4th Stage

Atmospheric Sciences Department

Numerical Weather Prediction

Laboratory

LAB 3 :NWP data Source and Data Types

- ERA5
- Example of NWP data from ECMWF
- Introduction to Jupyter Notebook
- NetCDF data Processing using Python Programming Language: Xarray Basics

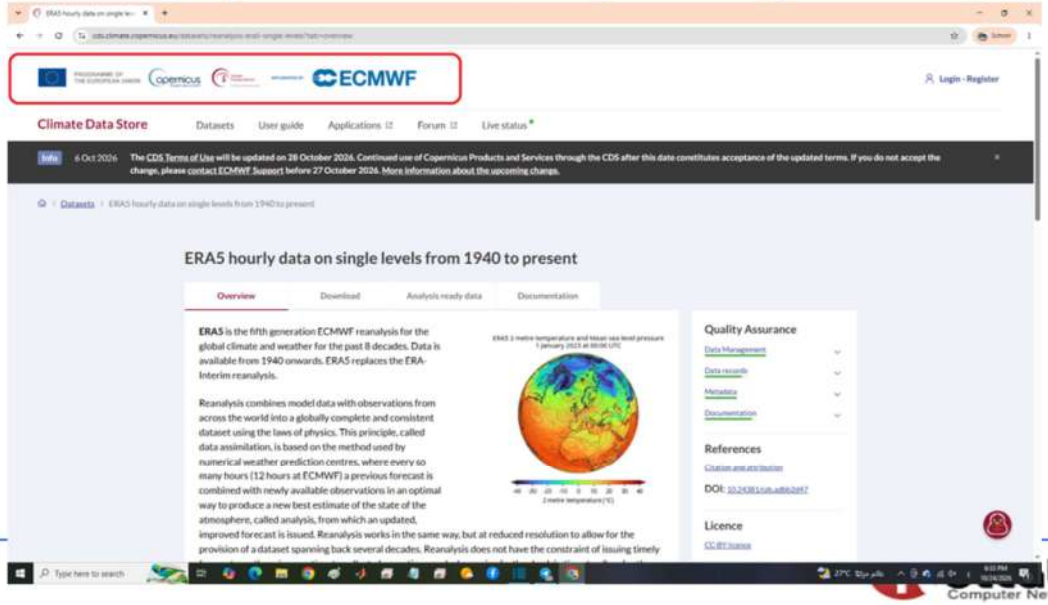
Prepared by

Lec. Firas Sabeeh Basheer

NetLab Computer Networks Research Lab

LAB 3 : Data Resources NetCDF

https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels?tab=overview



2

Computer Networks Research Lab

مجموعة المراكز البحثية العالمية المساهمة في بيانات التنبؤات العددية للطقس

LAB 3 : NetCDF Data Model

- ERA5:ERA5 is the **fifth-generation global reanalysis dataset** produced by the **European Centre for Medium-Range Weather Forecasts (ECMWF)** under the **Copernicus Climate Change Service (C3S)**.
- ERA5 provides hourly estimates of a large number of atmospheric, land and oceanic climate variables. The data cover the Earth on a 30km grid and resolve the atmosphere using 137 levels from the surface up to a height of 80km. ERA5 includes information about uncertainties for all variables at reduced spatial and temporal resolutions.
- Quality-assured monthly updates of ERA5 (1959 to present) are published within 3 months of real time. Preliminary daily updates of the dataset are available to users within 5 days of real time.

3

NetLab Computer Networks Research Lab

ERA5

ERA5هو **الجيل الخامس من إعادة التحليل العالمي** الذي ينتجه **المركز الأوروبي للتنبؤات الجوية متوسطة المدى (ECMWF)** تحت إشراف **خدمة تغيّر المناخ التابعة لك وبرنيكوس .(C3S)**

يوفّر **ERA5** تقديراتٍ كل ساعة لعددٍ كبير من متغيرات المناخ الجوية والبرية والبحرية . تغطي البيانات الكرة الأرضية بشبكة تبلغ دقتها 30كم، وتُمثّل الغلاف الجوي باستخدام 137مستوى من السطح حتى ارتفاع 80كم .يتضمن ERA5معلومات عن عدم اليقين لجميع المتغيرات ولكن بدقة مكانية وزمانية أقل.يتم نشر تحديثات شهرية مضمونة الجودة لبيانات) ERA5من عام 1959حتى الوقت الحاضر (خلال ثلاثة أشهر من الزمن الحقيقي .كما تتوفر للمستخدمين تحديثات يومية أولية لمجموعة البيانات خلال خمسة أيام من الزمن الحقيقي.

LAB 3 : NetCDF Data Model

- A preliminary ERA5 dataset from 1950 to 1978 is also available on the Climate data store (CDS) (1959-1978 is superseded by the quality assured dataset).
- ERA5 combines vast amounts of historical observations into global estimates using advanced modelling and data assimilation systems.
- ERA5 replaces the ERA-Interim reanalysis which stopped being produced on 31 August 2019.

4

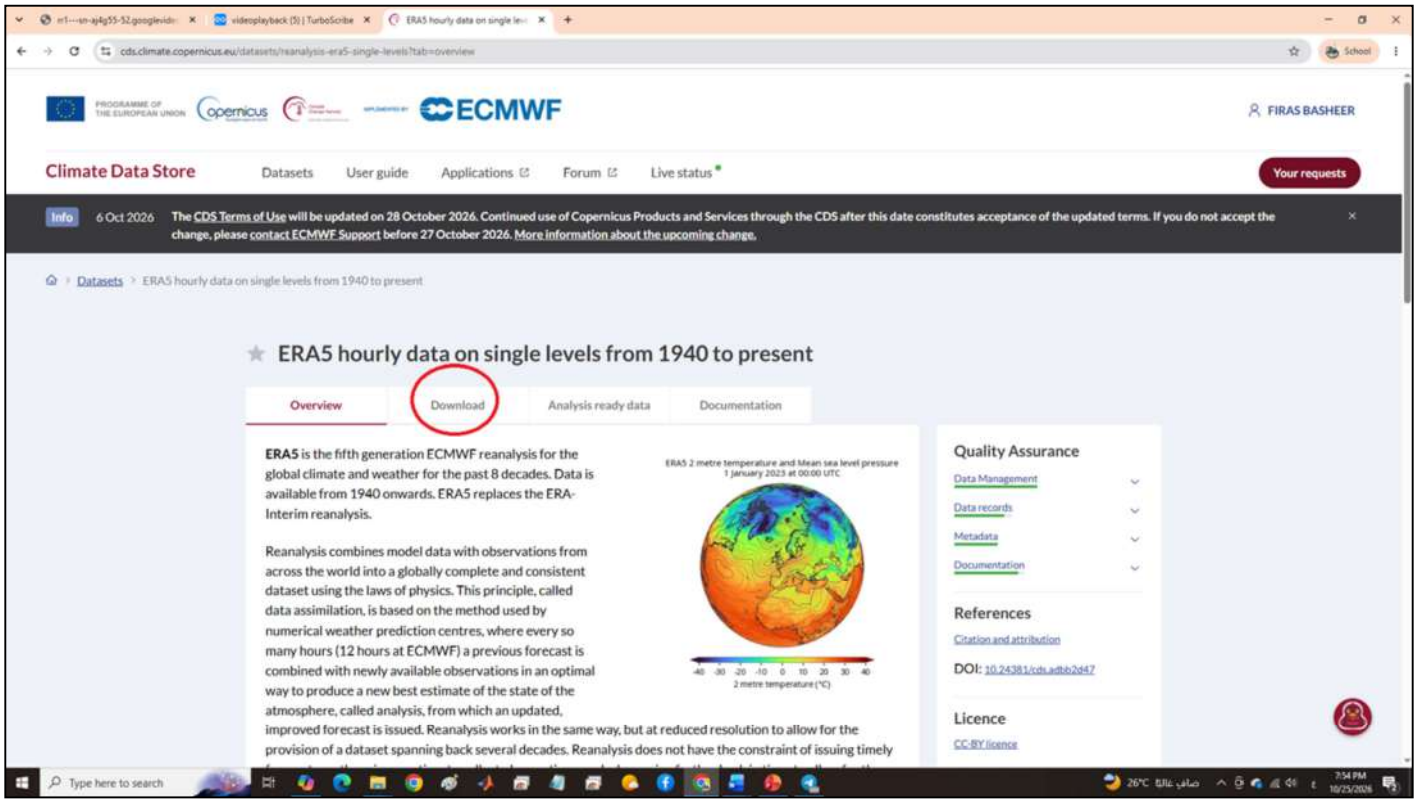
NetLab Computer Networks Research Lab

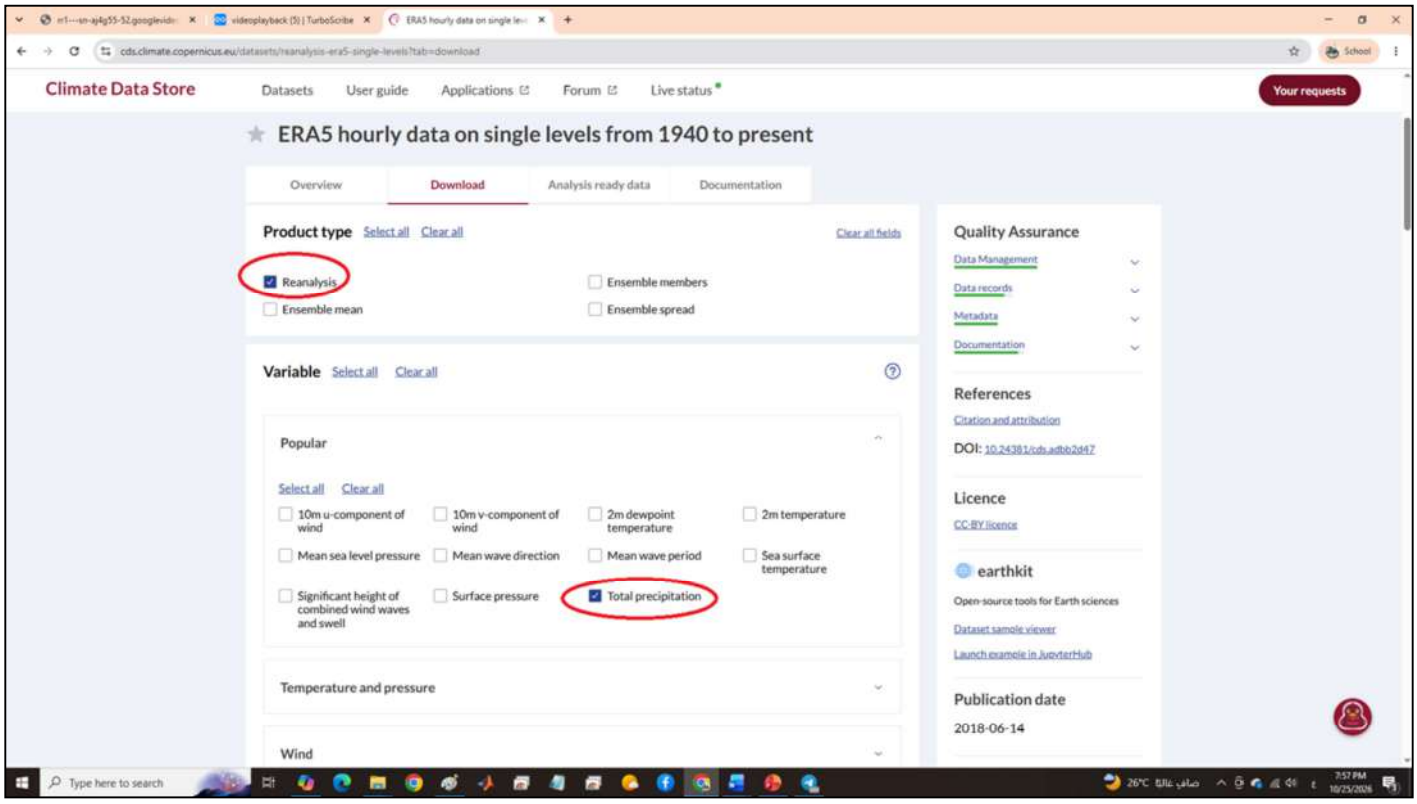
❑ **ERA5**

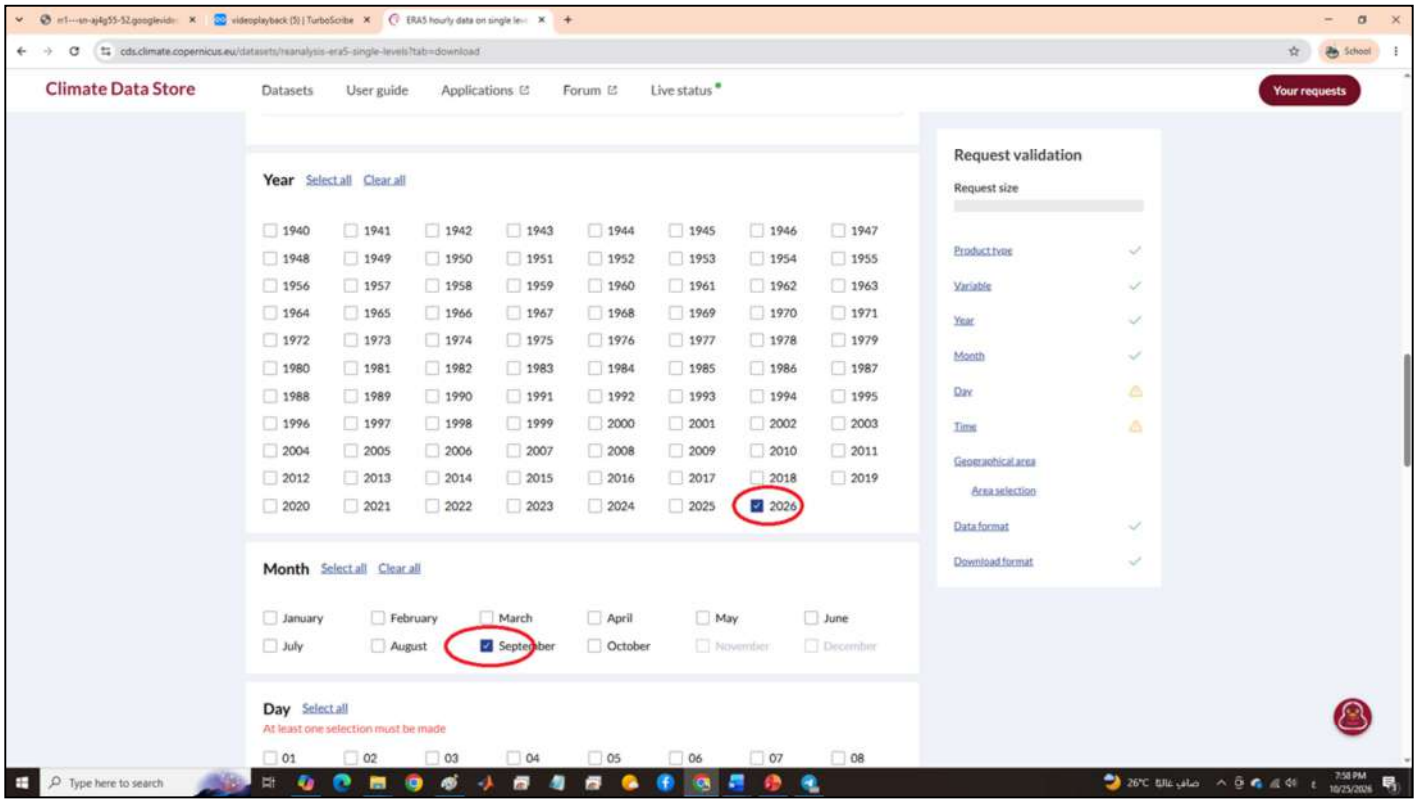
❑ يتوفر أيضًا **مجموعة بيانات أولية لـ ERA5** تغطي الفترة من 1950 إلى 1978 على **موقع بيانات المناخ (CDS)**، مع الإشارة إلى أن الفترة 1959–1978 قد تم استبدالها بمج موعة بيانات ذات دقة عالية

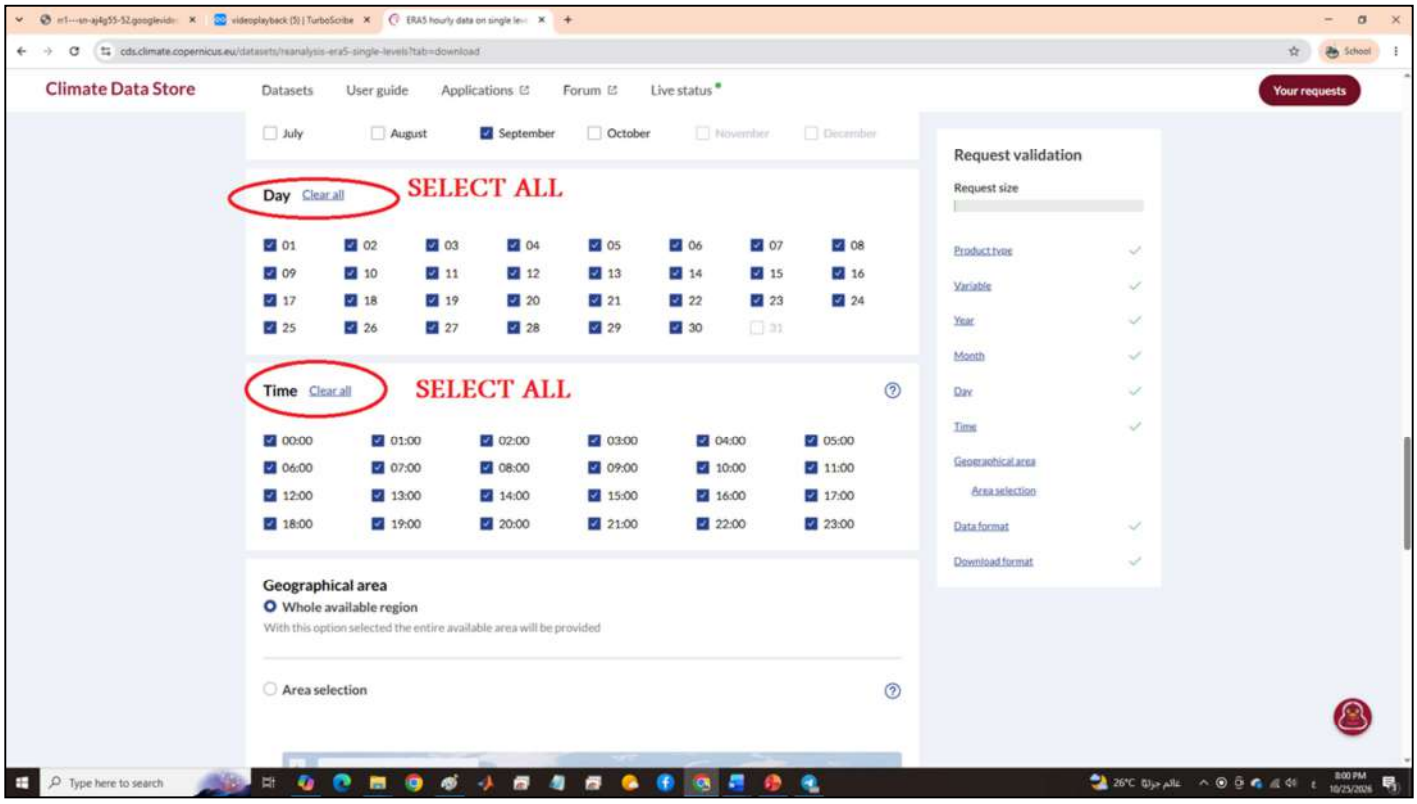
❑ يجمع ERA5 **كميات هائلة من الرصدات التاريخية** في تقديرات عالمية باستخدام نماذج متقدمة وأنظمة استيعاب البيانات.

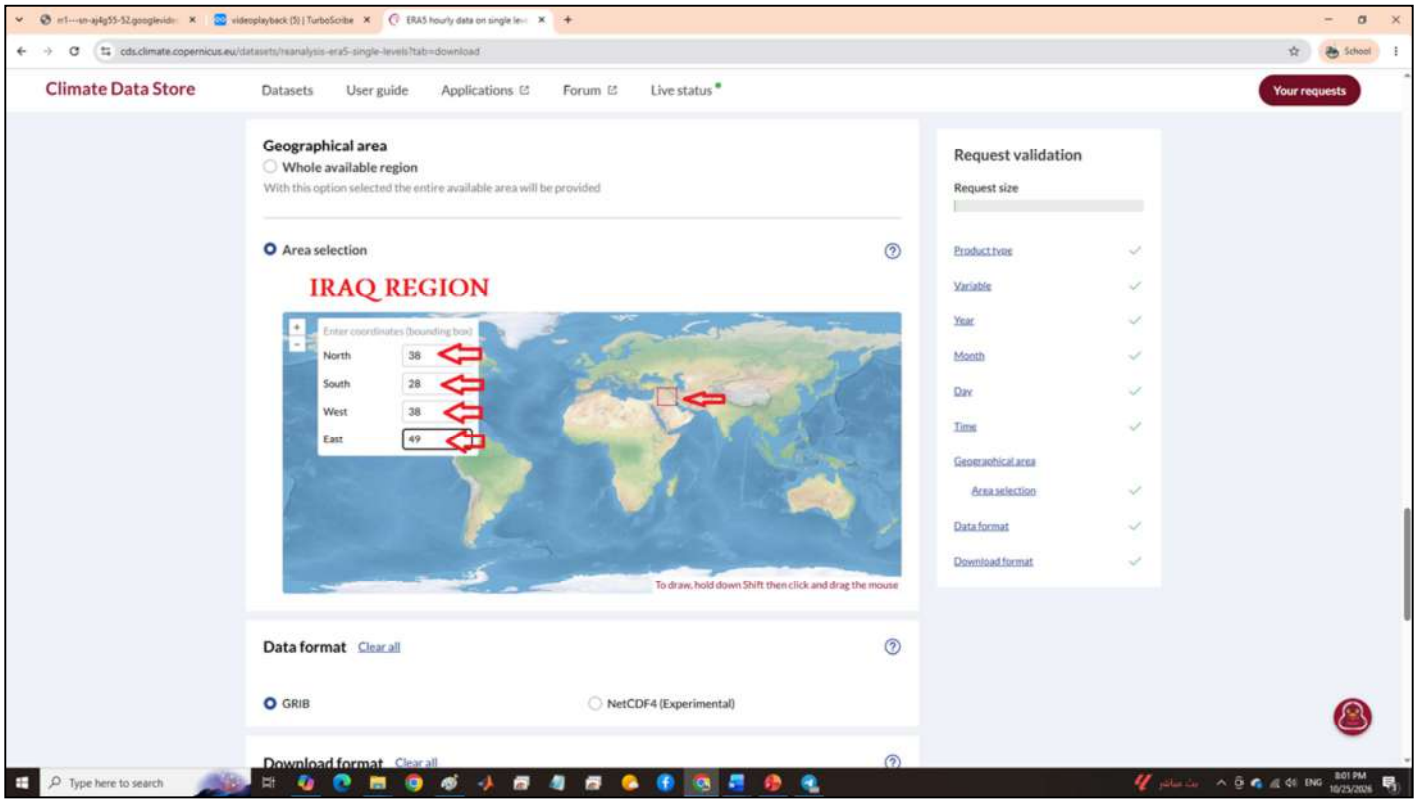
❑ يُعد ERA5 بديلاً عن إعادة التحليل **ERA-Interim** التي توقفت عن الإنتاج في 31 أغسطس 2019.

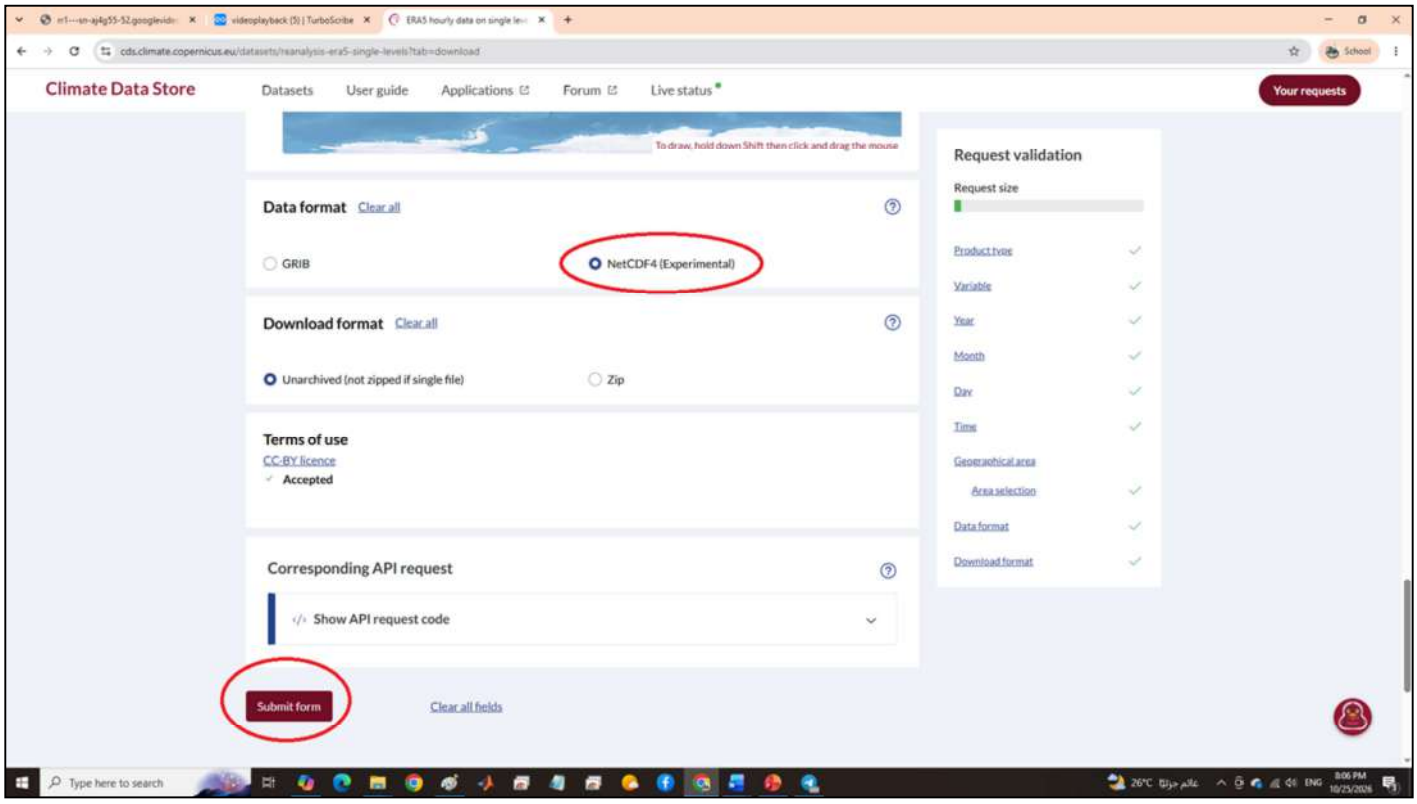

cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels?tab=overview
ECMWF
FIRAS BASHEER
Climate Data Store
Datasets
User guide
Applications
Forum
Live status
Your requests
Info
6 Oct 2026
The CDS Terms of Use will be updated on 28 October 2026. Continued use of Copernicus Products and Services through the CDS after this date constitutes acceptance of the updated terms. If you do not accept the change, please contact ECMWF Support before 27 October 2026. More information about the upcoming change.
Datasets
ERA5 hourly data on single levels from 1940 to present
ERA5 hourly data on single levels from 1940 to present
Overview
Download
Analysis ready data
Documentation
ERA5 is the fifth generation ECMWF reanalysis for the global climate and weather for the past 8 decades. Data is available from 1940 onwards. ERA5 replaces the ERA-Interim reanalysis.
Reanalysis combines model data with observations from across the world into a globally complete and consistent dataset using the laws of physics. This principle, called data assimilation, is based on the method used by numerical weather prediction centres, where every so many hours (12 hours at ECMWF) a previous forecast is combined with newly available observations in an optimal way to produce a new best estimate of the state of the atmosphere, called analysis, from which an updated, improved forecast is issued. Reanalysis works in the same way, but at reduced resolution to allow for the provision of a dataset spanning back several decades. Reanalysis does not have the constraint of issuing timely
ERA5 2 metre temperature and Mean sea level pressure
1 January 2023 at 00:00 UTC
2 metre temperature (°C)
Quality Assurance
Data Management
Data records
Metadata
Documentation
References
Citation and attribution
DOI: 10.24381/cds.adbb2d47
Licence
CC-BY licence
Type here to search
26°C
7:54 PM
10/25/2026

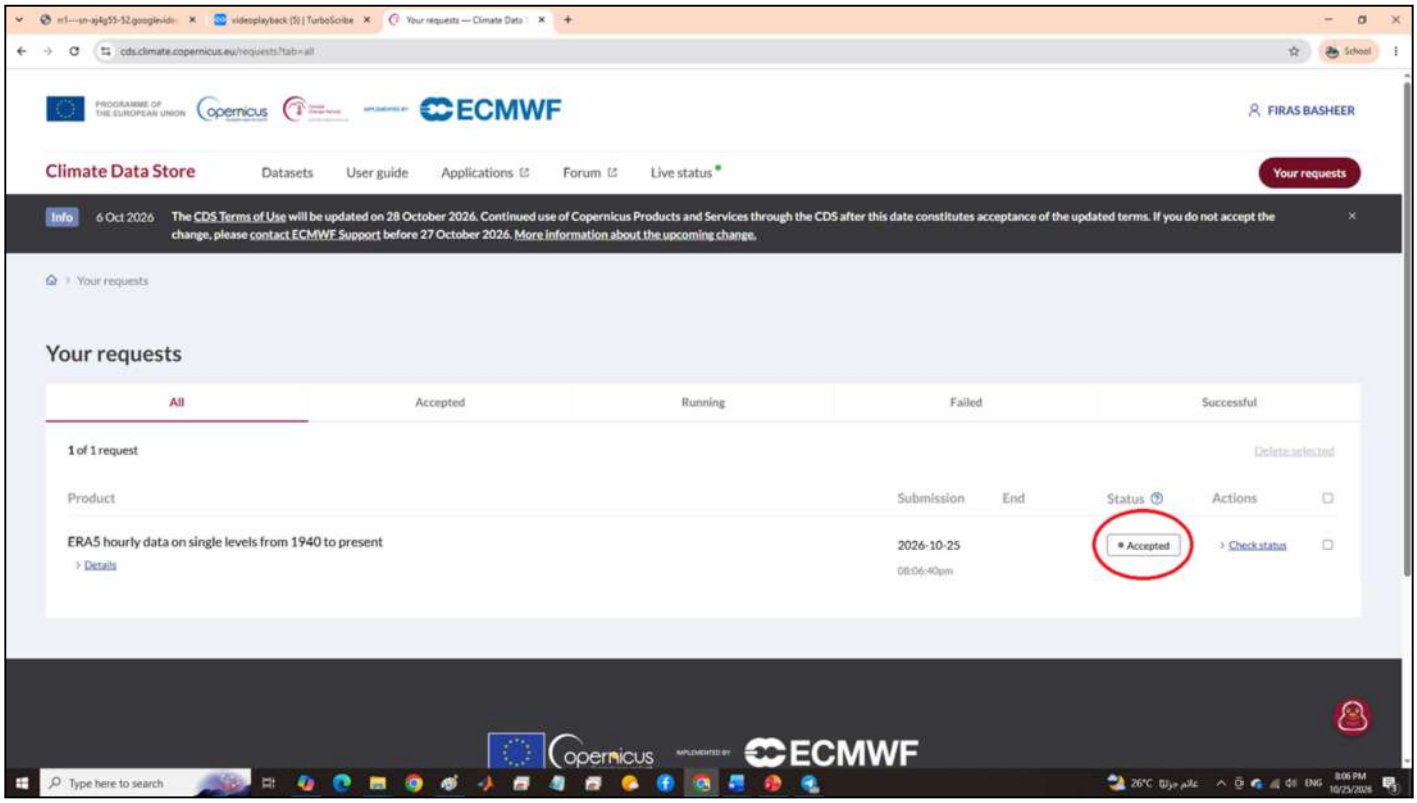

Climate Data Store
Datasets
User guide
Applications
Forum
Live status
Your requests
ERA5 hourly data on single levels from 1940 to present
Overview
Download
Analysis ready data
Documentation
Product type
Select all
Clear all
Clear all fields
Reanalysis
Ensemble members
Ensemble mean
Ensemble spread
Variable
Select all
Clear all
Popular
Select all
Clear all
10m u-component of wind
10m v-component of wind
2m dewpoint temperature
2m temperature
Mean sea level pressure
Mean wave direction
Mean wave period
Sea surface temperature
Significant height of combined wind waves and swell
Surface pressure
Total precipitation
Temperature and pressure
Wind
Quality Assurance
Data Management
Data records
Metadata
Documentation
References
Citation and attribution
DOI: 10.24381/cds.adbb2d47
Licence
CC-BY licence
earthkit
Open-source tools for Earth sciences
Dataset sample viewer
Launch example in JupyterHub
Publication date
2018-06-14
Type here to search

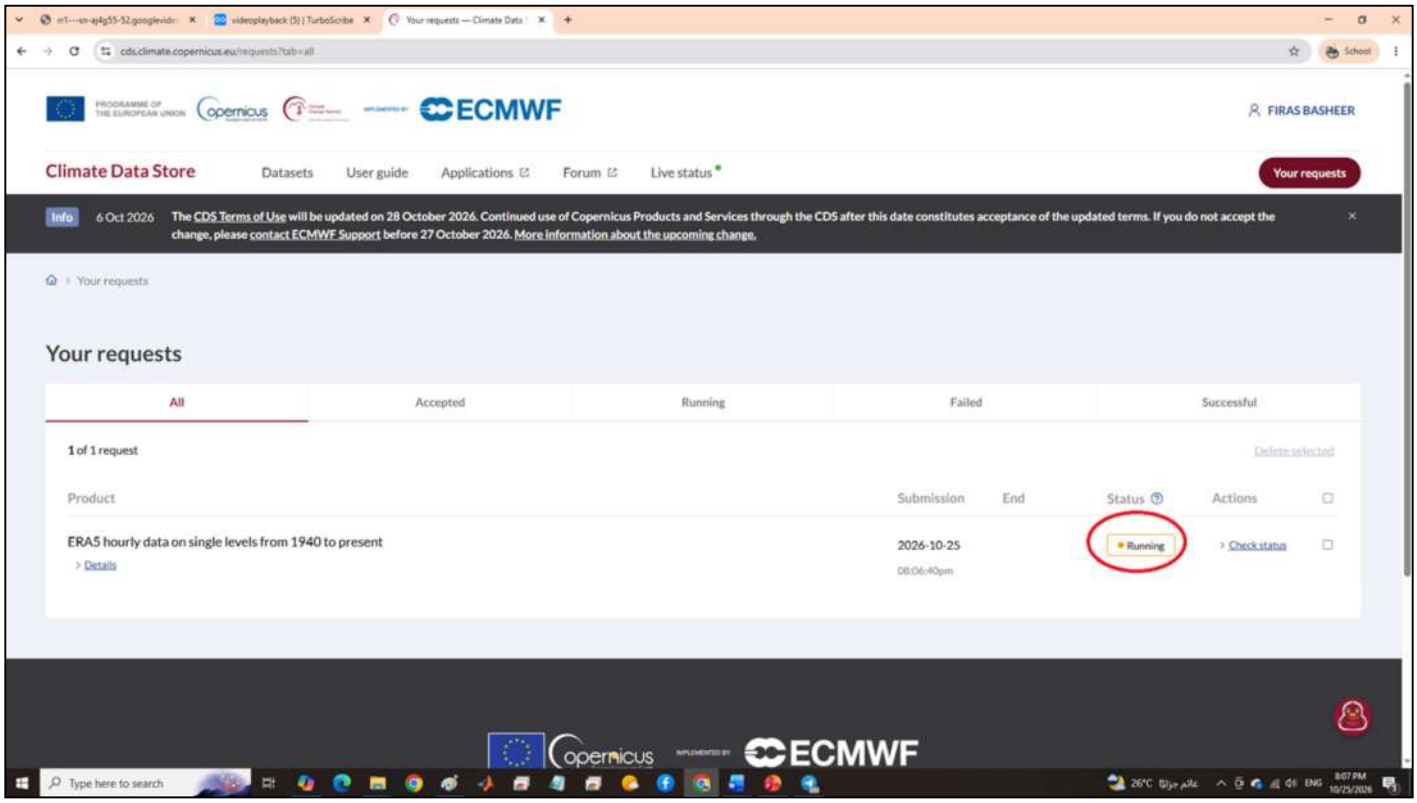

Climate Data Store
Datasets
User guide
Applications
Forum
Live status
Your requests
Year Select all Clear all
1940 1941 1942 1943 1944 1945 1946 1947
1948 1949 1950 1951 1952 1953 1954 1955
1956 1957 1958 1959 1960 1961 1962 1963
1964 1965 1966 1967 1968 1969 1970 1971
1972 1973 1974 1975 1976 1977 1978 1979
1980 1981 1982 1983 1984 1985 1986 1987
1988 1989 1990 1991 1992 1993 1994 1995
1996 1997 1998 1999 2000 2001 2002 2003
2004 2005 2006 2007 2008 2009 2010 2011
2012 2013 2014 2015 2016 2017 2018 2019
2020 2021 2022 2023 2024 2025 2026
Month Select all Clear all
January February March April May June
July August September October November December
Day Select all
At least one selection must be made
01 02 03 04 05 06 07 08
Request validation
Request size
Product type
Variable
Year
Month
Day
Time
Geographical area
Area selection
Data format
Download format

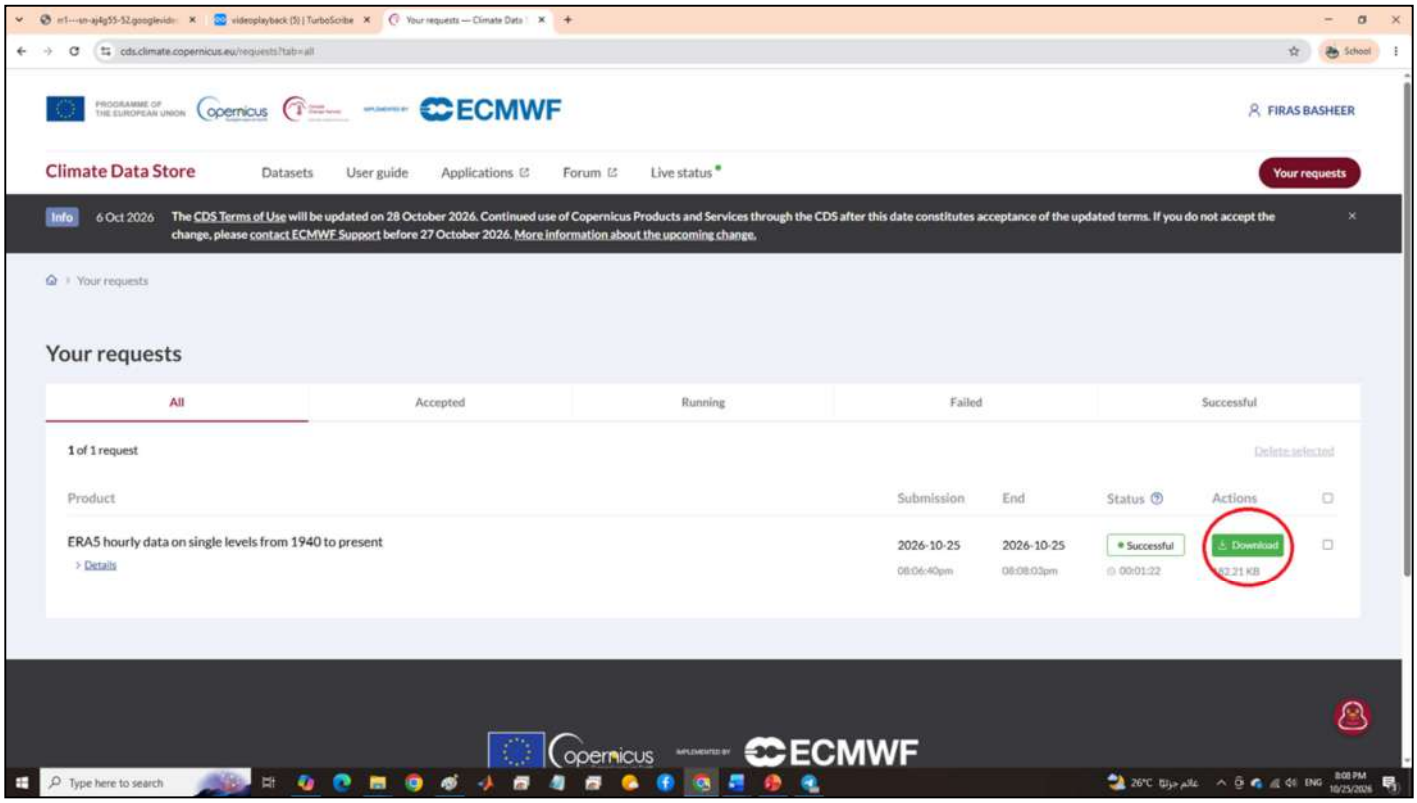

Climate Data Store
Datasets
User guide
Applications
Forum
Live status
Your requests
July
August
September
October
November
December
Day Clear all
SELECT ALL
01 02 03 04 05 06 07 08
09 10 11 12 13 14 15 16
17 18 19 20 21 22 23 24
25 26 27 28 29 30 31
Time Clear all
SELECT ALL
00:00 01:00 02:00 03:00 04:00 05:00
06:00 07:00 08:00 09:00 10:00 11:00
12:00 13:00 14:00 15:00 16:00 17:00
18:00 19:00 20:00 21:00 22:00 23:00
Geographical area
Whole available region
With this option selected the entire available area will be provided
Area selection
Request validation
Request size
Product type
Variable
Year
Month
Day
Time
Geographical area
Area selection
Data format
Download format
Type here to search

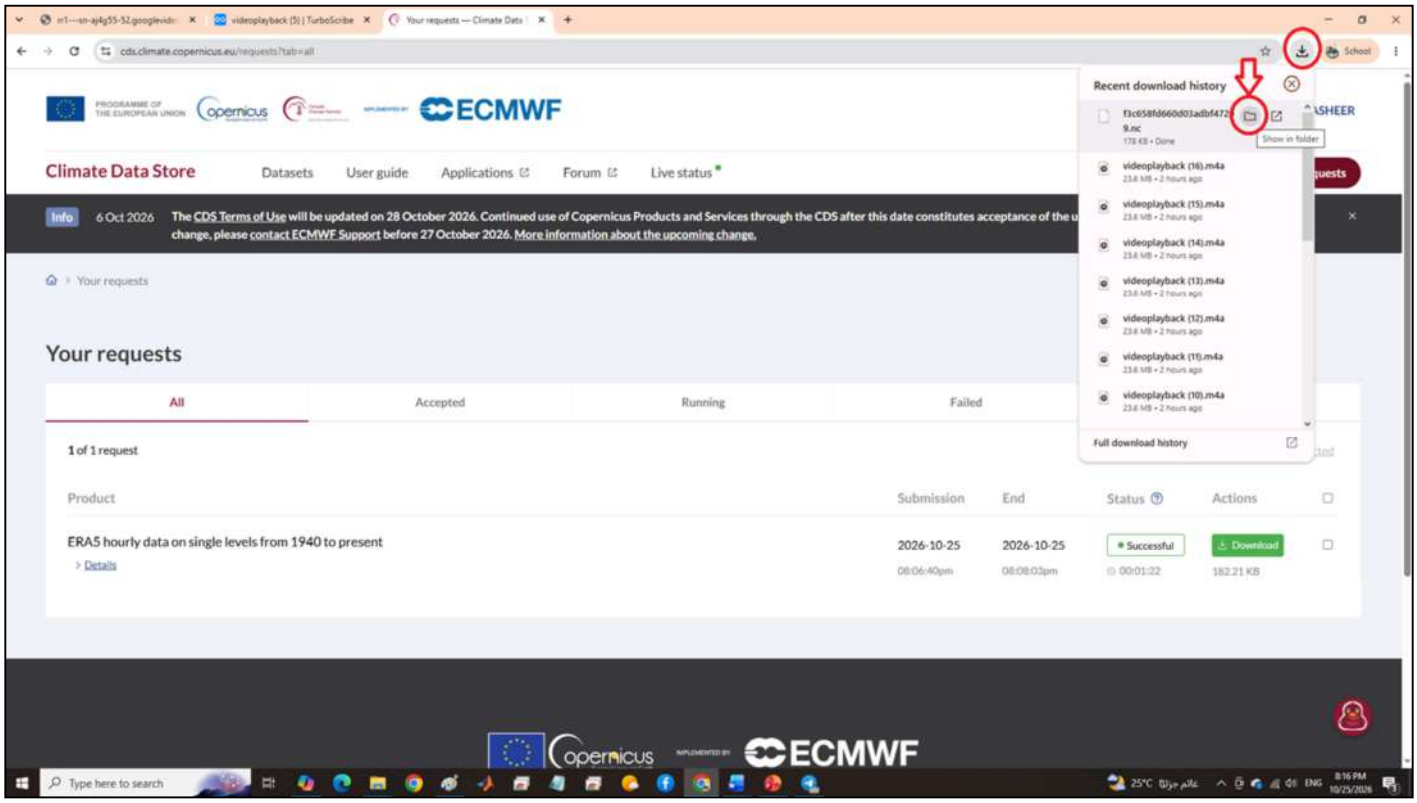

Climate Data Store
Datasets
User guide
Applications
Forum
Live status
Your requests
Geographical area
Whole available region
With this option selected the entire available area will be provided
Area selection
IRAQ REGION
Enter coordinates (bounding box)
North 38
South 28
West 38
East 49
To draw, hold down Shift then click and drag the mouse
Request validation
Request size
Product type
Variable
Year
Month
Day
Time
Geographical area
Area selection
Data format
Download format
Data format Clear all
GRIB
NetCDF4 (Experimental)
Download format Clear all
Type here to search

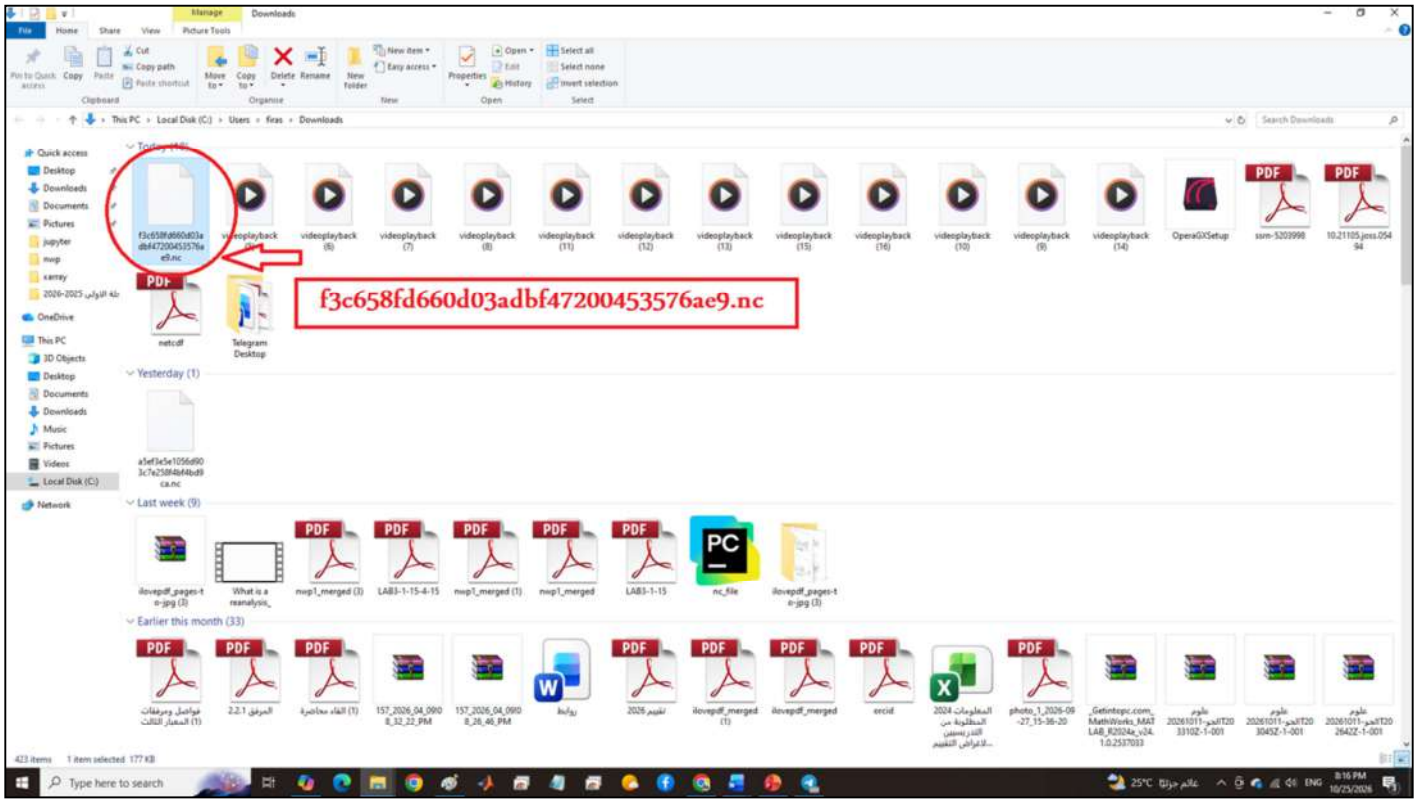

Climate Data Store
Datasets
User guide
Applications
Forum
Live status
Your requests
To draw, hold down Shift then click and drag the mouse
Data format
Clear all
GRIB
NetCDF4 (Experimental)
Download format
Clear all
Unarchived (not zipped if single file)
Zip
Terms of use
CC-BY licence
Accepted
Corresponding API request
Show API request code
Submit form
Clear all fields
Request validation
Request size
Product type
Variable
Year
Month
Day
Time
Geographical area
Area selection
Data format
Download format
Type here to search

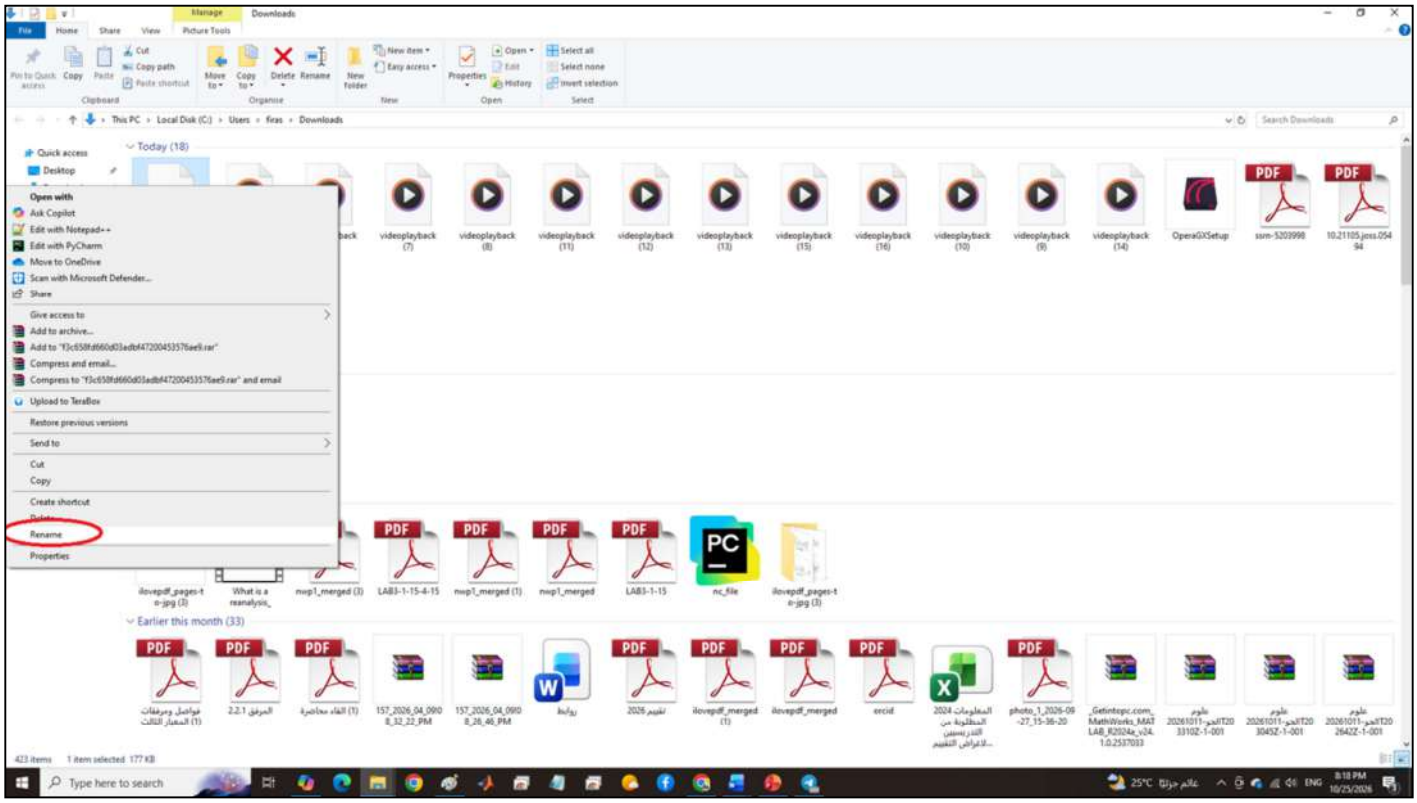

cds.climate.copernicus.eu/requests?tab=all
FIRAS BASHEER
Climate Data Store
Datasets
User guide
Applications
Forum
Live status
Your requests
Info
6 Oct 2026
The CDS Terms of Use will be updated on 28 October 2026. Continued use of Copernicus Products and Services through the CDS after this date constitutes acceptance of the updated terms. If you do not accept the change, please contact ECMWF Support before 27 October 2026. More information about the upcoming change.
Your requests
Your requests
All
Accepted
Running
Failed
Successful
1 of 1 request
Delete selected
Product
Submission
End
Status
Actions
ERA5 hourly data on single levels from 1940 to present
Details
2026-10-25
08:06:40pm
Accepted
Check status
Type here to search

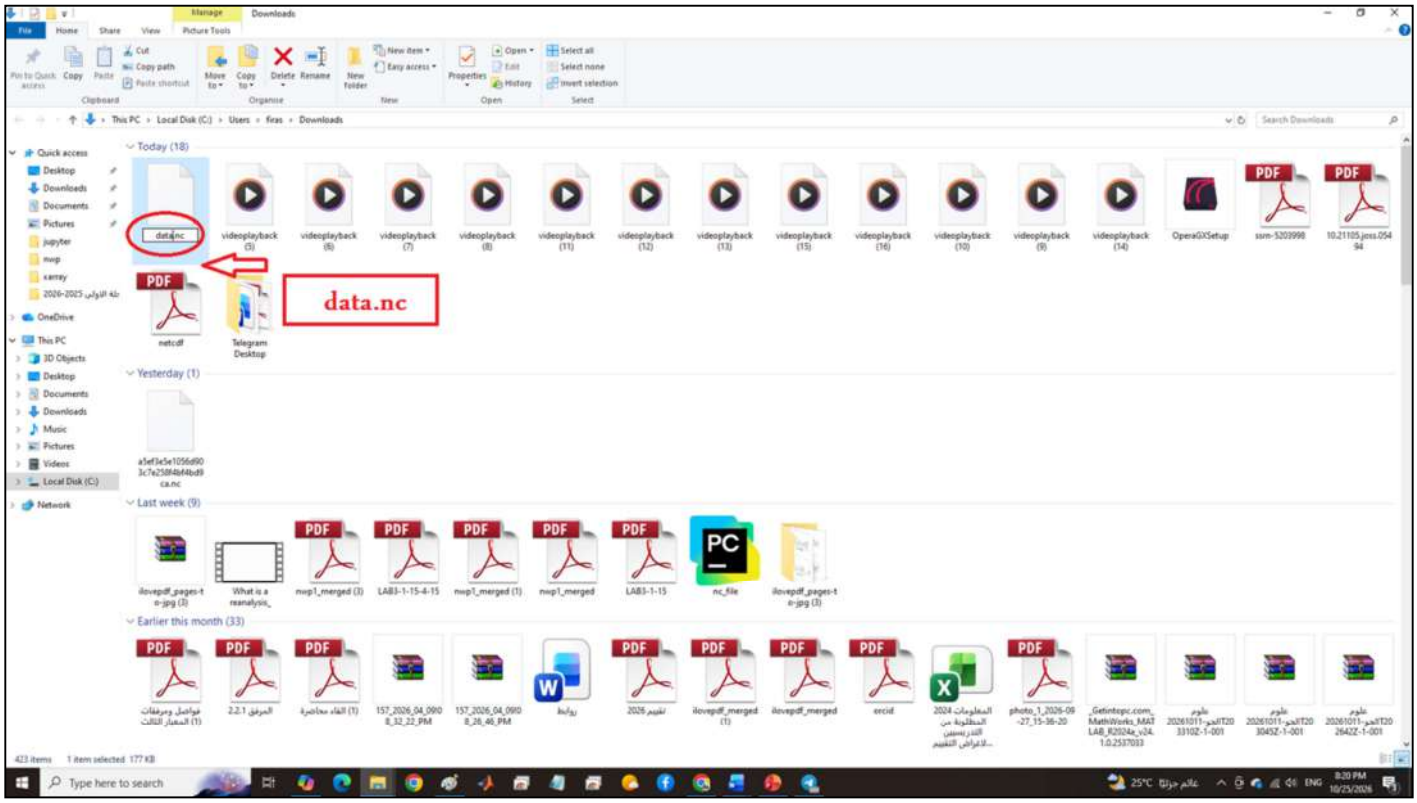

cds.climate.copernicus.eu/requests?tab=all
ECMWF
FIRAS BASHEER
Climate Data Store
Datasets
User guide
Applications
Forum
Live status
Your requests
Info
6 Oct 2026
The CDS Terms of Use will be updated on 28 October 2026. Continued use of Copernicus Products and Services through the CDS after this date constitutes acceptance of the updated terms. If you do not accept the change, please contact ECMWF Support before 27 October 2026. More information about the upcoming change.
Your requests
Your requests
All
Accepted
Running
Failed
Successful
1 of 1 request
Delete selected
Product
Submission
End
Status
Actions
ERA5 hourly data on single levels from 1940 to present
Details
2026-10-25
08:06:40pm
Running
Check status
Type here to search

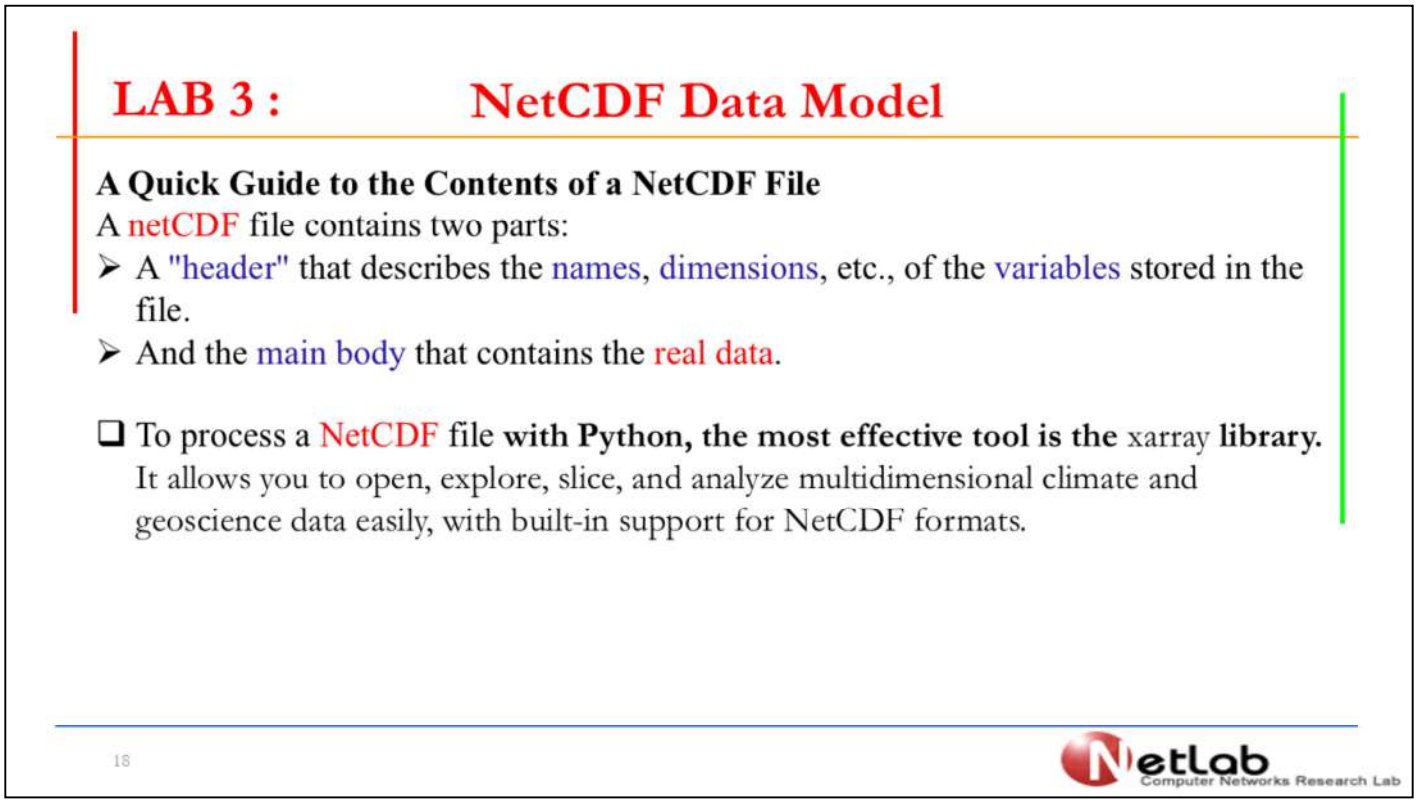

Climate Data Store
Datasets
User guide
Applications
Forum
Live status
Your requests
FIRAS BASHEER
Info
6 Oct 2026
The CDS Terms of Use will be updated on 28 October 2026. Continued use of Copernicus Products and Services through the CDS after this date constitutes acceptance of the updated terms. If you do not accept the change, please contact ECMWF Support before 27 October 2026. More information about the upcoming change.
Your requests
All
Accepted
Running
Failed
Successful
1 of 1 request
Product
Submission
End
Status
Actions
ERA5 hourly data on single levels from 1940 to present
Details
2026-10-25
08:06:40pm
2026-10-25
08:08:03pm
Successful
00:01:22
Download
82.21 KB
Type here to search


Recent download history
Show in folder
Full download history
Climate Data Store
Datasets
User guide
Applications
Forum
Live status
Info
6 Oct 2026
Your requests
All
Accepted
Running
Failed
1 of 1 request
Product
Submission
End
Status
Actions
ERA5 hourly data on single levels from 1940 to present
Details
2026-10-25
08:06:40pm
2026-10-25
08:08:03pm
Successful
00:01:22
Download
182.21 KB
ECMWF
Type here to search


f3c658fd660d03adbf47200453576ae9.nc


Open with
Ask Copilot
Edit with Notepad++
Edit with PyCharm
Move to OneDrive
Scan with Microsoft Defender...
Share
Give access to
Add to archive...
Compress and email...
Upload to TeraBox
Restore previous versions
Send to
Cut
Copy
Create shortcut
Rename
Properties
Today (18)
Earlier this month (33)


data.nc



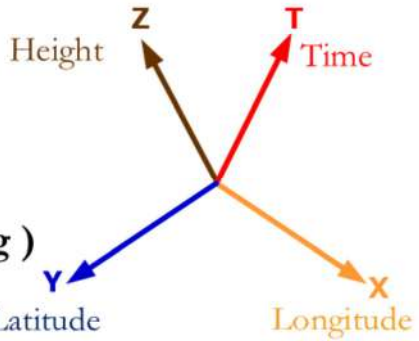
- دليل سريع لمحتويات ملف NetCDF
- يحتوي ملف NetCDF على جزأين:
- "الرأس" **header** يصف الأسماء والأبعاد ، وما إلى ذلك ، من المتغيرات التي يتم حفظها في الملف.
- والجسم الرئيسي **main body** الذي يحتوي على البيانات الحقيقية.
- تُعد مكتبة xarray الأداة الأكثر فعالية لمعالجة ملفات NetCDF باستخدام لغة بايثون ؛ فهي تتيح لك فتح بيانات المناخ والعلوم الجيولوجية متعددة الأبعاد واستكشافها وتقسيمها و تحليلها بسهولة، مع توفير دعم مدمج لتنسيقات NetCDF.

LAB 3 : Part(B) NetCDF Data Model(Contains)

A NetCDF dataset contains :

- ❑ **Dimensions**
- ❑ **Variables**
- ❑ **Attributes**

- ❑ **Dimensions**
  - ➢ **Time**
  - ➢ **Latitude**
  - ➢ **Longitude**



- ❑ **Variables(time, lat, long)**

Variables
long

lat		
	10	11
	12	13

For each location **one** value so for this **ex.** we have **four** locations .

NetLab Computer Networks Research Lab

LAB 3 : Part(B) NetCDF Data Model(Contains)

A NetCDF dataset contains :

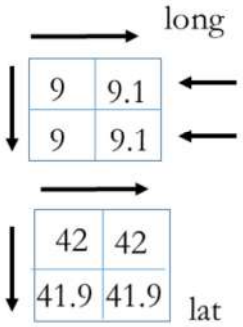
- ❑ Dimensions
- ❑ Variables
- ❑ Attributes

- ❑ Dimensions
  - ➢ Time
  - ➢ Latitude
  - ➢ Longitude

- ❑ Variables(time, lat, long)

Variables

lat \ long		
	10	11
	12	13



NetLab Computer Networks Research Lab

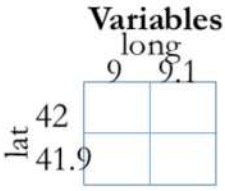
LAB 3 : Part(B) NetCDF Data Model(Contains)

A NetCDF dataset contains :

- ❑ Dimensions
- ❑ Variables
- ❑ Attributes

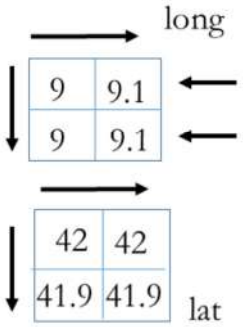
- ❑ Dimensions
  - ➢ Time
  - ➢ Latitude
  - ➢ Longitude

- ❑ Variables(time, lat, long)



$V_{(42,9)}$ $V_{(42,9.1)}$

$V_{(41.9,9)}$ $V_{(41.9, 9.1)}$



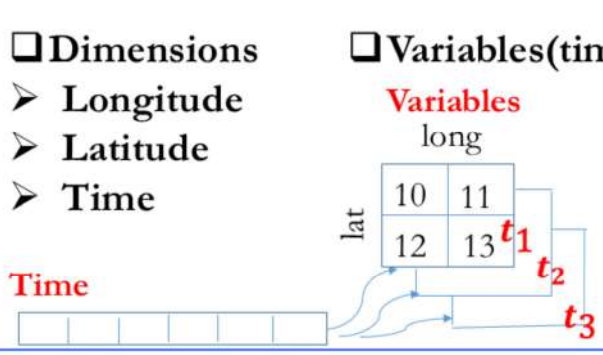
NetLab Computer Networks Research Lab

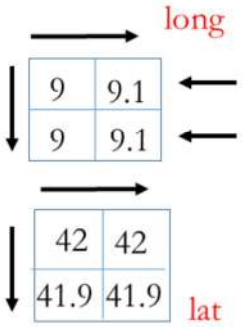
LAB 3 : Part(B) NetCDF Data Model(Contains)

A NetCDF dataset contains :

- ❑ Dimensions
- ❑ Variables
- ❑ Attributes

- ❑ Dimensions
  - ➢ Longitude
  - ➢ Latitude
  - ➢ Time
- ❑ Variables(time, lat ,long)



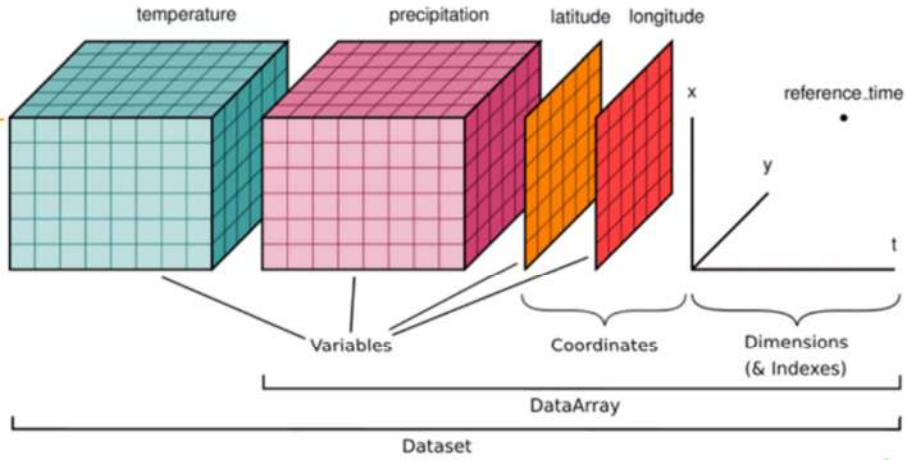


NetLab Computer Networks Research Lab

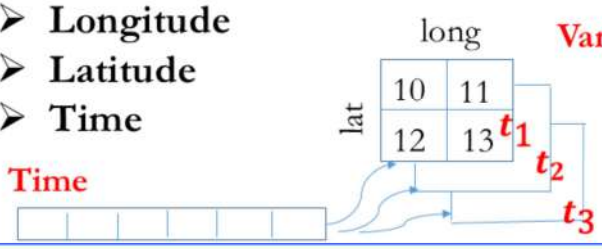
LAB 3 :

A NetCDF dataset contains :

- ❑ Dimensions
- ❑ Variables
- ❑ Attributes



- ❑ Dimensions
  - ➢ Longitude
  - ➢ Latitude
  - ➢ Time

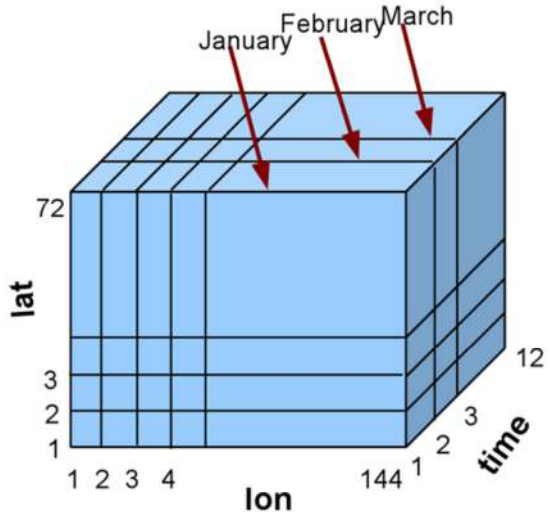


❑ Variables(time, lat , long)

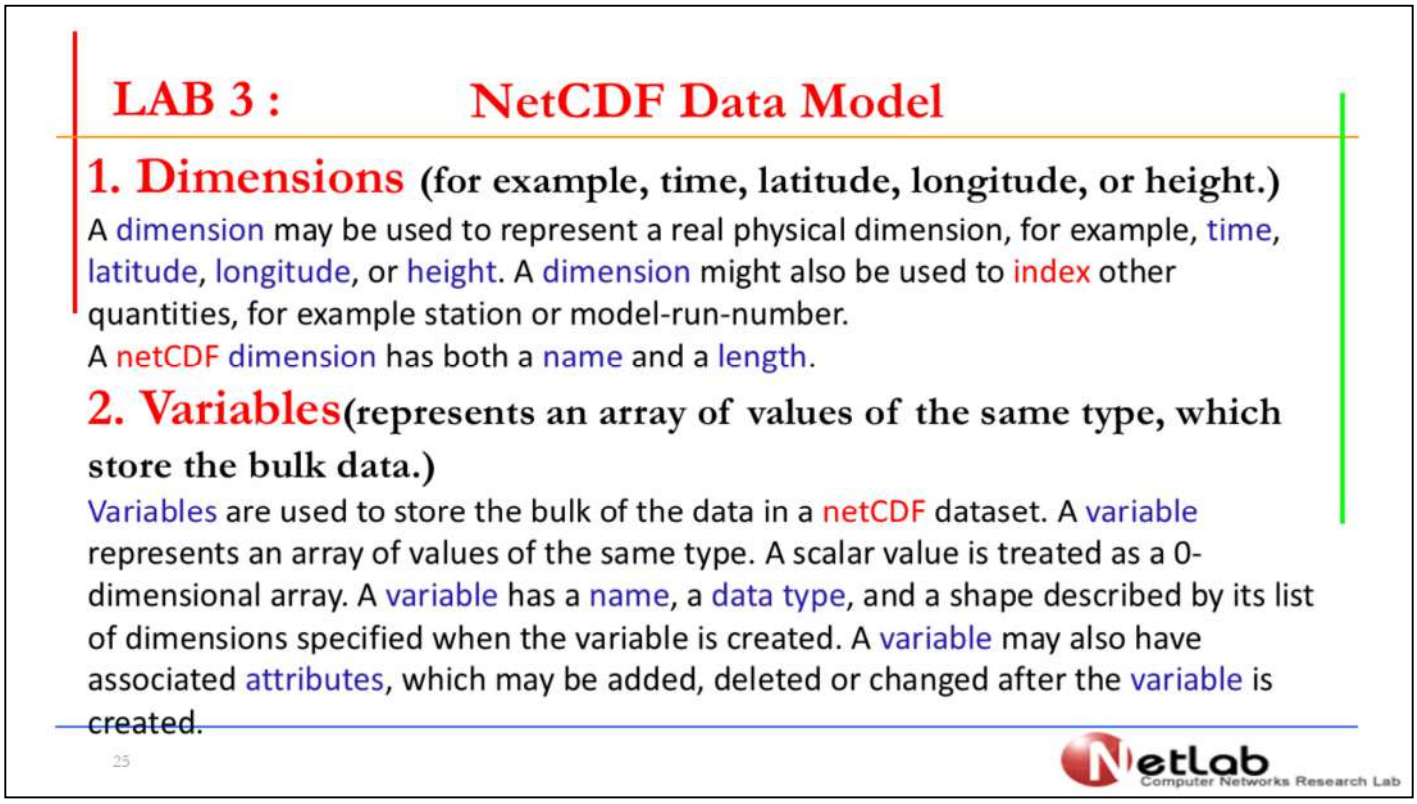
NetLab Computer Networks Research Lab

LAB 3 : Part(B) NetCDF Data Model(Contains)

- Here, we use the "normal" convention with each of the indices, i, j, and k, starting from 1.



NetLab Computer Networks Research Lab



❑1. الأبعاد (على سبيل المثال: الزمن، خط العرض، خط الطول، أو الارتفاع).

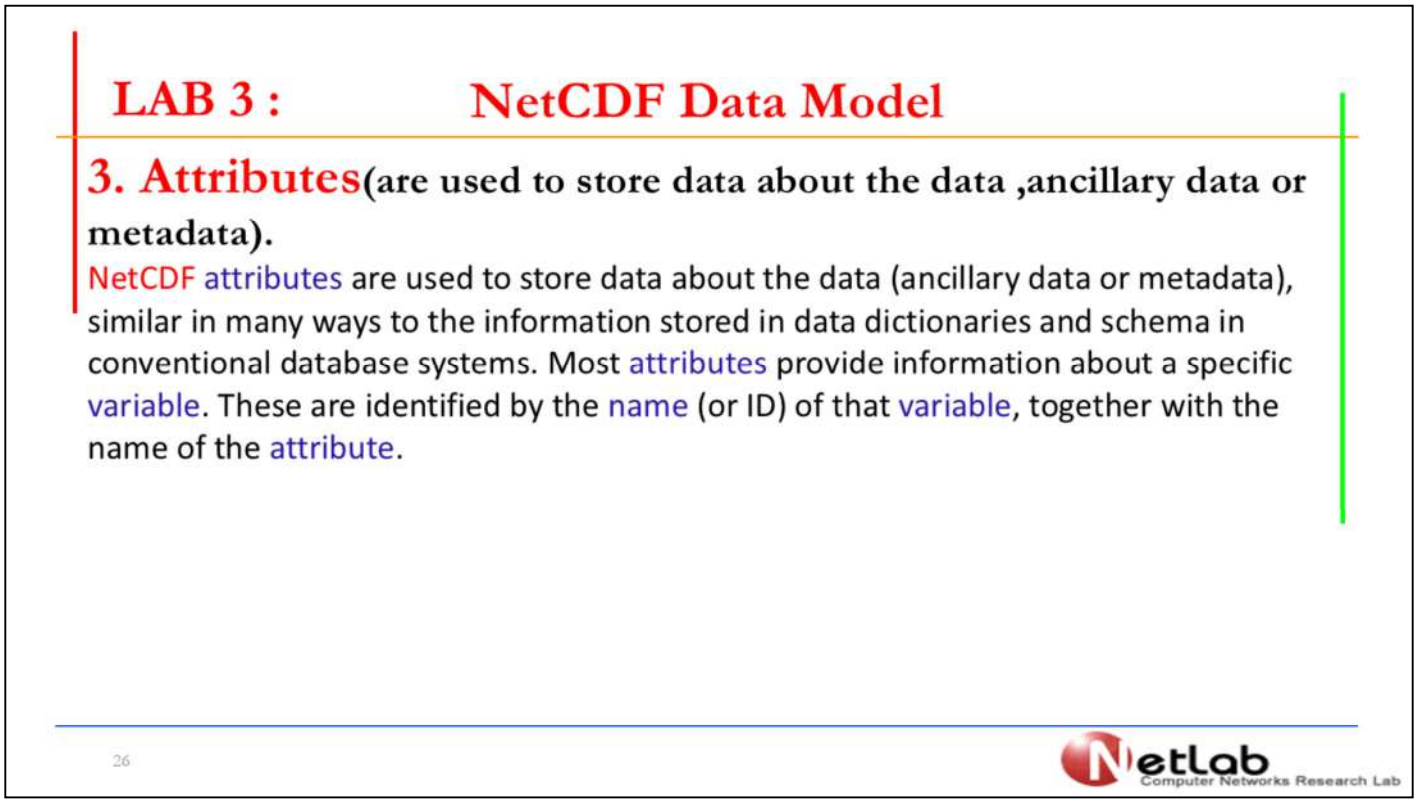
يمكن استخدام البعد لتمثيل بُعد فيزيائي حقيقي، مثل الزمن أو خط العرض أو

خط الطول أو الارتفاع. كما يمكن استخدامه لفهرسة كميات أخرى، مثل تحديد محطة رص د معينة أو رقم تشغيل لنموذج حاسوبي.

يتميز البعد في تنسيق netCDFبوجود اسم وطول محددين له.

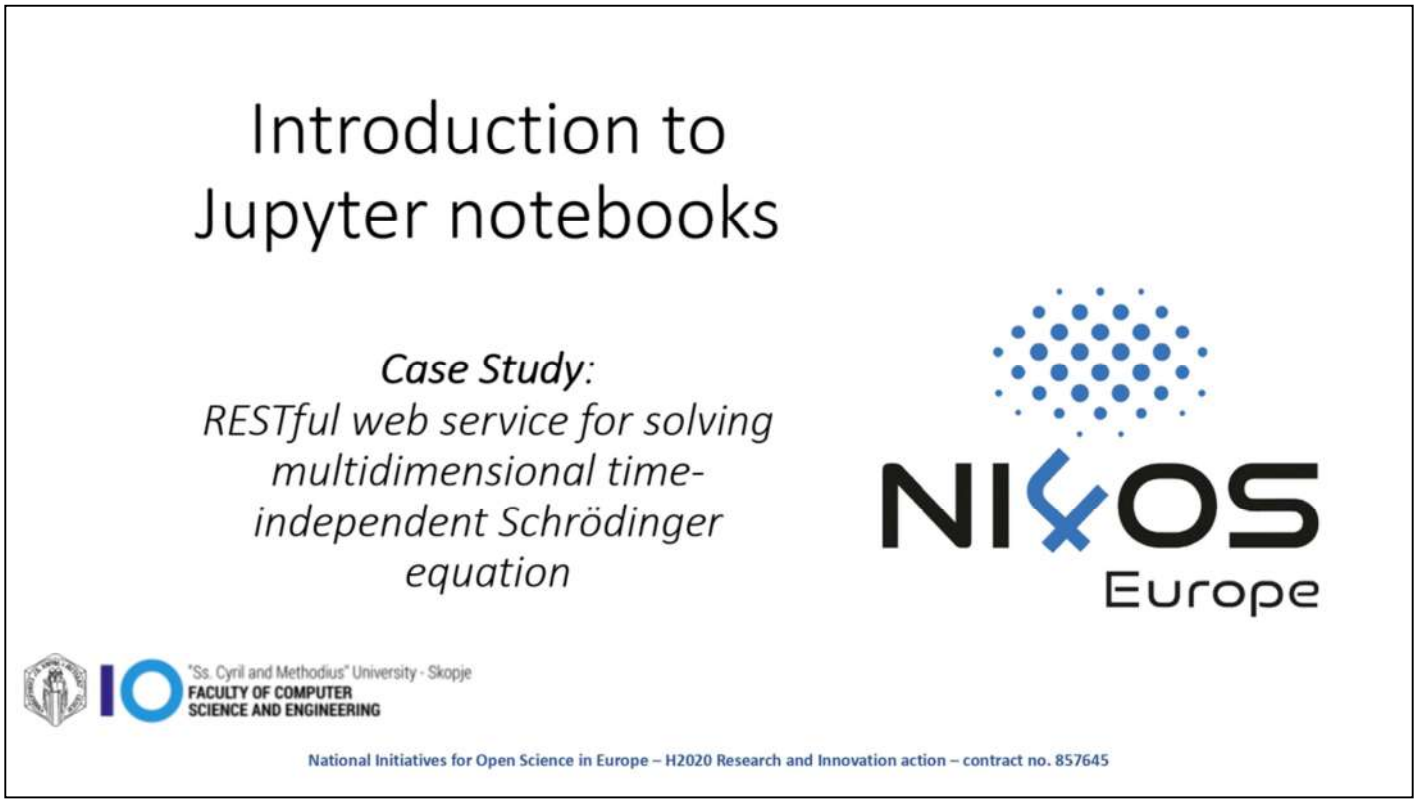
2. المتغيرات (تمثل مصفوفة من القيم من النوع نفسه، وتُستخدم لتخزين البيانات الأساسية) .

تُستخدم المتغيرات لتخزين الجزء الأكبر من البيانات في مجموعة بيانات netCDF. يمثل المتغير مصفوفة من القيم التي تنتمي إلى النوع نفسه، وتُعامل القيمة المفردة ((scalarعل ى أنها مصفوفة ذات بُعد صفري. لكل متغير اسم ونوع بيانات وشكل (بنية) يتحدد بناءً عل ى قائمة الأبعاد التي تُحدد عند إنشاء المتغير. كما يمكن أن يرتبط بالمتغير مجموعة من ال سمات ((attributes، والتي يمكن إضافتها أو حذفها أو تعديلها حتى بعد إنشاء المتغير.



3. الخصائص (تُستخدم لتخزين بيانات حول البيانات، أو بيانات إضافية، أو ما يُعرف بالبيانات الوصفية "metadata").

تُستخدم خصائص NetCDF لتخزين بيانات حول البيانات (بيانات إضافية أو بيانات وصفية)، وهي تشبه في جوانب كثيرة المعلومات المخزنة في قواميس البيانات والمخططات (schemas) في أنظمة قواعد البيانات التقليدية. توفر معظم الخصائص معلومات حول متغير محدد؛ ويتم تحديد هذه الخصائص من خلال اسم المتغير (أو مُعرِّفه) مقترنًا باسم الخاصية ذاتها.



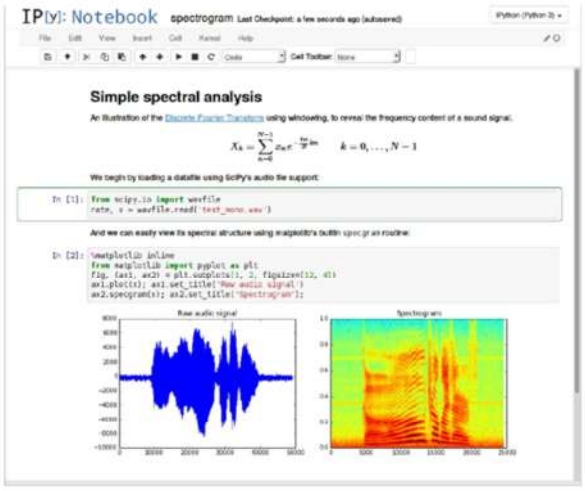
مدخل إلى بيئة Jupyter Notebook

What is a Notebook?

A *notebook* combines the functionality of

- a word processor — handles formatted text
- integrates code and its output into a single document
- combines visualizations, narrative text, mathematical equations, and other rich media

Single document where you can run code, display the output, add explanations, formulas, charts, and make your work more transparent, understandable, repeatable, and shareable.



National Initiatives for Open Science in Europe – H2020 Research and Innovation action – contract no. 857645

- ما هو الـ Notebook؟
- الـ Notebook هو بيئة تجمع بين وظائف متعددة:
- يشبه معالج النصوص في التعامل مع النصوص المنسّقة.
- يدمج الشيفرة البرمجية مع ناتجها في مستند واحد.
- يجمع بين التصورات البيانية، النصوص التوضيحية، المعادلات الرياضية، ووسائط غنية أخرى.
- هو مستند واحد يمكنك فيه تشغيل الشيفرة، عرض النتائج، إضافة الشرح، المعادلات، ال رسوم البيانية، مما يجعل عملك أكثر وضوحًا، قابلًا للفهم، التكرار، والمشاركة.
- في الجزء الأيمن من الشريحة يظهر مثال على Jupyter Notebook بعنوان Sim ple spectral analysis، حيث يحتوي على خلايا برمجية تعرض:
- قراءة ملف صوتي.
- رسم الإشارة الصوتية الخام.
- رسم مخطط الطيف (Spectrogram).

What is Jupyter Notebook?

- Incredibly powerful tool for interactively developing and presenting data science projects
- Open-source web application that you can use to create and share documents
- Can be installed locally (Python installed)
  - pip3 install jupyter (Linux, Mac OS x)
  - Anaconda Navigator (Windows)
- or run Online without installing anything
  https://jupyter.org/try



National Initiatives for Open Science in Europe – H2020 Research and Innovation action – contract no. 857645

❑**ما هو Jupyter Notebook؟**

❑**Jupyter Notebook** هو أداة قوية جدًا لتطوير وعرض مشاريع علوم البيانات بشكل تفاعلي.

❑هو تطبيق ويب مفتوح المصدر يمكن استخدامه لإنشاء المستندات ومشاركتها.

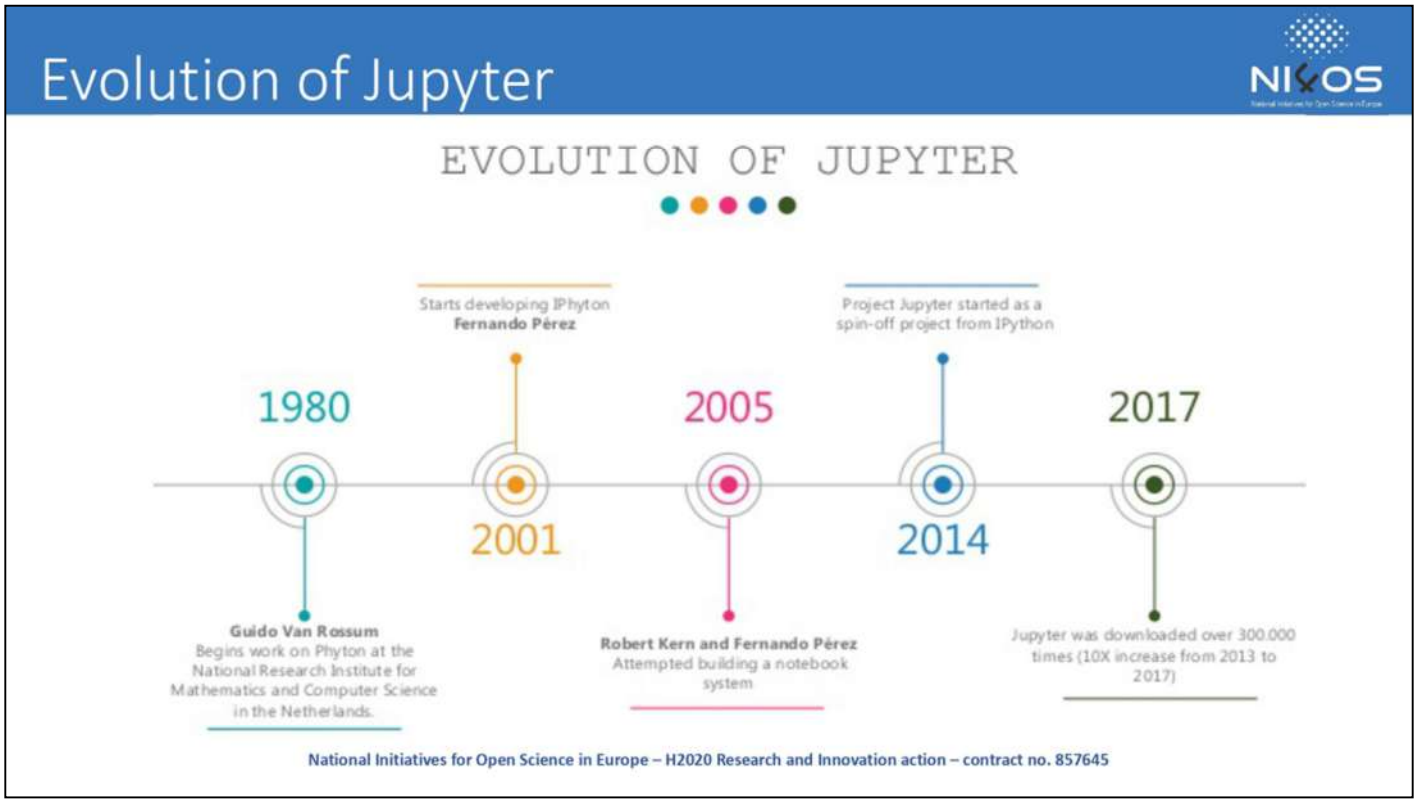
❑يمكن تثبيته محليًا)مع وجود Python)باستخدام:

➢) pip3 install jupyterلينكس، ماك OS X)

➢أو عبر (**Anaconda Navigator**ويندوز.(

❑يمكن تشغيله أيضًا عبر الإنترنت دون الحاجة لتثبيت أي شيء من خلال الرابط: jupyter.org/try.

الفكرة الأساسية**Jupyter Notebook** : يجمع بين النصوص، الشيفرة البرمجية، والرسوم البيانية في مستند واحد، مما يجعله أداة تعليمية وبحثية تفاعلية.



❑ **تطوّر مشروعJupyter**

❑ **1980**بدأ **غيدو فان روسوم** العمل على لغة **Python** في المعهد الوطني للبحوث في الرياضيات وعلوم الحاسوب في هولندا.

❑ **2001**بدأ **فرناندو بيريز** تطوير مشروع.**IPython**

❑ **2005**قام **روبرت كيرن** و**فرناندو بيريز** بمحاولة بناء نظام دفاترNotebook Syst) em).

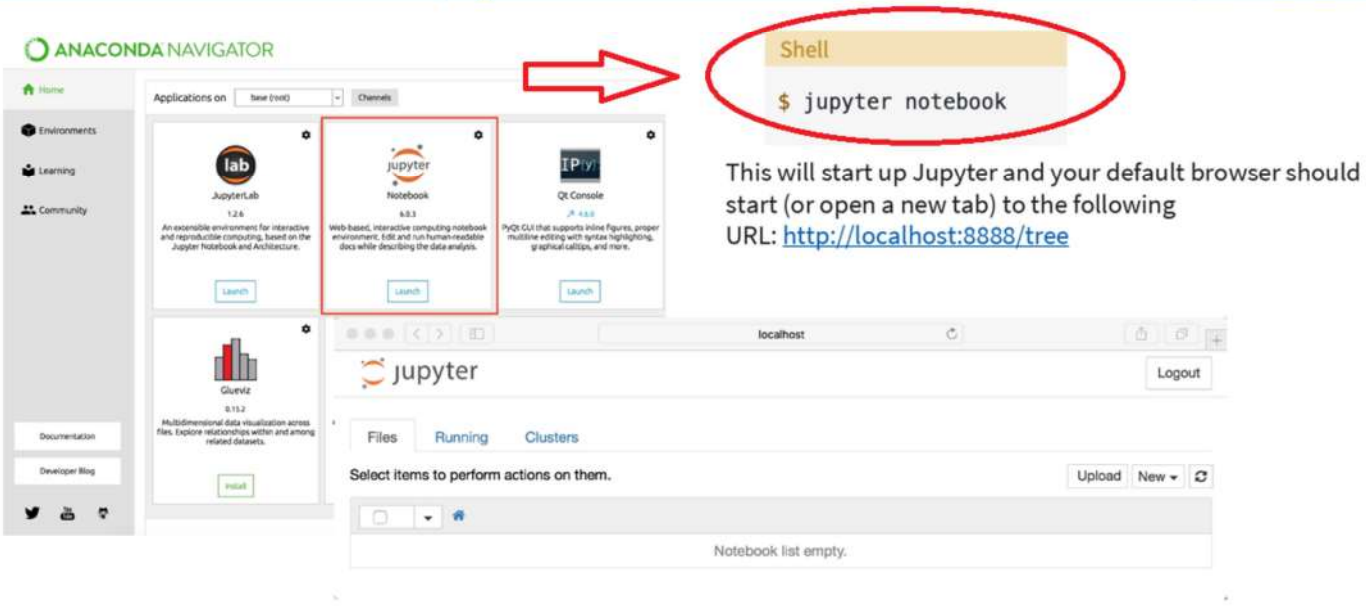
❑ **2014**انطلق مشروع **Jupyter** كمشروع مستقل من.**IPython**

❑ **2017**تم تحميل **Jupyter** أكثر من 300,000 مرة (زيادة بمقدار عشرة أضعاف بين عامي 2013 و2017).

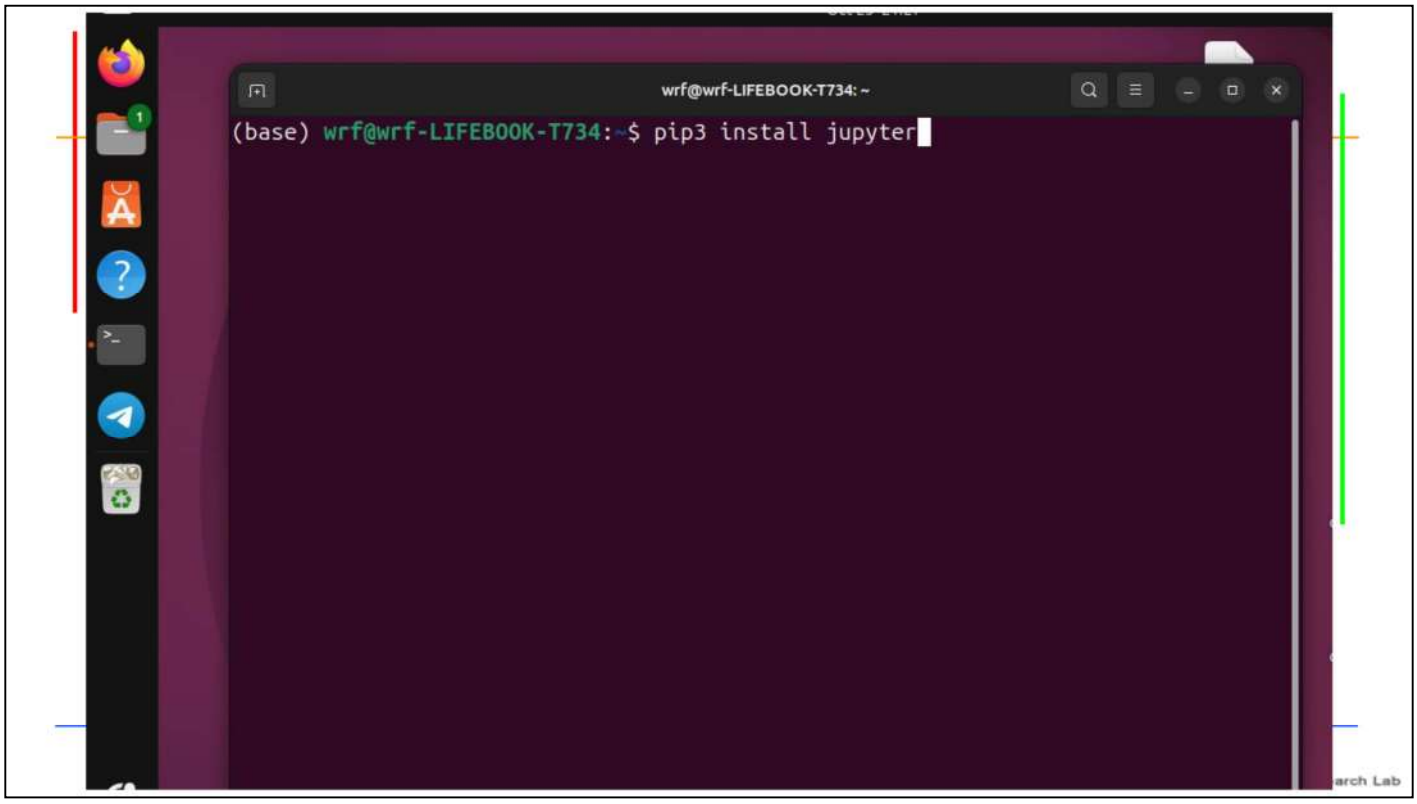


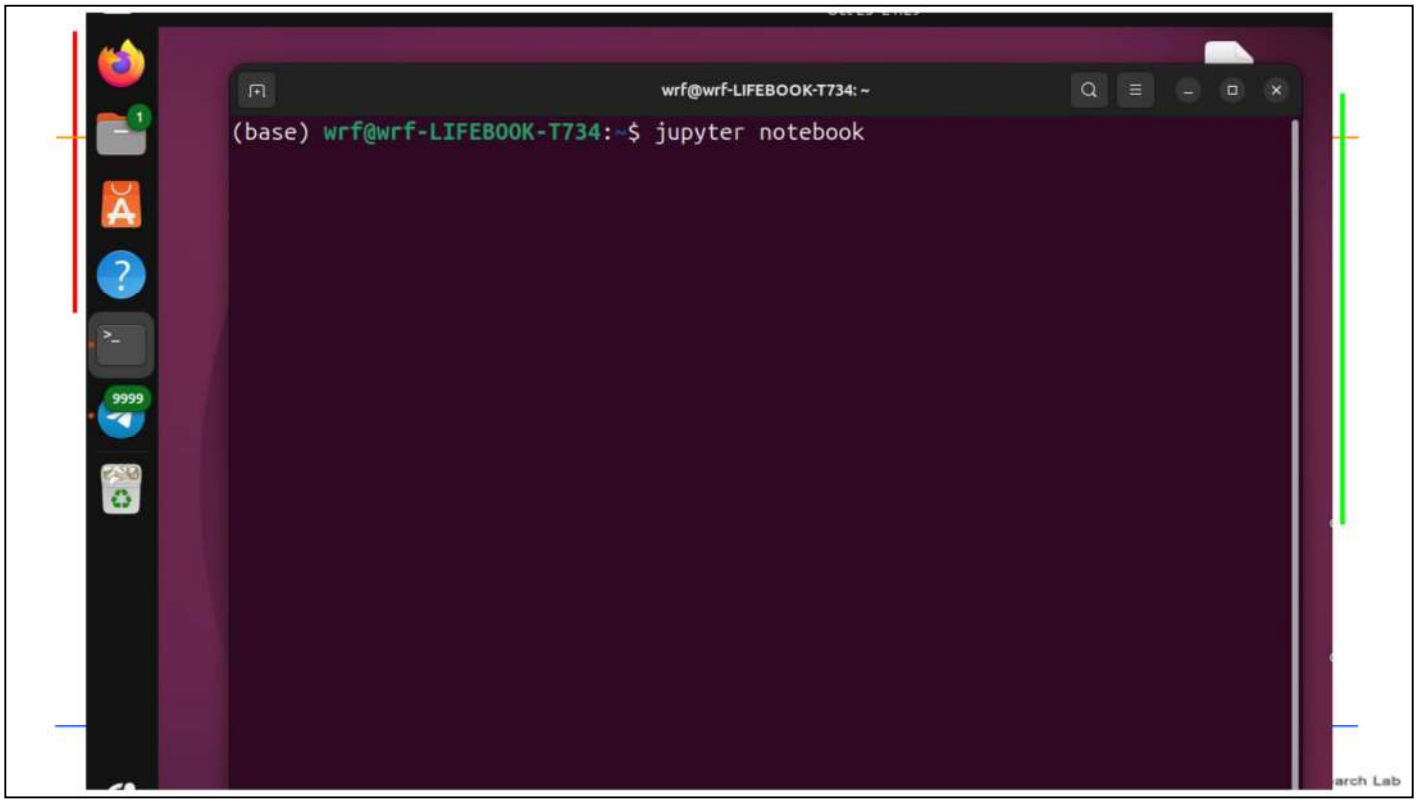
- **اختصار كلمة Jupyter**
- كلمة **Jupyter** هي اختصار غير رسمي يشير إلى ثلاث لغات برمجية رئيسية:
- **Julia**
- **Python**
- **R**
- وتُستخدم هذه اللغات بكثرة في علوم البيانات والحوسبة العلمية.

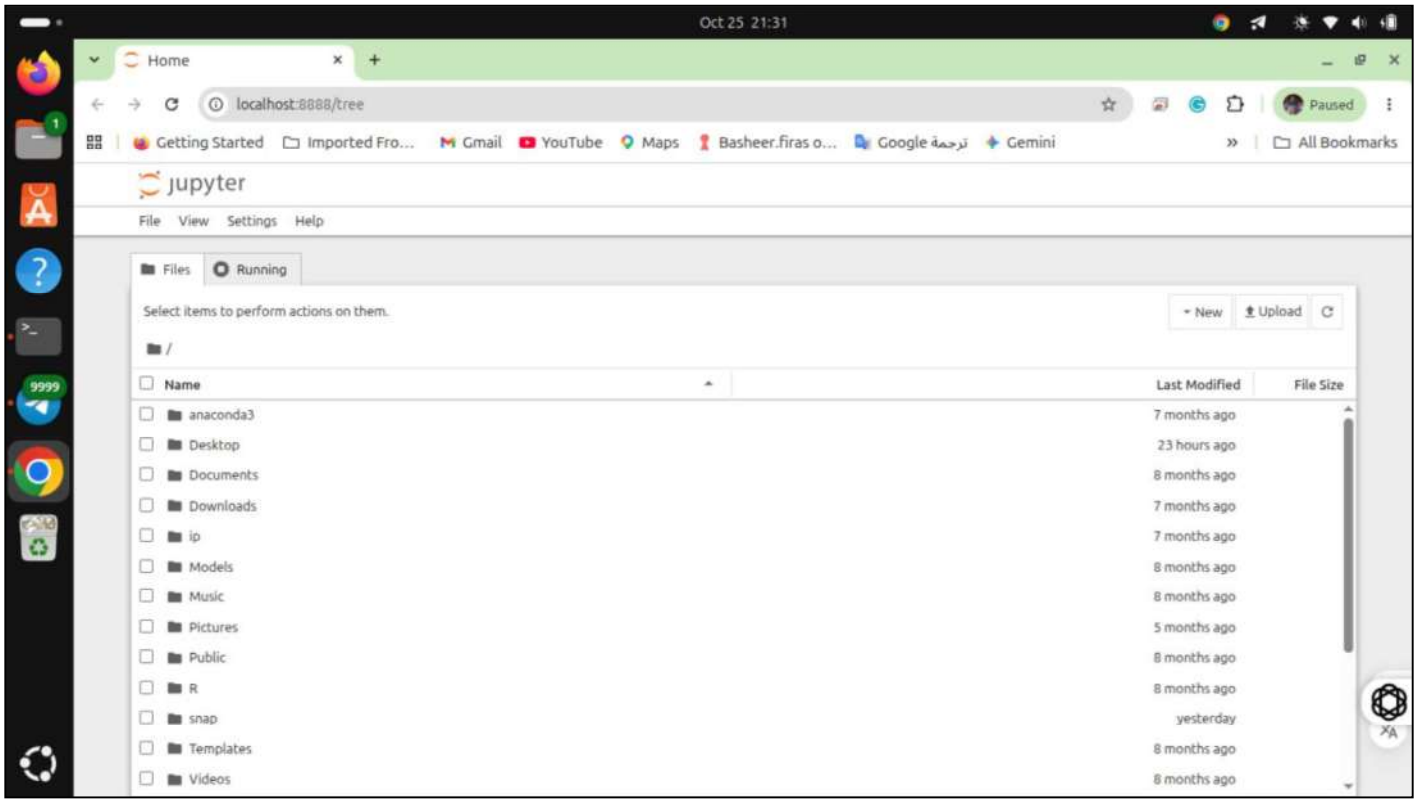
How to launch Jupyter Notebook

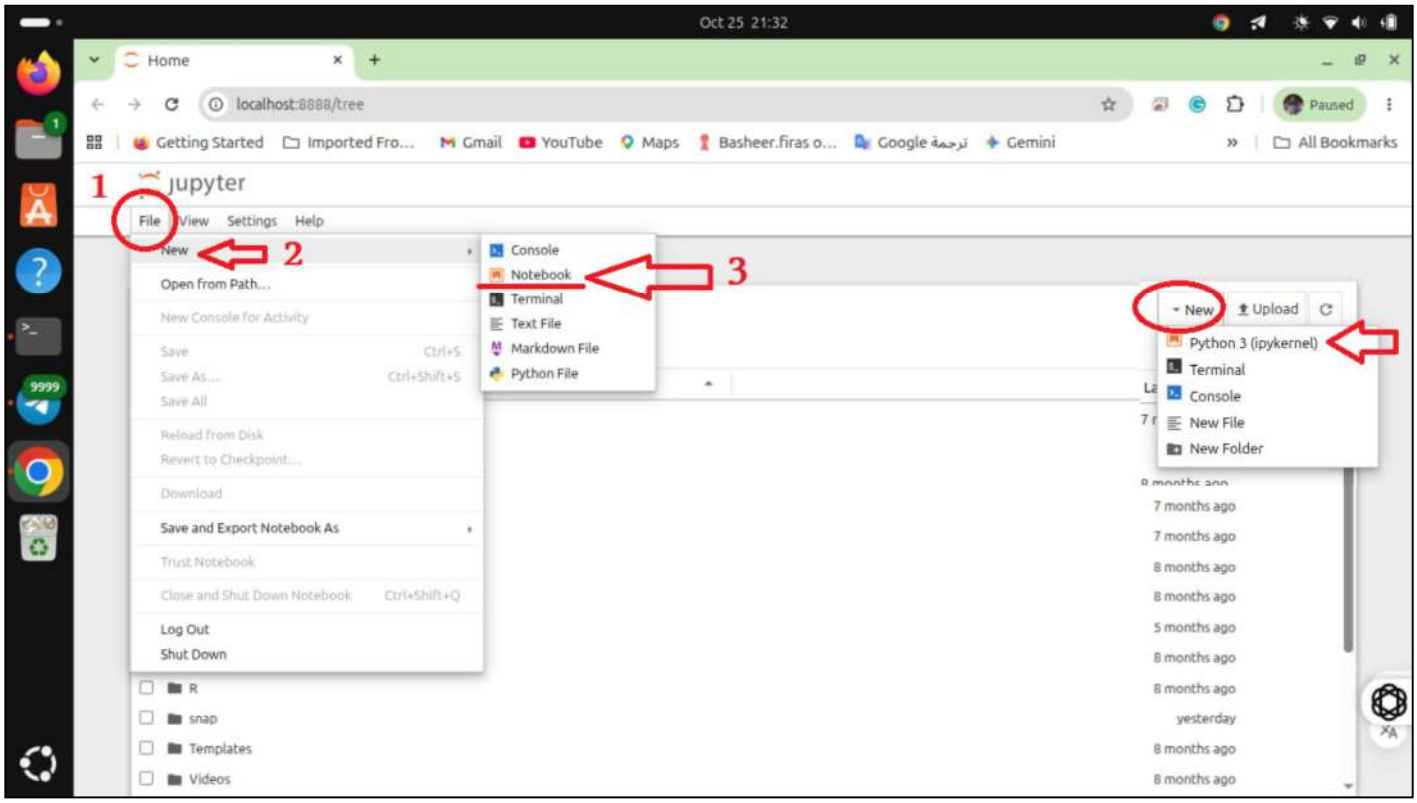


National Initiatives for Open Science in Europe – H2020 Research and Innovation action – contract no. 857645


wrf@wrf-LIFEBOOK-T734: ~
(base) wrf@wrf-LIFEBOOK-T734:~$ pip3 install jupyter

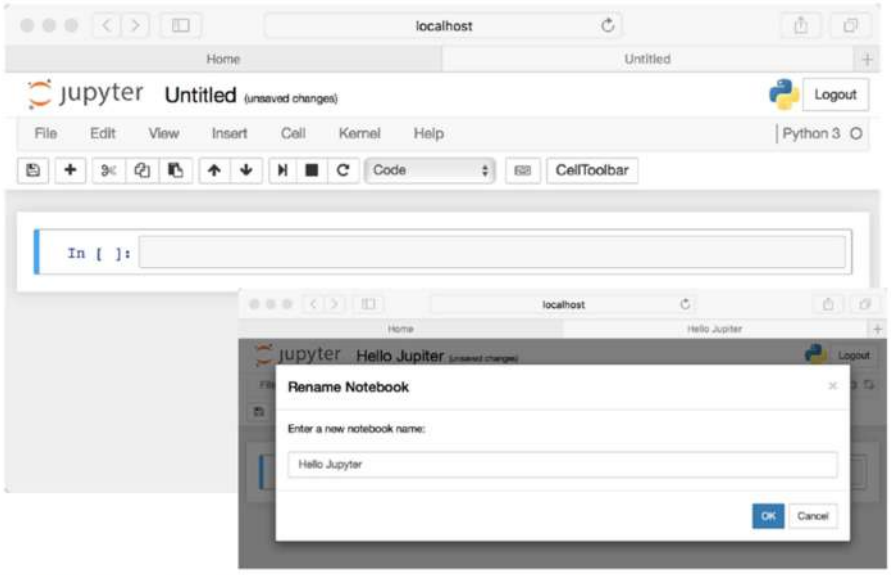

wrf@wrf-LIFEBOOK-T734: ~
(base) wrf@wrf-LIFEBOOK-T734:~$ jupyter notebook
9999


Oct 25 21:31
Home
localhost:8888/tree
Paused
Getting Started
Imported Fro...
Gmail
YouTube
Maps
Basheer.firas o...
Google ترجمة
Gemini
All Bookmarks
jupyter
File View Settings Help
Files
Running
Select items to perform actions on them.
New
Upload
/
Name
Last Modified
File Size
anaconda3 7 months ago
Desktop 23 hours ago
Documents 8 months ago
Downloads 7 months ago
ip 7 months ago
Models 8 months ago
Music 8 months ago
Pictures 5 months ago
Public 8 months ago
R 8 months ago
snap yesterday
Templates 8 months ago
Videos 8 months ago

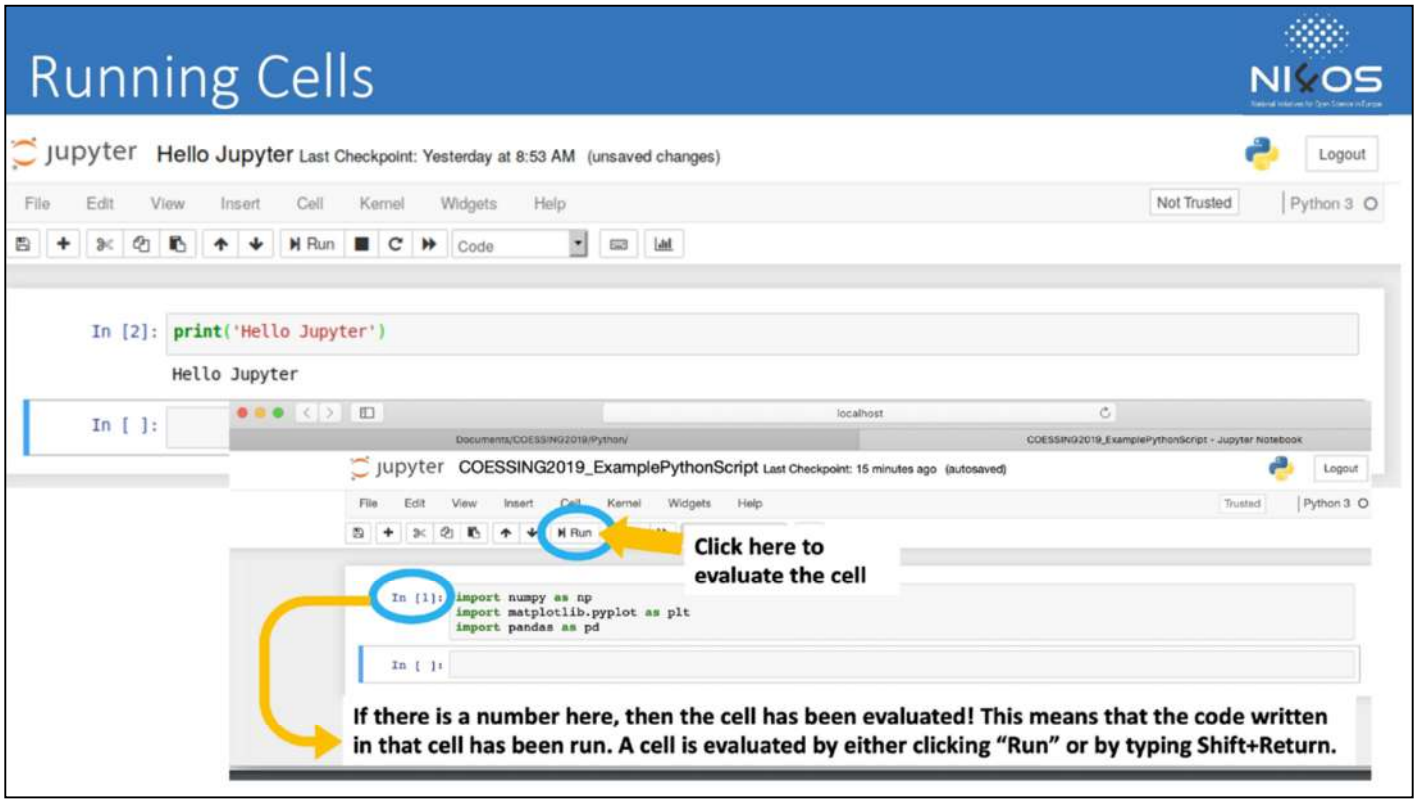

Oct 25 21:32
Home
localhost:8888/tree
Paused
Getting Started
Imported Fro...
Gmail
YouTube
Maps
Basheer.firas o...
Google ترجمة
Gemini
All Bookmarks
1
jupyter
File
View
Settings
Help
New
2
Open from Path...
New Console for Activity
Save
Ctrl+S
Save As...
Ctrl+Shift+S
Save All
Reload from Disk
Revert to Checkpoint...
Download
Save and Export Notebook As
Trust Notebook
Close and Shut Down Notebook
Ctrl+Shift+Q
Log Out
Shut Down
Console
Notebook
Terminal
Text File
Markdown File
Python File
3
New
Upload
Python 3 (ipykernel)
Terminal
Console
New File
New Folder
7 months ago
7 months ago
8 months ago
8 months ago
5 months ago
8 months ago
8 months ago
yesterday
8 months ago
8 months ago
R
snap
Templates
Videos

Creating Your First Notebook

- File-> New Notebook
- If you have Python 2 and Python 3 installed, you can create a Notebook that uses either of these
- If you switch back to the dashboard, you will see the new file Untitled.ipynb



National Initiatives for Open Science in Europe – H2020 Research and Innovation action – contract no. 857645



❑**الفكرة الأساسية**

❑في دفتر Jupyter، كل جزء من الكود يُكتب داخل **خلية .(Cell)**

❑لتشغيل الخلية)تنفيذ الكود المكتوب فيها(، لديك خياران:

➢الضغط على زر **Run** في الأعلى.

➢أو استخدام الاختصار **Shift + Enter** من لوحة المفاتيح.

❑**كيف تعرف أن الخلية تم تشغيلها؟**

❑إذا ظهر رقم بجانب الخلية)مثلاً [1]([، فهذا يعني أن الكود تم تنفيذه.

❑إذا لم يظهر رقم، فالخلية لم تُشغّل بعد.

❑**مثال عملي**

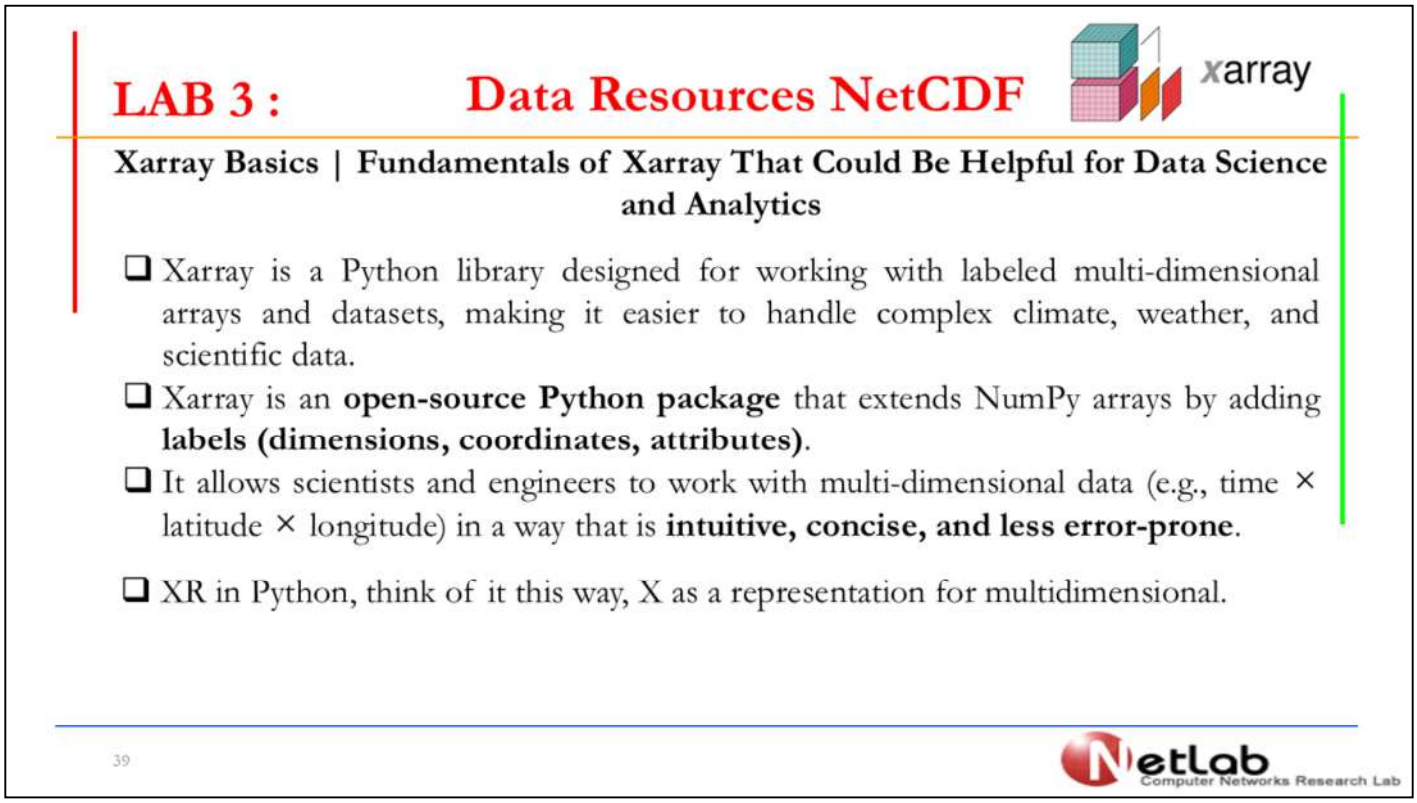
❑python

❑print("Hello Jupyter")

❑عند تشغيل هذه الخلية، سيظهر الناتج:

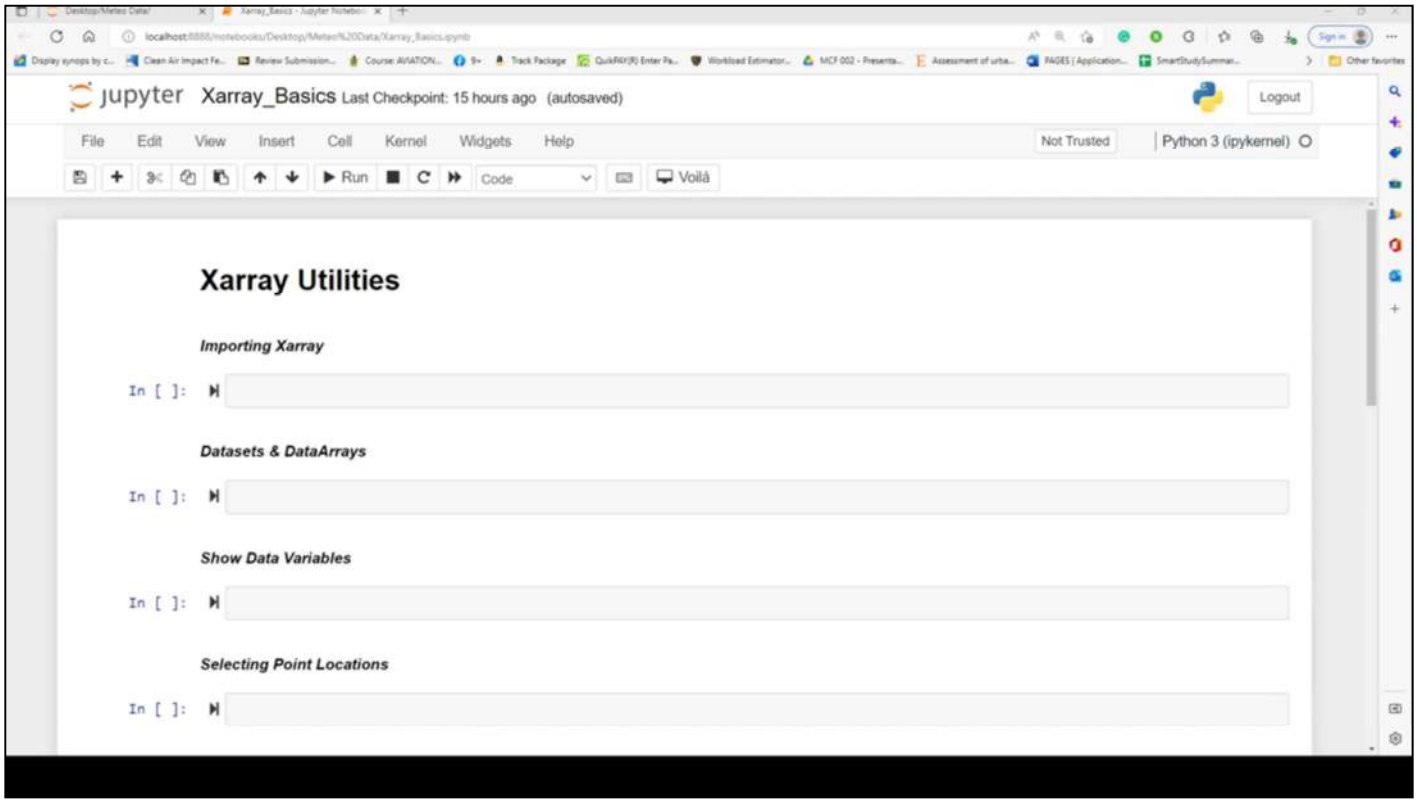
❑Code

❑Hello Jupyter



المبادئ الأساسية لـ Xarray لاستخداماتها في علم البيانات وتحليلها

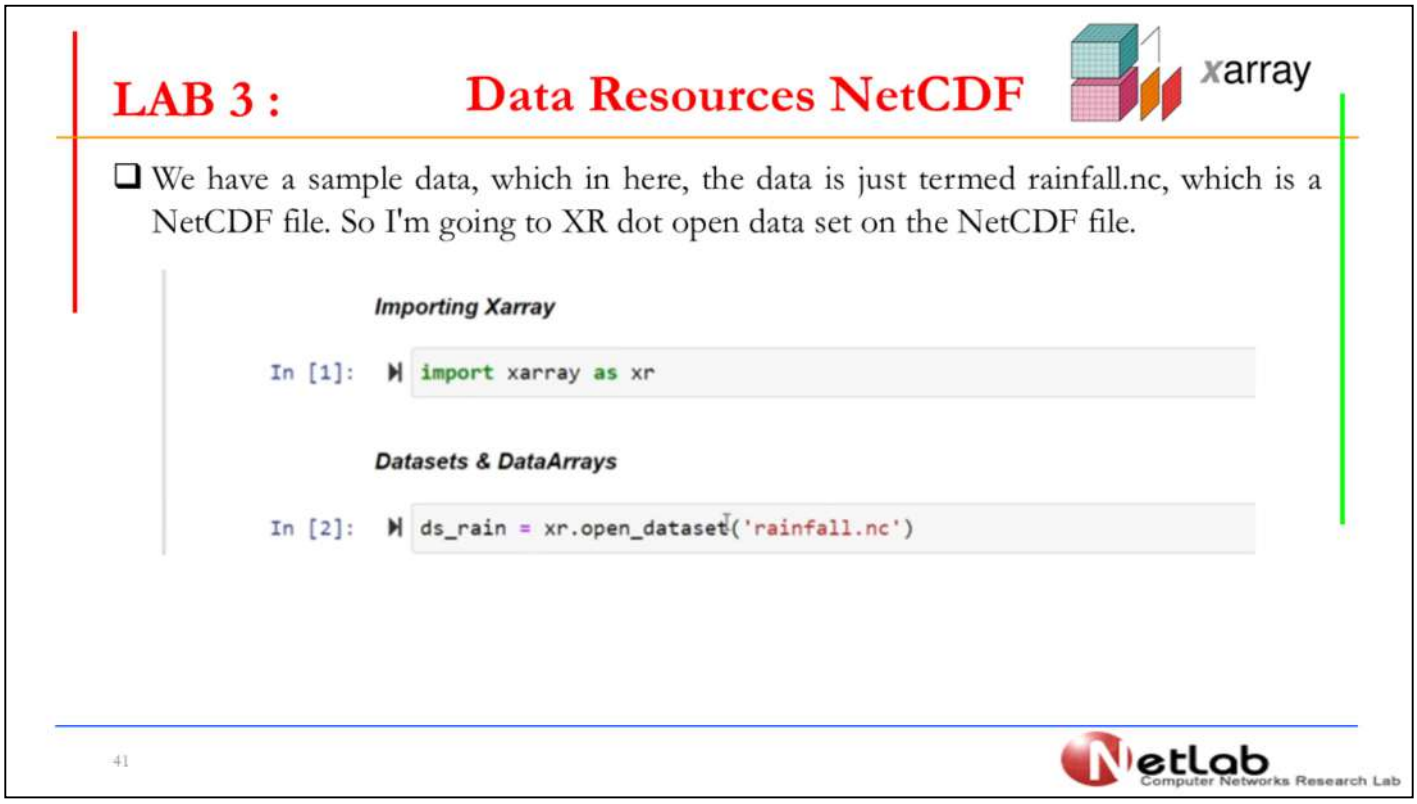
- **Xarray** هي مكتبة في لغة **Python** صُممت للعمل مع **المصفوفات والبيانات متعددة الأبعاد المعرفة لها** ، مما يجعل التعامل مع بيانات المناخ والطقس والبيانات العلمية المعقدة أكثر سهولة.
- تُعد **Xarray** حزمة مفتوحة المصدر في Python، وهي امتداد لمصفوفات **NumPy** بإضافة **تعريف (الأبعاد، الإحداثيات، الخصائص)**
- تتيح للعلماء والمهندسين العمل مع بيانات متعددة الأبعاد)مثل: **الزمن** × **خط العرض** × **خط الطول** (بطريقة **بديهية، موجزة، وأقل عرضة للأخطاء**.



Xarray Utilities

❑ادوات (Xarray)

1. **Importing Xarray** استيراد مكتبة (Xarray)
2. **Datasets & DataArrays** المجموعات والـ (DataArrays)
3. **Show Data Variables** عرض المتغيرات
4. **Selecting Point Locations** اختيار المواقع النقطية



نبدأ باستيراد حزمة XR، وهناك طرق عديدة لاستيرادها في بايثون .أستخدم طريقة بسيط ة جدًا، وهي استيراد اسم الحزمة، XR

Import xarray as xr

وعندما نريد استخدام حزمة فرعية، فإننا نستدعي الحزمة الأصلية متبوعةً بالحزمة الفرعية .في هذه الحالة، يمكن أن يكون الاسم XR متبوعًا باسم الحزمة الفرعية .ويمكننا أيضًا اس تخدام الأسماء المستعارة مثل XR , الأسماء المستعارة هي مجرد اختصارات للاسم الكامل .

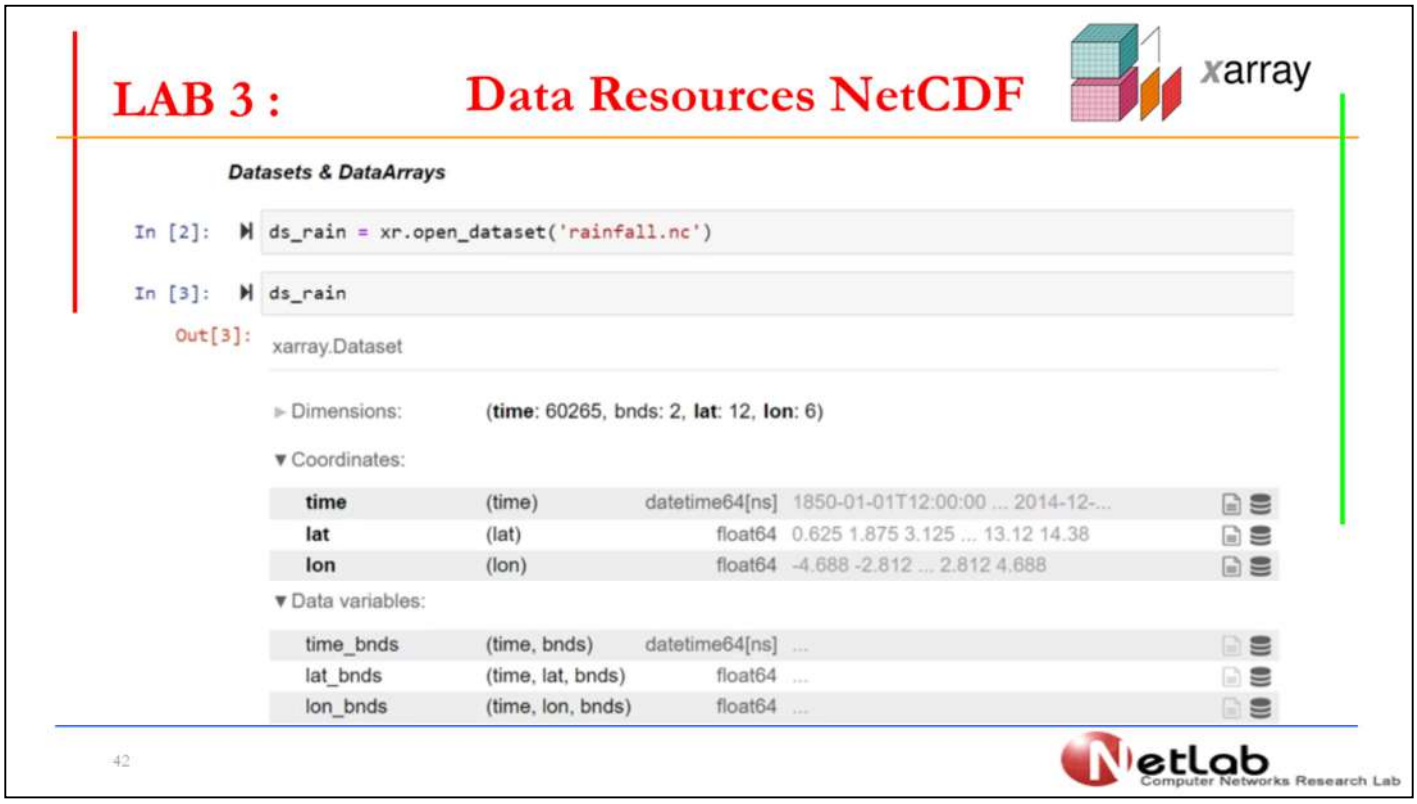
❑ استخدم مكتبة **Xarray (XR)**لفتح مجموعة البيانات الموجودة في ملف الـ **NetCD F:** (rainfall.nc)

ds = xr.open_dataset("rainfall.nc")

عند تنفيذ هذا الأمر، يتم فتح مجموعة البيانات وقراءة محتويات الملف، مما يتيح لنا استعرا ض المتغيرات والبيانات الوصفية metadata المخزنة بداخله وتحليلها بسهولة.

في العادة، نستخدم الاختصار ds للإشارة إلى مجموعة البيانات (Dataset) عند العمل باستخدام مكتبة Xarray.

لذلك يمكننا تخزين مجموعة بيانات هطول الأمطار في متغير باسم أكثر وضوحًا، مثل ds_rain: في هذا المثال، يشير المتغير ds_rain إلى مجموعة بيانات هطول الأمطار، مما يجعل الكود أكثر تنظيمًا وأسهل في الفهم، خاصة عند التعامل مع عدة مجموعات بيانات في البرنامج نفسه.



❑**عند عرض مجموعة البيانات، نلاحظ أنها تحتوي على عدد محدد من النقاط على كل بُعد .(Dimension)** فهناك عدد من النقاط على بُعد الزمن(Time) ، بالإضافة إلى بعدَ ي خطوط العرض (Latitude) وخطوط الطول .(Longitude) في البداية، لا تظهر ا لقيم الفعلية لهذه الإحداثيات بشكل مباشر ضمن ملخص البيانات.

❑عند الانتقال إلى قسم **الإحداثيات(Coordinates)** ، يمكن رؤية القيم الفعلية لكل بُعد بوضوح:

بُعد الزمن :(Time) يحتوي على قيم تاريخ ووقت بصيغة زمنية قياسية. على سبيل المثال ، تبدأ البيانات من:

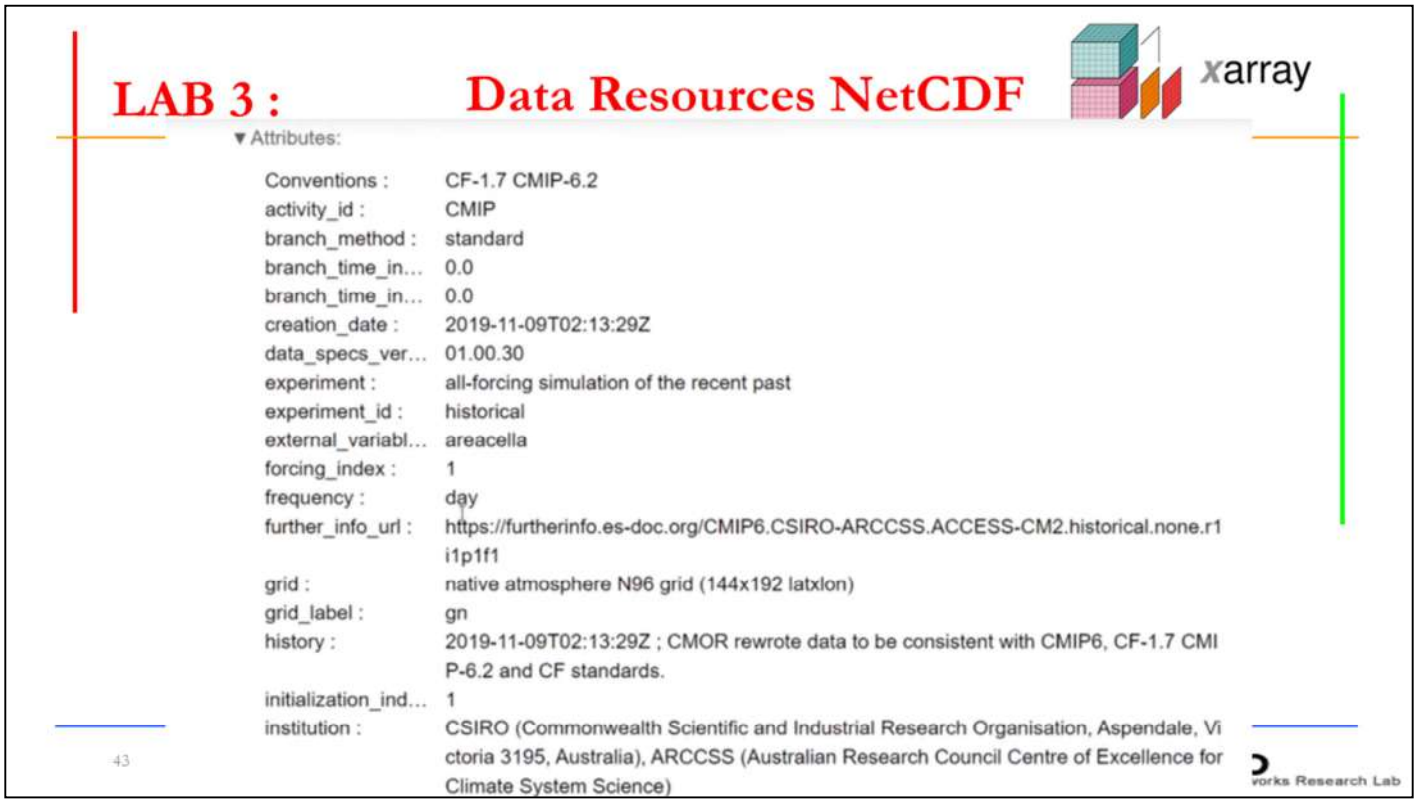
1يناير 1850، الساعة 12:00 بالتوقيت العالمي المنسق.(UTC)

❑عند توسيع قائمة الزمن، يمكن استعراض المزيد من التواريخ والأوقات المخزنة في مج موعة البيانات.

❑**خطوط العرض :(Latitude)** تعرض القيم العددية الدقيقة لمواقع خطوط العرض الم ستخدمة في البيانات.

❑**خطوط الطول :(Longitude)** تعرض كذلك القيم العددية الدقيقة لمواقع خطوط الطول .

❑أما **الحدود(Bounds)** ، فهي تمثل حدود الفترات الزمنية المرتبطة بالبيانات، وعادةً لا تكون ضرورية في التحليل الأولي،



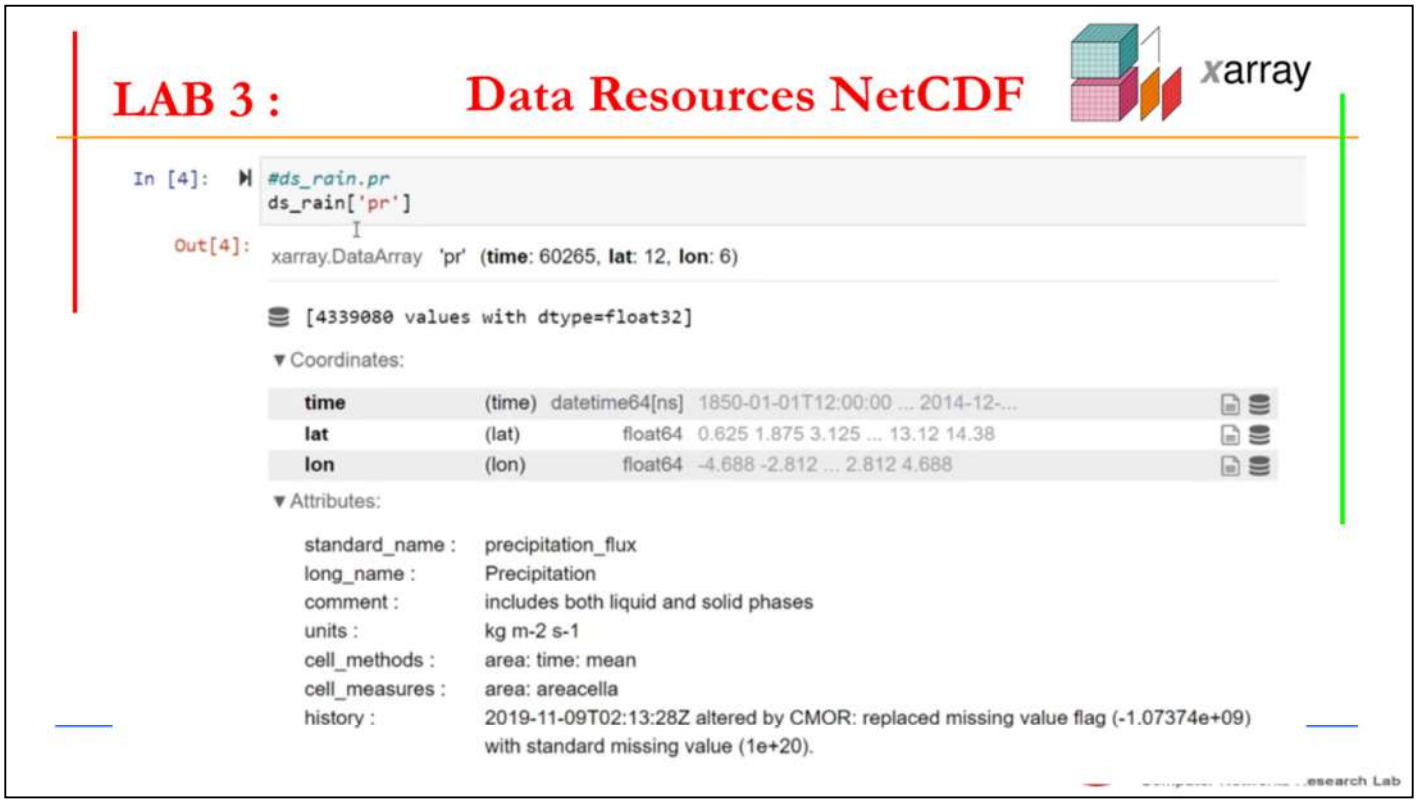
تحتوي مجموعة البيانات على متغيرات بيانات ((Data Variables، وهي المتغيرات ال تي تتغير قيمها وفقًا للأبعاد ((Dimensionsأو الإحداثيات ((Coordinatesالمرتبط ة بها.بمعنى آخر، يرتبط كل متغير بواحد أو أكثر من الأبعاد، وتتغير قيمه تبعًا لتغير هذه ا لأبعاد.

في مثالنا، المتغير الرئيس هو معدل هطول الأمطار ((Precipitation Rate، والذي يع تمد على ثلاثة أبعاد أساسية:

الزمن (Time)خط العرض (Latitude)خط الطول (Longitude)

لذلك فإن قيمة هطول الأمطار تتغير من وقت إلى آخر، كما تختلف من موقع جغرافي إلى آ خر وفقًا لخطوط العرض والطول.بالإضافة إلى متغيرات البيانات، تحتوي مجموعة البيانا ت على قسم الخصائص (Attributes)، والذي يوفر معلومات وصفية مهمة عن مجموع ة البيانات، مثل:

- كيفية إنتاج البيانات ومعالجتها.
- مصدر البيانات والنموذج المستخدم.
- الوحدات المستخدمة للمتغيرات. (جدا مهمة الوحدات)
- إجراءات وضوابط مراقبة الجودة .(Quality Control)
- معلومات إضافية تساعد على فهم البيانات واستخدامها بشكل صحيح.
- وتُعد هذه الخصائص مصدرًا مهمًا للمعلومات المرجعية التي تساعد الباحث على تفس ير البيانات وتحليلها بصورة دقيقة.



بعد فتح مجموعة البيانات، يمكننا الوصول إلى متغير هطول الأمطار (pr)بطريقتين شائعتين في Xarray:

1.استخدام طريقة النقطة (Dot Notation)

ds_rain.pr

في هذه الطريقة نكتب اسم مجموعة البيانات، ثم نقطة (.)، ثم اسم المتغير.

2.استخدام الأقواس المربعة (Bracket Notation)

ds_rain['pr']

عند استخدام الأقواس المربعة، يجب كتابة اسم المتغير كسلسلة نصية (String)بين علامتي اقتباس مفردة (' ')أو مزدوجة (" ").

عند استدعاء المتغير prبإحدى هاتين الطريقتين، ستلاحظ أن الناتج **لم يعد (Datasetمجموعة بيانات)**، بل أصبح (**DataArrayمصفوفة بيانات.)**

وذلك لأن مكتبة **Xarray** تمثل كل متغير منفرد داخل مجموعة البيانات على هيئة **DataArray** ، بينما تمثل مجموعة من المتغيرات معًا على هيئة **Dataset**.

بمعنى آخر:

Dataset: حاوية تحتوي على عدة متغيرات (DataArrays)وإحداثيات وخصائص.

DataArray: متغير واحد مع أبعاده وإحداثياته وخصائصه المرتبطة به.

في مثالنا، تمثل prبيانات الهطول، وهي مرتبطة بالأبعاد التالية:

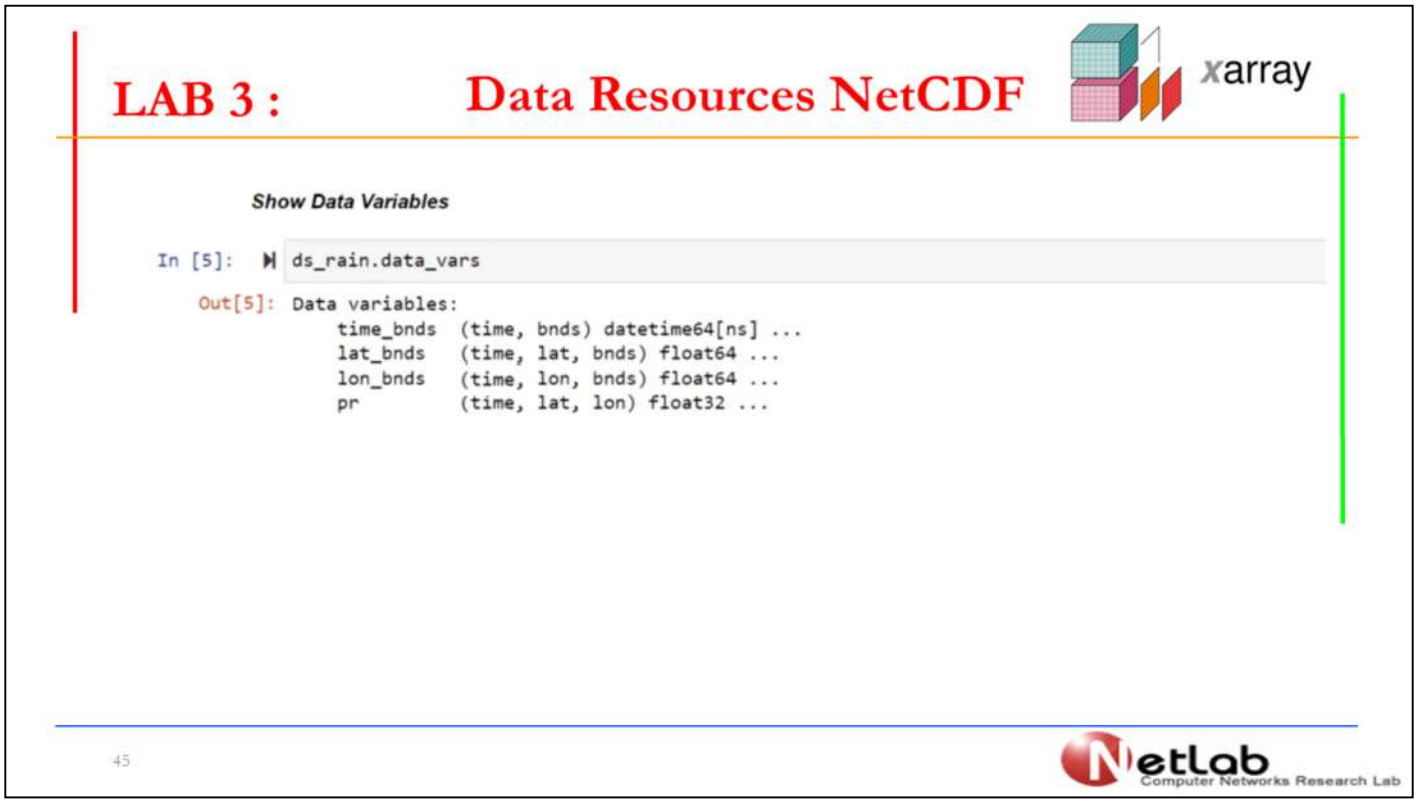
الزمن (**Time**),خط العرض (**Latitude**),خط الطول (**Longitude**)

كما تحتوي أيضًا على:

الإحداثيات (Coordinates),السمات أو الخصائص (Attributes)

لذلك فإن الأمر ds_rain['pr']:

يختار متغير الهطول من مجموعة البيانات ويُرجعه كمصفوفة بيانات (**DataArray**)جاهزة للتحليل والمعا لجة.



يمكننا أيضًا استعراض متغيرات البيانات (Show Data Variables) الموجودة داخل مجموعة البيانات باستخدام الخاصية :data_vars

ds_rain.data_vars

يعرض هذا الأمر جميع متغيرات البيانات الموجودة في مجموعة البيانات (Dataset)دو ن إظهار الإحداثيات (Coordinates)أو الخصائص .(Attributes)

وهذه طريقة سريعة للتعرف على المتغيرات التي تحتويها مجموعة البيانات، خصوصًا عند التعامل مع ملفات NetCDFكبيرة تضم عددًا من المتغيرات المختلفة.

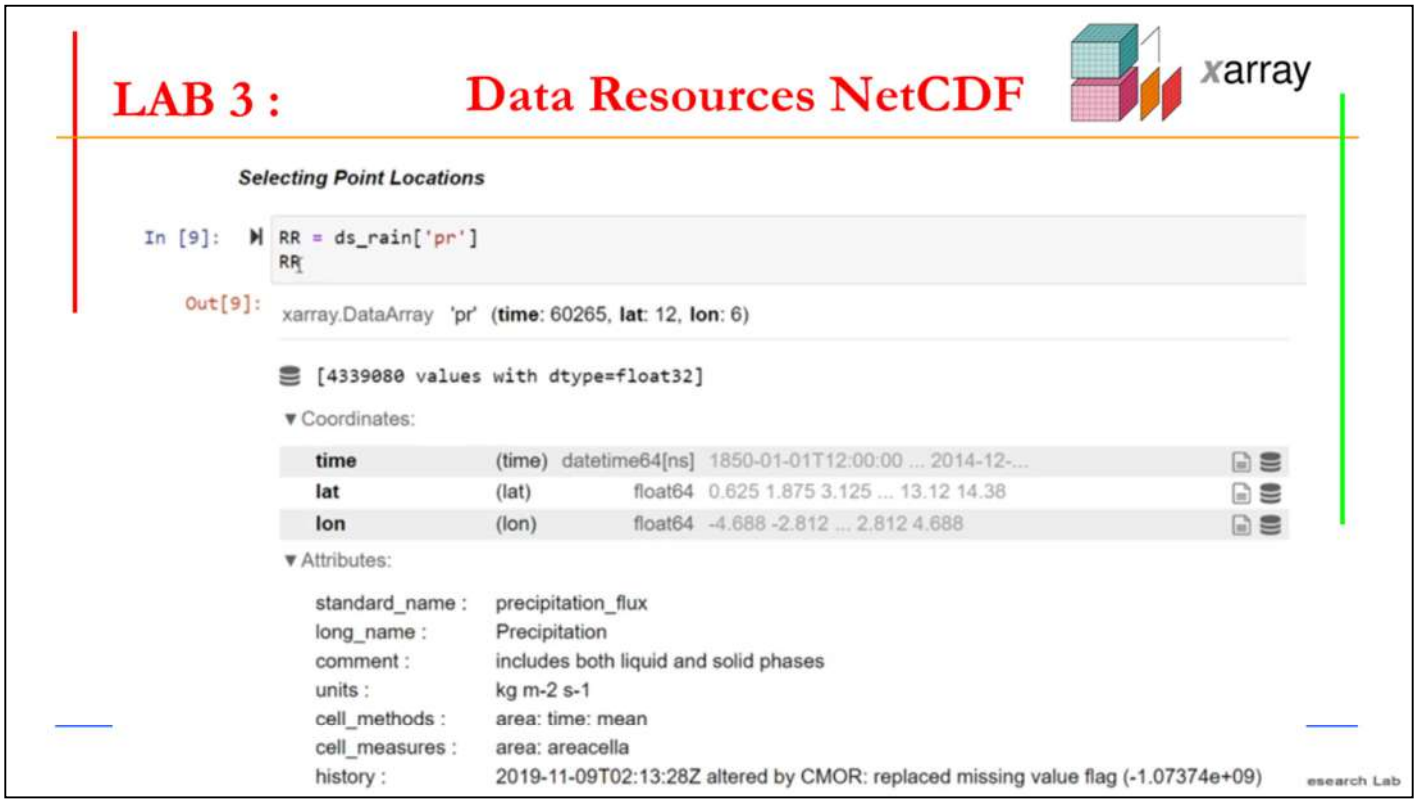
على سبيل المثال، إذا كانت مجموعة البيانات تحتوي على متغير واحد فقط للهطول، ومن ه ذا العرض يمكننا معرفة:

- اسم المتغير: pr(معدل هطول الأمطار).
- الأبعاد المرتبط بها المتغير: time(الزمن), lat(خط العرض), lon (خط الطول)
- نوع البيانات المخزنة (مثل .(float32

وبذلك يُعد الأمر:

ds_rain.data_vars

وسيلة سريعة لاستكشاف محتويات مجموعة البيانات ومعرفة المتغيرات المتاحة للتحليل ق بل اختيار أي متغير والعمل عليه.



❑**لتسهيل التعامل مع متغير هطول الأمطار، يمكننا تخزينه في متغير جديد يحمل اسمًا مختصرًا و واضحًا، مثل RRالذي يرمز إلى) Rainfall Rateمعدل هطول الأمطار:(**

RR = ds_rain.pr

أو بصورة مكافئة:

RR = ds_rain['pr']

❑بعد تنفيذ هذا الأمر، يصبح المتغير RRعبارة عن **مصفوفة بيانات (DataArray)**تحتوي ع لى بيانات الهطول فقط، بدلًا من التعامل مع مجموعة البيانات الكاملة .**(Dataset)**

❑يُسهّل هذا الأسلوب كتابة الأوامر اللاحقة وقراءة الشيفرة البرمجية، خاصةً عند إجراء عمليات ا لتحليل أو الرسم البياني .فعوضًا عن كتابة:

ds_rain.pr

❑في كل مرة، يمكننا ببساطة استخدام:

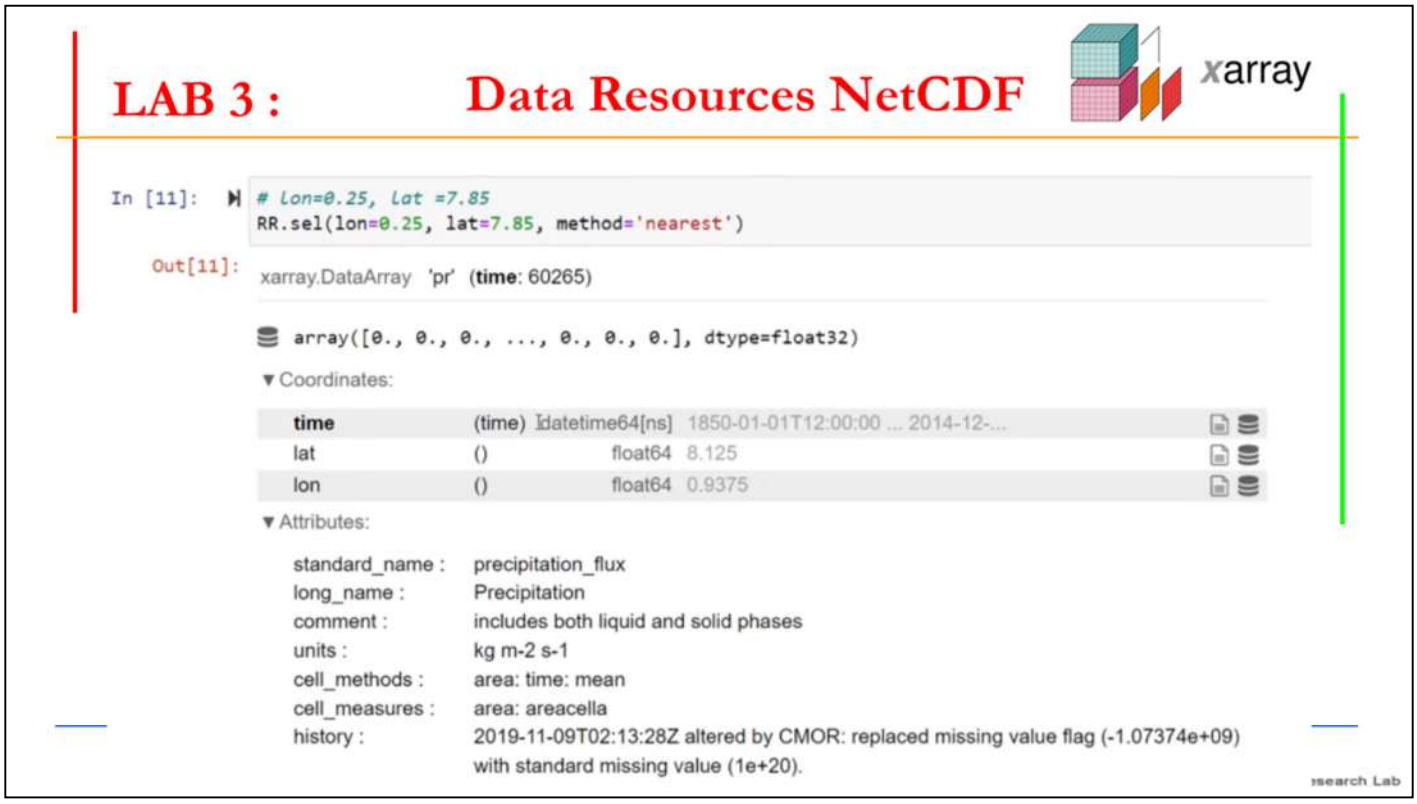
RR❑

❑كما توفر كائنات **DataArray**العديد من الخصائص والدوال التي يمكن استكشافها لمعرفة م علوماتها وأبعادها وإحداثياتها والعمليات الممكن تطبيقها عليها.

❑وباختصار:

RR = ds_rain.pr❑

❑يعمل على استخراج متغير الهطول (pr)من مجموعة البيانات ds_rainوتخزينه في المتغير RR، والذي يمثل الآن مصفوفة البيانات الخاصة بمعدل هطول الأمطار وجاهزة للاستخدام في ال تحليل والمعالجة.



اختيار نقطة مفردة واستخراج السلسلة الزمنية

الآن يمكننا اختيار نقطة جغرافية مفردة من مجموعة البيانات لأن لدينا بيانات مكانية أكثر تفصيلاً.لنفترض أننا نريد استخراج بيانات الهطول عند موقع محدد إحداثياته:

خط الطول (Longitude) = **0.25**

خط العرض (Latitude) = **7.85**

فإننا نستطيع اختيار هذه النقطة باستخدام:

RR.sel(lon=0.25, lat=7.85)

لكن قد لا تكون هذه الإحداثيات موجودة تمامًا ضمن نقاط الشبكة الخاصة بمجموعة البيانات .في هذه الحالة سيظهر خطأ (KeyError)لأن البرنامج لن يجد النقطة المطلوبة.

ولحل هذه المشكلة يمكن استخدام الخيار:

RR.sel(lon=0.25, lat=7.85, method='nearest')

حيث يقوم method='nearest' بالبحث عن أقرب نقطة شبكية إلى الإحداثيات المطلو بة، وإذا لم يجد النقطة بالضبط فإنه يختار أقرب نقطة متاحة تلقائيًا.

بهذه الطريقة نحصل على بيانات الهطول عند أقرب موقع ممكن للإحداثيات المطلوبة، ونس تخرج السلسلة الزمنية (Time Series)لذلك الموقع الثابت.

والنتيجة هي بيانات الهطول لجميع الأزمنة عند تلك النقطة المختارة، أي سلسلة زمنية تمثل تغير الهطول مع الزمن في ذلك الموقع.