

# Python - dátová analýza II (Pokročilé spracovanie dát)

Kód kurzu: PYTHON\_DATAN2

Počas týchto 5 dní budú mať účastníci príležitosť získať základné a pokročilé znalosti v oblasti spracovania dát v Pythone s použitím knižníc ako je numpy, pandas, modin, polars, vaex. Bude to kombinácia teoretických prednášok a praktických cvičení, aby sa účastníci stali schopnými efektívne pracovať s dátami a vykonávať analýzu dát v Pythone. Neoddeliteľnou súčasťou školenia u každého preberaného modulu je interaktívna vizualizácia analyzovaných dát.

## Pre koho je kurz určený

- Data Scientist, dátoví analytici, hlavne v Big Data prostredí sú primárnym auditóriom pre tento intenzívny kurz
- Software vývojári, ktorí ovládajú jazyk Python aspoň na strednej až pokročilej úrovni a ktorí majú za cieľ vytvárať data-intenzívne alokácie pomocou enginu SPARK v prostredí Big Data (Cloud)
- Dátoví architekti

## Požadované vstupné znalosti

- Znalosti jazyka Python a dátovej analýzy na úrovni kurzu
- PYTHON\_ADV
- a
- PYTHON\_DATAN

## Osnova kurzu

### Základy Pythonu a úvod do NumPy

- Úvod do Pythonu ako programovacieho jazyka pre analýzu dát
- Inštalácia a import modulu NumPy
- Práca s NumPy poliami a maticami
- Operácie s NumPy poliami (sčítanie, násobenie, indexovanie)
- Štatistické a matematické funkcie NumPy

### Pandas - Manipulácia s dátami

- Úvod do knižnice Pandas
- Načítanie a ukladanie dát v Pandas DataFrame
- Práca s dátami v DataFrame (výber, filtrovanie, zmena)
- Skupinové operácie a agregácia dát
- Merge a spájanie dátových rámcov
- Praktické cvičenia s Pandas

### Paralelné spracovanie dát s Modin

- Úvod do Modin - paralelné spracovanie dát v Pandas
- Inštalácia a konfigurácia Modin
- Porovnanie rýchlosti Pandas a Modin na reálnych dátach
- Praktické cvičenia pre optimalizáciu dátového spracovania

### Polars - Moderné dátové spracovanie

- Úvod do Polars - moderná knižnica pre manipuláciu s dátami
- Porovnanie Polars a Pandas
- Práca s dátami v Polars DataFrame
- Analytické funkcie a SQL dotazy v Polars
- Praktické cvičenia s Polars

### Vaex - Rýchle a efektívne spracovanie veľkých dát

- Úvod do Vaex - knižnica pre rýchle spracovanie veľkých dát
- Práca s Vaex DataFrame
- Čítanie a zapisovanie veľkých dátových súborov
- Optimalizácia výkonu vo Vaex

#### GOPAS Praha

Kodaňská 1441/46  
101 00 Praha 10  
Tel.: +420 234 064 900-3  
[info@gopas.cz](mailto:info@gopas.cz)

#### GOPAS Brno

Nové sady 996/25  
602 00 Brno  
Tel.: +420 542 422 111  
[info@gopas.cz](mailto:info@gopas.cz)

#### GOPAS Bratislava

Dr. Vladimíra Clementisa 10  
Bratislava, 821 02  
Tel.: +421 248 282 701-2  
[info@gopas.sk](mailto:info@gopas.sk)



Copyright © 2020 GOPAS, a.s.,  
All rights reserved

# Python - dátová analýza II (Pokročilé spracovanie dát)

- Praktické cvičenia na spracovanie veľkých dát s Vaex

## **GOPAS Praha**

Kodaňská 1441/46  
101 00 Praha 10  
Tel.: +420 234 064 900-3  
[info@gopas.cz](mailto:info@gopas.cz)

## **GOPAS Brno**

Nové sady 996/25  
602 00 Brno  
Tel.: +420 542 422 111  
[info@gopas.cz](mailto:info@gopas.cz)

## **GOPAS Bratislava**

Dr. Vladimíra Clementisa 10  
Bratislava, 821 02  
Tel.: +421 248 282 701-2  
[info@gopas.sk](mailto:info@gopas.sk)



Copyright © 2020 GOPAS, a.s.,  
All rights reserved