UNIVERSITY OF TWENTE.

CENTER OF EXPERTISE IN BIG GEODATA SCIENCE

Introduction to JupyterLab

dr. ing. Serkan Girgin MSc s.girgin@utwente.nl

Training Content

- Interactive Notebooks
- Project Jupyter
- JupyterLab
 - Installation
 - Kernels
 - Code Consoles
 - Text Editor
 - Terminals
 - Notebooks
 - User interface
 - Workspaces
 - File management
 - Debugging
 - Real-time collaboration
 - User settings
 - Extensions

Center of Expertise in Big Geodata Science

- Center of Expertise in Big Geodata Science (CRIB) is a horizontal facility established in March 2020 to enable the better use of geospatial cloud computing and big data technologies in education, research, and institutional strengthening activities at ITC.
- Our mission is to collect, develop, and share operational know-how on cloud computing and big data technologies to solve large-scale geospatial problems.
- Our vision is to position UT/ITC as a globally renowned center of excellence in geospatial cloud computing and big data science.

https://itc.nl/big-geodata

Interactive Notebooks

"The idea of a notebook is to have an interactive document that freely mixes code, results, graphics, text and everything else."

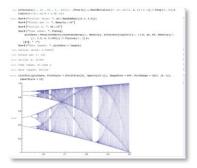
Stephen Wolfram

- Notebooks are traditionally used in science to document research procedures, data, calculations, and findings
- They track methodology to make it easier to reproduce results and calculations with different data sets
- Their utility for combining text with code makes them useful for education
- They can be used for interactive presentations
- Their ability to display data visually and retrieve data from different sources by modifying code makes them also useful for business intelligence



Interactive Notebooks

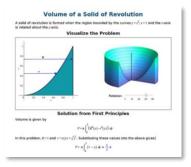
Mathematica





	$s_i = \text{if}\left(0 \le \varphi_i, s1_i, 0 \text{ m}\right)$	$s_i \mapsto if(\phi_r \le \varphi_i \cdot s_i(\phi_r) \cdot s_i^T \cdot s_i)$
	$s_{i} \!=\! \operatorname{if} \left(\phi_{i} \!+\! \phi_{d} \!\leq\! \varphi_{i} \!\cdot\! s \! 2_{i} \!\cdot\! s_{i} \! \right)$	$s_i \coloneqq \widehat{\mathrm{sf}} \left(\phi_i + \phi_d + \phi_f \leq \varphi_i, 0 \ \mathbf{m}, s_i \right)$
-	~ •	splacement, versus cam.angle
	$/ \setminus$	
•		
	Figure 4. Displacen	nenet profile throughout one rotation
	sure angle can be determined	$f(\varepsilon, s, v, R_0) = \operatorname{intain} \left(\frac{\varepsilon - v}{s + \sqrt[1]{N_0^2 - \varepsilon^2}} \right)$
	ng function.	$\left[s + \sqrt{R_0^2 - \sigma^2}\right]$
by following From equi	ng function: aity f(0, si, vi, ri) = ep nd the formula	$r_i = \frac{-w_i \cdot s}{\tan(\theta_i r)} - s_i$ $\left[s + \sqrt{R_0^2 - c^2}\right]$
by following from equilibrium one can fill the values.	aity 1(0, si, vi, n) = ep nd the formula	$r_{i} = \frac{-e_{i} \cdot a}{\tan(\theta_{i})} - \frac{1}{2}$ Throughout a given station to make sure the restriction is not exceed

Maple



Matlab Live Editor



Open-Source Projects

Jupyter

Apache Zeppelin

Apache Spark Notebook

SageMath

<u>Polynote</u>

Org Mode

Starboard

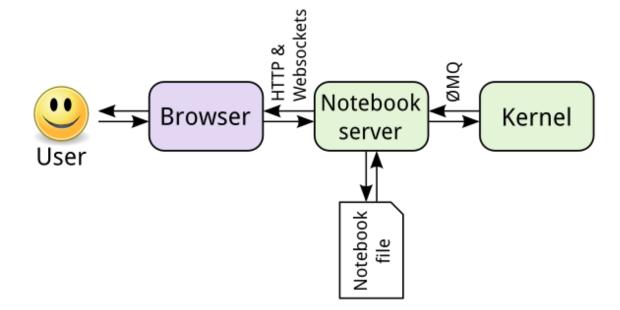
•••

IPython

- IPython (Interactive Python) is a command shell for interactive computing, first developed by <u>Fernando Perez</u> in 2001
 - An enhanced interactive Python shell (e.g. introspection, shell syntax, tab completion, history, etc.)
 - A decoupled two-process communication model, which allows for multiple clients to connect to a computation kernel
 - An architecture for interactive parallel computing (now part of <u>ipyparallel</u>)
- It supported additional interactive shells, a browser-based notebook interface, and interactive data visualization
- IPython 3.x was the last monolithic release and in 2014 the language-agnostic parts of the project have been moved to Project Jupyter

```
Python 3.6.3 | packaged
Type 'copyright', 'credi
IPython 7.0.0.dev -- An
In [1]: from numpy.fft i
   ...: from numpy impor
   \ldots: a = arange(32)
   \ldots: A = fft(a)
   \ldots: f = fftfreq(32)
In [2]: %matplotlib tk
In [3]: from matplotlib.
In [4]: stem(f, abs(A))
Out[4]: <Container objec</pre>
In [5]: _.
            add callback
            baseline
            count()
 500
 400
 300
 200
 100
       -0.4
           -0.2
                 0.0
                      0.7
```

Decoupled Two-process Communication



Project Jupyter

- Project Jupyter is a non-profit, open-source project that will always be 100% open-source software, free for all to use
- It is an ecosystem of a set of software applications and standards
 - <u>IPython</u> (interactive computing in Python)
 - Jupyter Server (main server application, server extensions, and related projects)
 - Jupyter Widgets (interactive widgets for the Jupyter Notebook)
 - <u>Jupyter Notebook</u> (Jupyter Interactive Notebook)
 - Jupyter Console (terminal-based console for interactive computing)
 - Jupyter QtConsole (Qt application for interactive computing with rich output)
 - JupterLab (web-based IDE for Jupyter notebooks, code, and data)
 - JupyterHub (Multi-user server for Jupyter notebooks)
 - <u>nbviewer</u> (web application to render notebooks as HTML pages)
 - <u>nbconvert</u> (converts notebooks to other formats)
 - and more...!

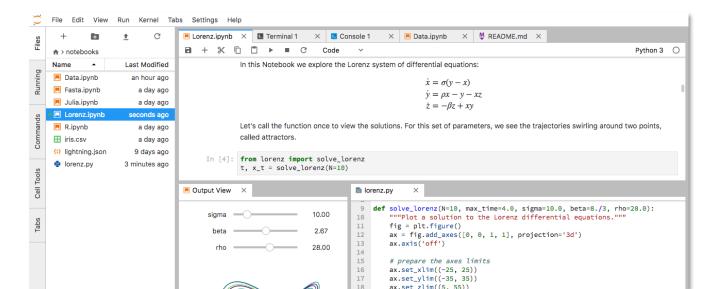


Open Standards for Interactive Computing

- <u>Notebook Document Format</u>
 - Open document format based in JSON to contain user sessions including code, narrative text, equations, and rich output
- Interactive Computing Protocol
 - Open network protocol to communicate with computational kernels based on JSON data over ZMQ and WebSockets.
- <u>Kernel API</u>
 - Open specifications and API for developing custom kernels

JupyterLab

- JupyterLab is the new web-based user interface for Project Jupyter
- It enables to work with documents and activities (e.g. notebooks, editors, terminals, and other tools in a flexible, integrated, and extensible way
- It can be run on a local desktop requiring no internet access or can be installed on a remote server and accessed through the Internet



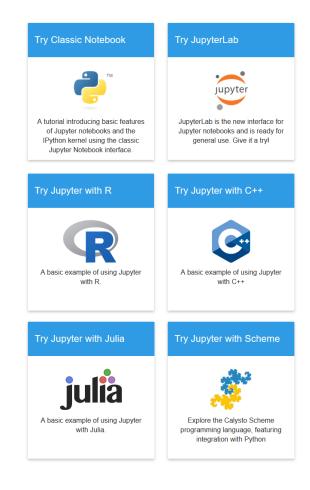
JupyterLab Release History

Initial Release	Release Data	Latest Release	Release Date
4.0.0.a0	27/07/2021	4.0.0.a7	01/09/2021
3.0.0	19/12/2020	3.1.10	01/09/2021
2.0.0	28/02/2020	2.3.2	05/08/2021
1.0.0	28/06/2019	1.2.21	05/08/2021
0.0.1	15/12/2015	0.35.6	05/05/2019

- Source code available at https://github.com/jupyterlab/jupyterlab
- 393 contributors, 23 442 releases, 1 865 open issues
- Major version upgrades were challenging for the community (e.g. some extensions are still not compatible with 3.x)
- Some extensions are still not compatible with JupyterLab (e.g. <u>nbgrader</u>)

Try Without Installation

- <u>Try Jupyter</u>
- JupyterLite
- <u>Google Colaboratory</u>*
- <u>Azure Machine Learning Workspace</u>*
- <u>Amazon SageMaker</u>*
- Microsoft Planetary Computer*
- <u>UT Computing Platform</u>* (not public, i.e. through VPN)
- ITC Geospatial Computing Platform*



ITC Geospatial Computing Platform





Scalable

purposes.

Powered by Dask and Apache Spark

Interactive

Interactive Jupyter notebooks provide easy to use and user friendly data analysis and visualization environment.





Containerized

Each working environment is containerized and isolated from each other and also the host unit to ered by Docke



Ready to Use

Software packages are ready to use out-ot-the-box, without any further setup required.



Distributed

Built on a cluster of computing

units, the platform scales

automatically and balances

workload among the units.

Powered by Docker Swarm





Multilingual All units in the platform and their Multiple programming languages GPUs are available for distributed can be used to access platform out-of-core geospatial data analysis resources and perform geospatial computations.

Replicated

Your assets are protected against

hardware failures through replicated storage with minimum

two copies.

Optimized

libraries.

ed by GlusterES and ZES

n python

iulia

ava

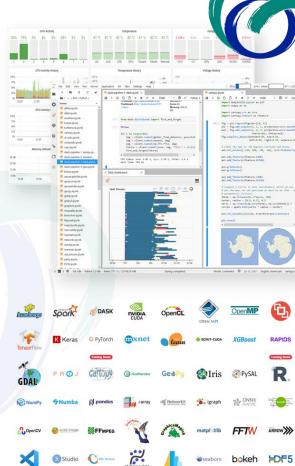
Powerful Each computing unit has 8-core ARM v8.2a 64-bit CPU, 512-core Volta GPU with Tensor Cores, and 32GB 256-bit LPDDR4 RAM.

d by Jetson AGX Xavie



Energy Efficient Computing units operate at

Software packages are fine-tuned for best performance utilizing high-10-30W ensuring low energy footprint, albeit high performance. performance, multi-threaded



https://crib.utwente.nl

Installation

- JupyterLab can be installed using <u>conda</u>, <u>mamba</u>, <u>pip</u>, <u>pipenv</u>, and <u>Docker</u>
- Basic steps:
 - Install Python
 - Download and run installer
 - Install node.js
 - **Download** and run installer
 - Install JupyterLab
 - <u>Follow installation notes</u>
 - e.g:pip install jupyterlab
- Detailed installation steps for Windows are available at CRIB website: <u>https://crib.utwente.nl/training/jupyterlab-installation.pdf</u>

Kernels

- Kernels are processes that run interactive code in a (programming) language and return output
- Kernels also respond to tab completion and introspection requests
- Kernels run until they are terminated (i.e. shutdown)
 - Closing a document or activitiy does not terminate related running kernel
- Core kernels
 - IPython
 - IRKernel
 - <u>Ijulia</u>

IP[y]:

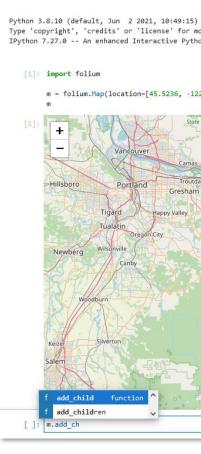




- More than 150 community maintained kernels
 - Some languages have multiple kernels
 - Not all of them are functional (e.g. prototype, not complete)
 - Not all of them have full kernel functionality (e.g. code completion)
 - Some of them are outdated

Code Consoles

- Code consoles allow to run code interactively in a kernel
- The cells of a code console show the order in which code was executed in the kernel
- Code consoles can display rich output
- Code consoles support code completion and tooltips
- Code consoles can be connected to running kernels
 - Allows interactively inspect and run code without disturbing other connected clients (e.g. notebooks)
 - Acts as a log of computations in the kernel
- If you close a console it continues to run in the background
 - You can access to the same console again
 - But you lose the console content



Text Editor

- The editor allows editing text files directly and supports:
 - Language-specific syntax highlighting
 - Theming
 - Key maps (e.g. vim, emacs, Sublime Text)
 - Configurable indentation
 - Auto close brackets
- Additional capabilities can be added by extensions (e.g. code completion, spellchecking)
- A file can be open simultaneously in multiple editors and viewers and they will remain in sync
 - Views may not be updated if file content is invalid
- Text files can be connected to code consoles or kernels
 - Code from a text file can be run interactively easily
 - Code blocks are automatically detected and run entirely

B lo	renz.py	×
	iciiz.py	
6	def solve_	lorenz(sigma=10.0, beta=8
7	"""Plo	t a solution to the Lorer
8		
9	max_ti	me = 4.0
10	N = 30	
11		
12	fig =	plt.figure()
13	ax = f	ig.add_axes([0, 0, 1, 1],
14	ax.axi	s('off')
15		
16		are the axes limits
17		_xlim((-25, 25))
18	ax.set	_ylim((-35, 35))
19	ax.set	_zlim((5, 55))
20		
21		<pre>renz_deriv(x_y_z, t0, sig</pre>
22		"Compute the time-derivat
23	×,	y, $z = x_y_z$
24	re	turn [sigma ★ (y - x), x
25		
26		se random starting points
27		dom.seed(1)
28	x0 = -	15 + 30 * np.random.rando
29		
30		e for the trajectories
31		.linspace(0, max_time, ir
32	x_t =	np.asarray([integrate.ode
33		for x0i in x0
34		
35		se a different color for
36	colors	<pre>= plt.cm.viridis(np.lins</pre>
37		
38		in range(N):
39	х,	y, z = x_t[i,:,:].T
40		<pre>nes = ax.plot(x, y, z, '-</pre>
41		t.setp(lines, linewidth=2
42	angle	= 104

Terminals

- Jupyter provides terminals for the system shells (e.g. bash, zsh on Linux/Mac and PowerShell on Windows)
- Any text-based program can be run by using a terminal
- You can copy-paste content from/to a terminal
 - Check the documentation for <u>OS-specific details</u>
- Terminals run with the privileges of your user on the system where Jupyter server is running
 - You may have limited privileges on remote systems
- If you close a terminal it continues to run in the background
 - You can access to the same terminal again
 - But you lose the screen content



Notebooks

- Jupyter notebooks are documents combining live runnable code with narrative text, equations, images, interactive visuals and other output
- Notebooks are composed of multiple cells with different types (i.e. code, narrative, and raw) and related outputs
- Each notebook is connected to a kernel (e.g. Python, R, Julia) that run code cells and return their outputs
- Cells can be run in any order, but the execution order is displayed explicitly

In [1]:	<pre>import matplotlib.pypl import pandas as pd pdversion</pre>
011111	'0 24 1'

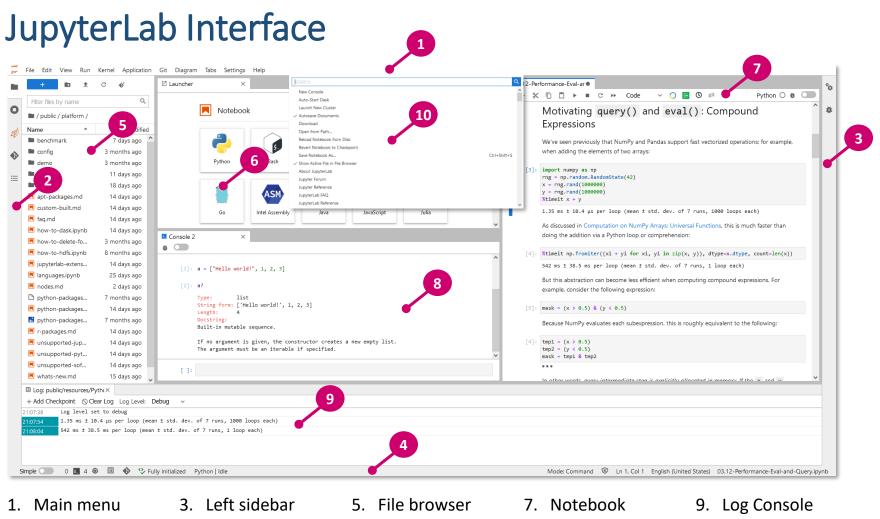
Dataset: Stanford

In [2]:		ri stand: = pd.rea		
In [3]:		what does .head()	s each ro	ow r
Out[3]:				
		stop_date	stop_time	cou
	0	2005-01- 02	01:55	
	1	2005-01- 18	08:15	
	2	2005-01- 23	23:15	
	3	2005-02- 20	17:15	
		2005 02		

Output Formats

- JupyterLab provides a unified architecture for viewing and editing data in different formats
 - For files, the data format is detected by the extension of the file
 - For outputs, the data format is set by Jupyter APIs using MIME types
- Formats supported by default
 - Markdown (.md)
 - lmages (.bmp .gif .jpeg .png .svg)
 - <u>Delimiter-separated values</u> (.csv)
 - A Maximum size depends on the web browser (e.g., 1.04GB for Firefox)
 - <u>JSON</u> (.json)
 - HTML(.html)
 - LaTeX* (.tex)
 - PDF*(.pdf)
 - Vega/Vega-Lite (.vl .vg.json)
 - <u>Virtual DOM</u> (.vdom .json)
- Extensions can add support for other formats

🖹 table.d	csv	×	
1	5,	7, 1, 0	9,4
2	4,	6, 5, 3	3, 3
3		7, 4, 0	9,4
4		10, 5,	
5			2,2
6			4, 5
7			1, 1
8			э, ө
9			4, 4
10			4,0
11 12			9,2
12			2,4
13			2, 2
15		6, 0, 5	5, 5
16			4,0
17			3,0
⊞ table. Delimiter	-	×	
5	7	1	0
4	6	5	3
4	7	4	0
4	10	5	5
5	6	5	2
4	9	1	4
5	7	2	1
2	9	3	0
0	9	5	4
1	8	0	4
1	9	3	0
3	6	1	2
2	10	4	2
0	8	2	2
1	6	0	5
4	8	1	4
-			
0	9	4	3



- 2. Right sidebar
- 4. Status bar
- 6. Launcher
- 8. Code Console

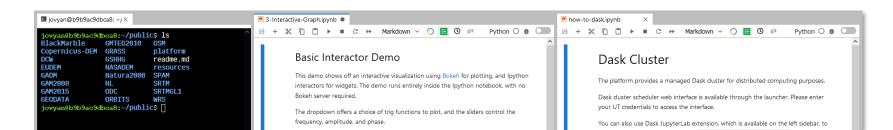
10. Command Palette

JupyterLab Interface

- Menu bar includes File, Edit, View, Run, Kernel, Tabs, Settings, Help menus
 - Extensions may create additional menus
- Left sidebar contains common tasks (e.g. file browser, list of running kernels, etc.)
 - Extensions may create additional sidebar tabs
- Main work area allows to arrange documents and activities (e.g., terminals, code consoles) into panels that can be resized and subdivided
 - Simple mode toggles single- and multi-activity layouts
- Context menu can be accessed by right mouse click
 - Native context menu can be accessed by Shift + right mouse click
- Keyboard shortcuts can be used to perform actions
 - Command palette can be accessed by Ctrl + Shift + C
- User interface language can be set by the user (new in JupyterLab 3.0)
 - Language pack of the language should be installed to display a language

Workspaces

- Each JupyterLab session resides in a workspace that contains its state (e.g. open files, layout of the open tabs, etc.)
- Terminals, code consoles, text editors, and notebooks can be layout
- Workspaces can be saved and restored another time
 - Default workspace is saved and restored automatically
- Workspaces can be shared between multiple users or browsers as long as they have access to the same server
- A workspace should only be open in a single browser tab at a time
 - Simultaneous use results in a new workspace assignment



File Management

- File browser allows opening, creating, deleting, copying, renaming, downloading, and sharing files and directories
- The directory you start JupyterLab matters; it becomes the root folder
- Directory listing is updated automatically if changes occur
- Ease of file management depends on deployment type
 - Local JupyterLab: easy file handling, difficult file sharing
 - Remote JupyterLab: difficult file handling, easy file sharing
 - Upload and downloads are performed through the browser
 - Some extension may facilitate file management
- When you create a file it is stored by using a default name
- Checkpoints are created for notebook files
 - <u>M</u> Hidden checkpoint directories may pollute your system

	+ 🗈 🛨	C		
Fi	lter files by name		Q	
	/			
Na	me	-	Last Modifie	d
	shared		10 days ago	^
	public		10 days ago	
	private		10 days ago	
	venv		6 days ago	
	Desktop		6 days ago	
	R		6 days ago	
	Downloads		6 days ago	
	Templates		6 days ago	
	Public		6 days ago	
	Documents		6 days ago	
	Music		6 days ago	
	Pictures		6 days ago	
	Videos		6 days ago	
	readme.md		10 days ago	
ß	local		10 days ago	
	Untitled.ipynb		10 days ago	
.	gcp-services.png		10 days ago	
⊞	untitled.csv		9 days ago	~

Sharing Notebooks

Markdown

- Notebook files are <u>JSON</u> files saved with the .ipynb extension
- They can be displayed or run using Jupyter or other compatible software
 - Not all output might be displayed properly (e.g. interactive widgets)
 - A Check your notebook before sharing by opening it in another session
 - Dependencies (i.e. packages, extensions) should be available to run them
 - External assests (e.g. files, datasets) should be shared separately
- They can be exported into other file formats

.md

- <u>AsciiDoc</u> .asciidoc
 <u>HTML</u> .html
 LaTeX .tex
 <u>PDF</u> .pdf
 <u>ReStructured Text</u> .rst
 Executable Script .py
 - <u>Reveal.js Slides</u>

.html

- Exporting options depend on <u>nbconvert</u> configuration
 - Some formats may require additional software packages
 - Some formats may allow customization by using cell metadata

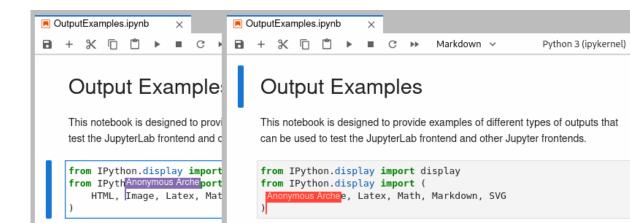
Debugger (new in JupyterLab 3.0*)

- Notebooks, code consoles, and files can be <u>debugged</u> <u>directly</u> from JupyterLab
- A kernel with support for debugging is required for the debugger to be enabled and visible
- Kernels should implement <u>Jupyter Debugger Protocol</u> to support debugging
- Kernels with debugging support
 - <u>xeus-python</u>
 - <u>xeus-robot</u>
 - ipykernel

VARI	ABLES						
► a	cals: : 1 : 2	2					
CALL	STACK	()⊳.		₹•	÷	†	
add							
BRE/	AKPOIN	ITS 🤅	Ð				
• /t	mp/xp	ython_	1866	0/211	4632	017.ру	
SOU	RCE {	🔒 /tm	ip/xpyt	hon_1	8660/2	114632017	рy
	RCE {	_		-	8660/2	114632017	.py
1 2		add(res	a, b = a): + b	8660/2	114632017	.py
1		add (a, b = a): + b	8660/2	114632017	.py

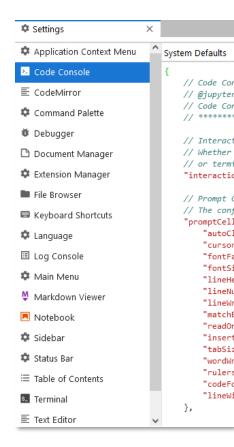
Real Time Collaboration (new in JupyterLab 3.1)

- Documents and notebooks have collaborative editing feature enabling collaboration in real-time between multiple clients **without user roles**
- Other users have access to the same environment you are working on
- Editors are not collaborative by default; to activate it, start JupyterLab with the --collaborative flag
- This feature is not well developed yet



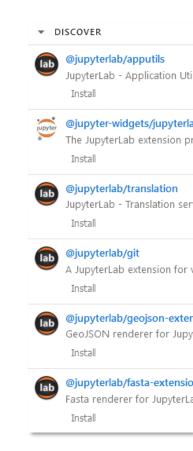
User Settings

- Some settings can be changed by using menu items (e.g. main theme, text editor indentation, text editor key map)
- All settings can be changed by using Advanced Settings Editor (under Settings menu)
 - System defaults (Read-only JSON with comments)
 - User preferences (JSON)
- Extensions may also have settings which can be modified by the editor
- Some changes may require restart (e.g. logout / login) (usually the case with extensions)
- See JupyterLab Documentation for more information on <u>user settings directory and files</u>



Extensions

- JupyterLab is designed as an extensible environment
- Extensions can customize or enhance any part of JupyterLab (e.g. new themes, file viewers, editors, or renderers for outputs)
- Types of extensions
 - Source extension (requires a rebuild of JupyterLab)
 - Prebuilt extension (does not require a rebuild of JupyterLab)
- Installation methods
 - Package managers (e.g. pip, conda)
 - Extension Manager in JupyterLab
 - jupyter labextension install command
 - Some extensions may require additional software packages
 - A Check extension documentation before installation
 - <u>A Restart JupyterLab after installation</u>



Popular Extensions

npm

Search

Name	Description
dask-labextension	JupyterLab extension for Dask
jupyter-archive	Jupyter/JupyterLab extension to create, download and extract archive files
jupyterlab-tensorboard	Tensorboard extension for JupyterLab
jupyterlab-chart-editor	JupyterLab extension for Plotly's react chart editor
jupyterlab-drawio	A standalone embedding of the drawio / mxgraph package into JupyterLab
jupyterlab-git	A Git extension for JupyterLab
jupyterlab-lsp	Coding assistance for JupyterLab using Language Server Protocol
jupyterlab-spellchecker	Spellchecker for JupyterLab notebook markdown cells and file editor
jupyterlab-spreadsheet-editor	JupyterLab spreadsheet editor for tabular data (e.g. csv, tsv)
jupyterlab-proxy-gui	A JupyterLab extension to control the configurable-http-proxy of JupyterHub

A Check <u>npm repository</u> for more by using the query "keywords:jupyterlab-extension"

Further Resources

- <u>https://jupyter.org/</u>
- Jupyter Project Documentation
- Jupyter Notebook Documentation
- JupyterLab Documentation
- IPython Documentation 🐵
- Jupyter Client Documentation 🐵
- <u>Awesome Jupyter</u> A curated list of awesome Jupyter projects, libraries and resources
- Jupyter Discourse

A Discourse Forum for a multitude of Jupyter topics

• Jupyter GitHub 🐵

A place where the community collaborates on the development of Jupyter software

• Contribute to Jupyter

It doesn't matter if you are new, Jupyter community welcomes you!



Contact CRIB



https://crib.utwente.nl



https://itc.nl/big-geodata



crib-itc@utwente.nl



<u>@BigGeodata</u>



Big Geodata Newsletter



Big Geodata Channel



Figurit. Regiones lineares Regiones lineares Regiones	UNIVERSITY OF TWENTE.	BIG GEODATA
<text></text>	Center of Expection in Big Geodeta Science	
<section-header>Contraction of the series of the series</section-header>		
In the second		
In the second	and a second second second	BIG GEODATA SCIENCE
 Accord and a constrained of co	Eserta	DIG GEODATA SOIENO
 invariant invariant	Links	
 mail <li< th=""><th></th><th></th></li<>		
 Find <li< td=""><td>y</td><td></td></li<>	y	
 Explor Explor	ල් Home	
 Institutions Institutions Institutions	# Explore	IN OWNER ON A DESCRIPTION OF A DESCRIPTI
Measaps Measaps </td <td></td> <td>ATTACT TO O</td>		ATTACT TO O
	Messages	
Intermediation of the second seco		
		@BigGeodata
<text><text><text><image/><section-header><section-header><section-header><section-header></section-header></section-header></section-header></section-header></text></text></text>		🐵 Enschede, the Netherlands 🔗 Itch/ibig-geodata 🖽 Joined June 2020
EVENENTINE Eventsion OF THE DESCRIPTION FOR SUBJECT THE ALL ADDRESS AND ADDRESS AN	- Prome	125 Fallowing 74 Followers
	applications of spatial phenological modelli computing, our new databases, and a ne	al data cubes, eScience Center-ITG collaboration on large-scale ing, recent developments in GPU-accelerated distributed web portal, a tool to extend spatial analysis capabilities of key-value w method to fili gaps in earth observation data implemented in the
	C. Dr. Castilla Service	a bond that more parameter busine the part bond
	No. of Concession, name	