



PRAHA **CODING** SCHOOL

**Příručka pro instalaci nástrojů ke kurzu**

**DATOVÁ ANALYTIKA**

# Obsah

Úvod .....	3
Instalace nástrojů.....	4
Typ operačního systému .....	4
Windows.....	4
MacOS .....	5
Instalace Python.....	5
Instalace Git .....	6
Instalace Visual Studio Code.....	7
Nástroj pro vytváření `pdf` .....	9
Vytvoření on-line účtů .....	10
GitHub účet a repository.....	10
Databáze SQL.....	13
Lokální inicializace projektů .....	14
Volitelné /využijeme dle potřeby/ .....	15
Grafana .....	15
Tableau Public.....	15
Instalace Tableau Public .....	16

# Úvod

V tomto dokumentu najdete návod pro instalaci nástrojů využívaných v kurzu **Datová analytika – Python, Big Data, AI a ML**.

Během kurzu budeme pracovat v programátorském prostředí **Visual Studio Code** v programovacím jazyku **Python**. Jako zdroj dat využijeme lokální soubory, SQL databázi i datové sady „na internetu“. Pro sdílení pracovních souborů a vlastní projekt budeme využívat **Github**.

Průběh instalace jednotlivých balíčků se může lišit podle verze operačního systému, přizpůsobte tedy jednotlivé kroky v tomto návodu své konkrétní situaci.

Instalaci a funkčnost jednotlivých nástrojů prověříme a doladíme společně na první lekci a dále v průběhu kurzu.

# Instalace nástrojů

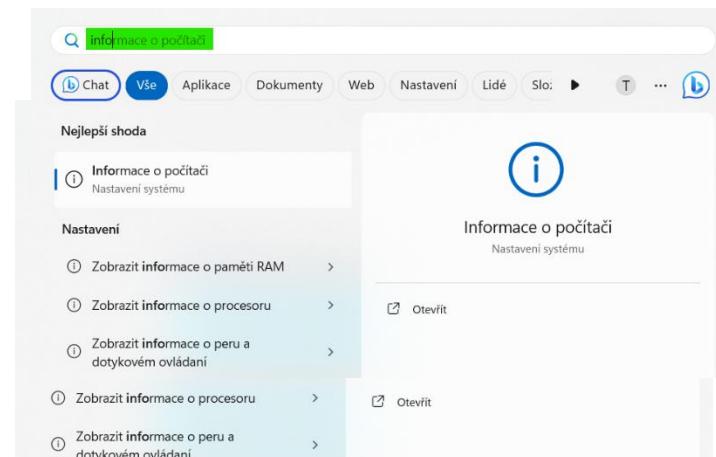
Před tím, než začneme práci v kurzu a na vlastním projektu, je nutné projít několika kroky pro nastavení prostředí a nástrojů v počítači.

## Typ operačního systému

Při instalaci jednotlivých nástrojů je potřeba znát typ operačního systému **Windows** či **MacOS**, který používáte. Pokud jej neznáte, můžete si jej nejprve zjistit následujícím způsobem:

### Windows

1. Klikněte na tlačítko Start nebo stiskněte klávesu Windows na klávesnici ☰
2. Když se otevře postranní panel, napište **Informace o počítači** a vyberte navrhovanou možnost **Nastavení systému**.



3. V novém okně v části Specifikace zařízení (O aplikaci) → Typ systému (Specifikace Windows) najdete typ systému vašeho počítače (64 nebo 32 bit).



## MacOS

1. Klikněte na logo **Apple** v levém horním rohu vaší obrazovky a zvolte možnost **About this Mac**.
2. V následujícím okně **Overview** hledejte položku **Processor** nebo **Chip**. Jestliže je uveden **Intel** chip, zvolte při stahování možnost **x64** nebo **Intel**, pokud **M1** či **M2**, zvolte možnost **ARM** nebo **Apple Silicon**.



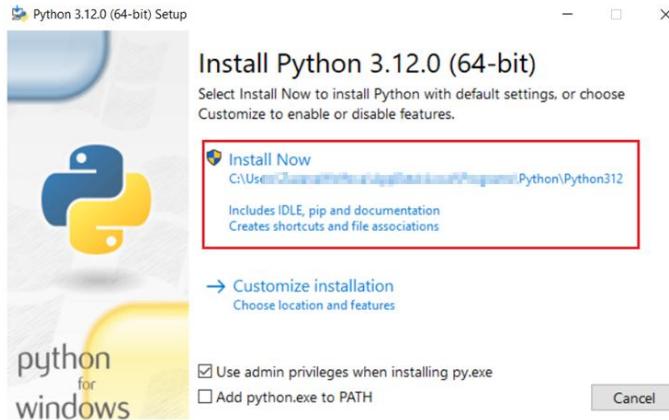
## Instalace Python

Instalační balíček si stáhněte zde: <https://www.python.org/downloads/>

1. Klikněte na možnost **Download Python 3.xx.x**:



2. Dvojklikem myší na stažený instalaci soubor zahájíte instalaci. Objeví se následující obrazovka, kde kliknete na možnost **Install Now**: Druhá možnost, pro pokročilé, **Customize installation** umožní výběr cílového adresáře a dalších volitelných položek.

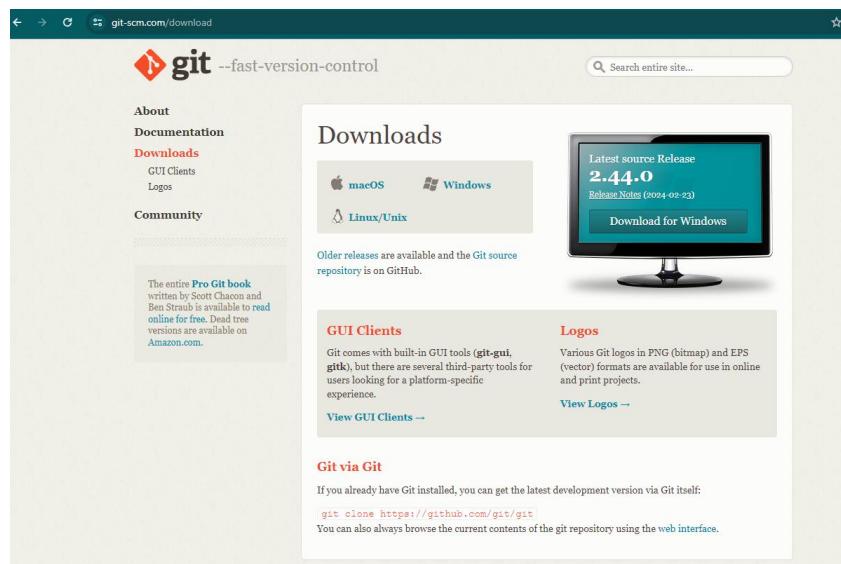


3. Po dokončení instalace stačí zavřít okno kliknutím na tlačítko **Close**

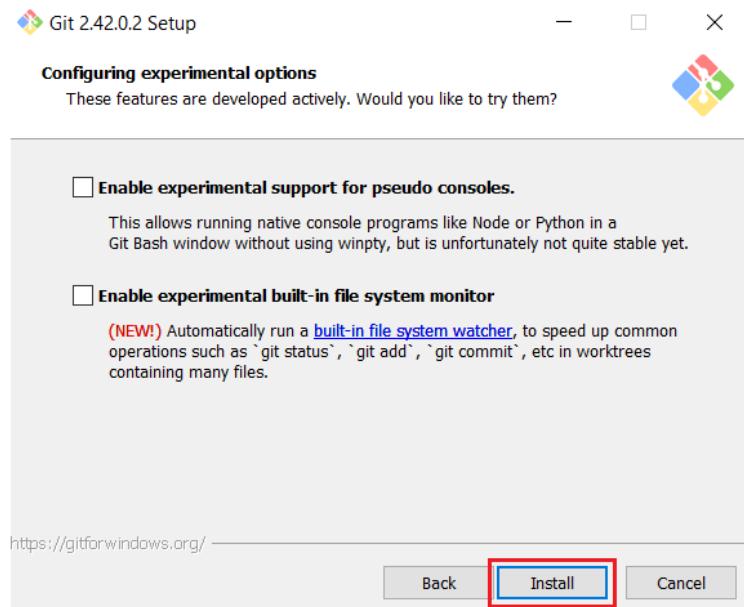
## Instalace Git

Instalační balíček si stáhněte zde: <https://git-scm.com/download>

1. Zvolte možnost na základě verze Vašeho operačního systému:



2. Instalaci zahájíte dvojitým poklepáním na stažený soubor. Setup je velice jednoduchý, neprobíhají zde žádné změny základního nastavení – v podstatě jen stále klikáte **Next**, než dorazíte k samotné instalaci. Klikněte na tlačítko **Install**:



3. Po dokončení instalace stačí zavřít okno kliknutím na tlačítko Finish

## Instalace Visual Studio Code

Instalační balíček si stáhněte zde: <https://code.visualstudio.com/download>

1. Klikněte na tlačítko (**modrý obdélník**) odpovídající vašemu operačnímu systému:

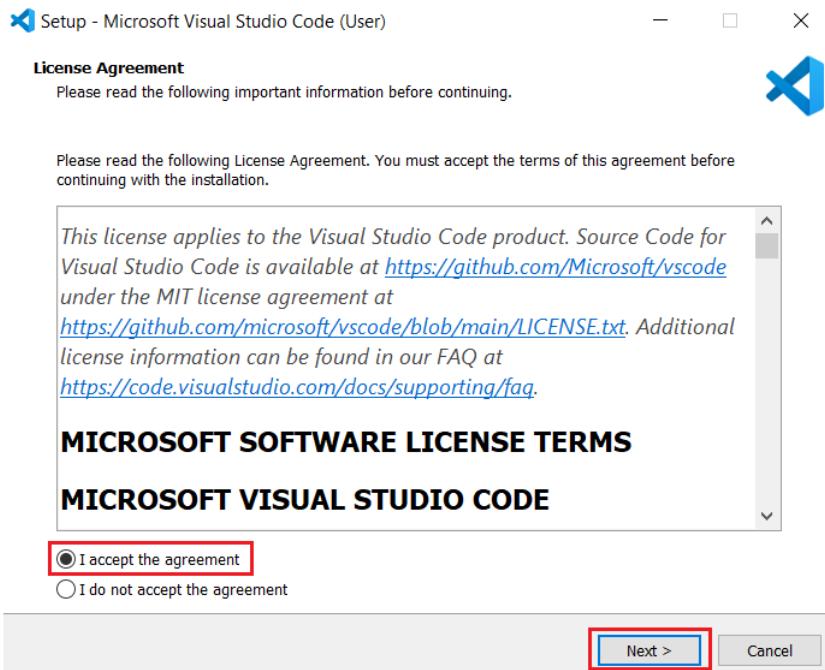
A screenshot of the Visual Studio Code download page. At the top, there is a navigation bar with links for Docs, Updates, Blog, API, Extensions, FAQ, Learn, and a search bar. A blue 'Download' button is visible. Below the navigation bar, a message says 'Version 1.87 is now available! Read about the new features and fixes from February.' The main section is titled 'Download Visual Studio Code' and describes it as 'Free and built on open source. Integrated Git, debugging and extensions.' It shows icons for Windows, Linux (Tux), and macOS. Below each icon is a blue button with a download arrow and the operating system name. Under each button, there is a table of download links for different architectures (x64, Arm64, etc.).

	x64	Arm64
User Installer	x64	Arm64
System Installer	x64	Arm64
.zip	x64	Arm64
CLI	x64	Arm64

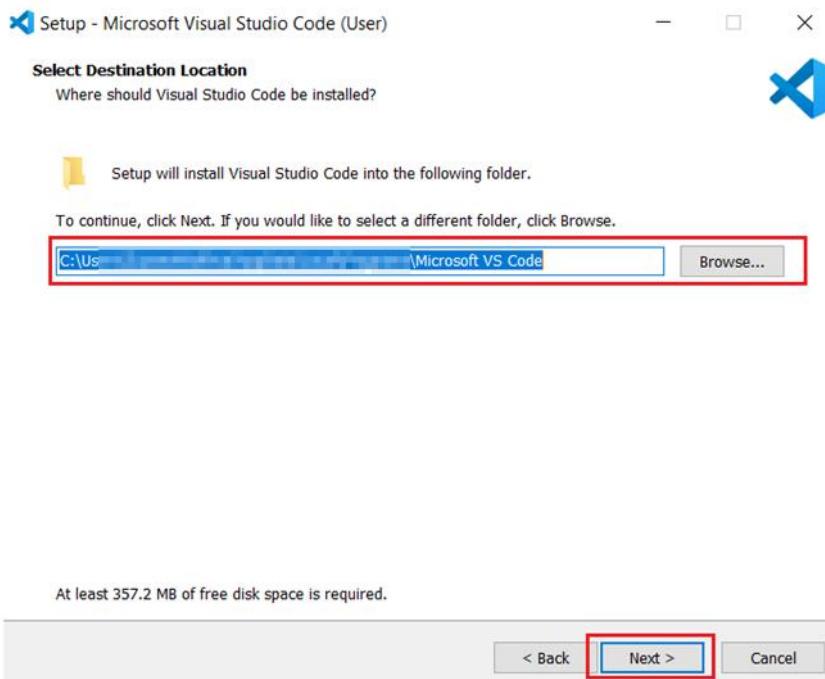
	x64	Arm32	Arm64
.deb	x64	Arm32	Arm64
.rpm	x64	Arm32	Arm64
.tar.gz	x64	Arm32	Arm64
Snap	Snap Store		

	x64	Arm32	Arm64
.zip	Intel chip	Apple silicon	Universal
CLI	x64	Arm32	Arm64

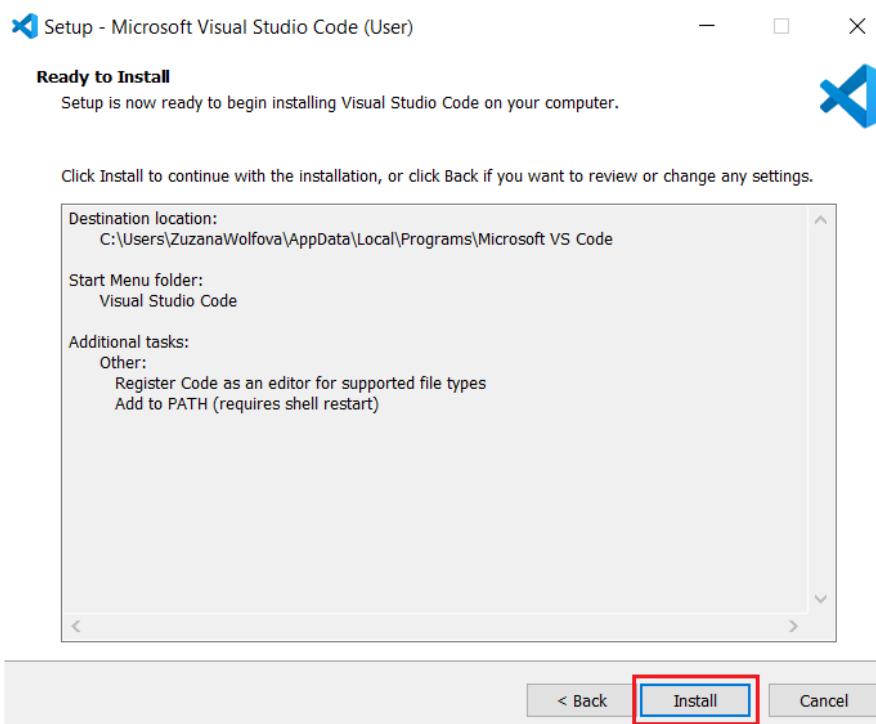
2. Uložte si nabídnutý instalacní balíček na místo ve Vašem počítači dle Vašeho přání. Dvojitým poklepáním myši na daný soubor spusťte instalaci. Zvolte možnost **I accept the agreement** a klikněte na tlačítko **Next**:



3. Zvolte si **místo instalace** a klikněte na **Next**:



4. Instalační procedura je opět velice jednoduchá. Na následujících obrazovkách vždy pouze odklikněte pomocí tlačítka **Next**. Nakonec se dostanete k samotné instalaci:



5. Nakonec potvrďte tlačítkem **Finish** a máte instalaci hotovou.

## Nástroj pro vytváření `pdf`

Pro generování `pdf` souborů programem v Python je nutné nainstalovat nástroj `wkhtmltopdf` z download stránky <https://wkhtmltopdf.org/downloads.html>

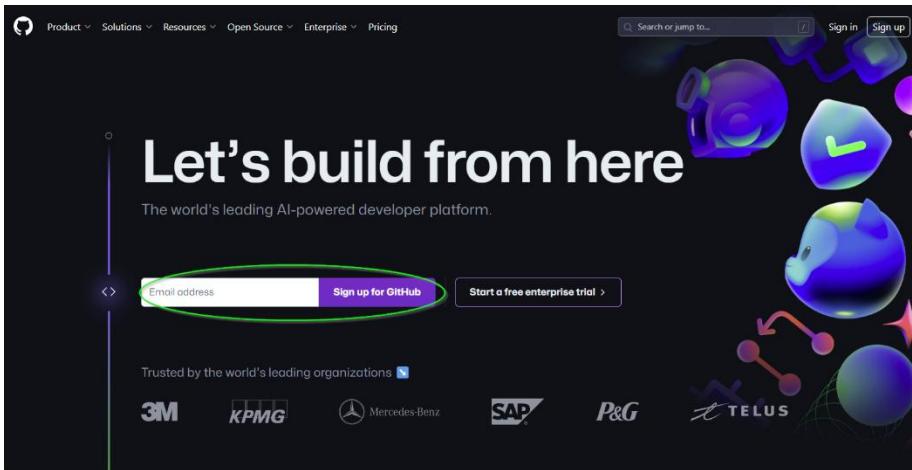
OS/Distribution	Supported on	Architectures			
Windows	Installer (Vista or later)	64-bit	32-bit		
	7z Archive (XP/2003 or later)	64-bit	32-bit		
macOS	Installer (10.7 or later)	64-bit			
	11 (bullseye)	amd64	i386	arm64	ppc64el
	10 (buster)	amd64	i386	arm64	ppc64el
Ubuntu	9 (stretch)	amd64	i386	arm64	raspberrypi
	22.04 (jammy)	amd64	arm64		ppc64el
	20.04 (focal)	amd64	arm64		ppc64el
	18.04 (bionic)	amd64	i386	arm64	ppc64el
	16.04 (xenial)	amd64	i386	arm64	

# Vytvoření on-line účtu

## GitHub účet a repository

V kurzu budeme využívat GitHub pro uložení zdrojových kódů, sledování změn a také předávání dalších materiálů ke kurzu.

1. Pokud nemáte účet na [github.com](https://github.com), je třeba si ho vytvořit...



2. Pro kurz si vytvořte nové (privátní) repository

Two screenshots of the GitHub interface. The top screenshot shows the user's repositories page for "PCS-Training". On the left, there is a sidebar with filters for All, Public, Private, Sources, Forks, Archived, and Templates. The main area shows a list of repositories, with one repository named "DAPYTPCS20240819" highlighted. A green oval highlights the "New repository" button in the top right corner of the header. The bottom screenshot shows the "Create a new repository" dialog. It includes fields for "Repository name" (set to "DAPYTPCS20240819"), "Owner" (set to "DAPY Training"), and "Description (optional)". There are also sections for selecting the repository's visibility ("Public" or "Private") and initializing it with a README file. A green oval highlights the "Repository name" field.

Poznámka: repo musí být prázdné, tj. bez `README` případně jiných souborů

### 3. Pro další práci se zdrojovými kódy v repository je dobré vytvořit si access token

The screenshot shows the GitHub developer settings page under 'Personal access tokens'. A specific token named 'DAPYTPCS20240819' is highlighted with a green circle. The token details show it was created on 2024-08-19, has no expiration date, and has 'Read & write' permissions for the repository 'DAPYTPCS20240819'. There is also a 'Generate new token' button.

### 4. V dalším kroku se definují parametry a práva tokenu ... zvolte maximální dobu platnosti 1 rok a read-write přístup k repository pro kurz

The screenshot shows the 'New fine-grained personal access token' form. The token name is 'DAPYTPCS20240819'. The expiration date is set to 'Custom...' with '29/06/2025'. The description is 'Access token for project DAPYTPCS20240819'. Under 'Resource owner', there is a dropdown for 'Organization' with 'DAPYTPCS' selected. In the 'Repository access' section, the 'Only select repositories' option is selected, and a single repository 'DAPYTPCS20240819' is chosen. In the 'Permissions' section, several permissions are selected: Actions, Administration, Attestations, Code scanning alerts, Codepaces, Codepaces lifecycle admin, Codepaces metadata, Codepaces secrets, Commit activities, Contents, and Custom properties. The 'Commit activities' and 'Contents' sections are circled in green.

5. Po stisknutí tlačítka 'Generate token' je důležité na další obrazovce **zkopírovat** text tokenu a **bezpečně** si ho **uložit**

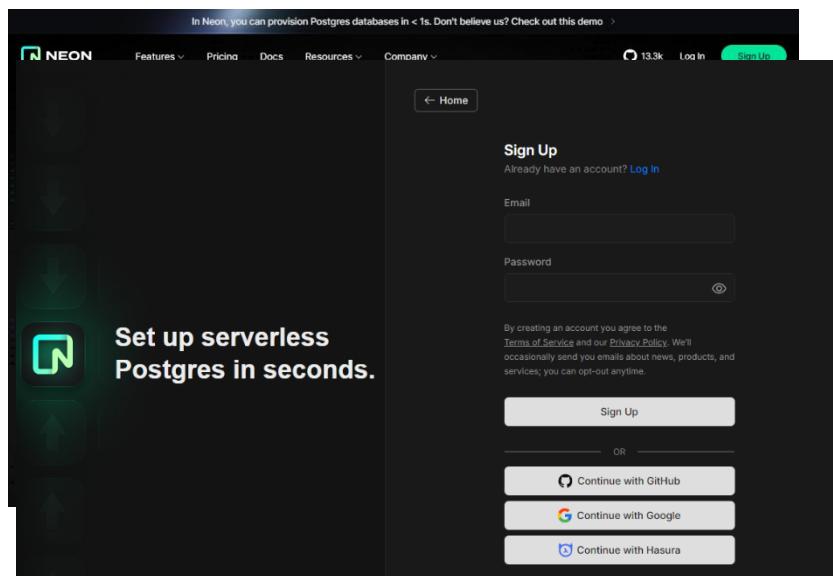
The screenshot shows the GitHub Developer Settings page for generating a personal access token. At the top, it says 'Selected 1 repository: PCS-Training/DAPYTPCS20240819'. Below that is a 'Permissions' section with 'Repository permissions' selected. Further down are sections for 'Organization permissions' and 'Overview' showing 3 permissions for repositories and 0 organization permissions. A note states the token will expire on Sun, Jun 29 2025. At the bottom, there is a green button labeled 'Generate token' with a circled 'Generate' button, and a note that the token will be ready for use immediately.

Below this, another screenshot shows the 'Personal access tokens' list. It lists several tokens, including one recently generated for the PCS-Training repository. The token itself is highlighted with a green circle around the first few characters: 'github\_pat\_13466617f...'. Other tokens listed include 'GitHub API', 'DAPYTPCS20240819', 'DAPYTPCS20240819', 'DAPYTPCS20240819', 'DAPYTPCS20240819', 'DAPYTPCS20240819', 'DAPYTPCS20240819', 'DAPYTPCS20240819', and 'PetrSeb'.

6. Token se používá v repo access URL takto:  
`<https://<token>@github.com/PCS-Training/DAPYTPCS20240819.git>`

# Databáze SQL

Budeme pracovat s on-line databází na SQL serveru NeonDB <https://neon.tech/>, kde si vytvořte účet (pokud ještě nemáte)



1. Po stisknutí tlačítka `Sign-Up` se se otevře stránka pro vytvoření nového účtu pomocí emailovou adresou nebo existujícího účtu Google.
2. v dalším kroku je nutné vytvořit si projekt, ve kterém budou spravovány databáze. Po úspěšném vytvoření se otevře `Project Dashboard`
3. teď je třeba vytvořit si databázi

A screenshot of the Project Dashboard. On the left sidebar, there are sections for PROJECT (Dashboard, Branches, SQL Editor, Restore, Monitoring, Integrations, Settings, Quickstart) and BRANCH (Overview, Tables). In the center, there's a 'Project Dashboard' card with 'Connection Details' and 'Branches' sections. A modal window titled 'Database creation' is open, showing a 'Create' button (labeled 1), a 'Database name' field containing 'DAPYTPCS' (labeled 2), and an 'Owner' dropdown. To the right of the modal, there's a summary of resources: Compute (0.25 CU, ACTIVE), Storage (457.66 MB), and Compute time since July 1 (5 h). A note says: 'any preview branches yet. Improve by adding database branching to your reviews.'

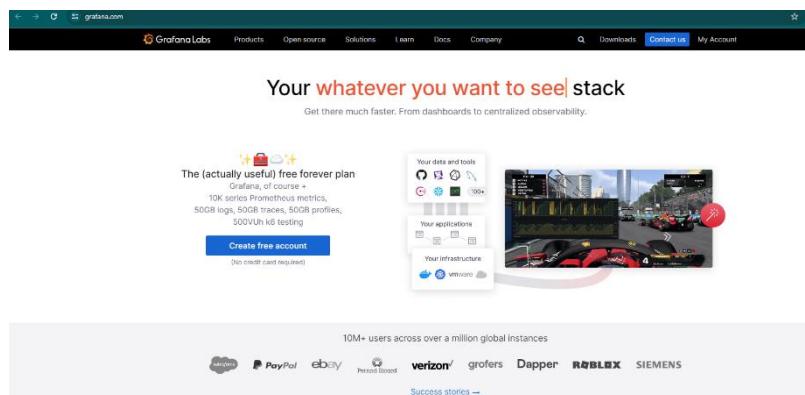
# Lokální inicializace projektů

Pro vytvoření prostředí projektů využijeme `template` z GitHub. To si projdeme společně na začátku kurzu.

# Volitelné /využijeme dle potřeby/

## Grafana

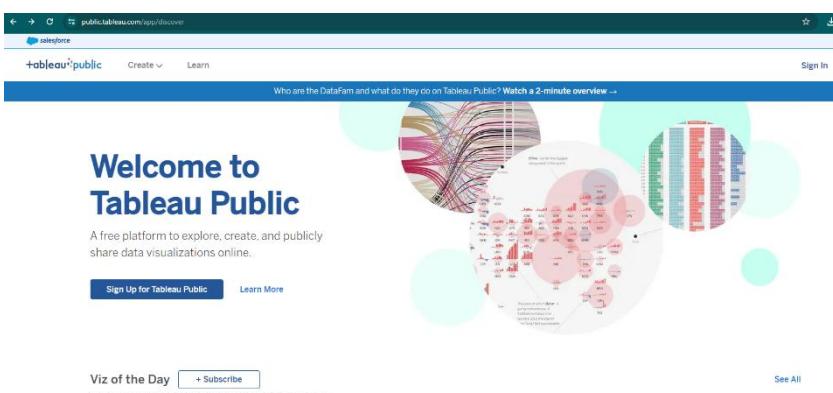
Nástroj pro grafickou prezentaci dat je vhodný zejména pro interaktivní zobrazení dat v reálném čase. Free on-line verze je dostupná na [grafana.com](https://grafana.com). Pro vytvoření účtu je možné použít některý z populárních cloud účtů nebo použitím jakékoli emailové adresy



Po úspěšné aktivaci účtu není třeba nic dalšího řešit, vše si projdeme společně

## Tableau Public

Enterprise nástroj pro analýzu a grafickou prezentaci data. Ve free verzi je vhodný pro práci se statickými datovými sadami. Desktop aplikaci je možné propojit s free public serverem, na kterém je možné vytvořit účet na [public.tableau.com](https://public.tableau.com).

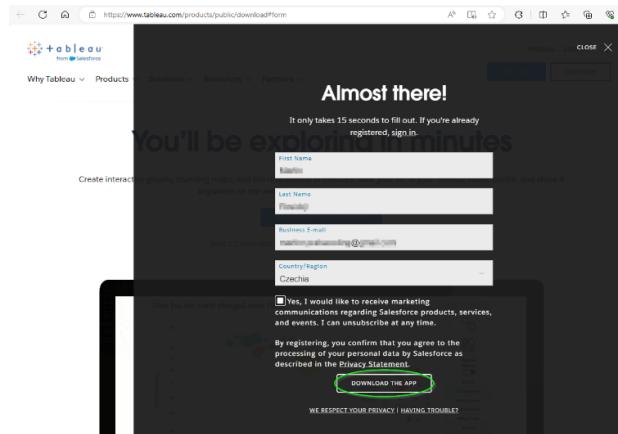


Na tomto serveru najdete nepřeberné množství inspirace od ostatních datových analytiků

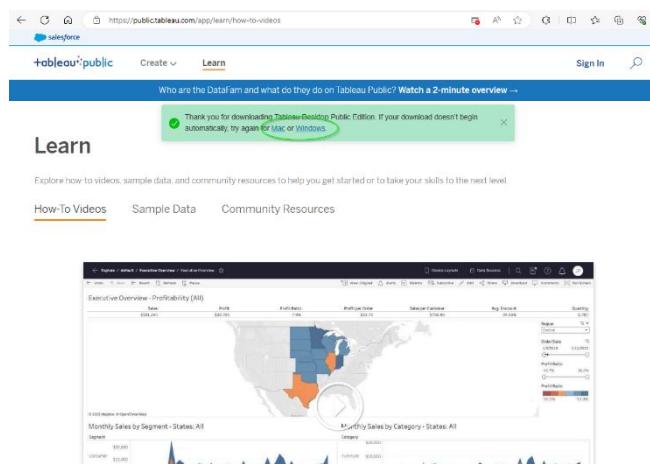
# Instalace Tableau Public

Stažení instalačního balíčku vyžaduje několik kroků

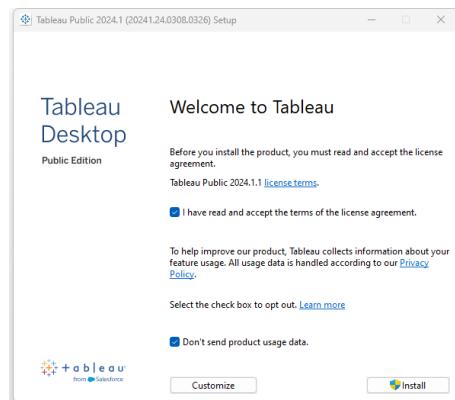
1. Odkaz pro stažení je zde: <https://www.tableau.com/products/public/download#form>



2. Ve formuláři je třeba vyplnit požadované údaje a kliknout na **Download App**
3. Pokud se stahování nespustí automaticky je třeba vybrat operační systém



4. Dvojklikem na stažený soubor se spustí vlastní instalace. Je nutné potvrdit licenci a pokračovat tlačítkem **Install**



5. Po dokončení instalace je na ploše vytvoří ikona **Tableau Public** a program se spustí