Tyrone Hunt - Project Update

Cryptocurrency trading decision support

I started my investigation with two hypotheses… my focus is now on the second…

1. *I can visualise my trading portfolio and calculate ROI*

Progress
- Built an ETL process for Bitcoin trades from Coinbase and GDAX – provides a single view of a users Bitcoin wallet since sign up
- Setup Plotly API as a dynamic graphing tool
- Generalised Coinbase ETL process for any cryptocurrency

Key challenges overcome
- Translating Unix time into datetime format
- Interpolating a Bitcoin price for a time that wasn't in the dataframe
- Identifying (and removing) transfer data between Coinbase/GDAX

Remaining challenges
- Overlaying markers for buys/sells on the Plotly graph
- Generalising GDAX ETL process for any currency
- Finding a source of price data for other cryptocurrencies
- Moving from csv files for price data to APIs

Conclusion – not feasible to finish this hypothesis in remaining time
- I can't calculate portfolio level ROI unless I build/combine wallets for all currencies
- I can't calculate ROI on Bitcoin wallet without building a new process to match BTC wallet transactions to transactions in other wallets

2. *I can predict cryptocurrency trading outcomes*

Progress
- Created a dataset of OHLC price data and 'raw' trading signal features
- Created 2 simple trading simulations –*approach too time consuming!*

Proposed next steps
1. Create a categorisation problem, where each instant in time is classified as a buy or sell moment
   - ➢ Create a target column, i.e. has price gone up/down 24h later
   - ➢ Turn raw trading signal data into 'meaningful' features. E.g. turn 'moving average' indicator into categorical data (is price above/below indicator), or into calculated continuous data (how much has indicator moved since previous period)

2. Use linear regression to predict the price in 24h time *(same steps)*

Challenges
- Adding 'real-world meaning' to trading signal features is likely to yield a sparse data frame of 1s and 0s – which may not work
- Unclear how to deal with time-series data in linear regression model
- Unclear if forcing trading signals to be continuous (rather than categorical) means anything, and there are infinite possibilities…