, faz-se necessário aplicar cortes de qualidades nos objetos físicos presentes no estado final do nosso sinal (elétrons, múon e jatos).



Caso eu queira selecionar somente pessoas altas, devo especificar o valor da altura! Algo como:

- Pessoas maiores de 1,90 cm

# Corte de qualidade dos objetos

- No caso dos objetos físicos os cortes são feitos em atributos diferentes:

Múon	Elétron	Jatos
<ul style="list-style-type: none"><li>• <math>p_T &gt; 30 \text{ GeV}</math></li><li>• <math> \eta  &lt; 2,1</math></li><li>• Identificação de rigor alto</li><li>• <math>R_{\text{eliso}} &lt; 0,15</math></li><li>• <math>\text{SIP3D} &lt; 4</math></li></ul>	<ul style="list-style-type: none"><li>• <math>p_T &gt; 30 \text{ GeV}</math></li><li>• <math> \eta  &lt; 2,1</math></li><li>• Identificação de rigor alto</li><li>• <math>\text{SIP3D} &lt; 4</math></li></ul>	<ul style="list-style-type: none"><li>• <math>p_T &gt; 30 \text{ GeV}</math></li><li>• <math> \eta  &lt; 2,4</math></li><li>• Jatos com discriminador de b-tag maior que 0,8 passam a ser considerados jatos de bottom</li></ul>

# Como são aplicados os cortes no código

```
selected_electrons = events.electron[  
    (events.electron.pt > 30) &  
    (abs(events.electron.eta)<2.1) &  
    (events.electron.isTight == True) &  
    (events.electron.sip3d < 4)  
]  
  
selected_muons = events.muon[  
    (events.muon.pt > 30) &  
    (abs(events.muon.eta)<2.1) &  
    (events.muon.isTight == True) &  
    (events.muon.sip3d < 4) &  
    (events.muon.pfRelIso04DBCorr < 0.15)  
]  
  
selected_jets = agc_events.jet[  
    (agc_events.jet.corrrt > 30) &  
    (abs(agc_events.jet.eta)<2.4)  
]
```

Nessa etapa estamos criando arrays com True e False dentro deles.

Usaremos essa informação mais pra frente na seleção de sinal!

# Seleção de sinal

Agora devemos aplicar cortes visando manter o máximo possível de sinal e excluir o máximo de *background*.

Os cortes adotados foram:

- Pelo menos um elétron ou um múon;
- Pelo menos quatro jatos;
- Pelo menos dois dos jatos devem ser selecionados, pelo critério do b-tag.

# Seleção de sinal

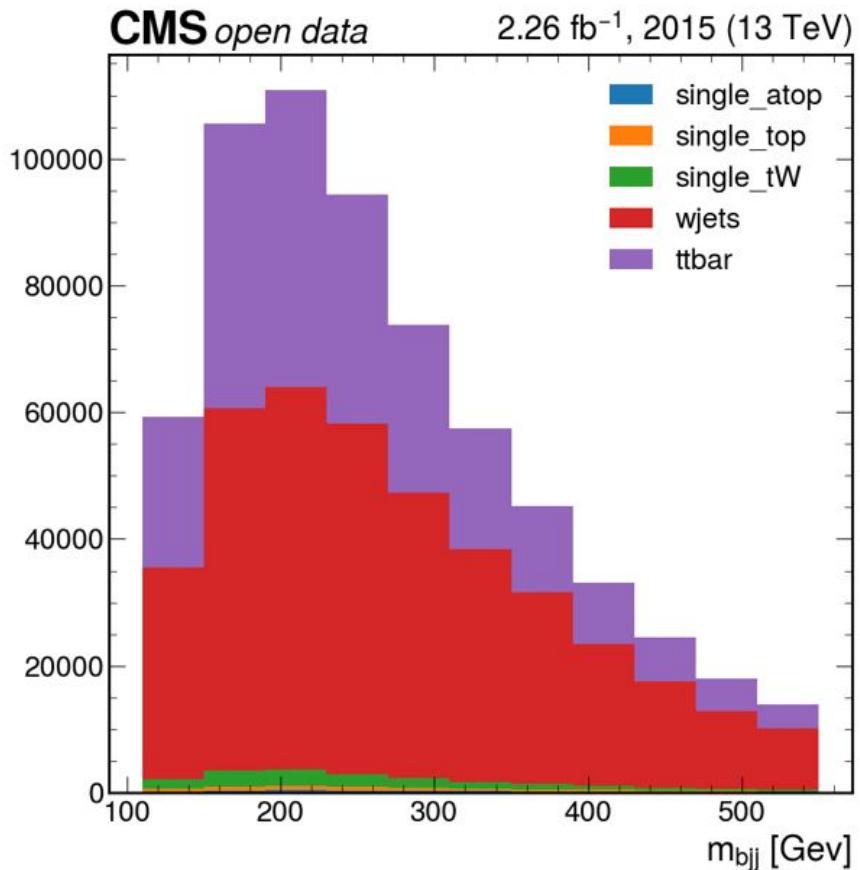
Agora devemos aplicar cortes visando manter o máximo possível de sinal e excluir o máximo de *background*.

Os cortes adotados foram:

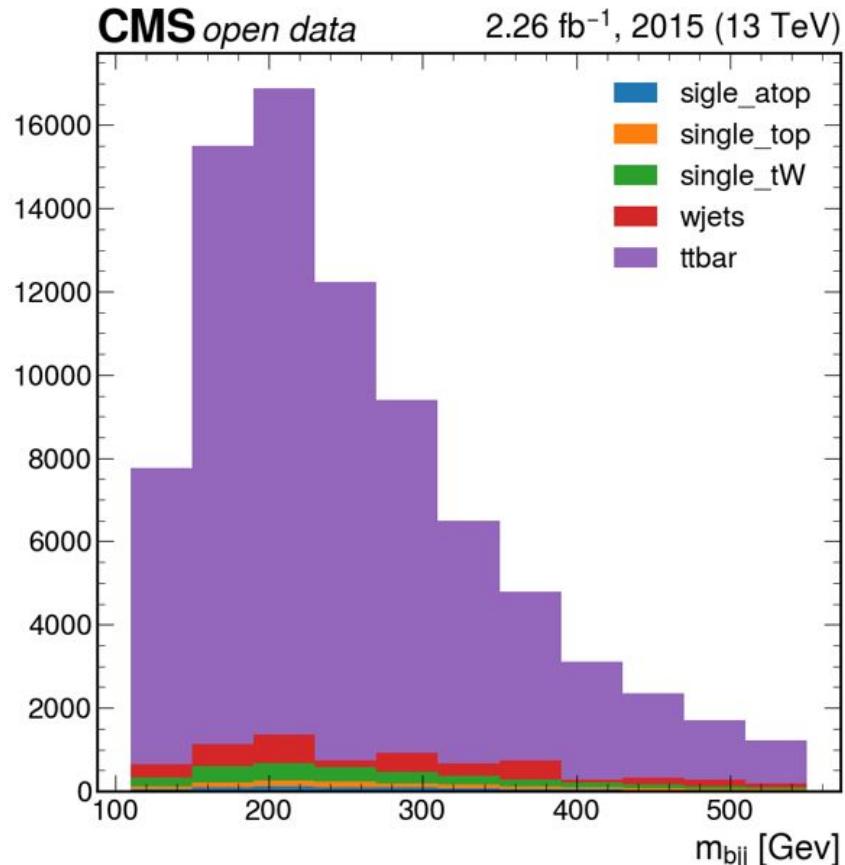
- Pelo menos um elétron ou um múon;
- Pelo menos quatro jatos;
- Pelo menos dois dos jatos devem ser selecionados, pelo critério do b-tag.

```
event_filters = ((ak.count(selected_electrons.pt, axis=1) + ak.count(selected_muons.pt, axis=1)) == 1)
event_filters = event_filters & (ak.count(selected_jets.corrpt, axis=1) >= 4)
B_TAG_THRESHOLD = 0.8
event_filters = event_filters & (ak.sum(selected_jets.btag >= B_TAG_THRESHOLD, axis=1) >= 1)
```

# Efeito da seleção



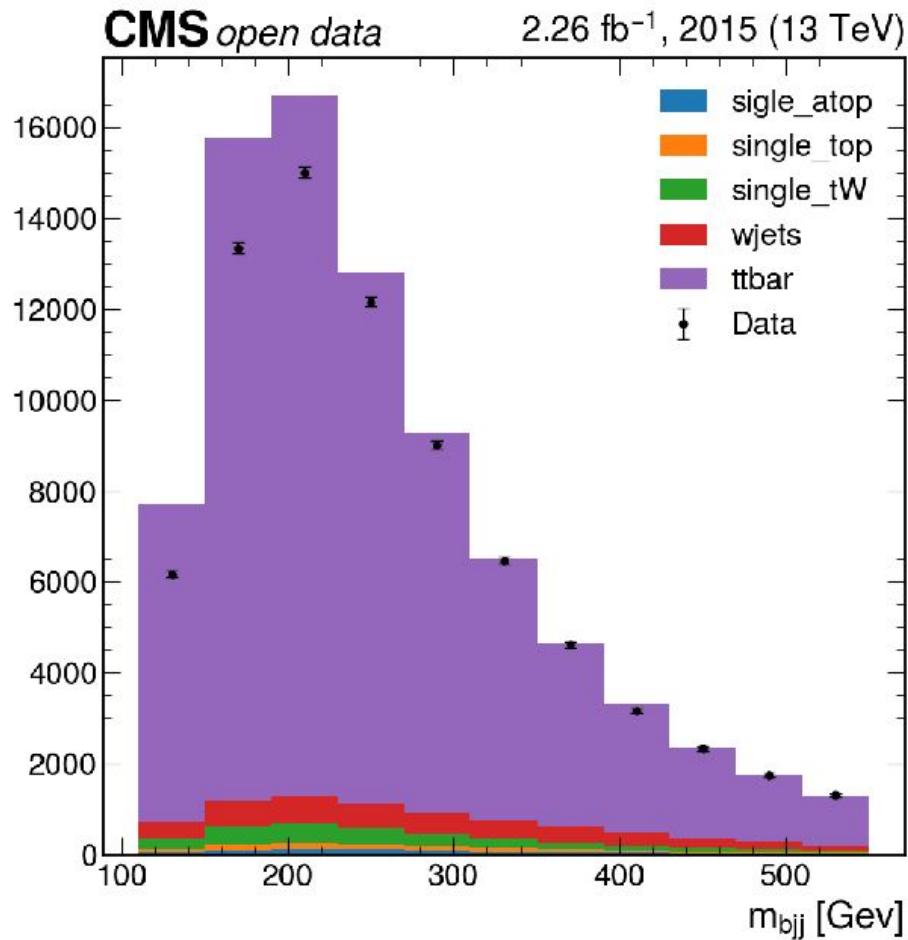
ANTES



DEPOIS

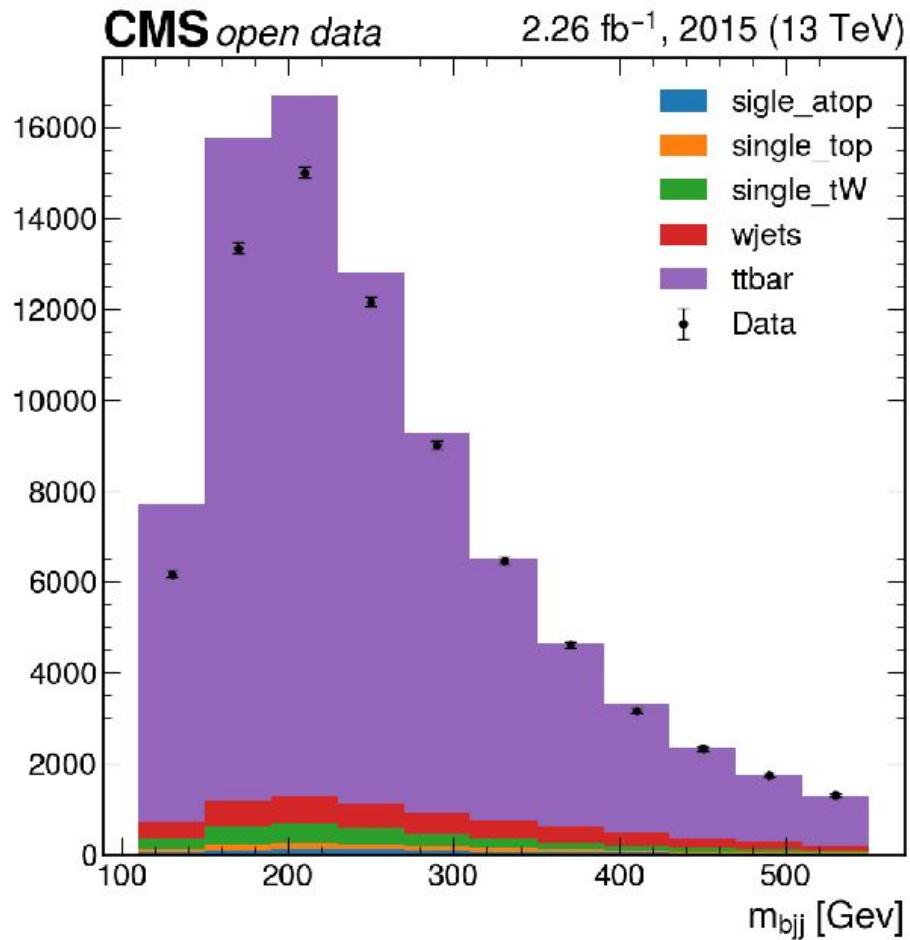
# Comparando a simulação com os dados das colisões

- No resultado obtido a simulação não representa os dados de forma fiel na região de baixa massa;



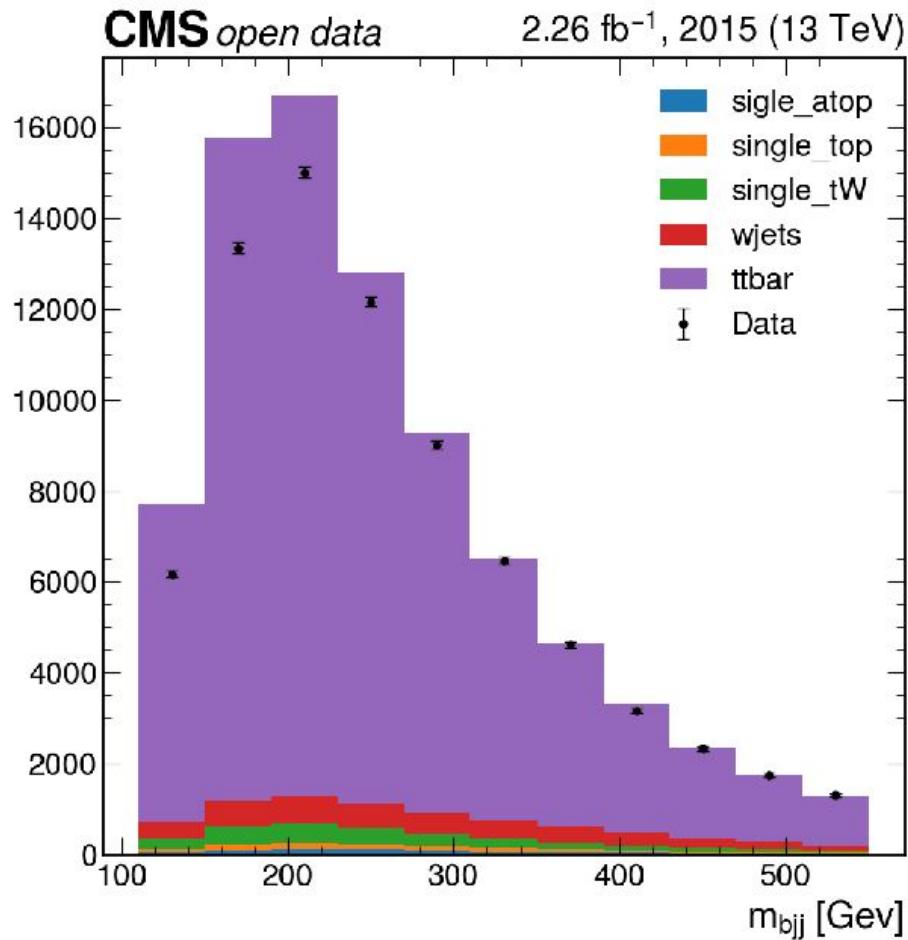
# Comparando a simulação com os dados das colisões

- No resultado obtido a simulação não representa os dados de forma fiel na região de baixa massa;
- Isso indica que, ou os modelos não funcionam, ou existem mais etapas a serem feitas;



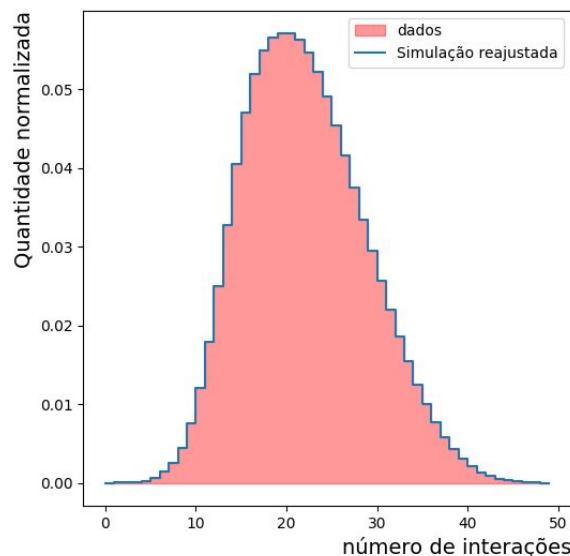
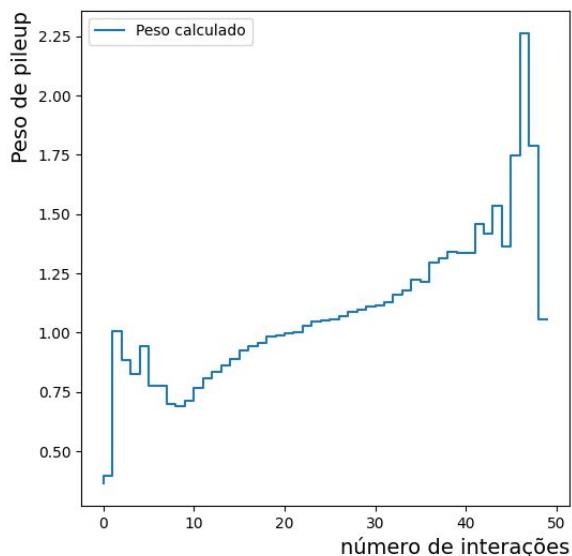
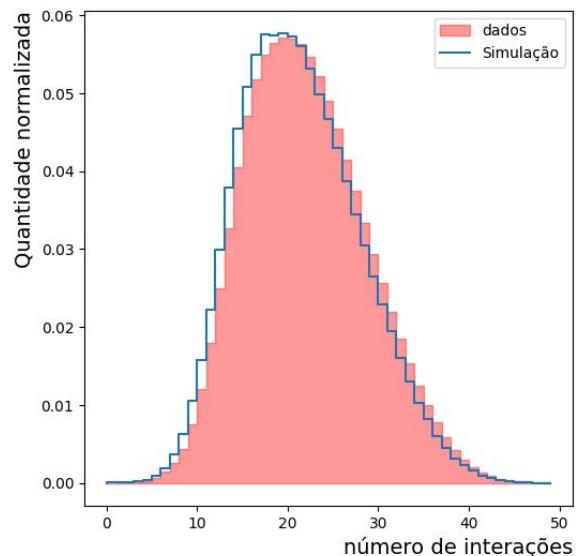
# Comparando a simulação com os dados das colisões

- No resultado obtido a simulação não representa os dados de forma fiel na região de baixa massa;
- Isso indica que, ou os modelos não funcionam, ou existem mais etapas a serem feitas;
- Na realidade, a próxima etapa da análise consiste em aplicar correções na simulação;
- Essas correções são baseadas em correções de energias, condições do detector e muito mais!



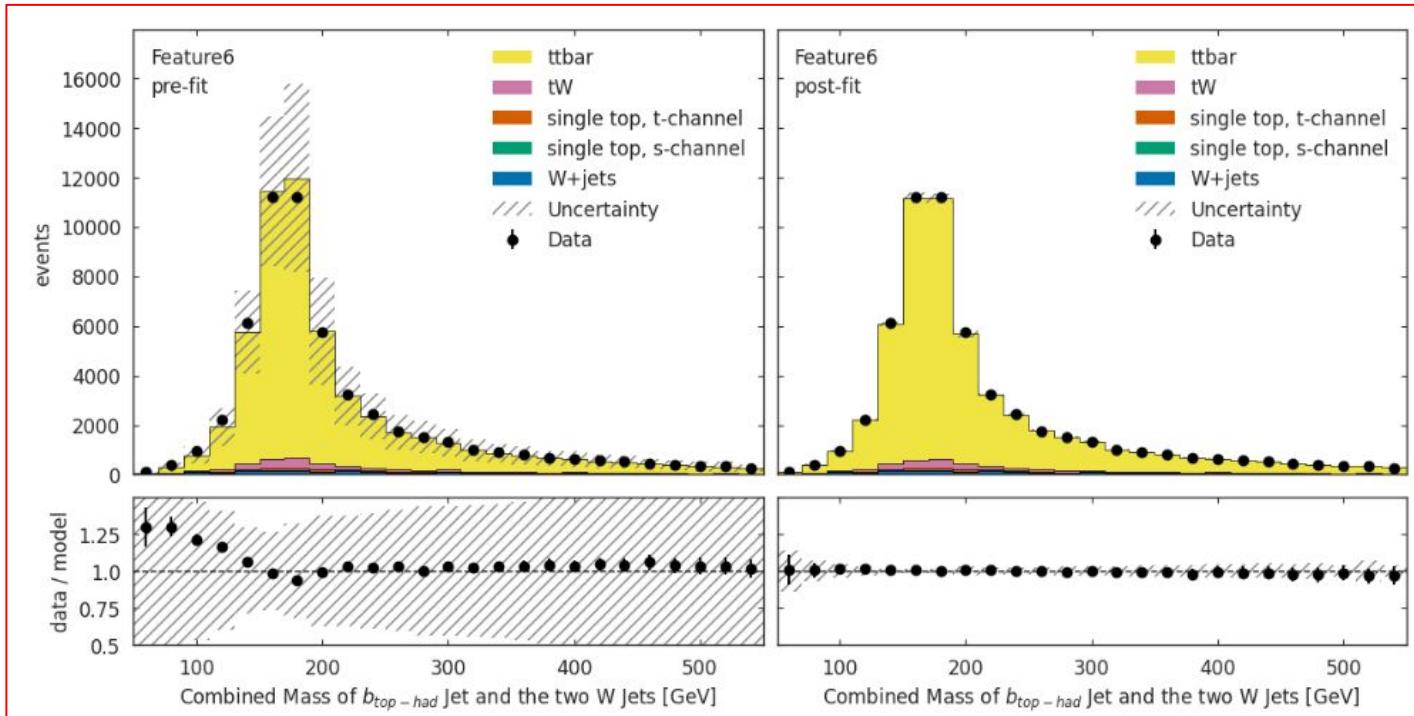
# Exemplo: Correção de pileup

*Pileup* é o fenômeno onde múltiplos eventos de colisões de partículas ocorrem simultaneamente ou quase simultaneamente em uma mesma área do detector.



# Resultado final

Após ser aplicado às correções na simulação, bem como fazer o tratamento do erro estatístico e sistemático\*, o resultado passa a ficar muito melhor!



\* O tratamento dos erros é algo que toma muito tempo pra se explicar, e requer um conhecimento avançado em estatística bem como conhecimento sobre o experimento que se está extraíndo os dados. Quaisquer dúvidas, sinta-se a vontade para enviar um e-mail ou perguntar.



**Agora é a sua vez!**

# Acessando o serviço do HepGrid Brazil

- Acesse o site: <https://jupyterhub.hepgrid.uerj.br/>



The image shows a 'Sign in' form with an orange header bar containing the text 'Sign in'. Below the header, there are two input fields: one for 'Username' and one for 'Password', both represented by empty rectangular boxes. At the bottom of the form is an orange button labeled 'Sign in' in white text. Below the button, there is a link in blue text that reads 'Quero trocar a senha'.

Sign in

Username:

Password:

Sign in

Quero trocar a senha

# Acessando o serviço do HepGrid Brazil

- Acesse o site: <https://jupyterhub.hepgrid.uerj.br/>

Sign in

Username:

Password:

**Sign in**

[Quero trocar a senha](#)

Sign in

Username:

Password:

**Sign in**

[Quero trocar a senha](#)



# Acessando o serviço do HepGrid Brazil

- Acesse o site: <https://jupyterhub.hepgrid.uerj.br/>
- Escolha a opção de imagem “Aula - Análise de Dados -2024.2”

## Server Options

Select an image:

Aula - Analise de Dados - 2024.2

Start

# Acessando o serviço do HepGrid Brazil

- Acesse o site: <https://jupyterhub.hepgrid.uerj.br/>
- Escolha a opção de imagem “Aula - Análise de Dados -2024.2”
- Clique em “Start”

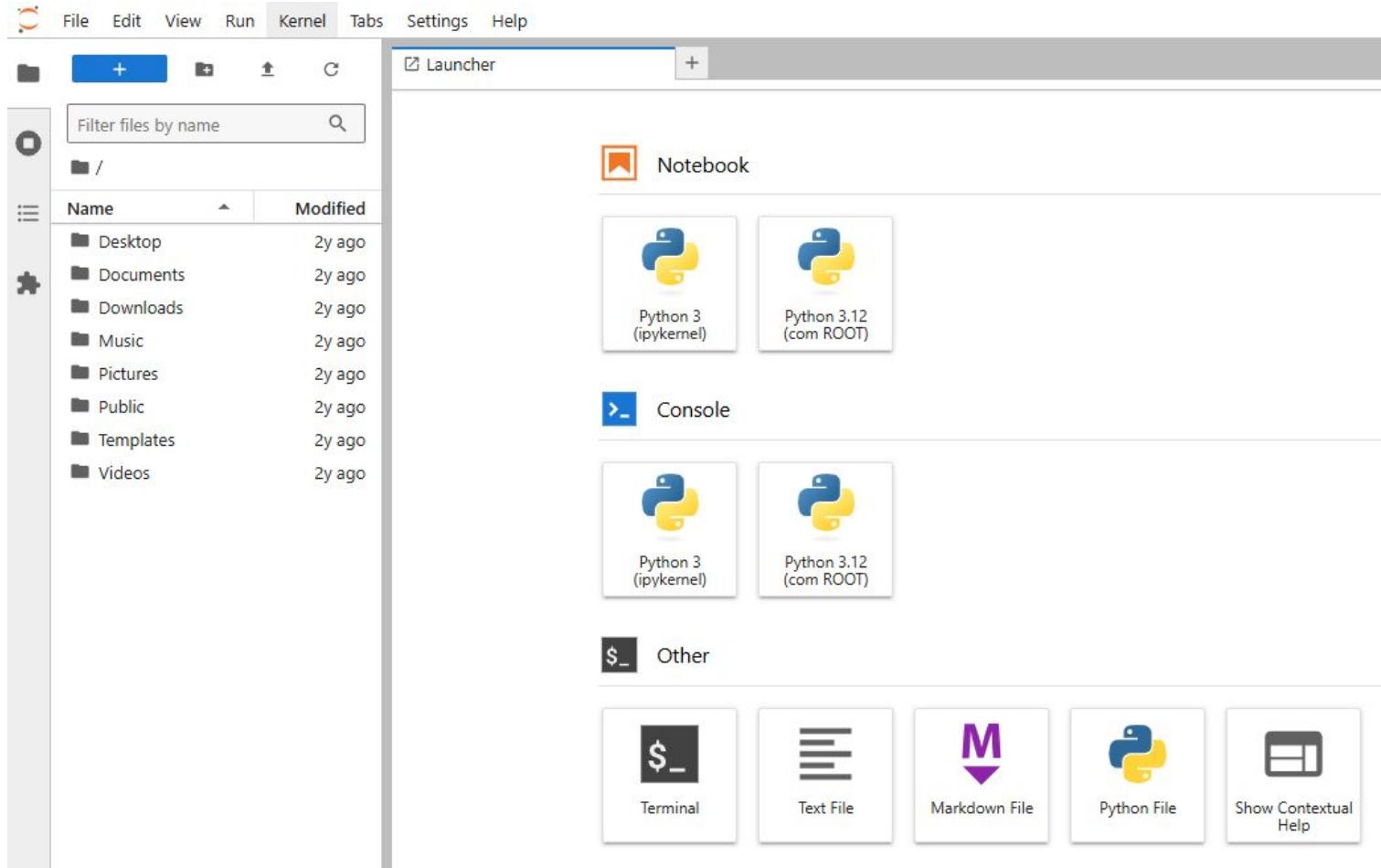
## Server Options

Select an image:

Aula - Analise de Dados - 2024.2

Start

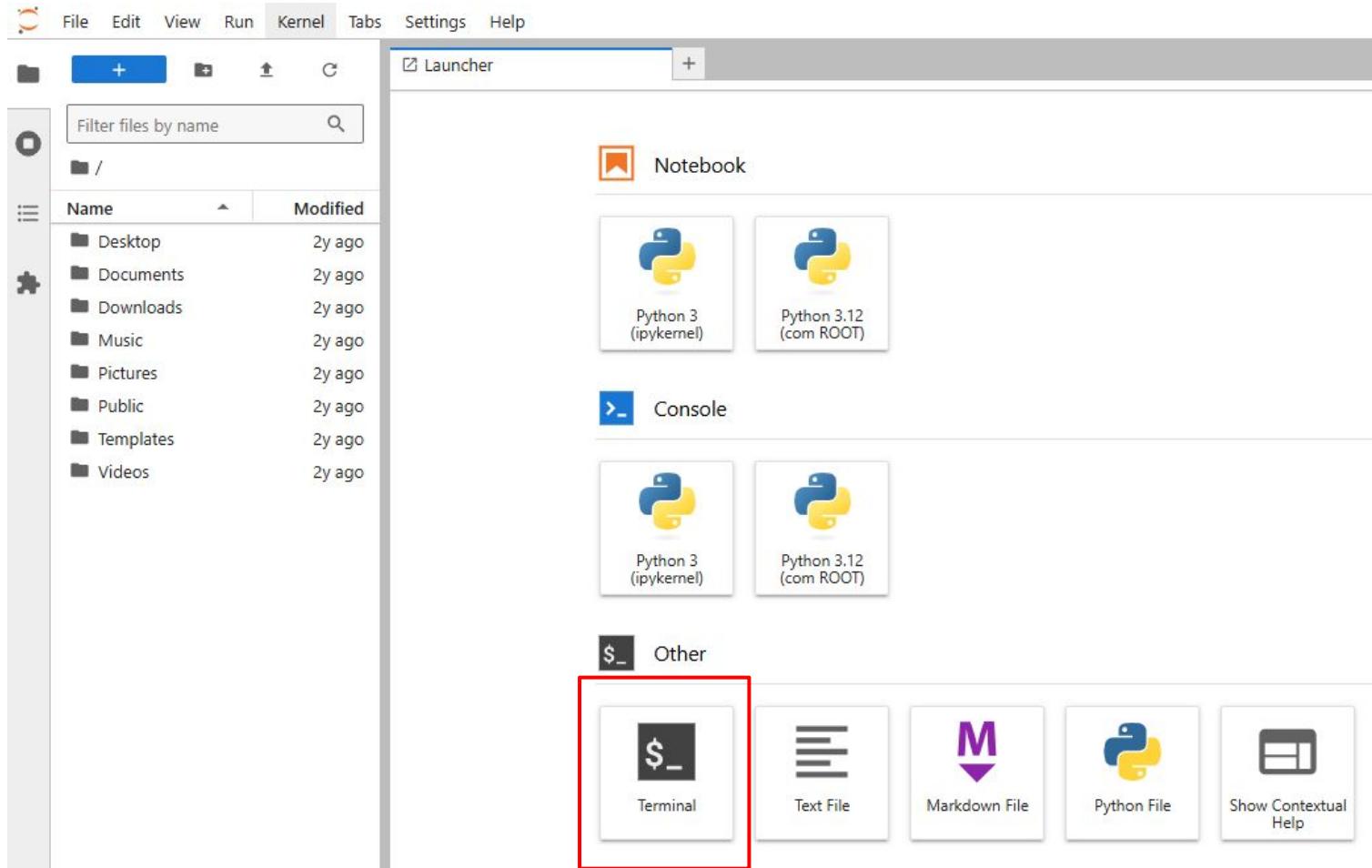
# Configurando o serviço do HepGrid Brazil



The screenshot shows a Jupyter Notebook interface with the following elements:

- File Edit View Run Kernel Tabs Settings Help**: The top navigation bar.
- Launcher**: A sidebar on the left containing:
  - A folder icon with a plus sign (+) to add new files.
  - A search bar labeled "Filter files by name".
  - A list of local directories:
    - Desktop (modified 2y ago)
    - Documents (modified 2y ago)
    - Downloads (modified 2y ago)
    - Music (modified 2y ago)
    - Pictures (modified 2y ago)
    - Public (modified 2y ago)
    - Templates (modified 2y ago)
    - Videos (modified 2y ago)
- Notebook**: A section containing two kernel icons:
  - Python 3 (ipykernel)
  - Python 3.12 (com ROOT)
- Console**: A section containing two kernel icons:
  - Python 3 (ipykernel)
  - Python 3.12 (com ROOT)
- Other**: A section containing icons for:
  - Terminal
  - Text File
  - Markdown File
  - Python File
  - Show Contextual Help

# Configurando o serviço do HepGrid Brazil



# Configurando o serviço do HepGrid Brazil

- No terminal digite:

```
pip install vector==0.8 hist mplhep==0.3.26 cabinetry coffea==0.7.21
```

- Instale o repositório git:

```
git clone https://github.com/cms-opendata-workshop/workshop2022-lesson-ttbarjetsanalysis-payload.git
```

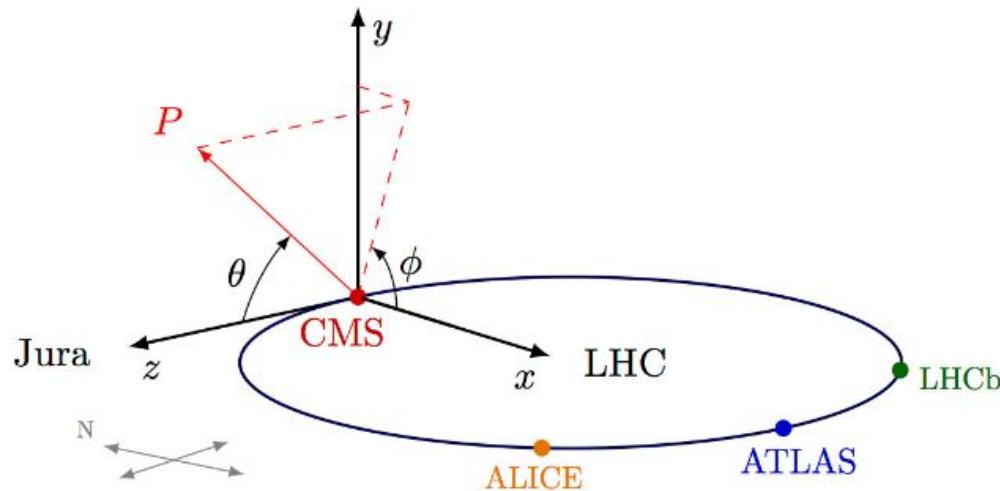
- Acesse os dados pelo diretório:

/opendata/cms/derived-data/POET/23-Jul-22\*

- Lembre-se de como acessar os dados! (slide 60-62)
  - Qualquer dúvida, pode perguntar!

# BACKUP

# Sistema de coordenadas do CMS e variáveis cinemáticas

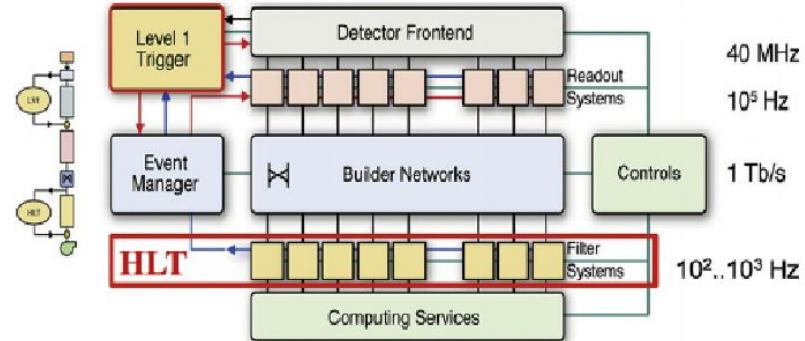


Grandeza	Equação
Momentum Transverso ( $p_T$ )	$\vec{p_T}^2 = \vec{p_x}^2 + \vec{p_y}^2$
Rapidez ( $y$ )	$y = \frac{1}{2} \ln\left(\frac{E+p_z}{E-p_z}\right)$
Pseudorapidez ( $\eta$ )	$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]$
Ângulo azimutal ( $\phi$ )	$\phi = \arcsin\left(\frac{p_y}{p_T}\right)$

# Sistema de triggers do CMS

- As colisões ocorrem a uma taxa de 40 MHz, tornando impossível armazenar todos os eventos. Para selecionar eventos relevantes são aplicados filtros (*triggers*) [Collaboration, 2008], sendo eles o L1 e o HLT.
- O L1 é um trigger de qualidade, enquanto que o HLT são *triggers* vinculados aos interesses das análises feitas pela colaboração CMS. Dos 40 milhões de eventos, produzidos a cada segundo, cerca de apenas 1000 são salvos.
- Após passar pelos filtros os dados são armazenados em Centros de processamento de Dados, onde neles estará disponível informações sobre que *triggers* do HLT foi ativado.

Esquema de gatilhos do CMS



Fonte: Jeitler, 2013