

May 2025
Irvine, CA

Going public with your CPS code and data

Jonathan Sprinkle (Vanderbilt University)

Cathy Wu (MIT)

Riley Wagner (Vanderbilt University)

Katie Dey (Vanderbilt University)

Stephen Rees (Vanderbilt University)

Matt Bunting (Vanderbilt University)

Junyi Ji (Vanderbilt University)

Tutorial, CPS-IoT Week 2025

Going public with your CPS code and data

Tutorial agenda

Session 1: Making your code do the talking

- 0900 - 0930 Introductions, the state of reproducibility
- 0930 - 1030 Code tutorial: make your code pretty and reusable
- 1030 - 1100 Coffee Break

Session 2: Creating an accessible and active resource

- 1100 - 1130 Demonstration of posting various types of resources with the CPS-VO
- 1130 - 1200 CPS open data sharing practice: I-24 MOTION testbed demonstration
- 1200 - 1230 CPS open data sharing practice: LADDMS demonstration



Quick survey

Make your code pretty and reusable

Going public with your CPS code and data

Junyi Ji - Vanderbilt University

CPS-IoT Week 2025

Tutorial

Prerequisites: <https://qtext.io/6sf0>

This tutorial is designed to be accessible. We welcome all skill levels.

You will need:

- A [Github](#) account (see [tutorial](#))
- Basic command line knowledge (see [tutorial](#))
 - If you can change directories, create & delete directories, and run a python script, you'll be fine.
 - Otherwise, you can take a few minutes to learn how. Here are some tutorials for [Unix](#) (Mac & Linux) and [Windows](#).
- Software requirements
 - A command line program (E.g., Terminal or [iTerm2](#) for MacOS, Command Prompt for Windows)
 - git (see [tutorial](#))
 - [Anaconda Python](#)
 - 🖱️ For Mac users with M1/M2 chips, please still install the Intel version
 - [Miniconda](#) should work fine, if you prefer.
 - A [Python IDE](#)
 - 🖱️ We will be using [Visual Studio Code](#) (VSCode). Install VSCode to best follow along.
- Knowledge of [git](#) and [python](#) are helpful but not required. Part 3 of the tutorial ("Make it pretty") will benefit from proficiency in Python.

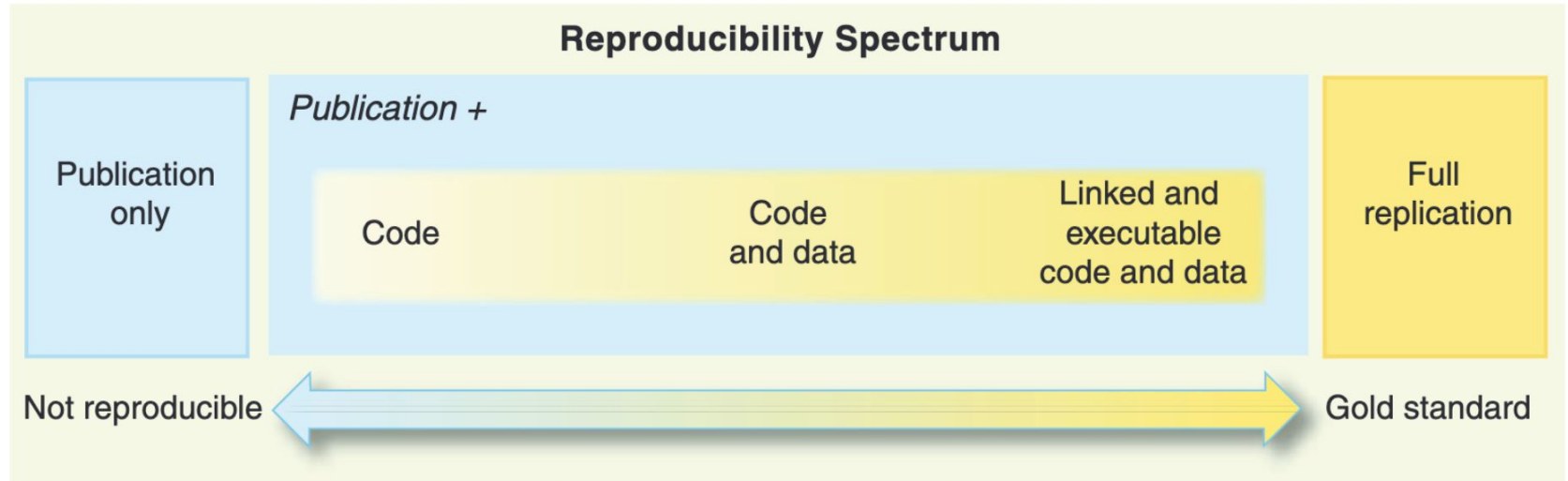
Reproducibility



Quick survey

- Reproducibility is obtaining **consistent results** using the same input data, computational steps, methods, and code, and conditions of analysis. This definition is synonymous with “computational reproducibility.”
- Replicability is obtaining consistent results across studies aimed at **answering the same scientific question**, each of which has obtained its own data. Two studies may be considered to have replicated if they obtain consistent results given the level of uncertainty inherent in the system under study.

Reproducibility levels



Where we are?

Transformation in the publication pipeline

1660's:

1. Enough detail on equipment, materials, and procedures for reproducibility
2. “Communal witnessing”
3. Exhaustive details on experimental settings, false starts, failures, etc.

1900's

Standards for journal publication: e.g. Introduction, Methods, Results, Discussion.

2025 and future

Executable workflows with links to corresponding data and results. Open accessibility.

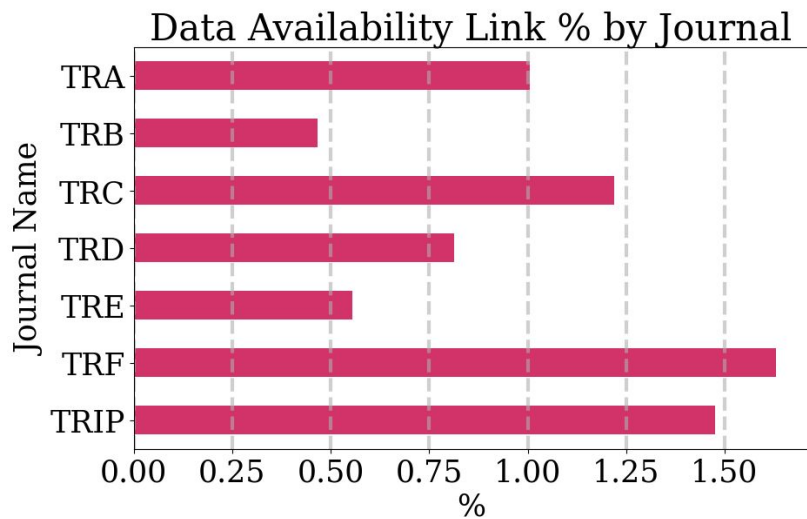


The question is: Who is the reader?

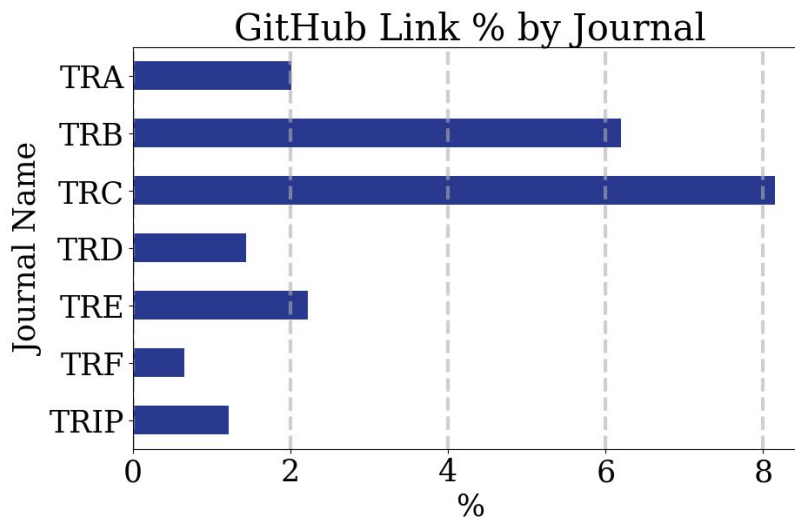
How to measure the credibility of the results?

Reproducibility - A closer look at empirical data

- Retrieved full-text data for 10,000+ papers (2019–2024) in TR journals



Data availability*: **1.0%**



Code availability*: **3.1%**

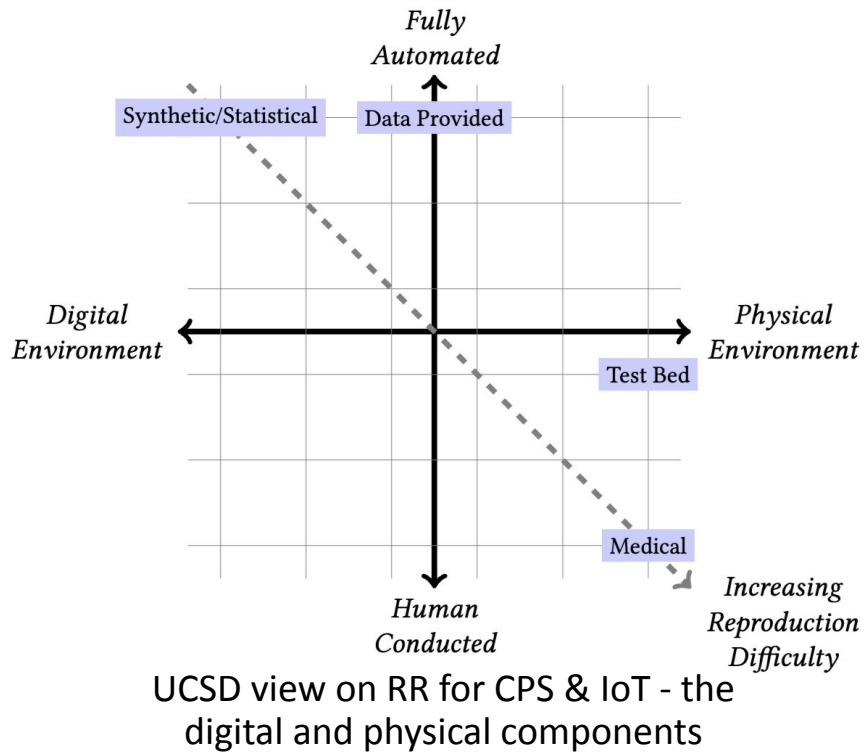
* Preliminary estimate. Could include use of open data or code from other projects, and exclude data or code mentioned elsewhere.

Is that our fault?

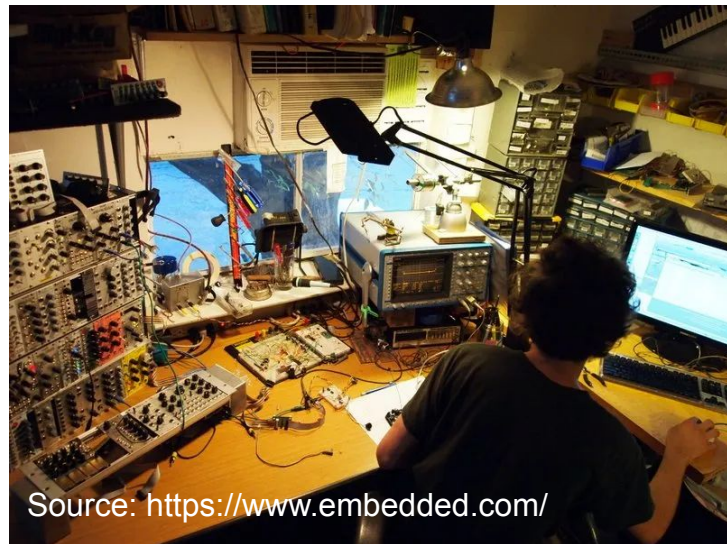
No!

The problem lies in the complexity
in the context of CPS.

Reproducibility in CPS



A typical hardware work deck



Source: <https://www.embedded.com/>

Instruments with software/hardware in the loop

My code before after this tutorial is like...



A small step further will bring in huge changes!

*Philosophy: **Get your feet wet with reproducibility***

From start to finish, create a **reproducible project on Github**

To **empower you** to make **your next research paper** reproducible.

Session goals & learning objective

- **Main goal:** Create a project that someone can reproduce in less than 5 minutes.
- **Stretch goal:** Organize the project and make it understandable.
- Learn principles and practical tools for making your research project reproducible. Including:
 - Fundamentals
 - Principles of organizing projects, code, & data
 - Working knowledge of Github
 - Advanced skills
 - Automatically extract the computing environment for reproducibility
 - VS Code and powerful extensions (tools for Jupyter, Markdown, csv, lint)

Prerequisites: <https://qtext.io/6sf0>

This tutorial is designed to be accessible. We welcome all skill levels.

You will need:

- A [Github](#) account (see [tutorial](#))
- Basic command line knowledge (see [tutorial](#))
 - If you can change directories, create & delete directories, and run a python script, you'll be fine.
 - Otherwise, you can take a few minutes to learn how. Here are some tutorials for [Unix](#) (Mac & Linux) and [Windows](#).
- Software requirements
 - A command line program (E.g., Terminal or [iTerm2](#) for MacOS, Command Prompt for Windows)
 - git (see [tutorial](#))
 - [Anaconda Python](#)
 - 🖱️ For Mac users with M1/M2 chips, please still install the Intel version
 - [Miniconda](#) should work fine, if you prefer.
 - A [Python IDE](#)
 - 🖱️ We will be using [Visual Studio Code](#) (VSCode). Install VSCode to best follow along.
- Knowledge of [git](#) and [python](#) are helpful but not required. Part 3 of the tutorial ("Make it pretty") will benefit from proficiency in Python.

Overview: Making your research project reproducible

- Key lesson: Don't worry about making things perfect!
- Priorities for research code:
 1. Make it run
 2. Make it available
 3. Make it pretty
- Available is better than perfect!

Outline

1. Make it run
2. Make it available
3. Make it pretty

Outline

1. Make it run

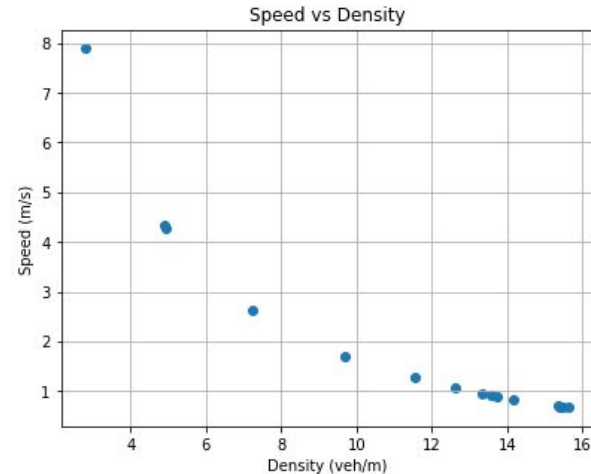
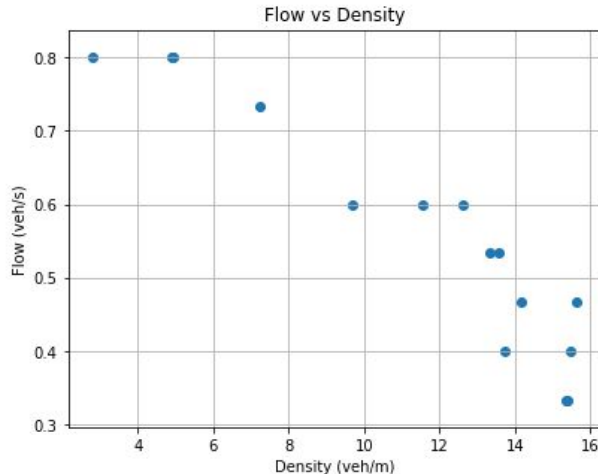
- a. For you: an ITS example; handling randomness
- b. For others, including future you: README, installation files

2. Make it available

3. Make it pretty

Make it run

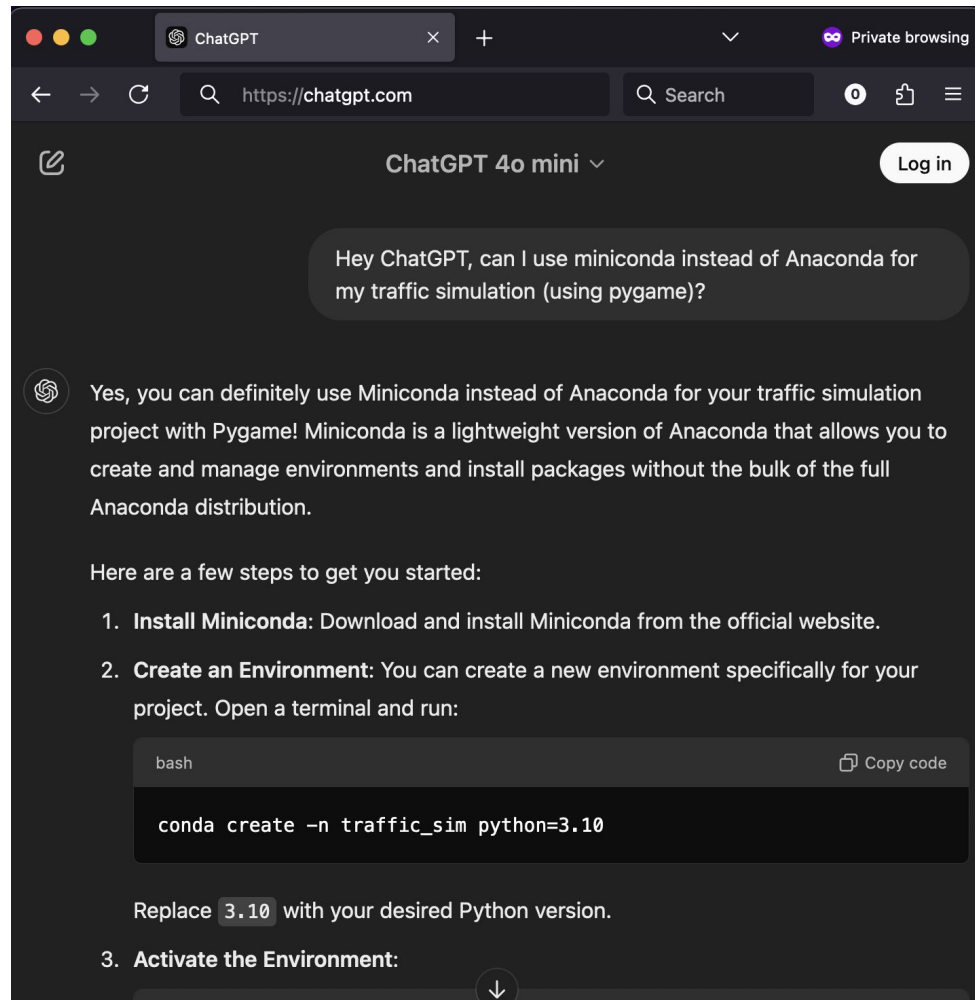
- **Exercise: Reproduce the car following simulation results**
- Download project files: <https://qtext.io/6sf0>
- **In 2 minutes, try to reproduce my plots. Go go go!**
 - Hint: `simulator.py` is the starting point.
- Fundamental diagrams generated from single lane traffic modeled using the Intelligent Driver Model (IDM)



By the way, if you get stuck...

- Try ChatGPT for resolving issues or clarify instructions.

- This will allow you to unblock yourself in the future in your own research project.



Fine, let me help you reproduce...

- `$ conda create -n cps25-RR-tutorial python=3.8`
- `$ conda activate cps25-RR-tutorial`

- `$ pip install matplotlib==3.1.0`
- `$ pip install pygame==2.6.0`
- `$ pip install pandas==2.0.3`
- `$ pip install numpy==1.24.4`
- `$ pip install jupyterlab==4.2.5`

- `$ python simulator.py --run-idm --no-render`

That was a lot of steps though. Easier...

- `$ conda env create -f environment.yml`
- `$ conda activate RR`

- `$ python simulator.py --run-idm --no-render`

Get the same result each time you run the code

- Handle randomness
- **Exercise: Reproduce the exact car following simulation results**
 - The project code we gave you will produce slightly different data and figures each time you run it.
- Solution: Fix random seed in Python and libraries (add to any `simulator.py`)
 - `import random`
 - `import numpy as np`
 - `random.seed(my_seed)`
 - `np.random.seed(my_seed)`
- Our reference example uses random seed **175175175**
- Check that there are no changes to the data/figures when you re-run `simulator.py`.

Outline

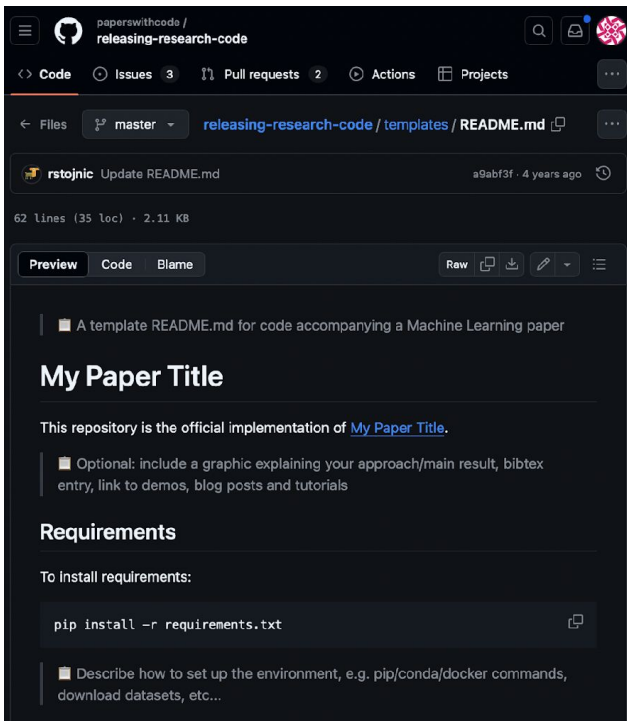
1. Make it run
 - a. For you: an ITS example; handling randomness
 - b. For others, including future you: README, installation files
2. Make it available
3. Make it pretty

Make it run, next time

Your closest collaborator is you six months ago,
but you don't reply to emails.

Overall project README

■ Exercise: Assemble a project README.md



Paper
reference

Any setup
commands

Training

To train the model(s) in the paper, run this command:

```
python train.py --input-data <path_to_data> --alpha 10 --beta 20
```

■ Describe how to train the models, with example commands on how to train the models in your paper, including the full training procedure and appropriate hyperparameters.

Evaluation

To evaluate my model on ImageNet, run:

```
python eval.py --model-file mymodel.pth --benchmark imagenet
```

■ Describe how to evaluate the trained models on benchmarks reported in the paper, give commands that produce the results (section below).

Pre-trained Models

You can download pretrained models here:

- [My awesome model](#) trained on ImageNet using parameters x,y,z.

■ Give a link to where/how the pretrained models can be downloaded and how they were trained (if applicable). Alternatively you can have an additional column in your results table with a link to the models.

Results

Our model achieves the following performance on :

Image Classification on ImageNet

Model name	Top 1 Accuracy	Top 5 Accuracy
My awesome model	85%	95%

■ Include a table of results from your paper, and link back to the leaderboard for clarity and context. If your main result is a figure, include that figure and link to the command or notebook to reproduce it.

Contributing

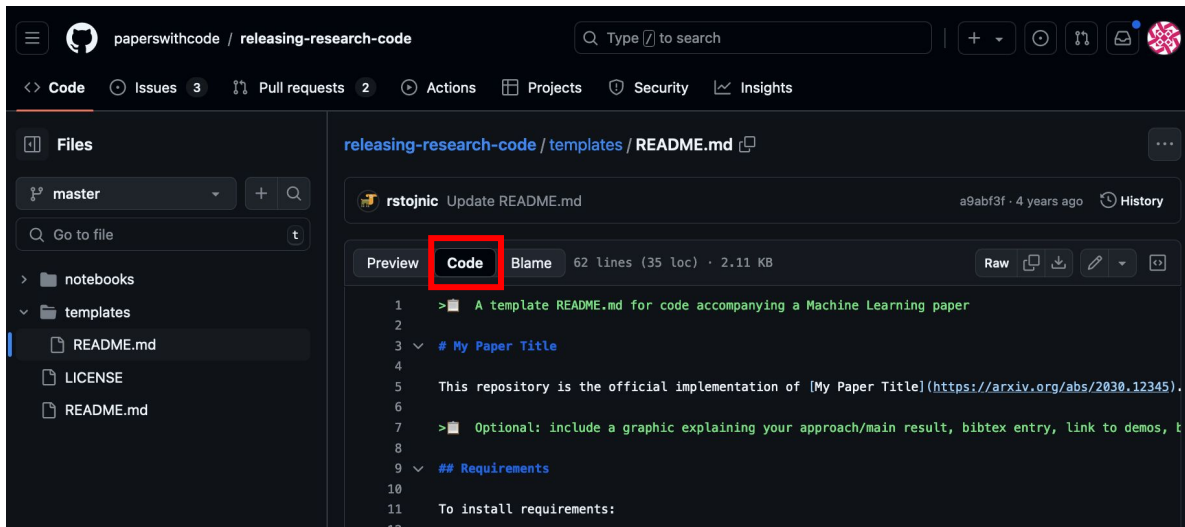
■ Pick a licence and describe how to contribute to your code repository.

Commands to
reproduce data

Commands to
reproduce plots

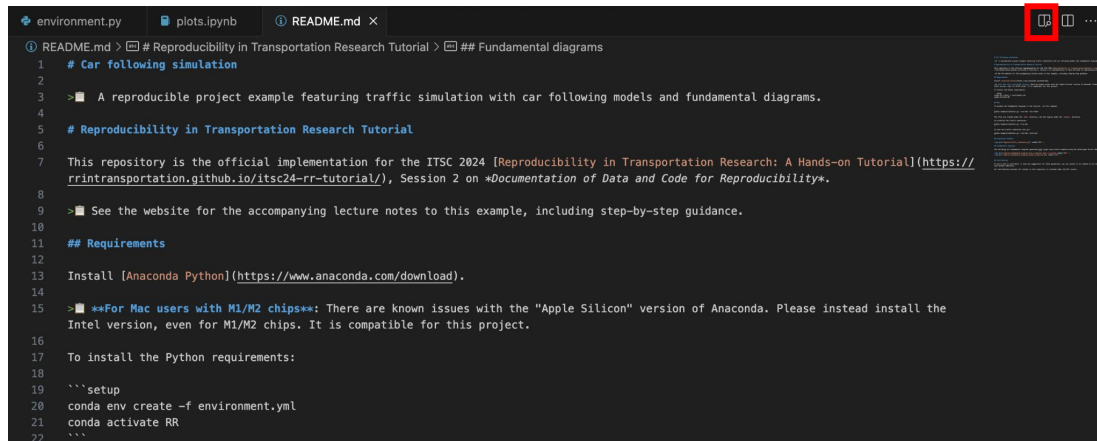
Optional tips: Copy-paste a README template

- Go to <https://github.com/paperswithcode/releasing-research-code/tree/master>
- Within the templates folder, open README.md
- Select the Code tab
- From here you can copy and paste the template into your project, and adapt it to your liking.
- Templates usually include more than what you need, so delete or adjust the sections based on what others need to reproduce your work.



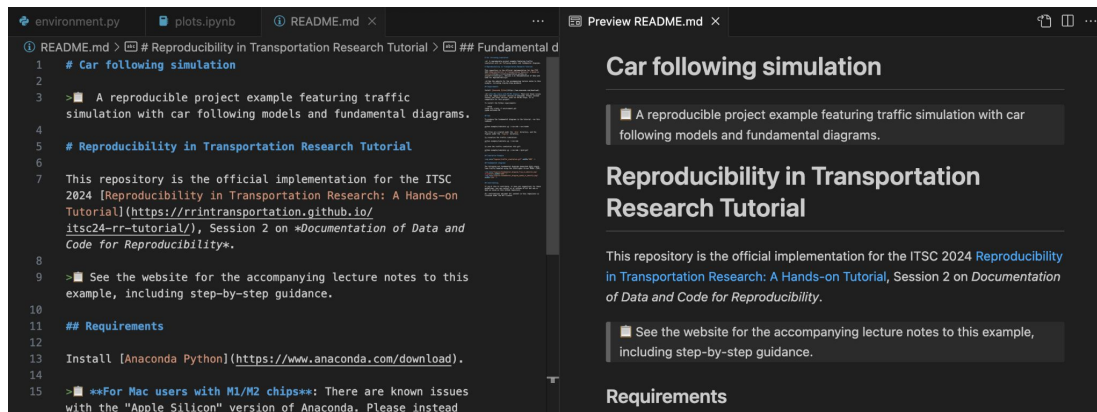
Optional tips: Editing a README.md in VS Code

Open Markdown preview



```
environment.py  plots.ipynb  README.md X
1  README.md > # Reproducibility in Transportation Research Tutorial > ## Fundamental diagrams
2  1 # Car following simulation
3  2
4  3 >■ A reproducible project example featuring traffic simulation with car following models and fundamental diagrams.
5  4
6  5 # Reproducibility in Transportation Research Tutorial
7  6
8  7 This repository is the official implementation for the ITSC 2024 [Reproducibility in Transportation Research: A Hands-on Tutorial](https://
9  8 rrintransportation.github.io/itsc24-rr-tutorial/), Session 2 on *Documentation of Data and Code for Reproducibility*.
10 9
11 10 >■ See the website for the accompanying lecture notes to this example, including step-by-step guidance.
12 11
13 12 ## Requirements
14 13
15 14 Install [Anaconda Python](https://www.anaconda.com/download).
16 15
17 16 >■ **For Mac users with M1/M2 chips**: There are known issues with the "Apple Silicon" version of Anaconda. Please instead install the
18 17 Intel version, even for M1/M2 chips. It is compatible for this project.
19 18
20 19 To install the Python requirements:
21 20
22 21 ```setup
23 22 conda env create -f environment.yml
24 23 conda activate RR
25 24 ```
```

Side-by-side Markdown +
preview



```
environment.py  plots.ipynb  README.md X  Preview README.md X
1  README.md > # Reproducibility in Transportation Research Tutorial > ## Fundamental d
2  1 # Car following simulation
3  2
4  3 >■ A reproducible project example featuring traffic
5  4 simulation with car following models and fundamental diagrams.
6  5
7  6 # Reproducibility in Transportation Research Tutorial
8  7
9  8 This repository is the official implementation for the ITSC
10 9 2024 [Reproducibility in Transportation Research: A Hands-on
11 10 Tutorial](https://rrintransportation.github.io/
12 11 itsc24-rr-tutorial/), Session 2 on *Documentation of Data and
13 12 Code for Reproducibility*.
14 13
15 14 >■ See the website for the accompanying lecture notes to this
16 15 example, including step-by-step guidance.
17 16
18 17 ## Requirements
19 18
20 19 Install [Anaconda Python](https://www.anaconda.com/download).
21 20
22 21 >■ **For Mac users with M1/M2 chips**: There are known issues
23 22 with the "Apple Silicon" version of Anaconda. Please instead
```

Car following simulation

■ A reproducible project example featuring traffic simulation with car following models and fundamental diagrams.

Reproducibility in Transportation Research Tutorial

This repository is the official implementation for the ITSC 2024 [Reproducibility in Transportation Research: A Hands-on Tutorial](https://rrintransportation.github.io/itsc24-rr-tutorial/), Session 2 on [Documentation of Data and Code for Reproducibility](#).

■ See the website for the accompanying lecture notes to this example, including step-by-step guidance.

Requirements

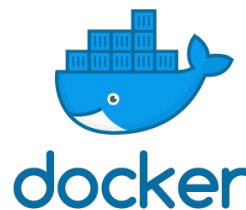
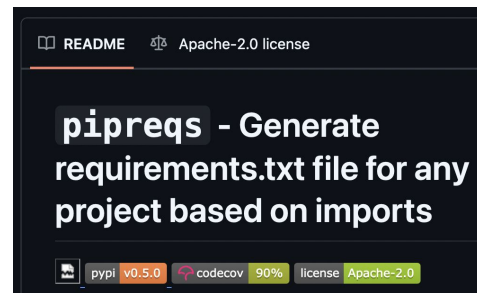
Project setup files

■ Exercise: Create installation files for the project

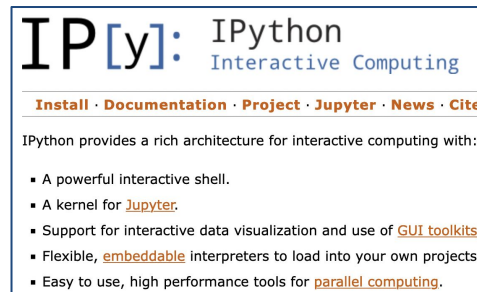
- Option 1: Extract environment info
 - Often good enough for research
 - User A: Programmatically extract the environment information from the kernel (e.g., anaconda Python)
 - User B: Install environment information in local system, then run code

■ Option 2: Docker

- More advanced, but simpler for the user
- User A: Create a Docker image of your project
- User B: Run image in any environment (Windows, Linux, MacOS, etc.)



docker



Extract installation information

```
l/itsc24-rr-tutorial-example-messy$ python --version  
Python 3.8.0
```

Python version

Packages based
on imports

```
setup-info > requirements-imports.txt  
1 matplotlib==3.1.0  
2 numpy==1.24.4  
3 pandas==2.0.3  
4 pygame==2.6.0  
5
```

```
setup-info > requirements-all-dependencies.txt  
1 anyio==4.4.0  
2 appnope==0.1.4  
3 argon2-cffi==23.1.0  
4 argon2-cffi-bindings==21.2.0  
5 arrow==1.3.0  
6 asttokens==2.4.1  
7 async-lru==2.0.4  
8 attrs==24.2.0  
9 babel==2.16.0  
10 backcall==0.2.0  
11 beautifulsoup4==4.12.3  
12 bleach==6.1.0  
13 certifi==2024.8.30  
14 cffi==1.17.1  
15 charset-normalizer==3.3.2  
16 comm==0.2.2  
17 cycler==0.12.1  
18 debugpy==1.8.5  
19 decorator==5.1.1  
20 defusedxml==0.7.1  
21 docopt==0.6.2  
22 exceptiongroup==1.2.2  
23 executing==2.1.0  
24 fastjsonschema==2.20.0  
25 fqdn  
26 h11  
27 httpx  
28 httpx  
29 idna  
30 ipykernel==6.29.5  
31 importlib_resources==6.4.5
```

All packages in
conda env

Create a conda environment.yml file (1/3)

- Create a blank file `environment.yml`. This file maintains the software dependencies for easy installation by others (& future you).

```
! environment.yml
1  name: RR
2
3  dependencies:
4      - pip
5      - python==3.8
6      - pip:
7          - matplotlib==3.7.5
8          - pygame==2.6.0
9          - pandas==2.0.3
10         - numpy==1.24.4
11         - pillow==10.4.0
```

A short name for the conda environment of your project (no spaces)

Python dependencies

Format: `library_name==version_number`

A user will install and activate the environment by:

```
$ conda env create -f environment.yml
$ conda activate RR
```

Create a conda environment.yml file (2/3)

Commands to extract the installation information:

- Get python version

- `$ python --version`

- Get module information

- `$ pipreqs . --force --savepath requirements-imports.txt`
 - To install it: `$ pip install pipreqs`
- `$ pip3 freeze > requirements-all-dependencies.txt`

- [pipreqs](#) vs pip3 freeze

- pipreqs: Based on imports, excludes package dependencies
- pip3 freeze: Anything installed by pip in the current conda env, including package dependencies

Create a conda environment.yml file (3/3)

- Fill in the `environment.yml` with the extracted installation information.

```
! environment.yml
1  name: RR
2
3  dependencies:
4      - pip
5      - python==3.8
6      - pip:
7          - matplotlib==3.7.5
8          - pygame==2.6.0
9          - pandas==2.0.3
10         - numpy==1.24.4
11         - pillow==10.4.0
```

Fill in from `$ python --version`. Some platforms require excluding the patch version (e.g. use 3.8 instead of 3.8.0).

Fill in from `requirements-imports.txt`. Add whatever else you'd like to. For example, add `jupyterlab`, to run more code snippets.

Outline

1. Make it run
2. **Make it available**
 - a. Open-source licenses
 - b. Get it on Github
3. Make it pretty

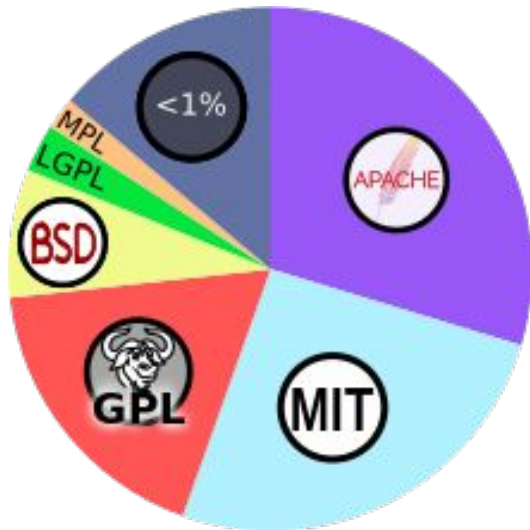
Code

“An article about computational science in a scientific publication is not the scholarship itself, it is merely advertising of the scholarship. The actual scholarship is the complete software development environment and the complete set of instructions which generated the figures.”

– [Buckheit and Donoho, 1995](#)

Make it available

1. Choose an open-source license.
2. Put your project on Github.




Open-source license, Wikipedia (2021)

tl;dr: Use the MIT license

Most common for research code



						
Type	Permissive	Permissive	Permissive	Copyleft	Copyleft	Copyleft
Provides copyright protection	✓ TRUE	✓ TRUE	✓ TRUE	✓ TRUE	✓ TRUE	✓ TRUE
Can be used in commercial applications	✓ TRUE	✓ TRUE	✓ TRUE	✓ TRUE	✓ TRUE	✓ TRUE
Provides an explicit patent license	✓ TRUE	✗ FALSE	✗ FALSE	✗ FALSE	✗ FALSE	✗ FALSE
Can be used in proprietary (closed source) projects	✓ TRUE	✓ TRUE	✓ TRUE	✗ FALSE	✗ FALSE partially	✗ FALSE for web
Popular open-source and free projects	Kubernetes Swift Firebase	Django React Flutter	Angular.js jQuery, .NET Core Laravel	Joomla Notepad++ MySQL	Qt SharpDevelop	SugarCRM Launchpad

35

Outline

1. Make it run
2. **Make it available**
 - a. Open-source licenses
 - b. **Get it on Github**
3. Make it pretty

Exercise: Get the code up on Github!

1. Set up git
2. Create a new repo on Github
3. Download the empty repo to your local machine
4. Add files to a new workspace for the repo
5. Commit & sync files to Github

If you're familiar with Github and your IDE, feel free to try without the step-by-step directions!

1. Installing Git (1/2)

- First check if you already have Git installed
 - In terminal/ command line: `git --version`
 - If you see a git version, you are all set. If not, follow the instructions below.
- On Mac
 - Via Homebrew: `brew install git`
 - Via Xcode: Apple ships a binary package of Git with Xcode. Therefore, just download Xcode from App store and install it.
- On Linux
 - Via Apt: `apt-get install git`
- On Windows
 - Download the standalone installer and install it:
<https://git-scm.com/download/win>

1. Setting up Git

- Open terminal /command line
 - Configure your details
 - `git config --global user.email "you@example.com"`
 - `git config --global user.name "Your Name"`

1. Setting up Git (2/6)

■ Git authentication

- To push/pull from a repository in Github, you will need to authenticate with a username and password.
 - Username: your Github username
 - Password: generate an authentication as follows
 - Settings → Developer settings → Personal access tokens → Tokens (classic)
 - Click Generate new token and select generate new token (classic option)

GitHub Apps

OAuth Apps

Personal access tokens

Fine-grained tokens Beta

Tokens (classic)

Personal access tokens (classic)

Generate new token Revoke all

Tokens you have generated that can be used to access the [GitHub API](#).

M1-new — `admin:enterprise, admin:pgp_key, admin:org, admin:org_hook,` Last used within the last week Delete
`admin:public_key, admin:repo_hook, admin:ssh_signing_key, delete:packages, delete_repo, gist, notifications, project,`
`repo, user, workflow, write:discussion, write:packages`

⚠ [This token has no expiration date.](#)

Personal access tokens (classic) function like ordinary OAuth access tokens. They can be used instead of a password for Git over HTTPS, or can be used to [authenticate to the API over Basic Authentication](#).

1. Setting up Git (3/6)

- Git authentication (cont...)
 - Provide a name for the token
 - Configure token expiration day
 - Select token scope (ex: what can someone do with the token - generally 'repo' scope is sufficient)

New personal access token (classic)

Personal access tokens (classic) function like ordinary OAuth access tokens. They can be used instead of a password for Git over HTTPS, or can be used to [authenticate to the API over Basic Authentication](#).

Note

What's this token for?

Expiration *

30 days

The token will expire on Thu, Oct 17 2024

Select scopes

Scopes define the access for personal tokens. [Read more about OAuth scopes](#).


<input type="checkbox"/> repo	Full control of private repositories
<input type="checkbox"/> repo:status	Access commit status
<input type="checkbox"/> repo_deployment	Access deployment status
<input type="checkbox"/> public_repo	Access public repositories
<input type="checkbox"/> repo:invite	Access repository invitations
<input type="checkbox"/> security_events	Read and write security events
<input type="checkbox"/> workflow	Update GitHub Action workflows
<input type="checkbox"/> write:packages	Upload packages to GitHub Package Registry
<input type="checkbox"/> read:packages	Download packages from GitHub Package Registry
<input type="checkbox"/> delete:packages	Delete packages from GitHub Package Registry
<input type="checkbox"/> admin:org	Full control of orgs and teams, read and write org projects
<input type="checkbox"/> write:org	Read and write org and team membership, read and write org projects
<input type="checkbox"/> read:org	Read org and team membership, read org projects
<input type="checkbox"/> manage_runners:org	Manage org runners and runner groups
<input type="checkbox"/> admin:public_key	Full control of user public keys
<input type="checkbox"/> write:public_key	Write user public keys
<input type="checkbox"/> read:public_key	Read user public keys



1. Setting up Git (4/6)

- Git authentication (cont...)
 - Copy the token.
 - Note: once you refresh the page, you can not no longer view the token and will have to regenerate.
 - Keep this token saved somewhere as we will use this later

Personal access tokens (classic) Generate new token ▾ Revoke all

Tokens you have generated that can be used to access the [GitHub API](#).

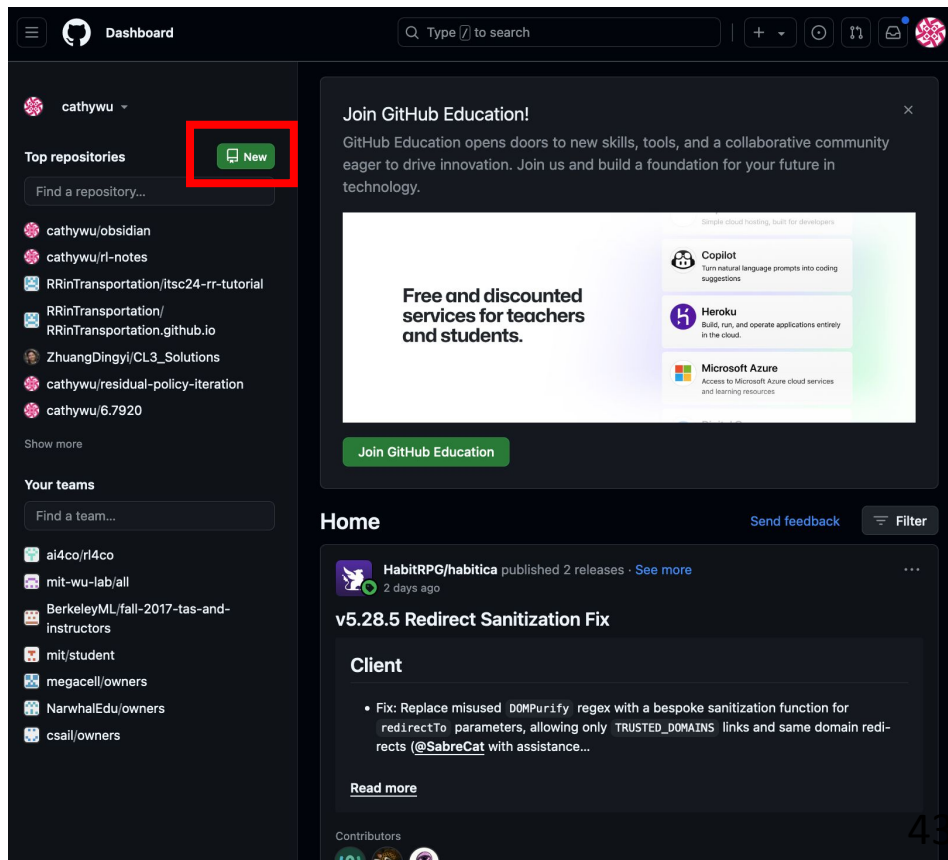
 Make sure to copy your personal access token now. You won't be able to see it again!

✓ 	Delete
itsc — repo Expires on Thu, Oct 17 2024.	Never used Delete
M1-new — admin:enterprise, admin:gpg_key, admin:org, admin:org_hook, admin:public_key, admin:repo_hook, admin:ssh_signing_key, delete:packages, delete_repo, gist, notifications, project, repo, user, workflow, write:discussion, write:packages  This token has no expiration date.	Last used within the last week Delete

Personal access tokens (classic) function like ordinary OAuth access tokens. They can be used instead of a password for Git over HTTPS, or can be used to [authenticate to the API over Basic Authentication](#).

2. Create a new repo on Github (1/2)

- Sign in to github.com



2. Create a new repo on Github (2/2)

- Choose a repo name and select an open-source license
- `.gitignore` files indicate which files NOT to add to the repo, like config or password files.

Create a new repository

A repository contains all project files, including the revision history. Already have a project repository elsewhere? [Import a repository](#).

Required fields are marked with an asterisk (*).

Repository template

No template ▾

Start your repository with a template repository's contents.

Owner * cathywu ▾

Repository name * itsc24-rr-tutorial-example
✔ itsc24-rr-tutorial-example is available.

Great repository names are short and memorable. Need inspiration? How about [glowing-lamp](#) ?

Description (optional)

☒ **Public**
Anyone on the internet can see this repository. You choose who can commit.

☐ **Private**
You choose who can see and commit to this repository.

Initialize this repository with:

☒ **Add a README file**
This is where you can write a long description for your project. [Learn more about READMEs](#).

Add .gitignore

.gitignore template: Python ▾

Choose which files not to track from a list of templates. [Learn more about ignoring files](#).

Choose a license

License: MIT License ▾

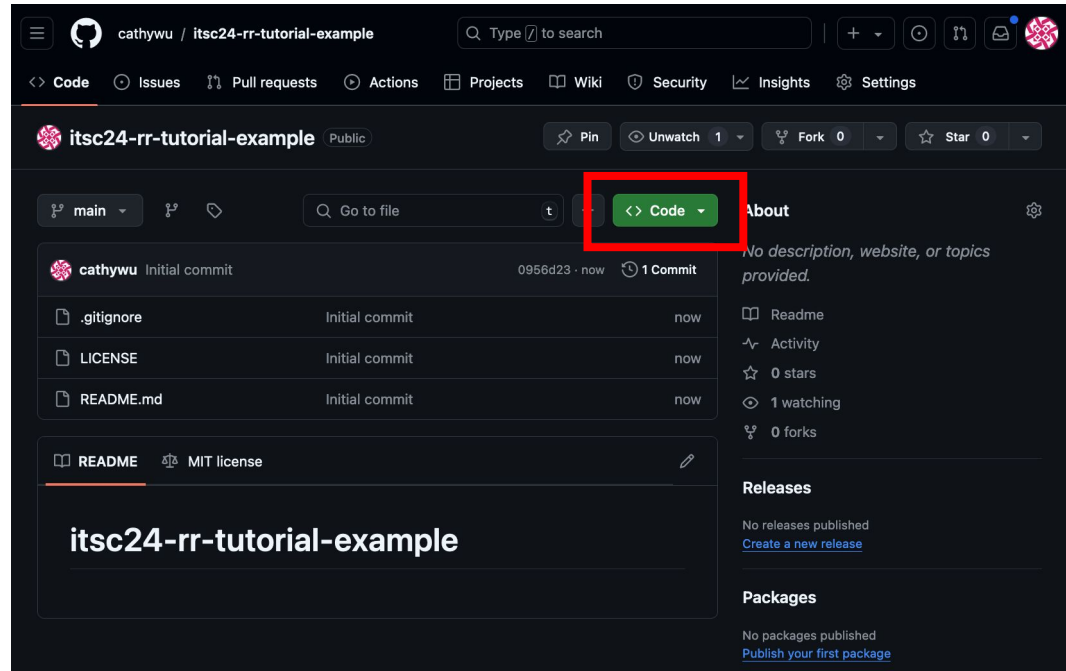
A license tells others what they can and can't do with your code. [Learn more about licenses](#).

This will set `main` as the default branch. Change the default name in your [settings](#).

🔔 You are creating a public repository in your personal account.

[Create repository](#)

3. Download the empty repo to your local machine (1/2)



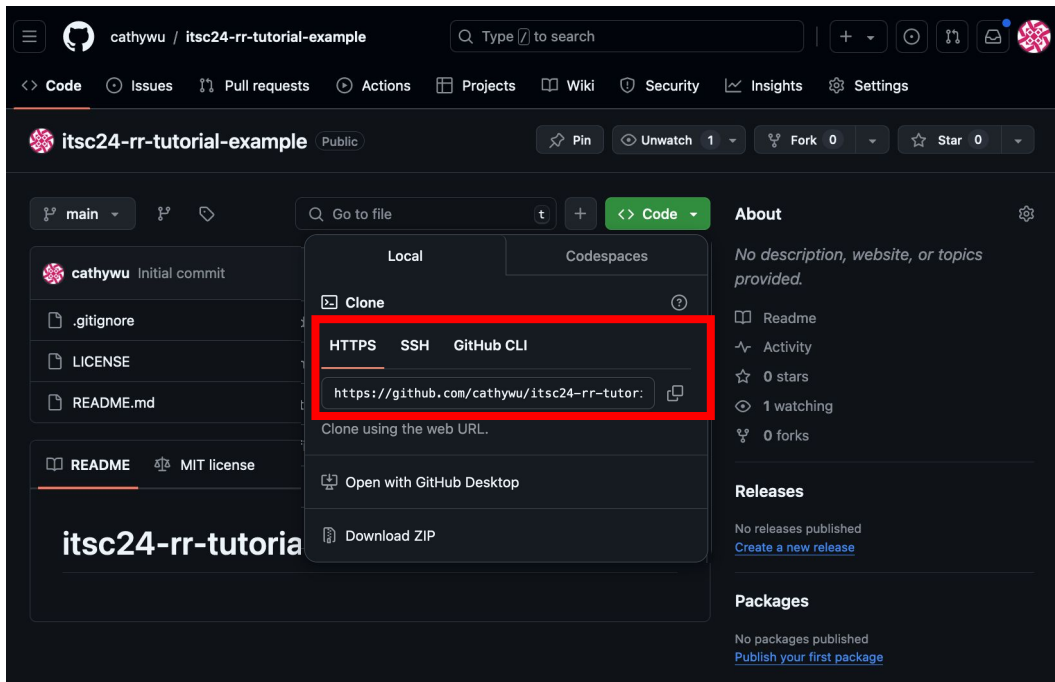
3. Download the empty repo to your local machine (1/2)

- On the website:

- Copy the `https://github.com/`
...
- Use the HTTPS option

- In the local terminal

- Navigate to where you want to download the repo (`$ cd`
...)
- `$ git clone`
`https://github.com/`
...

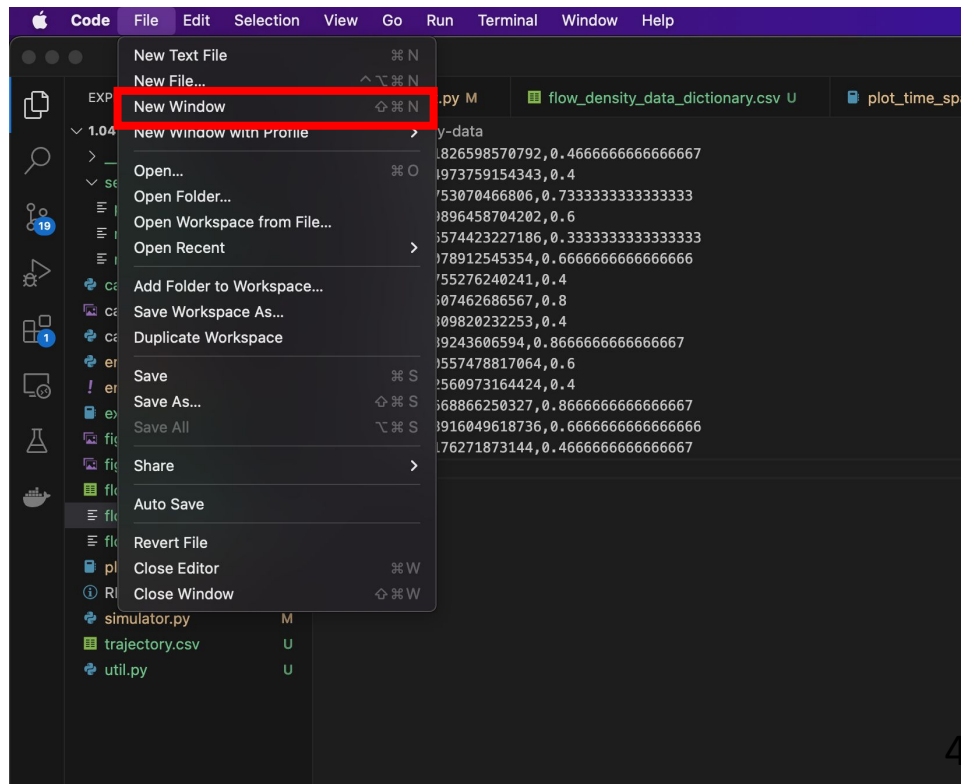


4. Add files to a new workspace for the repo (1/2)

- Open up the repo in Visual Studio Code

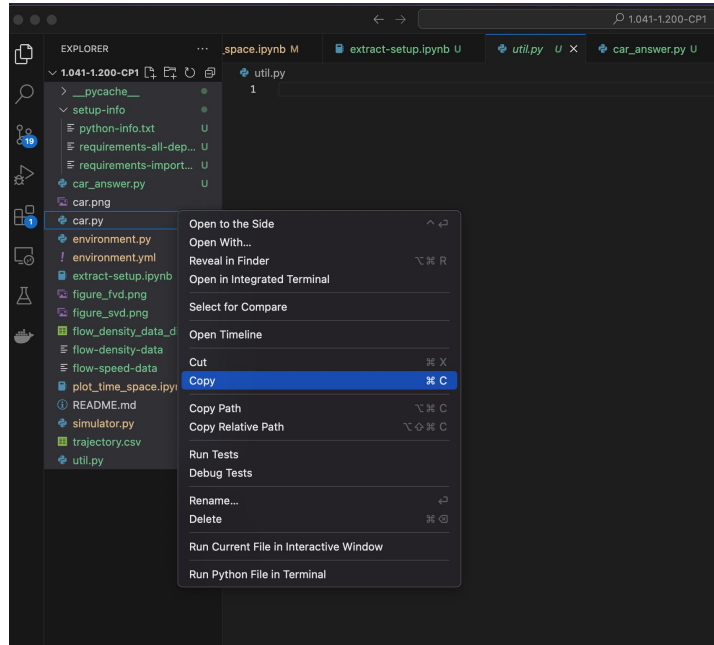
- New Window
- Open Folder...

- Navigate to where you put the repo

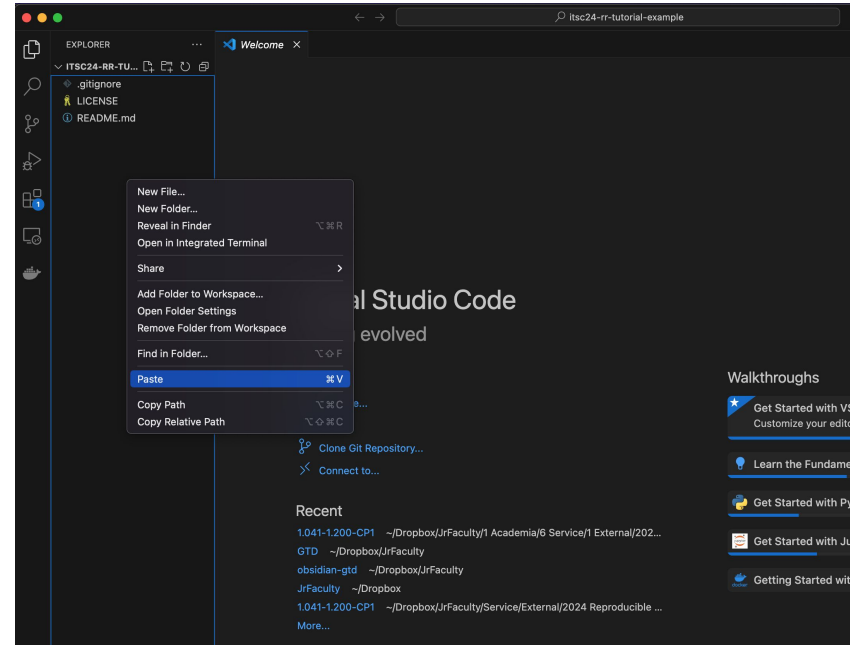


4. Add files to a new workspace for the repo (2/2)

- Copy all the files from the workspace you've been working in
- Paste them into the new workspace you just opened



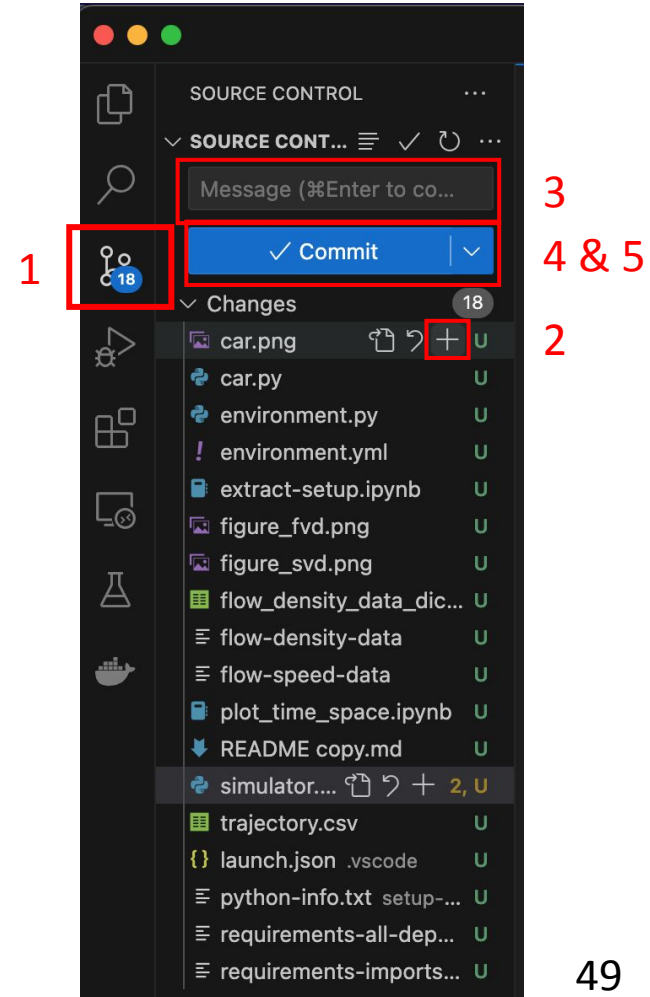
Old workspace



New workspace

5. Commit & sync files to Github

- Commit = designate the files you want to add/change to the repo (i.e., upload to Github)
- First, make sure the code in the new folder runs as expected
- Then, add the files you want to commit to your repo by:
 1. Navigate to the source control tab (left side panel)
 2. Click the “+” on the corresponding files
- Commit and sync!
 3. Add a commit message describing what you are uploading, like “A fully working car following simulation example.”
 4. Click Commit
 5. Click Sync Changes
 6. VSCode will prompt a window and ask you to sign into your Github account. Once signed, the commit will be synced.
- Check that your changes now appear in your repo on Github!
 - Refresh the page



5. Commit & sync files to Github (terminal)

- If you prefer, you can do all this in the terminal
- `$ git status`
 - See which files have changed
- `$ git add <file>`
 - Add the files you wish to commit
- `$ git status`
 - See which files you have staged for committing, remaining files which have changed
- `$ git commit -m "<commit message>"`
 - Commit the changes with a commit message describing what you are uploading
- `$ git pull`
 - In this case, this won't do anything, but it's always good practice to pull in changes to the repo (that your collaborators may have made) before pushing your new changes
- `$ git push`
 - This pushes your commit to Github!

Show time!

- Double check that the project works and reproduces the plots
 - Exit conda environment: `$ conda deactivate`
 - Clone a fresh copy of the project from Github into a new directory
 - Navigate to where you want to put the new directory. You want to explicitly give the new directory name so that it does not conflict with the repo you original cloned
 - `$ git clone git@github.com:... <new directory name>`
 - Follow the README and confirm that it works
 - The data and plots should look **exactly** the same as in the Github repo!
- Show us that it works
- **Bonus hard mode:** Ask someone else to confirm that it works.

Optional tips

- Your project probably won't work the first time you try to run your own instructions. Here are some suggestions on what to do when it doesn't work.
- If you conda environment creation fails, you'll need to delete the partially created conda environment before trying again
 - Command to delete conda env: `$ rm -rf /opt/anaconda3/envs/RR`

Outline

1. Make it run
2. Make it available
3. **Make it pretty**
 - a. Organizing data and code
 - b. Making your code readable

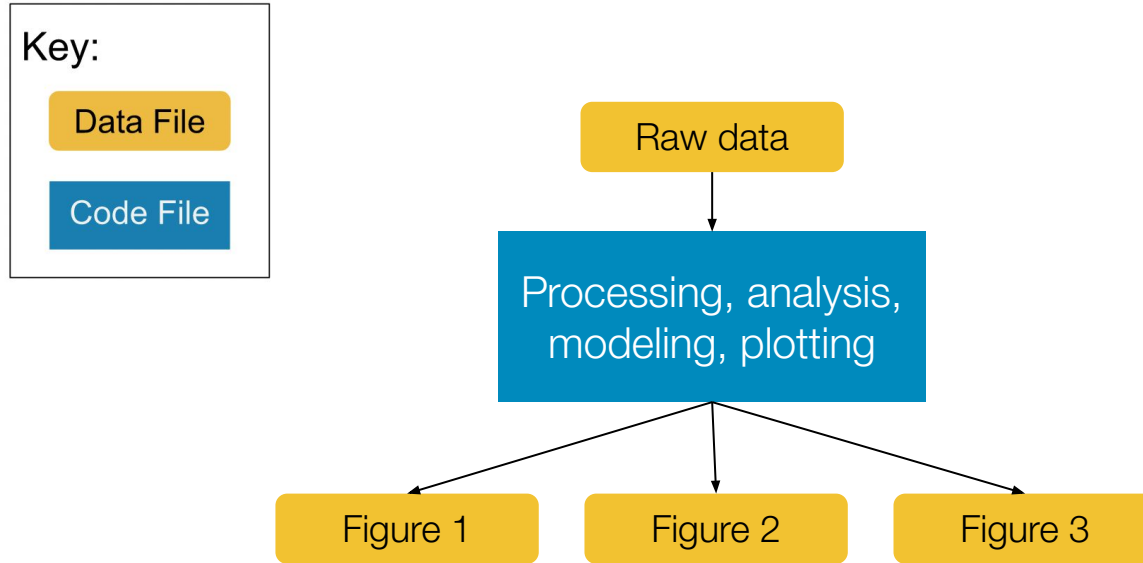
Outline

1. Make it run
2. Make it available
3. Make it pretty
 - a. Organizing data and code
 - b. Making your code readable

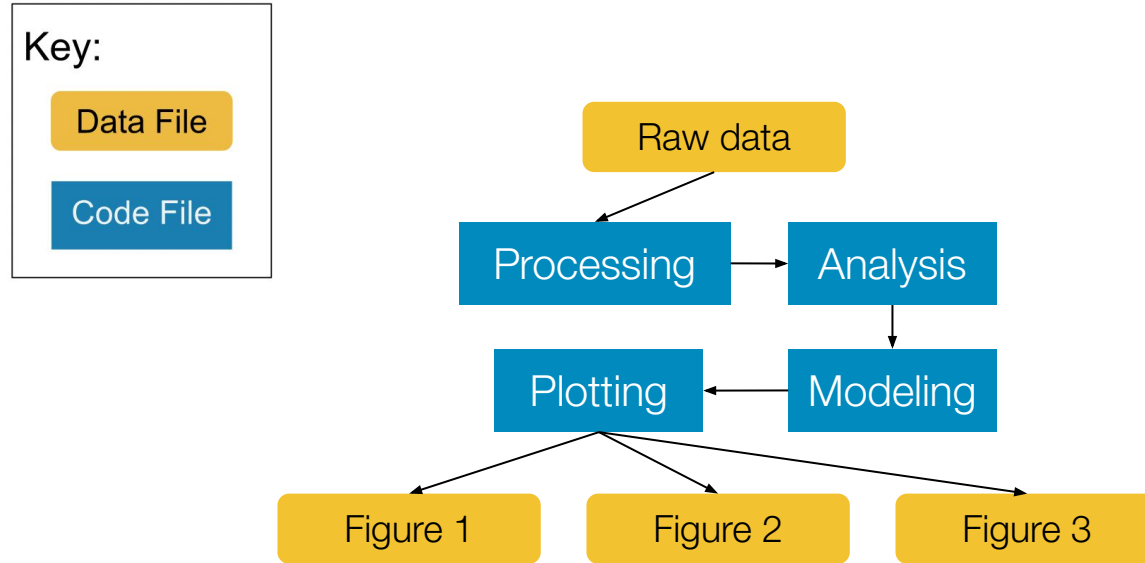
Organizing data and code

The best time to figure out how you want your project to be organized is before you begin. The second best time is now!

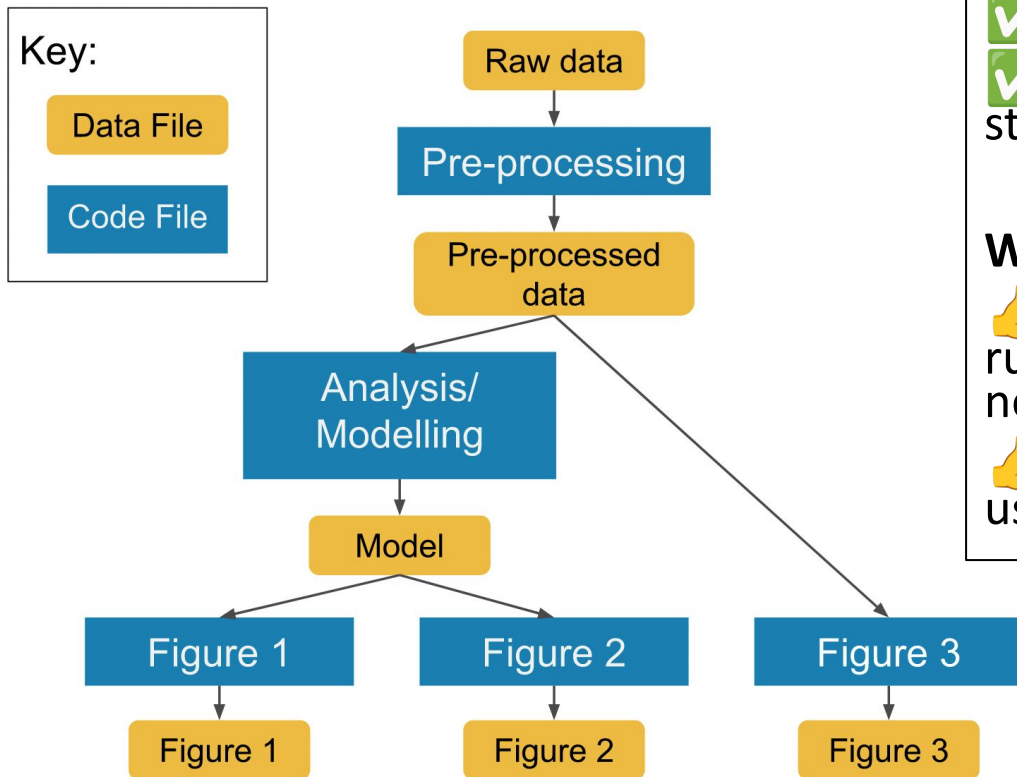
Organizing code and data



Organizing code and data



Organizing code and data



Organization principles

- ✓ As modular as possible
- ✓ Save the output of every step

Why?

- 👍 Research productivity: avoid running the full pipeline unless necessary
- 👍 Ease reproducibility: can use intermediate outputs

Exercise: Organize the ITS code example into directories

Organization takes time.

Be prepared to refactor as the project evolves.
The sooner the better, to minimize “tech debt.”

Car following simulation

■ Discussion:

- (a) Is it easy to understand what everything is?
- (b) Did you document the process of data preparation and analysis?
- (c) How might you re-organize things, to make it more clear to someone else, or to you a year from now?
- (d) Do you need more documentation (e.g., README files) or are the file names and organization self-explanatory?

Suggestions for organizing code and data

- Modularize the codebase by splitting the simulation code from the plotting code.
- Provide an example of how to plot the figures in a Jupyter Notebook, separate from the python scripts.
- Create separate directories for assets, data, figures.
- Create a config.py for shared parameters.

Outline

1. Make it run
2. Make it available
3. Make it pretty
 - a. Organizing data and code
 - b. Making your code readable

What does this code do?

```
import numpy as np
M, s1 = 0, 0.1
S2 = np.random.normal(M, s1, 1000)
import matplotlib.pyplot as plt; x, Y, zZ = plt.hist(S2,
30, density=True); plt.plot(Y, 1/(s1 * np.sqrt(2 *
np.pi)) * np.exp( - (Y - M)**2 / (2 * s1**2) ),
linewidth=2, color='r'); plt.show()
```

Python

What does this code do?

```
import numpy as np
import matplotlib.pyplot as plt

# generate random normal distribution
mu, sigma = 0, 0.1 # mean and standard deviation
sample = np.random.normal(mu, sigma, 1000)

# plot histogram with density curve
count, bins, ignored = plt.hist(sample, 30, density=True)
plt.plot(bins, 1/(sigma * np.sqrt(2 * np.pi)) *
         np.exp( - (bins - mu)**2 / (2 * sigma**2) ),
         linewidth=2, color='r')
plt.show()
```

Python

Make it pretty: making your code readable

Principles for making your code readable

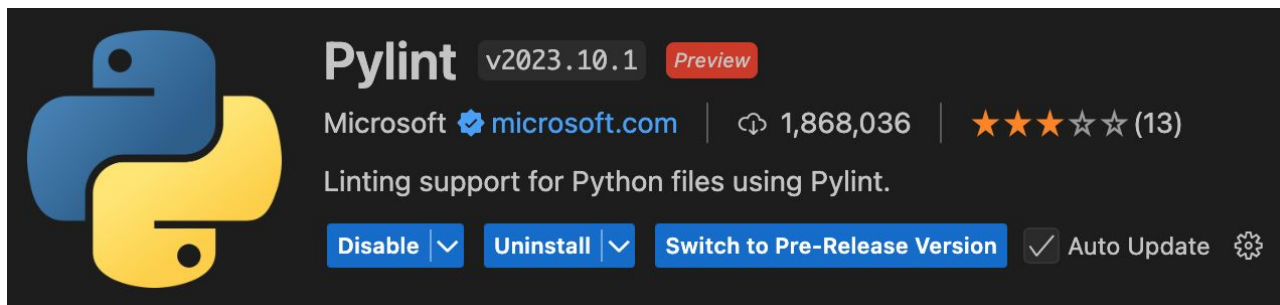
- Help readers follow the logic
 - Logical and human readable variable names
 - Modular classes and functions
 - Comment your code! Explain your thought process & design choices
- Ease readability by following stylistic conventions
 - Put all package imports at the top of the file
 - Use white space for readability
 - Break up long lines (keep lines <120 characters)
 - Consistent patterns of capitalization (e.g., ALL CAPS for constants, `lower_case_with_underscores` for functions and variables, `lower-case-with-dashes` for directories)

Style guides

- Style guides for academic writing:
 - MLA, APA or Chicago/Turabian
- Style guides for code (Python)
 - [PEP 8](#), [Google style guide](#)
- **[Linter](#)**: Tools that analyze source code to flag programming errors, bugs, stylistic errors, and suspicious constructs.
 - Like a grammar checker but for code
 - Ex. [Pylint](#) for Python



Pylint + Visual Studio Code demo

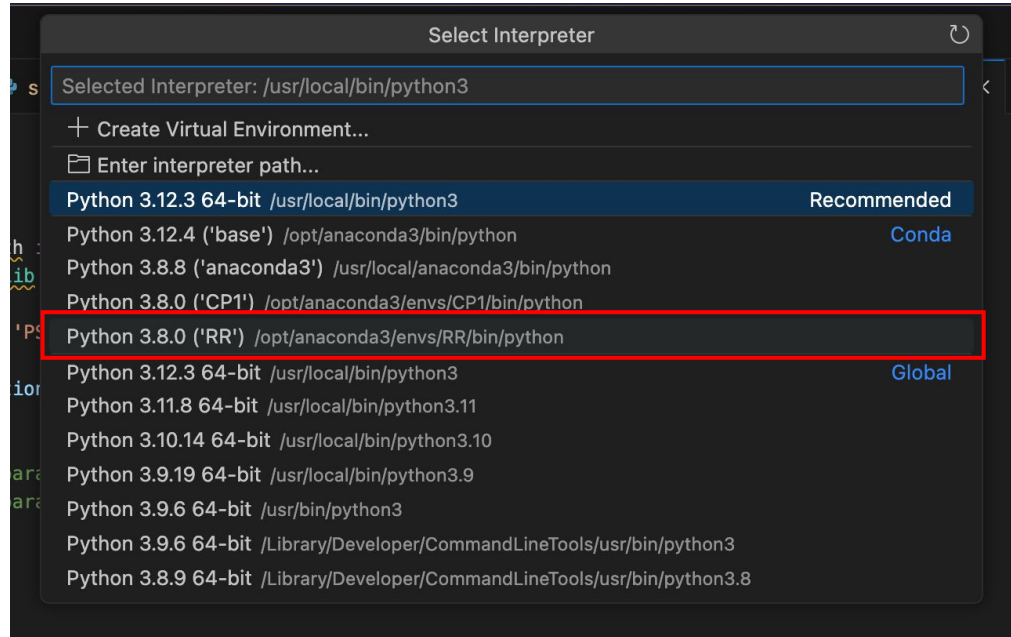
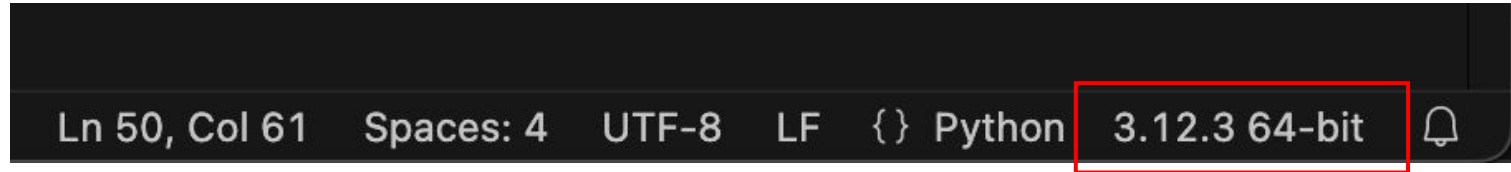


Adding your first VS Code extension

The screenshot shows the VS Code interface with the Extensions Marketplace open. The search bar at the top of the marketplace contains the text "pylint". The "Pylint" extension by Microsoft is highlighted in the list on the left. The extension details for "Pylint" are shown on the right, including the version "v2023.10.1" and the "Install" button. Red boxes and arrows highlight the following steps:

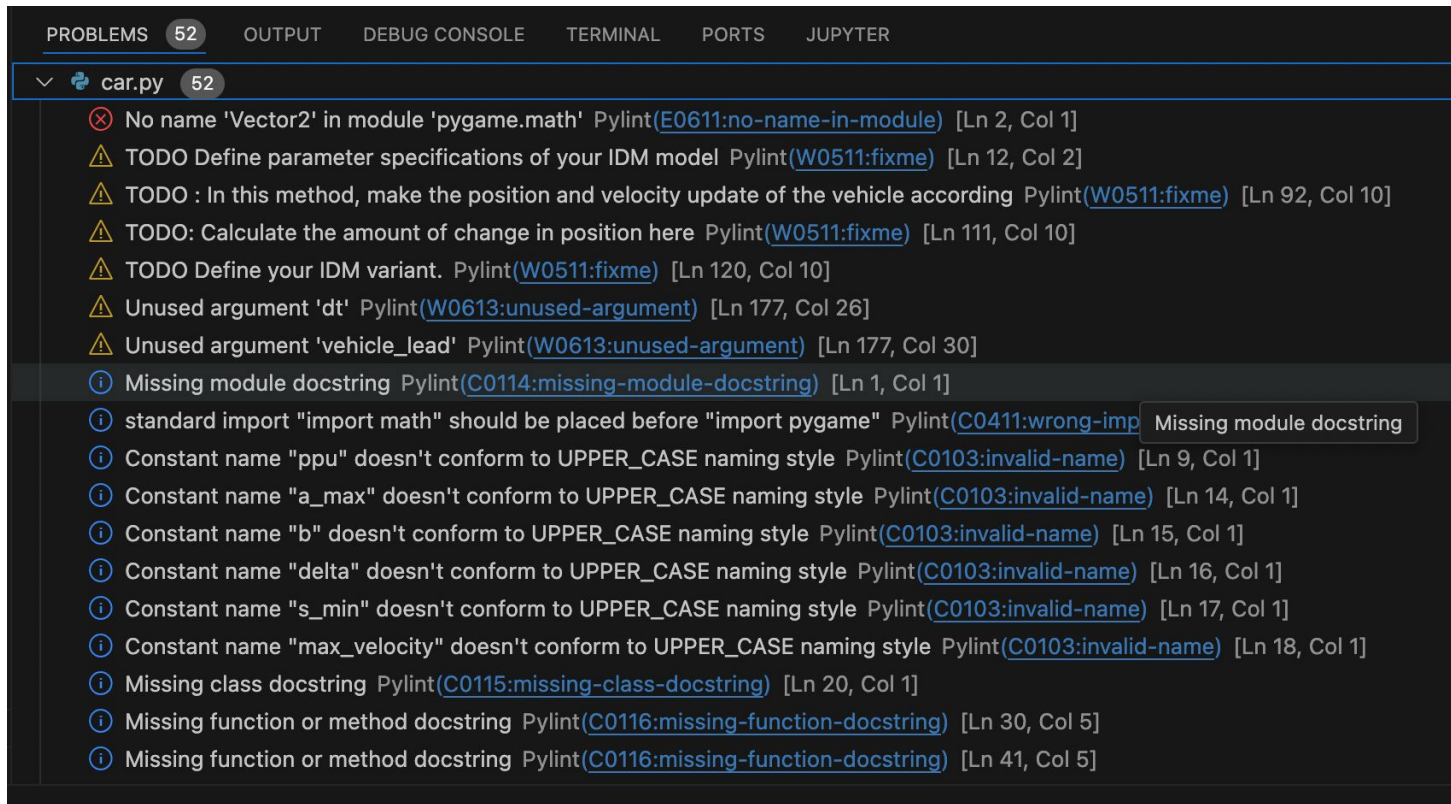
- 1) Extensions tab: A red box highlights the Extensions icon in the left sidebar.
- 2) Find extension: A red box highlights the search bar containing "pylint", with an arrow pointing to the "Pylint" extension in the list.
- 3) Install extension: A red box highlights the "Install" button for the "Pylint" extension, with an arrow pointing to it.

Setting up pylint: select the python interpreter



Select python interpreter

Using pylint: Problems pane

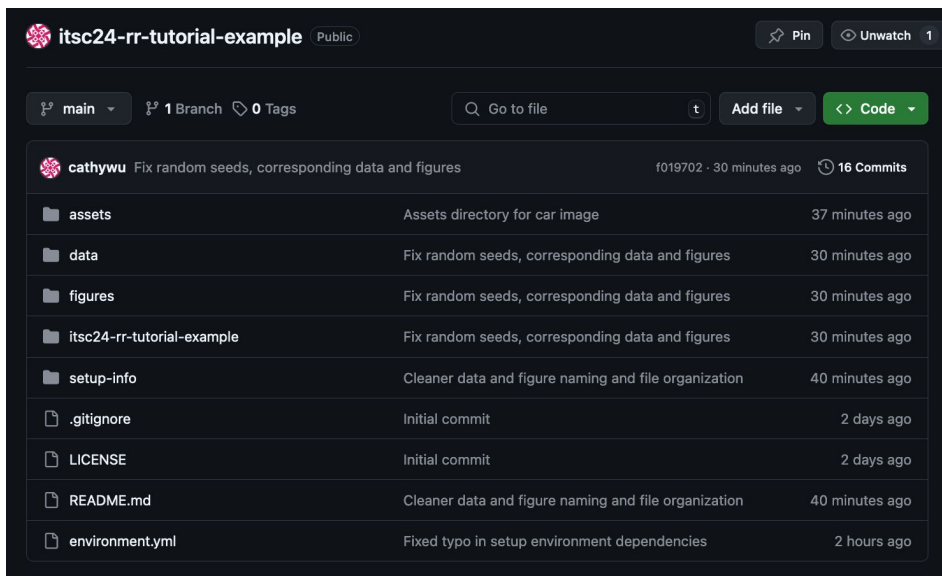


Show time!

- Commit your changes (same as before)
- Show us that it works (same as before)
- Exchange projects with someone and ask them to critique yours.
What organization could help someone more easily build on the project?
- Remember, that someone could be you in 6 months!

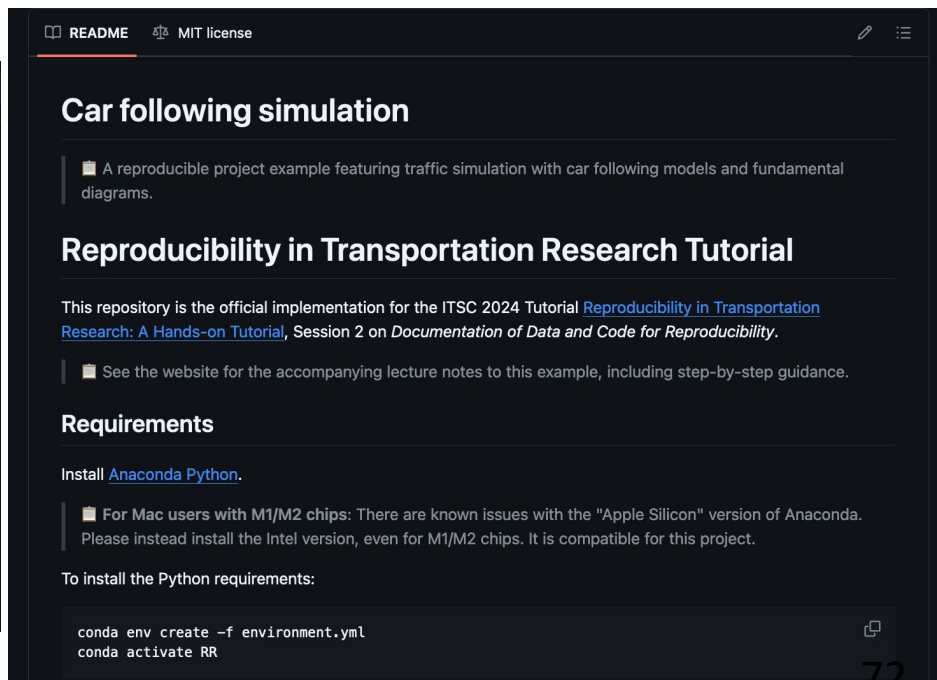
Reference implementation (try it yourself first!)

- See <https://github.com/cathywu/itsc24-rr-tutorial-example> for a reference implementation of this session's hands on activity.



The screenshot shows the GitHub repository page for 'itsc24-rr-tutorial-example'. The repository is public and has 1 branch and 0 tags. The file list includes:

File	Description	Time
assets	Assets directory for car image	37 minutes ago
data	Fix random seeds, corresponding data and figures	30 minutes ago
figures	Fix random seeds, corresponding data and figures	30 minutes ago
itsc24-rr-tutorial-example	Fix random seeds, corresponding data and figures	30 minutes ago
setup-info	Cleaner data and figure naming and file organization	40 minutes ago
.gitignore	Initial commit	2 days ago
LICENSE	Initial commit	2 days ago
README.md	Cleaner data and figure naming and file organization	40 minutes ago
environment.yml	Fixed typo in setup environment dependencies	2 hours ago



The screenshot shows the README file for the 'itsc24-rr-tutorial-example' repository. The README is titled 'Car following simulation' and describes a reproducible project example featuring traffic simulation with car following models and fundamental diagrams. It also mentions the repository is the official implementation for the ITSC 2024 Tutorial 'Reproducibility in Transportation Research: A Hands-on Tutorial', Session 2 on 'Documentation of Data and Code for Reproducibility'. The README includes a section for 'Requirements' and instructions on how to install Anaconda Python. At the bottom, it provides the commands to create and activate the environment:

```
conda env create -f environment.yml
conda activate RR
```


Acknowledgements

- Main contributor:
 - Cathy Wu
- Contributing content
 - Vindula Jayawardana, Jung-Hoon Cho
- Testers
 - Cathy Wu's research group, Xiaoyi Wu, Kate Sanborn
- Brainstorming
 - Bidisha Ghosh, Irene Martínez, Nicholas Saunier
- Support
 - [RERITE working group](#)
- CPS-IoT Week
 - For giving us an opportunity to broadcast the efforts

