

Vysokovýkonné počítanie

Python



Ako Začneme?

1. Stiahnite si Cvičný Súbor

- https://github.com/miroslav-reiter/Kurzy_SAV_Analytika_Python_R

2. Pridajte si ma na LinkedIn

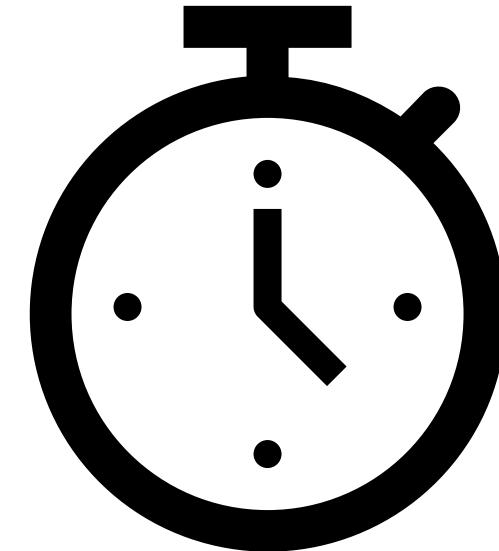
- www.linkedin.com/in/miroslav-reiter/

3. Prezentácia a materiály po prednáške

The screenshot shows a LinkedIn profile page for Miroslav Reiter. At the top, there's a navigation bar with 'SlideShare' and other LinkedIn icons. Below it, a banner for 'Automatizácia a optimalizácia Firemné procesy' is displayed, along with a small Python icon. Another presentation titled 'Najväčšie faily na sociálnych sietach a Google reklame' is also listed, featuring a Facebook icon. Both presentations were uploaded by Ing. Mgr. Miroslav Reiter, DIS., and have their respective URLs: [miroslav.reiter/slideshare/automatizacia-a-optimalizacia-firemn%C3%A9-procesov](#) and [miroslav.reiter/slideshare/najv%C3%A1c%C5%88ie-faily-na-soci%C5%88lnych-sietach-a-google-reklame](#).

Úvodné Informácie

- Časový rozvrh (9:00-13:30)
 - Programujeme a vizualizujeme (50 min)
 - Prestávky (10 min)
 - Obedová prestávka
 - Mobilné telefóny a zariadenia
-
- Priprav si otázky a rovno sa pýtaj
 - Interaktívna forma



O mne - Miroslav Reiter

4

40000+
klientov a
1000+ firiem

IT Architekt
Programátor
Manažér

Microsoft
Google
ISTQB tréner

134
certifikácií

151 príručiek
a publikácií

14 škôl

62 projektov

Vlastná firma



MOTIVÁCIA

Študuje 5 odborov a absolvoval už 12 univerzít. Ako zvláda stres a manažovanie času?



Foto: Jakub Kovalík pre FMK UCM | Miroslav Reiter na prednáške Grow with Google na FMK UCM.

Nikola Kotláriková

19. júl 2022 · 8 min. čítania





Miroslav Reiter

1. PhDr. VŠM (Podnikovný manažment)
2. Ing. STU FEI (**Aplikovaná informatika**)
3. Mgr. UK FM (**Strategický manažment a marketing**)
4. Mgr. VŠM (**Manažérstvo kvality**)
5. Mgr. VŠEMVŠ (Verejná správa)
6. Mgr. DTI (**Učiteľstvo ekonomických predmetov**)
7. DiS. AMOS (Cestovný ruch)
8. MBA LIGS (**Executive management**)
9. DBA Humanum (**IT manažment**)
10. MPA IES (Verejná správa a samospráva krajov)
11. MSc. Humanum (**Bezpečnosť inf. systémov**)
12. Ing. Paed. IGIP STU
13. Mgr. PEVŠ (**Bezpečnosť informačných systémov**)
14. RNDR. PEVŠ (**Bezpečnosť informačných systémov**)



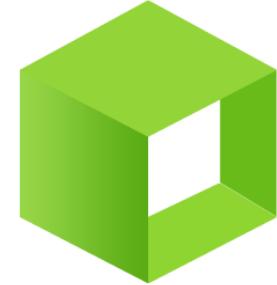
FAKULTA MANAGEMENTU
Univerzita Komenského
v Bratislave

**DIGITÁLNA
UNIVERZITA**

STU
FIIT



**VITA
ACADEMY**



IT ACADEMY

Most Valuable Professionals

[About](#) [Events](#) [Find an MVP](#)[Profile](#) [Events](#)

Headline

🌟 IT Architect and Programmer 📚 Hard worker 🎓
Lecturer and Certified Trainer

Biography

👉 I'm a hard worker, intellectual and joker. I love learning and teaching as well. My main objective is to teach people and improve their IT knowledge. To create truly practical knowledge necessary for life and present it in an interesting way. I don't like snake charmers and people who cannot...

[▼ Read more](#)

Miroslav Reiter

 Slovakia IT Academy s.r.o.

Most Valuable Professionals

High Impact Activities

Award Category

M365

Technology Area

Visio, Excel

Languages

English, Slovak

Social



This community leader has not added a high impact activity yet.

Znalec

PhDr. Ing. Miroslav Reiter

Evidenčné číslo: 915864

Miesto výkonu činnosti

Tomášikova 50
83104 Bratislava
Slovenská republika
[Zobrazit na mape](#)

Kontaktné údaje

Mobil: +421 908 163 084
E-mail: znapec@it-academy.sk

Odbory a odvetvia

| Odbor / Odvetvie | História | Stav |
|-----------------------------------------------------------|---------------------------|----------------|
| 100000 - Elektrotechnika | | |
| 100400 - Riadiaca technika, výpočtová technika (hardware) | 14.02.2023 - Zápis | AKTÍVNY |
| 100800 - Nosiče zvukových a zvukovoobrazových záznamov | 14.02.2023 - Zápis | AKTÍVNY |
| 100900 - Počítačové programy (software) | 14.02.2023 - Zápis | AKTÍVNY |
| 101000 - Bezpečnosť a ochrana informačných systémov | 14.02.2023 - Zápis | AKTÍVNY |

Vyberte si online kurz

Naučte sa programovať, tvoriť webstránky a grafiku, manažovať alebo sa zamerajte na osobný rozvoj. Všetko jednoducho vďaka našim online kurzom z pohodlia domova.

Ročné Predplatné na všetky Online Kurzy

2290 €

490 €

Pristup pre Vás do všetkých Aktuálnych aj Pripravovaných Online Kurzov

12 mesačná platnosť

Kúpiť teraz

Zistíť viac



541 kurzov v ponuke



Zábavné online lekcie



Akreditované kurzy



12 rokov skúseností



Certifikovaní profesionálni lektori

Top kurzy Najobľúbenejšie kurzy

BESTSELLER

Online kurz SAP I.
Začiatočník
Miroslav Reiter
 294 € 390 €

BESTSELLER

Online kurz Python a Jupyter I. Začiatočník
Miroslav Reiter
 152 € 214 €

BESTSELLER

Online kurz JavaScript I.
Začiatočník
Miroslav Reiter
 168 € 226 €

BESTSELLER

Online kurz Java I.
Začiatočník
Miroslav Reiter
 186 € 254 €

BESTSELLER

Online kurz SQL I.
Začiatočník
Miroslav Reiter
 168 € 226 €

BESTSELLER

Online kurz Microsoft Excel III. Pokročilý
Miroslav Reiter
 196 € 223 €

[Všetky kurzy kategórie Najobľúbenejšie kurzy →](#)



Luigi, Mário
a Yoshi

Čo Robíte?

1. Študent/učiteľ'

2. Zamestnanec

3. Podnikateľ'

4. Nezamestnaný/materská

5. Dievča pre všetko



Vaše Ciele a Očakávania

13

1. Doplniť si znalosti z jazyka Python

2. Doplniť si znalosti z HPC

3. Používanie knižnice Polars

4. Používanie knižnice PySpark

5. Práca s big data

Zábava je v zaručená v každom bode :-)



Ako začať s HPC a Big Data

AKREDITOVANÝ KURZ



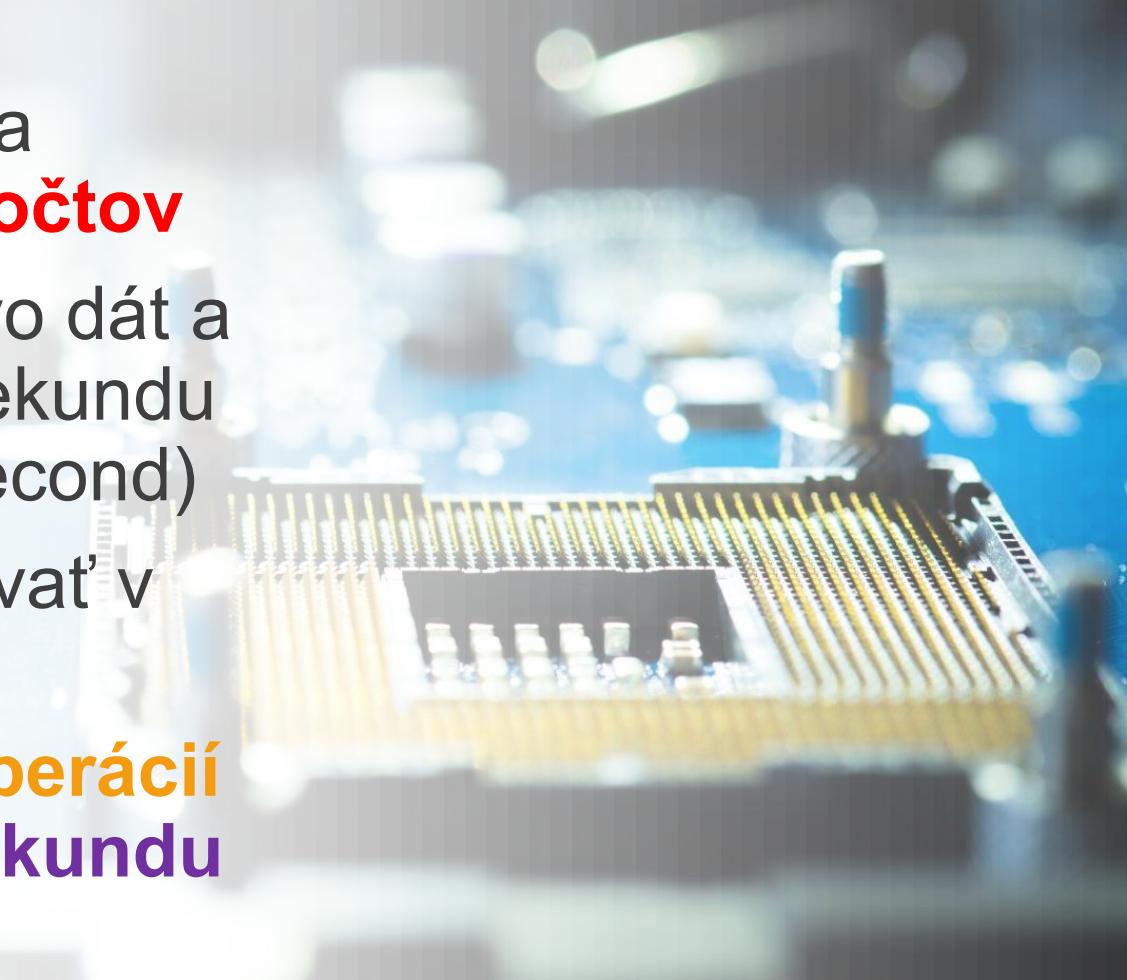
Vysokovýkonné počítanie HPC



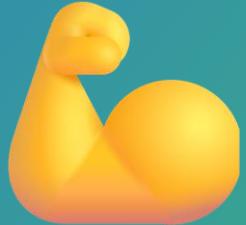
- High-Performance Computing
- Zameriava sa na **využívanie výkonných výpočtových systémov**
- Umožňuje riešiť **veľmi náročné výpočtové problémy**, ktoré by inak trvali príliš dlho na **bežných počítačoch**
- Cieľom je **optimalizovať** výpočty tak, aby sa **dosiahli čo najrýchlejšie výsledky**
- Často využitím **parallelného spracovania** a špecializovaných hardvérových komponentov ako **grafické procesory (GPU)**, **klastrami serverov** alebo **superpočítačmi**

Superpočítače

- Výkonné výpočtové systémy navrhnuté na **vykonávanie extrémne náročných výpočtov**
- Sú schopné spracovať obrovské množstvo dát a vykonávať tisíce až milióny operácií za sekundu (**FLOP - Floating Point Operations per Second**)
- Moderné superpočítače sa môžu pohybovať v rozsahu exaflopov (10^{18} FLOP)
- **1 teraflop (TFLOP) = 1 bilión (10^{12}) operácií s pohyblivou desatininou čiarkou za sekundu**



Superpočítače



1. **Fugaku (Japonsko):**
cca 442 petaflopov (PFLOP) =
442 000 teraflopov
2. **Summit (USA):** cca 200 petaflopov
(PFLOP) = 200 000 teraflopov
3. **Sierra (USA):** cca 125 petaflopov
(PFLOP) = 125 000 teraflopov
4. **LUMI (Európa):** cca 38 petaflopov
(PFLOP) = 38 000 teraflopov
5. **Devana (Slovensko):**
cca 800 teraflopov (TFLOP)



FRONTIER

AT
1.1 EXAFLOPS
IT'S ALSO THE
WORLD'S FASTEST
SUPERCOMPUTER.



SLUŽBY

[Sietové služby a infraštruktúra ▶](#)[Webové služby a propagácia ▶](#)[Vysokovýk. výpočtová technika ▾](#)[Základné informácie](#)[» Superpočítač Devana](#)[Časopis HPC Focus](#)[Výzvy a ďalšie dokumenty](#)[Dokumentácia ext](#)[Projekt Superhrdinovia vedy ext](#)[Ostatné služby](#)[Objednávkový formulár služieb](#)[Cenník služieb](#)

Superpočítač Devana

VYSOKOVÝKONNÁ VÝPOČTOVÁ TECHNIKA
NA VÝPOČTOVOM STREDISKU SAV

VVT systém Devana (Development and enabling) bol obstaraný ako náhrada za superpočítač Aurela v rámci Operačného Programu Integrovaná Infraštruktúra (OPII-2020/7/55-NP, projekt Národné kompetenčné centrum pre vysokovýkonné počítanie, kód 311071AKF2). Technológia od firmy Lenovo bola nasadená v konfigurácii 140 univerzálnych výpočtových uzlov a 8 gpu akcelerovaných uzlov pre úlohy z oblasti umelej inteligencie a strojového učenia (AI/ML). Výkonovou jednotkou univerzálneho výpočtového modulu je server SD630 s procesormi Intel Icelake 6338 32C/2.0 GHz/205 W a 256 GB operačnej pamäte. Akcelerované moduly pozostávajú zo serverov typu SR670 v2 so štyrmi GPGPU akcelerátormi Nvidia A100 40 GB. Viac podrobností nájdete v časopise **HPC Focus** (ročník 2022).

Devana je bezplatne prístupná pre všetkých vedcov a výskumníkov zo slovenských univerzít

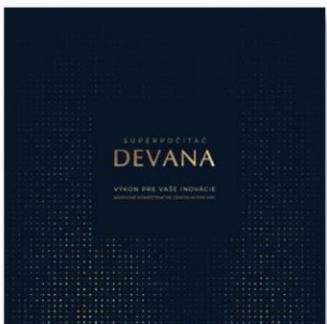




HARDVÉROVÁ KONFIGURÁCIA DEVANY

| | |
|------------------|----------------------------------------|
| Platforma | Lenovo ThinkSystem SD630 v2 + SR670 v2 |
| Operačný systém | CentOS 7.x |
| Architektúra | x86_64 |
| CPU uzly | 140 |
| GPU uzly | 8 |
| CPU jadrá | 9472 |
| GPU jadrá | 221184 CUDA + 13824 Tensor |
| Teoretický výkon | ~800 TFlops |
| Operačná pamäť | 38 TB |
| Dátové úložiská | 1 PB + 600 TB + 400 TB |
| Súborový systém | BeeGFS + Linux native |
| Výpočtová siet' | HDR Infiniband |
| Topológia siete | Fat tree (neblokujúca) |
| Príkon | 130 kW |

SUPERPOČÍTAČ DEVANA



Superpočítač Devana

PDF | 27 MB



Nástroje HPC

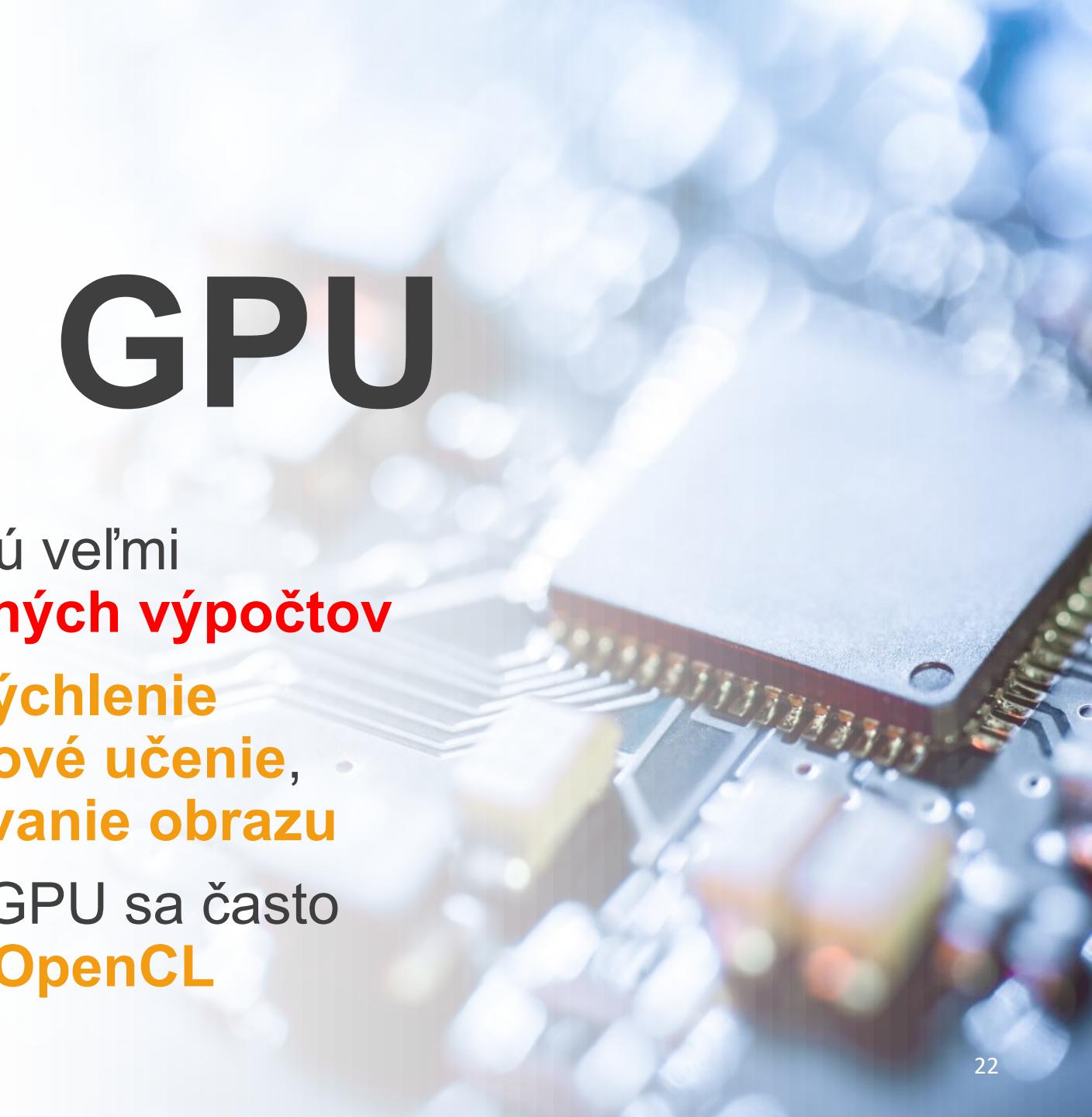


1. Polars
2. Parquet
3. NumPy
4. Apache Spark
5. SciPy
6. Dask
7. TensorFlow
8. PyTorch
9. MPI for Python (mpi4py)
10. CUDA



Grafické procesory GPU

- Sú **špecializované čipy**, ktoré sú veľmi efektívne pri vykonávaní **paralelných výpočtov**
- V HPC sa často používajú na **urýchlenie výpočtov** v oblastiach ako **strojové učenie, numerické simulácie a spracovanie obrazu**
- Na vývoj aplikácií využívajúcich GPU sa často používajú knižnice ako **CUDA** a **OpenCL**



Úlohy HPC a BD

1. Pochopenie čo je HPC
2. Prehľad nástrojov pre HPC a BD
3. Pochopenie úzkych hrdiel BD
4. Používanie techník na prácu s HPC



Ako začať s Big Data

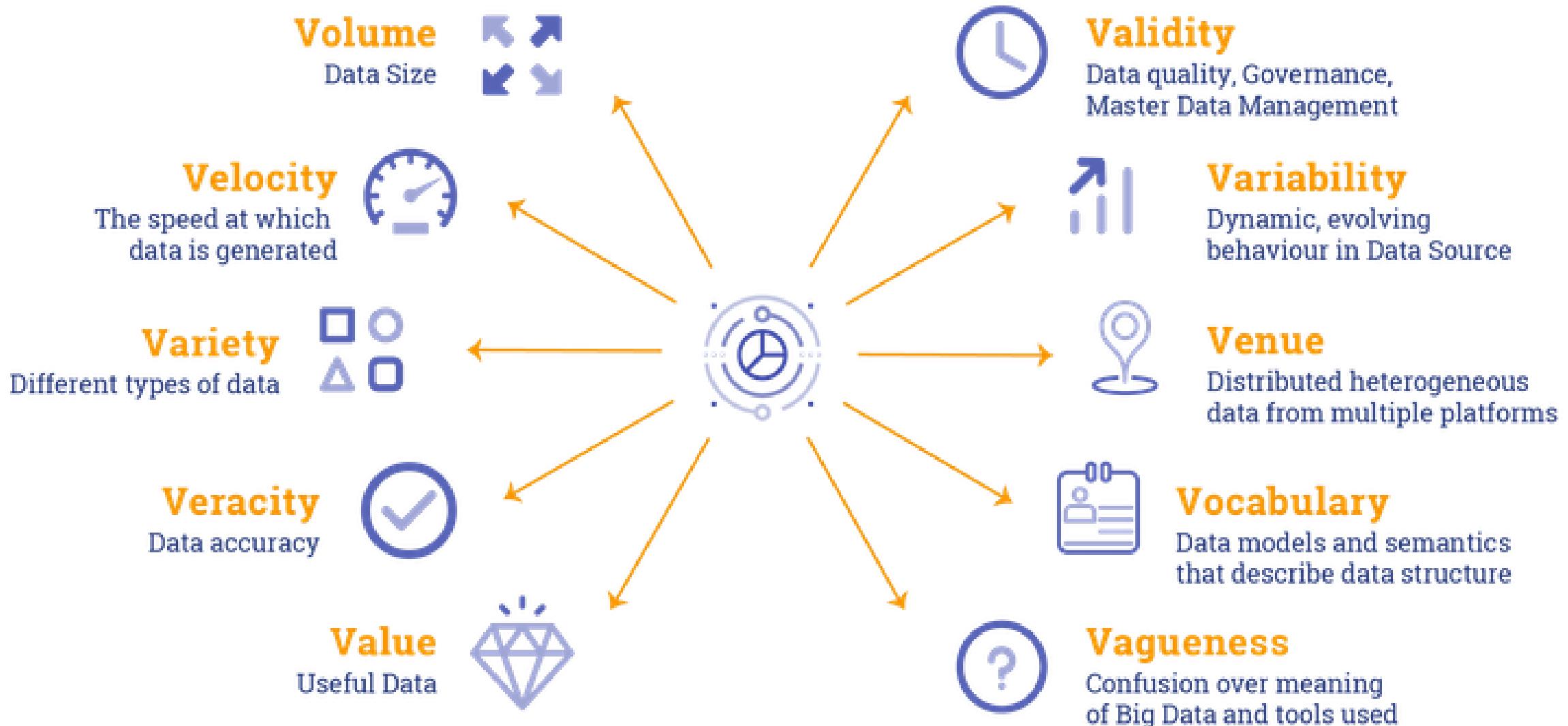
AKREDITOVANÝ KURZ



Veľké dáta (Big Data)

- Sú **obrovské objemy dát**, ktoré je potrebné **analyzovať** a **spracovávať pomocou HPC technológií**
- Tieto dáta môžu pochádzať z rôznych oblastí ako **sociálne siete, medicína, financie a prírodné vedy**
- **Charakteristika 3V-10V**

THE 10 Vs OF BIG DATA



Nástroje Big Data

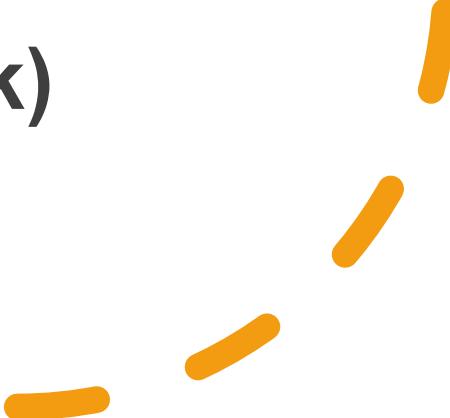


1. **Apache Hadoop:** Platforma na distribuované spracovanie veľkých dát
2. **Apache Spark:** Rýchly systém pre analýzu veľkých dát
3. **NoSQL databázy:** Ako MongoDB alebo Cassandra, na spracovanie nestruktúrovaných alebo polostruktúrovaných dát
4. **MapReduce:** Programovací model pre spracovanie veľkých dát v paralelných úlohách

Knižnice python



1. Pandas
2. Polars
3. Dask
4. PySpark (Apache Spark)
5. Vaex
6. Modin
7. Apache Arrow
8. Hadoop (PyArrow, PySpark)
9. Zeromq & Kafka
10. Blaze



Pandas

Knižnica na manipuláciu s dátami, vhodná pre menšie až stredné dátu.

Dask

Distribuované spracovanie dát, kompatibilné s Pandas a NumPy.

PySpark (Apache Spark)

Distribuované spracovanie veľkých dát cez Apache Spark.

Vaex

Efektívna analýza veľkých dát s využitím "lazy evaluation".

Modin

Paralelný výkon pre Pandas operácie.

Apache Arrow

Rýchla výmena dát medzi rôznymi knižnicami a programovacími jazykmi.

Hadoop (via PyArrow,
PySpark, etc.)

Distribuované spracovanie a ukladanie dát cez Hadoop.

Zeromq & Kafka

Pre streamovanie dát v reálnom čase.

Blaze

Práca s heterogénnymi dátovými zdrojmi (CSV, SQL, HDF5).

Polars

Rýchla knižnica na spracovanie veľkých dát s podobnou syntaxou ako Pandas, ale efektívnejšia pri paralelnej práci.

Riadky a Stípce

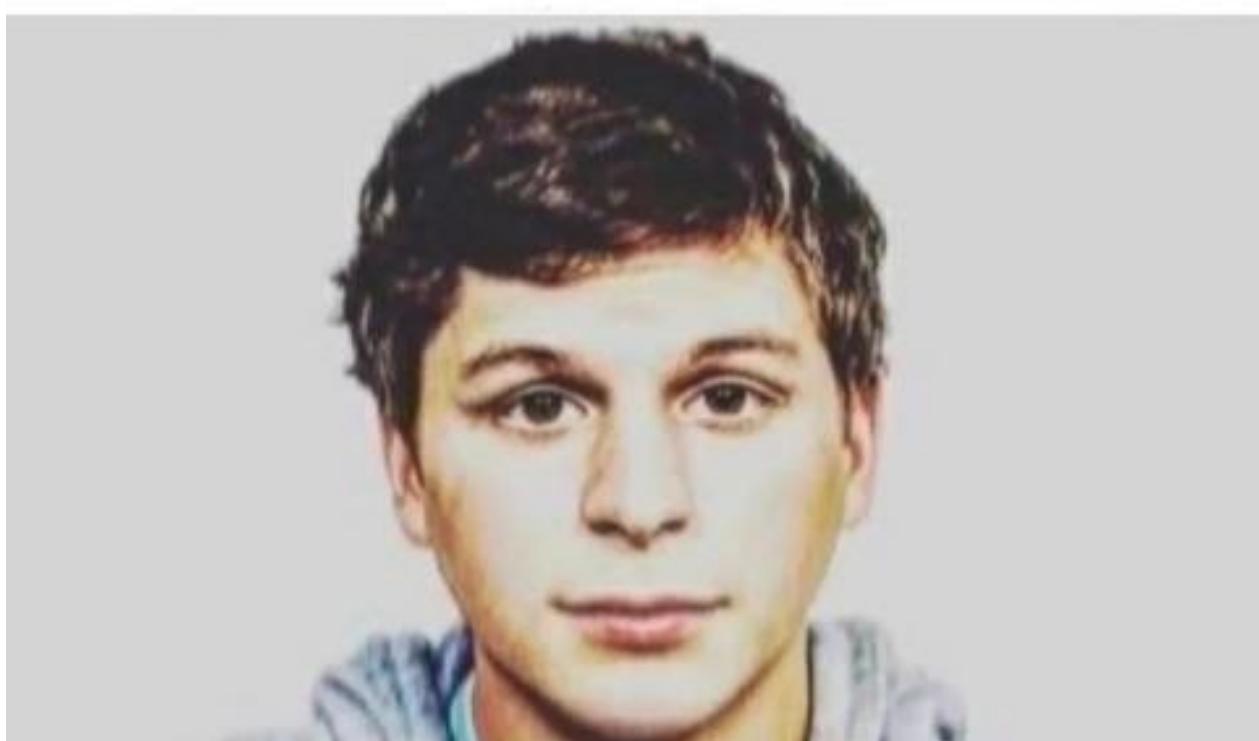
2007-2021 a Microsoft 365

- Počet riadkov: **1 048 576** (2^{20})
- Počet Stípcov: **16 384** (2^{14})(**XFD**)
- Celkovo Buniek: **17 179 869 184**
(cca 17 Miliárd)
- Súbory: **XLSX**, XLS, **XLSM**, **XLSB**

Google Tabuľky

- Počet riadkov: Maximálny počet buniek: **10 miliónov**
- Počet Stípcov: **18 278** ($2^{14,16}$) (**ZZZ**)
- Celkovo Buniek: **10 000 000**
(cca 10 Miliónov)
- Súbory: **XLSX**, **ODS**

Is the Cure to Slow Bad Code using faster hardware?



| Počet Stĺpcov | Umiestnenie | Veľkosť Súboru v MB | Počet Buniek | Počet Buniek v Mil | Čas |
|---------------|-------------|---------------------|---------------|--------------------|-----------------------------------|
| 10 | J | 28,9 | 10 485 760 | 10 | sekundy |
| 20 | T | 57,3 | 20 971 520 | 21 | sekundy |
| 30 | AD | 86,7 | 31 457 280 | 31 | sekundy |
| 40 | AN | 117 | 41 943 040 | 42 | sekundy |
| 50 | AX | 142 | 52 428 800 | 52 | sekundy |
| 60 | BH | 176 | 62 914 560 | 63 | sekundy |
| 70 | BR | 205 | 73 400 320 | 73 | sekundy |
| 80 | CB | 234 | 83 886 080 | 84 | sekundy |
| 90 | CL | 269 | 94 371 840 | 94 | sekundy |
| 100 | CV | 290 | 104 857 600 | 105 | sekundy |
| 110 | DF | 308 | 115 343 360 | 115 | sekundy |
| 120 | DP | 344 | 125 829 120 | 126 | sekundy |
| 130 | DZ | 362 | 136 314 880 | 136 | do minúty |
| 140 | EJ | 400 | 146 800 640 | 147 | do minúty |
| 150 | ET | 419 | 157 286 400 | 157 | do minúty |
| 160 | FD | 444 | 167 772 160 | 168 | do minúty |
| 170 | FN | 477 | 178 257 920 | 178 | do minúty |
| 180 | FX | 493 | 188 743 680 | 189 | do minúty |
| 190 | GH | 530 | 199 229 440 | 199 | do minúty (do 52 sekúnd) |
| 200 | GR | 553 | 209 715 200 | 210 | do minúty |
| 220 | HL | 607 | 230 686 720 | 231 | do minúty |
| 240 | IF | 655 | 251 658 240 | 252 | nad minútou |
| 260 | IZ | 719 | 272 629 760 | 273 | nad minútou (1:10) |
| 280 | JT | 752 | 293 601 280 | 294 | nad minútou |
| 300 | KN | 837 | 314 572 800 | 315 | nad minútou |
| 350 | ML | 971 | 367 001 600 | 367 | nad minútou |
| 370 | NF | 990 | 387 973 120 | 388 | nad minútou |
| 380 | NP | 1020 | 398 458 880 | 398 | nad minútou (1:39) |
| 400 | OJ | 1070 | 419 430 400 | 419 | nad minútou |
| 450 | QH | 1200 | 471 859 200 | 472 | nad 2 minútou |
| 500 | SF | 1360 | 524 288 000 | 524 | nad 2 minútou (2:11) |
| 550 | UD | 1500 | 576 716 800 | 577 | nad 2 minútou |
| 600 | WB | 1630 | 629 145 600 | 629 | nad 2 minútou (2:39) |
| 700 | ZX | 1890 | 734 003 200 | 734 | nad 3 minútou (3:04) |
| 800 | ADT | 2200 | 838 860 800 | 839 | nad 3 minútou |
| 900 | AHP | 2480 | 943 718 400 | 944 | 4 minúty (4:00) do 25 GB RAM |
| 1000 | ALL | 2800 | 1 048 576 000 | 1 049 | nad 4 minútou (4:00) do 25 GB RAM |
| 1100 | APH | 3100 | 1 153 433 600 | 1 153 | nad 4 minútou (4:52) do 28 GB RAM |
| 1200 | ATD | 3370 | 1 258 291 200 | 1 258 | nad 5 minútou (5:32) do 30 GB RAM |
| 1300 | AWZ | 3640 | 1 363 148 800 | 1 363 | nad 5 minútou (5:59) do 33 GB RAM |
| 1400 | BAV | 3970 | 1 468 006 400 | 1 468 | nad 6 minútou (6:10) do 35 GB RAM |
| 1500 | BER | 4340 | 1 572 864 000 | 1 573 | nad 7 minútou (7:01) do 37 GB RAM |

Prehľad Veľkostí Údajov, Počtu Buniek a Časy

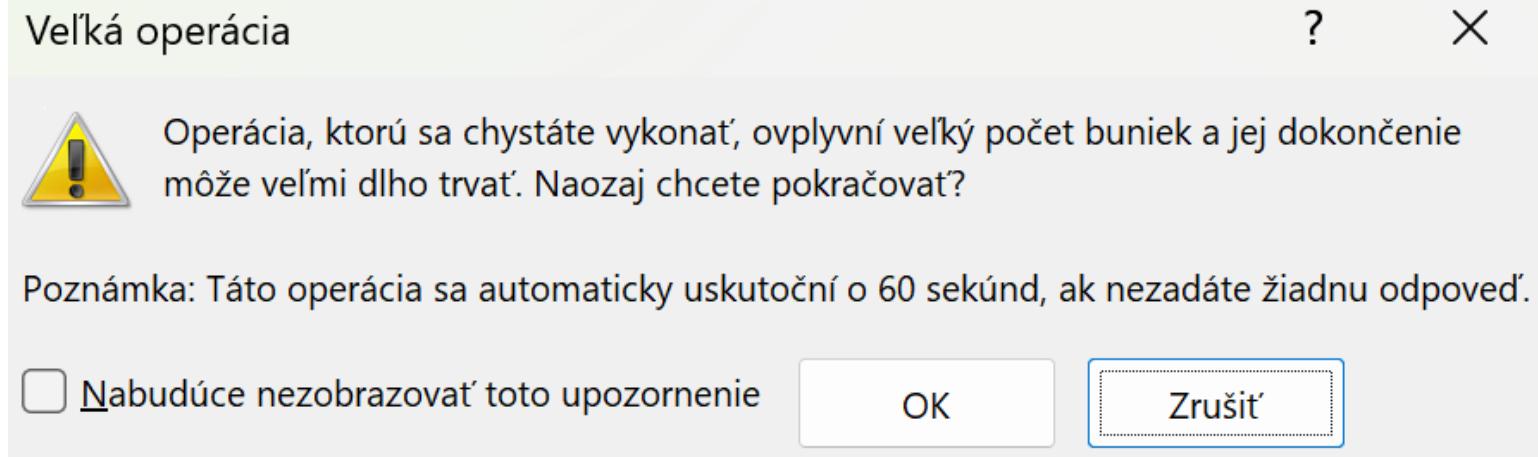
- A. 230 mil. Údajov do Minúty
- B. 400 mil. Údajov do 2 Minút
- C. 600 mil. Údajov do 3 Minút
- D. 900 mil. Údajov do 4 Minút
- E. 1,1 miliardy Údajov do 5 Minút
- F. 1,5 miliardy Údajov do 7 Minút

230 mil. Údajov do Minúty

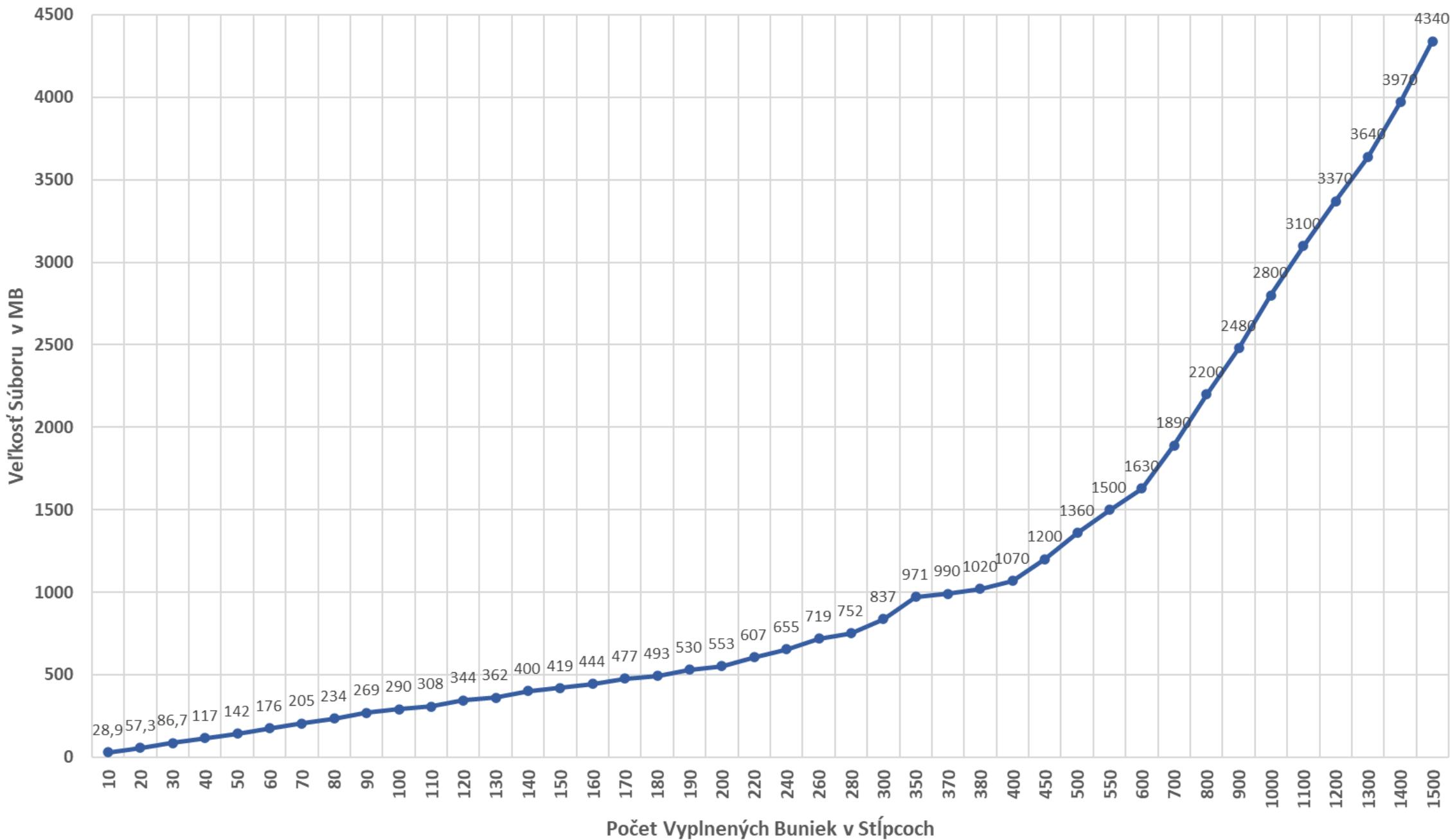
- A. 10 Stípcov po stípec J - 28,9 MB - 10485760 buniek - sekundy
- B. 20 Stípcov po stípec T - 57,3 MB - 20971520 buniek - sekundy
- C. 30 Stípcov po stípec AD - 86,7 MB - 31457280 buniek - sekundy
- D. 40 Stípcov po stípec AN - 117 MB - 41943040 buniek - sekundy
- E. 50 Stípcov po stípec AX - 142 MB - 52428800 buniek - sekundy
- F. 60 Stípcov po stípec BH - 176 MB - 62914560 buniek - sekundy
- G. 70 Stípcov po stípec BR - 205 MB - 73400320 buniek - sekundy
- H. 80 Stípcov po stípec CB - 234 MB - 83886080 buniek - sekundy
- I. 90 Stípcov po stípec CL - 269 MB - 94371840 buniek - sekundy
- J. 100 Stípcov po stípec CV - 290 MB - 104857600 buniek - sekundy
- K. 110 Stípcov po stípec DF - 308 MB - 115343360 buniek - sekundy
- L. 120 Stípcov po stípec DP - 344 MB - 125829120 buniek - sekundy
- M. 130 Stípcov po stípec DZ - 362 MB - 136314880 buniek - do minuty
- N. 140 Stípcov po stípec EJ - 400 MB - 146800640 buniek - do minuty
- O. 150 Stípcov po stípec ET - 419 MB - 157286400 buniek - do minuty (do 40 sek.)
- P. 160 Stípcov po stípec FD - 444 MB - 167772160 buniek - do minuty
- Q. 170 Stípcov po stípec FN - 477 MB - 178257920 buniek - do minuty
- R. 180 Stípcov po stípec FX - 493 MB - 188743680 buniek - do minuty
- S. 190 Stípcov po stípec GH - 530 MB - 199229440 buniek - do minuty (do 52 sek.)
- T. 200 Stípcov po stípec GR - 553 MB - 209715200 buniek - do minuty
- U. 220 Stípcov po stípec HL - 607 MB - 230686720 buniek - do minuty

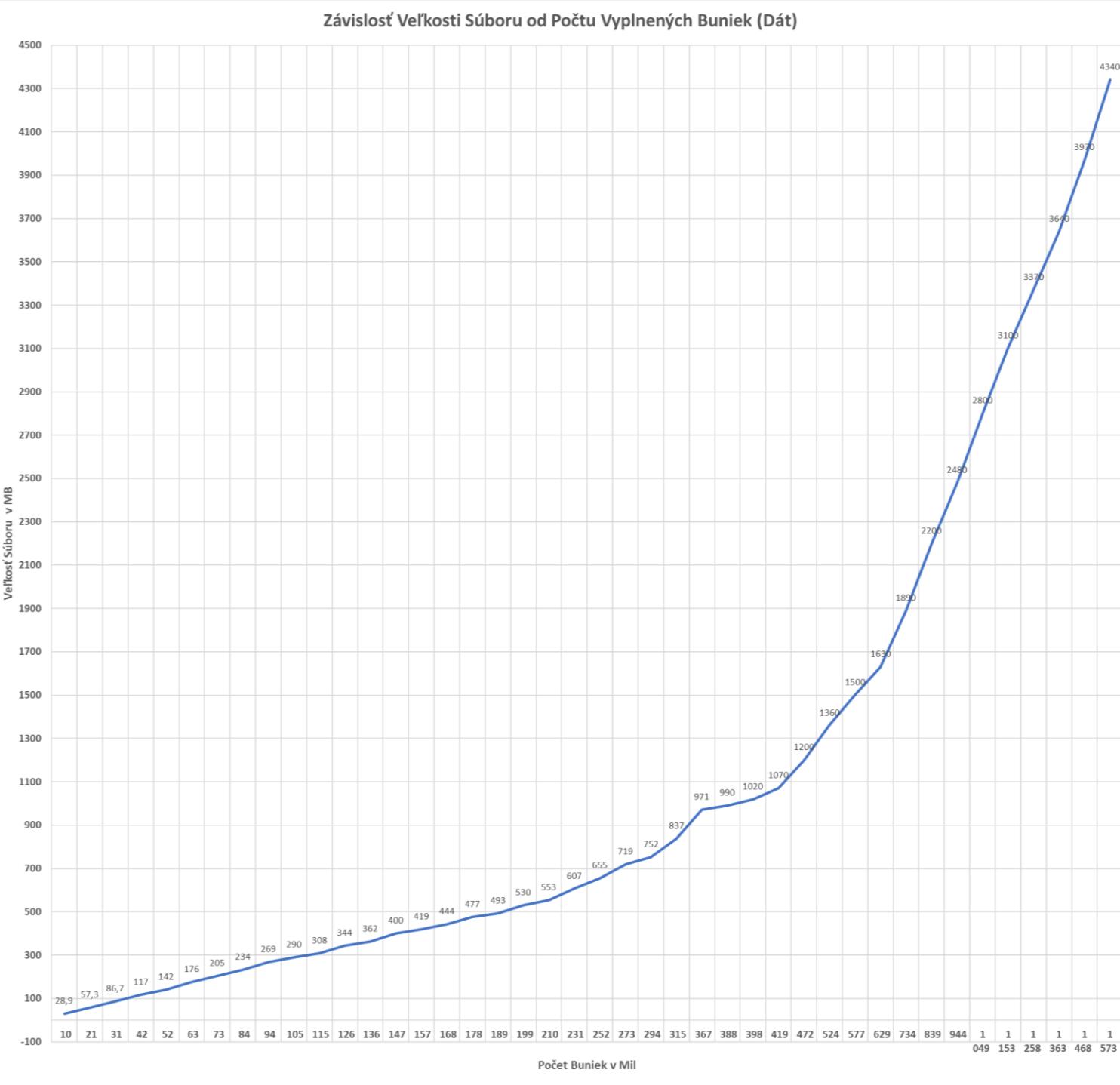
400 mil. Údajov do 2 Min.

- A. 240 Stĺpcov po stĺpec IF - 655 MB - 251658240 buniek - nad minútu
- B. 260 Stĺpcov po stĺpec IZ - 719 MB - 272629760 buniek - nad minútu (1:10)
- C. 280 Stĺpcov po stĺpec JT - 752 MB - 293601280 buniek - nad minútu
- D. 300 Stĺpcov po stĺpec KN - 837 MB - 314572800 buniek - nad minútu
- E. 350 Stĺpcov po stĺpec ML - 971 MB - 367001600 buniek - nad minútu
- F. 370 Stĺpcov po stĺpec NF - 0,99 GB - 387973120 buniek - nad minútu
- G. 380 Stĺpcov po stĺpec NP - 1,02 GB - 398458880 buniek - nad minútu (1:39)
- H. 400 Stĺpcov po stĺpec OJ - 1,07 GB - 419430400 buniek - nad minútu



Závislosť Veľkosti Súboru od Stĺpcov





Highlight Points

- Ctrl + * je rýchlejšie
(Okamžitá odozva 1-1,2 GB)
- Ctrl + A (4-5 sekúnd)
- Ideálne generovať dávkovo po 20-30/31 stĺpcach
(Inak Veľká Operácia)
- Na otvorenie 4,34 GB súboru treba 18 GB RAMky
- Na uloženie 4,34 GB súboru sa spotrebuje do 37 GB RAM

Big Data DBMS



| Databázový systém | Typ | Maximálna veľkosť databázy | Vhodnosť pre Big Data |
|----------------------|--------------------------------|-------------------------------------|--------------------------------------------------------|
| Microsoft Access | Relačná | 2 GB | Nízka |
| Microsoft SQL Server | Relačná | 524 PB (petabajty) | Vysoká |
| MySQL | Relačná | 256 TB | Stredná |
| MariaDB | Relačná | 256 TB | Stredná |
| PostgreSQL | Relačná | 32 TB na tabuľku | Vysoká |
| Oracle Database | Relačná | 8 ExB (exabajty) | Veľmi vysoká |
| MongoDB | NoSQL (dokumentová) | Neobmedzená | Vysoká |
| CassandraDB | NoSQL (wide-column) | Neobmedzená | Vysoká |
| SQLite | Relačná (embeddable) | 140 TB (maximálne pre súbory) | Nízka až stredná (pre malé až stredne veľké aplikácie) |
| IBM Db2 | Relačná | 100 TB (štandardne) až 12 PB a viac | Vysoká |
| SAP HANA | In-memory relačná databáza | Neobmedzená (závisí od hardvéru) | Veľmi vysoká |
| Amazon Aurora | Relačná (cloudová) | 64 TB per database | Vysoká |
| CouchDB | NoSQL (dokumentová) | Neobmedzená | Stredná |
| Redis | In-memory NoSQL (key-value) | Neobmedzená | Stredná až vysoká (na caching a real-time aplikácie) |
| Elasticsearch | NoSQL (full-text vyhľadávanie) | Neobmedzená | Vysoká |

Prehľad Veľkostí Údajov

- A. 1 mil. - S CPU: 4 GB: 26 MB
- B. 10 mil. - L CPU: 16 GB: 259,8 MB
- C. **100 mil.** - CPU XL : **32 GB: 2,6 GB**
- D. 1000 mil. CPU XXL: 96 GB: 20,1 GB
- E. ... Miliardy – GPU: 100+GB

DL Variable viewer

File Tools Kernel View Code Run Help

Build report Run all MR

df = {DataFrame} DataFrame[Meno: string, Vek: int, Me...
spark = {SparkSession} <pyspark.sql.session.SparkSess...

Vysokovýkonné počítanie v Pythone

Apache Spark Knižnice Plotly Interaktívne Plotly Príklady Bokeh Prehľad Bokeh Interaktívne Bokeh Komplex Pandas Sheet 9 +

```
18.3s
1 import pandas as pd
2 import numpy as np
3
4 # Vytvorenie datasetu s 10 miliónmi riadkov
5 rows = 10_000_000
6
7 # Generovanie dát
8 data_large = {
9     "Meno": np.random.choice(["Ján", "Anna", "Peter", "Eva",
10    "Marek"], size=rows),
11    "Vek": np.random.randint(18, 70, size=rows),
12    "Mesto": np.random.choice(["Bratislava", "Košice", "Žilina",
13    "Nitra", "Prešov"], size=rows),
14    "Nakupy": np.round(np.random.uniform(50.0, 1000.0, size=rows),
15    2)
16 }
17
18 # Uloženie do CSV súboru
19 df_large.to_csv("large_test_data_spark.csv", index=False)
20
```

Your machine seems to run out of memory soon

Running

| | |
|------------|--------|
| CPU | 47% |
| RAM | 512 MB |
| Calculated | 1 |
| In process | 1 |
| Errors | 0 |

CPU S

Vysokovýkonné počítanie v Pythone

Apache Spark

Sheet 9

Knižnice

Plotly Interaktívne

Plotly Príklady

Bokeh Prehľad

Bokeh Interaktívne

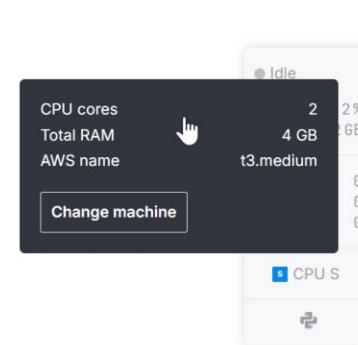
Bokeh Komplex

Pandas

+

▼

```
[2] > 1 import pandas as pd
2 import numpy as np
3
4 # Vytvorenie datasetu s 10 miliónmi riadkov
5 zaznamy = 100_000_000
6
7 # Generovanie dát
8 data_klienti = {
9     "Meno": np.random.choice([
10        "Adriana", "Alena", "Andrea", "Anna", "Barbora", "Beáta",
11        "Daniela", "Dominika", "Elena", "Emília", "Erika", "Eva",
12        "Gabriela", "Hana", "Ivana", "Jana", "Katarína", "Kristína",
13        "Laura", "Lenka", "Lucia", "Magdaléna", "Mária", "Martina",
14        "Monika", "Natália", "Petra", "Renáta", "Silvia", "Simona", "Soňa",
15        "Tamara", "Veronika", "Zuzana"], size = zaznamy),
16     "Vek": np.random.randint(18, 70, size = zaznamy),
17     "Mesto": np.random.choice([
18        "Bratislava", "Košice", "Prešov", "Žilina", "Nitra",
19        "Banská Bystrica", "Trnava", "Martin", "Trenčín", "Poprad"
20    ], size = zaznamy),
21     "Nakupy": np.round(np.random.uniform(150.0, 1000.0, size =
22                         zaznamy), 2)
23 }
24
25 # Vytvorenie DataFrame
26 df_data_klienti = pd.DataFrame(data_klienti)
27
28 # Uloženie do CSV súboru
29 df_data_klienti.to_csv("dataset-100M-klienti-SK.csv", index=False)
```



Attached data

File Tools Kernel View Code Run Help

Build report Run all

DL MR

Files Databases Cloud storages

Notebook files

| Name | Last modified | Size |
|----------------------------|---------------|----------|
| .private | seconds ago | |
| bokeh_plot.html | 20 hours ago | 5.5 kB |
| environment.yml | 20 hours ago | 110.0 B |
| zamestnanci.csv | 20 hours ago | 154.0 B |
| export_data.xml | 20 hours ago | 56.4 kB |
| pandas_zamestnanci.csv | 20 hours ago | 195.0 B |
| export_data.xlsx | 20 hours ago | 13.4 kB |
| moje_pole.npy | 20 hours ago | 168.0 B |
| export_data.html | 20 hours ago | 36.1 kB |
| expexport_data.sql | 20 hours ago | 56.4 kB |
| export_data.json | 20 hours ago | 21.4 kB |
| test_data_spark.csv | 20 hours ago | 135.0 B |
| large_test_data_spark.csv | 20 hours ago | 232.5 MB |
| dataset-10M-klienti-SK.csv | 19 hours ago | 259.8 MB |
| dataset-1M-klienti-SK.csv | 19 hours ago | 26.0 MB |

Private workspace files

Vysokovýkonné počítanie v Python

Apache Spark Sheet 9 Knižnice Plotly Interaktívne Plotly Príklady Bokeh Prehľad Bokeh Interaktívne Bokeh Komplex Pandas +

```
[2] 27.6s
1 import pandas as pd
2 import numpy as np
3
4 # Vytvorenie datasetu s 10 miliónmi riadkov
5 zaznamy = 100_000_000
6
7 # Generovanie dát
8 data_klienti = {
9     "Meno": np.random.choice([
10         "Adriana", "Alena", "Andrea", "Anna", "Barbora", "Beáta",
11         "Daniela", "Dominika", "Elena", "Emília", "Erika", "Eva",
12         "Gabriela", "Hana", "Ivana", "Jana", "Katarína", "Kristína",
13         "Laura", "Lenka", "Lucia", "Magdaléna", "Mária", "Martina",
14         "Monika", "Natália", "Petra", "Renáta", "Silvia", "Simona", "Soňa",
15         "Tamara", "Veronika", "Zuzana"], size = zaznamy),
16     "Vek": np.random.randint(18, 70, size = zaznamy),
17     "Mesto": np.random.choice([
18         "Bratislava", "Košice", "Prešov", "Žilina", "Nitra",
19         "Banská Bystrica", "Trnava", "Martin", "Trenčín", "Poprad"
20     ], size = zaznamy),
21     "Nakupy": np.round(np.random.uniform(150.0, 1000.0, size =
22     zaznamy), 2)
23 }
24
25 # Vytvorenie DataFrame
26 df_data_klienti = pd.DataFrame(data_klienti)
27
28 # Uloženie do súboru
29 df_data_klienti.to_csv("dataset-100M-klienti-SK.csv", index=False)
```

Kernel failed

CPU cores 4 1%
Total RAM 16 GB 4 GB
AWS name m6i.xlarge
Change machine

CPU L CPU R



Vysokovýkonné počítanie v Pythone

Apache Spark

Generovanie Dát

Knižnice

Plotly Interaktivne

Plotly Príklady

Bokeh Prehľad

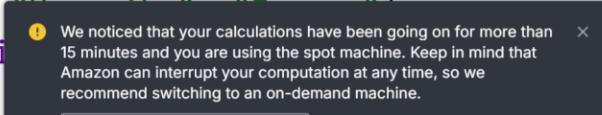
Bokeh Interaktivne

Bokeh Komplex

Pandas



```
1 import pandas as pd
2 import numpy as np
3
4 # Počet záznamov
5 zaznamy = 1_000_000_000
6 # Veľkosť jednej časti
7 velkost_casti = 10_000_000
8
9 # Cesta k súboru
10 csv_soubor = "dataset-1000M-klienti-SK.csv"
11
12 # Nastavíme, aby sa súbor prepisoval od začiatku, keďže budeme
    zapisovať po častiach
13 with open(csv_soubor, mode="a") as f:
14     # Vytvorenie dát a zápis po častiach
15     for i in range(zaznamy // velkost_casti):
16         data_klienti = {
17             "Meno": np.random.choice([
18                 "Adriana", "Alena", "Andrea", "Anna", "Barbora",
19                 "Beáta", "Daniela", "Dominika",
20                 "Elena", "Emília", "Erika", "Eva", "Gabriela",
21                 "Hana", "Ivana", "Jana", "Katarína",
22                 "Kristína", "Laura", "Lenka", "Lucia", "Magdaléna",
23                 "Mária", "Martina", "Monika",
24                 "Natália", "Petra", "Renáta", "Silvia", "Simona",
25                 "Soňa", "Tamara", "Svetlana", "Viktória", "Zuzana"],
```



| | |
|------------|-------|
| Running | |
| CPU | 2% |
| RAM | 83 GB |
| Calculated | 0 |
| In process | 1 |
| Errors | 0 |
| CPU XXL | |



We noticed that your calculations have been going on for more than 15 minutes and you are using the spot machine. Keep in mind that Amazon can interrupt your computation at any time, so we recommend switching to an on-demand machine.



Switch to on-demand machine

Don't ask again

Upozornenie na otvorenie príliš veľkého súboru

X

Otvorenie príliš veľkého súboru s veľkosťou viac ako 2 GB môže trvať niekoľko minút.

Chcete ho otvoriť?

Áno

Nie

Plugin Crash

X



CSVLint.dll just crashed in
`notify(SCNotification *notification) :`
`notification->nhdr.code == 1010`
`notification->nhdr.hwndFrom == 000000000390A7E`
`notification->nhdr.idFrom == 1941852424624`

OK

| Faktor | Pandas | Polars | Apache Spark |
|-----------------|----------------------------------------------------------------------------------------------|-----------------------------------------------------------------------|-----------------------------------------------------------------|
| Typ | Knižnica pre analýzu dát v pamäti | Knižnica pre analýzu dát v pamäti a na disku (paralelne) | Distribuovaný rámec pre spracovanie veľkých dát |
| Použitie | Menšie až stredne veľké dáta, analytika na jednom počítači | Veľké dátá, paralelné spracovanie a vyšší výkon v pamäti | Veľké dátá, distribuované spracovanie na viacerých uzloch |
| Výhody | - Jednoduché API, bohatá dokumentácia - Široké využitie v analýze dát | - Veľmi rýchle pre veľké dátá - Podporuje viacvláknové spracovanie | - Škálovateľnosť na clustery - Silná podpora pre SQL a ML |
| Nevýhody | - Slabšia škálovateľnosť pri veľkých dátach - Nízka rýchlosť pri veľmi veľkých datasetoch | - Mladšia knižnica, menej rozšírená - môže mať menšiu komunitu | - Vyžaduje zložité nasadenie - Vyššie nároky na konfiguráciu |
| Optimalizácia | - Často obmedzené RAM, vyžaduje optimalizáciu pre veľké dátá | - Rýchle operácie v pamäti, paralelné spracovanie | - Vysoce škálovateľné pre veľké clustery |
| Paralelizácia | - môže využiť len jedno jadro (súčasná verzia) | - Plne paralelizované spracovanie dát v rámci jedného počítača | - Paralelné spracovanie na viacerých uzloch distribučne |
| Spracovanie dát | - Pre dátá, ktoré sa zmestia do pamäte počítača | - Efektívne spracovanie veľkých dát, podporuje diskové operácie | - Efektívne spracovanie a ukladanie dát na HDFS |

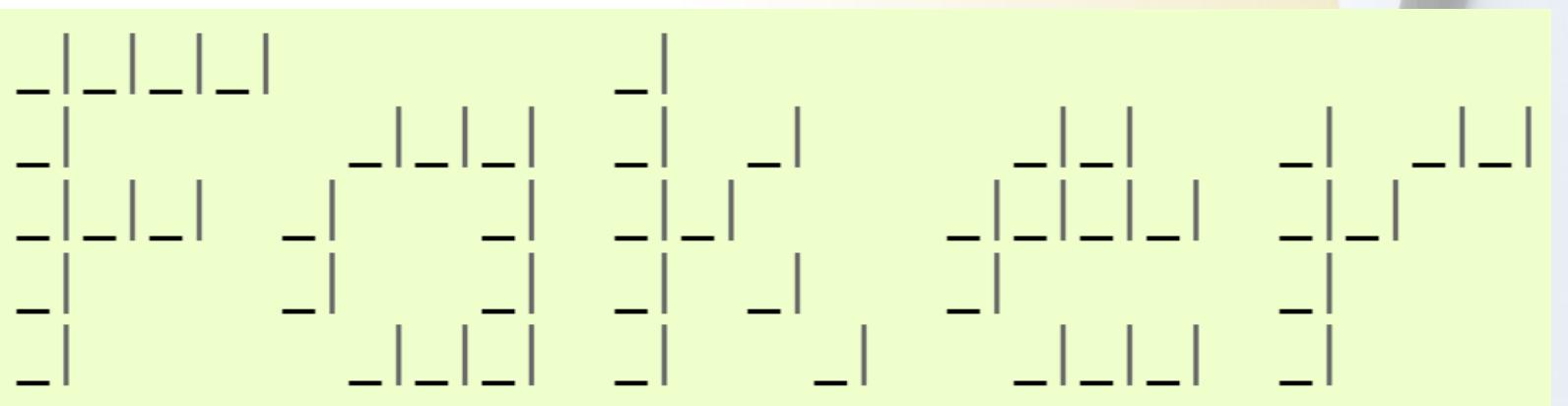
| Faktor | Pandas | Polars | Apache Spark |
|----------------------|-----------------------------------------------------------------------|-----------------------------------------------------------|------------------------------------------------------------|
| Kompatibilita | - Funguje s inými knižnicami ako Matplotlib, Scikit-learn | - Kompatibilné s Pandas API, môže byť náhrada za Pandas | - Podporuje prácu s HDFS, Hive, Cassandra, MLlib a ďalšími |
| Podpora pre SQL | - Obmedzená, vyžaduje externé knižnice (napr. <code>pandasql</code>) | - Obmedzená podpora SQL, ale môže byť rozšírená cez PySQL | - Silná podpora SQL (Spark SQL) |
| Prípad použitia | - Rýchla analýza a vizualizácia malých až stredne veľkých dát | - Práca s veľkými datasetmi, rýchla analýza, vysoký výkon | - Distribuované spracovanie obrovských dát na clusteroch |
| Knižnice/Technológie | - NumPy, Matplotlib, Scikit-learn | - NumPy kompatibilita, Rust backend pre vyšší výkon | - Spark MLlib, Hadoop, Hive, Kafka |

Úlohy Big Data

1. Pochopenie čo je big data
2. Prehľad nástrojov pre BD
3. Prehľad knižníc pre BD
4. Pochopenie úzkych hrdiel
5. Používanie techník na prácu s BD

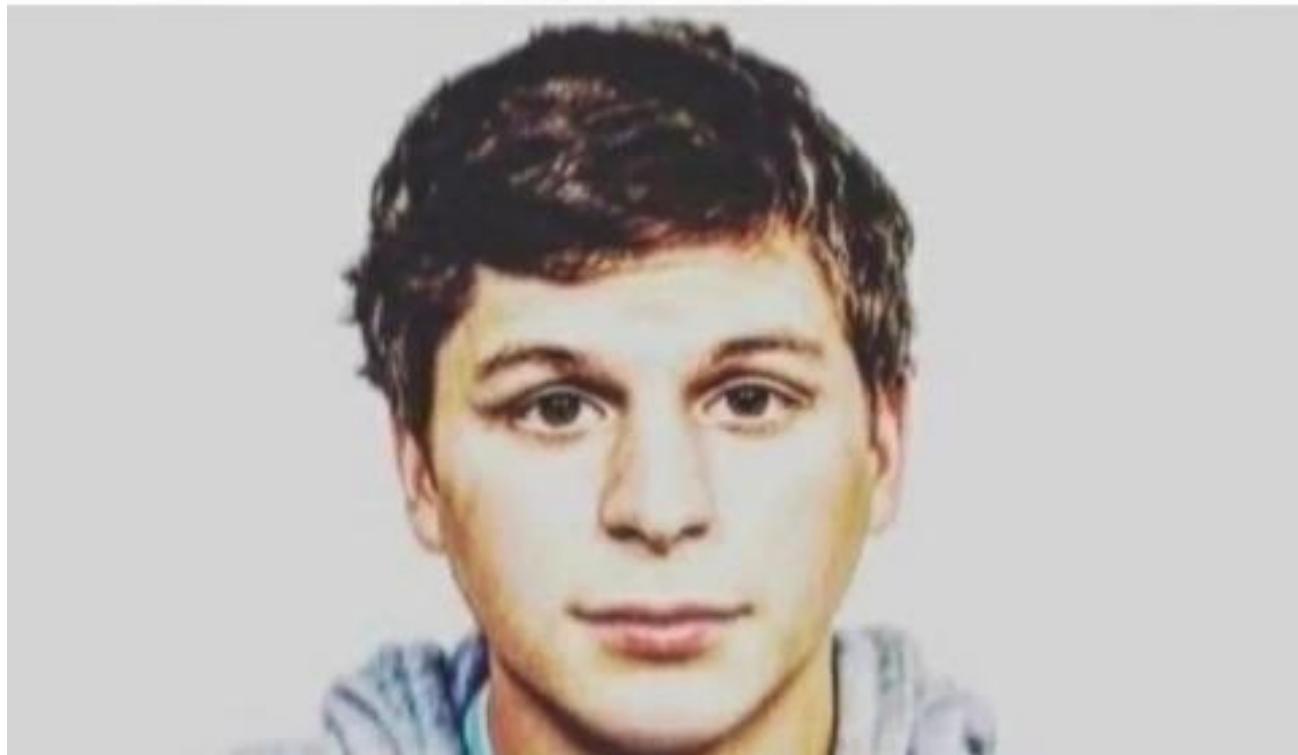


Testovacie Dáta a Generovanie



29. A B C D E
30. A B C D E
31. A B C D E
32. A B C D E
33. A B C D E
34. A B C D E
35. A B C D E
36. A B C D E
37. A B C D E
38. A B C D E
39. A B C D E
40. A B C D E
41. A B C D E
42. A B C D E
43. A B C D E
44. A B C D E
45. A B C D E
46. A B C D E
47. A B C D E
48. A B C D E
49. A B C D E
50. A B C D E
51. A B C D E
52. A B C D E
53. A B C D E
54. A B C D E
55. A B C D E
56. A B C D E
57. A B C D E
58. A B C D E
59. A B C D E
60. A B C D E
61. A B C D E
62. A B C D E
63. A B C D E
64. A B C D E
65. A B C D E
66. A B C D E
67. A B C D E
68. A B C D E
69. A B C D E
70. A B C D E
71. A B C D E
72. A B C D E
73. A B C D E
74. A B C D E
75. A B C D E
76. A B C D E
77. A B C D E
78. A B C D E
79. A B C D E
80. A B C D E
81. A B C D E
82. A B C D E
83. A B C D E
84. A B C D E
85. A B C D E
86. A B C D E
87. A B C D E
88. A B C D E
89. A B C D E
90. A B C D E
91. A B C D E
92. A B C D E
93. A B C D E
94. A B C D E
95. A B C D E
96. A B C D E
97. A B C D E
98. A B C D E
99. A B C D E
100. A B C D E

Is the Cure to Slow Bad Code using faster hardware?



Knižnica Faker



- Populárna **knižnica v Pythone**, ktorá slúži na **generovanie falošných údajov**, ako sú **mená, adresy, e-maily, telefónne čísla** atď.
- Je veľmi užitočná **pri testovaní aplikácií, generovaní testovacích dát alebo pri simulácii reálnych používateľov**
- Rýchlo a jednoducho **vytvárať realisticky vyzerajúce**, ale **úplne náhodné údaje**



Search projects



itacademy ▾

Faker 30.8.2

pip install Faker



Latest version

Released: Oct 31, 2024

Faker is a Python package that generates fake data for you.

Navigation

Project description

Release history

Download files

Verified details ✓

These details have been [verified by PyPI](#)

Maintainers



curella.org



joke2k

Unverified details

*These details have **not** been verified by PyPI*

Project links



Project description

Faker is a Python package that generates fake data for you. Whether you need to bootstrap your database, create good-looking XML documents, fill-in your persistence to stress test it, or anonymize data taken from a production service, Faker is for you.

Faker is heavily inspired by [PHP Faker](#), [Perl Faker](#), and by [Ruby Faker](#).

```
-|-|-|-|      _|
 -|       _|-|-| _| _|      _|-| _| _|-|
 _|-|-| _| _| _|-| _|-| _|-| _|-|
 _|       _| _| _|-| _|      _|
 _|       _|-|-| _| _| _|-| _|-| _|
```

version v30.8.2 Python Tests passing coverage 96% license MIT

Compatibility

Starting from version 4.0.0, Faker dropped support for Python 2 and from version 5.0.0 only supports Python 3.7 and above. If you still need Python 2 compatibility, please install version 3.0.1 in the meantime, and please consider updating your codebase to support Python 3 so you can enjoy the latest features Faker has to offer. Please see the



Table of Contents

Welcome to Faker's documentation!

- Compatibility
 - Basic Usage
 - Pytest fixtures
 - Providers
 - Localization
 - Optimizations
 - Command line usage
 - How to create a Provider
 - How to create a Dynamic Provider
 - How to customize the Lorem Provider
 - How to use with Factory Boy
 - Accessing the *random* instance
 - Unique values
 - Seeding the Generator
 - Tests
 - Contribute
 - License
 - Credits
 - Contents
- Indices and tables

Next topic

[Using the Faker Class](#)

This Page

[Show Source](#)

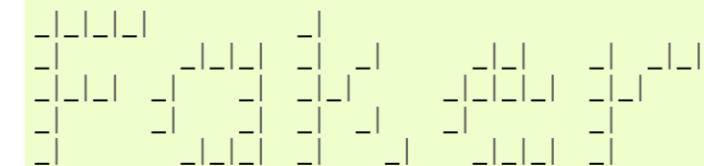
Quick search

Go

Welcome to Faker's documentation!

Faker is a Python package that generates fake data for you. Whether you need to bootstrap your database, create good-looking XML documents, fill-in your persistence to stress test it, or anonymize data taken from a production service, Faker is for you.

Faker is heavily inspired by [PHP Faker](#), [Perl Faker](#), and by [Ruby Faker](#).



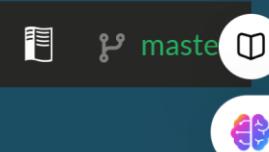
version [v30.8.2](#)  [Python Tests](#)  [passing](#) coverage  [96%](#) license  [MIT](#)

Compatibility

Starting from version `4.0.0`, Faker dropped support for Python 2 and from version `5.0.0` only supports Python 3.7 and above. If you still need Python 2 compatibility, please install version `3.0.1` in the meantime, and please consider updating your codebase to support Python 3 so you can enjoy the latest features Faker has to offer. Please see the [extended docs](#) for more details, especially if you are upgrading from version `2.0.4` and below as there might be breaking changes.

This package was also previously called `fake-factory` which was already deprecated by the end of 2016, and much has changed since then, so please ensure that your project and its dependencies do not depend on the old package.

Basic Usage



Používanie Faker



```
from faker import Faker
fake = Faker()

fake.name()
# 'Lucy Cechtelar'

fake.address()
# '426 Jordy Lodge
# Cartwrightshire, SC 88120-6700'

fake.text()
# 'Sint velit eveniet. Rerum atque repellat voluptatem quia rerum. Numquam excepturi
# beatae sint laudantium consequatur. Magni occaecati itaque sint et sit tempore. Nesciunt
# amet quidem. Iusto deleniti cum autem ad quia aperiam.
# A consectetur quo aliquam. In iste aliquid et aut similique suscipit. Consequatur qui
# quaerat iste minus hic expedita. Consequuntur error magni et laboriosam. Aut aspernatur
# voluptatem sit aliquam. Dolores voluptatum est.
# Aut molestias et maxime. Fugit autem facilis quo vero. Eius quibusdam possimus est.
# Ea quaerat et quisquam. Deleniti sunt quam. Adipisci consequatur id in occaecati.
# Et sint et. Ut ducimus quod nemo ab voluptatum.'
```

Faker



```
for _ in range(10):  
    print(fake.name())  
  
# 'Adaline Reichel'  
# 'Dr. Santa Prosacco DVM'  
# 'Noemy Vandervort V'  
# 'Lexi O'Conner'  
# 'Gracie Weber'  
# 'Roscoe Johns'  
# 'Emmett Lebsack'  
# 'Keegan Thiel'  
# 'Wellington Koelpin II'  
# 'Ms. Karley Kiehn V'
```

Prečo si vybrať Faker?

Rýchlosť generovania údajov

Široká škála údajov

Podpora viacerých jazykov a
regiónov

Flexibilita a prispôsobenie

Použitie v testovaní

Open source a zadarmo

Faker Limity



1. Pevne definované šablóny
2. Unikátnosť údajov
3. Výkon pri veľkých objemoch dát
4. Podpora pre špecifické jazyky a regióny
5. Náhodnosť a realita
6. Kvalita generovaných obrázkov
7. Závislosť na externej knižnici

Vyhnutie sa obmedzeniam

- Kontrola unikátnosti
- Optimalizácia výkonu
- Úprava šablón

Úlohy Faker

1. Generovanie testovacích dát
2. Pochopenie čo je faker
3. Používanie knižnice faker
4. Pochopenie limitov



Python Polars



Čo je to Polars?



- Knižnica pre spracovanie a analýzu dát v Pythone
- Zameraná na vysokú rýchlosť a efektívnosť, najmä pri práci s veľkými dátovými súbormi
- Navrhnuté ako alternatíva k populárnej knižnici Pandas
- Dôraz na optimalizáciu výkonu a využitie moderných technológií ako je paralelné spracovanie a využívanie viacerých jadier CPU

Kľúčové vlastnosti Polars



1. Rýchlosť
2. Efektívnosť pri spracovaní veľkých dát
3. Lazy Evaluation
4. Kategórie a optimalizácie pre textové dáta
5. Podpora pre paralelizáciu a distribúciu
6. Podpora pre čítanie a zápis do viacerých formátov
7. Kompatibilita s Pandas

Typické použitie Polars

1.  Analytické nástroje
2.  Strojové učenie
3.  Spracovanie logov
4.  Finančná analýza
5.  Práca s dátami IoT
6.  Analýza transakcií
7.  Geo dáta a mapy
8.  Bioinformatika
9.  Zdravotníctvo a med. analýza
10.  Logistika a optimalizácia trás
11.  E-commerce a analýza zákaz.
12.  Výskum a akademická analýza
13.  Siet'ová analytika a security
14.  Smart Cities a mestské dáta
15.  Agrikultúra a pol'nohospod. dáta

DataFrames for the new era

75M+ Downloads to date 29k+ Github stars

Get started

Polars is an open-source library for data manipulation, known for being one of the fastest data processing solutions on a single machine. It features a well-structured, typed API that is both expressive and easy to use.



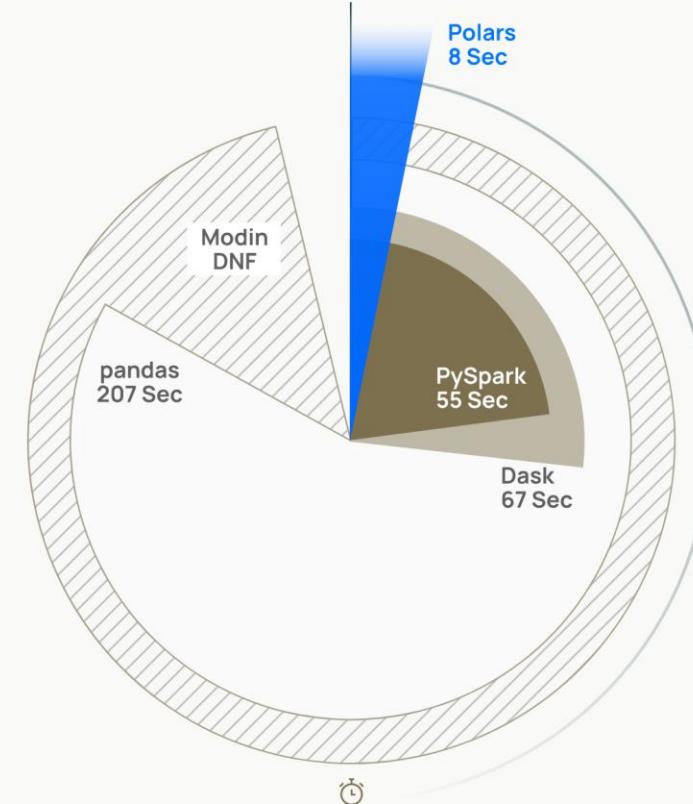
Built by developers for developers to achieve up to 50x performance

WHY USE POLARS

Polars was benchmarked in a derived version of the independent TPC-H benchmark against several other solutions. This benchmark aims to replicate data wrangling operations used in practice. Polars easily trumps other solutions due to its parallel execution engine, efficient algorithms and use of vectorization with SIMD (Single Instruction, Multiple Data). Compared to pandas, it can achieve more than 30x performance gains.

[Learn more →](#)

The benchmark was performed on a c3-highmem-22 with scale factor 10 including I/O. The queries are open source.



Compute Engine machine families



Univerzálny

- Najlepší pomer ceny a výkonu pre rôzne pracovné zaťaženia

Optimalizované úložisko

- Najlepšie pre pracovné zaťaženie, ktoré má nízku spotrebu jadra a vysokú hustotu úložiska

Výpočtovo optimalizované

- Najvyšší výkon na jadro na Compute Engine a optimalizované pre výpočtovo náročné pracovné zaťaženie

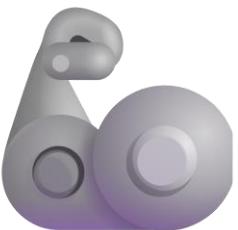
Pamäťovo optimalizované

- Ideálne pre úlohy náročné na pamäť, ponúka viac pamäte na jadro ako iné rodiny počítačov, s pamäťou až 12 TB

Optimalizované pre akcelerátor

- Ideálne pre masívne paralelizované výpočtové úlohy CUDA (Compute Unified Device Architecture), ako je strojové učenie (ML) a vysokovýkonné výpočty (HPC)
- Táto rodina je najlepšou voľbou pre úlohy, ktoré vyžadujú GPU

Predefined machine types



highcpu

- from 1 to 3 GB memory per vCPU; typically, 2 GB memory per vCPU

standard

- from 3 to 7 GB memory per vCPU; typically, 4 GB memory per vCPU

highmem

- from 7 to 14 GB memory per vCPU; typically, 8 GB memory per vCPU

megamem

- from 14 to 19 GB memory per vCPU

hypermem

- from 19 to 24 GB memory per vCPU; typically, 21 GB memory per vCPU

ultramem

- from 24 to 31 GB memory per vCPU

| General-purpose workloads | | | |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| N4, N2, N2D, N1 | C4A, C4, C3, C3D | E2 | Tau T2D, Tau T2A |
| Balanced price/performance across a wide range of machine types | Consistently high performance for a variety of workloads | Day-to-day computing at a lower cost | Best per-core performance/cost for scale-out workloads |
| <ul style="list-style-type: none"> • Medium traffic web and app servers • Containerized microservices • Business intelligence apps • Virtual desktops • CRM applications • Development and test environments • Batch processing • Storage and archive | <ul style="list-style-type: none"> • High traffic web and app servers • Databases • In-memory caches • Ad servers • Game Servers • Data analytics • Media streaming and transcoding • CPU-based ML training and inference | <ul style="list-style-type: none"> • Low-traffic web servers • Back office apps • Containerized microservices • Microservices • Virtual desktops • Development and test environments | <ul style="list-style-type: none"> • Scale-out workloads • Web serving • Containerized microservices • Media transcoding • Large-scale Java applications |

| Optimized workloads | | | |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Storage-optimized | Compute-optimized | Memory-optimized | Accelerator-optimized |
| Z3 | H3, C2, C2D | X4, M3, M2, M1 | A3, A2, G2 |
| Highest block storage to compute ratios for storage-intensive workloads | Ultra high performance for compute-intensive workloads | Highest memory to compute ratios for memory-intensive workloads | Optimized for accelerated high performance computing workloads |
| <ul style="list-style-type: none"> • File servers • Flash-optimized databases • Scale-out analytics • Other databases | <ul style="list-style-type: none"> • Compute-bound workloads • High-performance web servers • Game Servers • High performance computing (HPC) • Media transcoding • Modeling and simulation workloads • AI/ML | <ul style="list-style-type: none"> • Medium to extra-large SAP HANA in-memory databases • In-memory data stores, such as Redis • Simulation • High Performance databases such as Microsoft SQL Server, MySQL • Electronic design automation | <ul style="list-style-type: none"> • Generative AI models such as the following: <ul style="list-style-type: none"> • Large Language Models (LLM) • Diffusion Models • Generative Adversarial Networks (GAN) • CUDA-enabled ML training and inference • High-performance computing (HPC) • Massively parallelized computation • BERT natural language processing • Deep learning recommendation model (DLRM) • Video transcoding • Remote visualization workstation |

Leading companies using Polars

Optiver 

 RESEARCH

 Microsoft

NETFLIX
appian



CHECK.
showmax





ORACLE



SUPPORT

Support for all common data formats

Polars supports reading and writing to all common data formats. This allows you to easily integrate Polars with your existing data stack.

- ✓ Text: CSV & JSON
- ✓ Binary: Parquet, Delta Lake, AVRO & Excel
- ✓ IPC: Feather, Arrow
- ✓ Databases: MySQL, Postgres, SQL Server, Sqlite, Redshift & Oracle
- ✓ Cloud storage: S3, Azure Blob & Azure File



| Faktor | Pandas | Polars | Apache Spark |
|-----------------|----------------------------------------------------------------------------------------------|-----------------------------------------------------------------------|-----------------------------------------------------------------|
| Typ | Knižnica pre analýzu dát v pamäti | Knižnica pre analýzu dát v pamäti a na disku (paralelne) | Distribuovaný rámec pre spracovanie veľkých dát |
| Použitie | Menšie až stredne veľké dáta, analytika na jednom počítači | Veľké dátá, paralelné spracovanie a vyšší výkon v pamäti | Veľké dátá, distribuované spracovanie na viacerých uzloch |
| Výhody | - Jednoduché API, bohatá dokumentácia - Široké využitie v analýze dát | - Veľmi rýchle pre veľké dátá - Podporuje viacvláknové spracovanie | - Škálovateľnosť na clustery - Silná podpora pre SQL a ML |
| Nevýhody | - Slabšia škálovateľnosť pri veľkých dátach - Nízka rýchlosť pri veľmi veľkých datasetoch | - Mladšia knižnica, menej rozšírená - môže mať menšiu komunitu | - Vyžaduje zložité nasadenie - Vyššie nároky na konfiguráciu |
| Optimalizácia | - Často obmedzené RAM, vyžaduje optimalizáciu pre veľké dátá | - Rýchle operácie v pamäti, paralelné spracovanie | - Vysoce škálovateľné pre veľké clustery |
| Paralelizácia | - môže využiť len jedno jadro (súčasná verzia) | - Plne paralelizované spracovanie dát v rámci jedného počítača | - Paralelné spracovanie na viacerých uzloch distribučne |
| Spracovanie dát | - Pre dátá, ktoré sa zmestia do pamäte počítača | - Efektívne spracovanie veľkých dát, podporuje diskové operácie | - Efektívne spracovanie a ukladanie dát na HDFS |

| Faktor | Pandas | Polars | Apache Spark |
|----------------------|-----------------------------------------------------------------------|-----------------------------------------------------------|------------------------------------------------------------|
| Kompatibilita | - Funguje s inými knižnicami ako Matplotlib, Scikit-learn | - Kompatibilné s Pandas API, môže byť náhrada za Pandas | - Podporuje prácu s HDFS, Hive, Cassandra, MLlib a ďalšími |
| Podpora pre SQL | - Obmedzená, vyžaduje externé knižnice (napr. <code>pandasql</code>) | - Obmedzená podpora SQL, ale môže byť rozšírená cez PySQL | - Silná podpora SQL (Spark SQL) |
| Prípad použitia | - Rýchla analýza a vizualizácia malých až stredne veľkých dát | - Práca s veľkými datasetmi, rýchla analýza, vysoký výkon | - Distribuované spracovanie obrovských dát na clusteroch |
| Knižnice/Technológie | - NumPy, Matplotlib, Scikit-learn | - NumPy kompatibilita, Rust backend pre vyšší výkon | - Spark MLlib, Hadoop, Hive, Kafka |

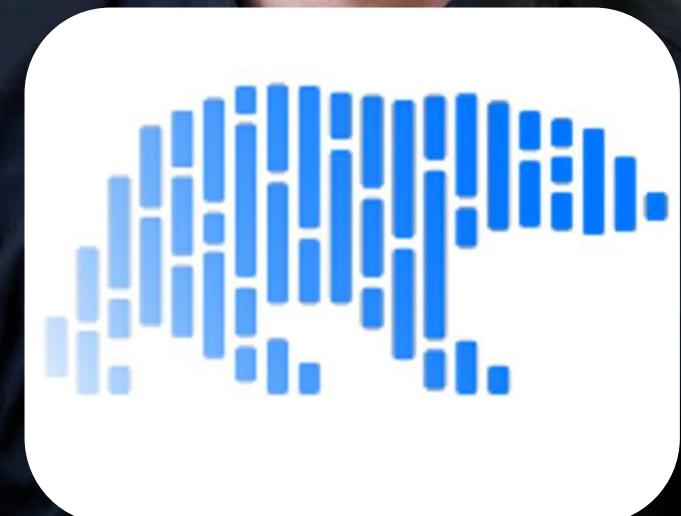
Úlohy Polars

1. Pochopenie čo je Polars
2. Používanie knižnice Polars
3. Filtrovanie, agregácie
4. Generovanie testovacích dát
5. Pochopenie limitov
6. Porovnanie s Pandas a Apache Spark



Ako pracovať s Big Data Súbormi

AKREDITOVANÝ KURZ



Čo je to Parquet?



- Formát súboru určený na ukladanie a spracovanie veľkých dát
- Je optimalizovaný pre rýchle čítanie a zápis dát v distribuovaných systémoch
- V Pythone sa používa najmä v spojení s knižnicami ako pandas, pyarrow, fastparquet na čítanie a zápis dát do tohto formátu
- Stĺpcový formát, čo znamená, že dáta sú uložené v stĺpcoch, čo zvyšuje efektívnosť pri analýze veľkých datasetov
- Používa sa najmä v Big Data aplikáciách



Search projects



itacademy ▾

parquet 1.3.1

pip install parquet



Latest version

Released: Apr 30, 2020

Python support for Parquet file format

Navigation

Project description

Release history

Download files

Verified details

These details have been [verified by PyPI](#)

Maintainers



jcrobak



turicas

Unverified details

*These details have **not** been verified by PyPI*

Project links

[Homepage](#)

Project description

parquet-python

build unknown

parquet-python is a pure-python implementation (currently with only read-support) of the [parquet format](#). It comes with a script for reading parquet files and outputting the data to stdout as JSON or TSV (without the overhead of JVM startup). Performance has not yet been optimized, but it's useful for debugging and quick viewing of data in files.

Not all parts of the parquet-format have been implemented yet or tested e.g. nested data—see Todos below for a full list. With that said, parquet-python is capable of reading all the data files from the [parquet-compatibility](#) project.

requirements

parquet-python has been tested on python 2.7, 3.6, and 3.7. It depends on [pythrift2](#) and optionally on [python-snappy](#) (for snappy compressed files, please also install [parquet-python\[snappy\]](#)).

getting started

parquet-python is available via PyPi and can be installed using `pip install parquet`. The package includes the `parquet` command for reading python files, e.g. `parquet test.parquet`. See `parquet -help` for full usage.



1 000 000 Záznamov

| | | |
|----------------------------------------------------------------------------------------------------------------|-------------------------------|----------|
|  export-kurzy-velke.csv | 3 minutes ago | 34.1 MB |
|  export-kurzy-velke.json | 2 minutes ago | 87.0 MB |
|  export-kurzy-velke.xml | 2 minutes ago | 182.9 MB |
|  export-kurzy-velke.parquet | 2 minutes ago | 2.2 MB |
|  export-kurzy-velke.arrow | 2 minutes ago | 67.3 MB |
|  export-kurzy-velke.orc | 2 minutes ago | 30.0 MB |
|  export-kurzy-velke.feather | 2 minutes ago | 19.1 MB |
|  export-kurzy-velke.xlsx | seconds ago <small>78</small> | 19.6 MB |

Porovnanie Súborov



| Vlastnosť | Parquet | ORC | Feather | IPC | CSV | JSON |
|---------------|---------------------------------------------------------------------------|-----------------------------------------------------------------------|----------------------------------------------------------------------|------------------------------------------------------------------|-----------------------------------------------------------|-----------------------------------------------------------------------|
| Použitie | Rýchle uloženie a načítanie veľkých dát | Veľmi efektívne pre analytické účely | Rýchla sérializácia dát medzi aplikáciami | Ukladanie dát vo forme stĺpcov, rýchle operácie | Bežne používaný formát pre jednoduché dátové súbory | Formát vhodný na prenos a archiváciu |
| Výhody | Vysoký výkon pri čítaní a zápisе, kompresia, efektívne pre veľké datasety | Kompresia, efektívne pre analytické operácie, optimalizované pre Hive | Rýchla sérializácia a deserializácia, kompatibilný s pandas a polars | Ideálne pre výpočty v reálnom čase, efektívny pri veľkých dátach | Jednoduchý a čitateľný pre ľudí, široká podpora | Podporuje flexibilitu dátového modelu, čitateľné |
| Nevýhody | Vyžaduje viac pamäte pri zápisе, menej čitateľný pre ľudí | Menej flexibilný, náročnejší na konfiguráciu | Nižšia podpora kompresie, obmedzené na malé dataset | Menej ľahko prenosný ako CSV alebo JSON | Pomalší zápis a čítanie veľkých dát | Môže byť pomalý pri veľmi veľkých datasetoch, nekomprimovaný |
| Kompatibilita | Kompatibilný s Hadoop, Spark, Dask, Polars | Používaný v ekosystéme Hadoop a pre analytiku | Dobrá podpora s pandas a Polars | Veľmi kompatibilný s Polars, Apache Arrow | Podporovaný takmer všetkými nástrojmi pre spracovanie dát | Vhodný pre prenos medzi aplikáciami, podporovaný vo väčšine nástrojov |
| Kompresia | Vysoká | Veľmi vysoká | Nie veľmi efektívna | Nie veľmi efektívna | Obvykle žiadna | Mierna, závisí od knižnice |

Porovnanie Súborov



| Vlastnosť | Parquet | ORC | Feather | IPC | CSV | JSON |
|------------------------|-------------------------------------|-----------------------------------------------------|---------------------------------------------------------------|------------------------------------------|------------------------------------|------------------------------------|
| Technológia | Columnar storage, Apache Arrow | Columnar storage, Hive | Feather je formát optimalizovaný pre prenos medzi aplikáciami | Ukladanie dát v stĺpcovom formáte | Bežný textový formát | Formát textových súborov |
| Formát dát | Binárny formát, stĺpcový | Binárny formát, stĺpcový | Binárny formát | Binárny formát | Textový formát | Textový formát |
| Podpora paralelizmu | Výborná (podporuje paralelný zápis) | Výborná (optimalizované pre distribučné prostredia) | Obmedzená paralelizácia | Vysoká (podporuje paralelné spracovanie) | Obmedzená (závisí od veľkosti dát) | Obmedzená pri veľkých datasetoch |
| Podpora pre veľké dátá | Vynikajúca | Vynikajúca | Dobrá | Vynikajúca | Obmedzená pri veľkých datasetoch | Obmedzená pri veľmi veľkých dátach |

+ Create

Home

Competitions

Datasets

Models

Code

Discussions

Learn

More

Your Work

VIEWED

Simulated Transactions



CONOR · UPDATED 3 YEARS AGO

▲ 17

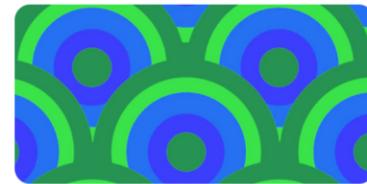
New Notebook

Download

⋮

Simulated Transactions

Randomly generated transaction data for 75000 customers

[Data Card](#) [Code \(0\)](#) [Discussion \(0\)](#) [Suggestions \(0\)](#)

About Dataset

NOTE: these transactions are randomly generated. The customers represented in the dataset are not real.

This is a large transaction dataset for data visualisation and processing tutorials. Transactions are generated for 75,000 customers and are classified into 12 expenditure types:

- Groceries
- Clothing
- Housing
- Education
- Health
- Motor/Travel
- Entertainment
- Gambling

Usability ⓘ

6.47

License

CC0: Public Domain

Expected update frequency

Not specified

Tags[Tabular](#) [Finance](#)
[Intermediate](#) [Banking](#)

CSV súbor s transakciami 22,29 GB (6 GB Zip)

260 mil. záznamov, 10 stĺpcov, 75 000 zákazníkov

Kompresné algoritmy parquet

| Algoritmus | Rýchlosť kompresie | Rýchlosť dekompresie | Kompresná miera | Použitie |
|------------|--------------------|----------------------|-------------------|-------------------------------------------------------|
| Snappy | Veľmi rýchla | Rýchla | Nízka | Systémy s nízkou latenciou, veľké dátá v reálnom čase |
| Gzip | Stredná | Veľmi rýchla | Stredná až vysoká | Bežné kompresné aplikácie, webové servery, databázy |
| Brotli | Pomalšia | Rýchla | Veľmi vysoká | Webové aplikácie, optimalizácia prenosu dát |

Úlohy Big Data Súbory

1. Práca s big data formátmi
2. Používanie knižnice polars a pyarrow (parquet)
3. Exporty do rôznych formátov
4. Porovnanie súborov
5. Pochopenie kompresne algoritmy



Dôležité Zdroje

Slovensko a Česko

- Albatrosmedia
- Kopp
- Grada
- Wolters Kluwer
- BEN
- Veda

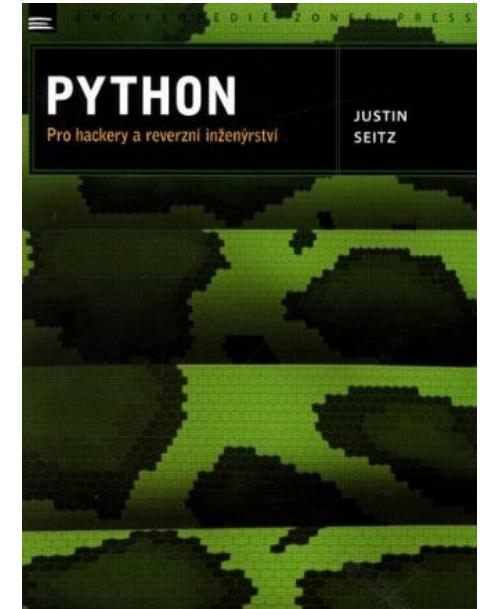
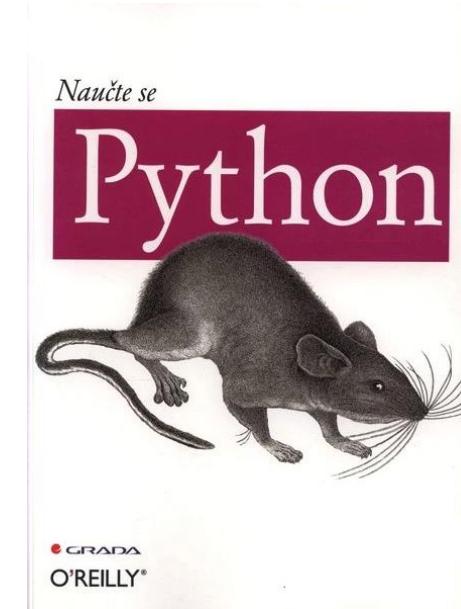
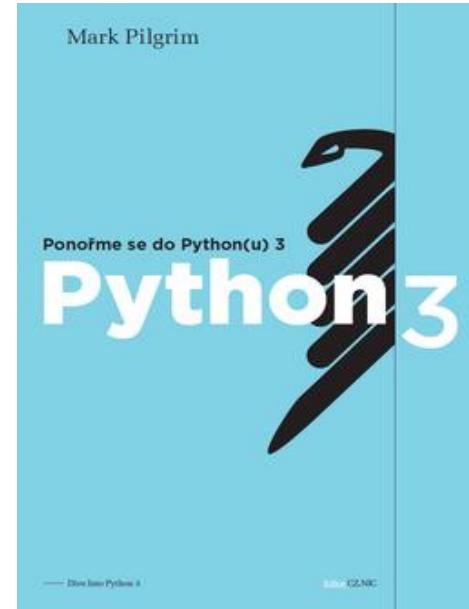
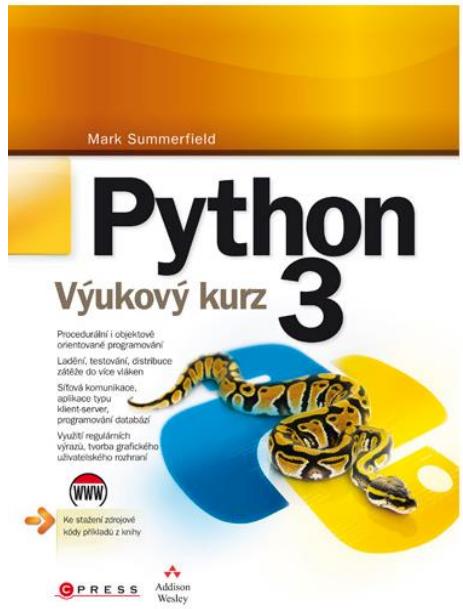
Zahraničie

- O'Reilly
- Manning
- Packt
- Apress
- Wiley
- No Starch Press

YouTube Tutoriály

- [IT Academy](#)

IT Academy



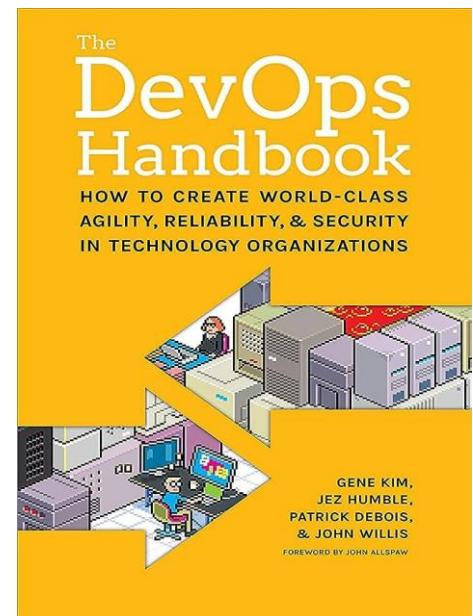
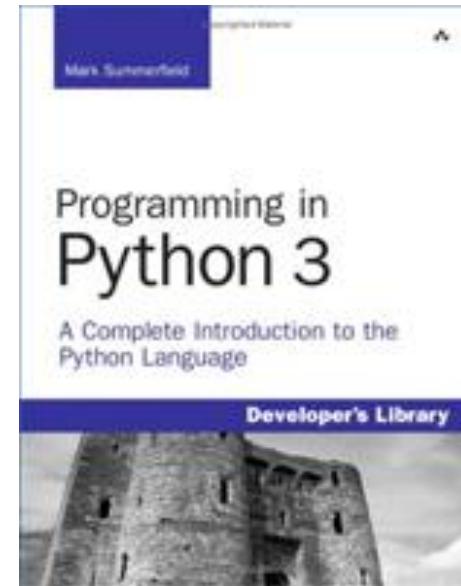
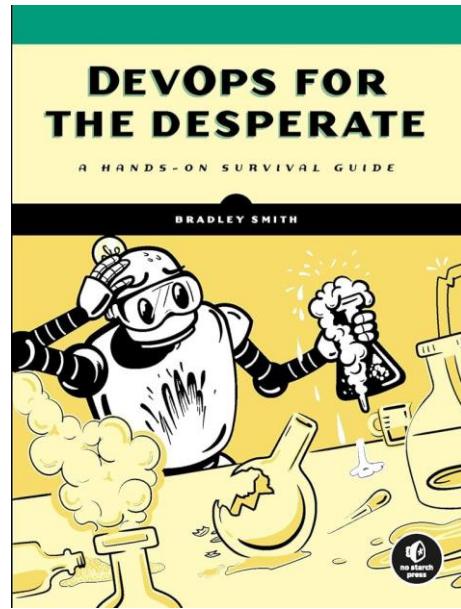
Čo sa oplatí/neoplatí prečítať SK/CZ?



Effective DevOps

BUILDING A CULTURE OF COLLABORATION,
AFFINITY, AND TOOLING AT SCALE

Jennifer Davis & Ryn Daniels



Čo sa oplatí/neoplatí prečítať EN?

Čo Odporúčam si Pozriet’?

1. <https://docs.python.org/3/>
2. <https://realpython.com/tutorials/best-practices/>
3. <https://google.github.io/styleguide/pyguide.html>
4. <https://docs.python.org/3/>
5. <http://python2013.input.sk/19prednaska>
6. <https://realpython.com/python3-object-oriented-programming/>
7. <https://jeffknupp.com/blog/2014/06/18/improve-your-python-python-classes-and-object-oriented-programming/>
8. <https://overiq.com/python-101/inheritance-and-polymorphism-in-python/>
9. <https://www.javatpoint.com/python-oops-concepts>
10. <https://www.programiz.com/python-programming/object-oriented-programming>



TOP

Klávesové Skratky



TOP Klávesové Skratky

Práca s IDE

- Ctrl + D Delete zmaž riadok
- **Ctrl + Space** Asistent kódu
- **Ctrl + /** Komentáre
- Ctrl + A Označ všetko
- **Alt + /** Dokonči slovo
- Ctrl + F Hľadanie a náhrady
- Ctrl + Shift + F Kompakt režim
- Ctrl + Shift + S Ulož všetko

Práca s browserom

- Ctrl + T Vytvor nový tab
- Ctrl + W Zatvor aktuálny tab
- Ctrl + Shift + W Zatvor všetky taby
- **Ctrl + Shift + T** Otvor posledný tab
- Ctrl + Shift + J/F12 Web console
- **F11** Fullscreen

Efektívne Používanie Klávesnice

Špeciálne znaky, kde ich nájst' na klávesnici

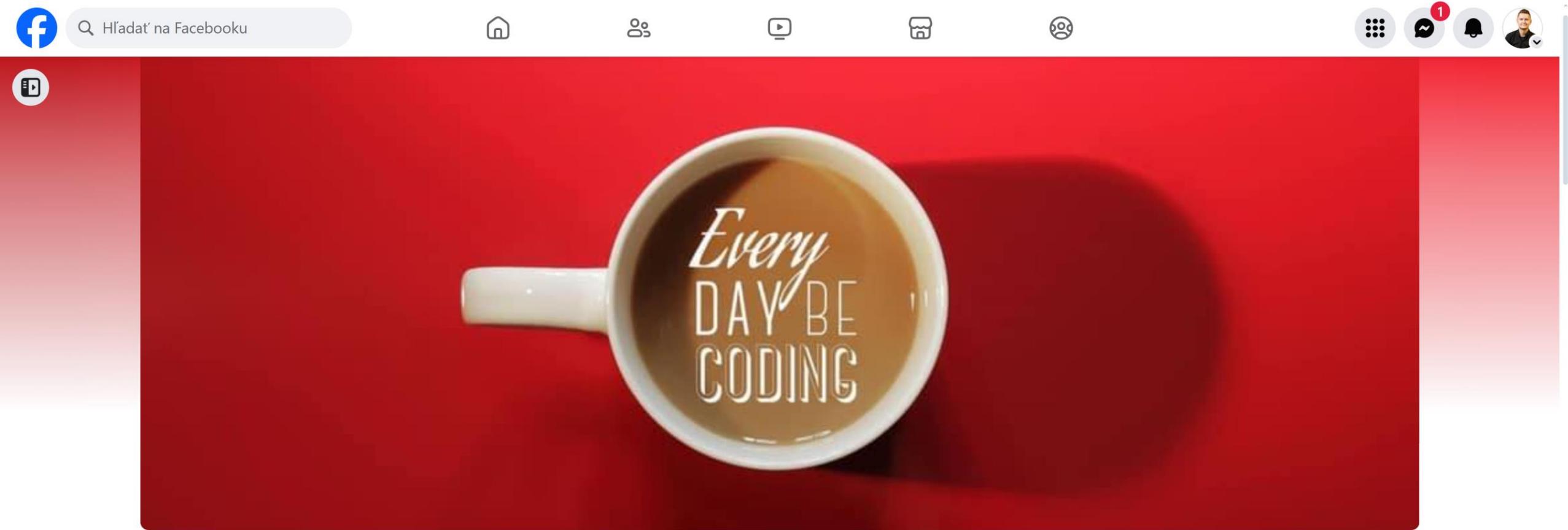
The diagram shows a standard QWERTY keyboard with various keys highlighted in different colors to represent different categories of special characters:

- Operátory (Operators):** +, *, -, /, %, @.
- Porovnávanie (Comparison):** <, >, =, !.
- Retázce (Text):** " (double quotes), \ (backslash).
- Oddelovače (Separators):**, ;, :.
- Bitové operácie (Bitwise Operations):** &, |, ^, ~, ?.
- Zátvorky (Brackets):** (,), {}, [].
- Poznámky (Annotations):** #, !, @, %, \$, ^, &, |, {, }, [,].
- Komentár (Comment):** /*, */.
- Pomocník (Helper):** ?.

Surrounding the keyboard are large, semi-transparent symbols representing common uses of these characters:

- #
- &
- !
- €

IT ACADEMY



Vývojári

Verejná skupina · 7,4 tis. členov

+ Pozvať

↗ Zdieľať

👤 Člen ▾

▼

Diskusia

Ludia

Podujatia

Médiá

Súbory



...



Napíšte niečo...



Reel



Fotka/video



Anketa

Informácie

Skupina softvérových vývojárov. Táto skupina by mala byť miestom, kde sa môžu českoslovení vývojári vzájomne spoznať, vyžiadať si konštrukívnu... [Zobrazit viac](#)

👤 Verejná

Členov skupiny a ich príspevky bude vidieť ktokoľvek.





Hľadať na Facebooku



```
52 </div>
53 </body>
54 <script type="text/javascript">
55 <!--
56 var currentImage = "bigImage1";
57 var pages = Math.ceil.photos.length / 9);
58 updatePages();
59 updateAllImages();
60 // document.getElementById('bigImage0').src = 'images/wieksza' + photos[page] + '1';
61 // document.getElementById('bigImage0').style.display = '';
62 changePhotoDescription('1');
63
64
65
66
67
68
69
70
71
72
73
```

Programátori

Verejná skupina · 11,1 tis. členov

+ Pozvat

Člen ▾



Diskusia

Vybrané

Ludia

Podujatia

Média

Súbory



Napíšte niečo...



Reel



Fotka/video



Anketa

Informácie

Táto skupina slúži na dohadzovanie si kšeftíkov a pre hľadačov programátorov / vývojárov.

Verejná

Členov skupiny a ich príspevky bude vidieť ktokoľvek.

Viditeľná

Príspevky sú viditeľní iba členom skupiny.



I am programmer



I have Life



I have
stackoverflow



IT ACADEMY

Home

PUBLIC

Questions

Tags

Users

COLLECTIVES

Explore Collectives

FIND A JOB

Jobs

Companies

TEAMS

Create free Team

Tags

A tag is a keyword or label that categorizes your question with other, similar questions. Using the right tags makes it easier for others to find and answer your question.

[Show all tag synonyms](#)

python

python

Python is a multi-paradigm, dynamically typed, multi-purpose programming language. It is designed to be quick to learn, understand, and...

1870168 questions 695 asked today, 6387 this week

python-3.x

USE ONLY IF YOUR QUESTION IS VERSION-SPECIFIC. For questions about Python programming that are specific to version 3+ of the language...

303562 questions 93 asked today, 836 this week

python-2.7

Python 2.7 is the last major version in the 2.x series, and is no longer maintained since January 1st 2020. Use the generic [python] tag on all Python...

94965 questions 24 asked this week, 106 this month

[Popular](#) [Name](#) [New](#)

python-requests

USE ONLY FOR THE PYTHON REQUESTS LIBRARY. Requests is a full-featured Python HTTP library with an easy-to-use, logical API.

18697 questions 8 asked today, 57 this week

python-imaging-library

The Python Imaging Library (PIL) provides the Python language with a de-facto standard foundation for image work. PIL's API is lightweight but...

7883 questions 5 asked today, 38 this week

wxpython

wxPython is a Python wrapper for the cross-platform C++ GUI API wxWidgets.

7047 questions 7 asked this week, 14 this month

ipython

IPython is a feature-rich interactive shell for Python, and provides a kernel for frontends such as IPython Notebook and Jupyter Notebook.

6886 questions 5 asked this week, 26 this month

python-3.6

Version of the Python programming language released in December 2016. For issues specific to Python 3.6. Use more generic [python] and [python-3....

5602 questions 11 asked this week, 24 this month

python-asyncio

to be used for the asyncio Python package which provides mechanisms for writing single-threaded concurrent code. The asyncio package provides...

5492 questions 29 asked this week, 125 this month

python-import

For questions about importing modules in Python

5119 questions 11 asked this week, 47 this month

python-multiprocessing

multiprocessing is a package that supports spawning processes using an API similar to the threading module in python programming language.

4036 questions 12 asked this week, 46 this month

python-3.7

Version of the Python programming language released in June 27, 2018. For issues that are specific to Python 3.7. Use the more generic [python] and...

4034 questions 5 asked this week, 21 this month

Ako Skončíme?

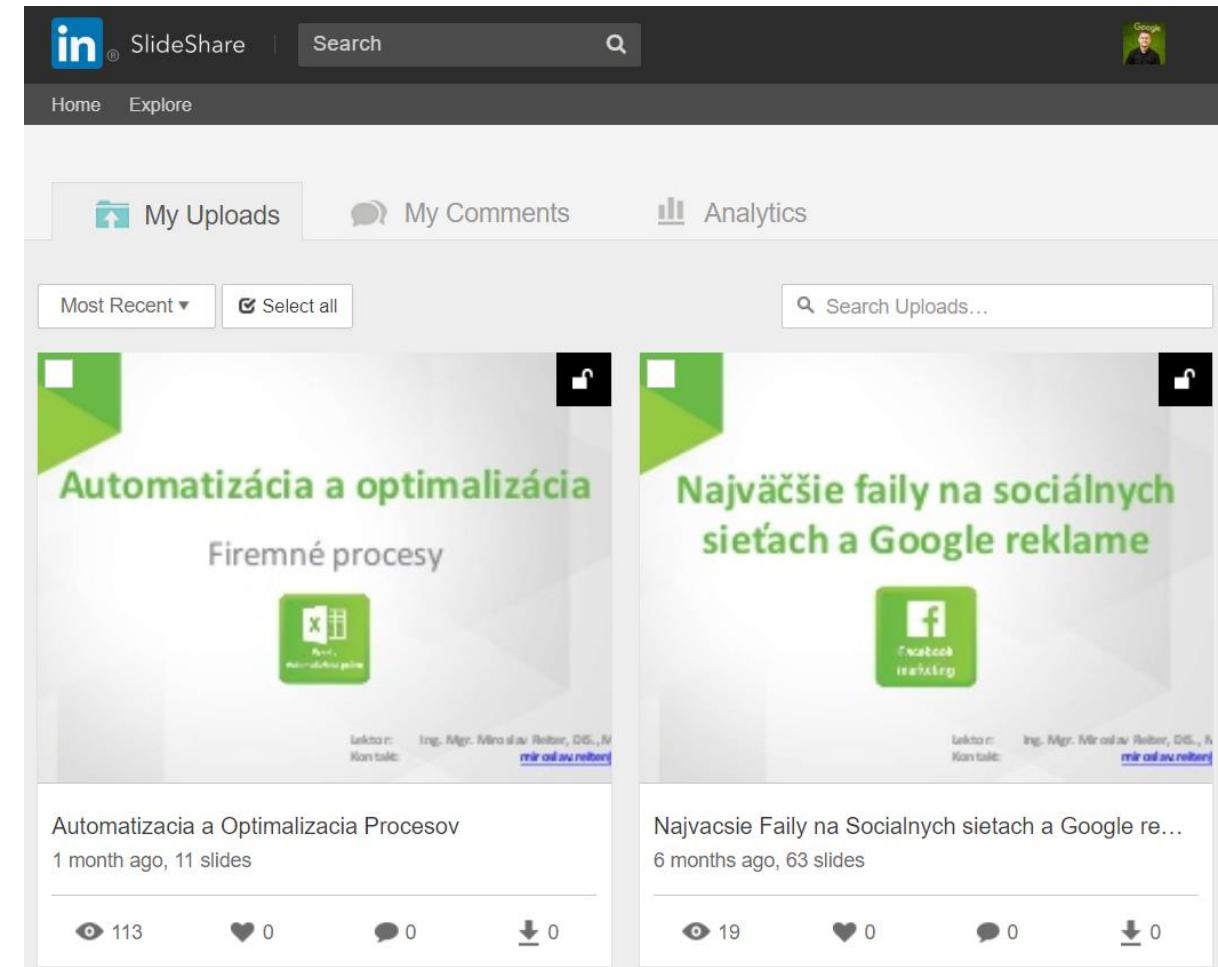
1. Stiahnite si Súbory

- https://github.com/miroslav-reiter/Kurzy_SAV_Analytika_Python_R

2. Pridajte si ma na LinkedIn

- www.linkedin.com/in/miroslav-reiter/

3. Prezentácia a materiály po prednáške



Vyber si online kurz

Nauč sa programovať, tvoriť webstránky a grafiku, manažovať alebo sa zameraj na osobný rozvoj. Všetko jednoducho vďaka našim online kurzom z pohodlia tvojho domova.

Ročné
predplatné na
všetky online
kurzy

~~2299.99€~~

399.99€

Priístup pre Teba do všetkých aktuálnych aj pripravovaných online kurzov

12 mesačná platnosť

🛒 Kúpiť teraz

Zadarmo

1. Kurzy SAV

2. YouTube kanál IT Academy

<https://www.youtube.com/c/IT-AcademySK>

Platené

Moje kurzy na www.vita.sk



Práca a
zamestnanosť



Rodina a sociálna
pomoc



Sociálne poistenie
a dôchodkový
systém



Ministerstvo

INŠPEKCIJA V SOCIÁLNYCH
VECIACH

Aktívne starnutie

Pomoc pre odídencov z Ukrajiny

Informácie pre médiá

Aktuality

Informácie pre verejnosť

Európsky orgán práce ELA

Pracovná mobilita cudzincov

Elektronické služby, formuláre a
žiadosti

Legislatíva

Organizácie ministerstva

Programy Európskej únie

Plán obnovy a odolnosti

Potravinová a základná
materiálna pomoc

Národné kompetenčné centrum
pre sociálne inovácie

Neprehliadnite

Podcasty o vývoji na trhu práce

Novačník milionára

19.03.2024

Úrady práce sociálnych vecí a rodiny podporia rozvoj zručnosti

Nový národný projekt Zručnosti pre trh práce prinesie už od 8.apríla možnosť bezplatne získať finančnú podporu na vzdelávanie či rekvalifikáciu. Je to prvý projekt svojho druhu, ktorý je určený pre ľudí bez práce, ale aj pre zamestnaných.

„Podobné projekty boli v minulosti obľúbené a fungovali, no vlane skončili. Jednou z našich priorit hned po nástupe na ministerstvo bolo opäťovne a v čo najkratšom čase podobný projekt opäť rozbehnuť. A tak sme pripravili národný projekt Zručnosti pre trh práce. Vzdelávanie je klúčovým nástrojom na podporu nezamestnaných pri ich ceste späť na trh práce, ako aj pre ľudí v ohrození stratou zamestnania. Tak ako pri ostatných nástrojoch, aj v tomto prípade sme urobili viac a priniesli inováciu s cieľom zvýšiť efektivitu a kvalitu služieb pre občanov a v jednom projekte spájame podporu pre ľudí bez práce aj zamestnaných. Okrem toho majú ľudia možnosť využiť tento projekt opakovane,“ povedal minister práce, sociálnych vecí a rodiny Erik Tomáš.

Cieľom nového národného projektu Zručnosti pre trh práce je podporiť uchádzačov o zamestnanie v snahе nájsť si prácu prostredníctvom rekvalifikácie. Projekt bude pôsobiť aj preventívne. Možnosť získať bezplatné vzdelávanie ponúkne aj zamestnaným záujemcom o zamestnanie, a to s cieľom ochrániť a zlepšiť ich postavenie na trhu práce. Projekt nie je určený pre samostatne zárobkovo činné osoby.

„Tento projekt sme nastavovali citlivо, na základe potrieb ľudí, ale aj údajov o trhu práce tak, aby bola pomoc čo najúčinnejšia. Väčší priestor a viac podpory sme v rámci neho dali ľuďom, ktorí čelia komplikovanejšej životnej situácii a lepšie vzdelanie u nich má väčší potenciál podstatne im zlepšiť kvalitu života. Aj preto sme kládli dôraz na znevýhodnených uchádzačov o zamestnanie a mladých ľudí do 30 rokov bez práce,“ dodal generálny riaditeľ a generálny tajomník služobného úradu Ústredia práce, sociálnych vecí a rodiny Peter Ormandy.

Žiadatelia si sami vyberú vzdelávanie a aj jeho poskytovateľa. Úrady práce, sociálnych vecí a rodiny budú akceptovať iba poskytovateľov pôsobiacich v oblasti vzdelávania najmenej jeden rok. O príspevok na vzdelávanie bude možné žiadať od 8.apríla na príslušnom úrade práce, sociálnych vecí a rodiny, ktorý každú žiadosť individuálne posúdi. Žiadosť o poskytnutie príspevku je potrebné podať najneskôr 30 kalendárnych dní pred začiatkom kurzu. Pri scénariu, že žiadost o poskytnutie príspevku je schvázená, poskytovateľ sa vzdelenie zaplatí vzdelávací kurz. V rámci tohto projektu sa poskytne aj finančná podpora pre ľudí bez práce, ktorí sú v súčasnosti vyučovaní v rámci iných projektov.

repas@itacademy.sk

Ako sa s nami Spojit'?



ADRESA: IT Academy, s. r. o.

Budova KOLOSEO prízemie
Tomášikova 50/A
831 04 Bratislava



WEB: www.it-academy.sk



E-MAIL: info@it-academy.sk



TELEFÓN: +421 917 095 406



Ako Vieme Pomôcť?

#Copywriting

#Školenia

#Zamestnanci

#Pomáhame

#Rast

#Projekty

#Certifikácie

#Kurzy

#Tréningy

#Vzdelávanie

#PPC Kampane

#Elearning

#Mentoring

#Konzultácie

#Online

#Programovanie

#Vývoj

#Marketing

#Reklama

#Prenájom Techniky

- [Domov](#)
- [Shorts](#)
- [Odbory](#)
- [Moje](#)
- [História](#)

Ak chcete k videám pridať označenie páči sa mi, komentovať alebo sa prihlásiť na odber, musíte sa prihlásiť.

[Prihlásiť sa](#)

Preskúmať

- [Trendy](#)
- [Hudba](#)
- [Filmy](#)
- [Hry](#)
- [Šport](#)
- [Prehliadat kanály](#)

Viac zo služby YouTube

- [YouTube Premium](#)
- [YouTube Music](#)
- [YouTube Kids](#)

Naštartuj sa s nami v IT!



PREZENČNÉ KURZY
www.it-academy.sk



ONLINE KURZY
www.vita.sk



Daj Odber

IT Academy


IT ACADEMY

@IT-Academy · 7,02 tis. odberateľov · 938 videí

[Online certifikované kurzy a školenia IT, marketingu a manažmentu.](#) >

[vita.sk a 8 ďalších odkazov](#)
[Odoberať](#)
[Domov](#) [Videá](#) [Shorts](#) [Naživo](#) [Zoznamy](#) [Komunita](#) [Hľadať](#)

Videá

[Prehrať všetko](#)

[Online kurz Data Science a Jazyk R - Jemný Úvod do...](#)

7 zhliadnutí · pred 1 hodinou


[Online kurz Python - Ako programuje Google v Python?](#)

51 zhliadnutí · pred 1 dňom


[Online Kurz Microsoft Excel Grafy a Typy Grafov](#)

61 zhliadnutí · pred 2 dňami


[Online Kurz Počítačové Siete - Protokol HTTP\(s\), Stavové...](#)

166 zhliadnutí · pred 3 dňami


[Online Kurz Efektívna Komunikácia - Názory a...](#)

60 zhliadnutí · pred 3 dňami


[Online kurz Microsoft Outlook - Ako na...](#)

142 zhliadnutí · pred 5 dňami

Oblíbené videá

[Prehrať všetko](#)

[SQL I. Začiatčník](#)

1:51:23


[Microsoft Excel - Kontingenčné Tabuľky](#)

1:51:58


[Siete I. Základy Siete](#)

1:40:33


[AKO SA STAŤ TESTEROM](#)

2:45:09


[Microsoft Access I. Začiatčník](#)

1:46:27


[Ako efektívne pracovať v Microsoft Excel?](#)

2:07:28

Dajte odber na IT Academy



www.youtube.com/c/IT-ACADEMYSK

