

Analytics Workbench (MayStreet) enables academic research team to gain unprecedented insight into odd lots

Overview

Odd lot orders, which comprise fewer than the standard unit of 100 shares, have become increasingly common on US exchange order books. This is due to a host of factors, including the unprecedented increase in the average stock price over the past decade, the rise of algorithmic trading and the growth of retail investing. However, because these orders are not publicly disseminated or included in the National Best Bid or Offer (NBBO) – the Securities Information Processors (SIPs) publish only round lots – no large-scale academic research project had yet been conducted on how odd-lot quotes compare to the NBBO for the stocks in question.

Three market structure researchers – Robert P. Bartlett of the University of California, Berkeley; Justin McCrary of Columbia University and Maureen O'Hara of Cornell University – sought to change that. They knew that finding a way to access high-quality, historical full-depth market data and robust analytical capabilities was a vital prerequisite for the project's success, but the limitations of public feeds and the prohibitive cost of proprietary feeds presented a challenge.

Bartlett had an existing relationship with MayStreet's senior leadership team, which had reached out to him years before to discuss one of his previous research papers. MayStreet shared the academics' desire to both educate the public on the evolving market landscape and support market structure research in the academic community, and so made pricing accommodations that enabled the research team to access its data and analytics while remaining within its budget.



Challenges

Taking a granular view of odd-lot quotes presented the research team with several data and analytics challenges that went beyond availability and cost. The project necessitated a high volume of large-scale data queries, so any product they used would require analytical capabilities that were both powerful and user-friendly, while the data itself needed to be clean and accurate. In addition, while the team was used to working with market data and analytics on conventional academic platforms, it had never worked within MayStreet's programming environment, so having access to an expert support team to assist throughout the process was a high priority.

Approach

The research trio settled on MayStreet's cloud-based analytics environment, Analytics Workbench (MayStreet), which eliminates the difficult and costly hassle of working with massive data sets like that the team would need to analyse for their study. Workbench gave the team seamless access to the Market Data Lake (MayStreet) – a 20+ petabyte repository of ultra-high-quality, full-depth global market data, all of which has been captured directly at the data centre level – letting them get to work quickly and confidently.

Since the success or failure of the study depended on the quality of the historical US equity feeds being analysed, it was critical that they verify that the data was as good as advertised. In addition to extensive conversations with MayStreet engineers to understand the firm's process for capturing, monitoring and normalising market data, the researchers performed their own validation testing. The researchers were delighted with the quality of the data, and in the few cases when the team did notice a data discrepancy, MayStreet was able to update the software and re-run the new code on the accurately captured data, quickly resolving the issue.

Execution

The team found the analytics environment easy to use and that it could support all necessary data queries. For example, because queries to access top-of-book data across all exchanges are relatively common, Analytics Workbench (MayStreet) has this capability built in. Instead of forcing users to send a program to acquire raw data from each individual source and sift through it manually to determine the best bid and offer, MayStreet automatically performs this work on the back end. The ability to skip this time-consuming data management work to determine the best bid and offer across all exchanges at any given time interval substantially simplified the computational challenge for the research team.

"As a non-Python user, I found it very easy to get up and running on the platform and perform the necessary queries," said Bartlett.

At times, the research team wanted to run queries that were too large for their virtual server. In these cases, MayStreet engineers were able to increase memory allocations so the researchers could get their queries done in a timely and cost-effective fashion, reverting them back to the original allocation upon completion.

"MayStreet not only provides the one-stop shop for all the data and analytics needed to truly understand the markets, but also a support team that works tirelessly to help users make the most of these capabilities," said Bartlett. "Their platform and team played a vital role in our work to shine a light on odd lots."



Results

The researchers publicly released their paper – [‘The Market Inside the Market: Odd-Lot Quotes’](#) – in February 2022, concluding that the US market system’s current approach to odd lot quotes has resulted in a large ‘inside’ market that often contains many prices better than those offered by the NBBO.

The paper immediately drew the attention of the equity market structure community, with multiple industry observers citing it as proof that the SIPs must include odd lots to ensure that all traders have full visibility into the best prices on the market. And perhaps uncoincidentally, shortly after publication the SIP Operating Committees announced they were re-evaluating a three-year-old proposal to add odd lot quotes to the UTP and CTA/CQ data feeds.

The research team is now working on a new project using MayStreet’s depth-of-book data.

Summary

Analytics Workbench (MayStreet) met the needs of the academic research team in terms of data availability, data quality, analytical capabilities, technical support and economics. This enabled the team to gain a level of insight into odd lot quotes that was unprecedented in the academic community, ultimately leading to the publication of a research paper explaining their precise impact. These capabilities will also form the basis for future research projects, one of which is already under way.

