RDataFrame Tutorial

Suvankar Proposing Cristina Alexe as the 2nd facilitator

RDF introduction

- 1. Input data format and layout assume NanoAOD
- 2. Introduction to RDF
 - a. Basic principles
 - b. Syntax compare to traditional approach
 - c. Manipulation of the input data
 - d. Results histograms, graphs, cut-flow report, skimming/Snapshot
 - e. Handling of multiple data samples
- 3. Requirements
 - a. ROOT >= v6.26
 - b. Input will be SingleMuon DATA and DY MC
 - c. Contact DESY for copying them to local site

RDF exercise set 1

1. Exercises

- a. Basic exercise on writing a script to plot some histograms
- b. Exercise on Filter
 - i. inspecting the cut flow report
 - ii. using json filtering on input data Muon sample from 2022
- c. Custom columns
 - i. Simple operations on columns
 - ii. user defined functions
 - iii. Operation with RVec
 - iv. Matching of object collections
- d. Applying corrections on the fly
 - i. Input PU weight file apply on MC
 - ii. Apply SF

RDF exercise set 2

1. Exercise

- a. Systematic variations use of Vary
- b. RDataFrame objects as function arguments and return values
- c. Visualizing the computation graph
- 2. Possible introduction of distributed RDF example...

CMSSW tutorial Oceane, Suvankar

CMSSW Introduction

- 1. Some basic introduction of CMSSW
 - Description of Event Data Model
 - Data tiers
 - Other suggestions ? Extremely vast subject and I don't know most of it
 - Release and what to use? Major, minor release etc, Master, PdMv pages for
 - i. SDT nigtlies,
 - ii. SCRAM_ARCH
- 2. Basic interfaces of user code
 - EDAnalyzer
 - EDFilter
 - EDProducer
- 3. EDMDumpEventContent, EDMConfigDump
 - Parameter description

CMSSW exercise

- 1. Setting up the CMSSW area
- 2. Checking out packages
- 3. Writing analyzer
 - a. Example of using MiniAOD data as input
 - b. Writing a DQM module
- 4. Generating MC events
 - a. Compare simulation to reconstruction
- 5. Use of runTheMatrix
- 6. Bonus how is your development integrated into CMSSW fwk?