TWiki > AtlasComputing Web > AtlasComputing > ComputingTutorials > SoftwareTutorial > SoftwareTutorialFindingDatasets
(2019-06-05, JamesWalder2)

Finding Datasets

Overview

There are several tools in ATLAS to help you find datasets and information associated to these datasets. The three major tools that we will cover here are:

- Run Query ⬀
- COMA ⬀
- AMI ⬀

We will get you to play around with each of these tools, and try to get you to find answers to certain questions. Please ask us for help if you can't find the answers!

Run Query tool

The Run Query tool provides detailed run-level information for the real data. It is particularly useful for detector and operation tasks, when you are trying to understand the LHC and ATLAS detector conditions and settings. You can use the search query either on the web interface or using the command line. On the web interface there are links to more expert pages for exact details on certain topics.

Exercise: Get Data Summary plots

Question: Use the Run Query web interface ⬀ to get the Data Summary Total Luminosity plots (vs day) for the entire 2018 13 TeV pp dataset and the 2017 13 TeV pp dataset.

These type of figures are shown very regularly in a diverse set of meetings, so it's nice to understand how to find them!

Exercise: Find luminosity blocks with stable beams

Use the Run Query web interface ⬀ to find which luminosity blocks had stable beams in Run 325558.

You may need to investigate and correlate trigger rates with LHC conditions, such as when stable beams were declared.

Exercise: Find run details

Use the Run Query web interface ⬀ to help you find the answers to the following questions about run **335302**:
   a. What date was it taken?
   b. What period was it taken in?
   c. What was the prescale of the trigger `HLT_g60_loose` during luminosity blocks 257-268?

COMA

COMA provides quick information about data in containers/periods/run/luminosity-blocks, and is a tool that aggregates information from across many runs. It is a useful tool to get an overview of the configurations of streams and triggers.

To access the COMA web portal you need a valid grid certificate in your browser and you need to be a part of the ATLAS virtual organization (VO) (tutorial instructions). When trying to access the web portal accept any site warnings and click OK (not CANCEL)

Exercise: Understanding data periods

Question: Use the COMA web portal ⧉ to find out what were the data periods taken in 2017 for 13 TeV pp collisions? And what happened in period E (i.e. what defines period E)?

Exercise: Determine lowest unprescaled trigger

Question: Use the COMA web portal ⧉ to find the lowest pt HLT unprescaled single muon trigger throughout the 2017 pp run. What was the corresponding L1 input item?

AMI and pyAMI

AMI is the dataset metadata catalogue. It is a "mediating" application, which means it correlates information from many sources.

pyAMI is the python client for AMI. It is in the ATLAS release, and also available standalone on lxplus, or as a tar file for download to your laptop. pyAMI documentation explains how to use pyAMI.

Exercise: Making an account in AMI

To use AMI you must have a valid grid certificate and be registered with the ATLAS VO.

If this is your first time accessing AMI through the web interface you will need to setup an account. If you have not already done so, follow the instructions described in Software Tutorial: Registering with AMI.

Exercise: Finding DESDs in 2017

Here we will search the AMI web interface ⧉ to find out which types of DESDM data formats were produced during the 2017 pp collision data at 13 TeV. The DESDM format is a derived (D) format from the ESD with some modifications (M) on top. (D)ESD(M) formats contain a lot of information about each event, including all the calorimeter cell information. The xAODs only contain calorimeter cell information for cells in em clusters.

Question: Which types of DESDM's were produced during the 2017 pp collision data at 13 TeV?

Exercise: Finding data formats

Question: Using the AMI web interface ⧉, how many different formats are available for run 333853 (taken in 2017 with pp collisions at 13 TeV)? Can you see the number of events associated to the different data formats? Can you guess why they might be different?

Exercise: Using getMetadata.py

Here we will use pyAMI and a nice script to get the metadata associated to MC files and dump that information to either your screen or a text file.

Go to this webpage:
https://twiki.cern.ch/twiki/bin/view/AtlasProtected/AnalysisMetadata#Sample_CrossSections_kFactors_Fi

And follow the instructions under the major heading "Sample CrossSections, kFactors, Filter Efficiencies", which includes the sub-headings "Exploring metadata with getMetadata" and "How you really should use getMetadata".

Getting help

If you need help about COMA or AMI you can email the following mailing list for help:
hn-atlas-physicsMetadata@cern.ch

Major updates:
-- LouiseHeelan - 2017-09-13

Responsible: LouiseHeelan
Last reviewed by: **Never reviewed**

I	Attachment	History	Action	Size	Date	Who	Comment
🖼	sumLumiByDay2016.png	r2 r1	manage	21.1 K	2018-01-18 - 15:15	LouiseHeelan	
🖼	sumLumiByDay2017.png	r1	manage	21.7 K	2018-01-18 - 15:01	LouiseHeelan	
🖼	sumLumiByDay2018.png	r1	manage	22.4 K	2019-06-05 - 12:13	JamesWalder2	2018 Luminosity file

Topic revision: r7 - 2019-06-05 - JamesWalder2