

<image/nation>
Future-proof Tech Skills

Specially designed for



HKU
BUSINESS
SCHOOL
港大經管學院

Masters
Programmes

2026/2027
Vibe Coding for Data Analysts

by

Sunny NG

sunny.ng@imagenation.com.hk

Sunny NG

Master of Fine Art City University of Hong Kong, **Master of Science** University of Hong Kong, **Bachelor of Science** University of Hertfordshire (UK)

training.imagenation.com.hk

www.linkedin.com/in/ngsunny/

<https://github.com/ngsanluk>

Founder & Master Trainer Image Nation 隨想國度

26-Year IT & New Media Experience

AI, Cloud, Blockchain, Web, Mobile, Database & IoT Applications Development

19-Year University Teaching & Tech Training

HKU / HKUSPACE / EdUHK / CityUHK

Certifications & Professional Qualifications

Microsoft Certified Azure Developer, AI Engineer & Github Copilot

AWS Certified Solution Architect, Developer, Data Engineer, Machine Learning Engineer & AI Practitioner

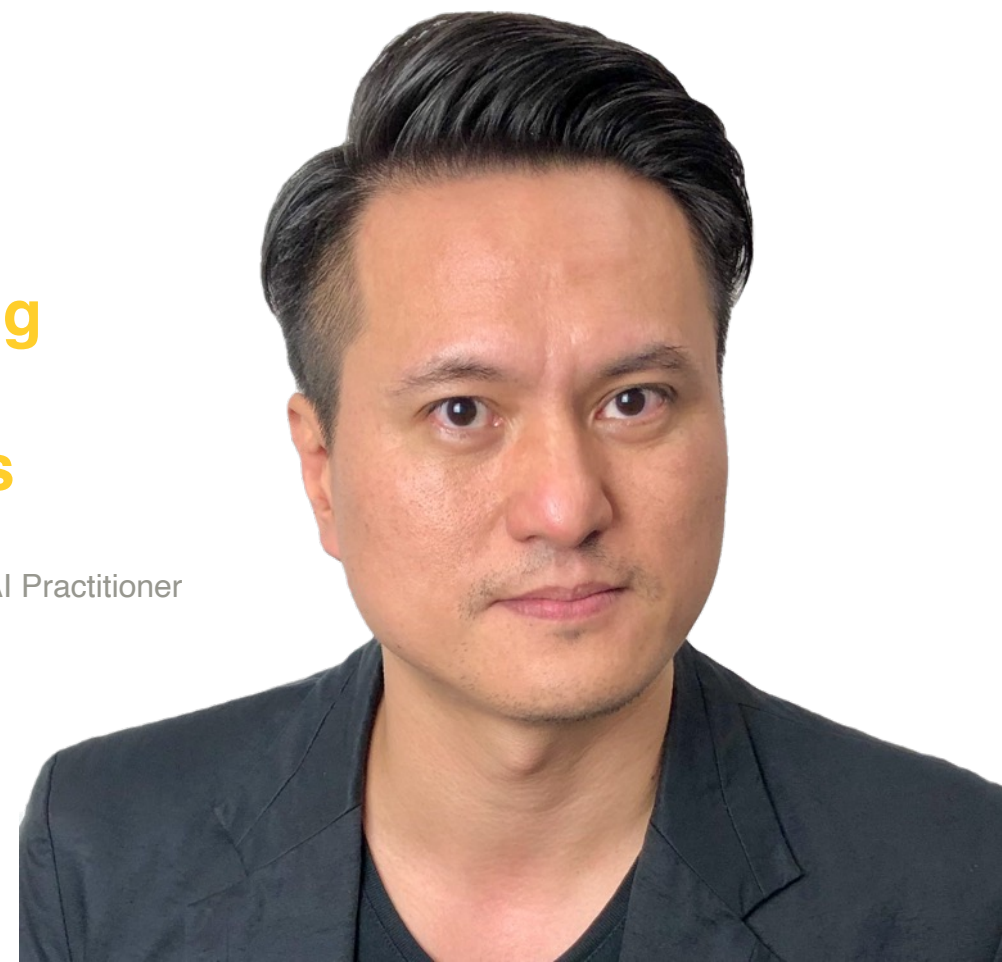
Alibaba Cloud Certified Professional in Cloud Computing

Accredited Educator for AWS Academy

Accredited Educator for Microsoft Learn for Educator (MSLE)

Accredited Educator for Oracle Academy

Accredited Educator for Cisco Networking Academy



Setup Check List

- Make sure you have a Gmail account. If you don't, create on
 - <https://gmail.com>
 - <https://gemini.google.com>
 - <https://colab.research.google.com/>
- Anaconda (Python Dev Environment)
 - <https://www.anaconda.com/download/success?reg=skipped>
- Visual Studio Code
 - <https://code.visualstudio.com/download>
- Create Github Account & Activate Github Copilot
 - <https://github.com/signup>
- Apply **FREE** Github Pro for Student
 - <https://github.com/education/students>

In this workshop (3 hours)

- Understanding Data Science Workflow
- Data Science Paradigm Shift: Where Does Vibe Coding Fit?
- Landscape of Vibe-coding Tools
- Quick and Casual EDA (Exploratory Data Analysis)
- Interactive Dashboard Generation
- Setting Up AI-Native Environment & Workflow
 - Performing **Data Import / Tidy / Transform / Visualization** by Natural Language

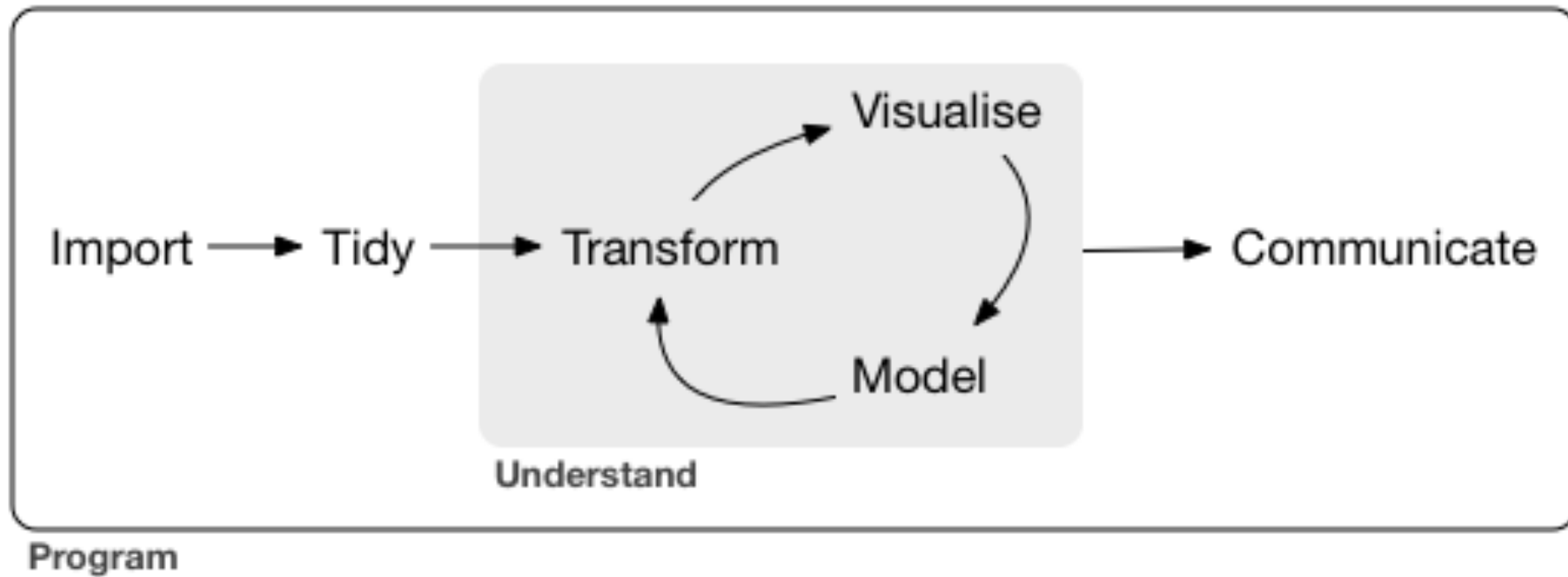
To download the slides

tinyurl.com/hkin-dl

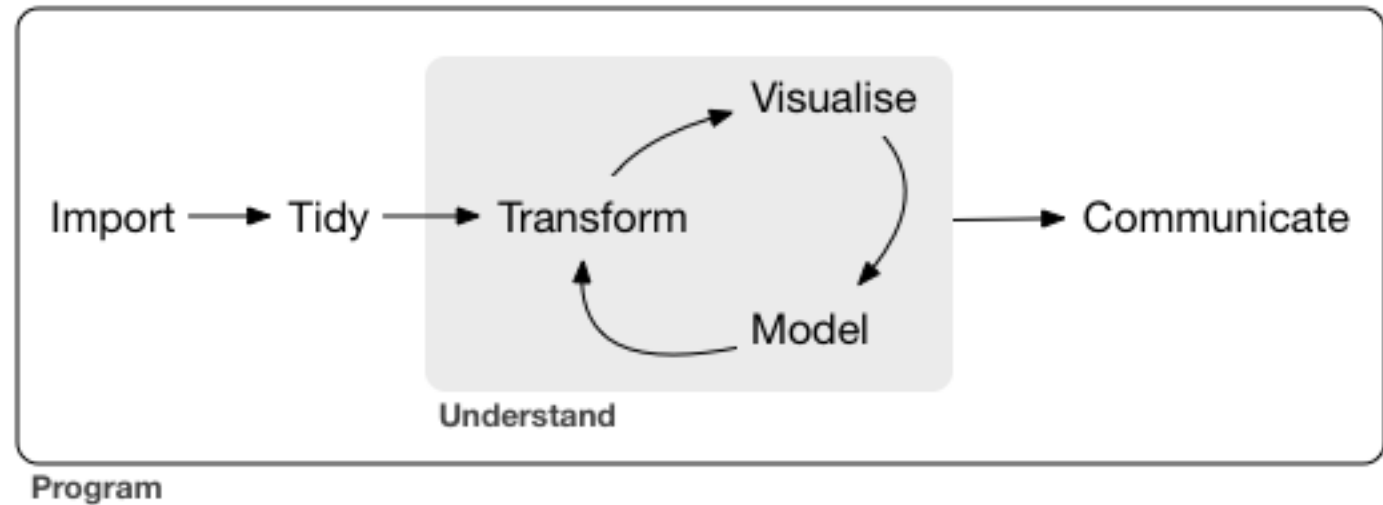


Data Science Workflow

Data Science Workflow



Data Science Workflow



- **Import:** Bring data into your programming environment.
- **Tidy:** Structure data into a consistent format.
- **Transform:** Filter and manipulate data of interest.
- **Visualize:** Display data graphically to find patterns.
- **Model:** Create statistical or machine learning models.
- **Communicate:** Share results and insights with others.
- **Program:** Write code to power the entire workflow.

Traditional Programming Barriers

- Syntax memorisation
- Too many programming languages
 - Python / R / JavaScript / SQL
- Complex library dependency
 - pandas, matplotlib, seaborn etc.
- Complicated application stack
 - Anaconda

Data Analysis Paradigm Shift

- Powered by AI, the data analysis paradigm has shifted from **Syntax Memorization** to **Intent Architecture**
- NEW roles for business analyst in the AI era
 - Specification Writer
 - Domain Evaluator
 - Quality Auditor

Tools Landscape of Vibe Coding

Popular AI Coding Tools

 Claude  Gemini  Copilot  GitHub Copilot

 CURSOR

 Lovable

bolt.new



 Base 44



Google
colab

 Google Antigravity

Top GenAI Tools

<https://a16z.com/100-gen-ai-apps-6/>

The Top 50 Gen AI Web Products, by Unique Monthly Visits				
1. ChatGPT	11. FREEPIK	21. Lovable	31. Arena	41. CURSOR
2. Gemini	12. Doubao	22. PolyBuzz	32. SEARAT.AI	42. CIVITAI
3. Canva	13. JanitorAI	23. ourdream.ai	33. Hugging Face	43. Midjourney
4. deepseek	14. Quark	24. Kimi	34. Crushon AI	44. manus
5. Grok	15. SUNO	25. Google Labs	35. Meta AI	45. KlingAI
6. Claude	16. remove.bg	26. Qwen	36. candy.ai	46. VEED
7. character.ai	17. CapCut	27. TurboScribe	37. Photoroom	47. Genspark
8. perplexity	18. grammarly	28. GAMMA	38. Pixelcut	48. SIGA CHAT
9. Notion	19. SPICYCHAT.AI	29. ElevenLabs	39. Adot	49. Poe
10. Google AI Studio	20. QuillBot	30. NotebookLM	40. Higgsfield	50. cutout.pro

Source: Similarweb, January 2026
Charts are for informational purposes only. Past performance is not indicative of future results.
None of the above should be taken as investment advice; see a16z.com/disclosures.



The Top 50 Gen AI Mobile Apps, by Monthly Active Users				
1. ChatGPT	11. Remini	21. Meta AI	31. Facemoji	41. Translate
2. CapCut	12. QQ浏览器	22. meitu	32. character.ai	42. BeautyCam
3. Gemini	13. deepseek	23. Copilot	33. Microsoft Bing	43. papago
4. Canva	14. Cici	24. Hypic	34. FaceApp	44. BRAINLY
5. AI Gallery	15. perplexity	25. Seekee	35. NOVA	45. PolyBuzz
6. Picsart	16. Adobe Lightroom	26. Notion	36. YouCut	46. AI Mirror
7. Doubao	17. Baidu AI Search	27. photomath	37. Polish	47. MIVI
8. Microsoft Edge	18. Grok	28. Gauth	38. B612	48. SNOW
9. Meituan	19. VN	29. Learna AI	39. Photoroom	49. VivaCut
10. Yandex	20. Edits, an instagram app	30. Wink	40. PixVerse	50. BeautyPlus

Source: Sensor Tower, January 2026
Charts are for informational purposes only. Past performance is not indicative of future results.
None of the above should be taken as investment advice; see a16z.com/disclosures.



Right Tools for Right Use Cases

- **Regular AI Chatbot** e.g.: Gemini, Copilot, Doubao. Good for quick and casual data exploration.
- **AI-Native IDEs** e.g.: Cursor, Claude Code, Github Copilot. Good for complex iterative data processing workflow.
- **Rapid App & Dashboard Generators** e.g.: V0, Lovable, Bolt
- **Enterprise “Vibe Analytics” Platforms** e.g.: Plotly, Base44

Reminder #1

AI products evolve rapidly.
Features and UI change very often.

Reminder #2

AI is NOT always right.
AI makes things up – hallucination.
Use multiple tools for validating.

Quick & Casual

Exploratory Data Analysis

Understand EDA

- **Exploratory Data Analysis**
- EDA is the process of investigating a dataset to understand:
 - What the data looks like (size, columns, data types)
 - Data quality issues (missing values, duplicates, errors)
 - Distributions of variables
 - Relationships between variables
 - Outliers and unusual patterns
 - Potential insights and hypotheses

Tools Good for EDA

- Gemini
- Copilot
- Grok
- Qwen
- Doubao

Common EDA Techniques Include

- Summary statistics (mean, median, standard deviation)
- Histograms and box plots
- Scatter plots
- Correlation matrices and heatmaps
- Data profiling and cleaning

If you have sales data, EDA might reveal

- Which products sell the most
- Seasonal trends
- Customer segments
- Unusual transactions that need investigation

A typical EDA workflow ...

- Load the data
- Inspect structure and data types
- Clean missing or incorrect values
- Generate summary statistics
- Create visualization
- Identify patterns and anomalies
- Draw initial conclusions

Let's Practice EDA

Go download the following dataset

<https://data.gov.hk/en-data/dataset/hk-ugc-ugc-student-graduates2>

EDA using Gemini / Copilot / Qwen / Doubao (10 mins)

- Upload the dataset to AI tools
- Practice with the following prompts
 - Describe the dataset
 - Change "Number of Graduates (Headcount)" column to "Headcount"
 - Retrieve only the "Taught Postgraduate" records
 - How many records are there?
 - Group the data by "Broad Academic Programme Category" and calculate the total headcount.
 - create a pie chart

Selecting the right AI tools

Many AI assistants can perform EDA tasks. But they are NOT equals

- Some AI assistants doesn't allow file upload by free account
- Some AI assistants restricts types of file upload
- Some has very small context window for free account and it will cut-off the excessive data without telling you and therefore your analysis results and visualization results are NOT correct.
- If you have a paid AI tools, priority should go to paid tools.
- Perform the same analysis tasks in multiple AI tools and compare their results for validation.

Mastering

Effective Prompting

AI Beginner User Pitfalls

- Only give simple instruction
- Doesn't provide context
- Doesn't provide clear direction
- Solution
 - AI prompting framework are structured components in a prompt that can direct AI to achieve the right goal

SPEC Prompt Framework

- **S (Situation & Role):** e.g., Senior Retail Strategist, Equity Research Analyst.
- **P (Problem & Objective):** e.g., Identify churn drivers among high-tier subscribers.
- **E (Expectation of Data/Schema):** Column formats, target variables, business metrics.
- **C (Constraints & Output Format):** Visual palette, interactive elements, tabular summaries.

Interactive Dashboard

Interactive Dashboard

- **Centralized Data Visualization:** Displays key metrics, charts, and graphs within a single web-based interface.
- **Dynamic User Control:** Empowers users to filter, sort, and drill down into real-time data on the fly.
- **Actionable Insights:** Simplifies complex information to facilitate quick tracking and informed decision-making.

Tools Good for Dashboard

■ v0 by Vercel 

■ Lovable  Lovable

■ bolt.new 

■ Base44  Base 44

■ Arena  Arena

Building Interactive Data Applications & Scenario Simulators

- Prompting AI to generate an interactive web-based dashboard
 - e.g. Using Gemini Canvas, React Chart components and etc.
- Adding user controls
 - Sliders for dynamic price elasticity simulations
 - Dropdown filters by region/product category
 - Dynamic hover tooltips

Let's Practice Interactive Dashboard Generation

We will use multiple tools and compare them:
v0, Lovable, Bolt, Base44 & Gemini

AI-powered Analytics Pipeline

AI-powered Analytics Pipeline

Import → Tidy & Transform

→ Statistical Computation

→ Visual Storytelling → Deployment

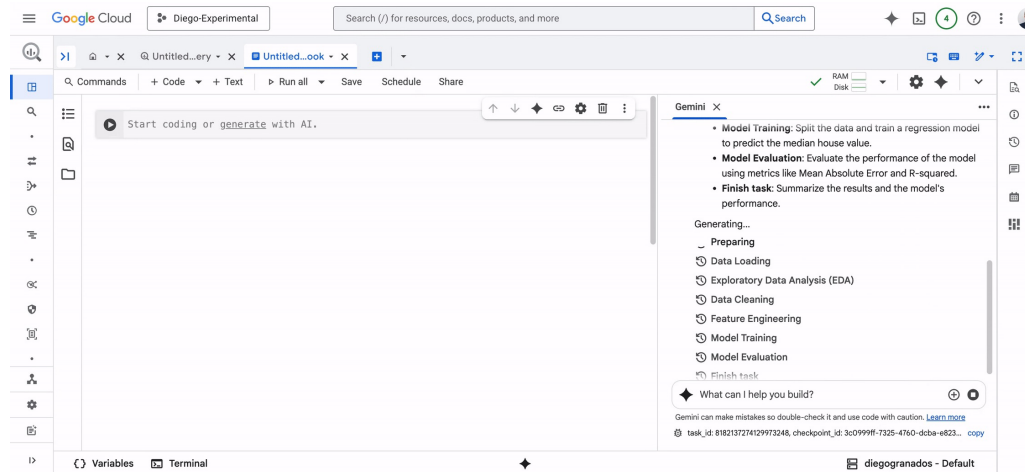
Tools Options

- **VS Code + Github Copilot** (Free student version)
- **Cursor** (Free version comes with limited AI usage)
- **Google Colab** (AI features not available to Hong Kong at the moment)

Google Colab + Gemini

Google Colab

<https://colab.research.google.com/>



- Google Colab is a free cloud-hosted Jupyter notebook environment that lets you write and execute Python code in your browser.
- The AI features in Google Colab is powered by Gemini. AI features are not available to Hong Kong at the moment.

Temporarily Fixing

Before Colab AI features are available to Hong Kong

- Google Colab + Gemini Assistant
- Use Gemini to generate python codes
 - Describe your data analysis task in natural language
 - Let Gemini generate python codes block
- Copy codes from Gemini to Google Colab cell and run

Let's Try (5 mins)

Colab + Gemini

Setting Up AI-Native Environment & Workflow

Setting Up AI-Native Environment & Workflow

1. Install Python Development Environment
2. Sign-up Github / Github Copilot
3. Install Visual Studio Code (VS Code) & Extensions
4. Configure Github Copilot in VS Code
5. Open Project in VS Code
6. Start Vibe Coding

1

Installing Python Environment

Installing Python

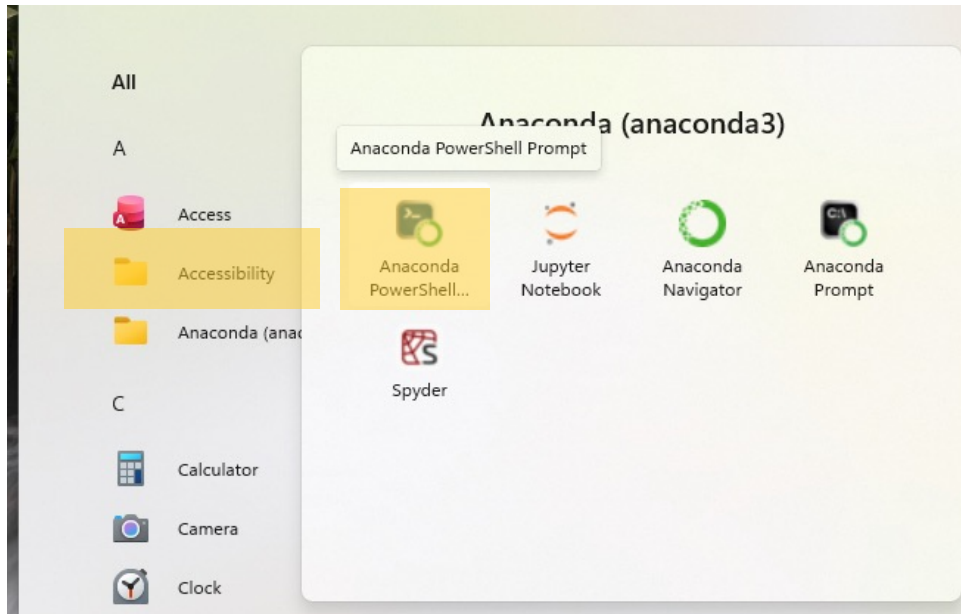
- Anaconda is a popular open-source distribution of Python and R that comes pre-loaded with hundreds of built-in data science and machine learning packages.
- Anaconda Download & Install
 - <https://www.anaconda.com/download/success?reg=skipped>



Test your Python Installation

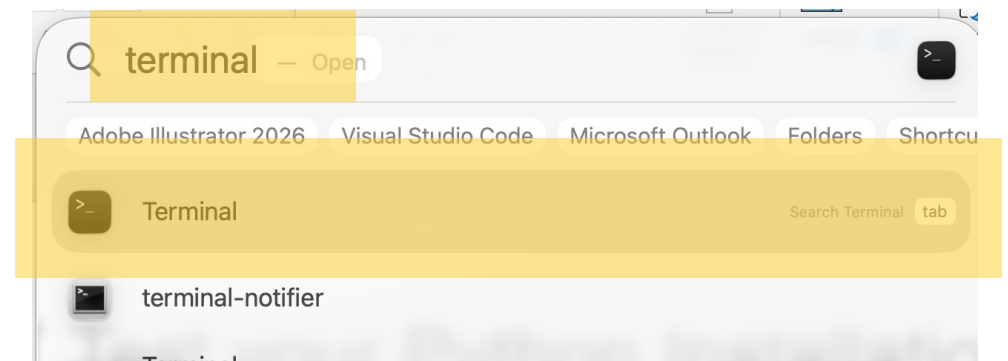
In Windows

Open Anaconda PowerShell



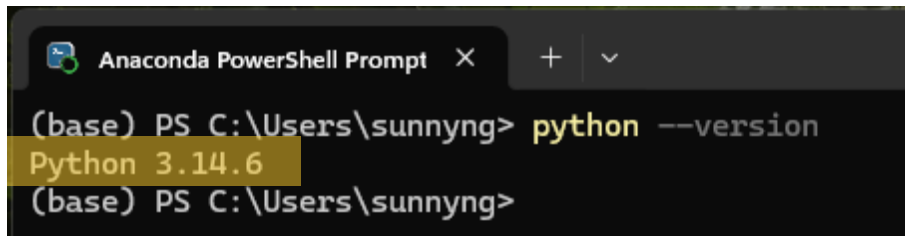
In Mac

Open **Terminal** app



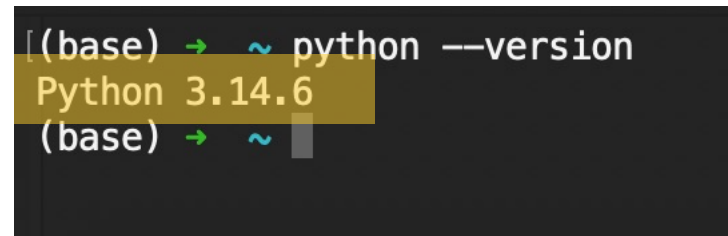
Type command: `python --version`

In Windows



```
Anaconda PowerShell Prompt
(base) PS C:\Users\sunnyng> python --version
Python 3.14.6
(base) PS C:\Users\sunnyng>
```

In Mac

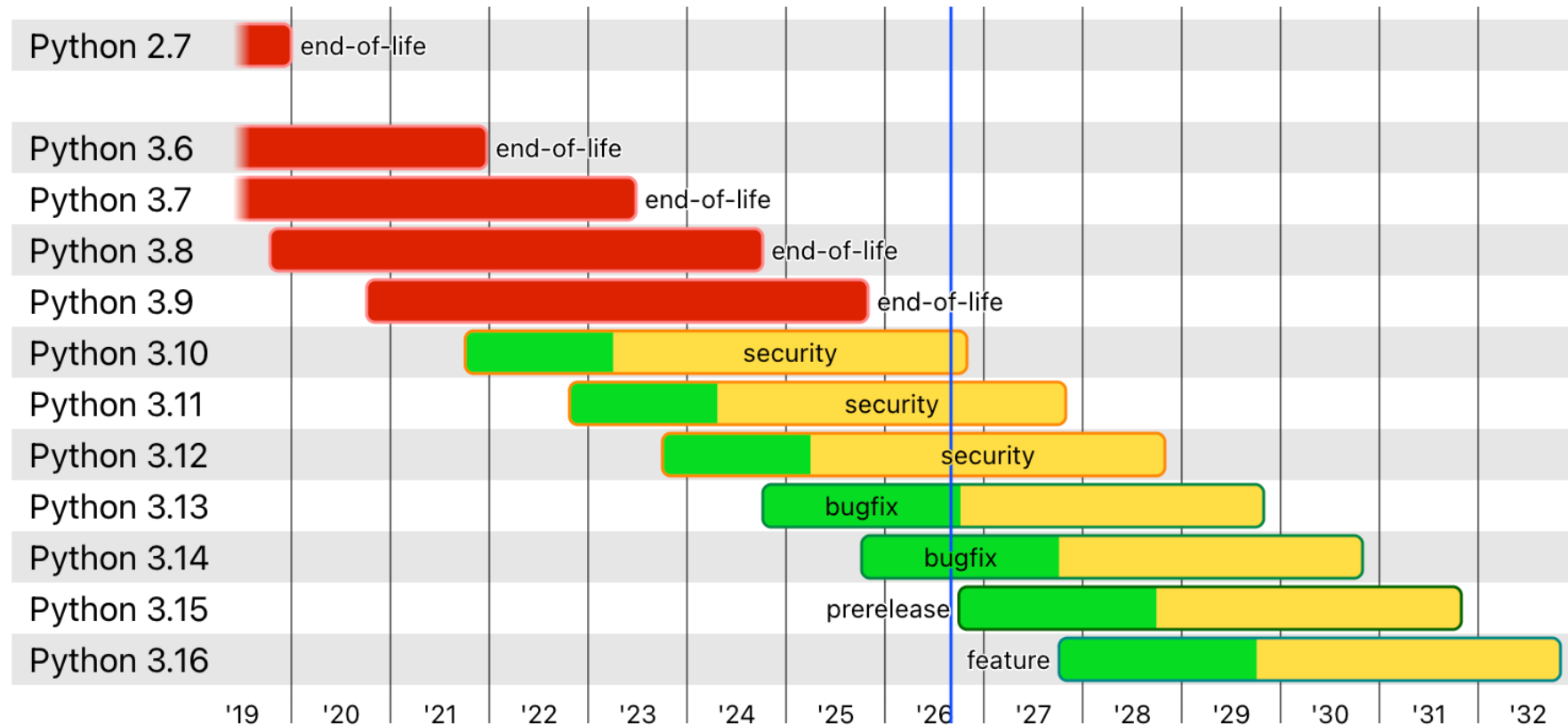


```
(base) → ~ python --version
Python 3.14.6
(base) → ~
```

It's okay if your version number show different as long as it greater than 3.10

Status of Python Versions

It's NOT safe to use versions shows end-of-life



2

Sign-up Github & Github Copilot

Google for “Github”

Or visit <https://github.com/>

Create your free account

Explore GitHub's core features for individuals and organizations.

See what's included ^

- ✓ **Access to GitHub Copilot**
Increase your productivity and accelerate software development.
- ✓ **Unlimited repositories**
Collaborate securely on public and private projects.
- ✓ **Integrated code reviews**
Boost code quality with built-in review tools.
- ✓ **Automated workflows**
Save time with CI/CD integrations and GitHub Actions.
- ✓ **Community support**
Connect with developers worldwide for instant feedback and insights.

Display a menu

Sign up for GitHub

 Continue with Google

 Continue with Apple

or

Email*

Email

Password*

Password

Password should be at least 15 characters OR at least 8 characters including a number and a lowercase letter.

Username*

Username

Username may only contain alphanumeric characters or single hyphens, and cannot begin or end with a hyphen.

Your Country/Region*

Hong Kong

For compliance reasons, we're required to collect country information to send you occasional updates and announcements.

GitHub Copilot

- ☒ Sign up for Copilot Free for coding assistance on the web, in the IDE, and in the CLI. [See terms](#)

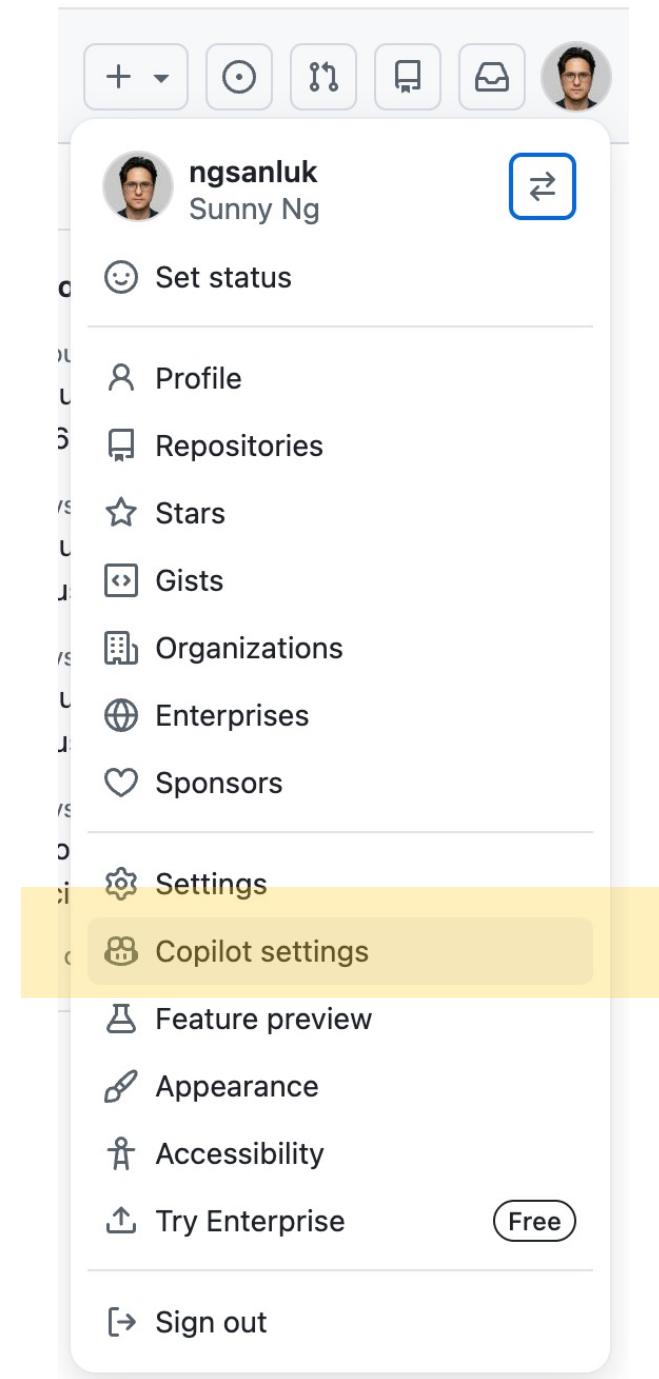
Email preferences

- ☐ Receive occasional product updates and announcements

Create account >

Github Copilot Setting

In Github website, click at your profile icon and reach to “**Copilot settings**”



Copilot Setting on Github Website

GitHub Copilot

Free

Getting started

0/3 complete

▼ Install Copilot in your editor

Ask about coding problems and get code completions while you work.

Visual Studio Code

▼

> Chat with Copilot anywhere

> Start building with Copilot

Usage

Inline suggestions

0% used

Included credits

0% used

Upgrade your Copilot plan for higher usage limits, premium models, AI reviews and more.

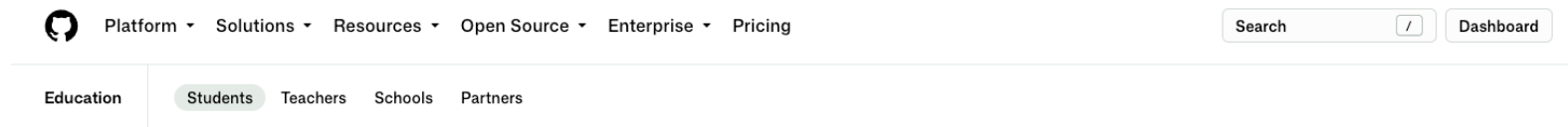
Upgrade plan

Github Copilot Pricing

			BEST VALUE
Free For lightweight chats \$0 USD Start for free Open in VS Code	Pro For medium agent tasks \$10 USD per user / month Choose Pro	Pro+ For complex tasks \$39 USD per user / month Choose Pro+	Max For sustained, high-volume tasks \$100 USD per user / month Choose Max
What's included: ▲ ✓ Access to entry level models ✓ 2,000 completions per month ✓ Community Support	Everything in Free and: ▲ ✓ Access to medium-complexity models ✓ Unlimited code completion and next edit suggestions ✓ Access to 3rd party agents (Claude Code and Codex) ✓ 1,500 monthly AI credits (\$15 value)	Everything in Pro and: ▲ ✓ Access to frontier models ✓ Audit logs ✓ Over 4.6x included monthly credits than Pro	Everything in Pro+ and: ▲ ✓ Early access to frontier models ✓ Over 13.3x included monthly credits than Pro

Github Education

<https://github.com/education/students>



GITHUB EDUCATION ■

Hey students, your tech career starts here

Open the doors to innovation and growth in the tech world

Join GitHub Education

Free GitHub Pro account for verified learners



3

Install & Configure Visual Studio Code

Installing VS Code

<https://code.visualstudio.com/download>

Download Visual Studio Code

Free and built on open source. AI agents, Git, debugging, and extensions included.



↓ Windows

Windows 10, 11

User Installer [x64](#) [Arm64](#)

System [x64](#) [Arm64](#)

Installer [x64](#) [Arm64](#)

.zip [x64](#) [Arm64](#)

CLI [x64](#) [Arm64](#)



↓ .deb

Debian, Ubuntu

↓ .rpm

Red Hat, Fedora, SUSE

.deb [x64](#) [Arm32](#) [Arm64](#)

.rpm [x64](#) [Arm32](#) [Arm64](#)

.tar.gz [x64](#) [Arm32](#) [Arm64](#)

Snap [Snap Store](#)

CLI [x64](#) [Arm32](#) [Arm64](#)



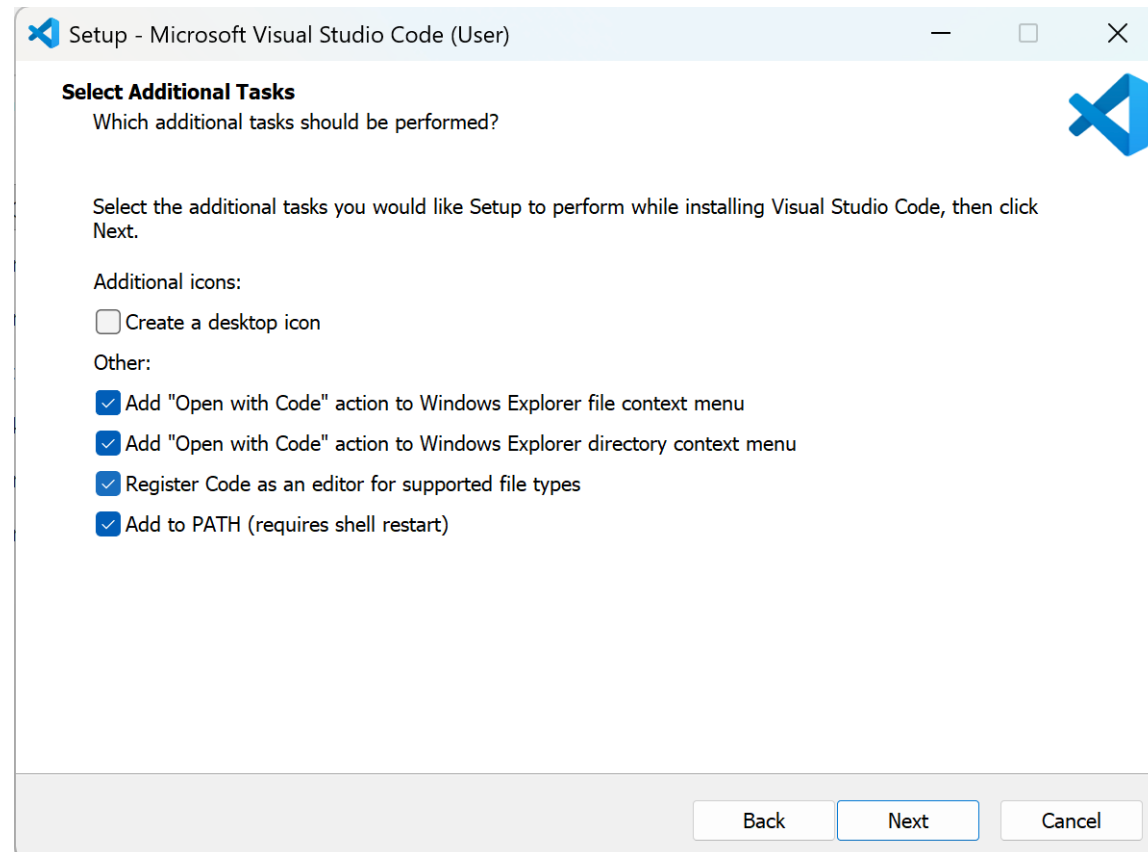
↓ Mac

macOS 12.0+

.dmg [Intel chip](#) [Apple silicon](#) [Universal](#)

CLI [Intel chip](#) [Apple silicon](#)

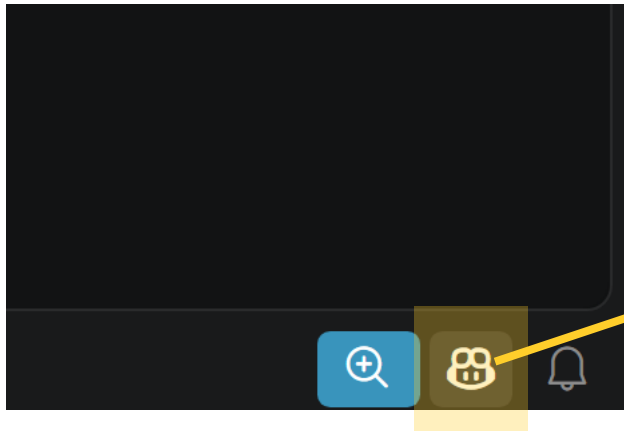
Configure VS Code



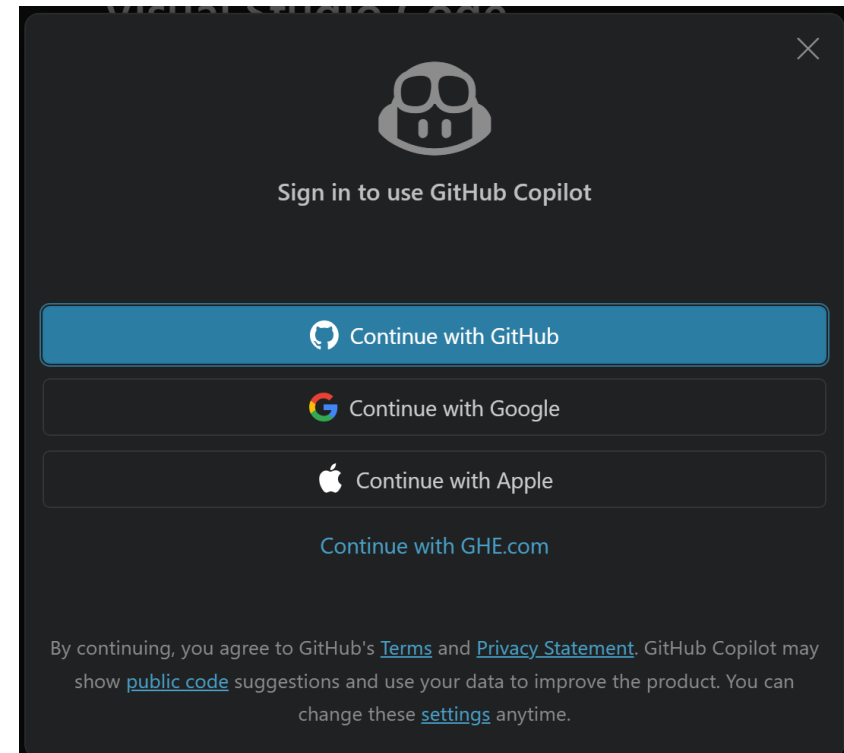
4

Configuring Github Account

Sign in your Github account



At lower-right corner





Authorize Visual Studio Code

From the options below, choose which account you would like to use to authorize this app.

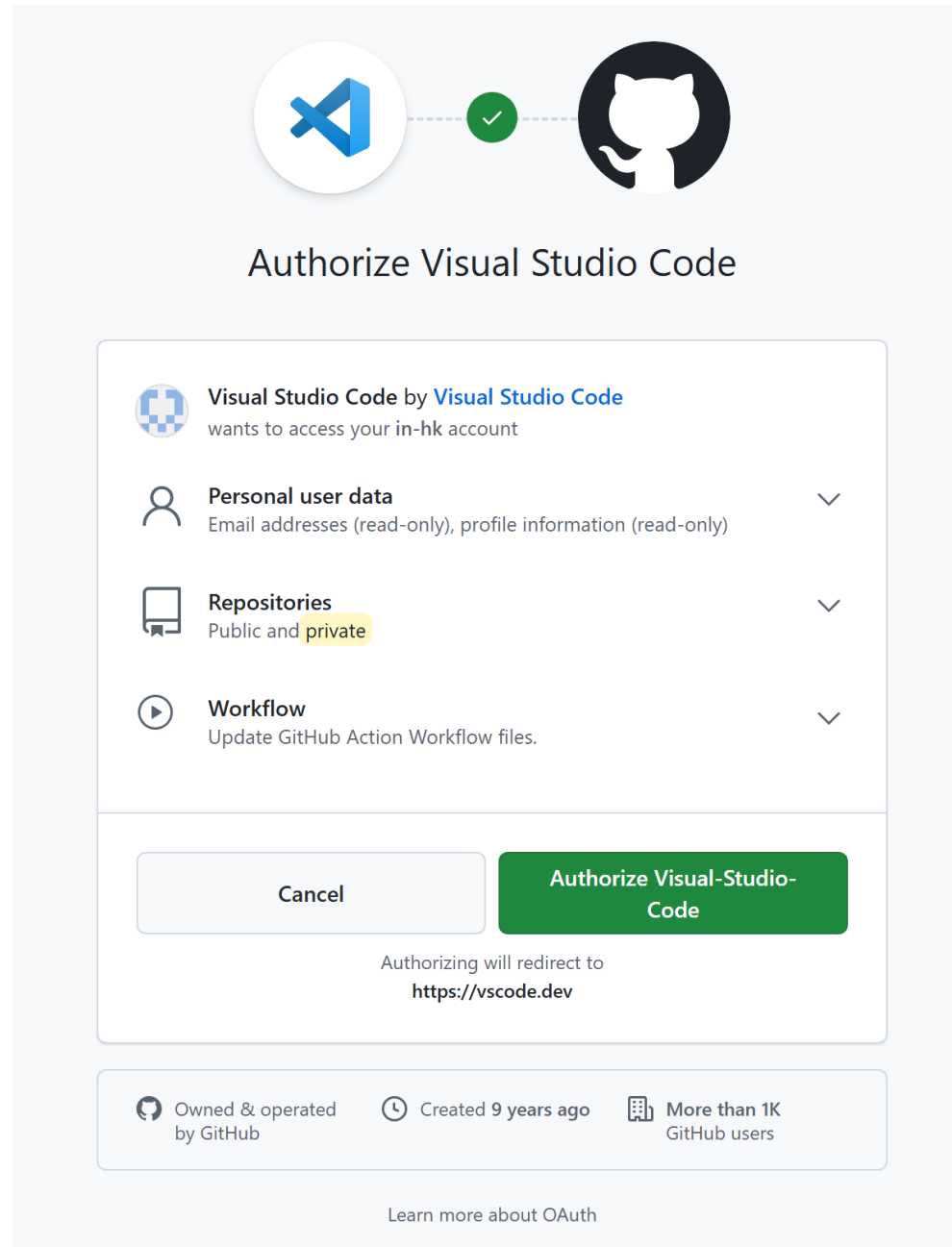


Signed in as
in-hk

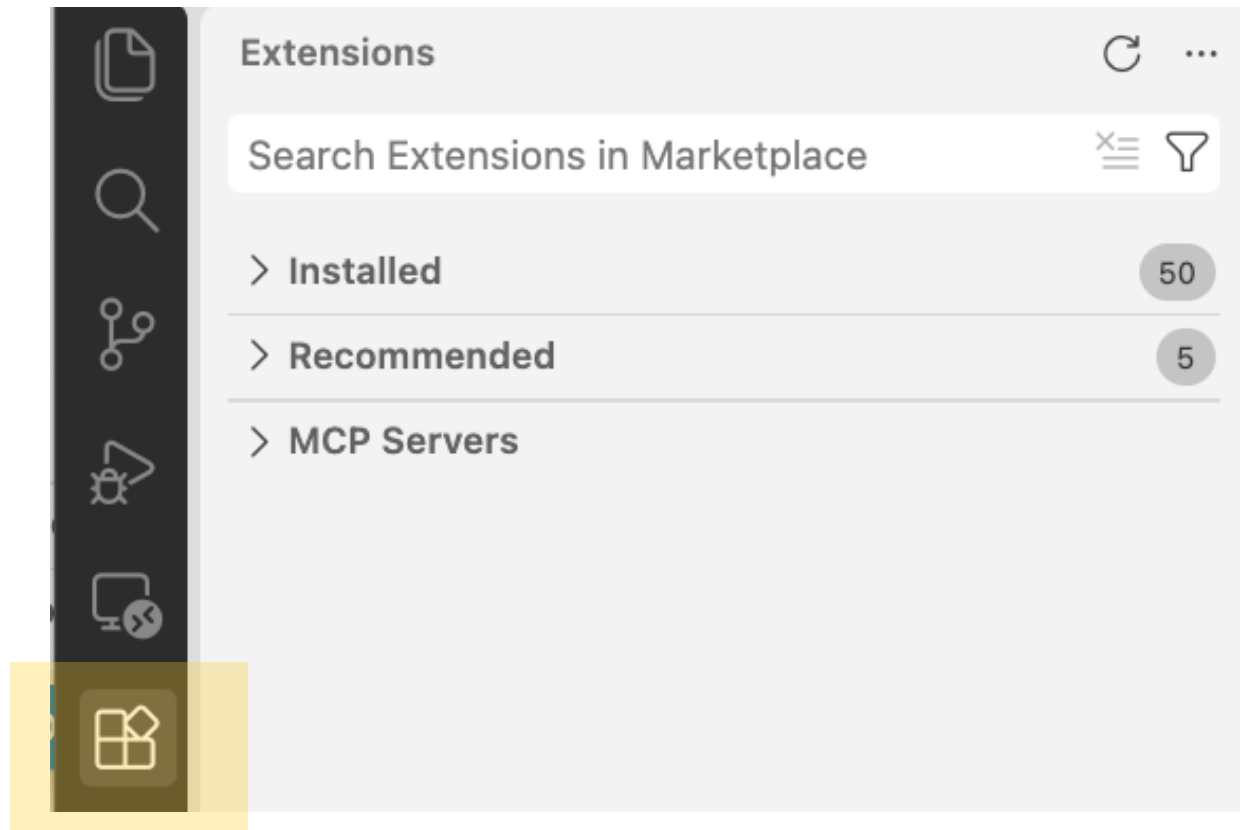
Continue

or

Use a different account

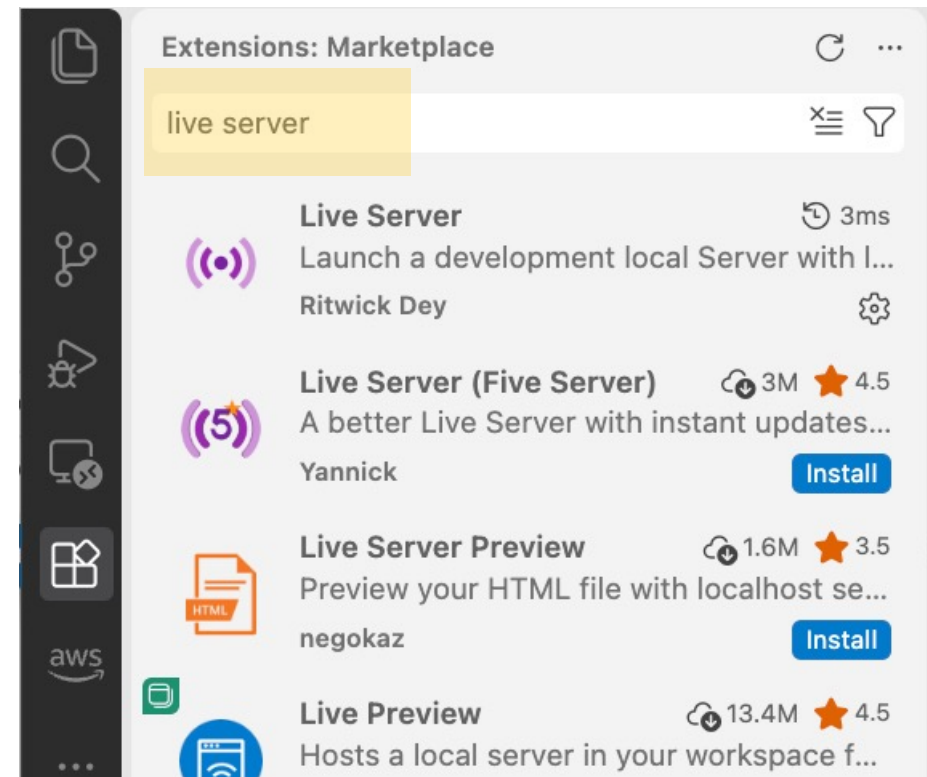


VS Code Extensions



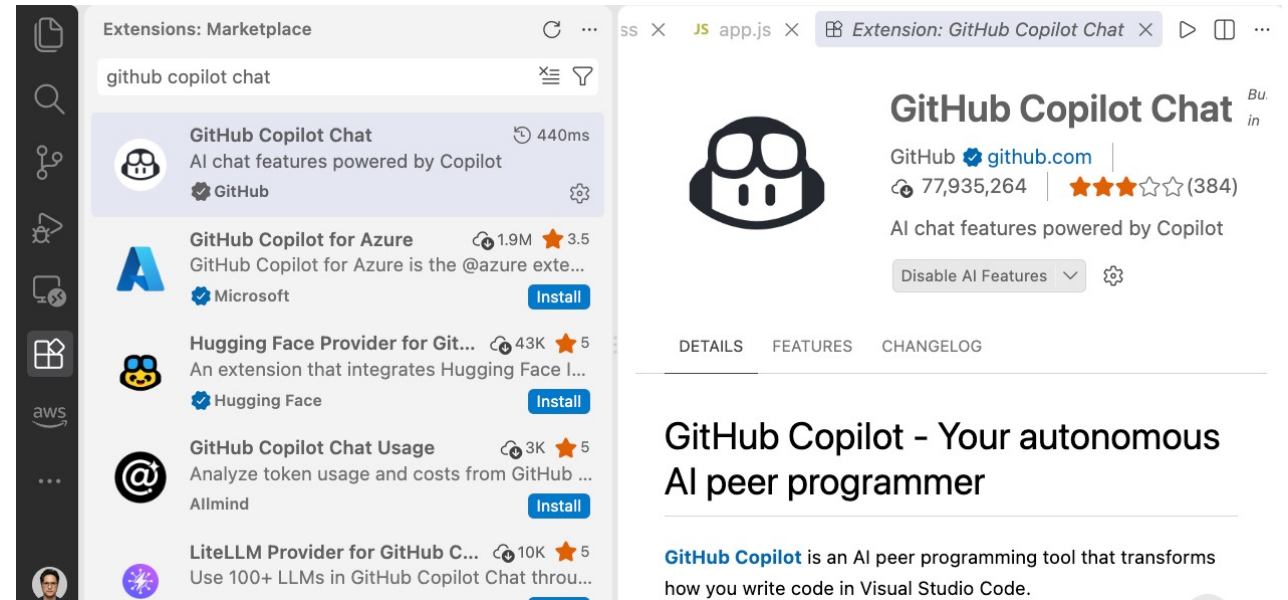
Installing “Live Server” Extension

- In the search box, type “**Live Server**” to search and install the extension for web development

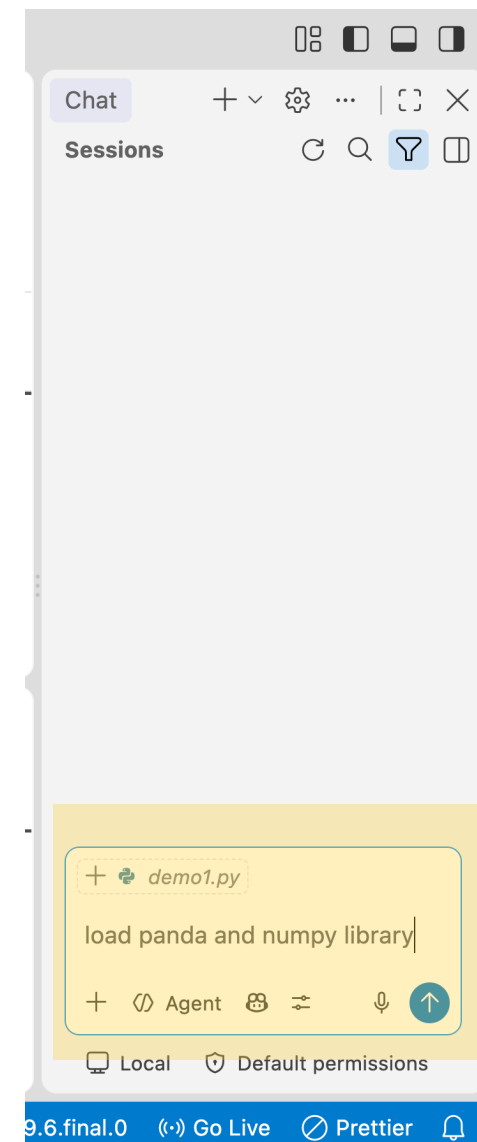
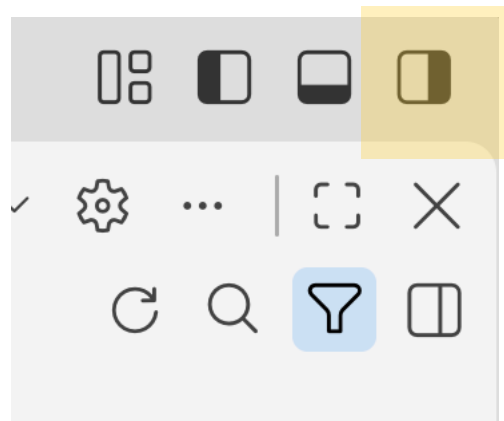


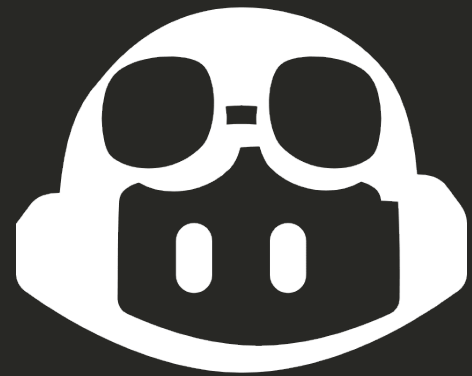
Check “Github Copilot Chat” Extension

- In the latest VS Code, this feature is pre-installed and therefore no need to install. You are suggested to install the latest version
- In case you are using older version of VS Code, please search and install it.



VS Code AI Pane





Getting Familiar with Github Copilot



Sarah's
Intent

Context
(Codebase, File Tree, Comments)

GitHub Copilot:
Intelligent Assistant

Sarah's Intent

Codium Assistant

Real-time Code
Suggestions

Contextual
Understanding

Security Scan

Optimized Code Snippets

Autocomplete

Refactoring Suggestions
Documentation

Global Cloud LLM

(Large Language Model)

OpenAI GPT-4, GitHub
Data Lake

Global Cloud LLM
(Large Language Model)

OpenAI GPT-4, GitHub Infrastructure,
Data Lake

Models in Github Copilot

<https://docs.github.com/en/copilot/reference/ai-models/supported-models>

 GitHub Docs

Version: Free, Pro, & Team

Search or ask Copilot



GitHub Copilot

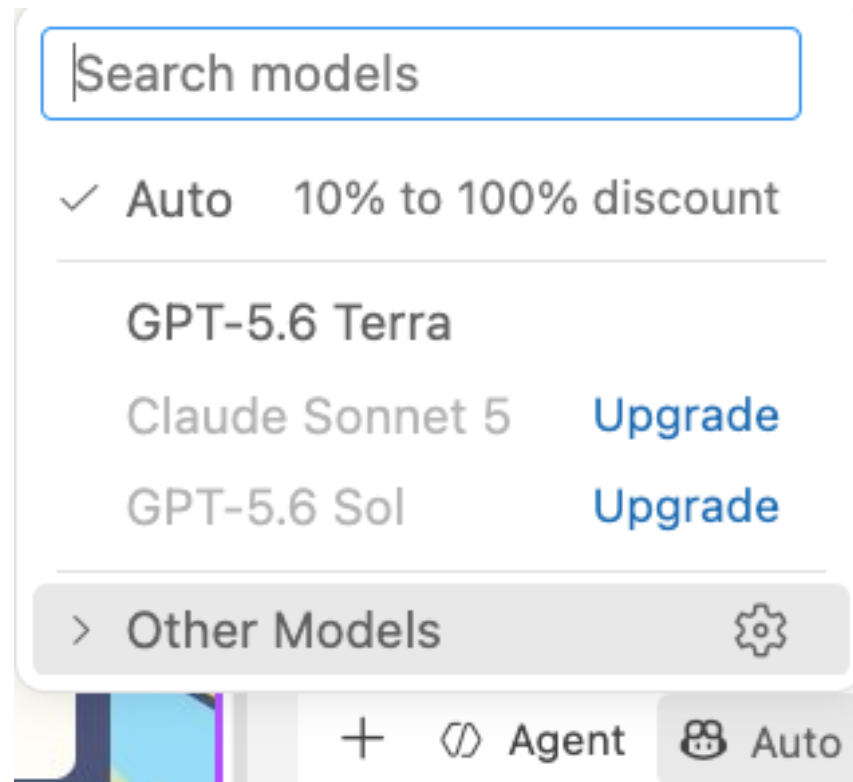
- Concepts
- How-tos
- Reference
 - Chat cheat sheet
 - Customization cheat sheet
 - AI models
 - Supported models**
 - Model comparison
 - Model hosting
 - Copilot feature matrix
 - Keyboard shortcuts
 - Copilot CLI reference
 - GitHub Copilot app reference
 - Custom agents configuration

Supported AI models in Copilot

This table lists the AI models available in Copilot, along with their release status.

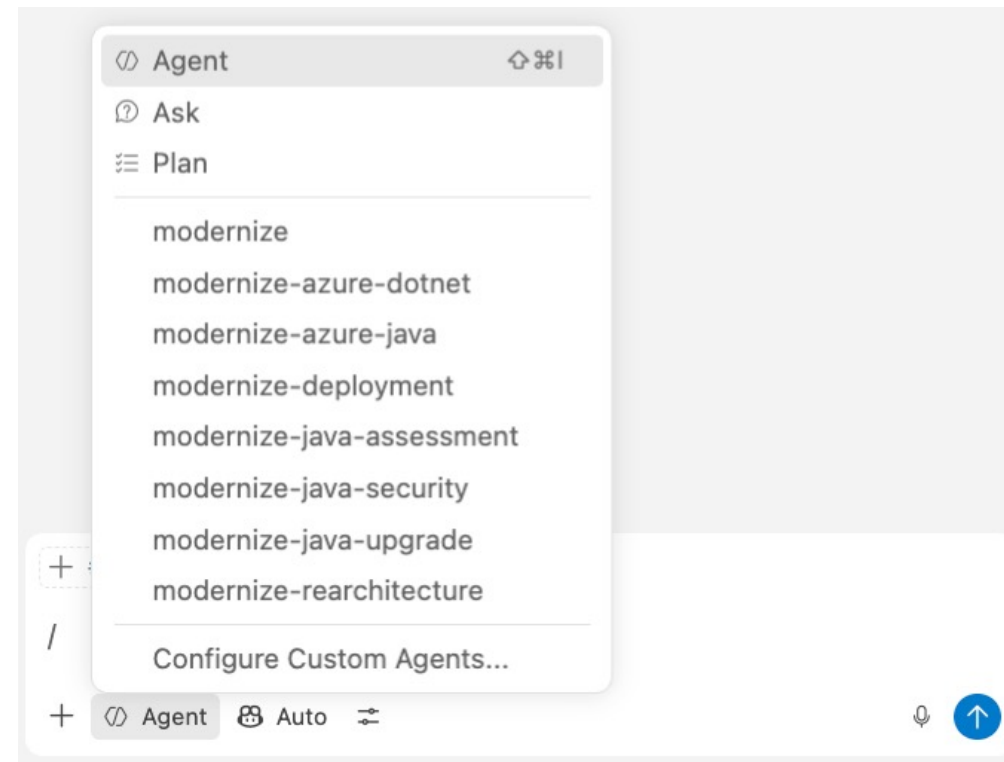
Model name	Provider	Release status
GPT-5 mini	OpenAI	GA
GPT-5.3-Codex	OpenAI	GA
GPT-5.4	OpenAI	GA
GPT-5.4 mini	OpenAI	GA
GPT-5.4 nano ^[1]	OpenAI	GA
GPT-5.5	OpenAI	GA
GPT-5.6 Luna	OpenAI	GA
GPT-5.6 Sol	OpenAI	GA
GPT-5.6 Terra	OpenAI	GA
Claude Fable 5 ^[2]	Anthropic	GA

Pick your favorite model



Interaction Modes

- **Agent** Mode: Autonomously executes multi-step tasks, edits files across your project, and runs commands to achieve a high-level goal.
- **Ask** Mode: Answers questions and explains code without modifying your workspace.
- **Plan** Mode: Explores your codebase to create a detailed, step-by-step implementation plan before any code is written.



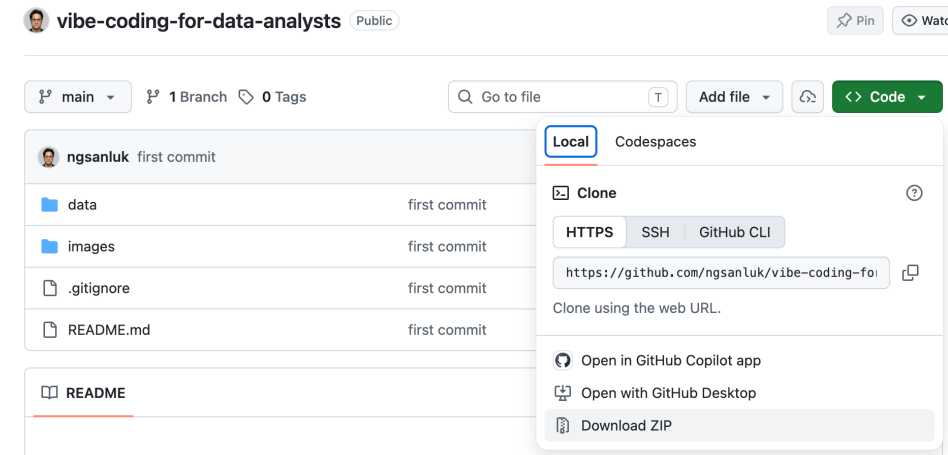
5

Opening Project

Download Project Codes

<https://github.com/ngsanluk/vibe-coding-for-data-analysts>

- Download the project codes
- Unzip the file
- Open the project folder in VS Code



6

Start Vibe Coding

Vibe Coding to do ...

Data Import

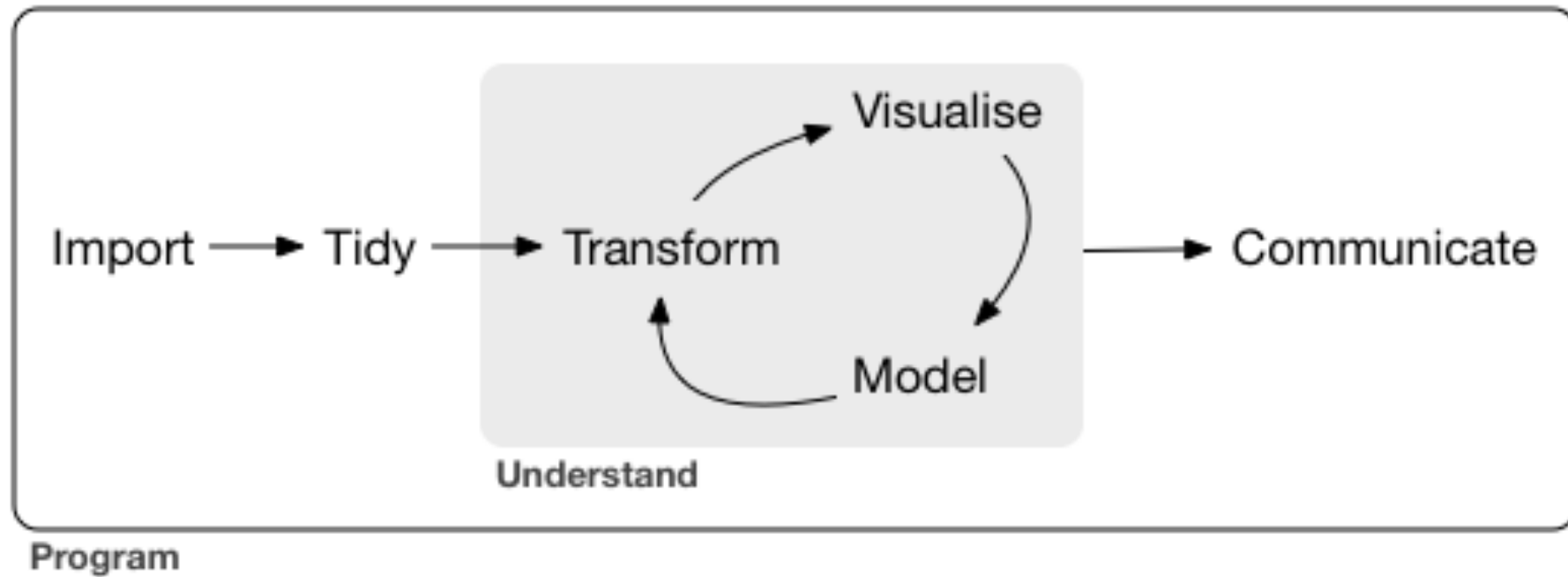
Tidy & Transform

Statistical Computation

Visual Story Telling

Executive Summary Generation

Data Science Workflow



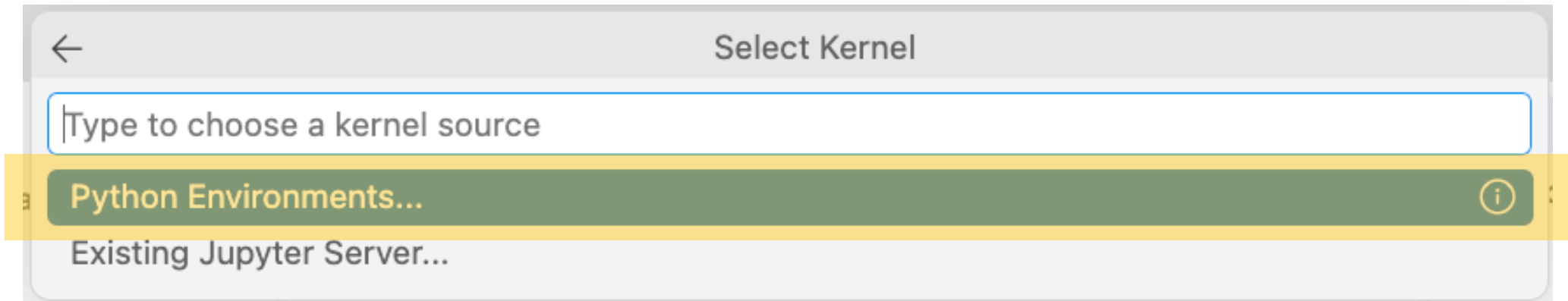
Working with Jupyter Notebook in Visual Studio Code

To create new Jupyter Notebook in VS Code

- There is existing Jupyter Notebook file in our project, you can double click to open it
- If you wish to create new notebook file, choose **View→Command Palette ...**
- In the popup command input box, Type and choose **Create: New Jupyter Notebook**
- By default, it is untitled notebook file, save the notebook file with your preferred file name before you start coding

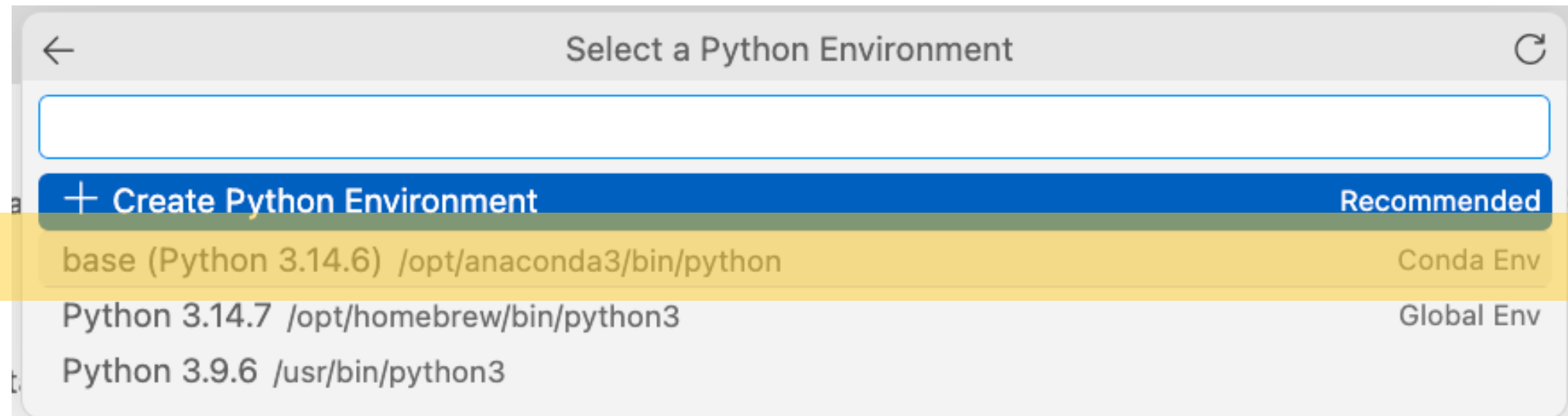
Choose your Python Environment 1/2

■ Choose Python Environments ...



Choose your Python Environment 2/2

- Choose Anaconda if you installed Anaconda
- Or choose any other python installed on your computer



Starting Building Up Your Github Portfolios

In an AI-driven market, an active GitHub portfolio is the ultimate proof that you can actually build and execute solutions, not just pitch them.

Enterprise AI Governance, Data Privacy & Verification

Enterprise AI Governance, Data Privacy & Verification

- Protecting Proprietary Corporate Data
 - Understanding enterprise data ingestion policies and anonymization best practices before feeding tables to public LLMs.
- Free AI tools often will feed your uploaded data for model training purpose
- Auditing AI calculations against ground-truth math to prevent hallucinations.

Thank You!