

Podobný pohyb kryptomien

Michal Porubský

RNDr. Ing. Peter Molnár PhD.

Príbeh

- Krypto je zaujímavá téma
- Data science je zaujímavá oblasť
 - ale nič som o nej nevedel
 - chcel som skúsiť
- Zlatá stredná cesta - táto bakalárka



Doterajší výskum

- Podobný pohyb je študovaný v iných oblastiach (**nie v krypte!**)
- Krypto market je **nový** - a neprebádaný

- Naše otázky:
 - Existujú závislosti aj medzi takou novou, *informaticou komoditou* akou sú kryptomeny?
 - Aj keď sa ceny hýbu tak rýchlo, že to na prvý pohľad pôsobí chaoticky?




Podobný pohyb kryptomien

= dlhodobá silná závislosť medzi cenovým vývinom

- Všetky ceny sú nejako prepojené
 - ten istý typ trhu
- Niektoré ale viac ako iné
 - spol. technológia, bezpečnostný incident, ...
- Cieľ:
Identifikácia dlhodobého stáleho relatívneho vývoja dvojice mien a vyhodnotenie porovnaním zisku ziskateľným automatizovaným obchodovaním

Zber dát

- **Výber zdroju dát a základnej meny (ETH/BTC)**
 - Program na zber dát (nastaviteľný):
 - Minimálne časové okno? - 2 roky (39 mien)
 - Základná mena? - BTC
 - Frekvencia cien? - 1h
 - Node.js, ES6, Babel, ESLint, Prettier, git
- 

Zber dát - výzvy

- Dlhý beh programu
 - Asynchrónnosť
- Veľa requestov
 - Semafór
- Sťahovanie dát znova a znova
 - Caching ekosystém



Analýza dát

- Predspracovanie dát
 - Hodiny bez obchodu - trebalo vyplniť
 - Normalizácia
- Grafy na mieru
 - lepšie 3x merať a raz strihať
- Testy podobného pohybu
- Python, Jupyter Notebook, Pandas, git



Testy spoločného pohybu

- Euklidovská vzdialenosť
 - Veľké výkyvy od normy veľmi ovplyvnia výsledok
 - Silná korelácia
 - Pointa: $\text{cena } X = \text{cena } Y$ (normalizované)

- Kointegrácia
 - Výkyvy neovplyvnia výsledok
 - Možnosť omylu vyššia
 - Pointa: $C + \mathbf{CHAOS} = \text{cena } X + a * \text{cena } Y$



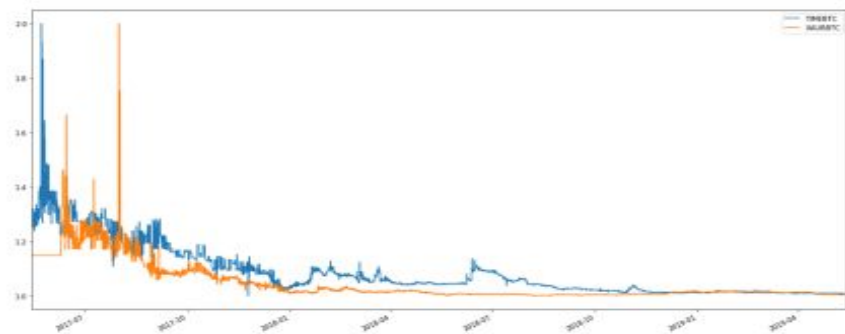
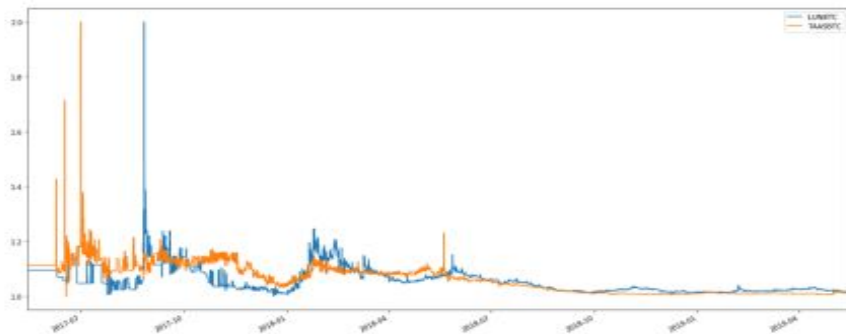


Figure 5.5: Two randomly chosen pairs from the top 20 pairs ranked by the lowest euclidean distance.

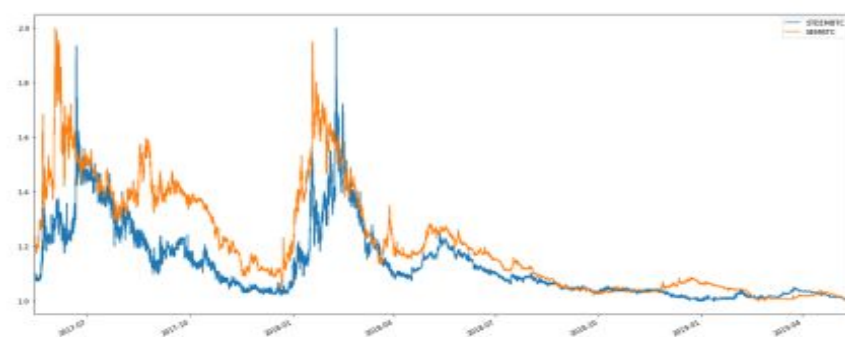
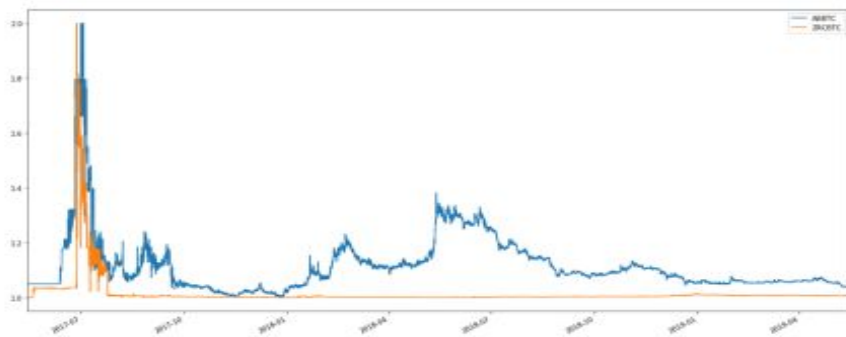
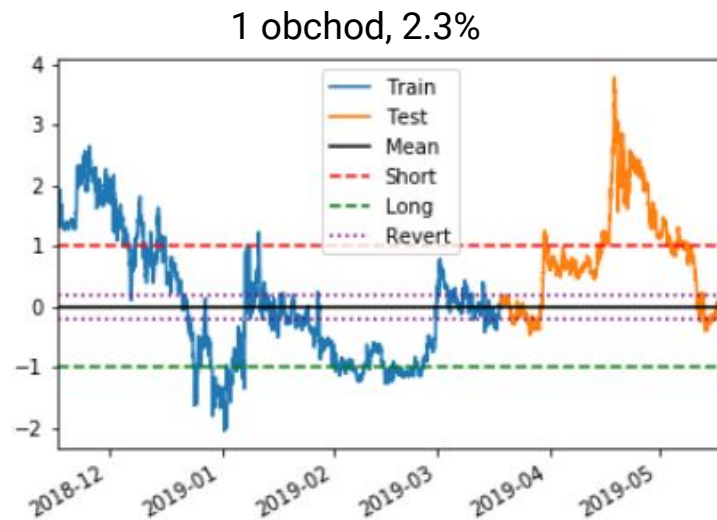


Figure 5.6: Two randomly chosen cointegrated pairs. Shown normalized, even though the cointegration test was not performed on the normalized prices.

Obchodná stratégia

- Párové obchodovanie (*pairs trading*)
- Kedy presne otvoriť pozície - *z-index*
- Rozdelenie dát
 - Dáta = c * disjunktná časť
 - chunk = $\frac{2}{3}$ trénovacie + $\frac{1}{3}$ testovacie



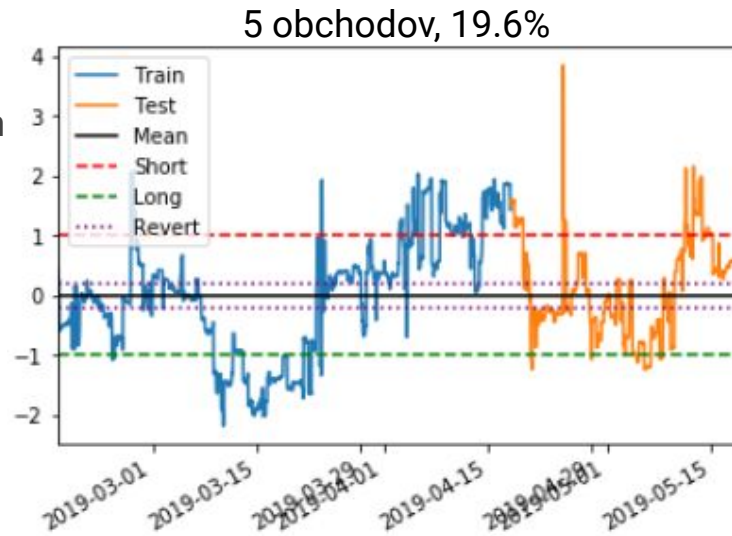
Obchodná stratégia

- **Detaily**

- Pár na trénovaných dátach, zisk na testovacých
- Vstupný kapitál vstupuje do každej časti
- Každá časť = viac párov, rovnomerne (fix)
- Zisk = súčet ziskov za časti

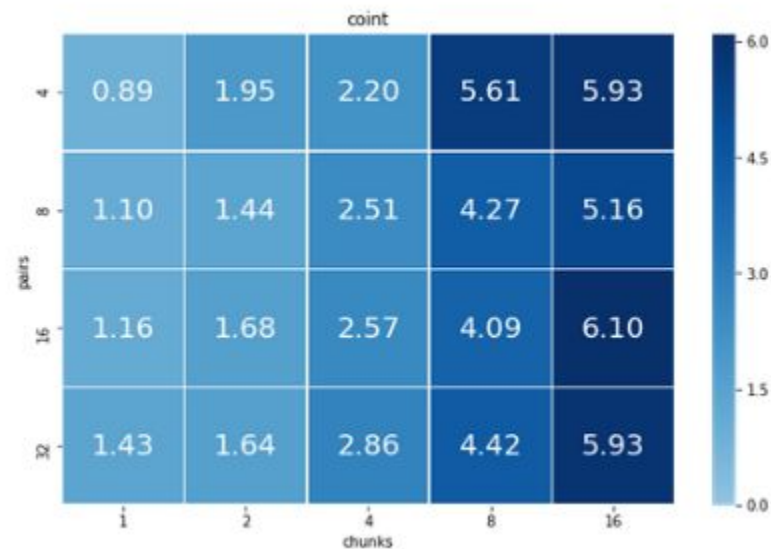
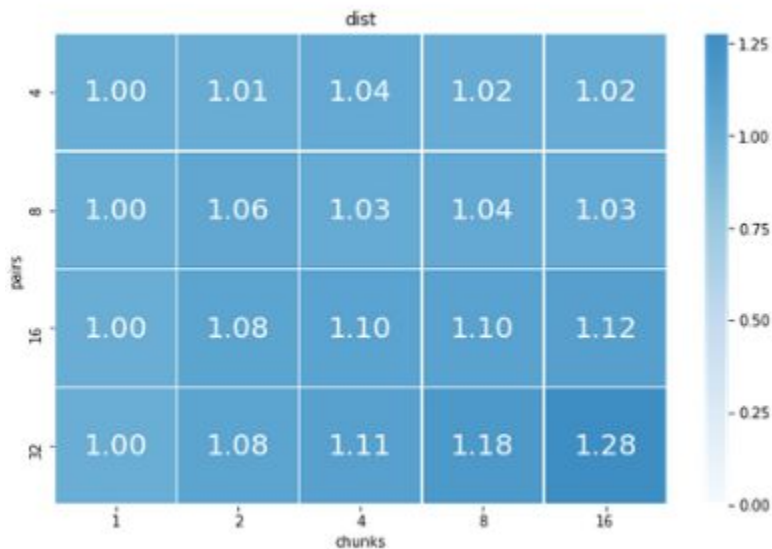
- **Implementované všeobecne**

- Počet častí
- Max. obchodovaných párov na časti
- Zvolená metóda
- = ľahko pridať novú metódu a/alebo upraviť celé fungovanie



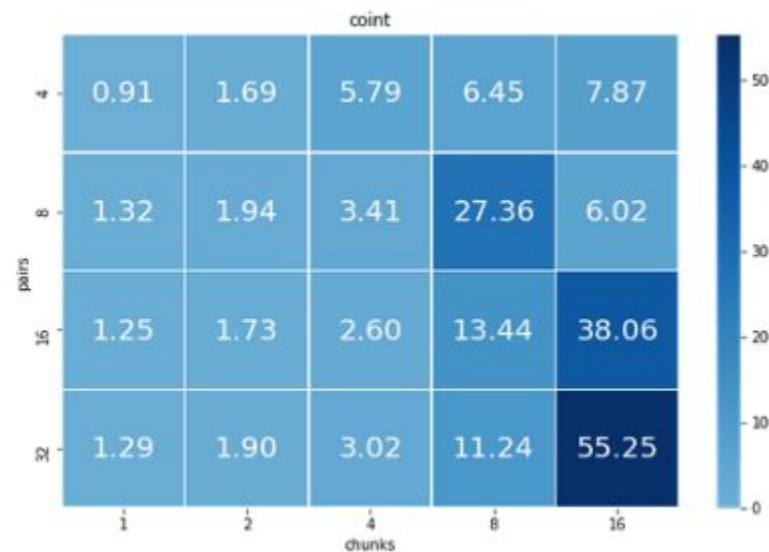
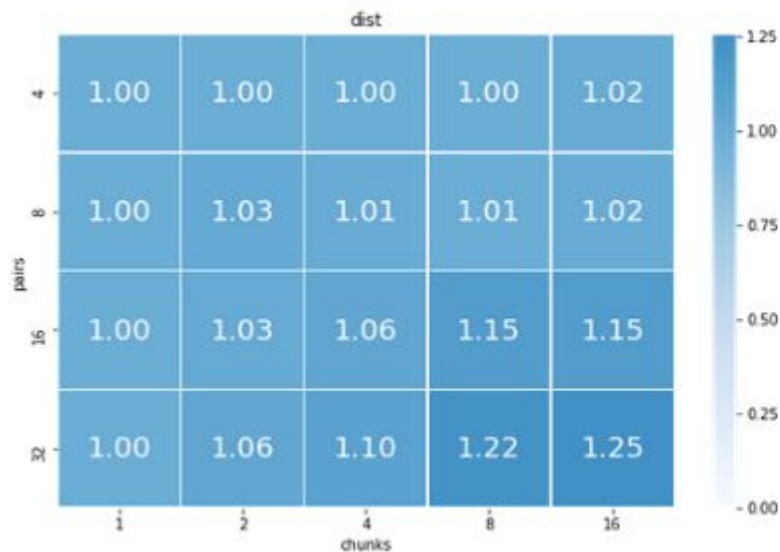
Náš výskum - porovnanie výsledkov

- Vplyv počtu distribuovanosti párov na zisk (dist. najlepšie, coint. náhodné)
- Dĺžka časti = dĺžka pozorovanej korelácie



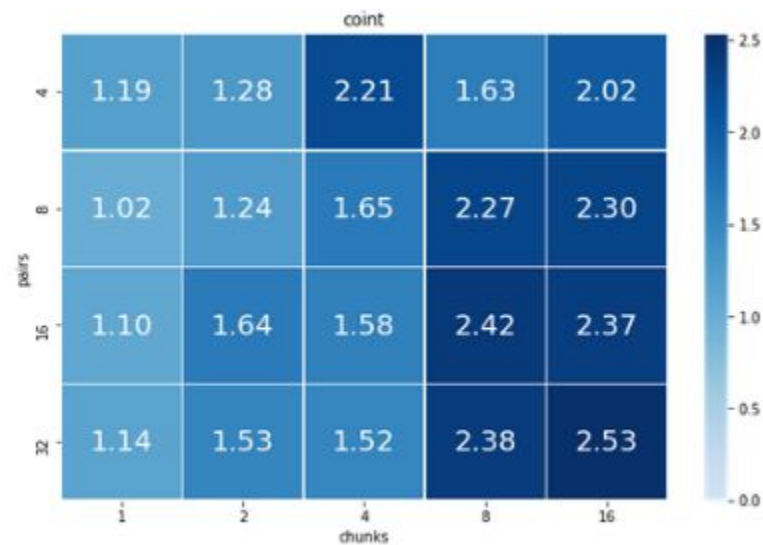
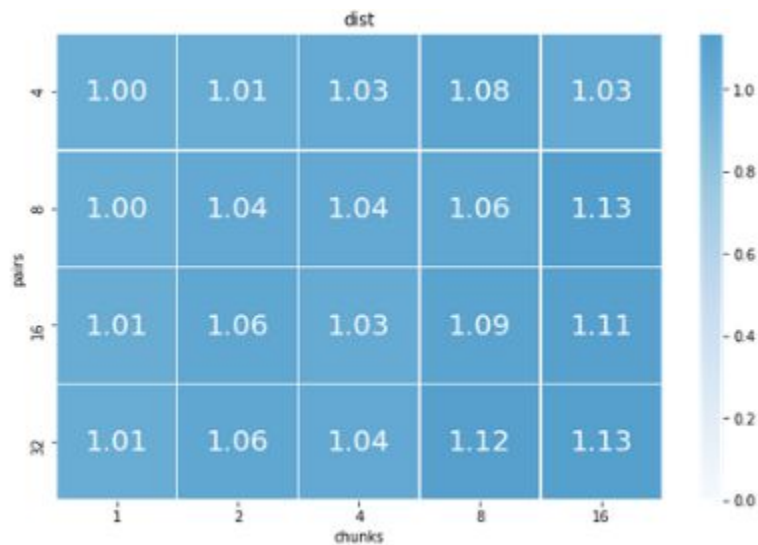
Náš výskum - porovnanie výsledkov

- 15 minútové dáta



Náš výskum - porovnanie výsledkov

- 1 dňové dáta



Do budúcnosti

- Do budúcnosti odporúčame pozrieť sa na:
 - Burzové poplatky
 - Ceny premeniť na logaritmy - pri dist. to môže pomôcť s výkyvmi
 - Obchodovanie "real-time", v tzv. Rolling window
 - Použitie viacerých metód



Ďakujem za pozornosť

