



# Test Mini Analytics

Version : 30/01/2017

This is a test for a junior analytics developer. The recommended time for the test would be 3h, but the candidate is free to spend time on the test. It is desirable to give us the time you've spend to do the test. Please send the result as a zip file to :

[andrew.louth@ancoa.com](mailto:andrew.louth@ancoa.com)

[simon.vella@ancoa.com](mailto:simon.vella@ancoa.com)

[chithanh.nguyen@ancoa.com](mailto:chithanh.nguyen@ancoa.com)

**General idea of the program** The program has two components : the data feeder in real time, and a analytics API that process the data in real time, analyse and trigger a warning message if it found appropriate. Since the real data feeder could be complicated, we use a 'fake' data feeder. It load the whole offline data from a csv file, then give out one by one element. The Analytics take the data element one by one given by the system, and process it.

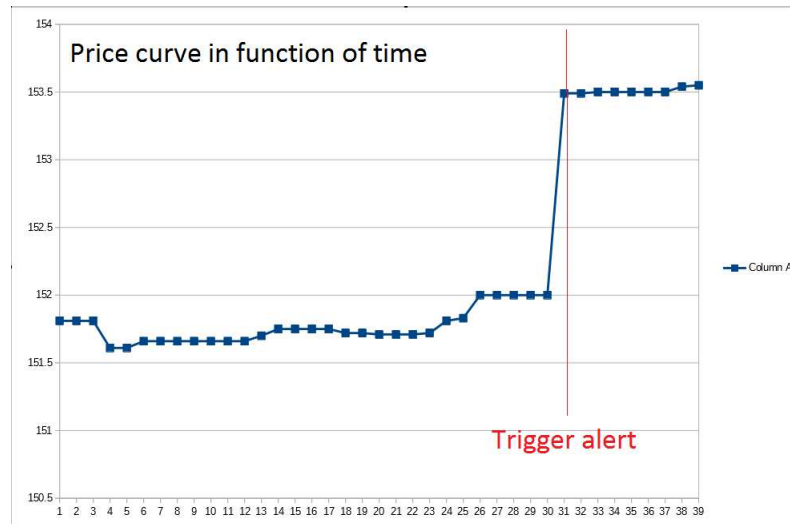


Figure 1: Abnormal Trade Data

**Example of data** In figure 1, we see a real Trade data. The vertical axe is the price, the horizontal axe does not exist, but can be viewed as time. At the vertical red line has occurred a abnormal price movement. We would like to write a algorithm to catch that high increase of price and printout a warning text, saying something abnormal happened at that moment and explaining the reason.

**To implement** The candidate has to fill the function `AbnormalTrade::process` take a real value in argument, it supposed to be a trading price at the current moment. Candidate can be free to implement this class `AbnormalTrade` as far as it inherit from `AnalyticsAPI`.

## Objective

- An csv example of Trade data is provided. Candidate firstly has to implement the function `AbnormalTrade::process` that discover when the Trade Price move abnormally, in that case, print to console a warning message.
- Candidate is then supposed to create a various different data, of different variation and different size (up to 1M points) to test the program. Can use Excel/R/Matlab or any tool to generate the data.
- At the end, write a short report (2 pages max) with some screenshots commenting how good is the algorithm, in term of accuracy, interm of performance. What was the challenges.