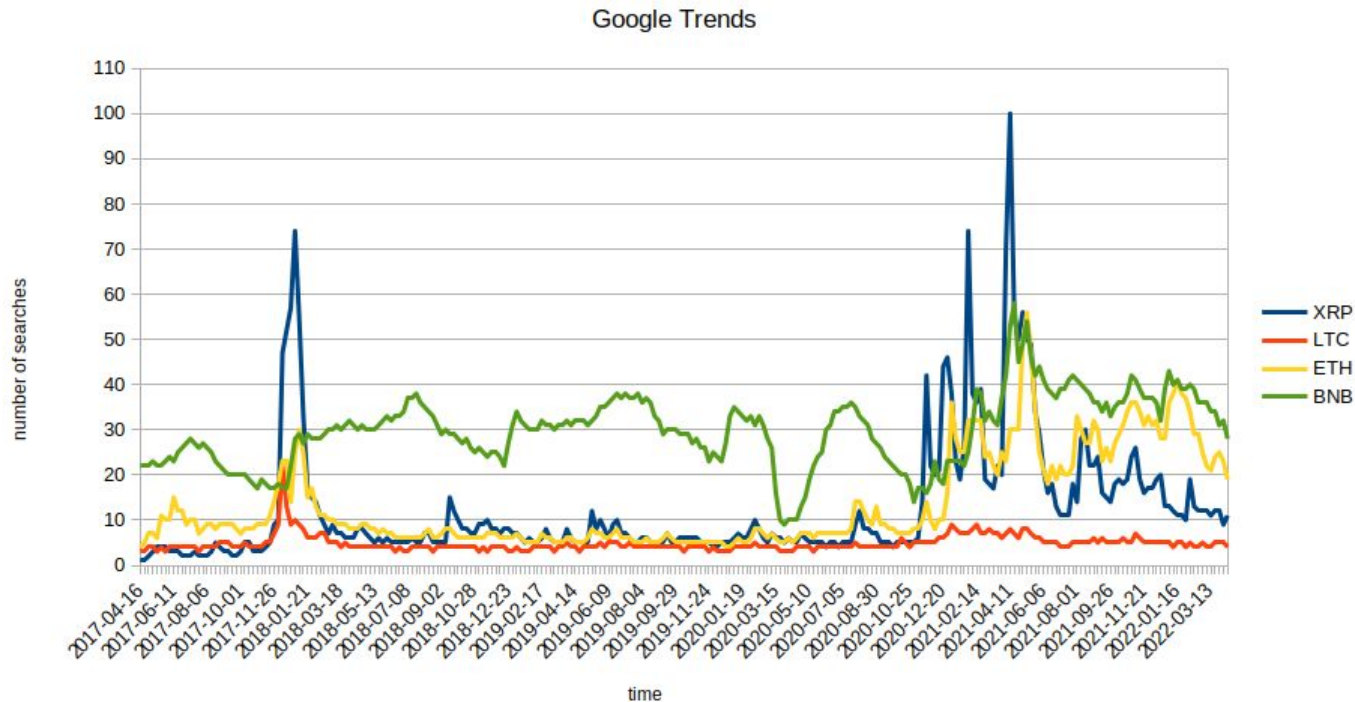


# Decentralized Finance Backtesting



Jakub Ryšavý  
28. 4. 2022  
PV174

# Table of Content

- Intro
- Aim
- Data
- Our Approach
  - description of steps
- Summary
- Discussion

	1INCH	AAVE	ACA1	ADA	ALGO	AMP	ANC1	ANKR	APE3	AR	ARDR
1INCH	0.00	0.15	0.25	0.67	0.52	0.70	0.27	0.55	0.66	0.94	0.44
AAVE	0.15	0.00	0.35	0.32	0.37	0.39	0.46	0.39	0.95	0.74	0.33
ACA1	0.25	0.35	0.00	0.27	0.30	0.60	0.50	0.45	0.70	0.48	0.58
ADA	0.67	0.32	0.27	0.00	0.13	0.17	0.89	0.18	0.80	0.22	0.14
ALGO	0.52	0.37	0.30	0.13	0.00	0.30	0.68	0.18	0.33	0.12	0.11
AMP	0.70	0.39	0.60	0.17	0.30	0.00	0.92	0.24	0.60	0.43	0.32
ANC1	0.27	0.46	0.50	0.89	0.68	0.92	0.00	0.33	0.39	0.88	0.35
ANKR	0.55	0.39	0.45	0.18	0.18	0.24	0.33	0.00	0.85	0.26	0.06
APE3	0.66	0.95	0.70	0.80	0.33	0.60	0.39	0.85	0.00	0.57	0.90
AR	0.94	0.74	0.48	0.22	0.12	0.43	0.88	0.26	0.57	0.00	0.22
ARDR	0.44	0.33	0.58	0.14	0.11	0.32	0.35	0.06	0.90	0.22	0.00

# Aim

- Identify a suitable **cryptocurrency portfolio** by clustering.
- Evaluate **long-term trends** (megatrends) to find profitable investment opportunities.
- Help with **portfolio rebalancing**.

# Data

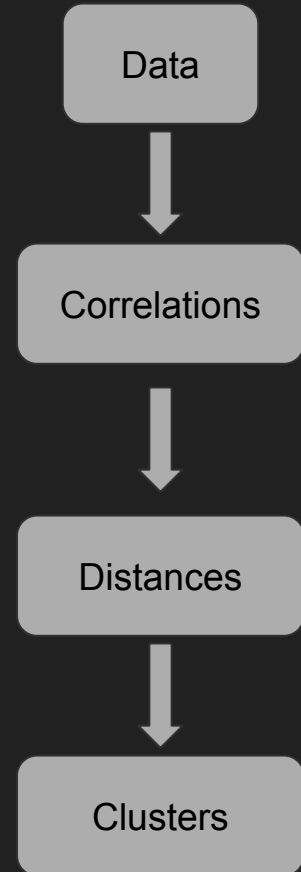
- prices
  - open, high, low, close (candle format)
- volumes
- swaps
- holders
- active addresses
- circulating supply (market\_cap)
- sentiment?
  - *Telegram, Discord, Reddit, Twitter*
  - raw messages
    - most are generated by bots
  - **buzz** - number of messages per time unit

## • Resolutions

- M1 (minute)
- M5
- M15
- M30
- H1 (hour)
- H4
- H12
- D1 (day)
- W1 (week)

# Our Approach

1. select type of data
2. calculate (pairwise) correlations
3. translate correlations to distances
4. calculate clusters on distance matrix
5. evaluate trends of clusters
6. create visualizations



## Our Approach: 1) select type of data

- only one type (*prices, circulating supply*)
- combine more types
  - equally weighted
  - searched weights by optimization
  
- Which timeframe to use? Daily?
- How long history to take?

## Our Approach: 2) calculate (pairwise) correlations

- *Pearson's correlation* is probably the most straightforward choice.
- test also other methods (*Spearman, ...*)?
  
- Calculate distance matrix without using correlation...?
- How often to recalculate?
- Use overlapping window? How much?

## Our Approach: 3) translate correlations to distances

- correlation coefficient has range  $[-1, 1]$
- distance is always **positive**
- **high correlation** corresponds to **small distance**

$$\mathit{distance} = 1 - \mathit{abs}(\mathit{corr\_coef})$$

- Is it correct?
- Is there another way?



## Our Approach: 4) calculate clusters on distance matrix

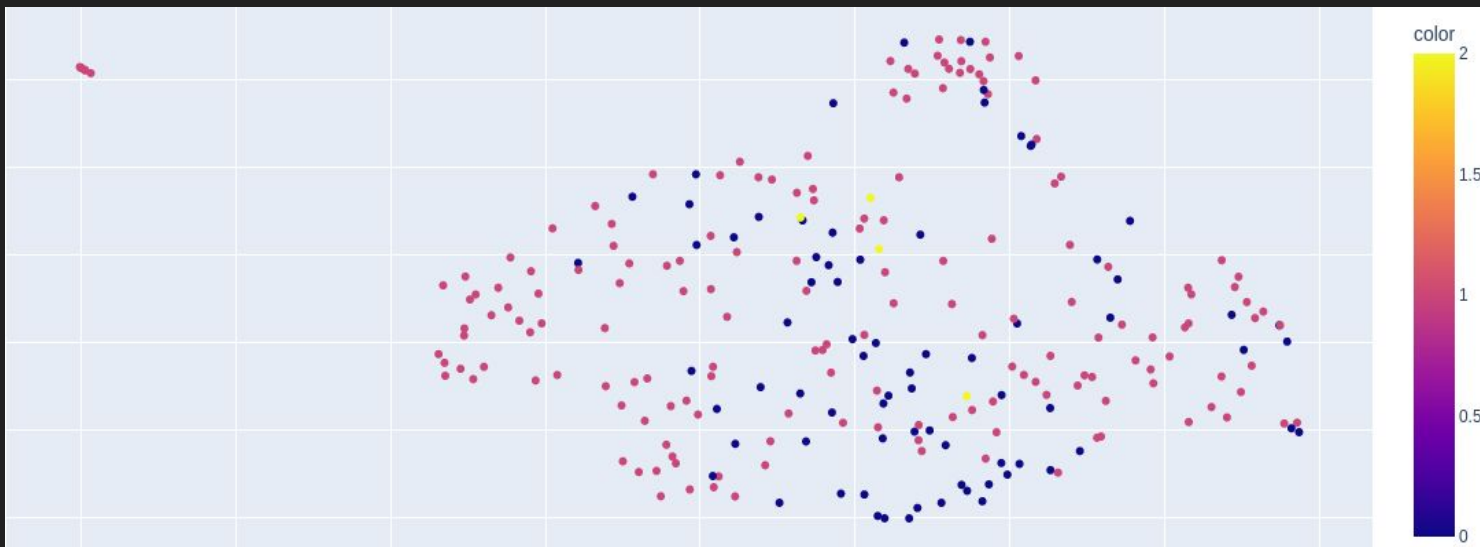
- Some clustering algorithms take distance matrix as input
  - parameter `affinity` or `metric` set to `"precomputed"` (available for `AffinityPropagation`, `SpectralClustering`, `DBSCAN` and others)
- Any way to transform distance matrix to “features” to calculate clusters on it?

## Our Approach: 5) evaluate trends of clusters

- to evaluate:
  - **baseline** - overall index (buy and hold everything)
  - **individual clusters**
- there are (probably) methods in existing packages for **portfolio evaluation**
- not yet implemented
  
- Which evaluation are expected?

# Our Approach: 6) create visualizations

- projection of distance matrix to 2D
  - parameter `metric` set to "precomputed" (available for TSNE, Isomap, UMAP, etc.)
- visualize trends
  - existing libraries can be used
  - baseline
  - individual clusters



# Summary

- We have designed an experiment pipeline.
  - data selection
  - correlation calculation
  - translation correlations to distances
  - cluster calculation
  - evaluation
  - visualization
- We do not yet know if this will lead to any positive result.
  - Clustering based on correlations seems to work.
  - However, it does not have to implicate good investment recommendations.

## Discussion

- See the questions on previous slides.

*Any feedback and suggestions are welcome!*

# Readings

- [Time-series clustering – A decade review]
- (<https://www.sciencedirect.com/science/article/pii/S0306437915000733>)
- 
- [Financial Times Series Forecasting of Clustered Stocks]
- (<https://link.springer.com/article/10.1007/s11036-020-01647-8>)
  
- Optional:
- [Clustering and Regression Techniques for Stock Prediction]
- (<https://www.sciencedirect.com/science/article/pii/S2212017316301931>)