Contribution ID: **649** Contribution code: **contribution ID 649**   Type: **Oral**

Using ROOT to analyse High-Frequency Finance Data

Wednesday, 1 December 2021 18:20 (20 minutes)

The analysis of high-frequency financial trading data faces similar problems as High Energy Physics (HEP) analysis. The data is noisy, irregular in shape, and large in size. Recent research on the intra-day behaviour of financial markets shows a lack of tools specialized for finance data, and describes this problem as a computational burden. In contrary to HEP data, finance data consists of time series. Each time series spans multiple hours from the start to the end of a trading session, and is related to others (i.e., multiple financial products are traded in parallel at an exchange).

This presentation shows how ROOT can be used in high-frequency finance analysis, which extensions are required to process time series data, and what the advantages are with regard to high-frequency finance data. We provide implementations for data synchronisation (i.e., zipping multiple files together), iterating over the data sequentially with a mutable state (i.e., each entry updates the state of a financial product), generating snapshots (i.e., resampling data based on the timestamps of the entries), and visualisation. These transformations make it possible to fold time series data into high-dimensional data points, where each data point contains an aggregation of recent time steps. This new dataset removes the need to process data serially as a time series, and instead allows the use of parallelised tools in ROOT, like RDataFrame.

Significance

Research in finance, and particularly fraud detection in financial markets, needs a boost in knowledge and capabilities to work with big data. There is a backlog of data to be investigated by regulators and researchers. Project HighLO contributes to this endeavour by analysing high-frequency trading data from electronic exchanges, introducing HEP tools to research in finance, and investigating fraud in commodity markets.

References

Verhulst E Marjolein, Philippe Debie, Stephan Hageboeck, Joost ME Pennings, Cornelis Gardebroek, Axel Naumann, Paul van Leeuwen, Andres A. Trujillo-Barrera, and Lorenzo Moneta. "When two worlds collide: Using particle physics tools to visualize the limit order book." Journal of Futures Markets (2021).

Speaker time zone

Compatible with Europe

Primary authors: DEBIE, Philippe (Wageningen University, Wageningen Economic Research); NAUMANN, Axel (CERN); PENNINGS, Joost (Wageningen University, Maastricht University, University of Illinois at Urbana-Champaign); TEKIN-ERDOGAN, Bedir (Wageningen University); CATAL, Cagatay (Qater University); REMBSER, Jonas (CERN); VERHULST, Marjolein; VAN LEEUWEN, Paul (Wageningen Economic Research); MONETA, Lorenzo (CERN); AL-SKAIF, Tarek (Wageningen University)

Presenter: DEBIE, Philippe (Wageningen University, Wageningen Economic Research)

Session Classification: Track 2: Data Analysis - Algorithms and Tools

Track Classification: Track 2: Data Analysis - Algorithms and Tools