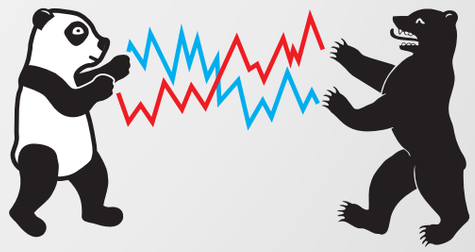
Rise of Machines?

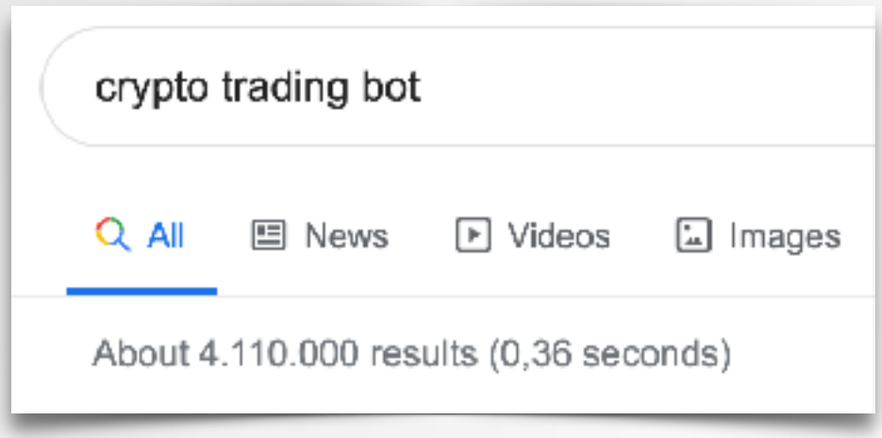
Intraday High-Frequency Trading Patterns of Cryptocurrencies

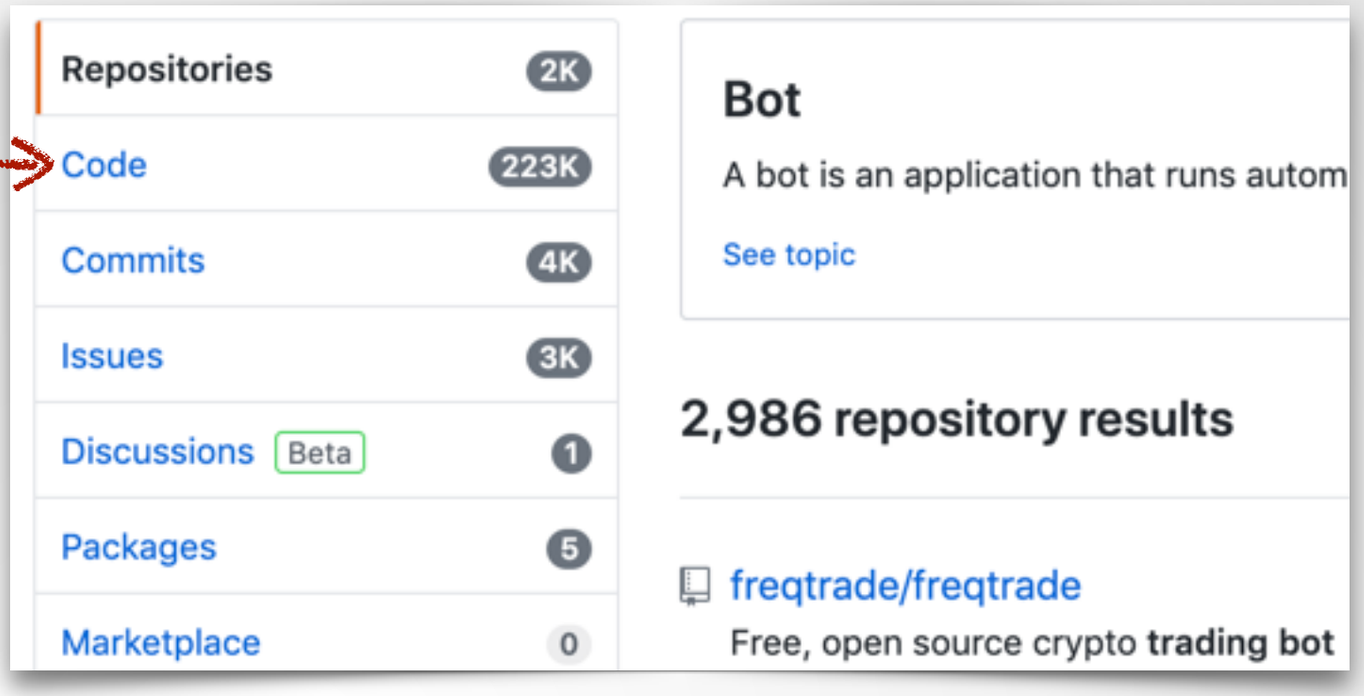
Alla Petukhina
Raphael Reule
Wolfgang Karl Härdle

Ladislaus von Bortkiewicz Professor of Statistics
Humboldt-Universität zu Berlin
IRTG 1792, BRC
irtg1792.hu-berlin.de
Charles University, WISE XMU, NCTU 玉山学者

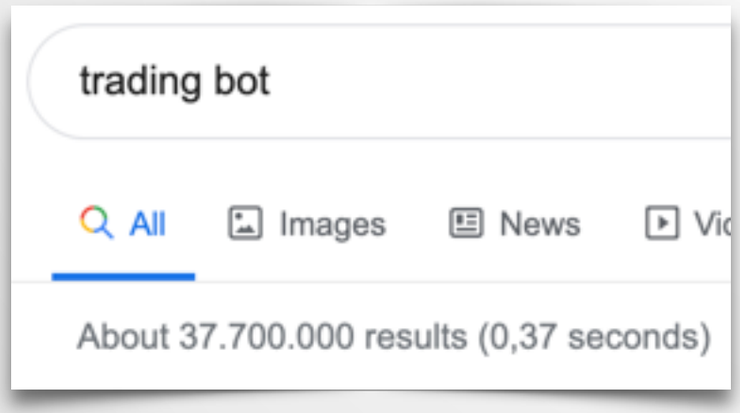


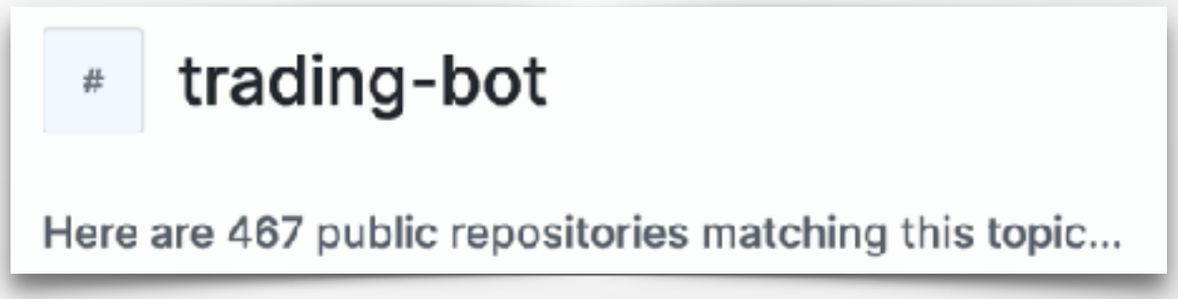
Algorithmic trading



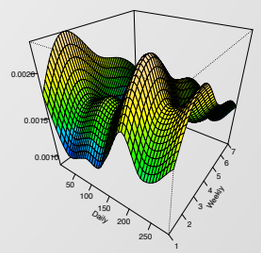


- Smart trading algorithms available, e.g. via Github
- Respective service providers available for 24/7 services
- Market must represent also machine made decisions ! (?)

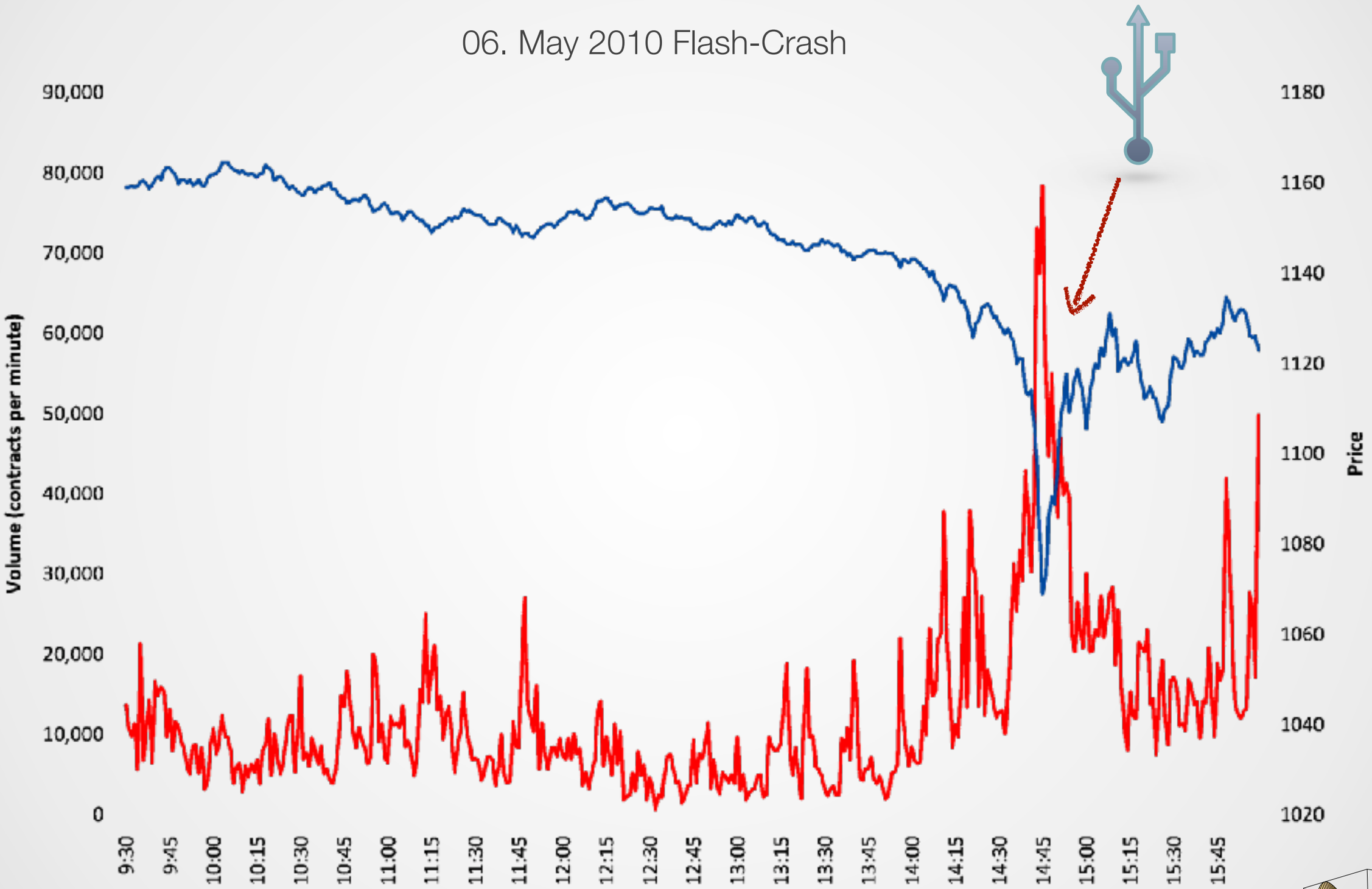




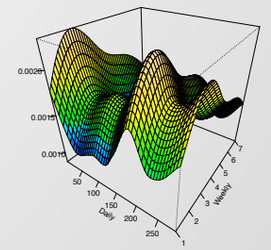
16. APR 2020



Machine made decisions are identifiable!



Price and trading volume of the E-mini S&P 500 stock index futures contract



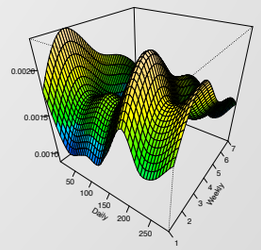
Behaviour Pattern

"That … it can't be […] *That is not a real price*. Oh well, just go buy Procter. Just go buy Procter […] *they reported a decent quarter, just go buy it.*" (Wetware unknowingly reacts to Software)

("*Is there a hedge fund that is liquidating?*") (Software reacts faster than Wetware)

"*I mean, this is ridi- … this is a good opportunity.*
When I walked down here it was at 61 […] It's at 47 […] (Wetware shuts down Software)
Procter just jumped seven points because I said I liked it at 49." (Wetware reacts to Software)

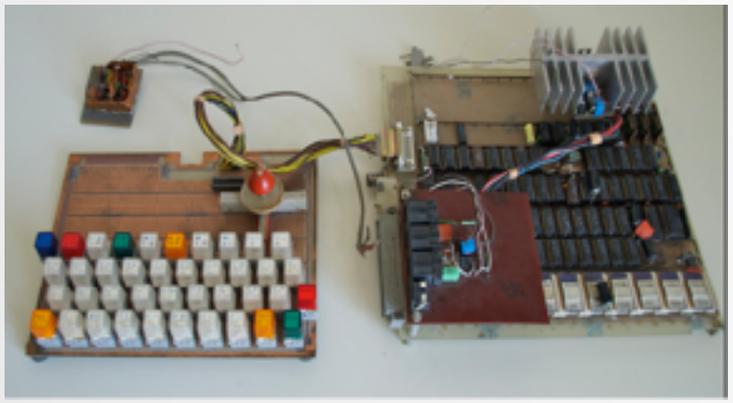
CNBC - Jim Cramer



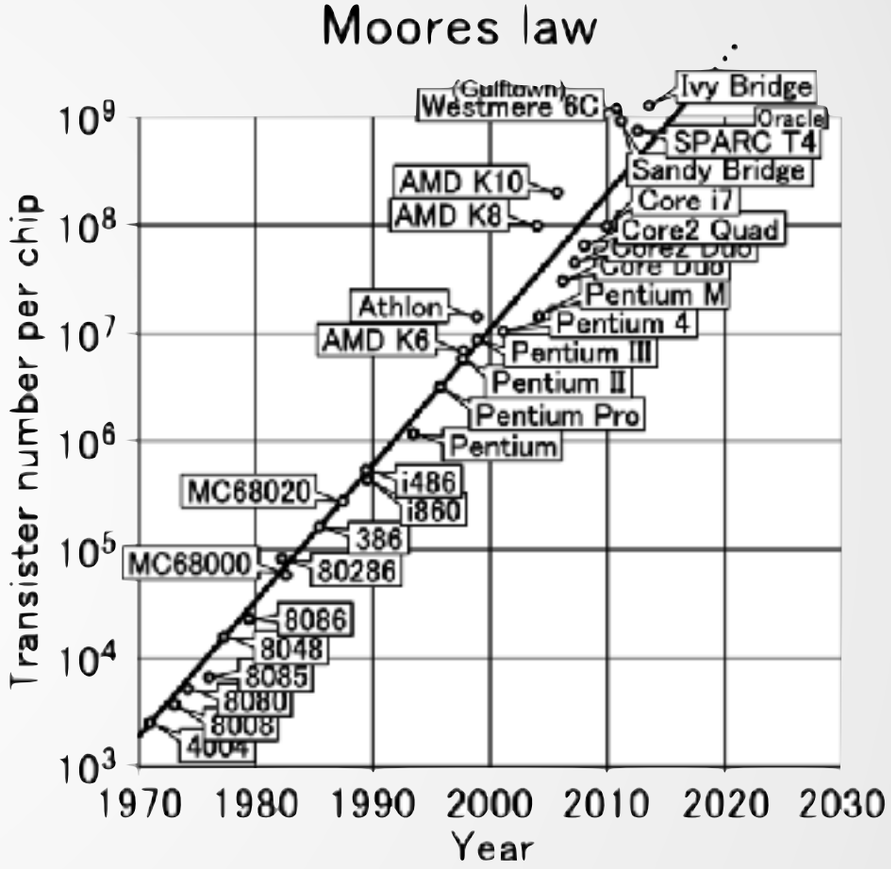
Information processing speed

Mechanical calculator MADAS
H.W. Egli S.A., Zurich, 1965

VEB Röhrenwerk Mühlhausen, 1975
8 digits

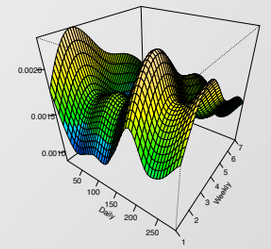


ZX Spectrum East / West 1982

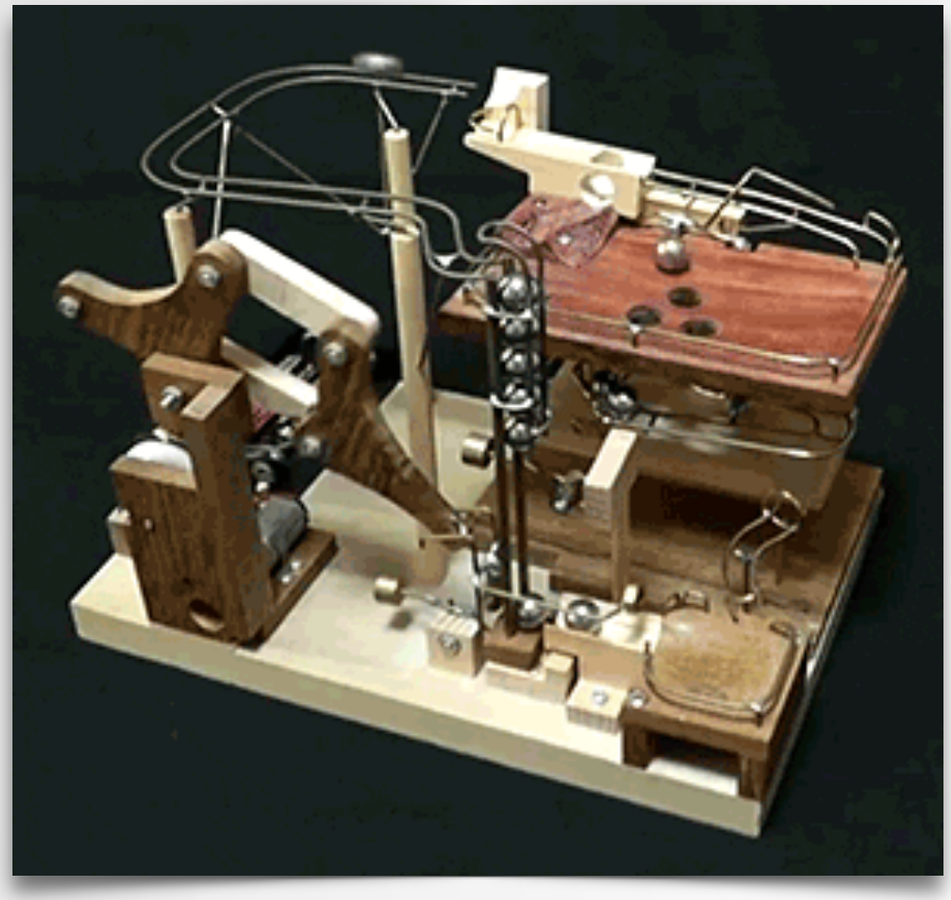


https://phys.org/news/2015-08-silicon-limits-power-electronics-revolution.html

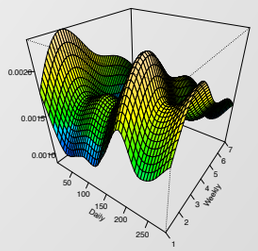
http://computermuseum.wiwi.hu-berlin.de



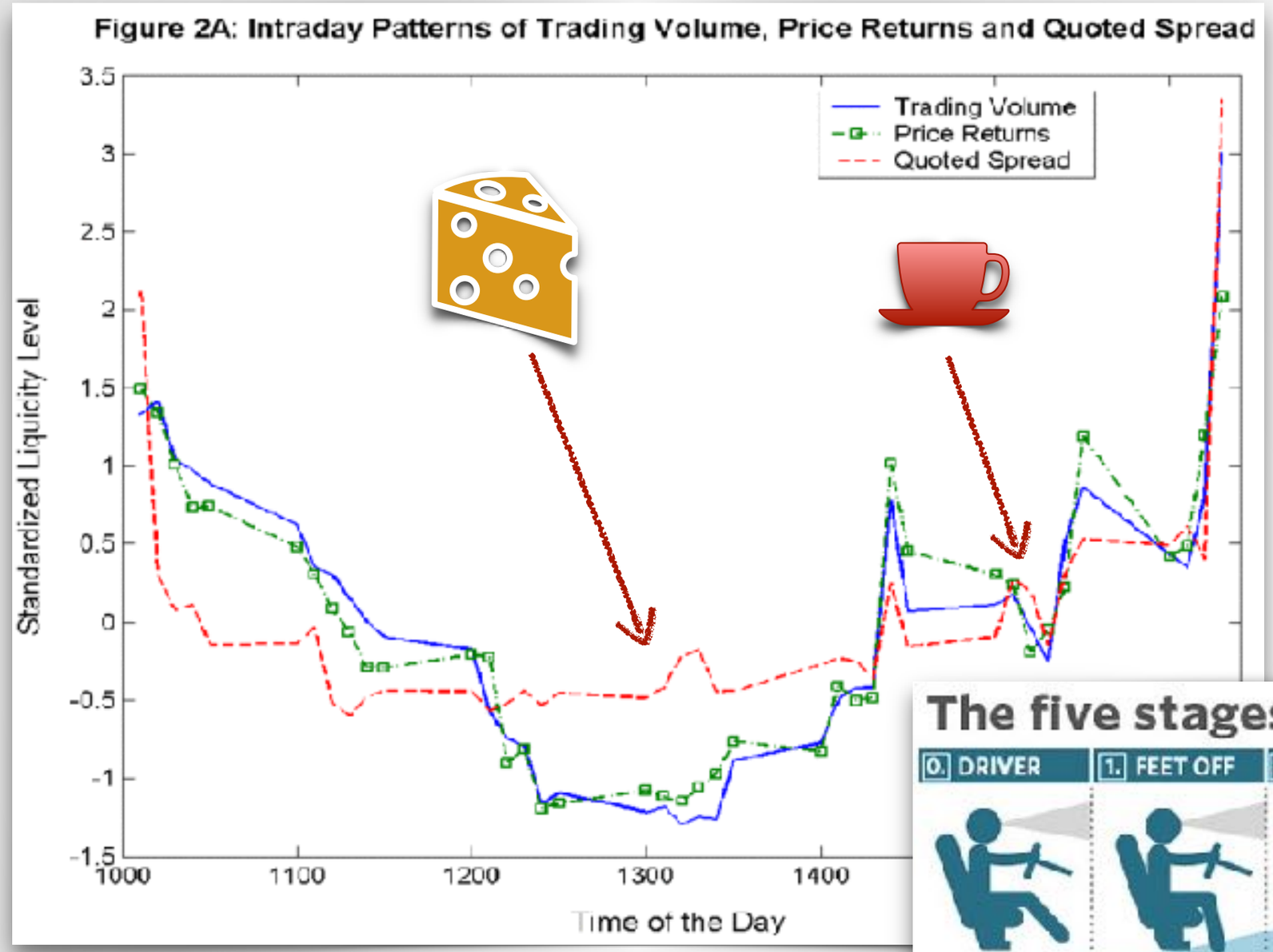
Wetware vs. Software

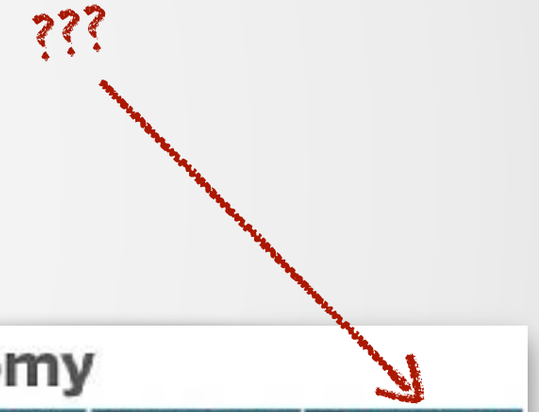


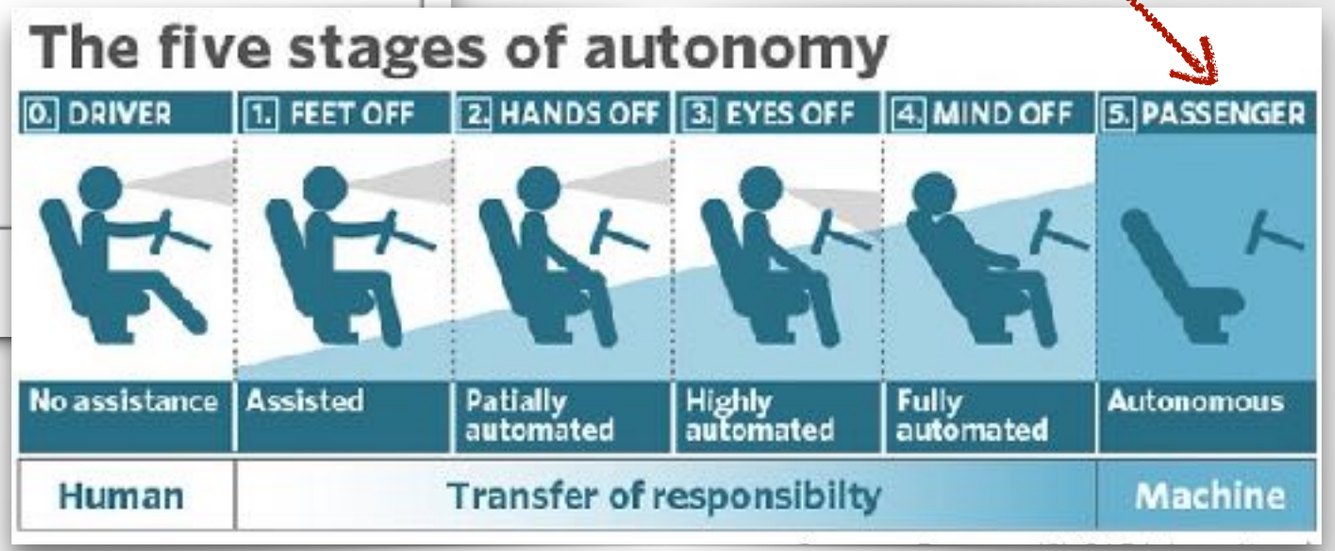
WKH

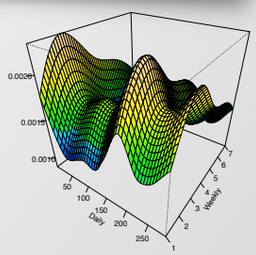


Recognisable intraday trading patterns

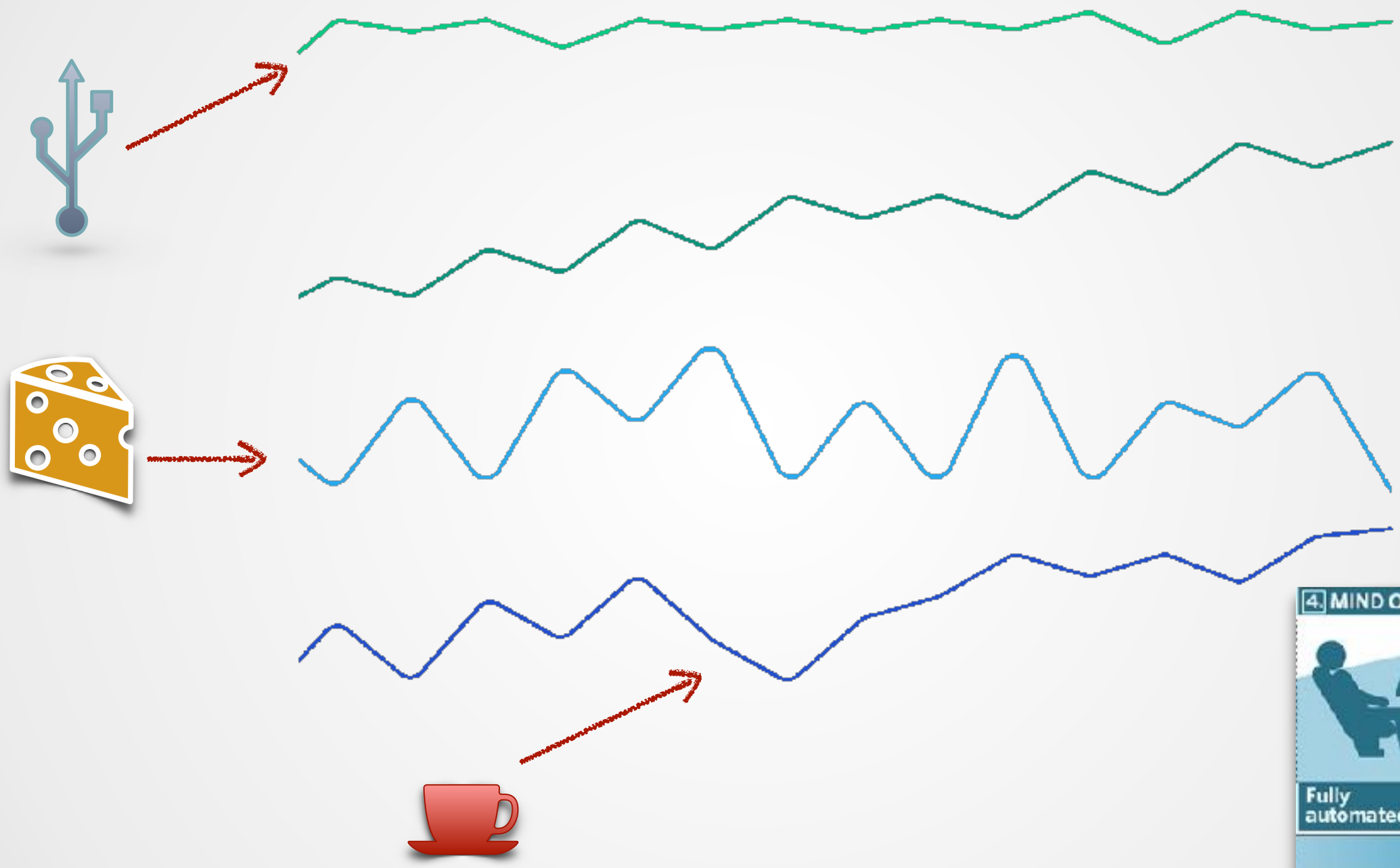


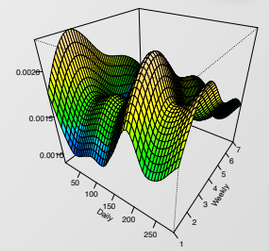






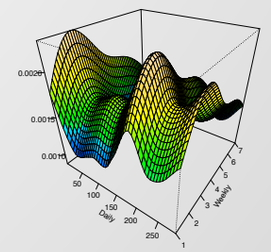
Recognisable intraday trading patterns





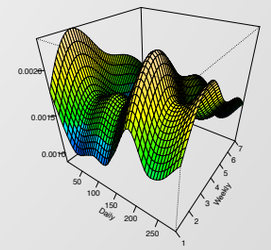
Statistical Challenges

- How to deal with non-classic high frequency cryptocurrency data
- Combine classic methods with digitised economy approaches
- Nonparam Approach? (Add Mod)
- high dimensional parameters, curse of dimensionality
- How 2 id it'd patterns



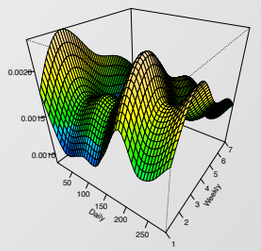
Outline

1. Motivation ✔
2. Decentralized vs. Classic Finance
3. DeFi Market Overview
4. Tech-side of DeFi
5. Data
6. Methodology
7. Empirical Results
8. Time of day effects
9. Conclusion



Systemic Similarities

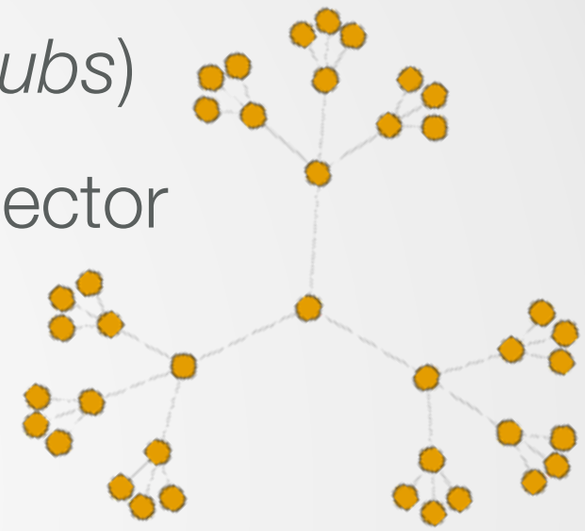
- technology (e.g. ATMs, BC constructions)
- infrastructure (e.g. Accounting, off-chain CC-related backends)
- markets (e.g. Stock / CC exchanges)
- methods (e.g. Sentiment Analysis)
- applications (e.g. algorithm-/oracle-backed Trading / SCs)



Systemic Differences

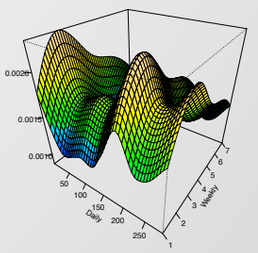
Classic Finance (ClaFi) centralisation (e.g. accounting purposes)

- traditional activity patterns create access points (*hubs*)
- concentration of sufficient tx volume & # in given sector
- concentration of expertise & ressources



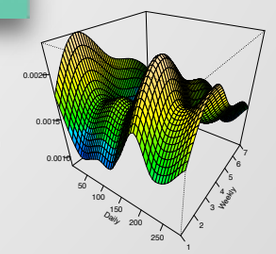
ClaFi de-centralisation meeting supply & demand (e.g. cash value)

- locally (e.g. also domestically)
- regionally (“)
- global

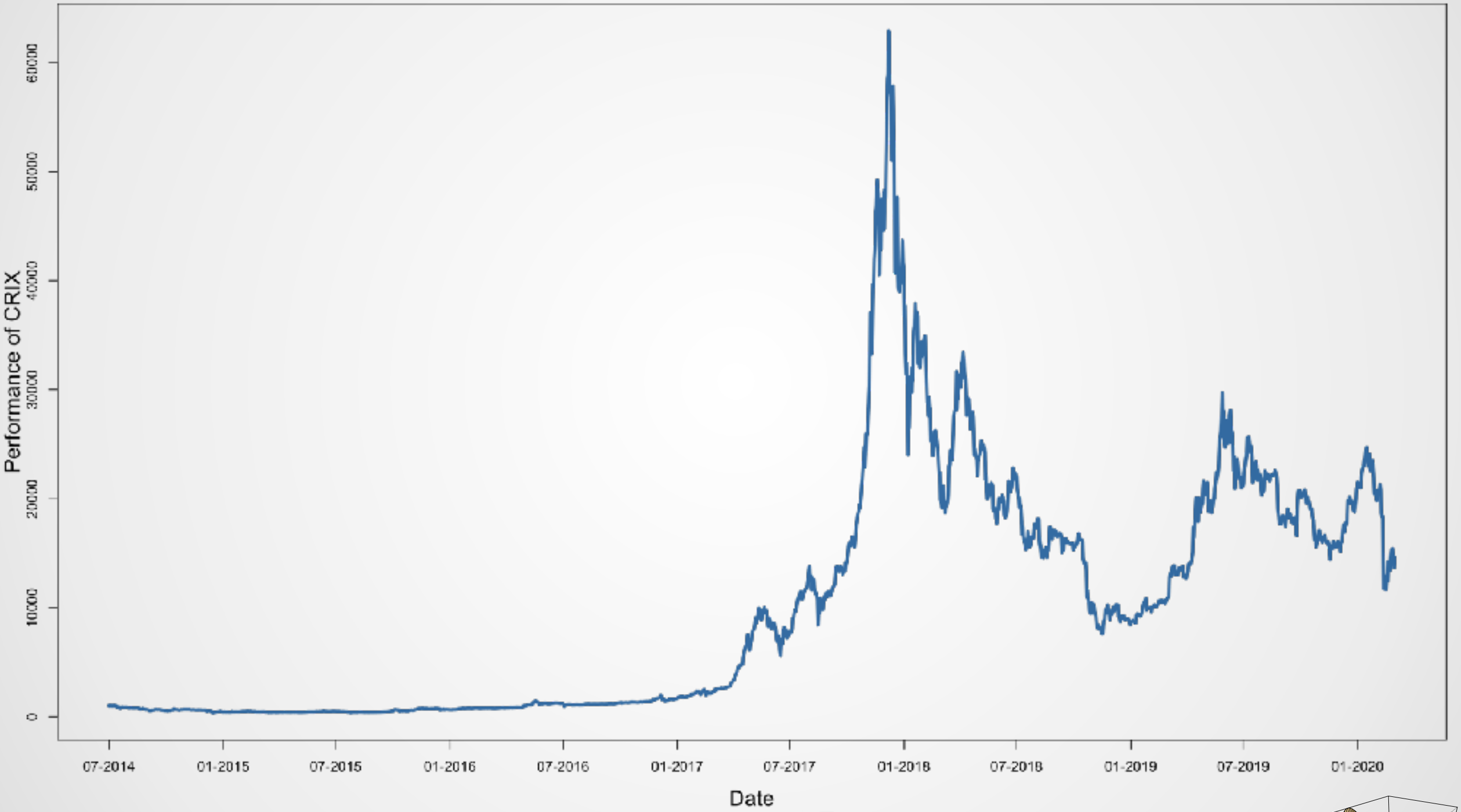


Decentralized Finance

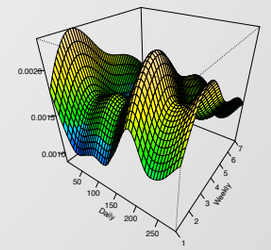
- scalability via tech, not through hubs
- no need to adjust to language, information, and law
- no need for costly means to access (e.g. bank branch)
- no geopolitical dependencies & influence over financial assets
- reduced human-based dependencies (e.g. Algo-driven trading)



Benchmark CRIX

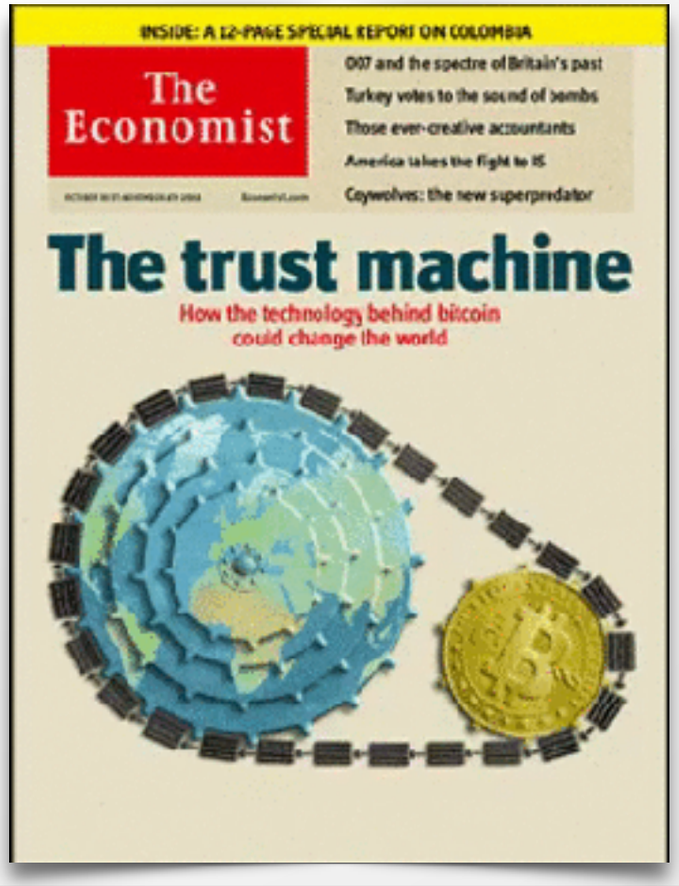


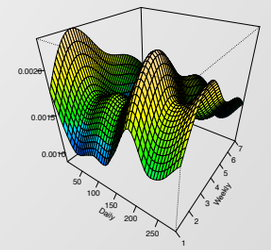
thecrix.de



DeFi B-Side = Pre-destined for machine rule ! (?)

- Hardware Visualisation (Moore's Law, Kryder's Law, Neuland Law)
- ABCD = DNA of DeFi
  - AI Artificial Intelligence (mimics human cognitive functions)
  - BC BlockChain, DLT Distributed Ledger Tech, SC Smart Contracts
  - Cloud Services (computing, storage)
  - Data Analytics (*Fourth Industrial Revolution*, SDA, IoT, etc.)
- CC Crypto-Currencies as DeFi

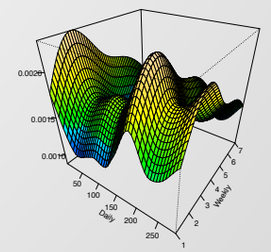




Data Sources

- CRIX 5 min data
  - Time frame: 20180601 - 20180930 (0 GMT)

- 8 CCs 5 min candles data (Europe-based source):
  - BCH, BTC, DASH, ETC, ETH, LTC, REP, STR
  - Time frame: 20180601 - 20180930 (+1 GMT)
  - Sources: dyos solutions GmbH dyos



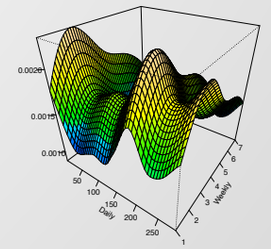
Generalized Additive Model (GAM)

- Hastie, T. J.; Tibshirani, R. J. (1990)

$$G\left\{E\left(y_i\right)\right\} = \beta_0 + f_1\left(x_{i1}\right) + \cdots + f_p\left(x_{ip}\right)$$

  - $y = \left(y_1, \ldots, y_n\right)^\top$ observation of dependent variable $\boldsymbol{Y}$, e.g. returns, volatility, volume
  - β_0 - an intercept
  - $(x_{i1}, \ldots, x_{ip})^\top$ predictor variables (daily and weekly effects)
  - link function $\boldsymbol{G}$ (identical, logarithmic or inverse)
  - $f_1\left(x_{i1}\right) \ldots f_p\left(x_{ip}\right)$ smooth functions (splines) with dimension k_q and basis $b_j^q(x)$ function

$$f_q(x) = \sum_{j=1}^{k_q} \beta_{q,j} b_j^q(x)$$



GAM ctd.

- Loss function

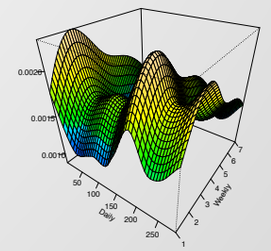
$$\sum_{i=1}^{n}\left(y_i - \sum_{q=1}^{p} f_q\left(x_i\right)\right)^2 + \sum_{q=1}^{p}\lambda_q \int \left\| f_q''(x) \right\|^2 dx$$

where λ_q is a smoothing parameter

$$\widehat{\beta} = \arg\min_{\lambda,\beta}\left\{ \|Y - B\beta\|^2 + \sum_{q=1}^{p}\lambda_q \beta^{\top} S_q \beta \right\}$$

where S_q is a matrix of known coefficients and B is a matrix of $b_j^q(x)$

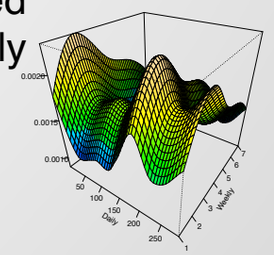
- Method - P-IRLS: Penalized Iteratively Re-weighted Least Squares



General Statistics

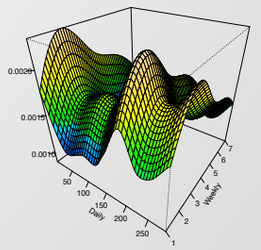
	$\widehat{\rho}_1(ret_t)$	$\widehat{\rho}_1(ret_t^2)$	$\widehat{\rho}_1(\lvert ret_t \rvert)$	$\widehat{S}$	$\widehat{e.Kurt}$	JB
BCH	-0.01	0.12	0.20	0.49	13.69	140148.24
BTC	-0.05	0.13	0.24	1.30	49.44	1823779.80
DASH	0.01	0.17	0.20	0.73	28.98	626596.64
ETC	-0.06	0.26	0.26	0.70	26.07	507374.39
ETH	-0.01	0.18	0.27	0.17	16.34	198777.58
LTC	-0.01	0.11	0.19	0.44	14.91	166121.81
REP	-0.08	0.22	0.19	0.35	21.89	356937.91
STR	-0.09	0.12	0.18	0.28	8.12	49354.96
XMR	-0.07	0.13	0.14	0.03	10.51	82241.48
XRP	-0.05	0.17	0.25	0.11	11.44	97390.58
ZEC	-0.07	0.25	0.22	1.30	26.66	534032.89

Estimated first order autocorrelation of the returns, squared returns, absolute returns, estimated skewness, estimated excess kurtosis, and the Jarque-Bera test statistic for the overall summed intraday high-frequency data from the 01. July 2018 to the 31. August 2018.

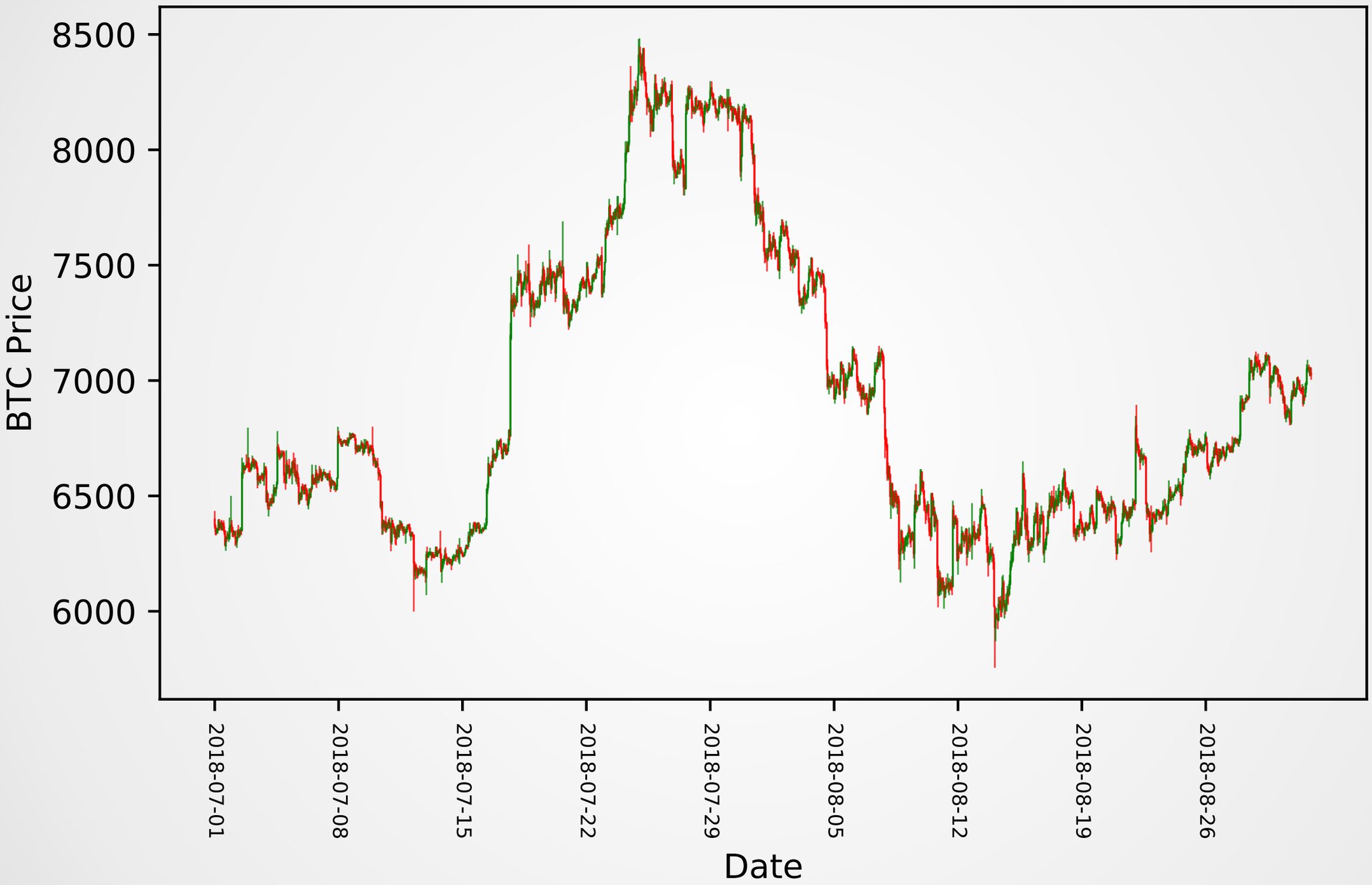


Candlestick Chart - Sample data

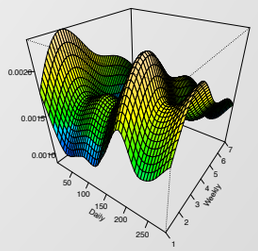
CRIX candlestick chart, 01.06.2018 - 30.09.2018

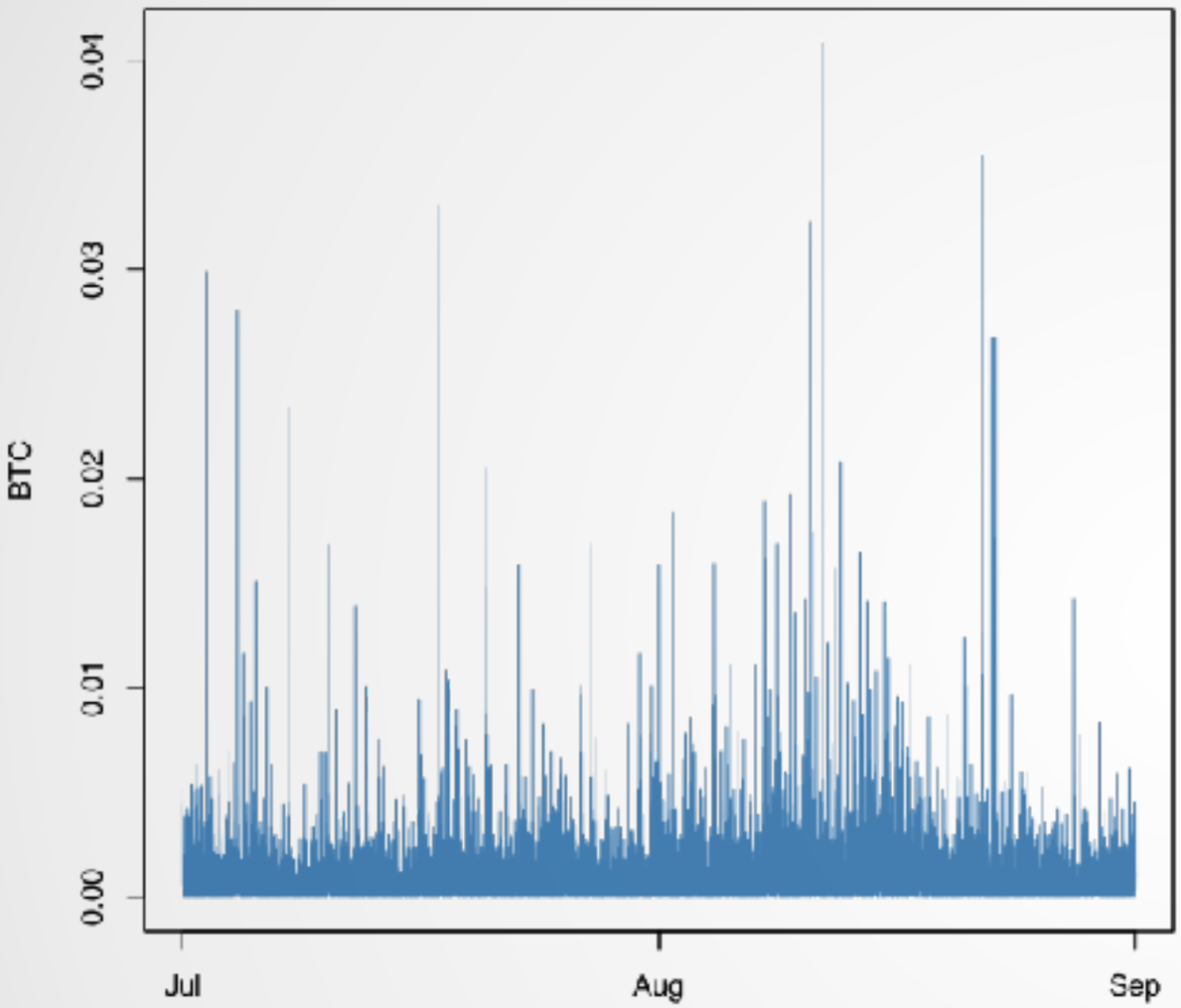


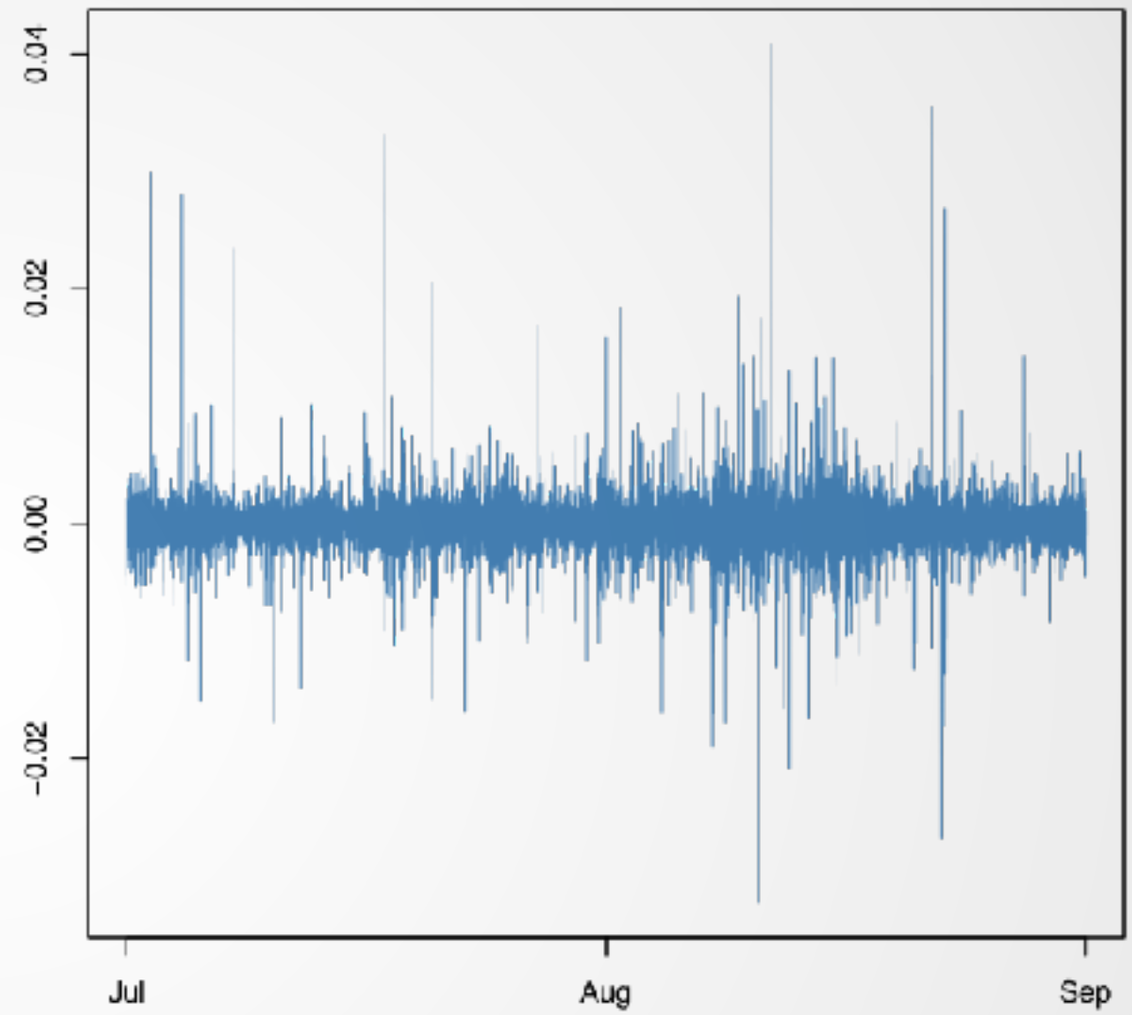
Trading characteristics: BTC



BTC candlestick chart, 01.06.2018 - 30.09.2018



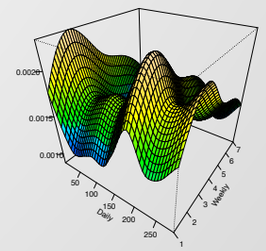




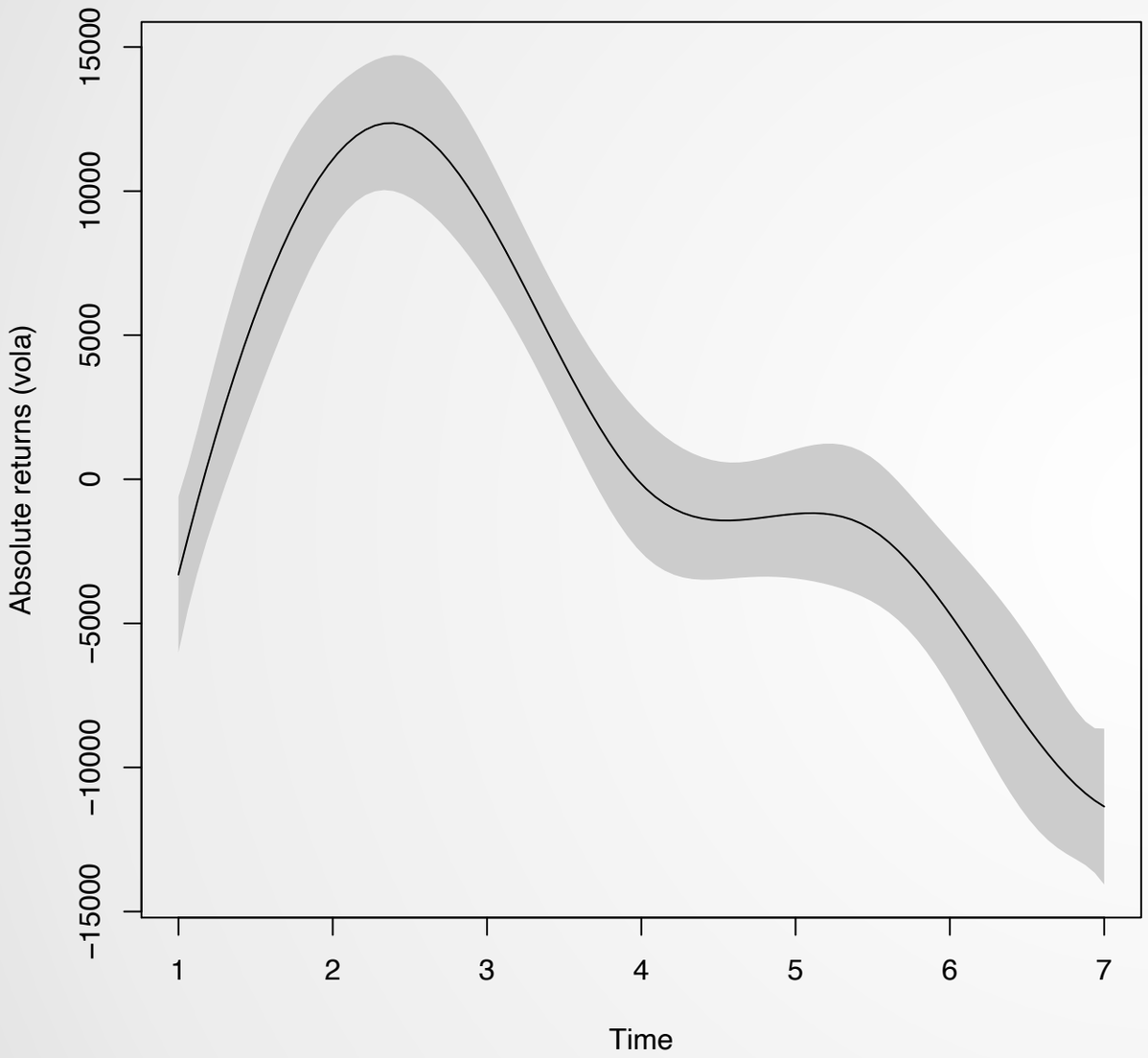
Volatility

Returns

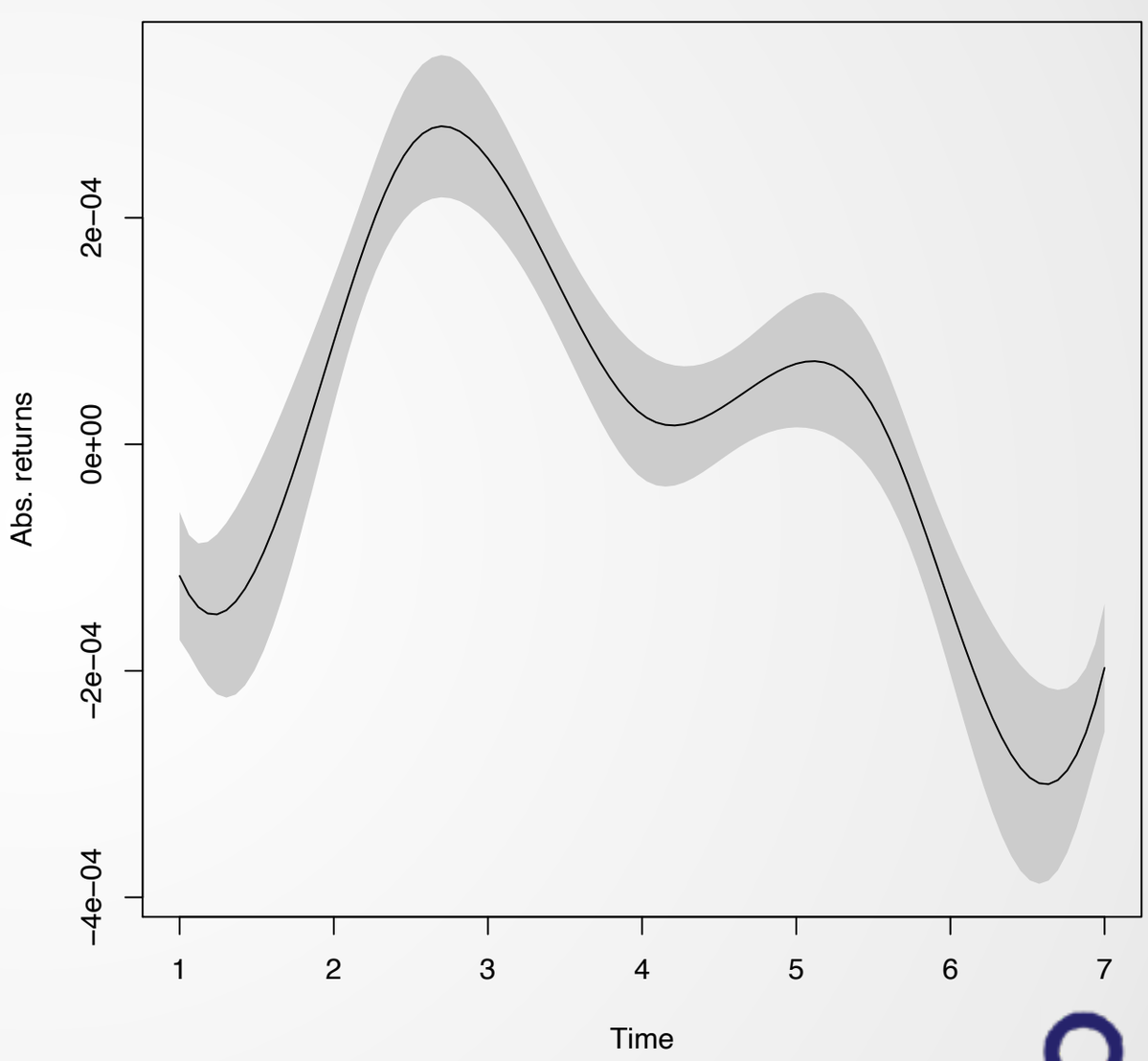
01.06.2018 - 30.09.2018



Trading characteristics: BTC

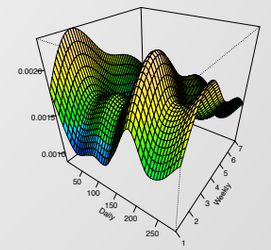


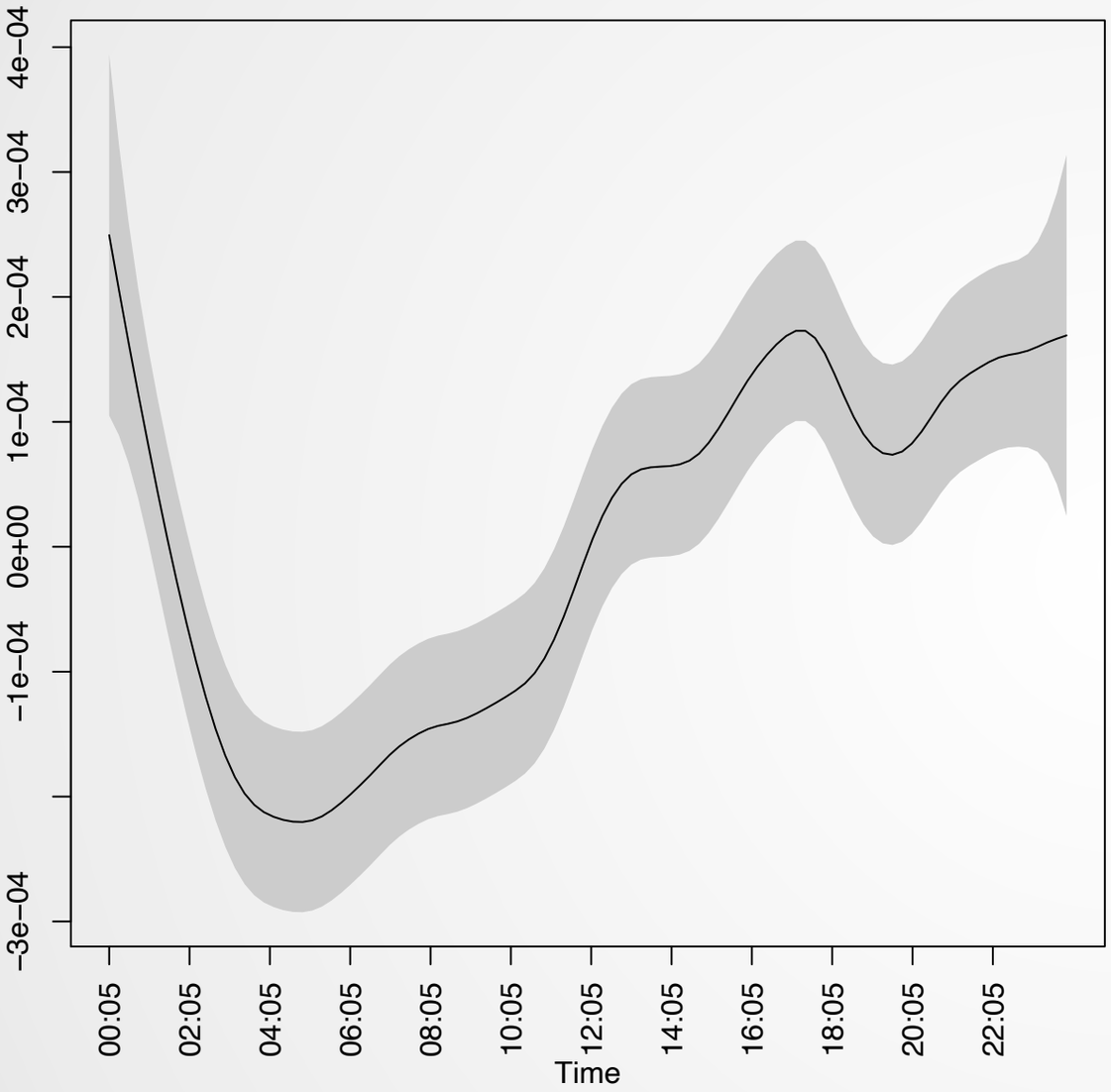
Weekly Volume (GAM)

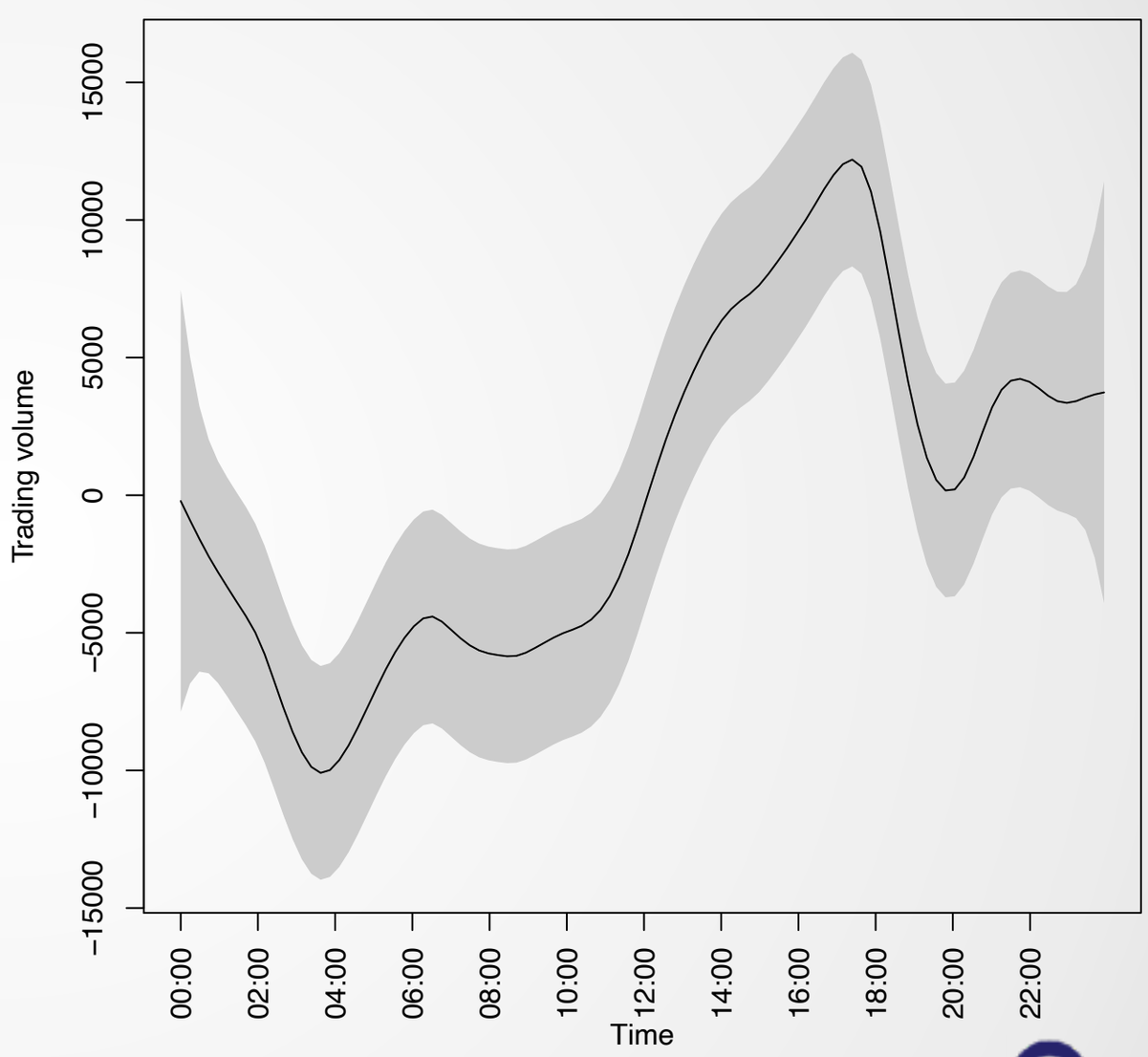


Weekly Vola (GAM)

01.06.2018 - 30.09.2018 (95% CI)



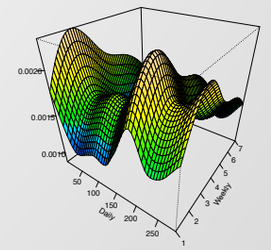


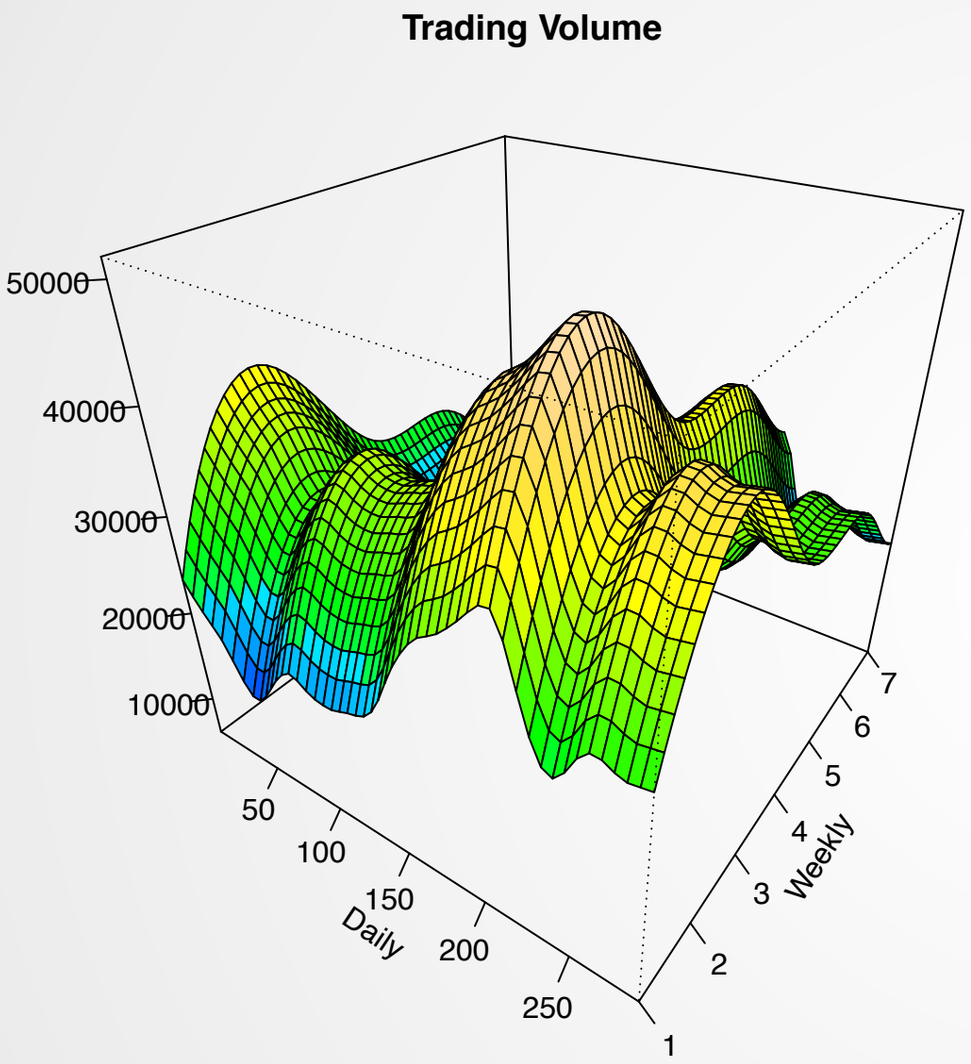


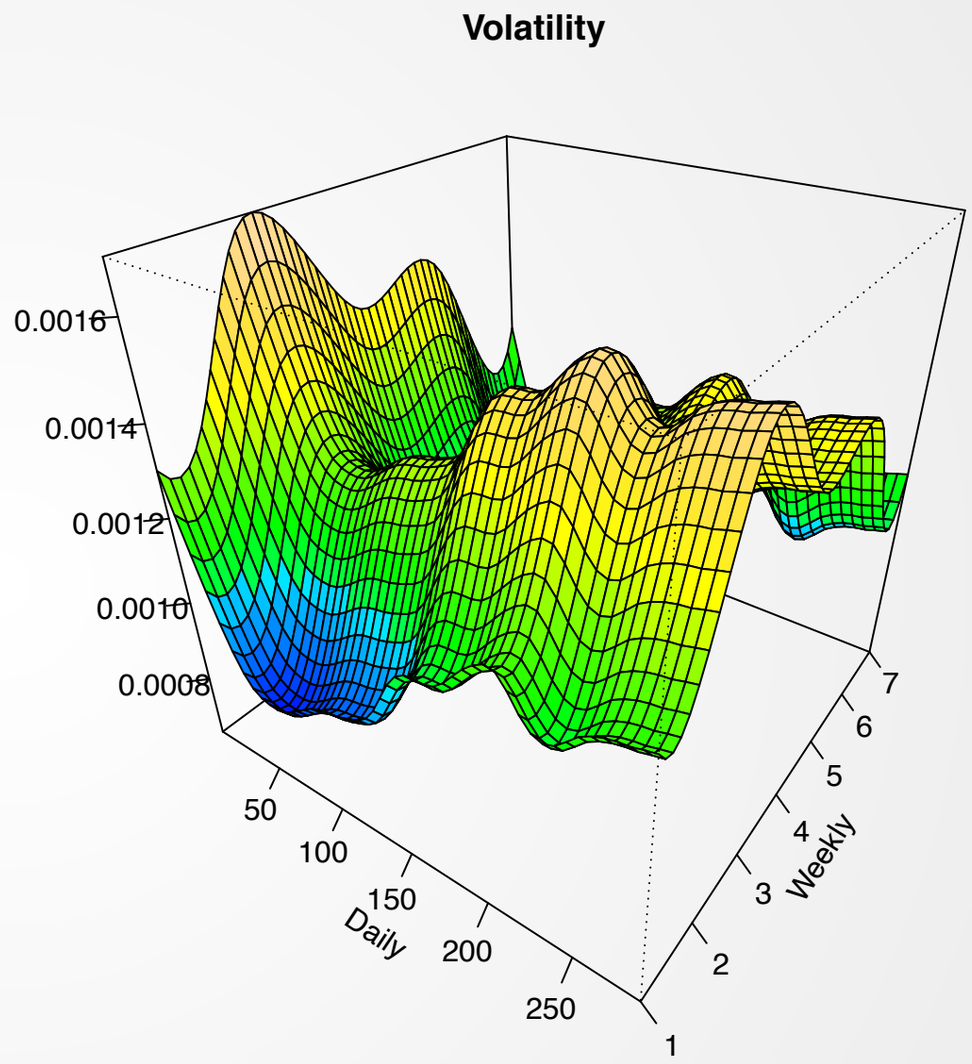
Daily Volume (GAM)

Daily Returns (GAM)

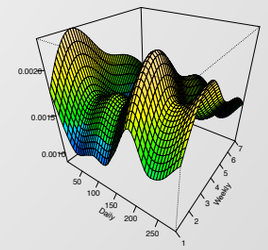
01.06.2018 - 30.09.2018 (95% CI)



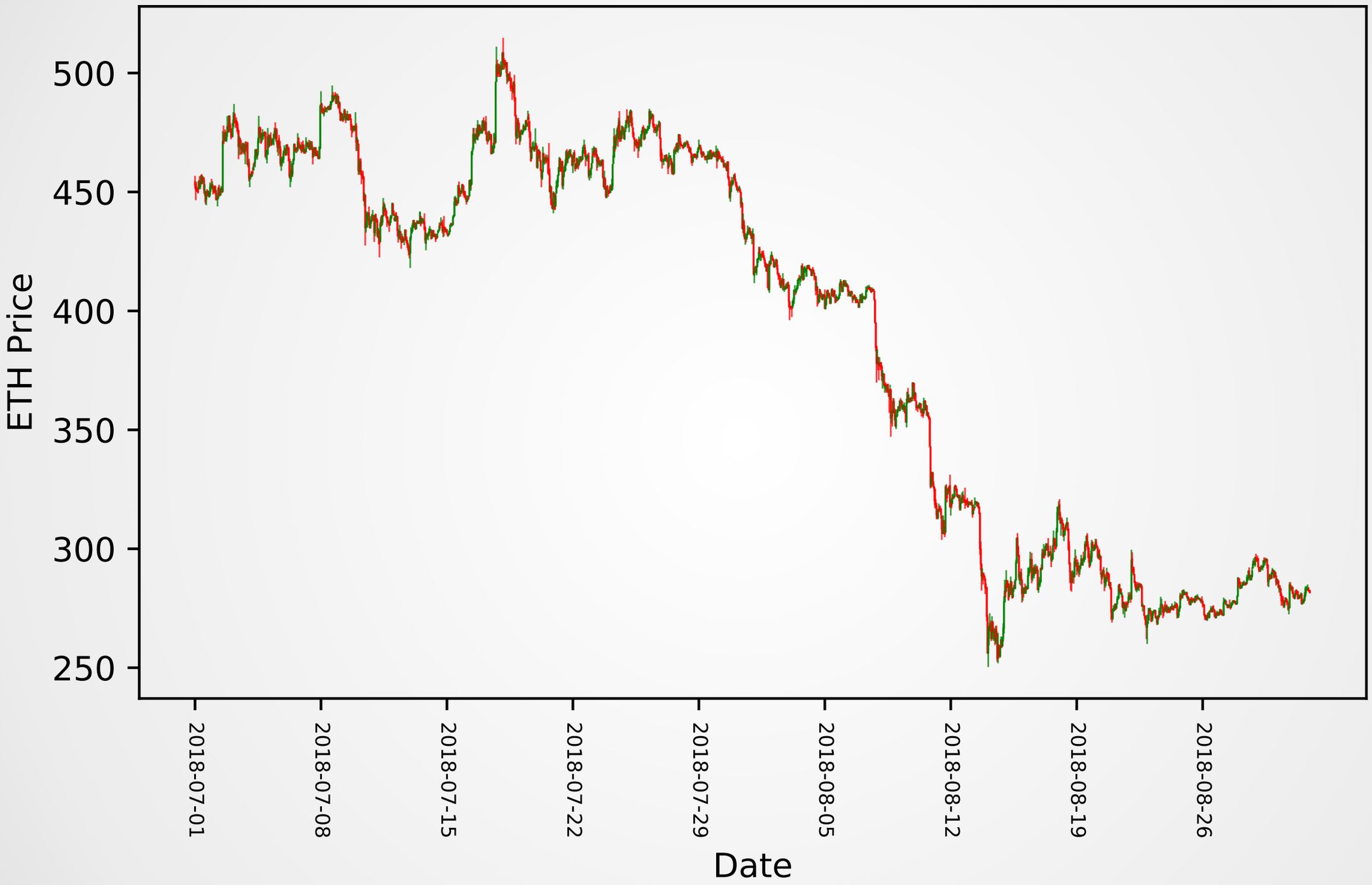




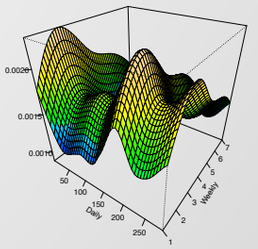
Daily Seasonality, 01.06.2018 - 30.09.2018

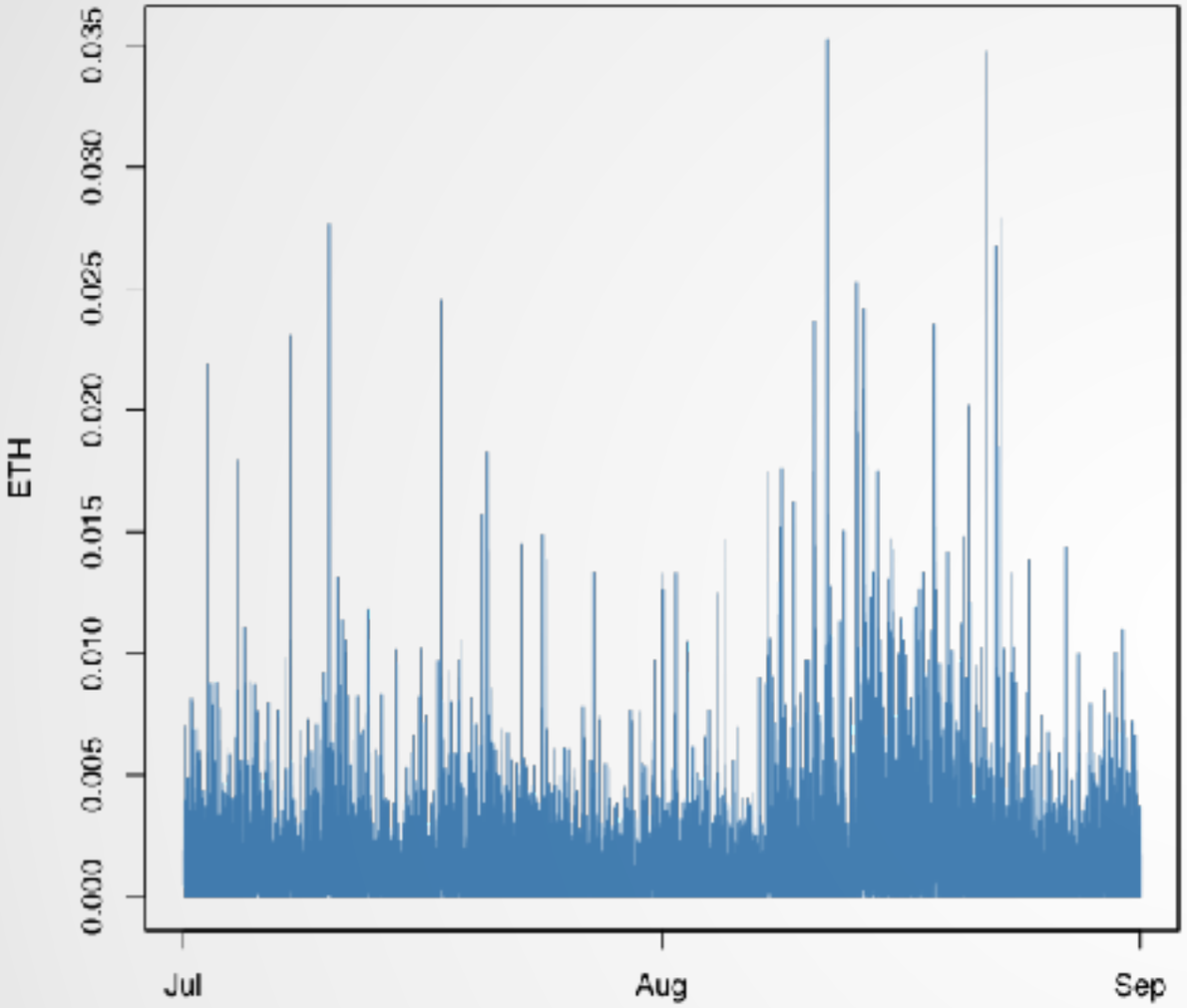


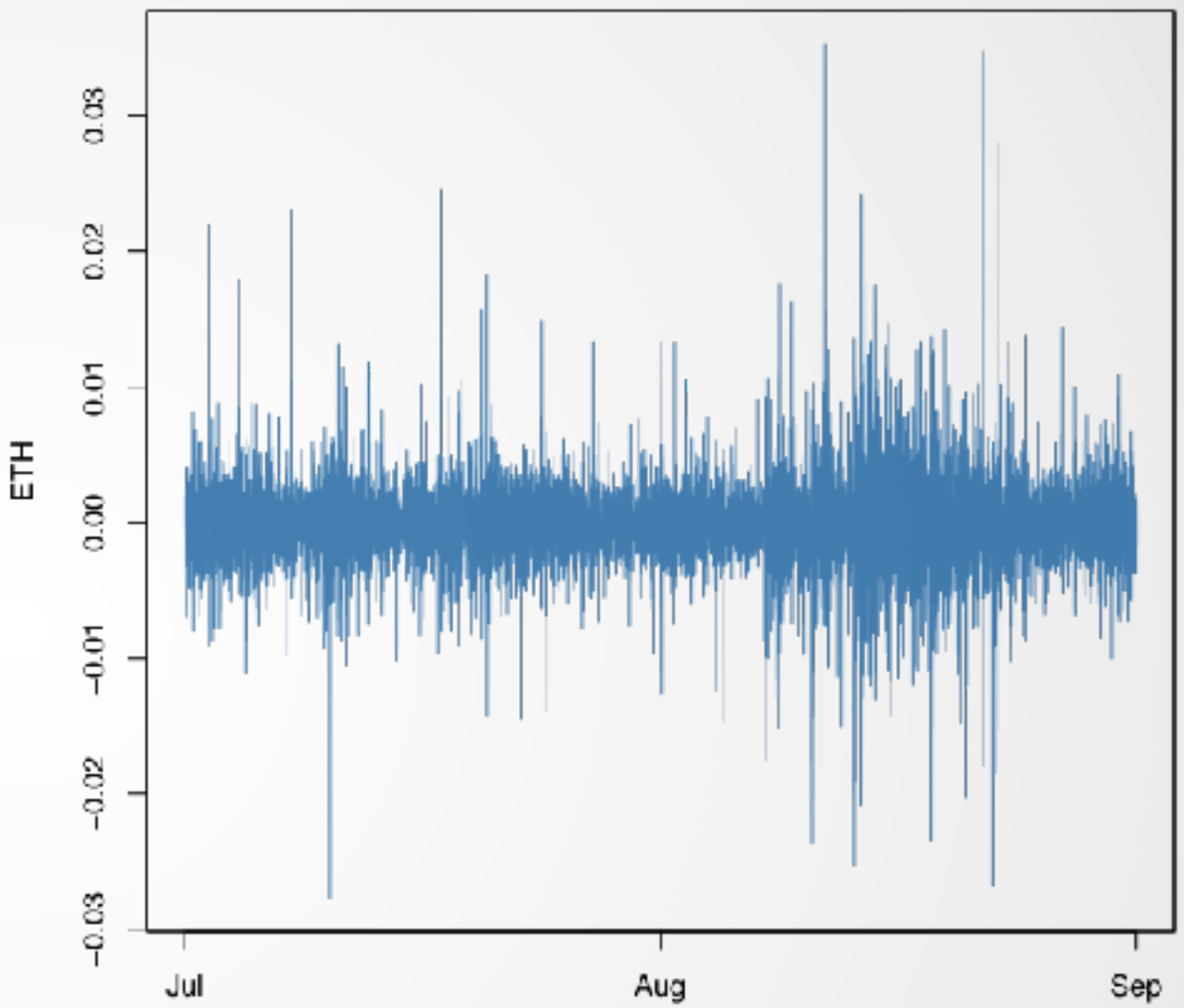
Trading characteristics: ETH



ETH candlestick chart, 01.06.2018 - 30.09.2018



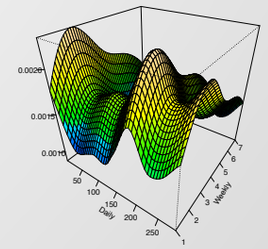


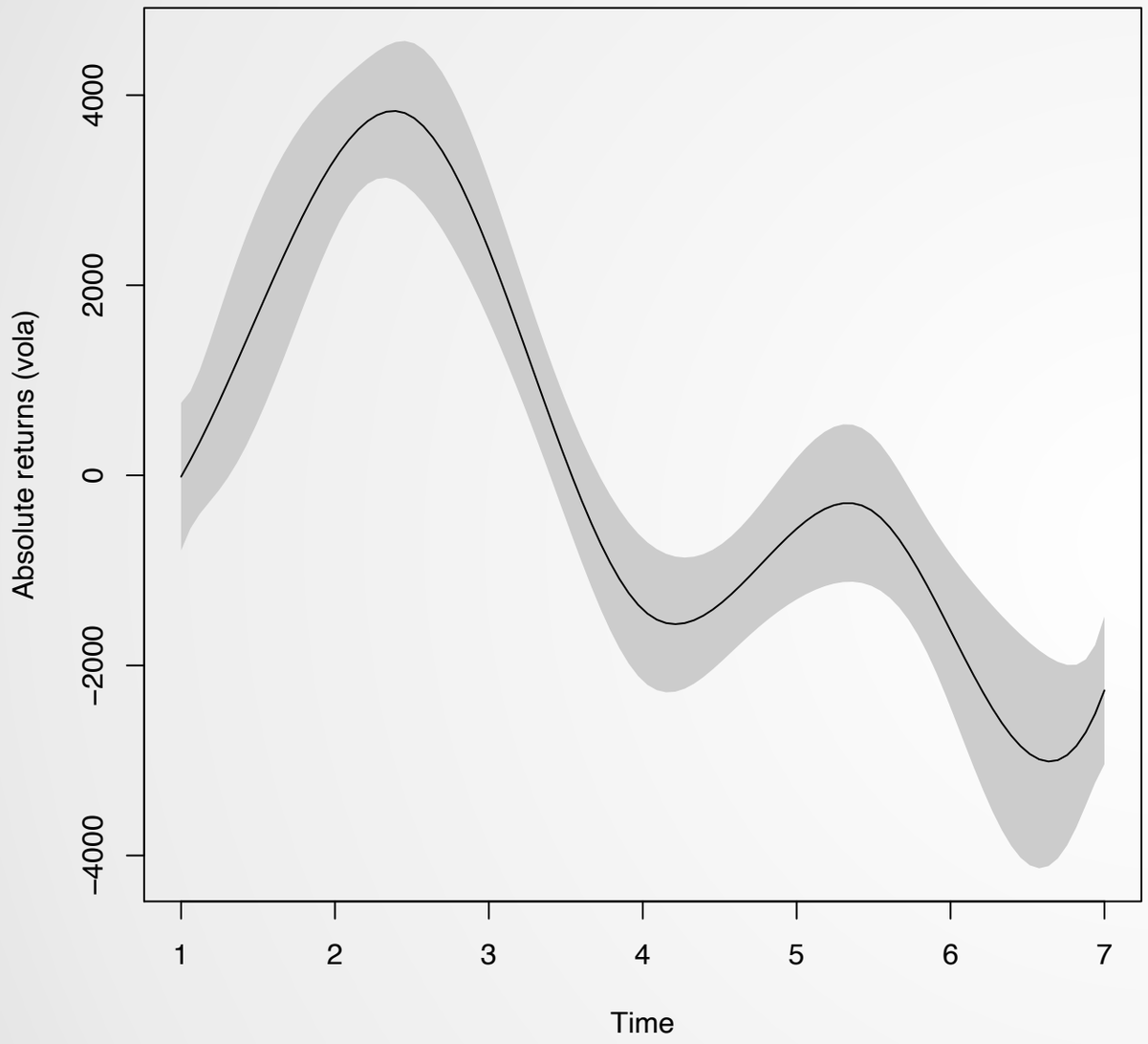


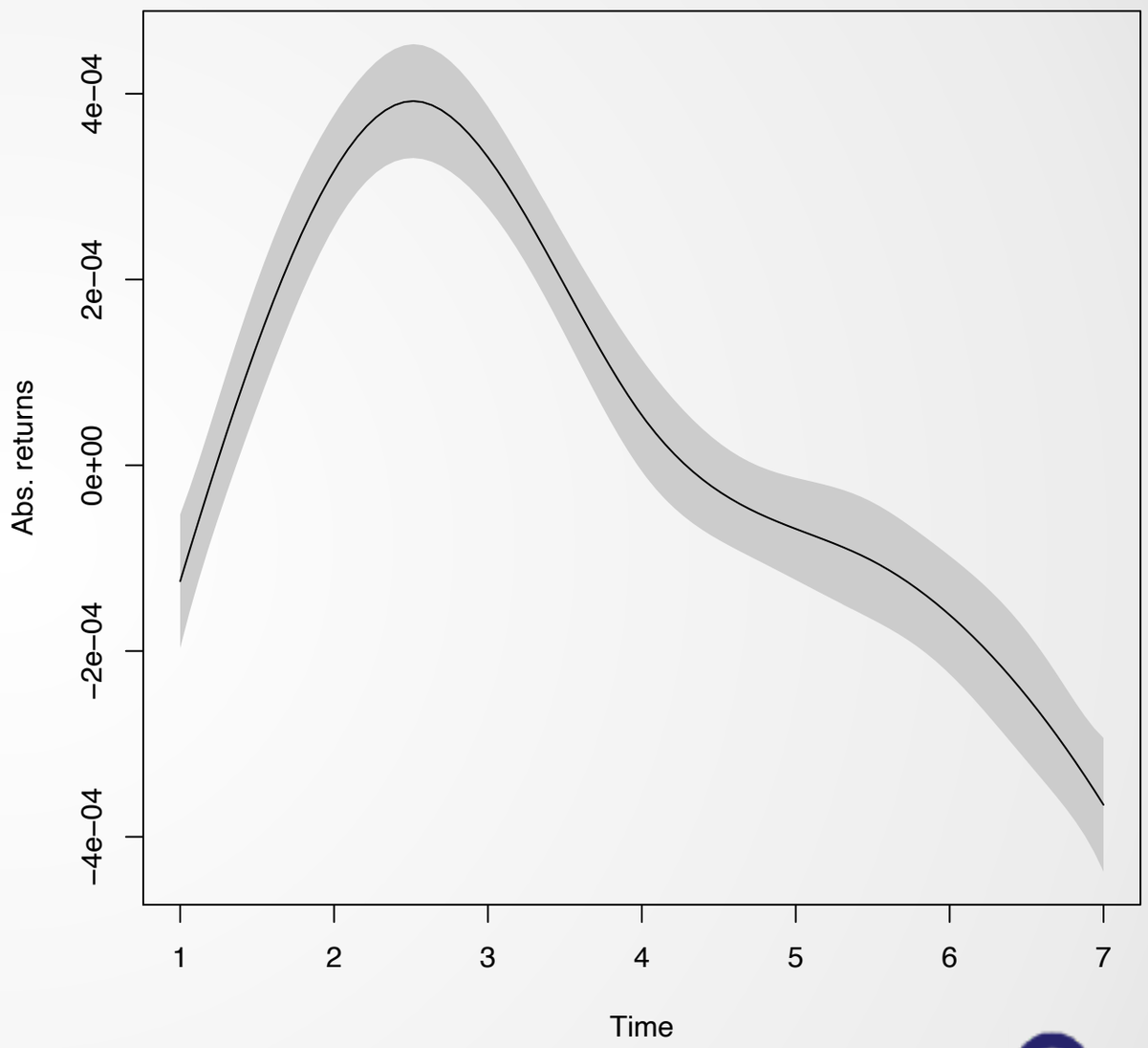
Volatility

Returns

01.06.2018 - 30.09.2018



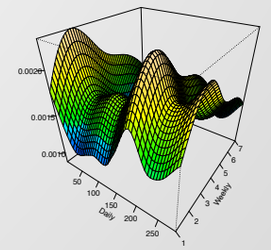




Weekly Volume (GAM)

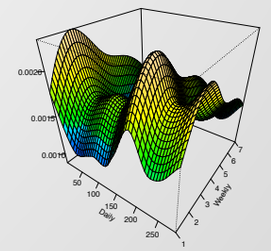
Weekly Returns (GAM)

01.06.2018 - 30.09.2018 (95% CI)

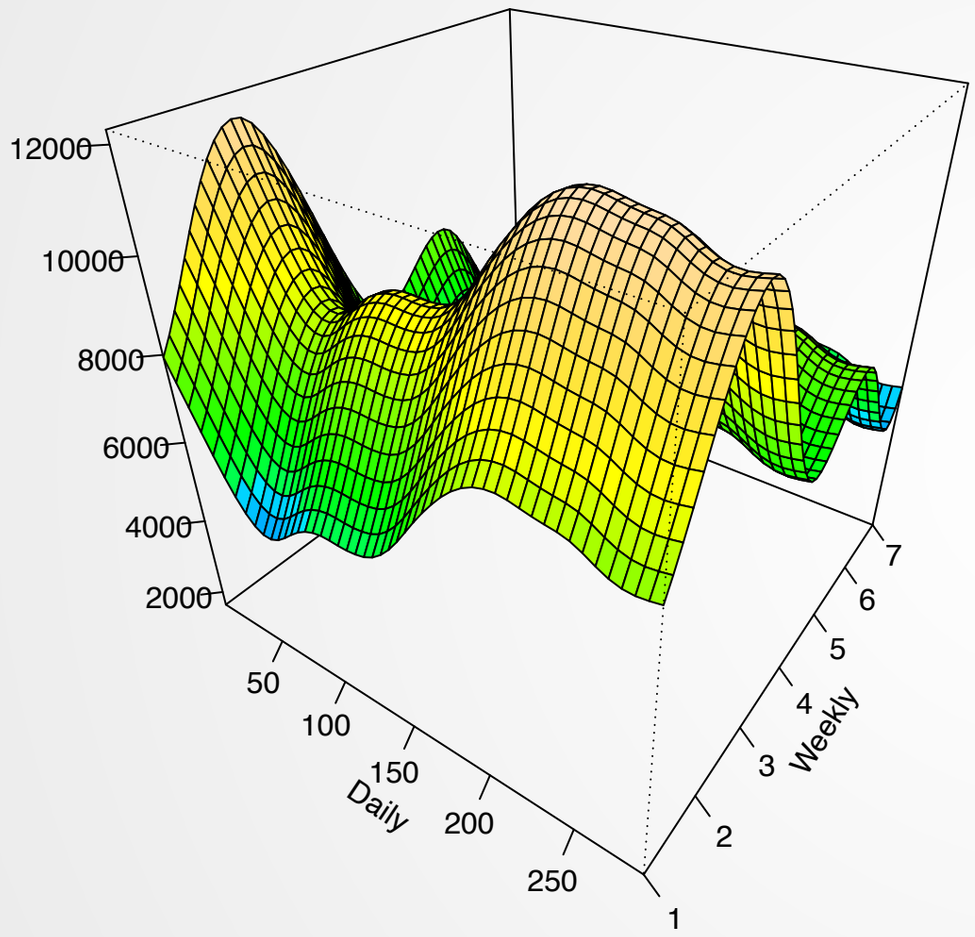


Daily Volume (GAM) Daily Returns (GAM)

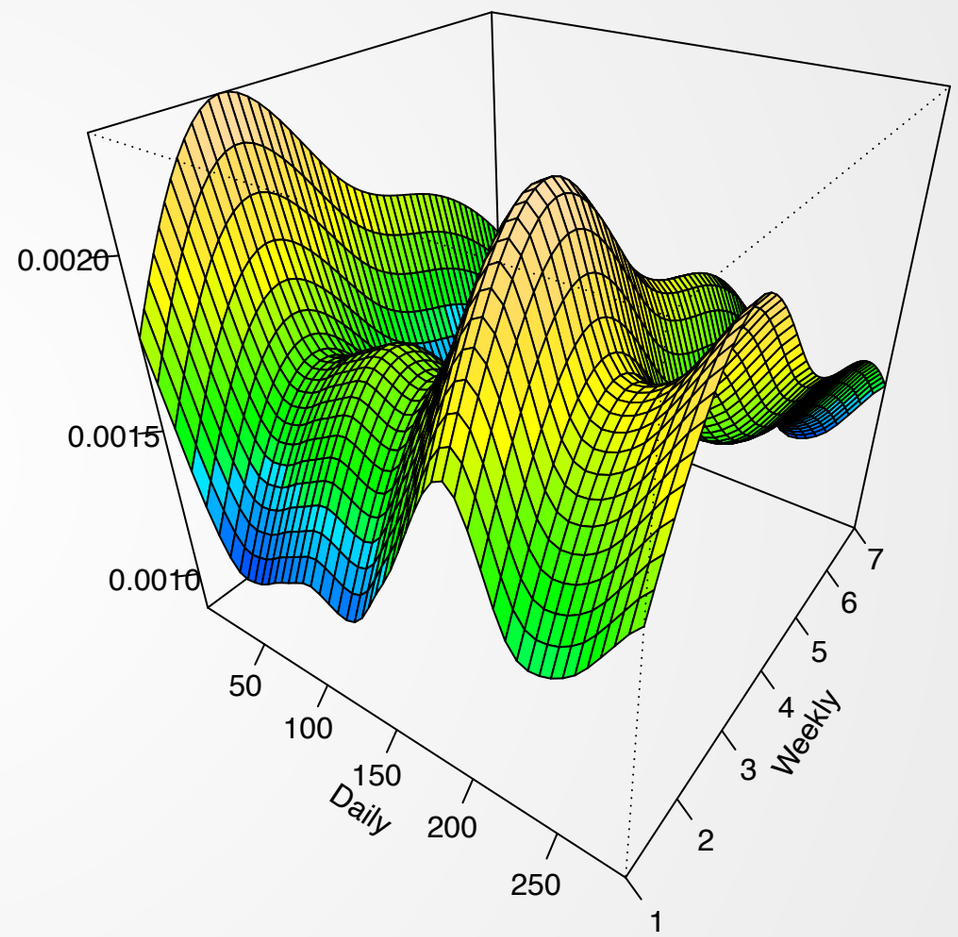
01.06.2018 - 30.09.2018 (95% CI)



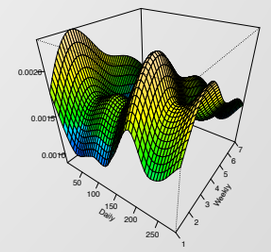
Trading Volume



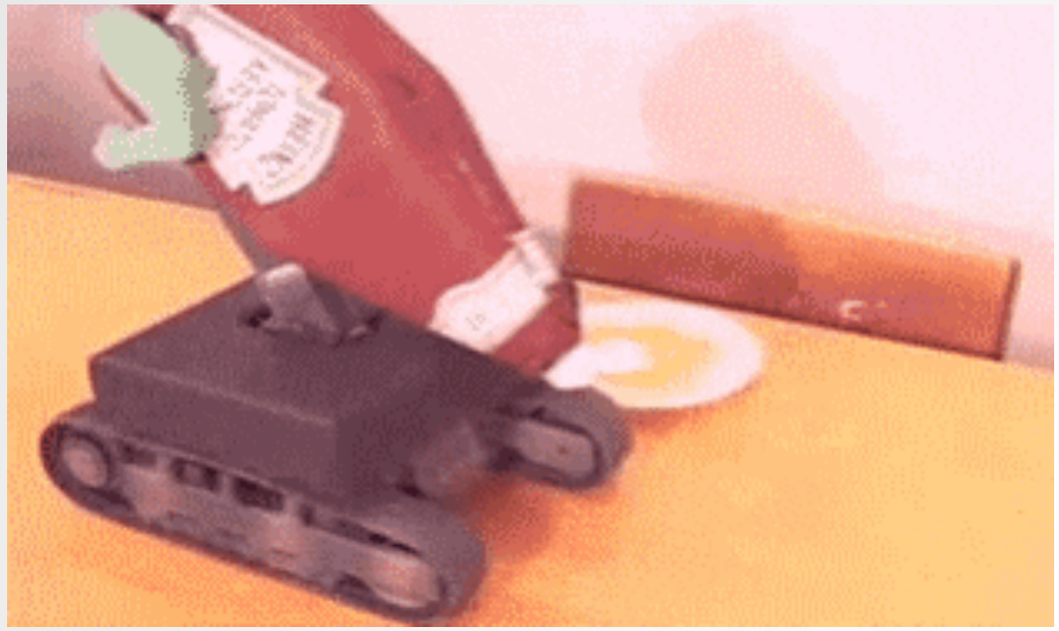
Volatility



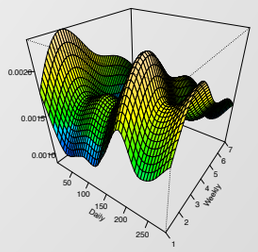
Daily Seasonality, 01.06.2018 - 30.09.2018



Conclusions

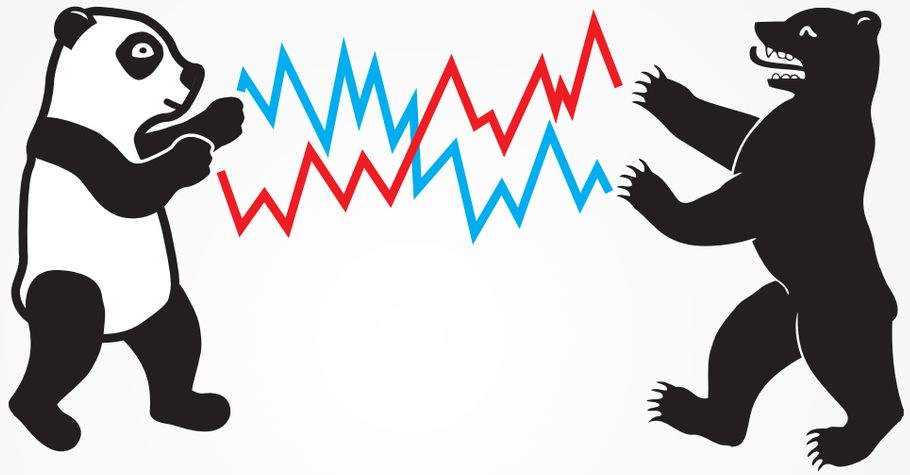


24/7 algorithmic cryptocurrency trade negligible.
People don’t trust the machines.

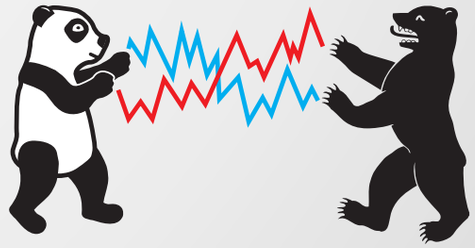


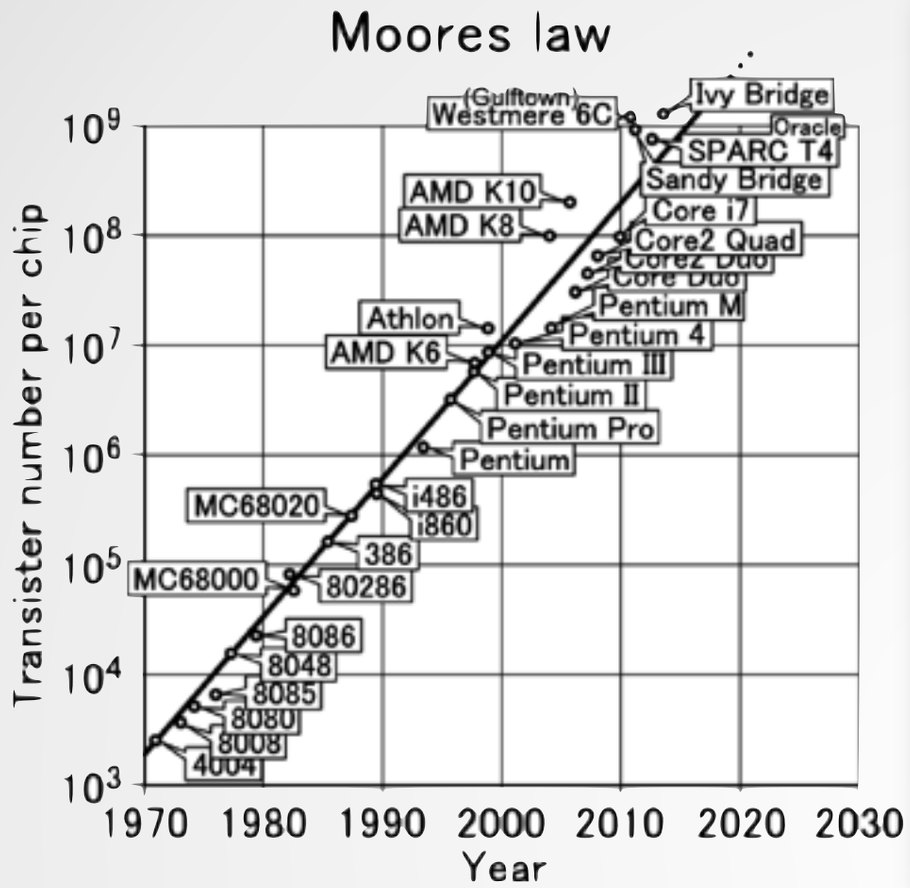
Thank you!

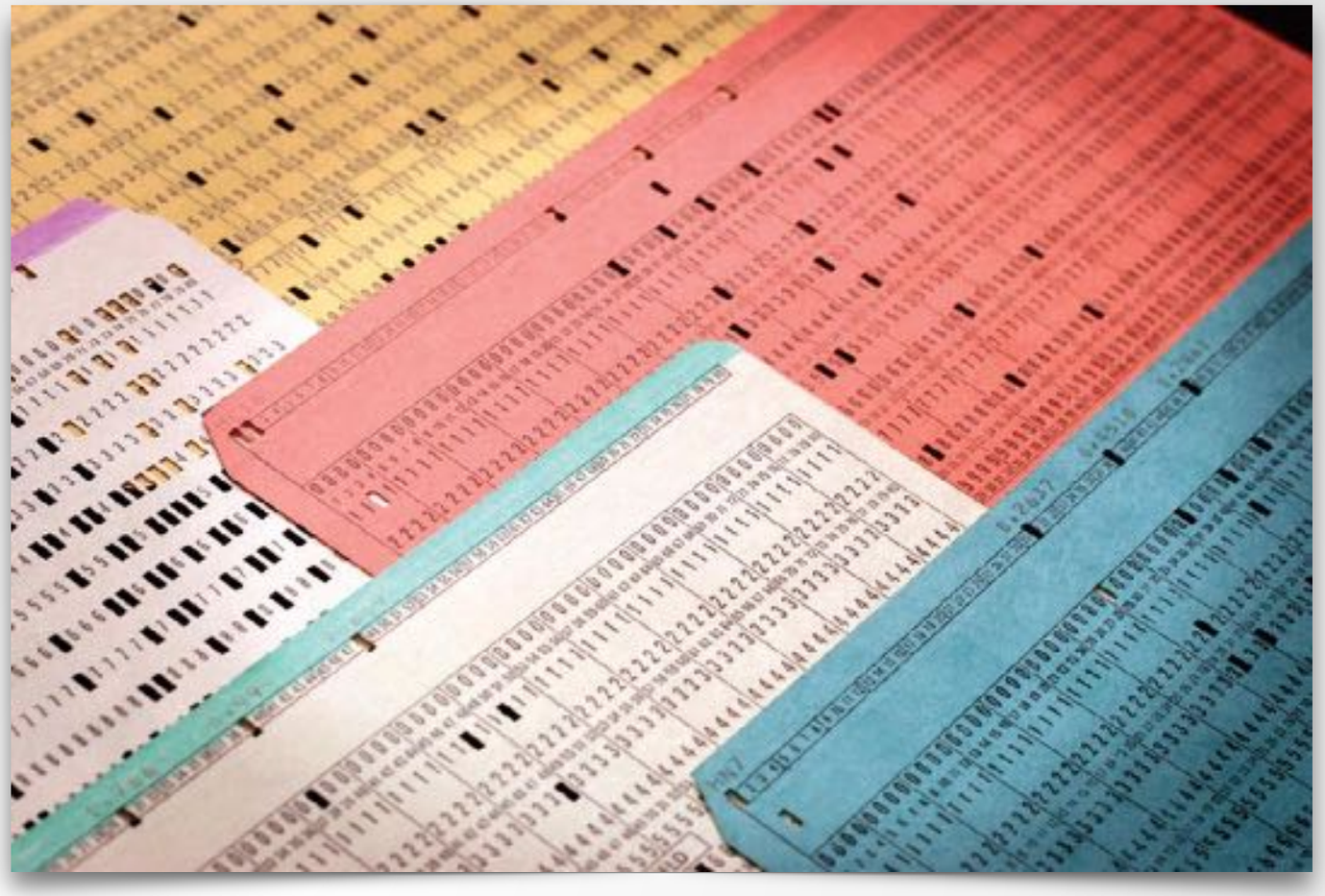
谢谢



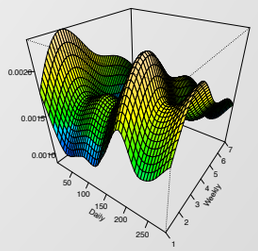
Ladislaus von Bortkiewicz Professor of Statistics
Humboldt-Universität zu Berlin
IRTG 1792, BRC
irtg1792.hu-berlin.de
Charles University, WISE XMU, NCTU 玉山学者



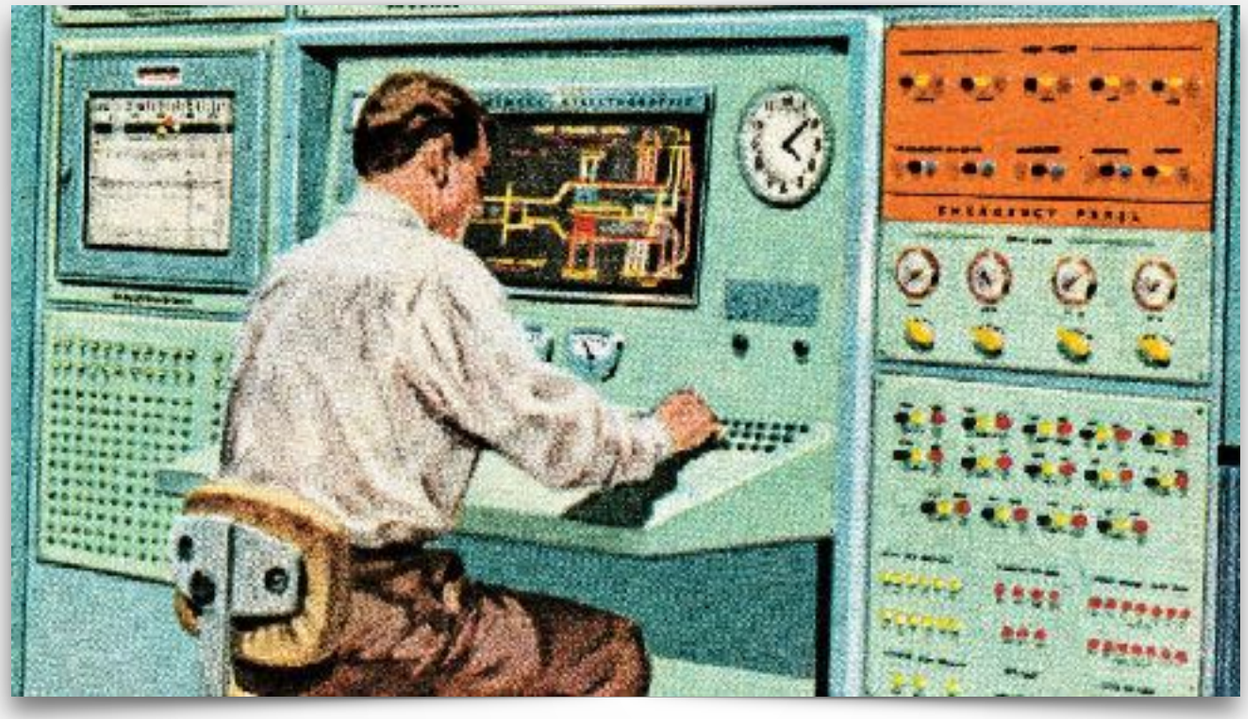




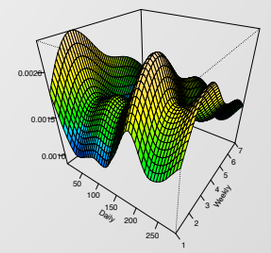
- Exponential increase in computing power
- Automation of whole industries
- 24/7 on-line services for and used by everyone ?



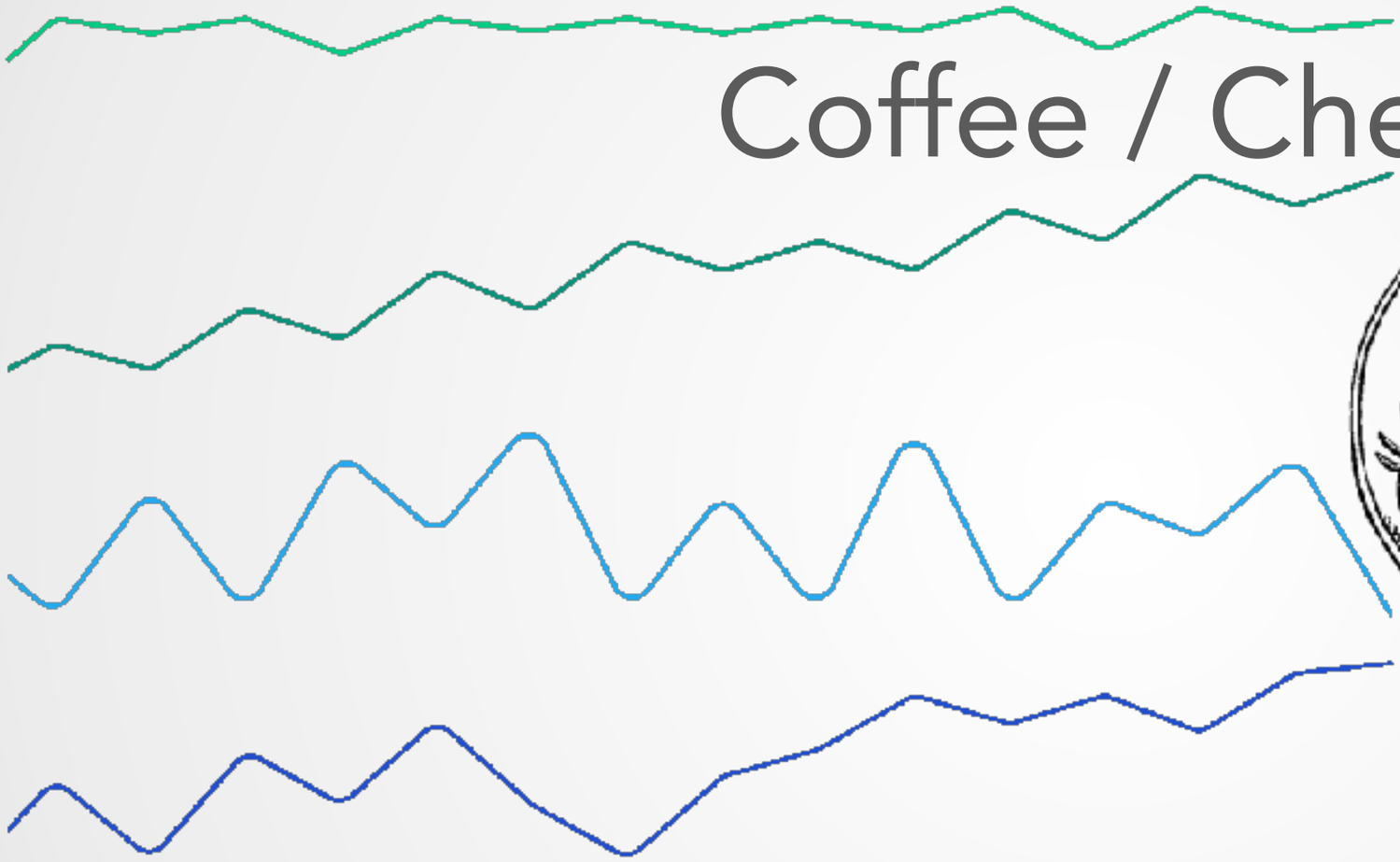
Algorithmic trading

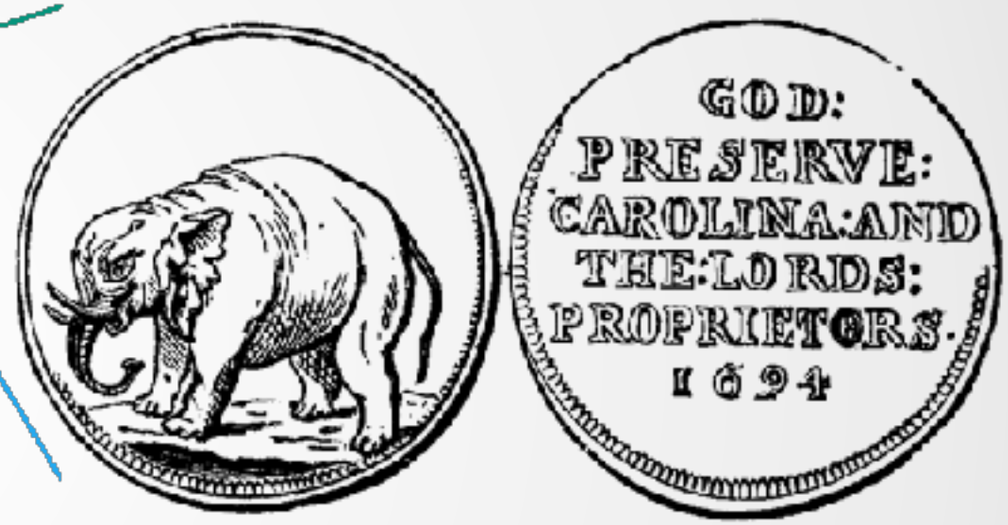


- Smart trading algorithms available, e.g. via Github
- Respective service providers available for 24/7 services
- Must entail machine made decisions ! (?)



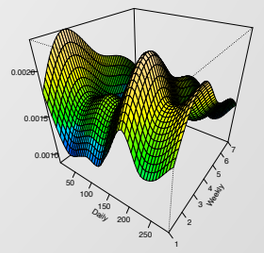
Elephant Test: Digital markets ruled by machines! (?)





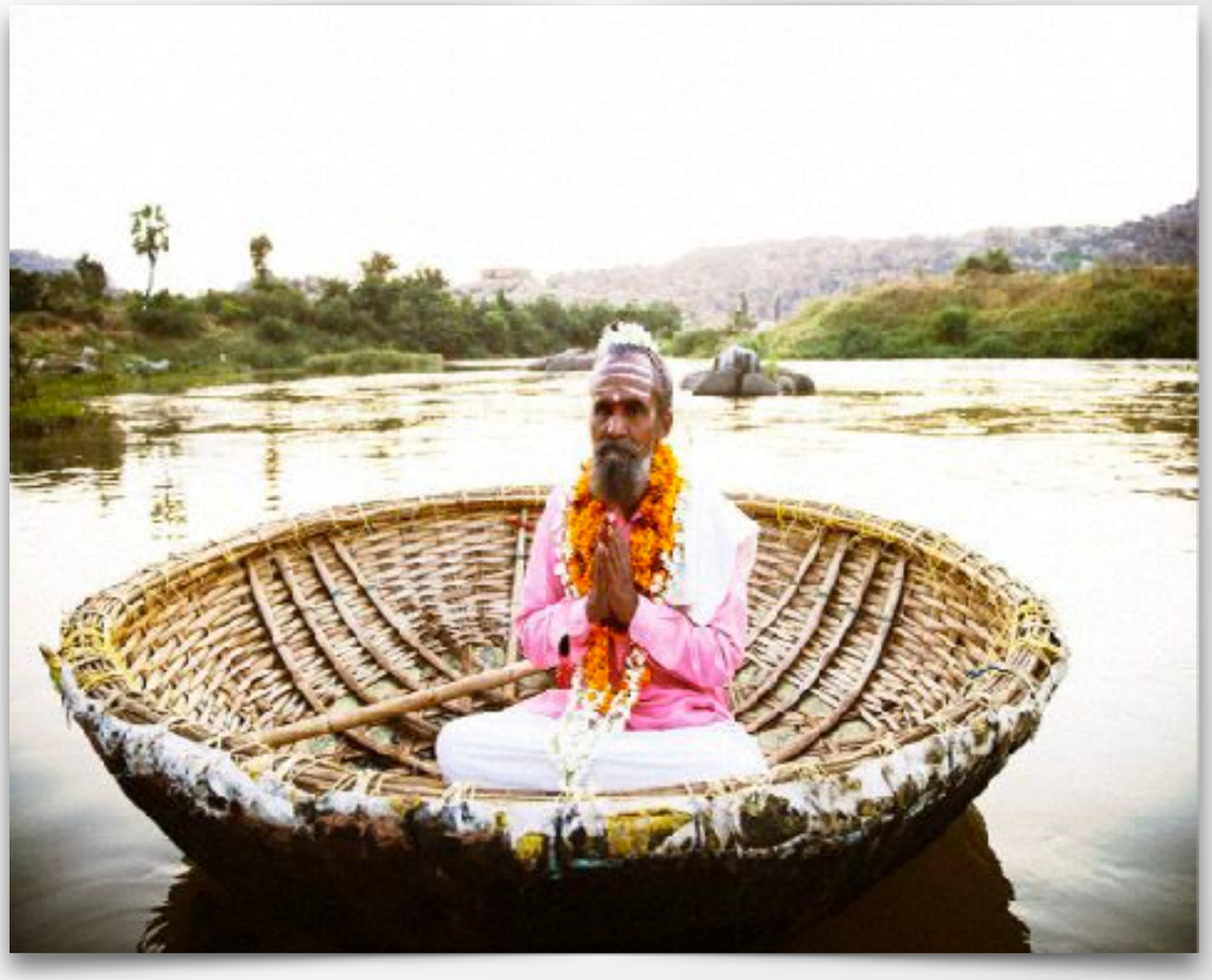
"The elephant, characterised more by recognition when encountered, than by definition"

Coin: Half Pence (1/2 pence) Obverse has the image of an elephant. Reverse bears the inscription.

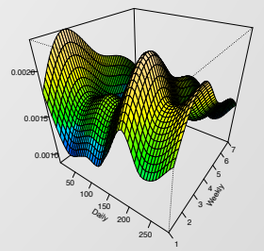


A priest and a scientist walk into a computer lab . . .

- commonly = *quant-monks* process info
- “Rechenknechte” (Gilbert Walker’s Indian helpers)
- “Computers” (i.e. women; HGL)



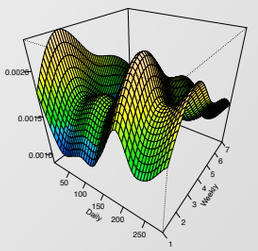
Fr. Roberto Busa (SJ): “The *difficult* we do right away; the *impossible* takes a little longer”.



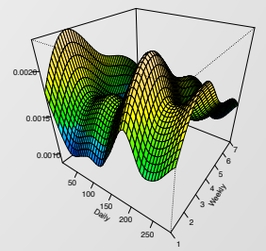
A priest and a scientist walk into a computer lab . . .

- do the *monks* process info?
- or do *machines* work autonomously?
- Wetware vs. Software

Father Roberto Busa: “The *difficult* we do right away; the *impossible* takes a little longer”.

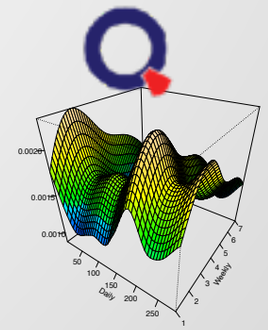


Intermezzo:
GAM or BEER?



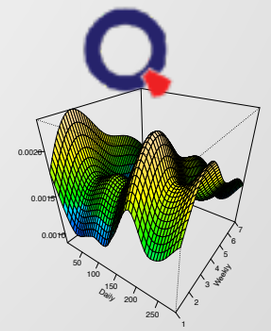
Correlation to CRIX with market movements upwards

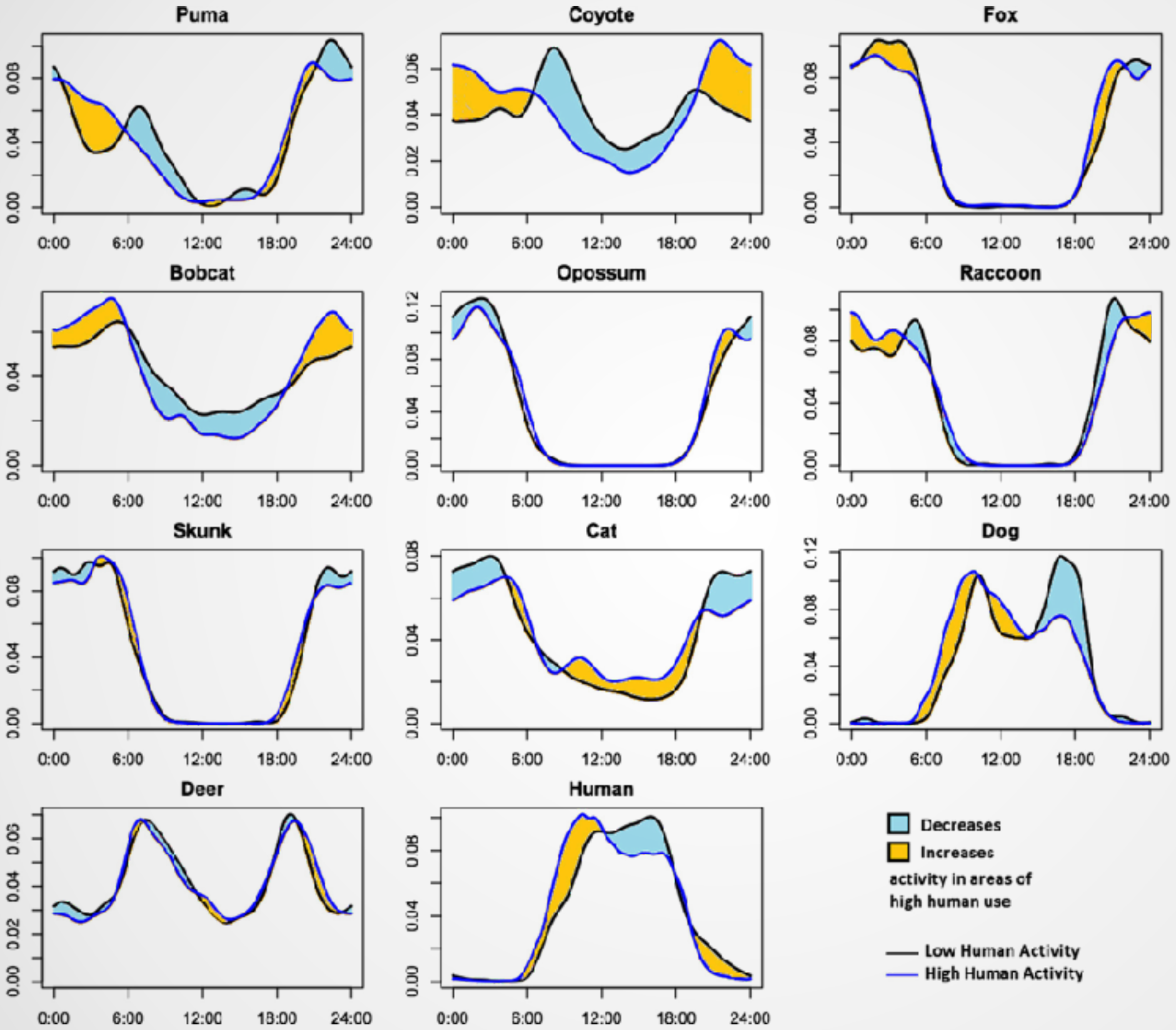
UP	BCH	BTC	DASH	ETC	ETH	LTC	REP	STR	XMR	XRP	ZEC
BCH		0.50	0.23	0.33	0.47	0.46	0.13	0.29	0.25	0.37	0.23
BTC	0.50		0.27	0.36	0.55	0.49	0.18	0.34	0.30	0.40	0.27
DASH	0.23	0.27		0.17	0.22	0.22	0.10	0.17	0.17	0.22	0.14
ETC	0.33	0.36	0.17		0.37	0.31	0.11	0.21	0.17	0.28	0.14
ETH	0.47	0.55	0.22	0.37		0.47	0.16	0.30	0.27	0.42	0.22
LTC	0.46	0.49	0.22	0.31	0.47		0.17	0.26	0.25	0.39	0.23
REP	0.13	0.18	0.10	0.11	0.16	0.17		0.12	0.11	0.11	0.11
STR	0.29	0.34	0.17	0.21	0.30	0.26	0.12		0.18	0.27	0.19
XMR	0.25	0.30	0.17	0.17	0.27	0.25	0.11	0.18		0.20	0.15
XRP	0.37	0.40	0.22	0.28	0.42	0.39	0.11	0.27	0.20		0.19
ZEC	0.23	0.27	0.14	0.14	0.22	0.23	0.11	0.19	0.15	0.19	

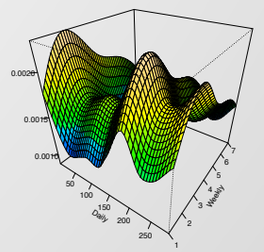


Correlation to CRIX with market movements downwards

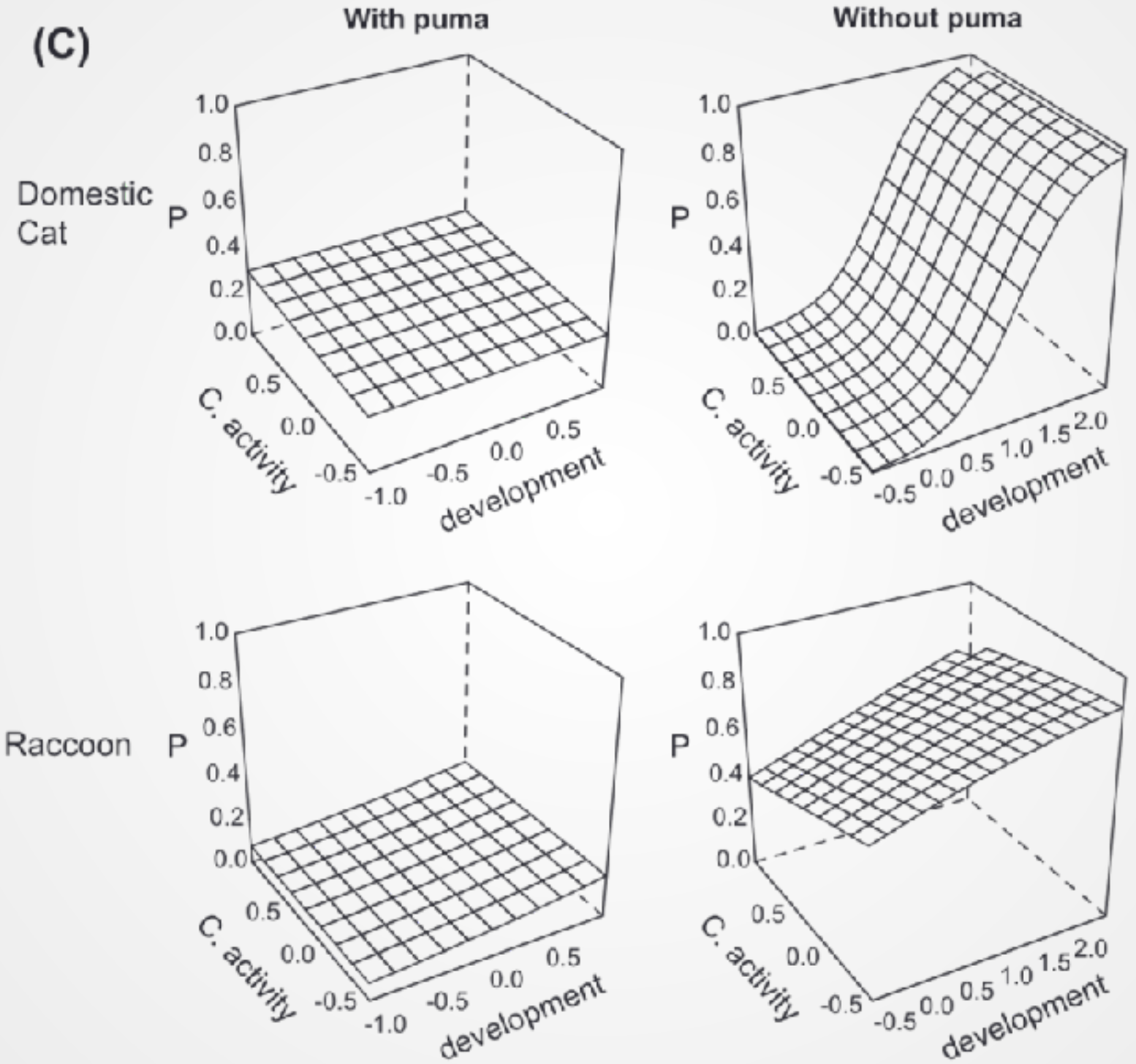
DOWN	BCH	BTC	DASH	ETC	ETH	LTC	REP	STR	XMR	XRP	ZEC
BCH		0.48	0.21	0.32	0.47	0.43	0.15	0.27	0.23	0.37	0.22
BTC	0.48		0.26	0.36	0.52	0.45	0.19	0.33	0.30	0.41	0.24
DASH	0.21	0.26		0.15	0.22	0.21	0.11	0.16	0.18	0.18	0.14
ETC	0.32	0.36	0.15		0.36	0.30	0.14	0.21	0.18	0.30	0.16
ETH	0.47	0.52	0.22	0.36		0.42	0.16	0.29	0.23	0.40	0.21
LTC	0.43	0.45	0.21	0.30	0.42		0.16	0.26	0.24	0.35	0.19
REP	0.15	0.19	0.11	0.14	0.16	0.16		0.11	0.12	0.13	0.08
STR	0.27	0.33	0.16	0.21	0.29	0.26	0.11		0.16	0.26	0.16
XMR	0.23	0.30	0.18	0.18	0.23	0.24	0.12	0.16		0.20	0.15
XRP	0.37	0.41	0.18	0.30	0.40	0.35	0.13	0.26	0.20		0.17
ZEC	0.22	0.24	0.14	0.16	0.21	0.19	0.08	0.16	0.15	0.17	



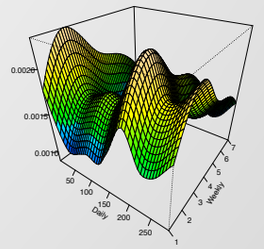




https://doi.org/10.1016/j.biocon.2015.05.007



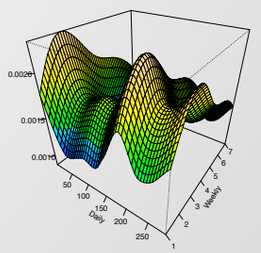
Algo-Puma (left) vs. Human-Puma (Right)



CC's = DeFi

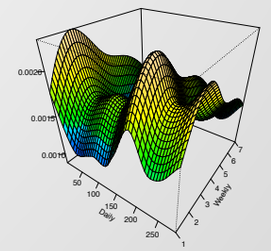
(a) Centralized (b) Decentralized (c) Distributed

Types of (finance) networks



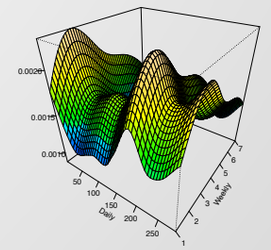
ABCD-DNA:

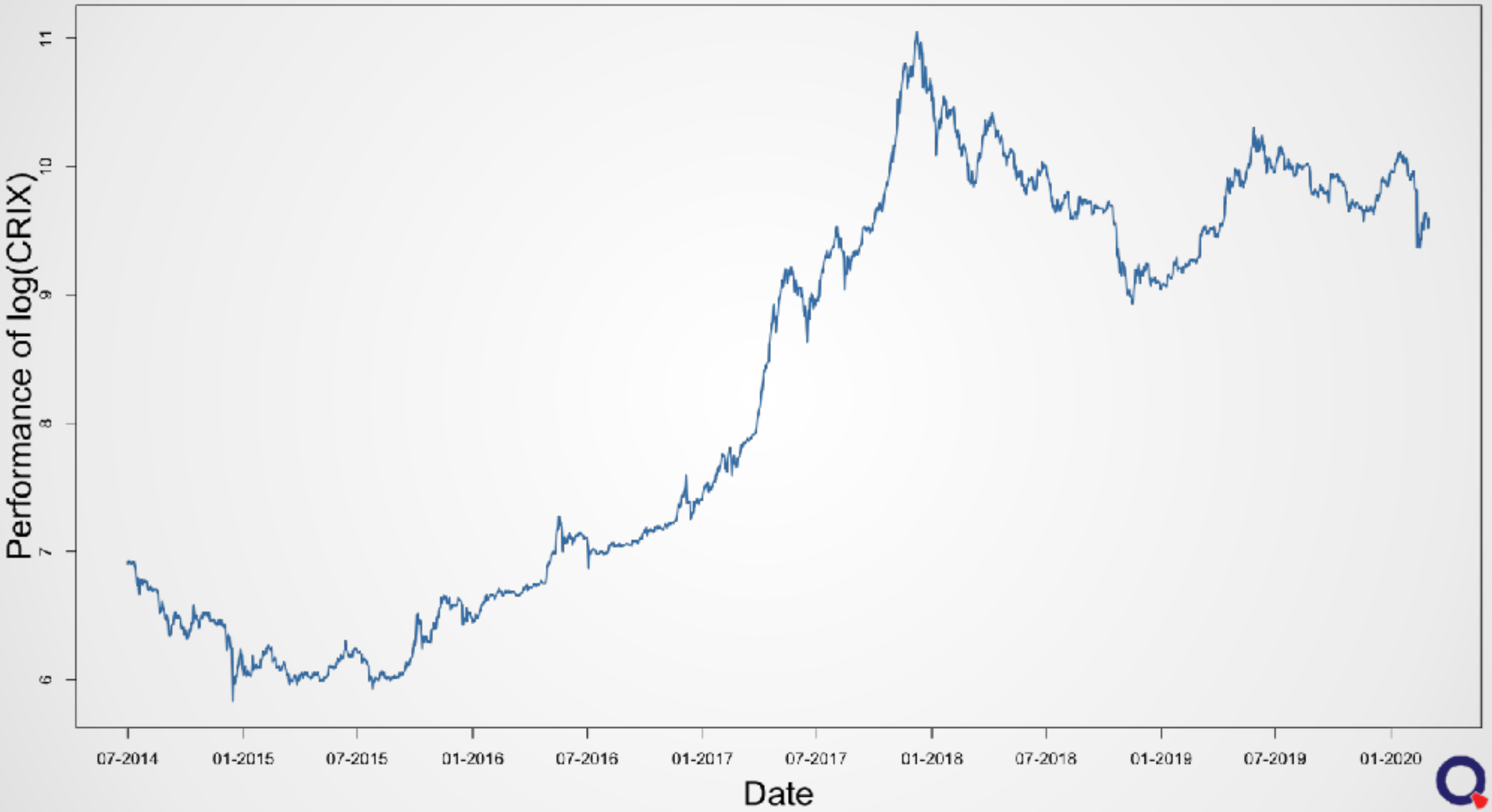
- *on-demand* structures
- AI = software that mimics human cognitive functions
- BC, DL, SC = Distributed Ledger Technologies (DLT)
- CS = cloud computing, decentralization of server capacity & availability of data storage, processing power
- Data = *Fourth Industrial Revolution*, SDA, IoT, etc.



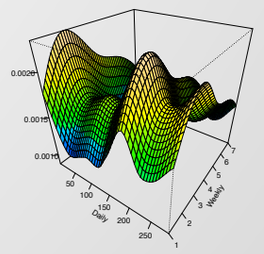
Hardware virtualization:

- software is hosted, updated, run decentrally
- only specific data that *needs* to be processed locally under specified conditions remains to be processed locally (e.g. hardware-bound access like hardware CC wallets)
- creation & set-up of DeFi architecture (“software as a service”)




Performance of log(CRIX)
6
7
8
9
10
11
07-2014
01-2015
07-2015
01-2016
07-2016
01-2017
07-2017
01-2018
07-2018
01-2019
07-2019
01-2020
Date

thecrix.de



Generalized Additive Model (GAM)

- Hastie, T. J.; Tibshirani, R. J. (1990)

$$G\left\{E\left(y_i\right)\right\} = \beta_0 + f_1\left(x_{i1}\right) + \cdots + f_p\left(x_{ip}\right)$$

  - $\boldsymbol{Y}$ dependent variable, e.g. returns, volatility, volume
  - β_0 regression coefficient estimated by multiple regression
  - x_p predictor variables (daily and weekly effects)
  - link function $\boldsymbol{G}$ (identical, logarithmic or inverse)
  - $i = 1 \ldots n$
  - $f_1\left(x_{i1}\right) \ldots f_p\left(x_{ip}\right)$ smooth functions (splines) with dimension k_q and basis function $b_j^q(x)$

$$f_q(x) = \sum_{j=1}^{k_q} \beta_{q,j} b_j^q(x)$$

Back