



TECHNICAL UNIVERSITY OF LIBEREC  
Faculty of Mechatronics, Informatics  
and Interdisciplinary Studies



# DATABÁZE PRO BIG DATA

## CVIČENÍ IX.

### ELASTICSEARCH

*Lukáš Matějů*

*29.4.2024 | DPB*





# CVIČENÍ

- zaměřené na tři oblasti
  - mapování
  - analýza
  - agregace
- využívá data z minulého cvičení
  - indexy products, orders a recipes
- k řešení využijte opět Kibanu
  - potřeba spustit kontejnery pro Elasticsearch a Kibanu  
<http://localhost:5601>



# ÚLOHY

- mapování
- 1. vytvořte mapování pro definici struktury indexu pro objednávky
  - index nazvěte orders-new
  - index bude umožňovat objednávky produktů z indexu products (viz minulé cvičení)
  - vymyslete vhodnou strukturu objednávky a vhodné datové typy
  - není potřeba vymýšlet až příliš hodnot, např. pro zákazníka stačí jméno
- 2. vložte záznam do vytvořeného indexu
  - objednávka musí obsahovat kávu a další dva libovolné produkty

# ÚLOHY

- analyzéry
- 1. vyzkoušejte si 3 analyzéry pomocí POST \_analyze
  - [seznam](#)
  - příklad použití najdete např. ve standardním analyzáru

# ÚLOHY

- agregace (nad indexem products)
- 1. zjistěte maximální, minimální a průměrnou cenu produktů
  - bez použití stats
- 2. zjistěte maximální, minimální a průměrný počet prodaných produktů
  - s použitím stats
  - a) pomocí dalšího dotazu zjistěte, který produkt se prodává nejvíce
- 3. pro každý tag zjistěte, v kolika dokumentech je obsažen
- 4. zjistěte cenové statistiky pro jednotlivé tagy
  - vnořená agregace
  - do 3. úlohy přidat stats

## BONUSOVÉ ÚLOHY

- agregace (nad indexem products)
  1. zjistěte cenové statistiky pouze pro produkty s tagem „Coffee“ a „Cake“ (filters)
- agregace (nad indexem orders)
  2. zjistěte minimální, maximální a průměrnou cenu objednávek podle jejich statusu
  3. zjistěte průměrnou cenu a počet objednávek pro jednotlivá čtvrtletí roku 2016
    - dokumenty bez data objednání skončí v bucketu „N/A“
    - přidejte takový dokument a ověřte