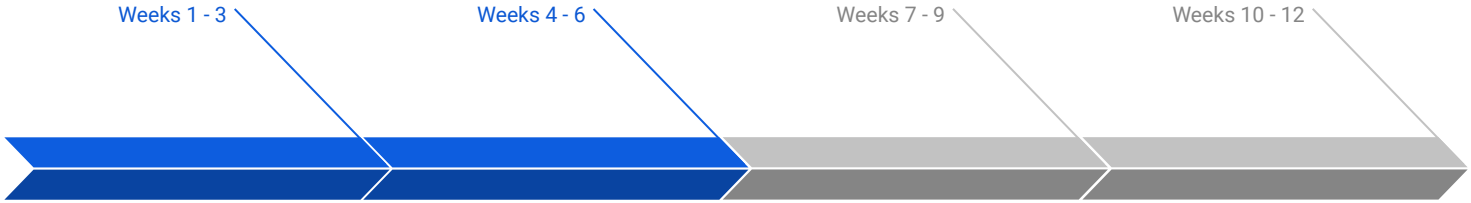
The Quest to Generate Alpha From Alternative Data

Group: Quantzz - Aryan Mishra, Emilio Alberini, Thomas Gerbeaud, Hippolyte Buisson, Eli Salk

Final Outcome

Final Outcome (For Real)

Timeline of Strategies



Simple Fama - French 3 Factor Model

Our initial strategy was to run a simple regression model based on the Fama - French 3 Factor model.

Incorporate Online Sentiment Factor

We expanded our focus to include scraped Twitter and Yahoo information to leverage market sentiment.

Increased Focus on Cryptocurrencies

Attempting to capture the alpha from short-term volatile swings, we decided to mostly restrict our universe to crypto.

Consolidation of Previous Strategies

Realizing that no one strategy completely dominates the other, we consolidated all our strategies in one model.

Simple Fama - French 3 Factor Model

- This strategy consisted of ranking the company size, value, and performance by normalizing all of the data with a Z-score and then ranking them in order from the average of all three Z-scores combined.
- However, results were poor since we didn't have enough relevant data. We restricted our time horizon to only include data from the previous three weeks.

Market Sentiment Analysis

Step One

Build a scraper in order to scrape data from Twitter and/or Yahoo Finance.

Step Two

Obtain the relevant tweets from Twitter and the “performance outlook” indicator from YF.

Step Three

Clean the data to minimize as much noise as possible (especially for the scraped tweets).

Step Four

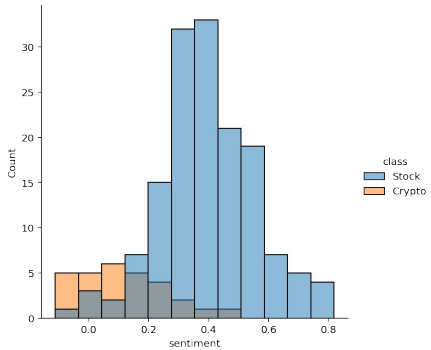
Perform sentiment analysis on the data.

Step Five

Aggregate the tweets and the performance indicators and calculate the sentiment factors.

Step Six

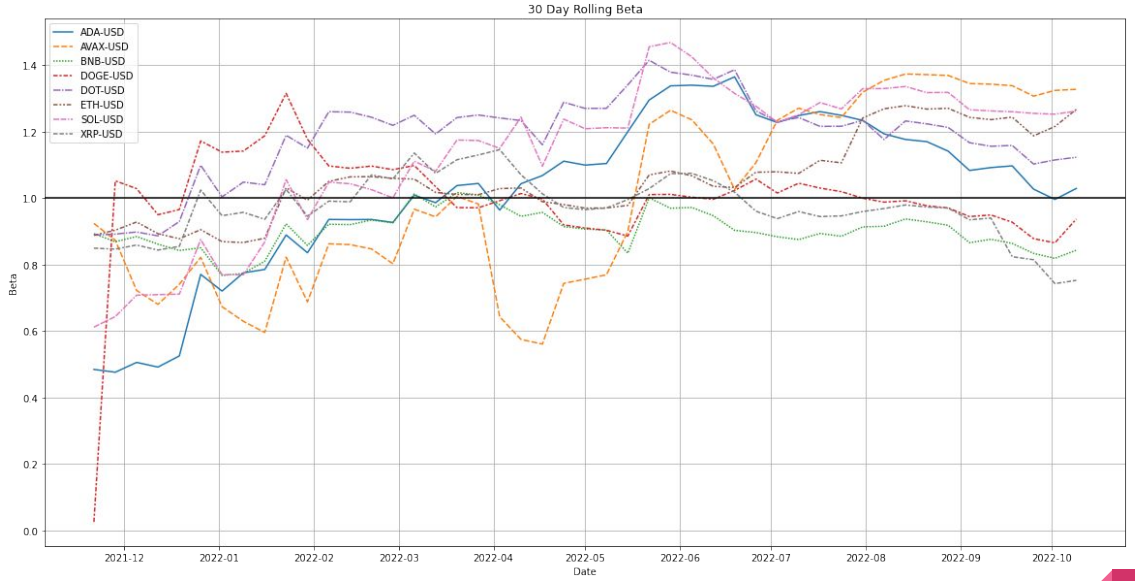
Forecast returns, iterate and optimize.



Crypto Strategy - Beta Signal

- This signal is based on the expected weekly move given the BETA of a crypto to BTC
- The BETA is calculated using 30 weeks worth of data on a rolling basis
- We use last week's BETA to calculate the expected move of a crypto given the move in BTC this week.
- If crypto moved 5% more (less) than expected we short (long) it for the following week

30 Week Rolling BTC BETA



Backtest

- L/S 100$ positions held for one week, benchmarked against a EW long only strategy

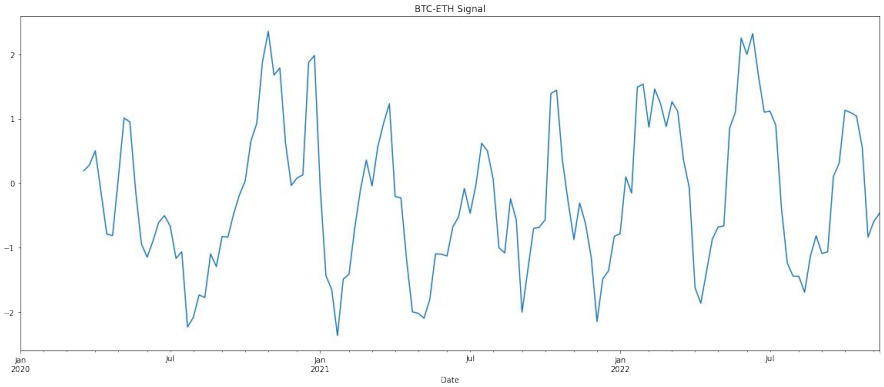
Crypto Strategy - Pairs Trading

- We consider each possible pair in our crypto universe and compute a times series of the ratio their prices
- We normalize this ratios by using a rolling mean and standard deviation with a 2 week lookback
- The normalized time series are then checked for stationarity using the ADF test

```
('XRP-USD', 'AVAX-USD') stationary
('SOL-USD', 'DOGE-USD') not stationary
('SOL-USD', 'DOT-USD') not stationary
('SOL-USD', 'SHIB-USD') not stationary
('SOL-USD', 'AVAX-USD') stationary
('DOGE-USD', 'DOT-USD') not stationary
('DOGE-USD', 'SHIB-USD') not stationary
('DOGE-USD', 'AVAX-USD') stationary
('DOT-USD', 'SHIB-USD') stationary
('DOT-USD', 'AVAX-USD') stationary
('SHIB-USD', 'AVAX-USD') stationary
```

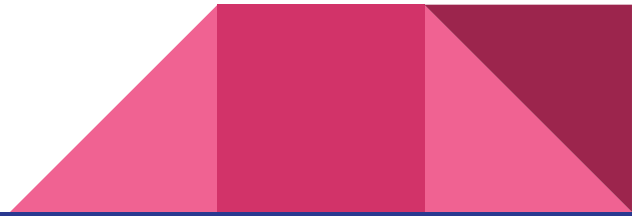
Crypto Strategy - Pairs Trading

- When the normalized time series reaches a value of 2 (-2) we short (long) the pair
- For the BTC-ETH pair a long position implies that the ratio will increase thus we want to be long BTC and short ETH, the opposite for short
- We do equal weight for both pairs.
- Could be improved by accounting for volatility in the weights of the paris



Final Strategy

- Consolidation of all previous factors used to create a more systematic strategy. Roughly divide the factors into four main sources of signal:
  - Sentiment from unstructured text data (Twitter, YF).
  - Common financial technical indicators + usual Fama-French factors.
  - Crypto Pairs Trade Signal.
  - Crypto - Beta Signal
- Each signal outputs a long/short/neutral position. We then regress these onto weekly returns to identify the strength of each signal.

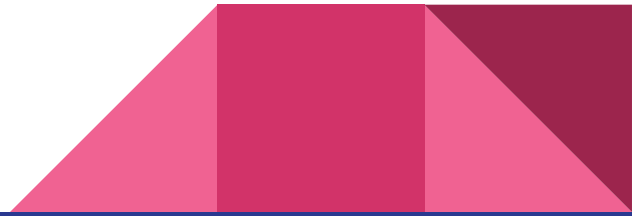


Conclusion

- While there were some weeks where our strategies performed better than the market, overall, our performance was sub-par. The reasons:
  - There might have been a lag in the sentiment strategy which could explain the below average performance.
  - The crypto market suffered historical losses this fall (especially November).
  - Unstable macroeconomic conditions contributed to the poor returns.
  - We didn't have enough time to fully back-test our final consolidated strategy.

Lessons Learned

- Markets are noisy.
- The importance of alternative data increases exponentially the closer the short-term horizon (and vice versa).
- We found that short history, low signal-to-noise ratio, and lack of stationarity makes it hard to achieve reliable and robust results.



Questions?