Anomaly Detection

Cryptocurrency Pump and Dumps

What?

Coordinated market manipulation targeting retail investors

How?

- Social media groups coordinate a time to increase an assets price
- Participants buy the asset when given the signal
- An insider group sells at that time
- Participants funds are drained

Wrangling

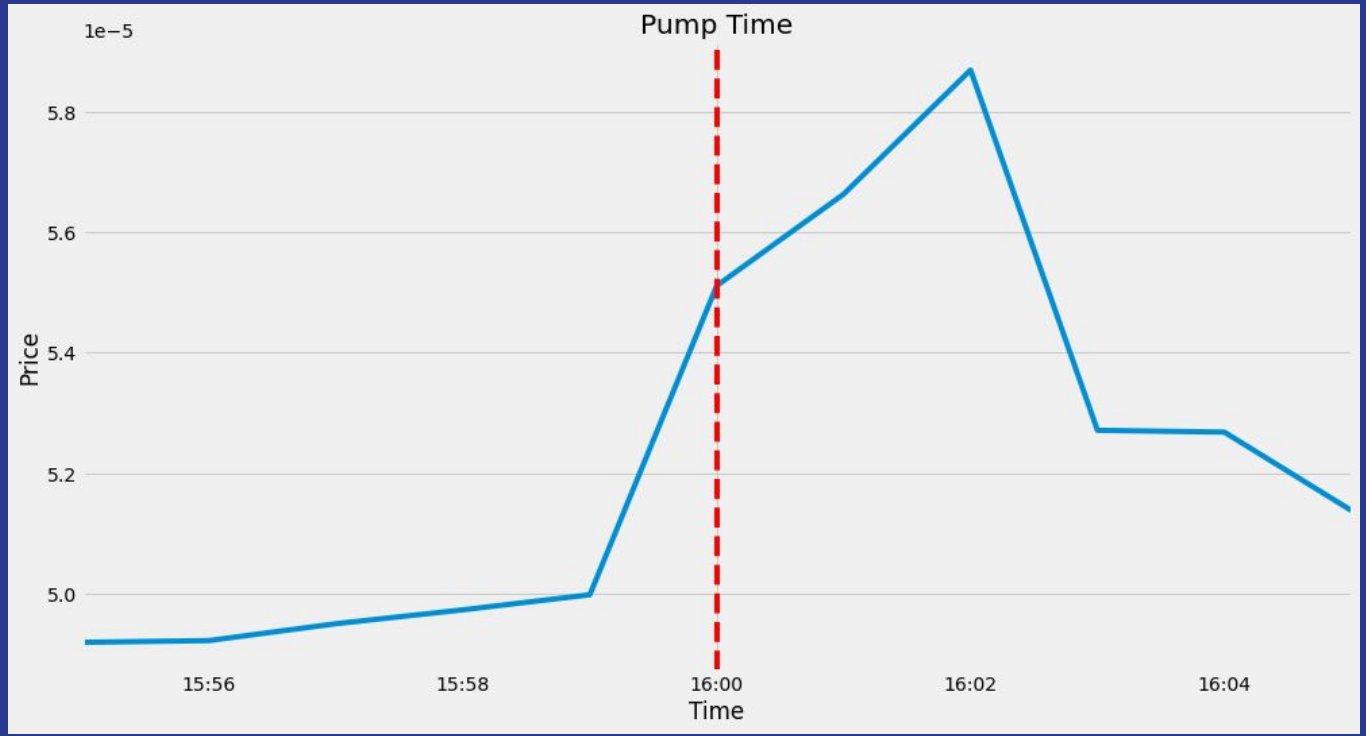
- 200 instances of manipulation gathered from Binance & Telegram
- 1 Month of asset trading data per pump event
- 1min K-lines
- Trading data to millisecond

Exploration

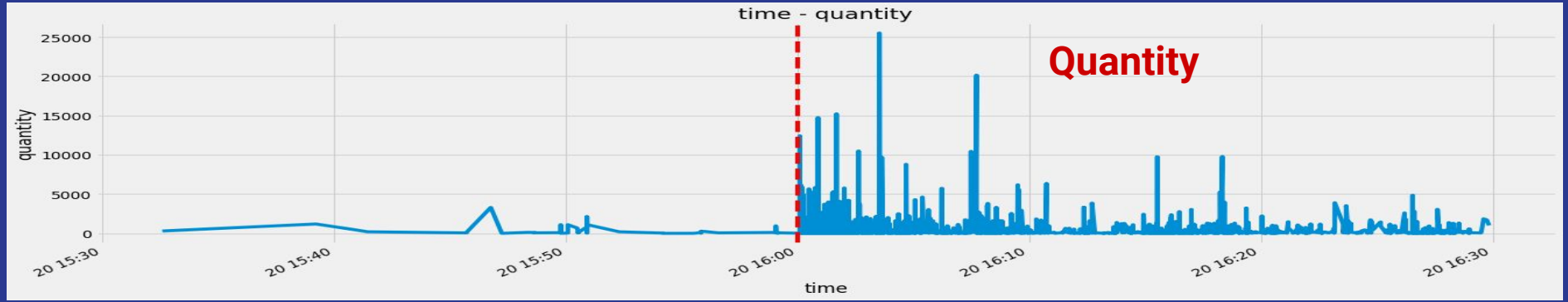
- Successful events will increase the price 2x-5x on average
- All coins are under a 50,000,000 market cap
- The number of individual quotes increases 10x, indicating more participants in the scheme

Cols:
Time - ID - Price - Quantity - QuoteQuantity - IsBuyerMaker

Exploration (cont.)



Exploration (cont.)



Preprocessing

- Extracted only 15 mins before/after the event
- Refocus on trading data only
- Each individual pump standardized

Preprocessing (cont.)

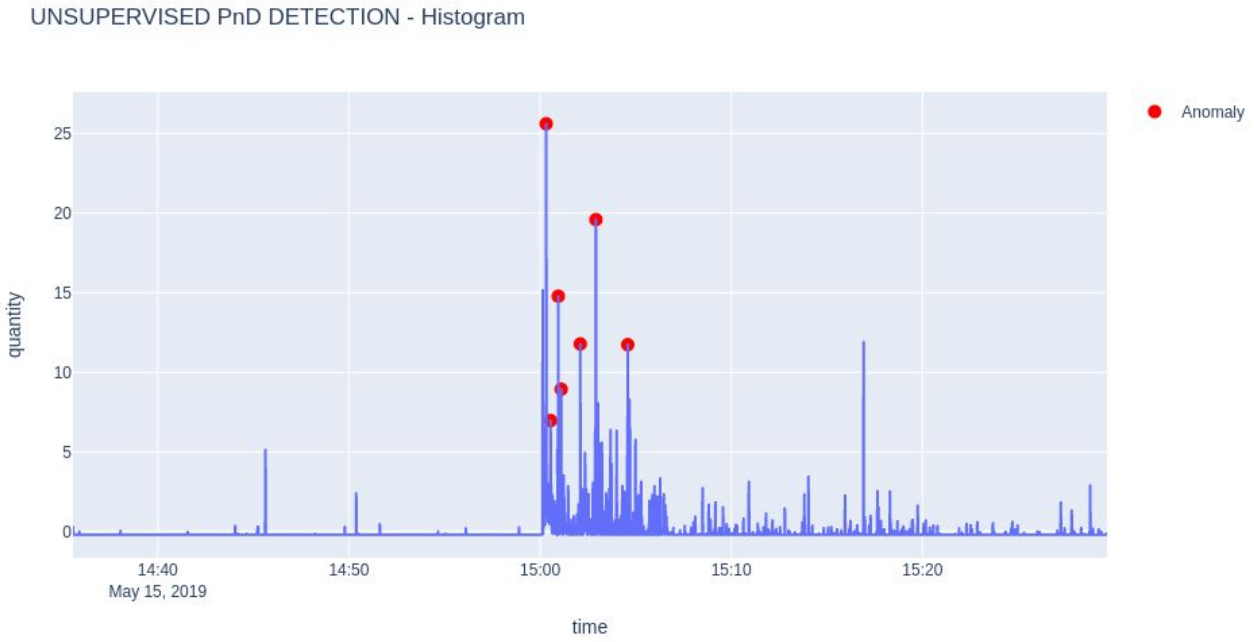
- Each pump resampled to 1s bins, data aggregated properly
- Missing data forward filled
- All pump events merged into a single dataframe

Modeling

3 Models for outlier detection were tested:

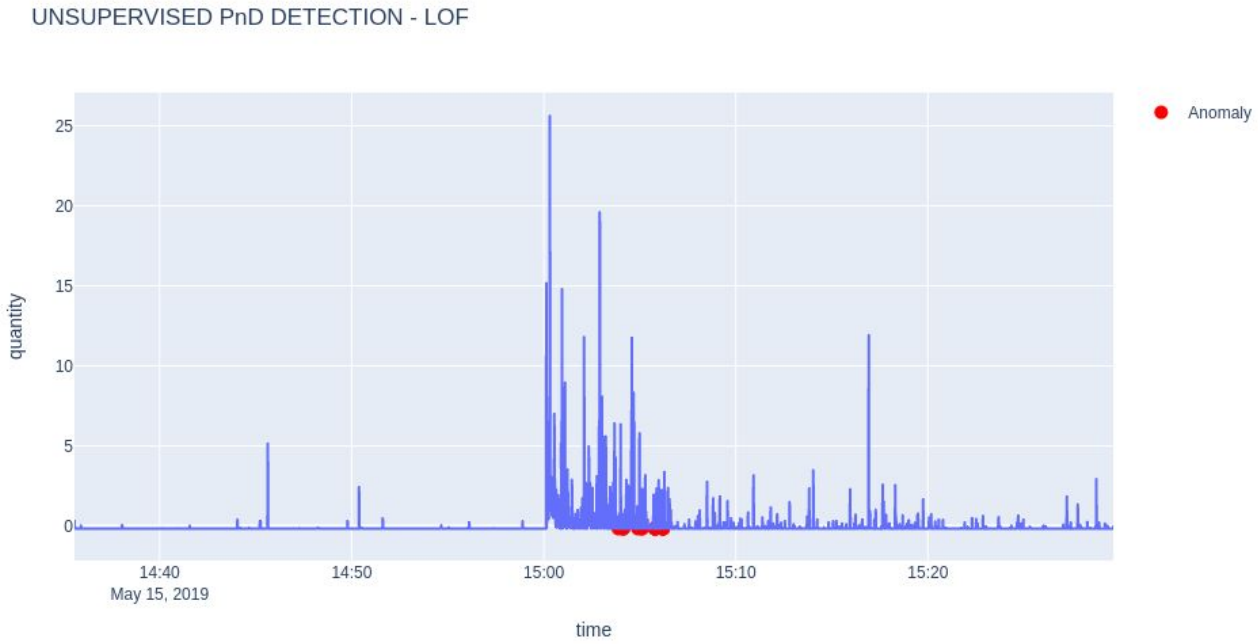
- Histogram: Simple and useful for sparse data
- Local Outlier Factor: Builds on KNN to find localized outliers
- Isolation Forest: Ease of use and targets imbalanced data

Modeling - Histogram



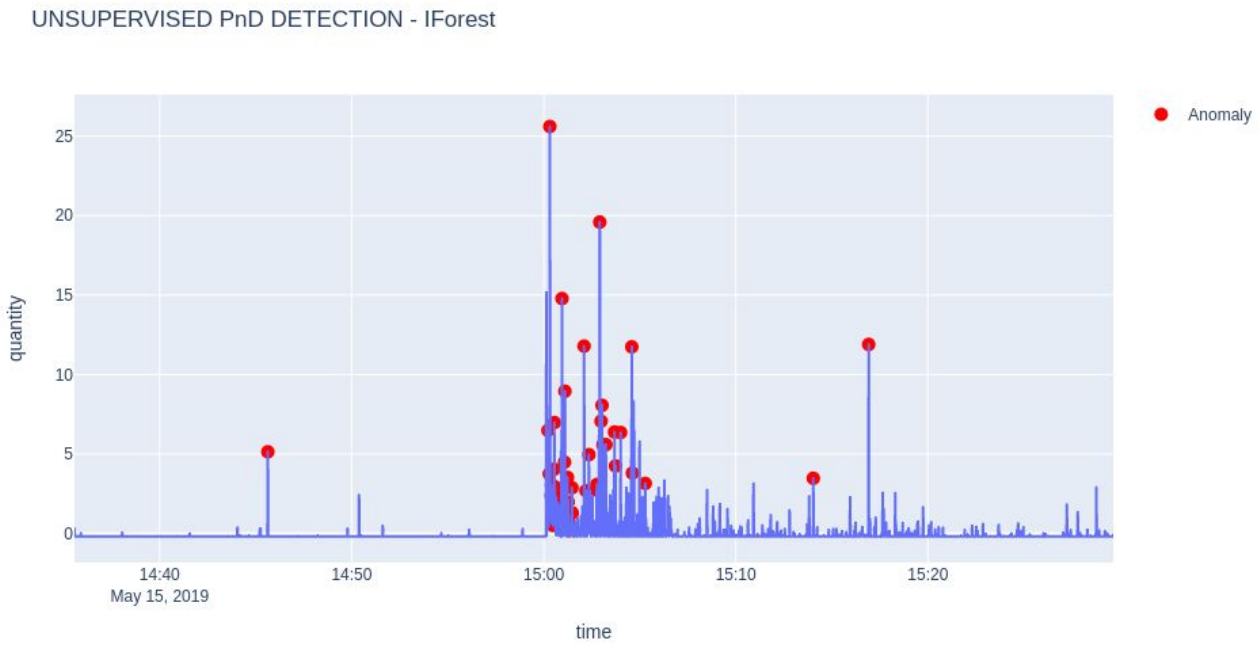
Few points found

Modeling - Local Outlier Factor



Not useful

Modeling - Isolation Forest



Many expected Points found

Several novel anomalies to explore

Why?

Millions in value is drained from hopeful novice investors each year

- Identify organizers
- Find likely target assets
- Better educate participants
- Increase trade safety for exchanges

Further Research

- Define features like the if limit order and volume of money
- Use different resampling times
- Adjust rolling mean window for resample

Conclusion

Isolation forest is found to be the most useful. It found numerous anomalies minutes to seconds before the pump signal