Enabling a Data-informed Public Sector: An Introductory Course to BDTI Essentials

Session 1: Data Access and Exploration

Welcome!
The webinar will begin shortly.

EC-BDTI-PILOTS@ec.europa.eu

DG CNECT
Directorate-General for Communications
Networks, Content and Technology

DG DIGIT
Directorate-General
for Digital Services

European
Commission

Some housekeeping

- The Chat and Q&A widgets are located to the right of the presenter screen
- Please use the Chat during the session, but wait until the dedicated Q&A portion to post your questions
- Microphones and video are off by default, attendees cannot turn them on
- The session will be recorded and available shortly after the end of the session on our website

Do

- Participate in the chat during the session
- Ask questions during the Q&A portion at the end
- Keep your questions concise
- Give feedback after the session
- Join the dedicated discussion board

Don’t

- Use chat as a watercooler (no personal conversations)
- Self-promote
- Disrespect anyone

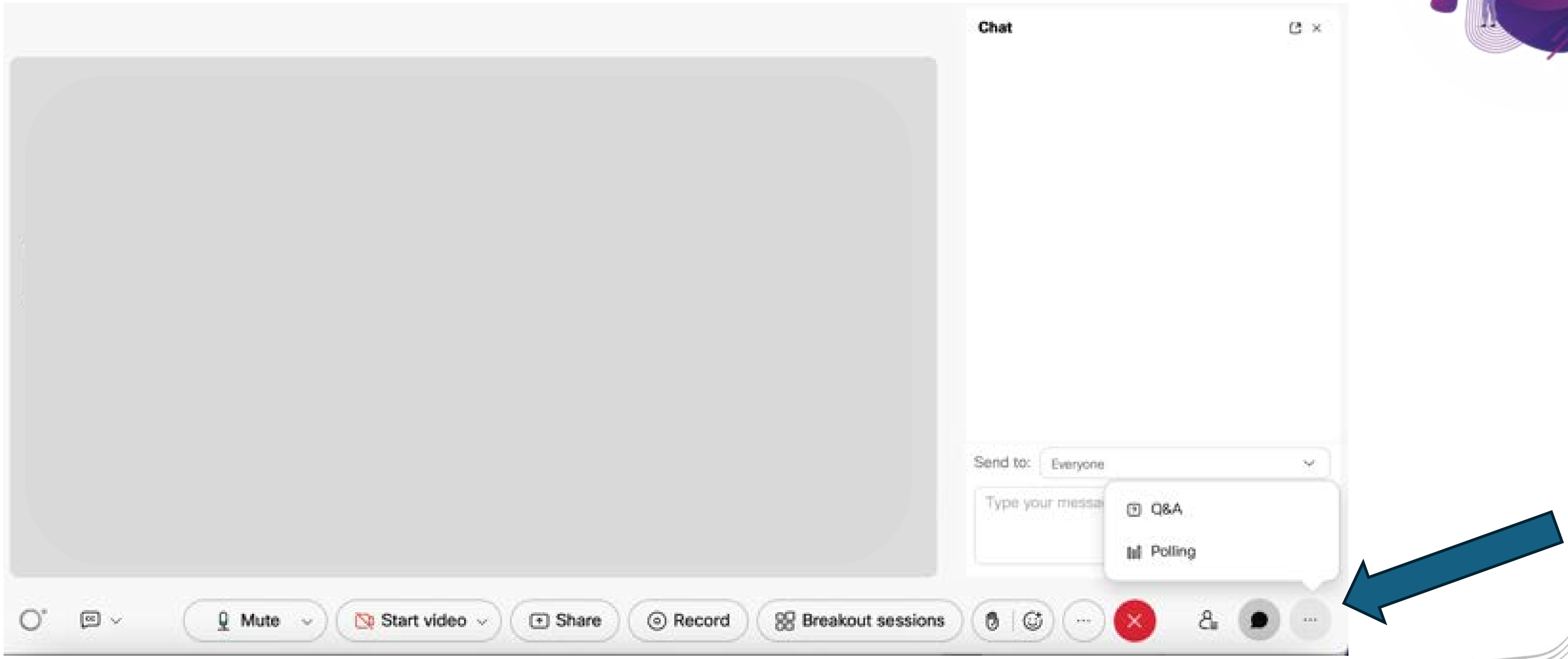
Enabling a Data-informed Public Sector: An Introductory Course to BDTI Essentials

Session 1: Data Access and Exploration

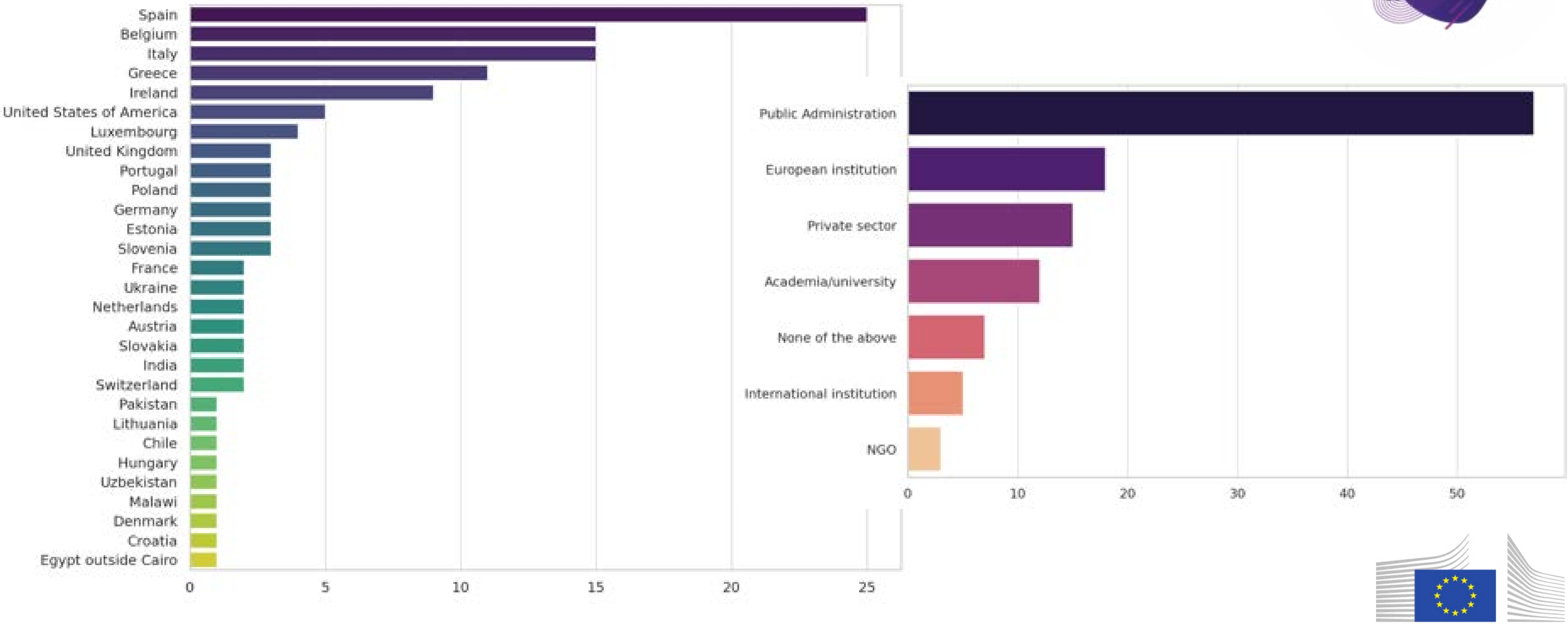
Agenda

- Welcome [11:00-11:05]
- The Big Data Test Infrastructure in nutshell [11.05-11.10]
- Data.Europa.eu: where to find EU open data for your projects [11.10-11.15]
- Data access and Exploration [11:15-12:00]
- Q&A [12:00-12:15]

Locate Chat and Q&A - Webex



Our data enthusiasts ☺



Intro

- The data and AI revolution: an increasing demand for professionals equipped with the skills to analyse and extract intelligence
- To draw attention to the new skills that the digital and green transition require, the European Commission has declared **2023 the European Year of Skills** (from May 2023 to May 2024)
- The Digital Europe Programme: Supports public sector in acquiring the necessary skills to extract value from data and new technologies
- The Big Data Test Infrastructure: A powerful tool offered to all EU public administrations
- This course is about **getting familiar with Data analytics**, building a concrete use case
- After this course, you will be ready to apply for BDTI and build a public sector data use case using the platform

The Data Analysis Process is like cooking a dish

Business challenge/question/problem statement

Ingredients + tools

1. **Find** the data you need → *gather the right ingredients, good quality*
2. **Get, clean** and **prepare** your data → *slice and dice*
3. **Analyse** your data → *mix ingredients together and try different combinations*
4. **Present** the results and **create** knowledge → *serve and consume*

Data→ Information → Knowledge

What is the Big Data Test Infrastructure (BDTI) ?

Not **only** for big data, for **public sector in general (open data)**

You have the key ingredients (datasets),
we equip you with the best **open-source tool**
to create amazing recipes for **public good.**

Six months free of charge service for EU public administrations *

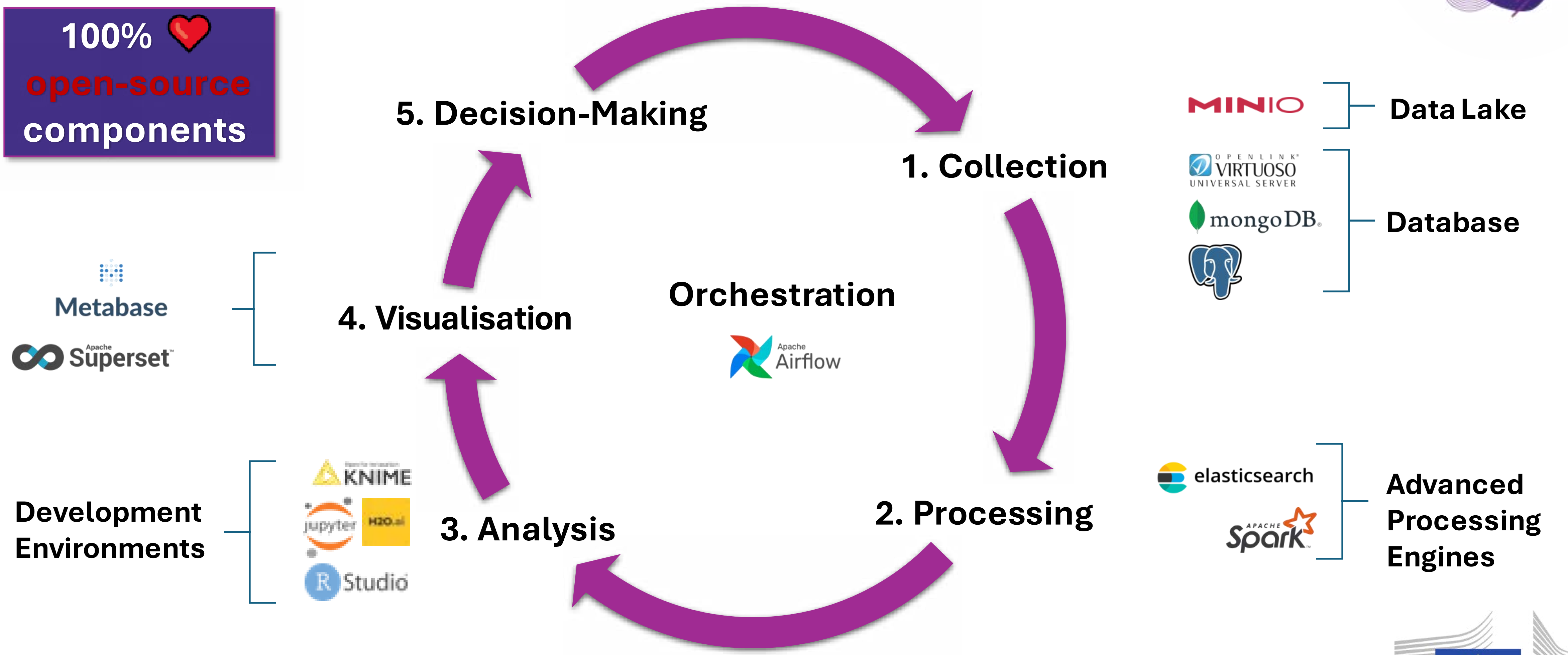
Ready-to-use data analytics stack and support

Cloud platform based on **open-source** tools

* The cost of the pilot project must fit within the funding boundaries of the BDTI pilot budget

Open-source tools to support your data journey

100% ❤
open-source
components



1. Collection

2. Processing

3. Analysis

4. Visualisation

5. Decision-Making

Orchestration

Data Lake

Database

Advanced Processing Engines

Development Environments

Who is the Big Data Test Infrastructure (BDTI) for?

European Public Administrations
All European Public Administrations at **local, regional and national level** can independently apply for a BDTI pilot project

Ecosystem with academia and private sector
Academia, spin-off, startups can apply for pilot projects once there is a **clear collaboration** with a Public Administration which will be the main point of contact for the project (**Master/PhD, GovTech startups)**

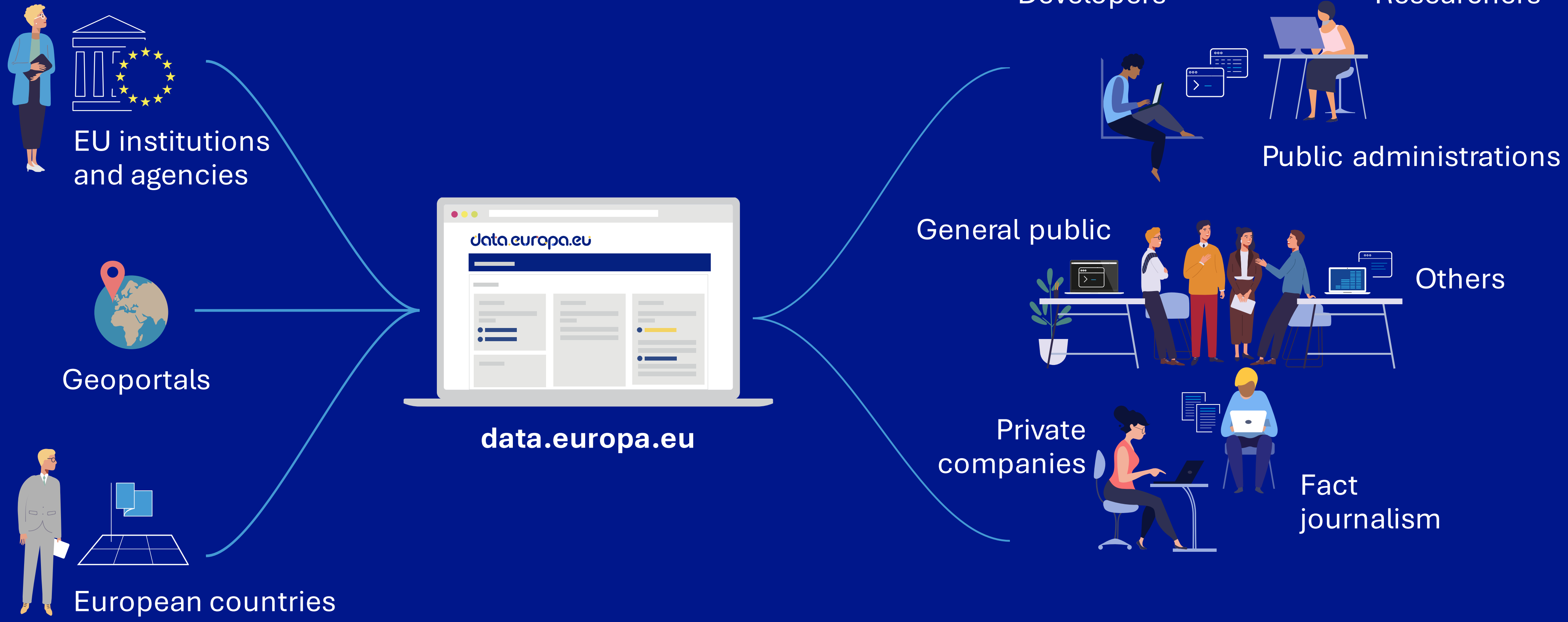
European Commission

data.europa.eu - your data source

Els Breedstraet
Publications office of the EU - Open data reuse and innovation

data.europa.eu

The portal is a bridge between the data providers and data users



Data providers

EU institutions and agencies

Geoportals

European countries

data.europa.eu

Data users

Developers

Researchers

Public administrations

General public

Others

Private companies

Fact journalism

data.europa.eu

Our services in a nutshell

Data

Providing access to free public data resources across Europe via a single platform (the portal).

Academy

Supporting the EU institutions, EU Member States and European countries to publish metadata of their open data in a harmonised manner.

Community

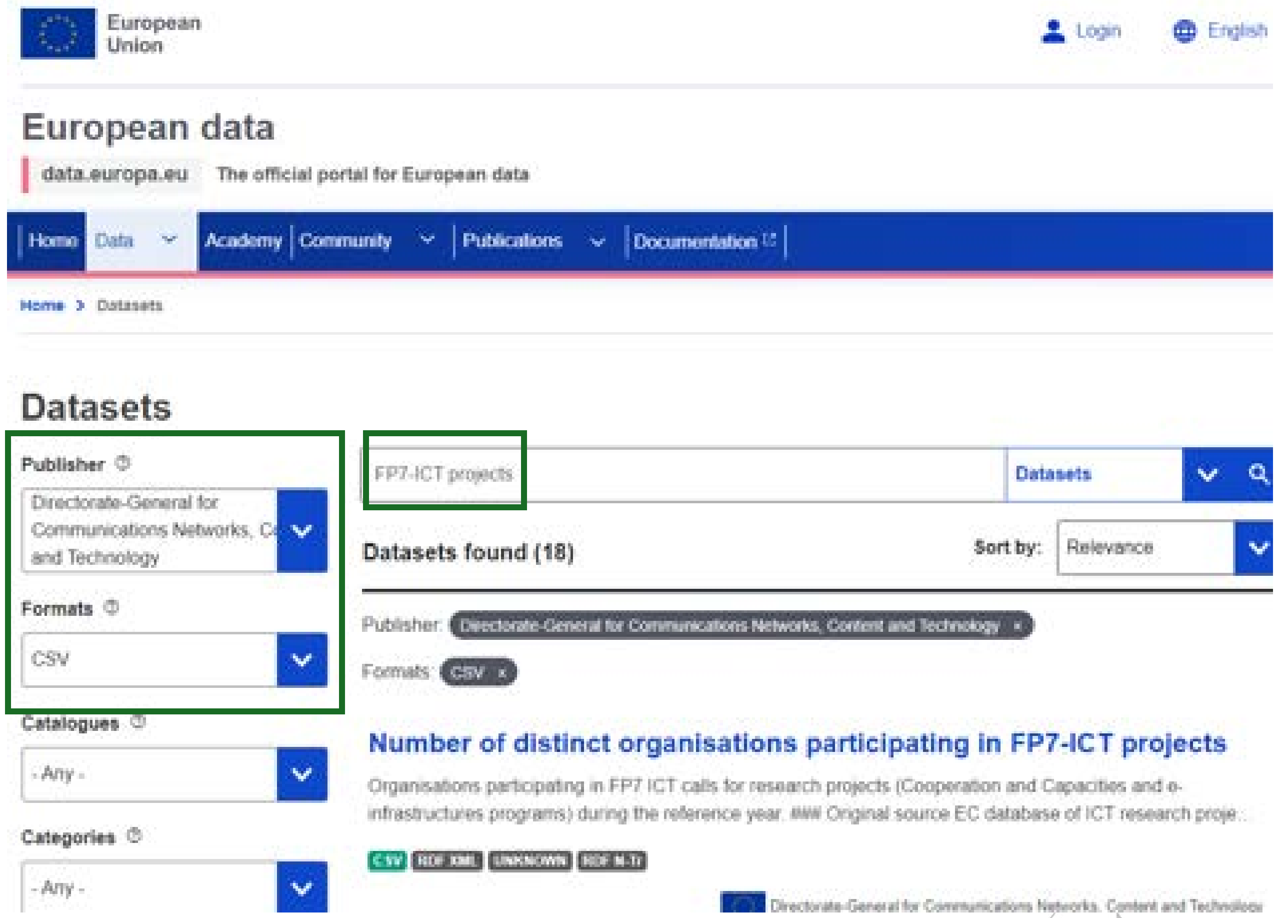
Organising open data competitions and conferences; communicating via social media and newsletters.

Publications

Assessing open data maturity in Europe; providing reports, studies and training via data.europa academy.

Data.europa.eu as data hub

- **More than 1.6 million datasets**, grouped in 183 data catalogues
- **Navigate or search to get to the data** or catalogue you are looking for
- Benefit from **many filters**
- **Metadata translations** in all EU languages, machine translations for other text
- **Download and transform CSV files** automatically in many different formats
- Get **quick visualisations** for geo datasets
- Get **feedback for the metadata quality** on how to improve it



European data

data.europa.eu The official portal for European data

Datasets

Publisher: Directorate-General for Communications Networks, Co… and Technology

Formats: CSV

FP7-ICT projects

Datasets found (18)

Sort by: Relevance

Publisher: Directorate-General for Communications Networks, Content and Technology ×

Formats: CSV ×

Number of distinct organisations participating in FP7-ICT projects

Organisations participating in FP7 ICT calls for research projects (Cooperation and Capacities and e-infrastructures programs) during the reference year. ### Original source EC database of ICT research proje…

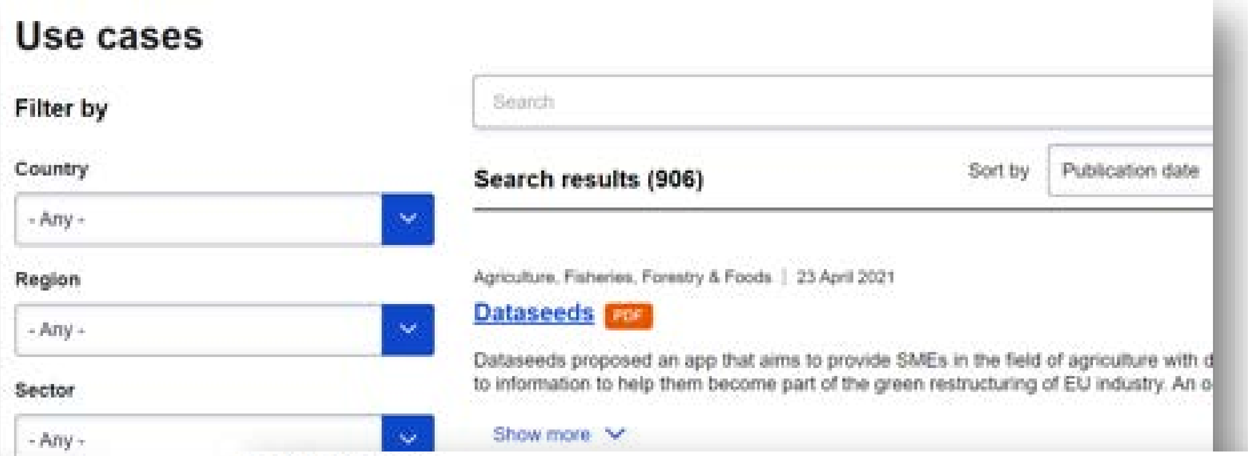
CSV RDF XML UNKNOWN RDF N-TI

Directorate-General for Communications Networks, Content and Technolog

European Commission

Data.europa.eu: get inspired

- Consult our list of use cases (more than 900 examples)
- Follow re-users through the **use case observatory**
- Read our data stories



Use cases

Filter by

Country

- Any -

Region

- Any -

Sector

- Any -

Use Case

Search

Search results (906)

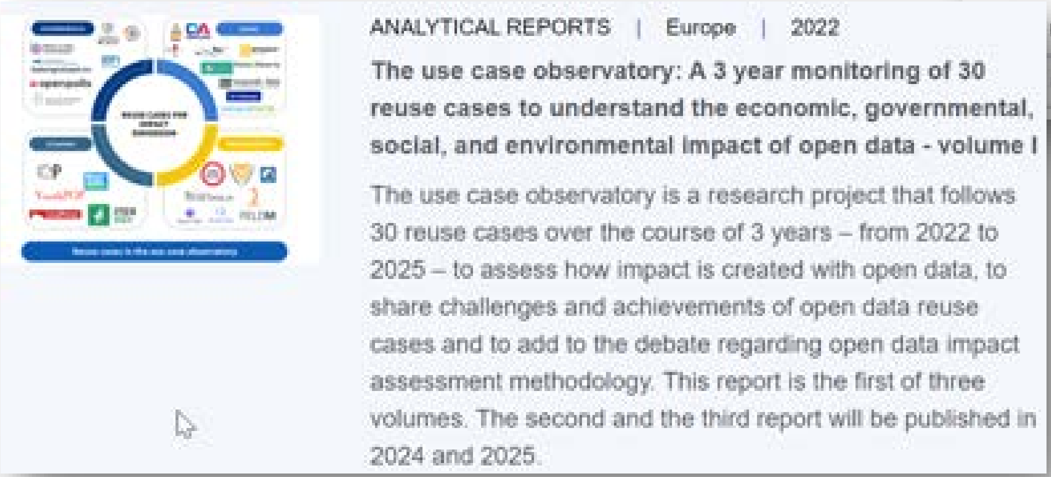
Sort by Publication date

Agriculture, Fisheries, Forestry & Foods | 23 April 2021

Dataseeds PDF

Dataseeds proposed an app that aims to provide SMEs in the field of agriculture with d to information to help them become part of the green restructuring of EU industry. An o

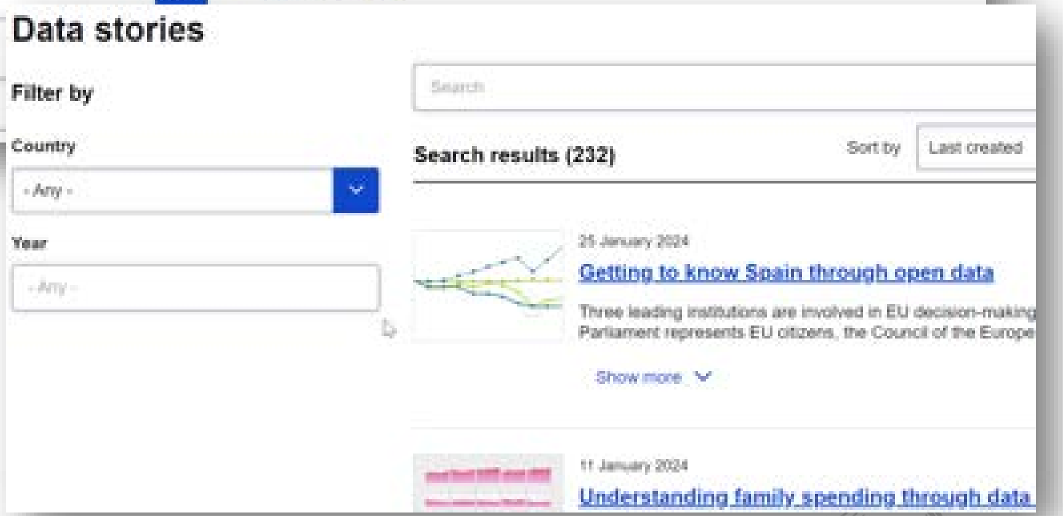
Show more



ANALYTICAL REPORTS | Europe | 2022

The use case observatory: A 3 year monitoring of 30 reuse cases to understand the economic, governmental, social, and environmental impact of open data - volume I

The use case observatory is a research project that follows 30 reuse cases over the course of 3 years – from 2022 to 2025 – to assess how impact is created with open data, to share challenges and achievements of open data reuse cases and to add to the debate regarding open data impact assessment methodology. This report is the first of three volumes. The second and the third report will be published in 2024 and 2025.



Data stories

Filter by

Country

- Any -

Year

- Any -

Search

Search results (232)

Sort by Last created

25 January 2024

Getting to know Spain through open data

Three leading institutions are involved in EU decision-making Parliament represents EU citizens, the Council of the Europe

Show more

11 January 2024

Understanding family spending through data

Learning Overview

Structure of sessions

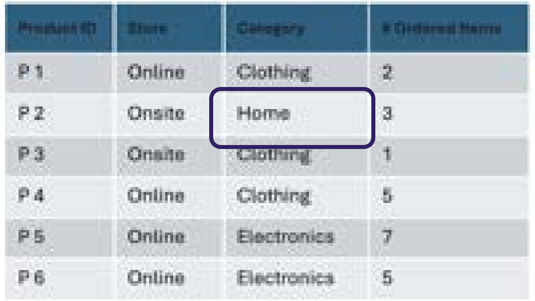
Session	Topic	Duration
Webinar 1	Data Access and Exploration	45 min
Webinar 2	Data Cleaning and Transformation	75 min
Webinar 3	Data Blending and Storage	75 min
Webinar 4	Analytics	75 min
Webinar 5	Advanced Module: Gathering Data from the Web and Geo Visualisation	75 min

- Slides, hands-on exercises, and solutions will be provided on the webinar
- Gitlab repository https://code.europa.eu/bdti/bdti-essentials-course

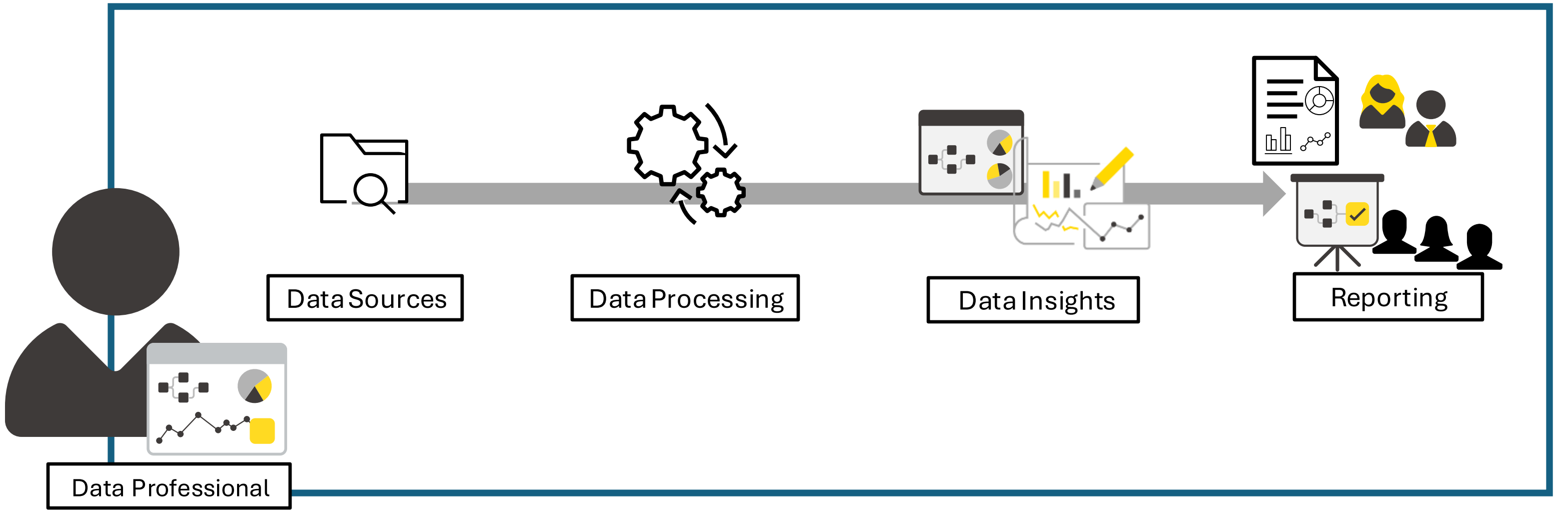
Target Audience

- Assume beginner level (spreadsheet users): KNIME
- Advanced users: Jupyter lab and R-studio

From tables and cells to process thinking



Product ID	Store	Category	# Ordered Items
P 1	Online	Clothing	2
P 2	Onsite	Home	3
P 3	Onsite	Clothing	1
P 4	Online	Clothing	5
P 5	Online	Electronics	7
P 6	Online	Electronics	5



Data Sources

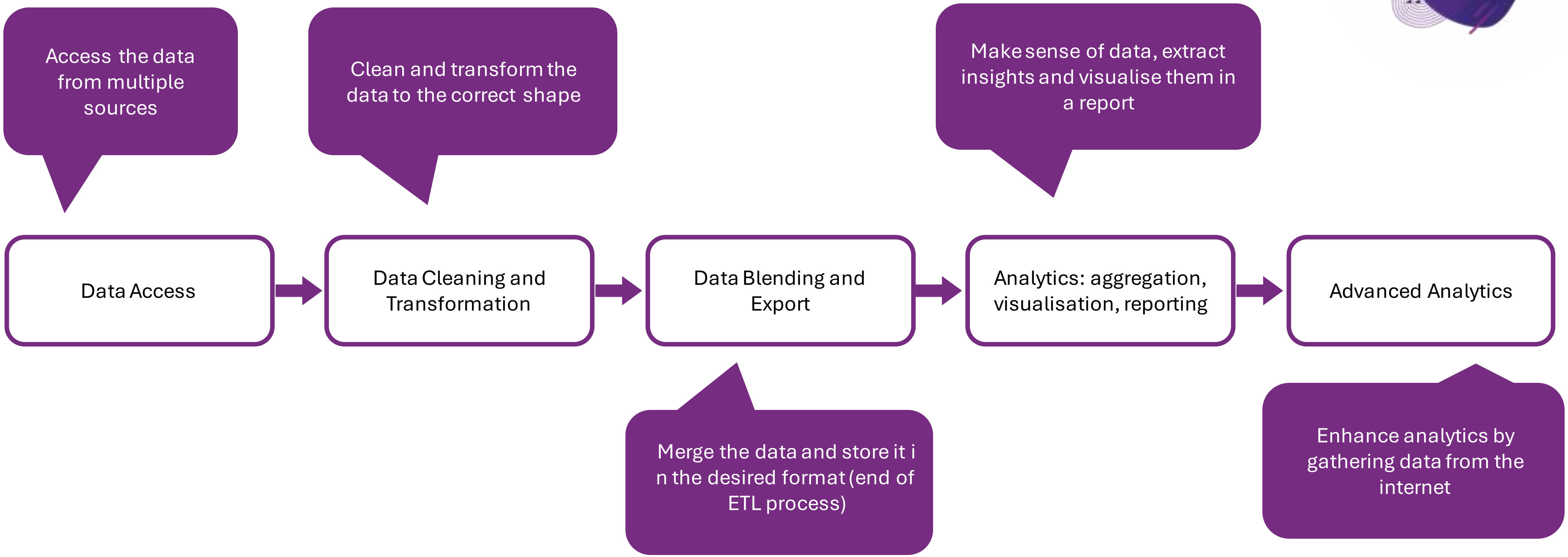
Data Processing

Data Insights

Reporting

Data Professional

The Data Process



Data Access

Access the data from multiple sources

Data Cleaning and Transformation

Clean and transform the data to the correct shape

Data Blending and Export

Merge the data and store it i n the desired format (end of ETL process)

Analytics: aggregation, visualisation, reporting

Make sense of data, extract insights and visualise them in a report

Advanced Analytics

Enhance analytics by gathering data from the internet

Use case: "Investigating Funding of Green Energy projects on EU Universities"

Use Case

Zoi and her team, work in the Educational Department of a highly polluted EU region.

- Their **purpose** is to identify potential ***"energy partnerships"*** between universities on a national and international level to address pollution issues.

- To achieve this purpose, they track **two main sets data**. The first concerns the research and innovation funding of ***energy*** projects at universities across EU member states. The second dataset contains CO2 emissions of these countries. Moreover, they use a matching table to relate these datasets to each other.

- Zoi reached out the BDTI team to help learn how to **create *insights*** from the data sets and produce a ***report*** accordingly for their department, which can be **easily updated** when new data is available.

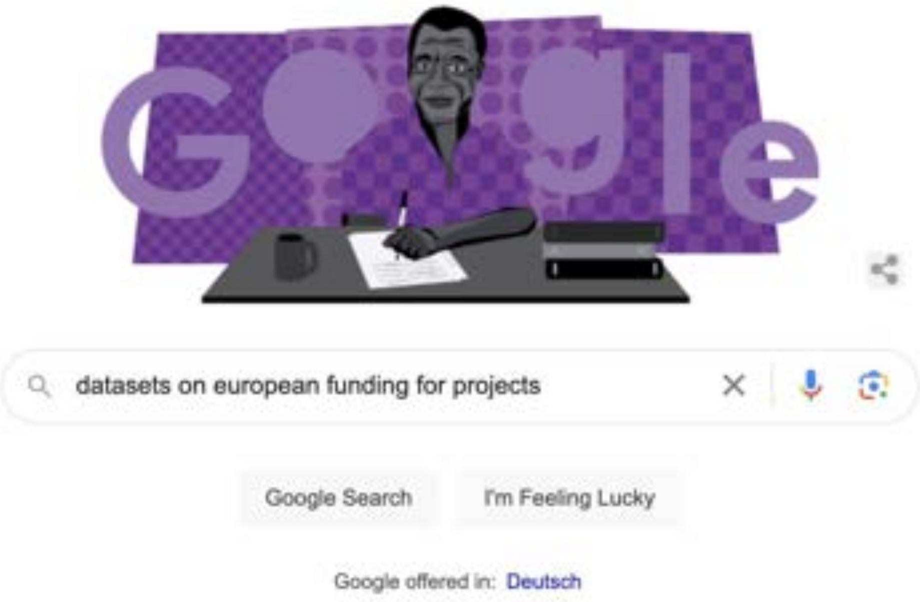
Breakdown of objectives

- Visualise Energy projects per Member State and for **Organizations** per **Country**
- Identify and Visualise Member States with the Most and Least **CO2 Emissions**
- Identify **Energy projects** per Country per **year**
- Explore the Relationship Between CO2 Emissions and Energy Project Funding
- Locate on a **map** the Universities that got funding for Energy projects from Horizon

Mapping objectives to data

- First, finding the correct data to achieve the objectives.
- Then, Zoi has to **explore** the data to make decisions according to the different objectives.
- Map her objectives with the data she has.

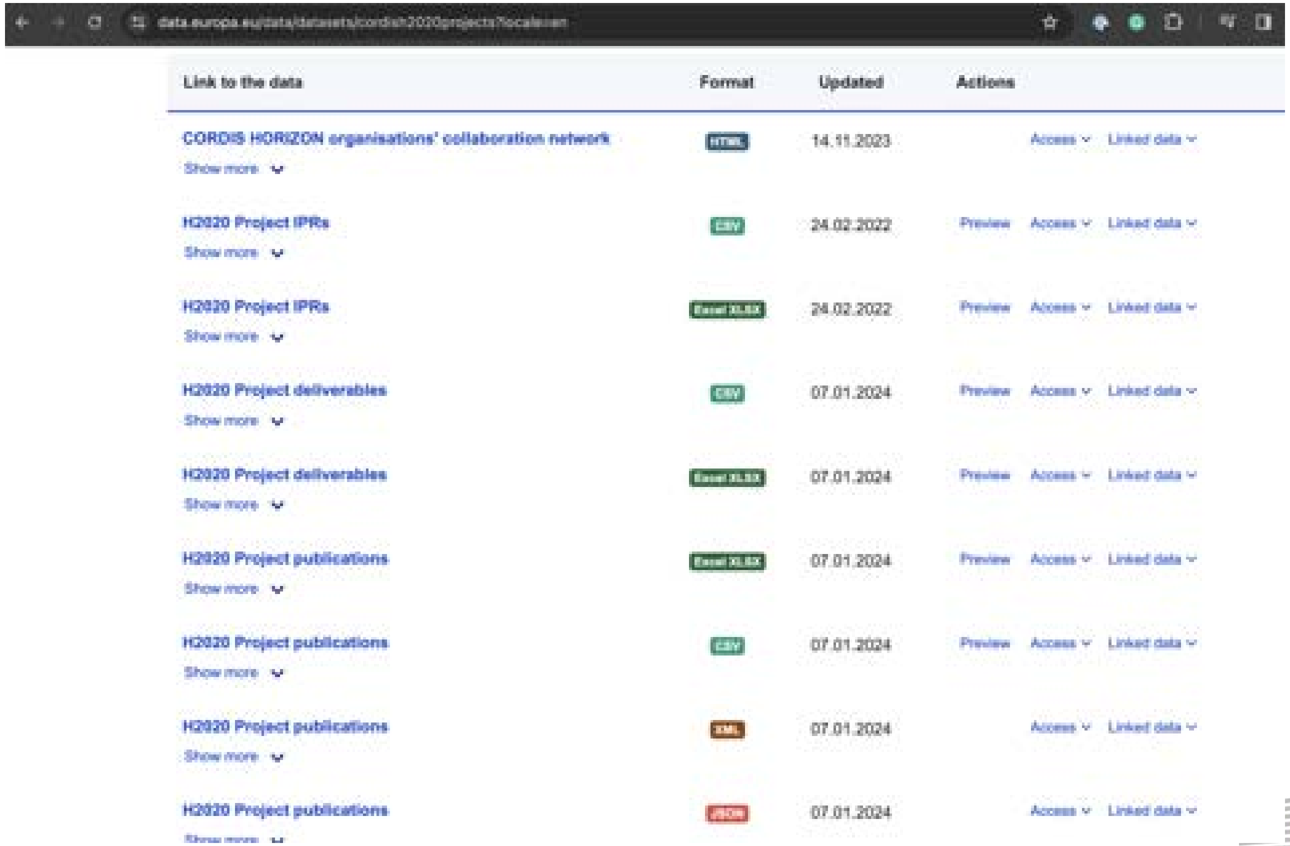
Searching for the (open) data



datasets on european funding for projects

Data Access - Sources

- Zoi found open data at https://data.europa.eu/en, she searched: "Horizon 2020 cordis" and "Horizon 2021".

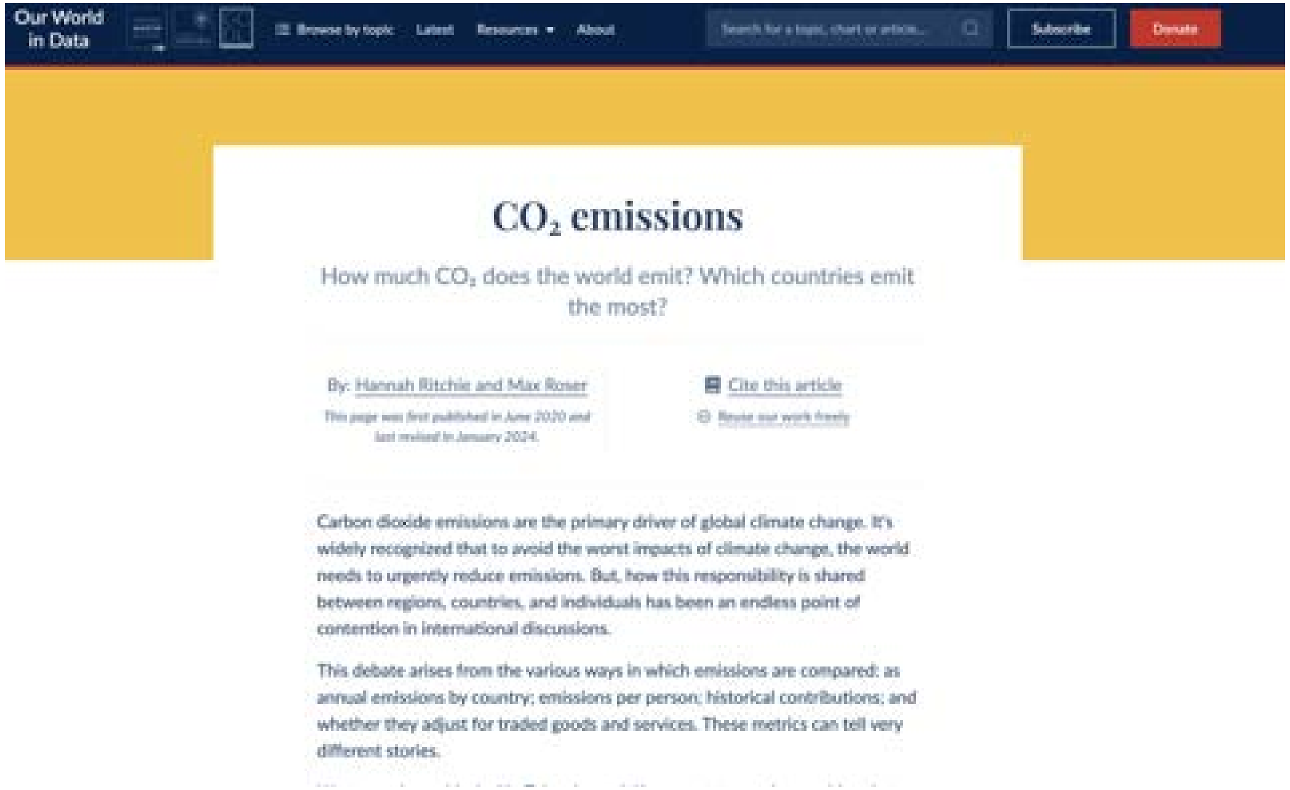


data.europa.eu/data/datasets/cordish2020projects?locale=en

Link to the data	Format	Updated	Actions
CORDIS HORIZON organisations' collaboration network Show more	HTML	14.11.2023	Access Linked data
H2020 Project IPRs Show more	CSV	24.02.2022	Preview Access Linked data
H2020 Project IPRs Show more	Excel XLSX	24.02.2022	Preview Access Linked data
H2020 Project deliverables Show more	CSV	07.01.2024	Preview Access Linked data
H2020 Project deliverables Show more	Excel XLSX	07.01.2024	Preview Access Linked data
H2020 Project publications Show more	Excel XLSX	07.01.2024	Preview Access Linked data
H2020 Project publications Show more	CSV	07.01.2024	Preview Access Linked data
H2020 Project publications Show more	XML	07.01.2024	Access Linked data
H2020 Project publications Show more	JSON	07.01.2024	Access Linked data

Data Access - Sources

- The team also searched for CO2 emissions and they landed on "Our World in Data" website https://ourworldindata.org/



CO_2 emissions

How much CO_2 does the world emit? Which countries emit the most?

By: Hannah Ritchie and Max Roser

This page was first published in June 2020 and last revised in January 2024.

Cite this article

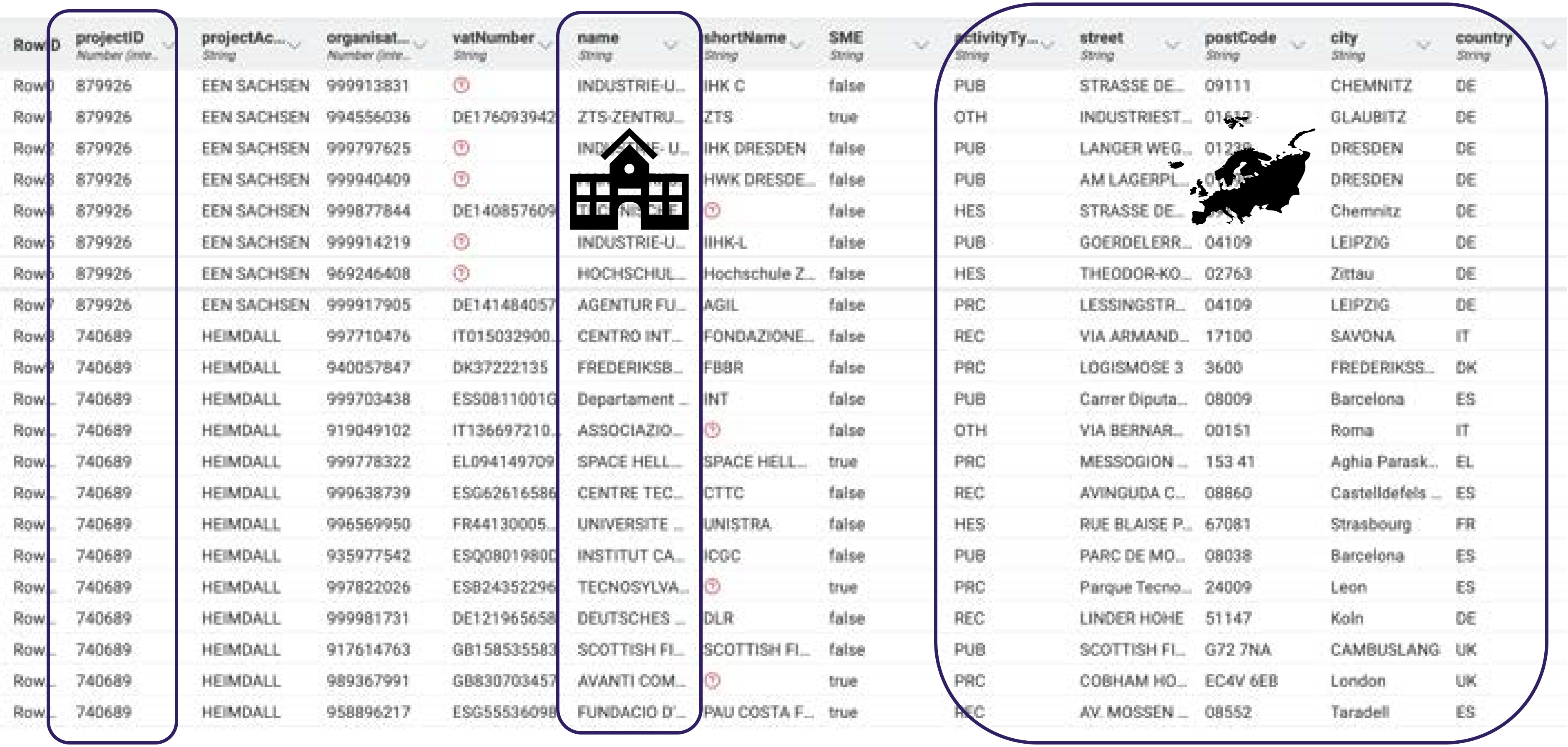
Reuse our work freely

Carbon dioxide emissions are the primary driver of global climate change. It's widely recognized that to avoid the worst impacts of climate change, the world needs to urgently reduce emissions. But, how this responsibility is shared between regions, countries, and individuals has been an endless point of contention in international discussions.

This debate arises from the various ways in which emissions are compared: as annual emissions by country; emissions per person; historical contributions; and whether they adjust for traded goods and services. These metrics can tell very different stories.

Horizon data

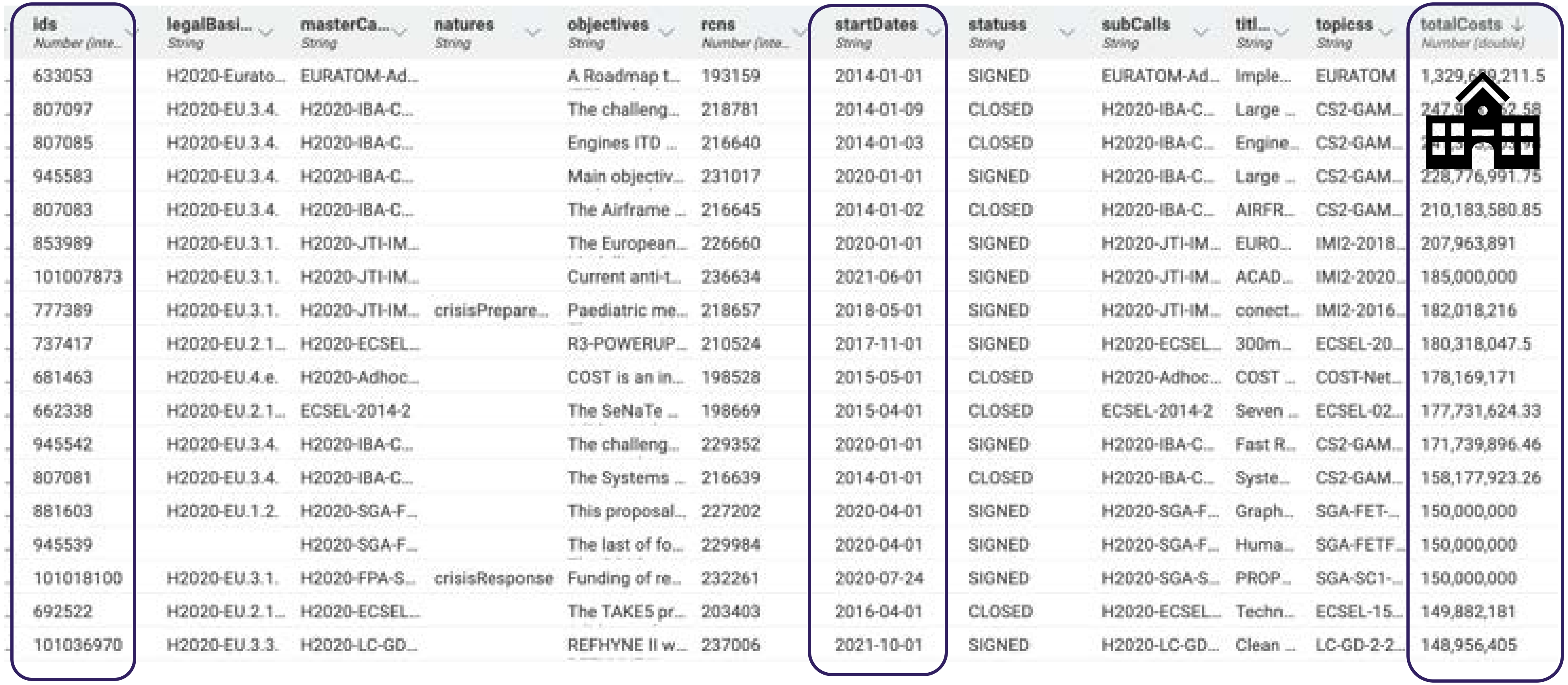
- Organizations dataset - *Visualise Energy projects per Member State and for* ***Organizations*** *per* ***Country***



RowID	projectID Number (inte…	projectAc… String	organisat… Number (inte…	vatNumber String	name String	shortName String	SME String	activityTy… String	street String	postCode String	city String	country String
Row0	879926	EEN SACHSEN	999913831	?	INDUSTRIE-U…	IHK C	false	PUB	STRASSE DE…	09111	CHEMNITZ	DE
Row1	879926	EEN SACHSEN	994556036	DE176093942	ZTS-ZENTRU…	ZTS	true	OTH	INDUSTRIEST…	01612	GLAUBITZ	DE
Row2	879926	EEN SACHSEN	999797625	?	INDUSTRIE- U…	IHK DRESDEN	false	PUB	LANGER WEG…	0123[illegible]	DRESDEN	DE
Row3	879926	EEN SACHSEN	999940409	?	[illegible]	HWK DRESDE…	false	PUB	AM LAGERPL…	[illegible]	DRESDEN	DE
Row4	879926	EEN SACHSEN	999877844	DE140857609	TECHNISCHE…	?	false	HES	STRASSE DE…	[illegible]	Chemnitz	DE
Row5	879926	EEN SACHSEN	999914219	?	INDUSTRIE-U…	IIHK-L	false	PUB	GOERDELERR…	04109	LEIPZIG	DE
Row6	879926	EEN SACHSEN	969246408	?	HOCHSCHUL…	Hochschule Z…	false	HES	THEODOR-KO…	02763	Zittau	DE
Row7	879926	EEN SACHSEN	999917905	DE141484057	AGENTUR FU…	AGIL	false	PRC	LESSINGSTR…	04109	LEIPZIG	DE
Row8	740689	HEIMDALL	997710476	IT015032900…	CENTRO INT…	FONDAZIONE…	false	REC	VIA ARMAND…	17100	SAVONA	IT
Row9	740689	HEIMDALL	940057847	DK37222135	FREDERIKSB…	FBBR	false	PRC	LOGISMOSE 3	3600	FREDERIKSS…	DK
Row…	740689	HEIMDALL	999703438	ESS0811001G	Departament …	INT	false	PUB	Carrer Diputa…	08009	Barcelona	ES
Row…	740689	HEIMDALL	919049102	IT136697210…	ASSOCIAZIO…	?	false	OTH	VIA BERNAR…	00151	Roma	IT
Row…	740689	HEIMDALL	999778322	EL094149709	SPACE HELL…	SPACE HELL…	true	PRC	MESSOGION …	153 41	Aghia Parask…	EL
Row…	740689	HEIMDALL	999638739	ESG62616586	CENTRE TEC…	CTTC	false	REC	AVINGUDA C…	08860	Castelldefels …	ES
Row…	740689	HEIMDALL	996569950	FR44130005…	UNIVERSITE …	UNISTRA	false	HES	RUE BLAISE P…	67081	Strasbourg	FR
Row…	740689	HEIMDALL	935977542	ESQ0801980D	INSTITUT CA…	ICGC	false	PUB	PARC DE MO…	08038	Barcelona	ES
Row…	740689	HEIMDALL	997822026	ESB24352296	TECNOSYLVA…	?	true	PRC	Parque Tecno…	24009	Leon	ES
Row…	740689	HEIMDALL	999981731	DE121965658	DEUTSCHES …	DLR	false	REC	LINDER HOHE	51147	Koln	DE
Row…	740689	HEIMDALL	917614763	GB158535583	SCOTTISH FI…	SCOTTISH FI…	false	PUB	SCOTTISH FI…	G72 7NA	CAMBUSLANG	UK
Row…	740689	HEIMDALL	989367991	GB830703457	AVANTI COM…	?	true	PRC	COBHAM HO…	EC4V 6EB	London	UK
Row…	740689	HEIMDALL	958896217	ESG55536098	FUNDACIO D'…	PAU COSTA F…	true	REC	AV. MOSSEN …	08552	Taradell	ES

Horizon data

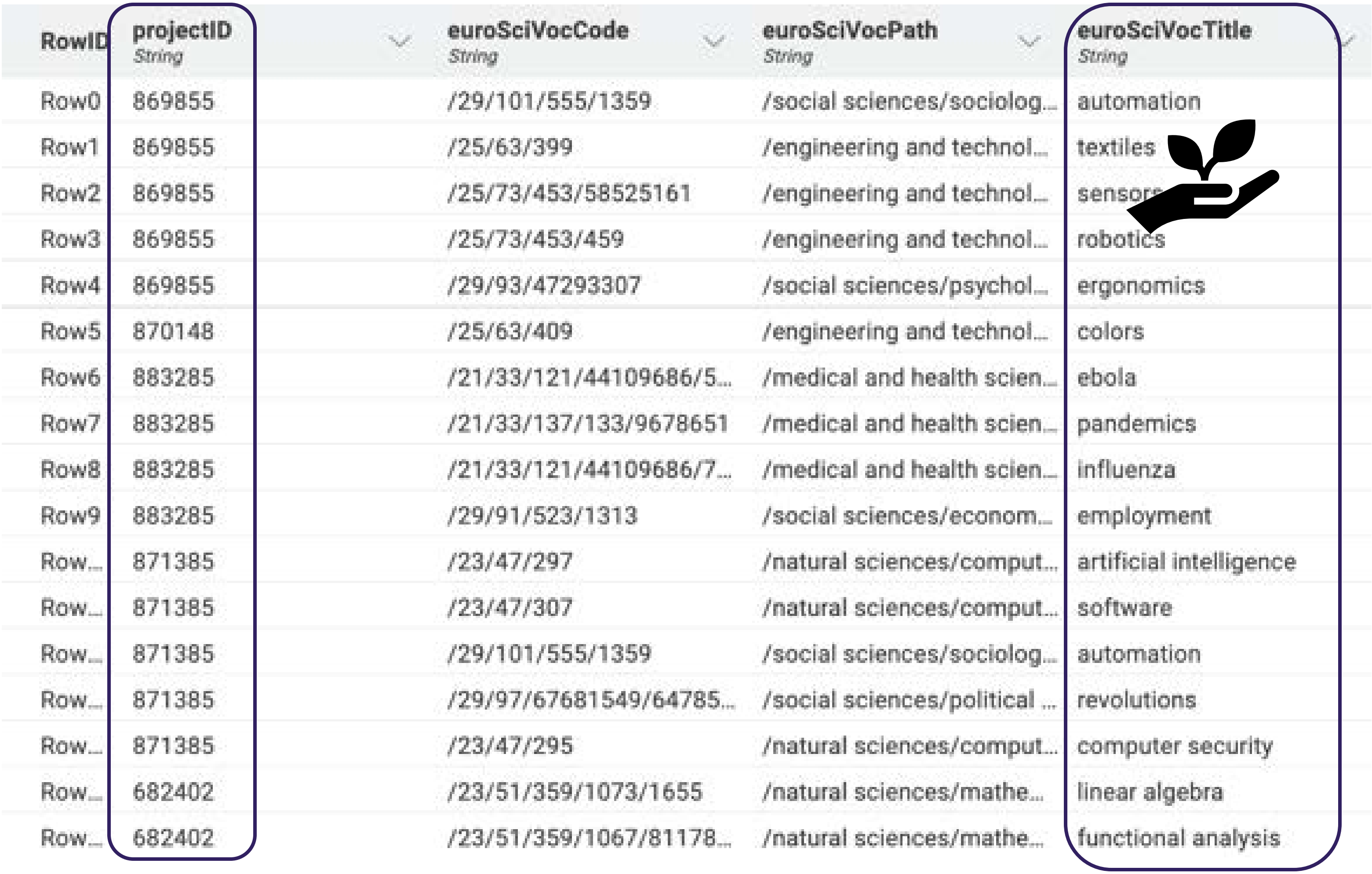
- Projects dataset - *Identify Energy **Projects** per Country per **year***



ids Number (inte...	legalBasi... String	masterCa... String	natures String	objectives String	rcns Number (inte...	startDates String	statuss String	subCalls String	titl... String	topicss String	totalCosts ↓ Number (double)
633053	H2020-Eurato...	EURATOM-Ad...		A Roadmap t...	193159	2014-01-01	SIGNED	EURATOM-Ad...	Imple...	EURATOM	1,329,6[illegible]9,211.5
807097	H2020-EU.3.4.	H2020-IBA-C...		The challeng...	218781	2014-01-09	CLOSED	H2020-IBA-C...	Large ...	CS2-GAM...	247,9[illegible]2.58
807085	H2020-EU.3.4.	H2020-IBA-C...		Engines ITD ...	216640	2014-01-03	CLOSED	H2020-IBA-C...	Engine...	CS2-GAM...	[illegible]
945583	H2020-EU.3.4.	H2020-IBA-C...		Main objectiv...	231017	2020-01-01	SIGNED	H2020-IBA-C...	Large ...	CS2-GAM...	228,776,991.75
807083	H2020-EU.3.4.	H2020-IBA-C...		The Airframe ...	216645	2014-01-02	CLOSED	H2020-IBA-C...	AIRFR...	CS2-GAM...	210,183,580.85
853989	H2020-EU.3.1.	H2020-JTI-IM...		The European...	226660	2020-01-01	SIGNED	H2020-JTI-IM...	EURO...	IMI2-2018...	207,963,891
101007873	H2020-EU.3.1.	H2020-JTI-IM...		Current anti-t...	236634	2021-06-01	SIGNED	H2020-JTI-IM...	ACAD...	IMI2-2020...	185,000,000
777389	H2020-EU.3.1.	H2020-JTI-IM...	crisisPrepare...	Paediatric me...	218657	2018-05-01	SIGNED	H2020-JTI-IM...	conect...	IMI2-2016...	182,018,216
737417	H2020-EU.2.1...	H2020-ECSEL...		R3-POWERUP...	210524	2017-11-01	SIGNED	H2020-ECSEL...	300m...	ECSEL-20...	180,318,047.5
681463	H2020-EU.4.e.	H2020-Adhoc...		COST is an in...	198528	2015-05-01	CLOSED	H2020-Adhoc...	COST ...	COST-Net...	178,169,171
662338	H2020-EU.2.1...	ECSEL-2014-2		The SeNaTe ...	198669	2015-04-01	CLOSED	ECSEL-2014-2	Seven ...	ECSEL-02...	177,731,624.33
945542	H2020-EU.3.4.	H2020-IBA-C...		The challeng...	229352	2020-01-01	SIGNED	H2020-IBA-C...	Fast R...	CS2-GAM...	171,739,896.46
807081	H2020-EU.3.4.	H2020-IBA-C...		The Systems ...	216639	2014-01-01	CLOSED	H2020-IBA-C...	Syste...	CS2-GAM...	158,177,923.26
881603	H2020-EU.1.2.	H2020-SGA-F...		This proposal...	227202	2020-04-01	SIGNED	H2020-SGA-F...	Graph...	SGA-FET-...	150,000,000
945539		H2020-SGA-F...		The last of fo...	229984	2020-04-01	SIGNED	H2020-SGA-F...	Huma...	SGA-FETF...	150,000,000
101018100	H2020-EU.3.1.	H2020-FPA-S...	crisisResponse	Funding of re...	232261	2020-07-24	SIGNED	H2020-SGA-S...	PROP...	SGA-SC1-...	150,000,000
692522	H2020-EU.2.1...	H2020-ECSEL...		The TAKE5 pr...	203403	2016-04-01	CLOSED	H2020-ECSEL...	Techn...	ECSEL-15...	149,882,181
101036970	H2020-EU.3.3.	H2020-LC-GD...		REFHYNE II w...	237006	2021-10-01	SIGNED	H2020-LC-GD...	Clean ...	LC-GD-2-2...	148,956,405

Horizon data

- EuroSciVoc dataset - *Identify **Energy** Projects per Country per year*

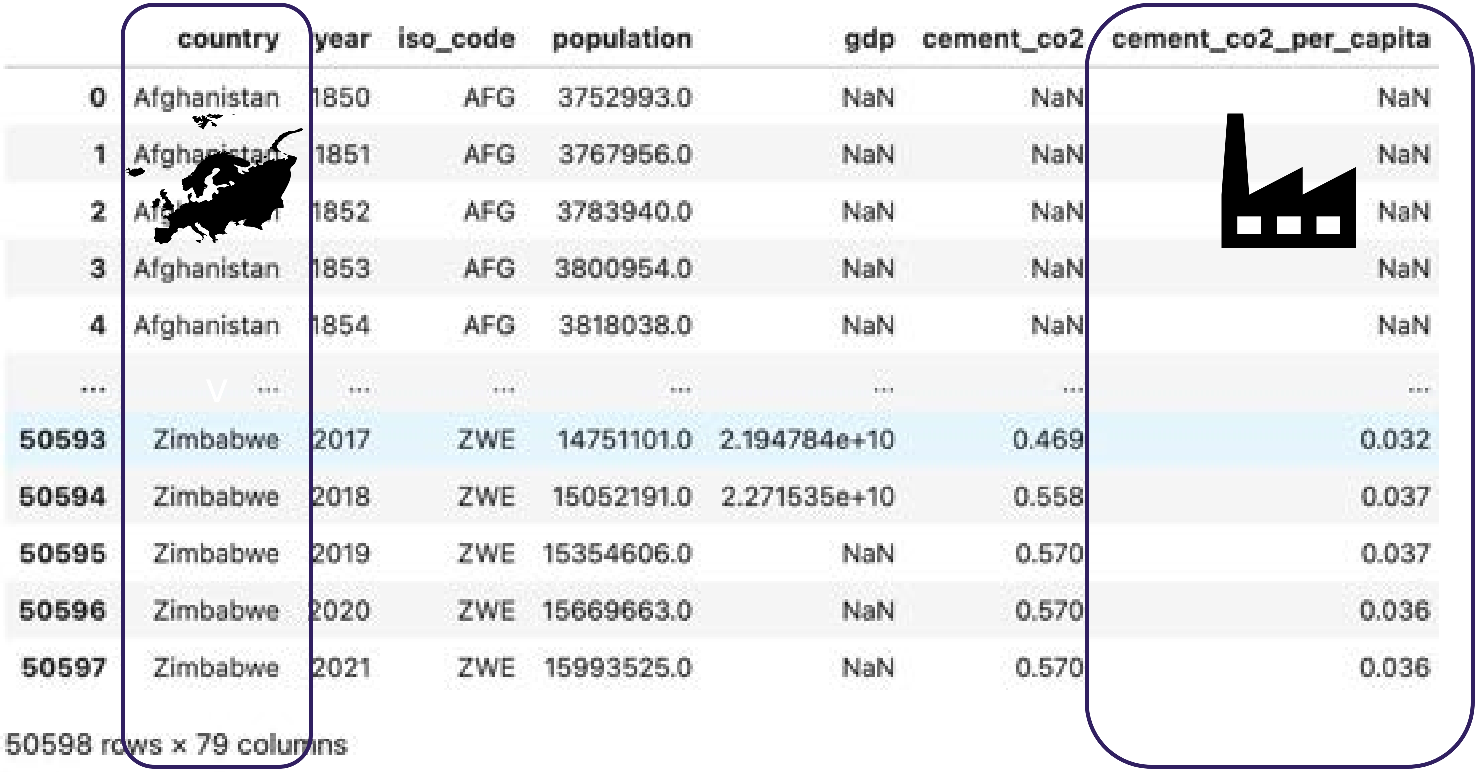


RowID	projectID String	euroSciVocCode String	euroSciVocPath String	euroSciVocTitle String
Row0	869855	/29/101/555/1359	/social sciences/sociolog…	automation
Row1	869855	/25/63/399	/engineering and technol…	textiles
Row2	869855	/25/73/453/58525161	/engineering and technol…	sensors
Row3	869855	/25/73/453/459	/engineering and technol…	robotics
Row4	869855	/29/93/47293307	/social sciences/psychol…	ergonomics
Row5	870148	/25/63/409	/engineering and technol…	colors
Row6	883285	/21/33/121/44109686/5…	/medical and health scien…	ebola
Row7	883285	/21/33/137/133/9678651	/medical and health scien…	pandemics
Row8	883285	/21/33/121/44109686/7…	/medical and health scien…	influenza
Row9	883285	/29/91/523/1313	/social sciences/econom…	employment
Row…	871385	/23/47/297	/natural sciences/comput…	artificial intelligence
Row…	871385	/23/47/307	/natural sciences/comput…	software
Row…	871385	/29/101/555/1359	/social sciences/sociolog…	automation
Row…	871385	/29/97/67681549/64785…	/social sciences/political …	revolutions
Row…	871385	/23/47/295	/natural sciences/comput…	computer security
Row…	682402	/23/51/359/1073/1655	/natural sciences/mathe…	linear algebra
Row…	682402	/23/51/359/1067/81178…	/natural sciences/mathe…	functional analysis

CO2 emissions data

- Identify and Visualise Member States with the Most and Least **CO2 Emissions**
- One single file with 79 columns, and information about all the countries in the world.

But Zoi only needs EU members!



	country	year	iso_code	population	gdp	cement_co2	cement_co2_per_capita
0	Afghanistan	1850	AFG	3752993.0	NaN	NaN	NaN
1	Afghanistan	1851	AFG	3767956.0	NaN	NaN	NaN
2	Af	1852	AFG	3783940.0	NaN	NaN	NaN
3	Afghanistan	1853	AFG	3800954.0	NaN	NaN	NaN
4	Afghanistan	1854	AFG	3818038.0	NaN	NaN	NaN
...	...	...	...	...	...	...	...
50593	Zimbabwe	2017	ZWE	14751101.0	2.194784e+10	0.469	0.032
50594	Zimbabwe	2018	ZWE	15052191.0	2.271535e+10	0.558	0.037
50595	Zimbabwe	2019	ZWE	15354606.0	NaN	0.570	0.037
50596	Zimbabwe	2020	ZWE	15669663.0	NaN	0.570	0.036
50597	Zimbabwe	2021	ZWE	15993525.0	NaN	0.570	0.036

50598 rows × 79 columns

Use case: Approach

- Use the Big Data Test Infrastructure (BDTI) initiative from the DEP program (European Commission)

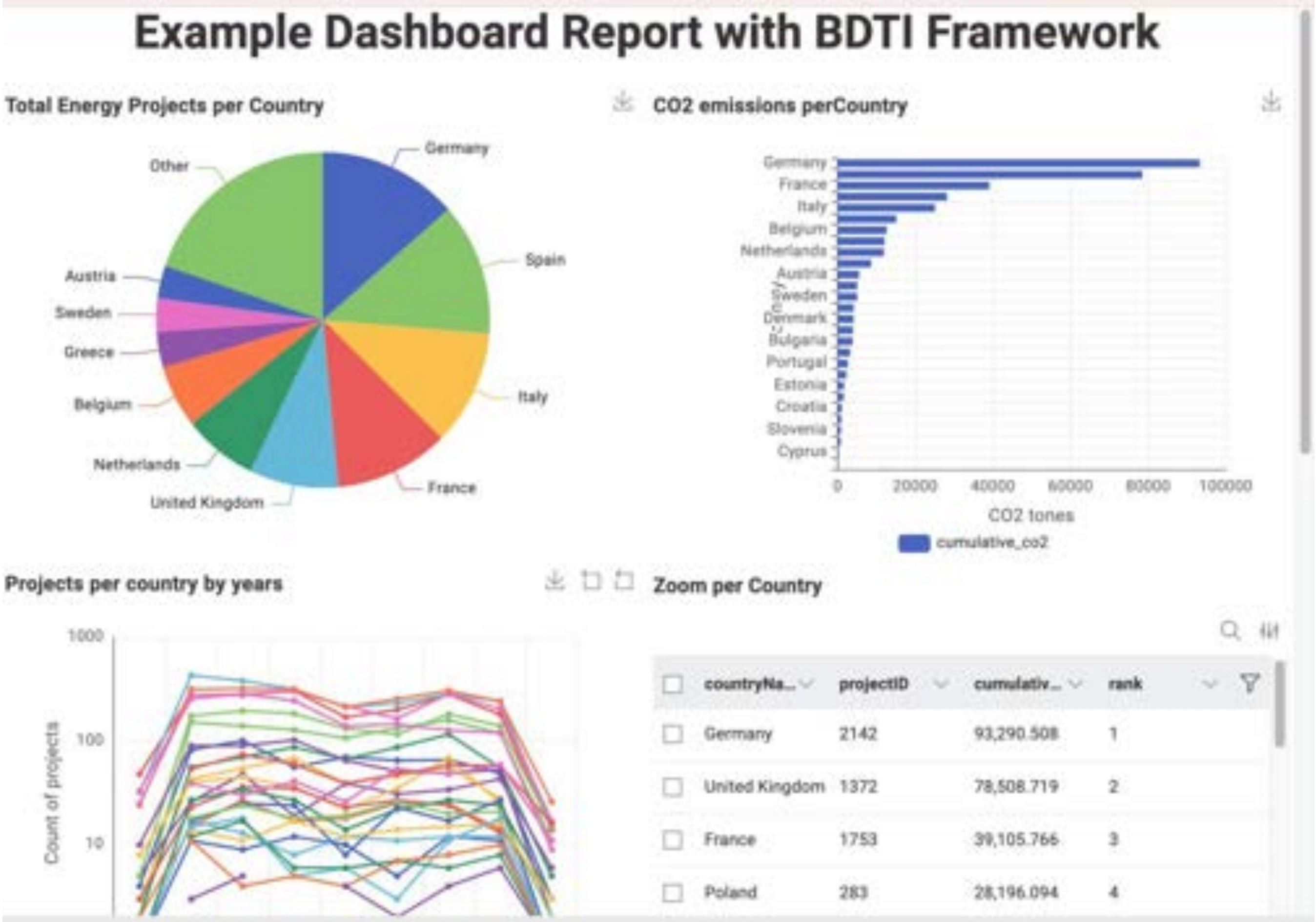
- Received specialised training for effective data analysis

- Leverage open data from different sources
    - EU Funding Program Data – Horizon2020/1: https://data.europa.eu/data/datasets/cordish2020projects?locale=en
    - CO2 Emissions Data (Our World in Data) https://ourworldindata.org/co2-emissions
    - Collect external data if needed

Outputs according to Zoi goals

The report/dashboard needs to be informative!

- Insightful plots
- Tables with a rank of countries according to CO2 emissions
- Bar chart with CO2 emissions
- Line plot with both Horizon 2020 and 2021
- Locate on a map the Universities that got funding for Energy projects from Horizon

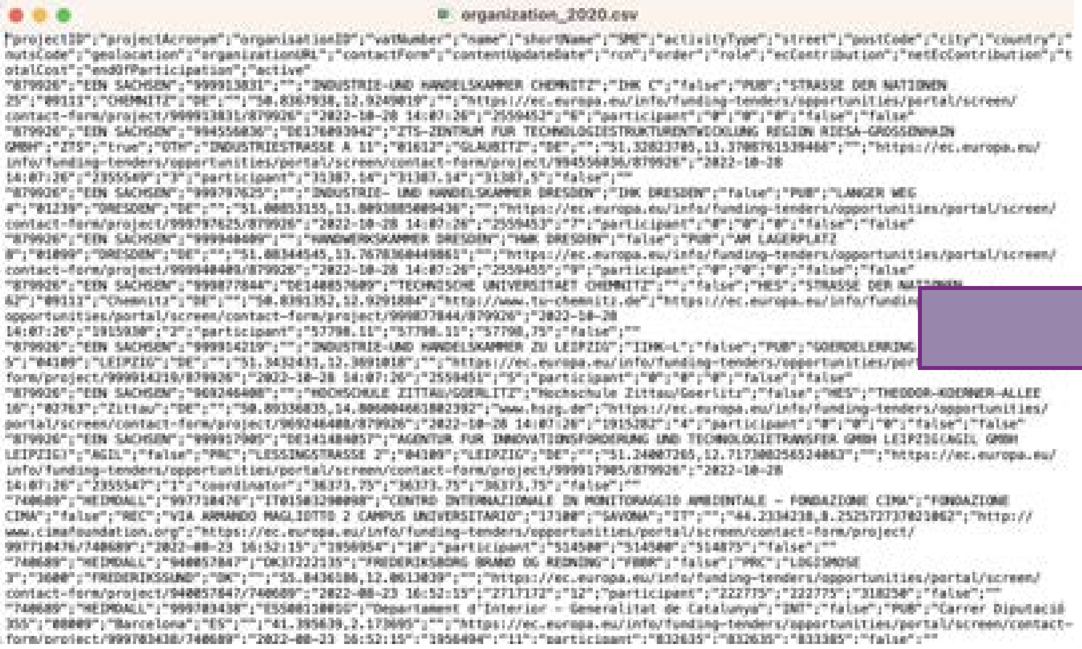


Example Dashboard Report with BDTI Framework

□	countryNa...	projectID	cumulativ...	rank
□	Germany	2142	93,290.508	1
□	United Kingdom	1372	78,508.719	2
□	France	1753	39,105.766	3
□	Poland	283	28,196.094	4

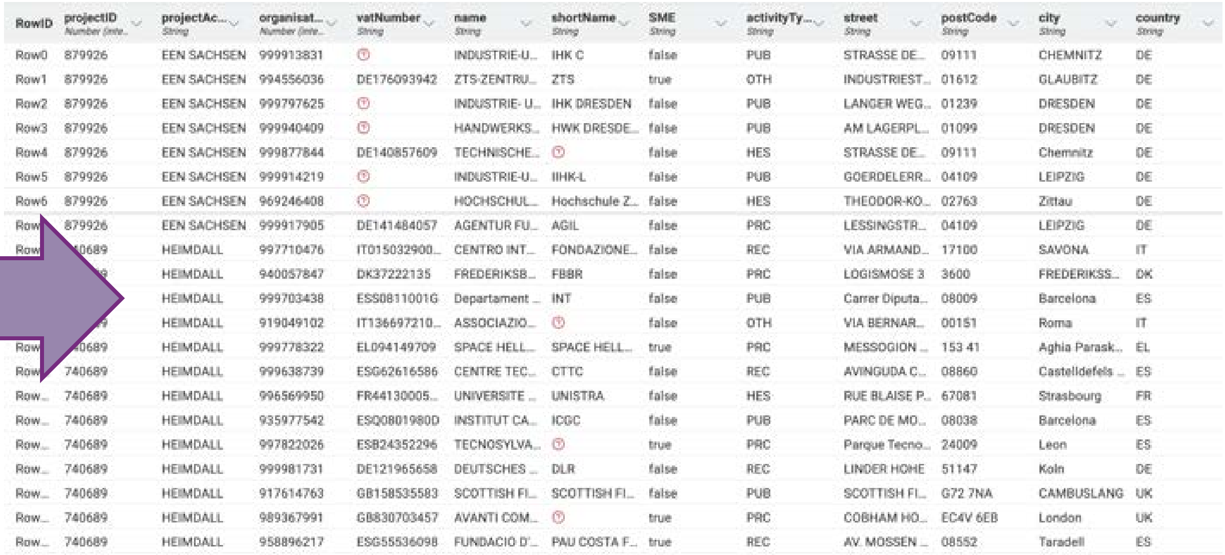
But…

- The flipside of the coin is the technical side
- Raw file (CSV) vs table



organization_2020.csv

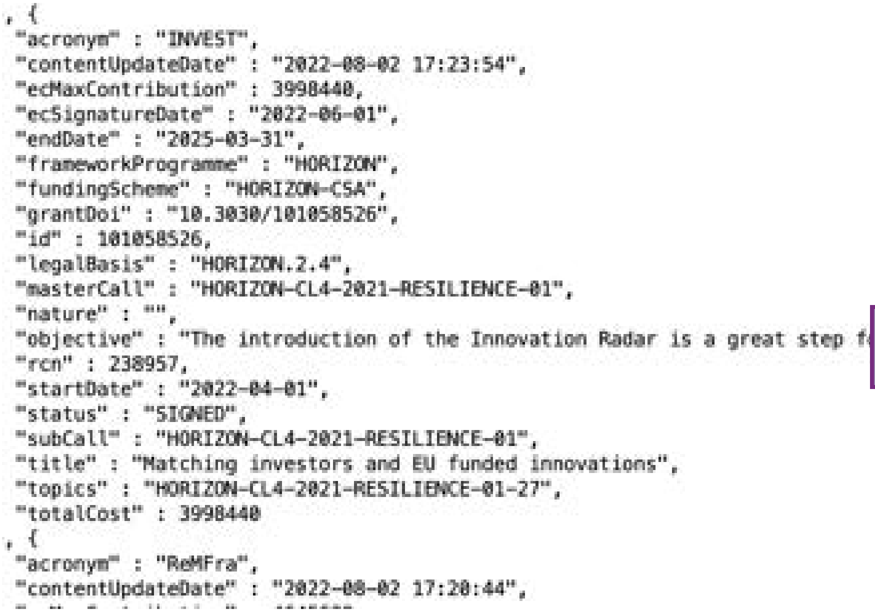
```
"projectID";"projectAcronym";"organisationID";"vatNumber";"name";"shortName";"SME";"activityType";"street";"postCode";"city";"country";"
nutsCode";"geolocation";"organizationURL";"contactForm";"contentUpdateDate";"rcn";"order";"role";"ecContribution";"netEcContribution";"t
otalCost";"endOfParticipation";"active"
"879926";"EEN SACHSEN";"999913831";"";"INDUSTRIE-UND HANDELSKAMMER CHEMNITZ";"IHK C";"false";"PUB";"STRASSE DER NATIONEN
25";"09111";"CHEMNITZ";"DE";"";"50.8367938,12.9249019";"";"https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/
contact-form/project/999913831/879926";"2022-10-28 14:07:26";"2559452";"6";"participant";"0";"0";"0";"false";"false"
"879926";"EEN SACHSEN";"994556036";"DE176093942";"ZTS-ZENTRUM FUR TECHNOLOGIESTRUKTURENTWICKLUNG REGION RIESA-GROSSENHAIN
GMBH";"ZTS";"true";"OTH";"INDUSTRIESTRASSE A 11";"01612";"GLAUBITZ";"DE";"";"51.32823705,13.3708761539466";"";"https://ec.europa.eu/
info/funding-tenders/opportunities/portal/screen/contact-form/project/994556036/879926";"2022-10-28
14:07:26";"2355549";"3";"participant";"31387.14";"31387.14";"31387,5";"false";""
"879926";"EEN SACHSEN";"999797625";"";"INDUSTRIE- UND HANDELSKAMMER DRESDEN";"IHK DRESDEN";"false";"PUB";"LANGER WEG
4";"01239";"DRESDEN";"DE";"";"51.00853155,13.8093885009436";"";"https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/
contact-form/project/999797625/879926";"2022-10-28 14:07:26";"2559453";"7";"participant";"0";"0";"0";"false";"false"
"879926";"EEN SACHSEN";"999940409";"";"HANDWERKSKAMMER DRESDEN";"HWK DRESDEN";"false";"PUB";"AM LAGERPLATZ
8";"01099";"DRESDEN";"DE";"";"51.08344545,13.7678360449861";"";"https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/
contact-form/project/999940409/879926";"2022-10-28 14:07:26";"2559455";"9";"participant";"0";"0";"0";"false";"false"
"879926";"EEN SACHSEN";"999877844";"DE140857609";"TECHNISCHE UNIVERSITAET CHEMNITZ";"";"false";"HES";"STRASSE DER NATIONEN
62";"09111";"Chemnitz";"DE";"";"50.8391352,12.9291884";"http://www.tu-chemnitz.de";"https://ec.europa.eu/info/fundin
opportunities/portal/screen/contact-form/project/999877844/879926";"2022-10-28
14:07:26";"1915930";"2";"participant";"57798.11";"57798.11";"57798,75";"false";""
"879926";"EEN SACHSEN";"999914219";"";"INDUSTRIE-UND HANDELSKAMMER ZU LEIPZIG";"IIHK-L";"false";"PUB";"GOERDELERRING
5";"04109";"LEIPZIG";"DE";"";"51.3432431,12.3691018";"";"https://ec.europa.eu/info/funding-tenders/opportunities/por
form/project/999914219/879926";"2022-10-28 14:07:26";"2559451";"5";"participant";"0";"0";"0";"false";"false"
"879926";"EEN SACHSEN";"969246408";"";"HOCHSCHULE ZITTAU/GOERLITZ";"Hochschule Zittau/Goerlitz";"false";"HES";"THEODOR-KOERNER-ALLEE
16";"02763";"Zittau";"DE";"";"50.89336835,14.806004661802392";"www.hszg.de";"https://ec.europa.eu/info/funding-tenders/opportunities/
portal/screen/contact-form/project/969246408/879926";"2022-10-28 14:07:26";"1915282";"4";"participant";"0";"0";"0";"false";"false"
"879926";"EEN SACHSEN";"999917905";"DE141484057";"AGENTUR FUR INNOVATIONSFORDERUNG UND TECHNOLOGIETRANSFER GMBH LEIPZIG(AGIL GMBH
LEIPZIG)";"AGIL";"false";"PRC";"LESSINGSTRASSE 2";"04109";"LEIPZIG";"DE";"";"51.24007265,12.717300256524063";"";"https://ec.europa.eu/
info/funding-tenders/opportunities/portal/screen/contact-form/project/999917905/879926";"2022-10-28
14:07:26";"2355547";"1";"coordinator";"36373.75";"36373.75";"36373,75";"false";""
"740689";"HEIMDALL";"997710476";"IT01503290098";"CENTRO INTERNAZIONALE IN MONITORAGGIO AMBIENTALE - FONDAZIONE CIMA";"FONDAZIONE
CIMA";"false";"REC";"VIA ARMANDO MAGLIOTTO 2 CAMPUS UNIVERSITARIO";"17100";"SAVONA";"IT";"";"44.2334238,8.252572737021062";"http://
www.cimafoundation.org";"https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/contact-form/project/
997710476/740689";"2022-08-23 16:52:15";"1950954";"10";"participant";"514500";"514500";"514875";"false";""
"740689";"HEIMDALL";"940057847";"DK37222135";"FREDERIKSBORG BRAND OG REDNING";"FBBR";"false";"PRC";"LOGISMOSE
3";"3600";"FREDERIKSSUND";"DK";"";"55.8436186,12.0613029";"";"https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/
contact-form/project/940057847/740689";"2022-08-23 16:52:15";"2717172";"12";"participant";"222775";"222775";"318250";"false";""
"740689";"HEIMDALL";"999703438";"ESS0811001G";"Departament d'Interior - Generalitat de Catalunya";"INT";"false";"PUB";"Carrer Diputació
355";"08009";"Barcelona";"ES";"";"41.395639,2.173695";"";"https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/contact-
form/project/999703438/740689";"2022-08-23 16:52:15";"1950494";"11";"participant";"832635";"832635";"833385";"false";""
```



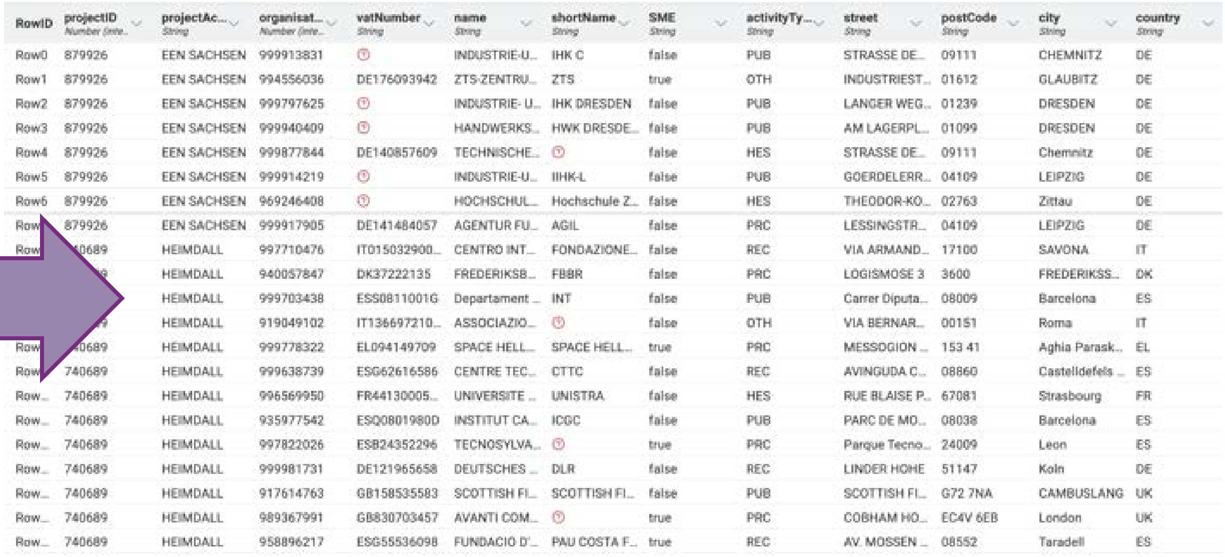
RowID	projectID Number (inte…	projectAc… String	organisat… Number (inte…	vatNumber String	name String	shortName String	SME String	activityTy… String	street String	postCode String	city String	country String
Row0	879926	EEN SACHSEN	999913831	?	INDUSTRIE-U…	IHK C	false	PUB	STRASSE DE…	09111	CHEMNITZ	DE
Row1	879926	EEN SACHSEN	994556036	DE176093942	ZTS-ZENTRU…	ZTS	true	OTH	INDUSTRIEST…	01612	GLAUBITZ	DE
Row2	879926	EEN SACHSEN	999797625	?	INDUSTRIE- U…	IHK DRESDEN	false	PUB	LANGER WEG…	01239	DRESDEN	DE
Row3	879926	EEN SACHSEN	999940409	?	HANDWERKS…	HWK DRESDE…	false	PUB	AM LAGERPL…	01099	DRESDEN	DE
Row4	879926	EEN SACHSEN	999877844	DE140857609	TECHNISCHE…	?	false	HES	STRASSE DE…	09111	Chemnitz	DE
Row5	879926	EEN SACHSEN	999914219	?	INDUSTRIE-U…	IIHK-L	false	PUB	GOERDELERR…	04109	LEIPZIG	DE
Row6	879926	EEN SACHSEN	969246408	?	HOCHSCHUL…	Hochschule Z…	false	HES	THEODOR-KO…	02763	Zittau	DE
Row	879926	EEN SACHSEN	999917905	DE141484057	AGENTUR FU…	AGIL	false	PRC	LESSINGSTR…	04109	LEIPZIG	DE
Row	0689	HEIMDALL	997710476	IT015032900…	CENTRO INT…	FONDAZIONE…	false	REC	VIA ARMAND…	17100	SAVONA	IT
[illegible]	9	HEIMDALL	940057847	DK37222135	FREDERIKSB…	FBBR	false	PRC	LOGISMOSE 3	3600	FREDERIKSS…	DK
[illegible]	[illegible]	HEIMDALL	999703438	ESS0811001G	Departament …	INT	false	PUB	Carrer Diputa…	08009	Barcelona	ES
[illegible]	9	HEIMDALL	919049102	IT136697210…	ASSOCIAZIO…	?	false	OTH	VIA BERNAR…	00151	Roma	IT
Row	0689	HEIMDALL	999778322	EL094149709	SPACE HELL…	SPACE HELL…	true	PRC	MESSOGION …	153 41	Aghia Parask…	EL
Row	740689	HEIMDALL	999638739	ESG62616586	CENTRE TEC…	CTTC	false	REC	AVINGUDA C…	08860	Castelldefels …	ES
Row…	740689	HEIMDALL	996569950	FR44130005…	UNIVERSITE …	UNISTRA	false	HES	RUE BLAISE P…	67081	Strasbourg	FR
Row…	740689	HEIMDALL	935977542	ESQ0801980D	INSTITUT CA…	ICGC	false	PUB	PARC DE MO…	08038	Barcelona	ES
Row…	740689	HEIMDALL	997822026	ESB24352296	TECNOSYLVA…	?	true	PRC	Parque Tecno…	24009	Leon	ES
Row…	740689	HEIMDALL	999981731	DE121965658	DEUTSCHES …	DLR	false	REC	LINDER HOHE	51147	Koln	DE
Row…	740689	HEIMDALL	917614763	GB158535583	SCOTTISH FI…	SCOTTISH FI…	false	PUB	SCOTTISH FI…	G72 7NA	CAMBUSLANG	UK
Row…	740689	HEIMDALL	989367991	GB830703457	AVANTI COM…	?	true	PRC	COBHAM HO…	EC4V 6EB	London	UK
Row…	740689	HEIMDALL	958896217	ESG55536098	FUNDACIO D'…	PAU COSTA F…	true	REC	AV. MOSSEN …	08552	Taradell	ES

But...

- File formats
- Raw (JSON) vs table



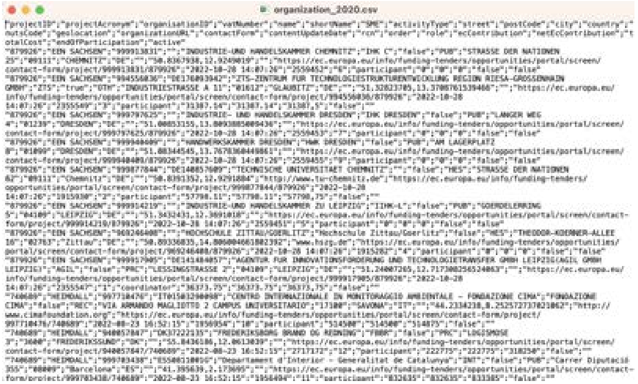
```
, {
 "acronym" : "INVEST",
 "contentUpdateDate" : "2022-08-02 17:23:54",
 "ecMaxContribution" : 3998440,
 "ecSignatureDate" : "2022-06-01",
 "endDate" : "2025-03-31",
 "frameworkProgramme" : "HORIZON",
 "fundingScheme" : "HORIZON-CSA",
 "grantDoi" : "10.3030/101058526",
 "id" : 101058526,
 "legalBasis" : "HORIZON.2.4",
 "masterCall" : "HORIZON-CL4-2021-RESILIENCE-01",
 "nature" : "",
 "objective" : "The introduction of the Innovation Radar is a great step fo
 "rcn" : 238957,
 "startDate" : "2022-04-01",
 "status" : "SIGNED",
 "subCall" : "HORIZON-CL4-2021-RESILIENCE-01",
 "title" : "Matching investors and EU funded innovations",
 "topics" : "HORIZON-CL4-2021-RESILIENCE-01-27",
 "totalCost" : 3998440
, {
 "acronym" : "ReMFra",
 "contentUpdateDate" : "2022-08-02 17:20:44",
```



RowID	projectID Number (inte...	projectAc... String	organisat... Number (inte...	vatNumber String	name String	shortName String	SME String	activityTy... String	street String	postCode String	city String	country String
Row0	879926	EEN SACHSEN	999913831	?	INDUSTRIE-U...	IHK C	false	PUB	STRASSE DE...	09111	CHEMNITZ	DE
Row1	879926	EEN SACHSEN	994556036	DE176093942	ZTS-ZENTRU...	ZTS	true	OTH	INDUSTRIEST...	01612	GLAUBITZ	DE
Row2	879926	EEN SACHSEN	999797625	?	INDUSTRIE- U...	IHK DRESDEN	false	PUB	LANGER WEG...	01239	DRESDEN	DE
Row3	879926	EEN SACHSEN	999940409	?	HANDWERKS...	HWK DRESDE...	false	PUB	AM LAGERPL...	01099	DRESDEN	DE
Row4	879926	EEN SACHSEN	999877844	DE140857609	TECHNISCHE...	?	false	HES	STRASSE DE...	09111	Chemnitz	DE
Row5	879926	EEN SACHSEN	999914219	?	INDUSTRIE-U...	IIHK-L	false	PUB	GOERDELERR...	04109	LEIPZIG	DE
Row6	879926	EEN SACHSEN	969246408	?	HOCHSCHUL...	Hochschule Z...	false	HES	THEODOR-KO...	02763	Zittau	DE
Row	879926	EEN SACHSEN	999917905	DE141484057	AGENTUR FU...	AGIL	false	PRC	LESSINGSTR...	04109	LEIPZIG	DE
Row	40689	HEIMDALL	997710476	IT015032900...	CENTRO INT...	FONDAZIONE...	false	REC	VIA ARMAND...	17100	SAVONA	IT
	9	HEIMDALL	940057847	DK37222135	FREDERIKSB...	FBBR	false	PRC	LOGISMOSE 3	3600	FREDERIKSS...	DK
		HEIMDALL	999703438	ESS0811001G	Departament ...	INT	false	PUB	Carrer Diputa...	08009	Barcelona	ES
	9	HEIMDALL	919049102	IT136697210...	ASSOCIAZIO...	?	false	OTH	VIA BERNAR...	00151	Roma	IT
Row	40689	HEIMDALL	999778322	EL094149709	SPACE HELL...	SPACE HELL...	true	PRC	MESSOGION ...	153 41	Aghia Parask...	EL
Row	740689	HEIMDALL	999638739	ESG62616586	CENTRE TEC...	CTTC	false	REC	AVINGUDA C...	08860	Castelldefels ...	ES
Row...	740689	HEIMDALL	996569950	FR44130005...	UNIVERSITE ...	UNISTRA	false	HES	RUE BLAISE P...	67081	Strasbourg	FR
Row...	740689	HEIMDALL	935977542	ESQ0801980D	INSTITUT CA...	ICGC	false	PUB	PARC DE MO...	08038	Barcelona	ES
Row...	740689	HEIMDALL	997822026	ESB24352296	TECNOSYLVA...	?	true	PRC	Parque Tecno...	24009	Leon	ES
Row...	740689	HEIMDALL	999981731	DE121965658	DEUTSCHES ...	DLR	false	REC	LINDER HOHE	51147	Koln	DE
Row...	740689	HEIMDALL	917614763	GB158535583	SCOTTISH FI...	SCOTTISH FI...	false	PUB	SCOTTISH FI...	G72 7NA	CAMBUSLANG	UK
Row...	740689	HEIMDALL	989367991	GB830703457	AVANTI COM...	?	true	PRC	COBHAM HO...	EC4V 6EB	London	UK
Row...	740689	HEIMDALL	958896217	ESG55536098	FUNDACIO D'...	PAU COSTA F...	true	REC	AV. MOSSEN ...	08552	Taradell	ES

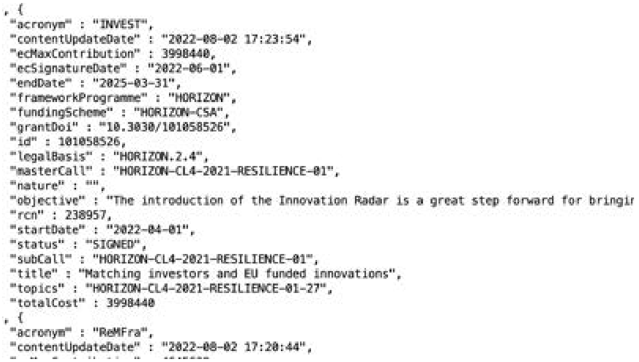
But…

- Need to work with all files in one environment
- Enter the BDTI Framework!

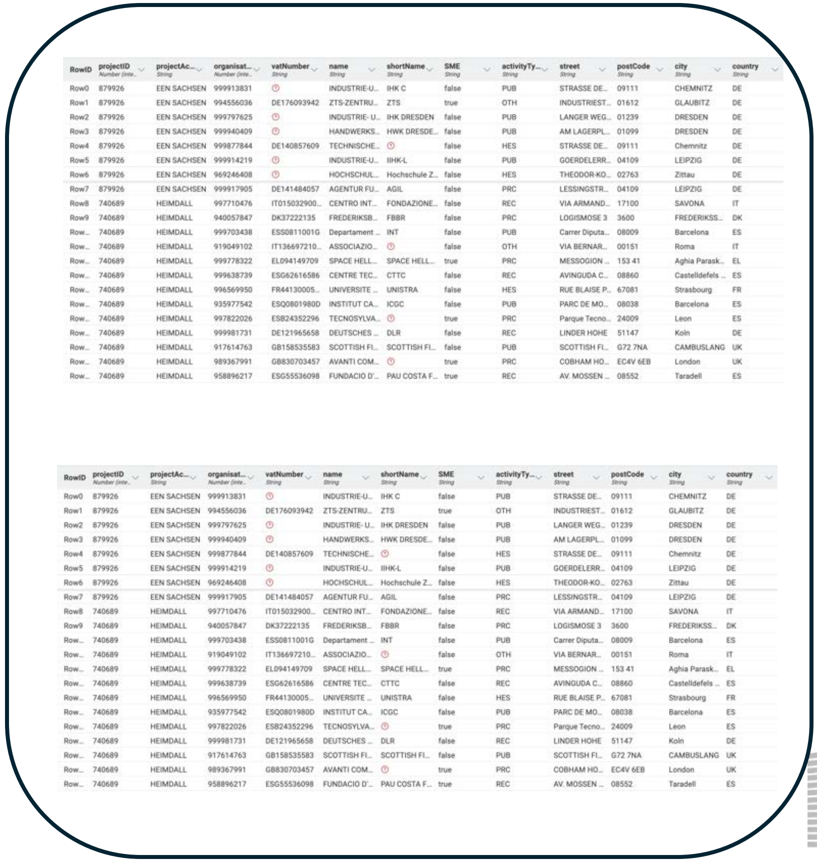


organization_2020.csv

```
"projectID";"projectAcronym";"organisationID";"vatNumber";"name";"shortName";"SME";"activityType";"street";"postCode";"city";"country";"
nutsCode";"geolocation";"organizationURL";"contactForm";"contentUpdateDate";"rcn";"order";"role";"ecContribution";"netEcContribution";"t
otalCost";"endOfParticipation";"active"
"879926";"EEN SACHSEN";"999913831";"";"INDUSTRIE-UND HANDELSKAMMER CHEMNITZ";"IHK C";"false";"PUB";"STRASSE DER NATIONEN
25";"09111";"CHEMNITZ";"DE";"";"50.8367938,12.9249019";"";"https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/
contact-form/project/999913831/879926";"2022-10-28 14:07:26";"2559452";"6";"participant";"0";"0";"0";"false";"false"
"879926";"EEN SACHSEN";"994556036";"DE176093942";"ZTS-ZENTRUM FUR TECHNOLOGIESTRUKTURENTWICKLUNG REGION RIESA-GROSSENHAIN
GMBH";"ZTS";"true";"OTH";"INDUSTRIESTRASSE A 11";"01612";"GLAUBITZ";"DE";"";"51.32823705,13.3706761539466";"";"https://ec.europa.eu/
info/funding-tenders/opportunities/portal/screen/contact-form/project/994556036/879926";"2022-10-28
14:07:26";"2355549";"3";"participant";"31387.14";"31387.14";"31387,5";"false";""
"879926";"EEN SACHSEN";"999797625";"";"INDUSTRIE- UND HANDELSKAMMER DRESDEN";"IHK DRESDEN";"false";"PUB";"LANGER WEG
4";"01239";"DRESDEN";"DE";"";"51.00853155,13.8093885009436";"";"https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/
contact-form/project/999797625/879926";"2022-10-28 14:07:26";"2559453";"7";"participant";"0";"0";"0";"false";"false"
"879926";"EEN SACHSEN";"999940409";"";"HANDWERKSKAMMER DRESDEN";"HWK DRESDEN";"false";"PUB";"AM LAGERPLATZ
8";"01099";"DRESDEN";"DE";"";"51.08344545,13.7678360448861";"";"https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/
contact-form/project/999940409/879926";"2022-10-28 14:07:26";"2559455";"8";"participant";"0";"0";"0";"false";"false"
"879926";"EEN SACHSEN";"999877844";"DE140857609";"TECHNISCHE UNIVERSITAET CHEMNITZ";"";"false";"HES";"STRASSE DER NATIONEN
62";"09111";"Chemnitz";"DE";"";"50.8391352,12.9291884";"http://www.tu-chemnitz.de";"https://ec.europa.eu/info/funding-tenders/
opportunities/portal/screen/contact-form/project/999877844/879926";"2022-10-28
14:07:26";"1915930";"2";"participant";"57798.11";"57798.11";"57798,75";"false";""
"879926";"EEN SACHSEN";"999914219";"";"INDUSTRIE-UND HANDELSKAMMER ZU LEIPZIG";"IIHK-L";"false";"PUB";"GOERDELERRING
5";"04109";"LEIPZIG";"DE";"";"51.3432431,12.3691018";"";"https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/contact-
form/project/999914219/879926";"2022-10-28 14:07:26";"2559451";"5";"participant";"0";"0";"0";"false";"false"
"879926";"EEN SACHSEN";"969246408";"";"HOCHSCHULE ZITTAU/GOERLITZ";"Hochschule Zittau/Goerlitz";"false";"HES";"THEODOR-KOERNER-ALLEE
16";"02763";"Zittau";"DE";"";"50.89336835,14.806004661802392";"www.hszg.de";"https://ec.europa.eu/info/funding-tenders/opportunities/
portal/screen/contact-form/project/969246408/879926";"2022-10-28 14:07:26";"1915282";"4";"participant";"0";"0";"0";"false";"false"
"879926";"EEN SACHSEN";"999917905";"DE141484057";"AGENTUR FUR INNOVATIONSFORDERUNG UND TECHNOLOGIETRANSFER GMBH LEIPZIG(AGIL GMBH
LEIPZIG)";"AGIL";"false";"PRC";"LESSINGSTRASSE 2";"04109";"LEIPZIG";"DE";"";"51.24007265,12.717300256524063";"";"https://ec.europa.eu/
info/funding-tenders/opportunities/portal/screen/contact-form/project/999917905/879926";"2022-10-28
14:07:26";"2355547";"1";"coordinator";"36373.75";"36373.75";"36373,75";"false";""
"740689";"HEIMDALL";"997710476";"IT01503290098";"CENTRO INTERNAZIONALE IN MONITORAGGIO AMBIENTALE - FONDAZIONE CIMA";"FONDAZIONE
CIMA";"false";"REC";"VIA ARMANDO MAGLIOTTO 2 CAMPUS UNIVERSITARIO";"17100";"SAVONA";"IT";"";"44.2334238,8.252572737821862";"http://
www.cimafoundation.org";"https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/contact-form/project/
997710476/740689";"2022-08-23 16:52:15";"1950954";"10";"participant";"514500";"514500";"514875";"false";""
"740689";"HEIMDALL";"940057847";"DK37222135";"FREDERIKSBORG BRAND OG REDNING";"FBBR";"false";"PRC";"LOGISMOSE
3";"3600";"FREDERIKSSUND";"DK";"";"55.8436186,12.0613039";"";"https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/
contact-form/project/940057847/740689";"2022-08-23 16:52:15";"2717172";"12";"participant";"222775";"222775";"318250";"false";""
"740689";"HEIMDALL";"999703438";"ESS0811001G";"Departament d'Interior - Generalitat de Catalunya";"INT";"false";"PUB";"Carrer Diputació
355";"08009";"Barcelona";"ES";"";"41.395639,2.173095";"";"https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/contact-
form/project/999703438/740689";"2022-08-23 16:52:15";"1956494";"11";"participant";"832635";"832635";"833385";"false";""
```



```
, {
 "acronym" : "INVEST",
 "contentUpdateDate" : "2022-08-02 17:23:54",
 "ecMaxContribution" : 3998440,
 "ecSignatureDate" : "2022-06-01",
 "endDate" : "2025-03-31",
 "frameworkProgramme" : "HORIZON",
 "fundingScheme" : "HORIZON-CSA",
 "grantDoi" : "10.3030/101058526",
 "id" : 101058526,
 "legalBasis" : "HORIZON.2.4",
 "masterCall" : "HORIZON-CL4-2021-RESILIENCE-01",
 "nature" : "",
 "objective" : "The introduction of the Innovation Radar is a great step forward for bringin
 "rcn" : 238957,
 "startDate" : "2022-04-01",
 "status" : "SIGNED",
 "subCall" : "HORIZON-CL4-2021-RESILIENCE-01",
 "title" : "Matching investors and EU funded innovations",
 "topics" : "HORIZON-CL4-2021-RESILIENCE-01-27",
 "totalCost" : 3998440
, {
 "acronym" : "ReMFra",
 "contentUpdateDate" : "2022-08-02 17:20:44",
```



RowID	projectID Number (inte…	projectAc… String	organisat… Number (inte…	vatNumber String	name String	shortName String	SME String	activityTy… String	street String	postCode String	city String	country String
Row0	879926	EEN SACHSEN	999913831	?	INDUSTRIE-U…	IHK C	false	PUB	STRASSE DE…	09111	CHEMNITZ	DE
Row1	879926	EEN SACHSEN	994556036	DE176093942	ZTS-ZENTRU…	ZTS	true	OTH	INDUSTRIEST…	01612	GLAUBITZ	DE
Row2	879926	EEN SACHSEN	999797625	?	INDUSTRIE- U…	IHK DRESDEN	false	PUB	LANGER WEG…	01239	DRESDEN	DE
Row3	879926	EEN SACHSEN	999940409	?	HANDWERKS…	HWK DRESDE…	false	PUB	AM LAGERPL…	01099	DRESDEN	DE
Row4	879926	EEN SACHSEN	999877844	DE140857609	TECHNISCHE…	?	false	HES	STRASSE DE…	09111	Chemnitz	DE
Row5	879926	EEN SACHSEN	999914219	?	INDUSTRIE-U…	IIHK-L	false	PUB	GOERDELERR…	04109	LEIPZIG	DE
Row6	879926	EEN SACHSEN	969246408	?	HOCHSCHUL…	Hochschule Z…	false	HES	THEODOR-KO…	02763	Zittau	DE
Row7	879926	EEN SACHSEN	999917905	DE141484057	AGENTUR FU…	AGIL	false	PRC	LESSINGSTR…	04109	LEIPZIG	DE
Row8	740689	HEIMDALL	997710476	IT015032900…	CENTRO INT…	FONDAZIONE…	false	REC	VIA ARMAND…	17100	SAVONA	IT
Row9	740689	HEIMDALL	940057847	DK37222135	FREDERIKSB…	FBBR	false	PRC	LOGISMOSE 3	3600	FREDERIKSS…	DK
Row…	740689	HEIMDALL	999703438	ESS0811001G	Departament …	INT	false	PUB	Carrer Diputa…	08009	Barcelona	ES
Row…	740689	HEIMDALL	919049102	IT136697210…	ASSOCIAZIO…	?	false	OTH	VIA BERNAR…	00151	Roma	IT
Row…	740689	HEIMDALL	999778322	EL094149709	SPACE HELL…	SPACE HELL…	true	PRC	MESSOGION …	153 41	Aghia Parask…	EL
Row…	740689	HEIMDALL	999638739	ESG62616586	CENTRE TEC…	CTTC	false	REC	AVINGUDA C…	08860	Castelldefels …	ES
Row…	740689	HEIMDALL	996569950	FR44130005…	UNIVERSITE …	UNISTRA	false	HES	RUE BLAISE P…	67081	Strasbourg	FR
Row…	740689	HEIMDALL	935977542	ESQ0801980D	INSTITUT CA…	ICGC	false	PUB	PARC DE MO…	08038	Barcelona	ES
Row…	740689	HEIMDALL	997822026	ESB24352296	TECNOSYLVA…	?	true	PRC	Parque Tecno…	24009	Leon	ES
Row…	740689	HEIMDALL	999981731	DE121965658	DEUTSCHES …	DLR	false	REC	LINDER HOHE	51147	Koln	DE
Row…	740689	HEIMDALL	917614763	GB158535583	SCOTTISH FI…	SCOTTISH FI…	false	PUB	SCOTTISH FI…	G72 7NA	CAMBUSLANG	UK
Row…	740689	HEIMDALL	989367991	GB830703457	AVANTI COM…	?	true	PRC	COBHAM HO…	EC4V 6EB	London	UK
Row…	740689	HEIMDALL	958896217	ESG55536098	FUNDACIO D'…	PAU COSTA F…	true	REC	AV. MOSSEN …	08552	Taradell	ES

RowID	projectID Number (inte…	projectAc… String	organisat… Number (inte…	vatNumber String	name String	shortName String	SME String	activityTy… String	street String	postCode String	city String	country String
Row0	879926	EEN SACHSEN	999913831	?	INDUSTRIE-U…	IHK C	false	PUB	STRASSE DE…	09111	CHEMNITZ	DE
Row1	879926	EEN SACHSEN	994556036	DE176093942	ZTS-ZENTRU…	ZTS	true	OTH	INDUSTRIEST…	01612	GLAUBITZ	DE
Row2	879926	EEN SACHSEN	999797625	?	INDUSTRIE- U…	IHK DRESDEN	false	PUB	LANGER WEG…	01239	DRESDEN	DE
Row3	879926	EEN SACHSEN	999940409	?	HANDWERKS…	HWK DRESDE…	false	PUB	AM LAGERPL…	01099	DRESDEN	DE
Row4	879926	EEN SACHSEN	999877844	DE140857609	TECHNISCHE…	?	false	HES	STRASSE DE…	09111	Chemnitz	DE
Row5	879926	EEN SACHSEN	999914219	?	INDUSTRIE-U…	IIHK-L	false	PUB	GOERDELERR…	04109	LEIPZIG	DE
Row6	879926	EEN SACHSEN	969246408	?	HOCHSCHUL…	Hochschule Z…	false	HES	THEODOR-KO…	02763	Zittau	DE
Row7	879926	EEN SACHSEN	999917905	DE141484057	AGENTUR FU…	AGIL	false	PRC	LESSINGSTR…	04109	LEIPZIG	DE
Row8	740689	HEIMDALL	997710476	IT015032900…	CENTRO INT…	FONDAZIONE…	false	REC	VIA ARMAND…	17100	SAVONA	IT
Row9	740689	HEIMDALL	940057847	DK37222135	FREDERIKSB…	FBBR	false	PRC	LOGISMOSE 3	3600	FREDERIKSS…	DK
Row…	740689	HEIMDALL	999703438	ESS0811001G	Departament …	INT	false	PUB	Carrer Diputa…	08009	Barcelona	ES
Row…	740689	HEIMDALL	919049102	IT136697210…	ASSOCIAZIO…	?	false	OTH	VIA BERNAR…	00151	Roma	IT
Row…	740689	HEIMDALL	999778322	EL094149709	SPACE HELL…	SPACE HELL…	true	PRC	MESSOGION …	153 41	Aghia Parask…	EL
Row…	740689	HEIMDALL	999638739	ESG62616586	CENTRE TEC…	CTTC	false	REC	AVINGUDA C…	08860	Castelldefels …	ES
Row…	740689	HEIMDALL	996569950	FR44130005…	UNIVERSITE …	UNISTRA	false	HES	RUE BLAISE P…	67081	Strasbourg	FR
Row…	740689	HEIMDALL	935977542	ESQ0801980D	INSTITUT CA…	ICGC	false	PUB	PARC DE MO…	08038	Barcelona	ES
Row…	740689	HEIMDALL	997822026	ESB24352296	TECNOSYLVA…	?	true	PRC	Parque Tecno…	24009	Leon	ES
Row…	740689	HEIMDALL	999981731	DE121965658	DEUTSCHES …	DLR	false	REC	LINDER HOHE	51147	Koln	DE
Row…	740689	HEIMDALL	917614763	GB158535583	SCOTTISH FI…	SCOTTISH FI…	false	PUB	SCOTTISH FI…	G72 7NA	CAMBUSLANG	UK
Row…	740689	HEIMDALL	989367991	GB830703457	AVANTI COM…	?	true	PRC	COBHAM HO…	EC4V 6EB	London	UK
Row…	740689	HEIMDALL	958896217	ESG55536098	FUNDACIO D'…	PAU COSTA F…	true	REC	AV. MOSSEN …	08552	Taradell	ES

Session 1:
Data Access and Exploration

Session 1: Learning outcomes

At the end of this session, you should be able to:

- Access BDTI framework

- Navigate BDTI Service Catalog

- Launch various data analytic tools:
    - KNIME Analytics Platform,
    - R-studio,
    - Jupyter Notebook.

- Read different data file types into the analytic tools:
    - CSV,
    - Excel,
    - JSON.

- Complete basic data exploration in relation to the use case

How to access the BDTI framework?

- To have access, public administrations need to ***apply on the official website***

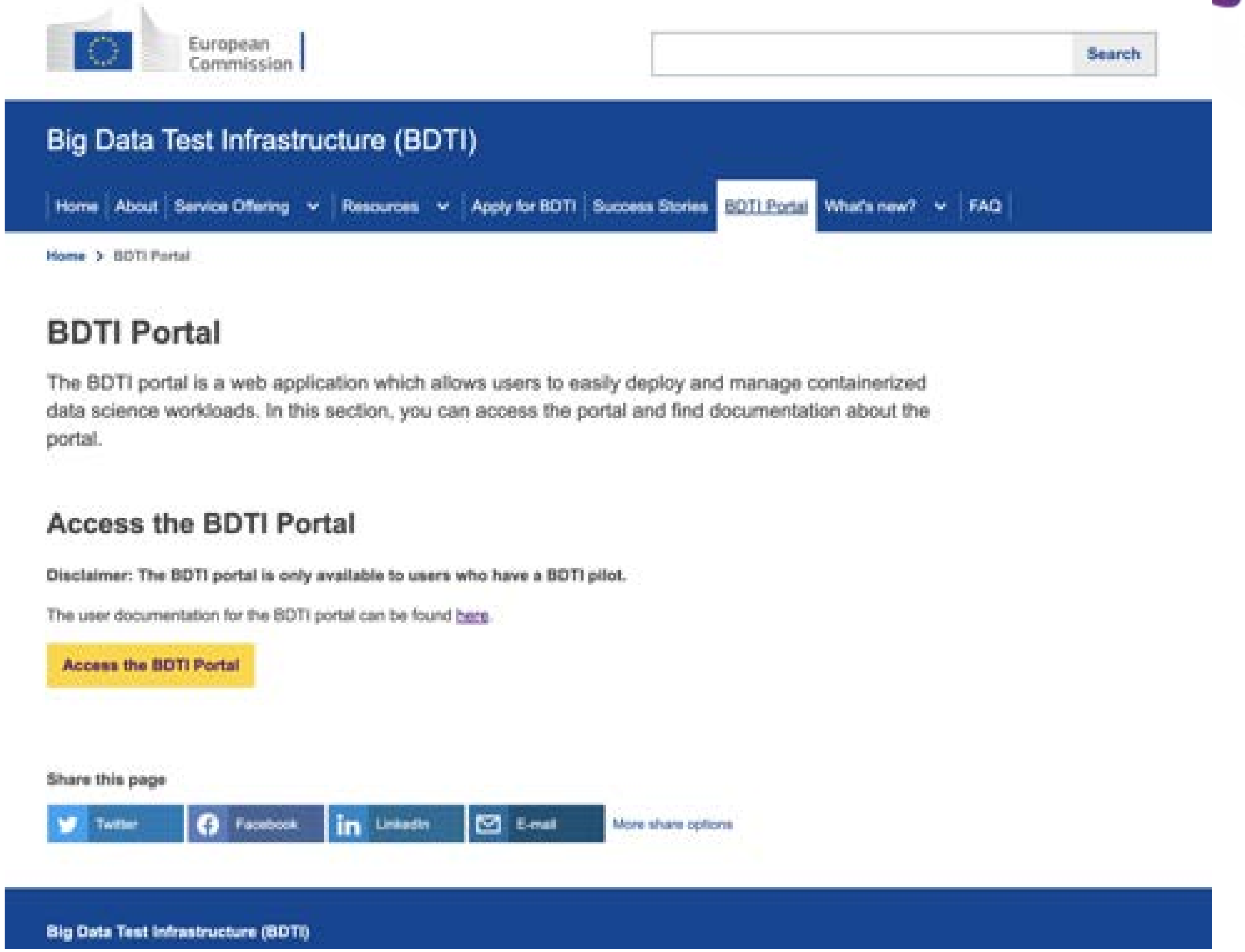
https://big-data-test-infrastructure.ec.europa.eu/apply-bdti_en

- Once you have access follow these steps:
  - Step 1 > Access the portal via the website
  - Step 2 > Navigate to the "Service Catalog" and launch the tools
  - Step 3 > Go to "My Services" and open the tool you created

Tour of the BDTI framework

Step 1:

Once you applied succesfully,
Access the portal via the website



European Commission

Search

Big Data Test Infrastructure (BDTI)

Home > BDTI Portal

BDTI Portal

The BDTI portal is a web application which allows users to easily deploy and manage containerized data science workloads. In this section, you can access the portal and find documentation about the portal.

Access the BDTI Portal

Disclaimer: The BDTI portal is only available to users who have a BDTI pilot.

The user documentation for the BDTI portal can be found here.

Access the BDTI Portal

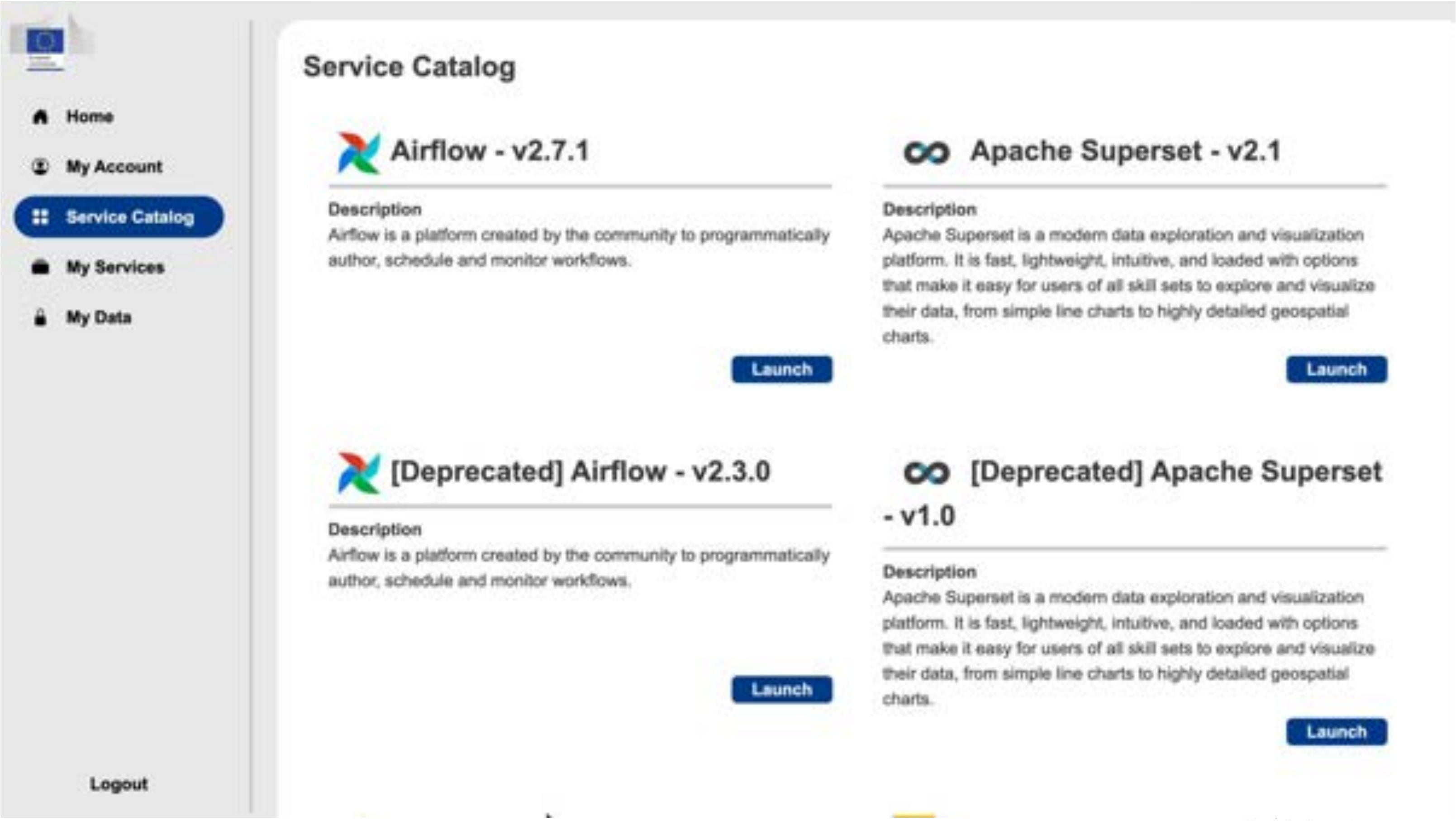
Share this page

Big Data Test Infrastructure (BDTI)

Tour of the BDTI framework

Step 2:

Navigate to the "Service Catalog" and launch the desired tools



Service Catalog

Airflow - v2.7.1

Description

Airflow is a platform created by the community to programmatically author, schedule and monitor workflows.

Apache Superset - v2.1

Description

Apache Superset is a modern data exploration and visualization platform. It is fast, lightweight, intuitive, and loaded with options that make it easy for users of all skill sets to explore and visualize their data, from simple line charts to highly detailed geospatial charts.

[Deprecated] Airflow - v2.3.0

Description

Airflow is a platform created by the community to programmatically author, schedule and monitor workflows.

[Deprecated] Apache Superset - v1.0

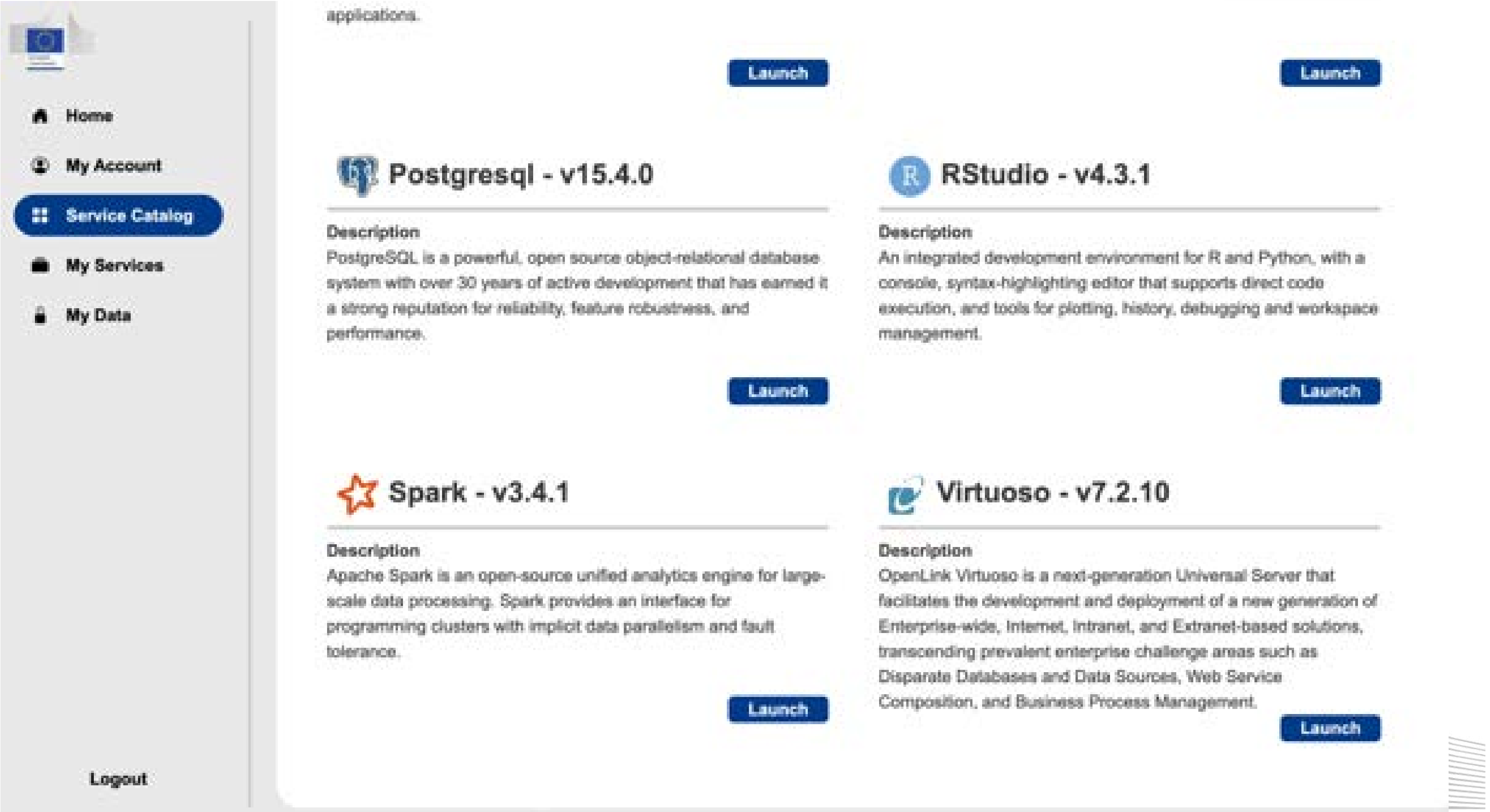
Description

Apache Superset is a modern data exploration and visualization platform. It is fast, lightweight, intuitive, and loaded with options that make it easy for users of all skill sets to explore and visualize their data, from simple line charts to highly detailed geospatial charts.

Tour of the BDTI framework

Step 3:

Go to "My Services" and open the tool you created



applications.

Launch

Launch

Postgresql - v15.4.0

Description
PostgreSQL is a powerful, open source object-relational database system with over 30 years of active development that has earned it a strong reputation for reliability, feature robustness, and performance.

Launch

RStudio - v4.3.1

Description
An integrated development environment for R and Python, with a console, syntax-highlighting editor that supports direct code execution, and tools for plotting, history, debugging and workspace management.

Launch

Spark - v3.4.1

Description
Apache Spark is an open-source unified analytics engine for large-scale data processing. Spark provides an interface for programming clusters with implicit data parallelism and fault tolerance.

Launch

Virtuoso - v7.2.10

Description
OpenLink Virtuoso is a next-generation Universal Server that facilitates the development and deployment of a new generation of Enterprise-wide, Internet, Intranet, and Extranet-based solutions, transcending prevalent enterprise challenge areas such as Disparate Databases and Data Sources, Web Service Composition, and Business Process Management.

Launch

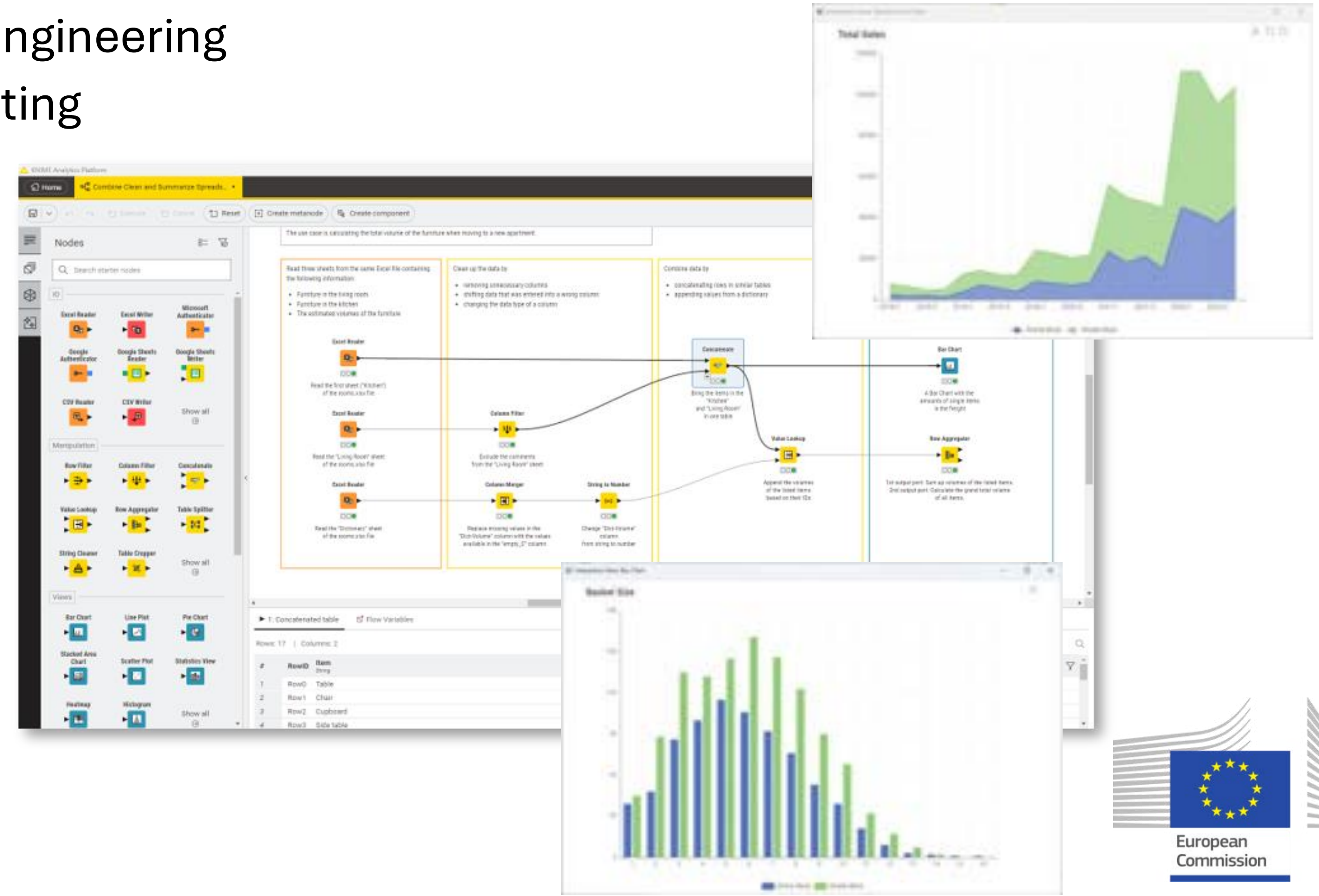
All these tools can be downloaded locally

As the nature of these tools is ***open-source***, in case you are not a public administration interested in apply for the BDTI playground, participants can download them separately before getting access to the BDTI, that way you can follow along with the sessions and practice at home!

- KNIME Analytic Platform: https://www.knime.com/downloads
- R-studio: https://cran.r-project.org/
- Jupyter Notebook: https://jupyter.org/install

What is KNIME Analytics Platform

- A tool to make sense of your data
  - Data analysis, Data Science, Data Engineering
  - Transformation, visualisation, reporting
- Open source
- Visual programming paradigm
  - No coding required



The use case is calculating the total volume of the furniture when moving to a new apartment.

Read three sheets from the same Excel file containing the following information:
- Furniture in the living room
- Furniture in the kitchen
- The estimated volumes of the furniture

Clean up the data by
- removing unnecessary columns
- shifting data that was entered into a wrong column
- changing the data type of a column

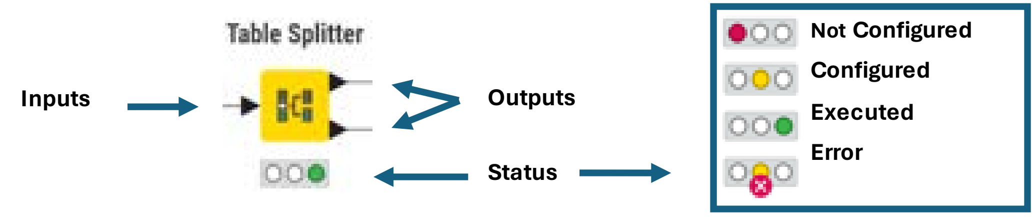
Combine data by
- concatenating rows in similar tables
- appending values from a dictionary

Rows: 17 | Columns: 2

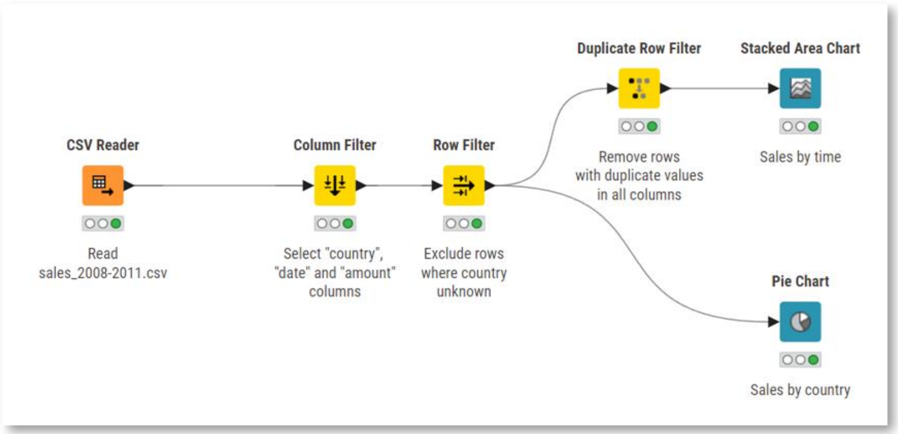
#	RowID	Item
1	Row0	Table
2	Row1	Chair
3	Row2	Cupboard
4	Row3	Side table

Nodes and Workflows

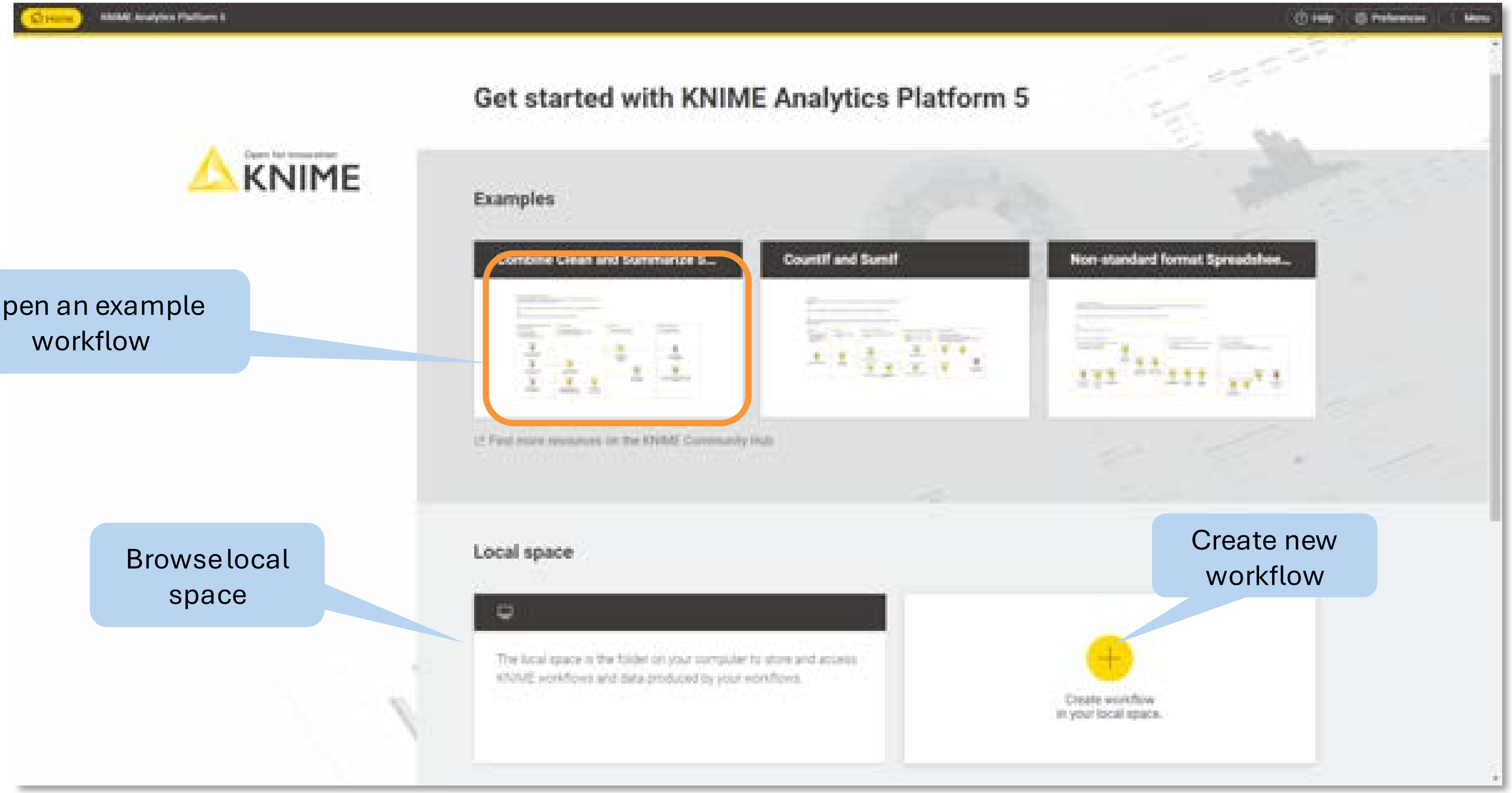
- **NODES** perform tasks on data



- **Nodes** are combined to create WORKFLOWS



Tour of the User Interface



Get started with KNIME Analytics Platform 5

Find more resources on the KNIME Community Hub

The local space is the folder on your computer to store and access KNIME workflows and data produced by your workflows.

Open an example workflow

Browse local space

Create new workflow

Data Access

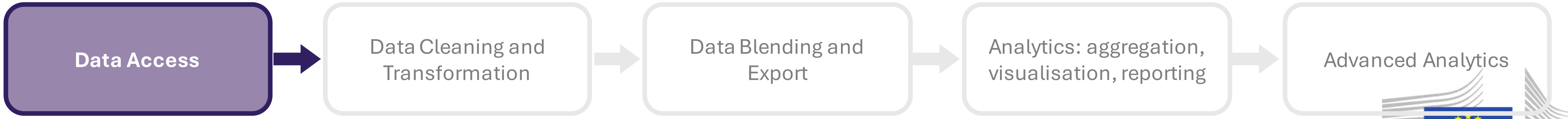
- The beginning of every data process
- Data can be stored in many ways
  - Locally
  - In different data format (.csv, .xls…)
  - On the cloud
  - In a database
  - …
- We need a way to access all of them

Use Case

Zoi and her team are excited to use the BDTI framework. They found the dataset "Horizon 2020", which contains the research and innovation funding of **energy** projects, and the matching table on data.europa.eu.

The dataset on the "CO2 emissions" was also downloaded online.

The first task they would like to learn is how to access and work with the different file types.


Data Access

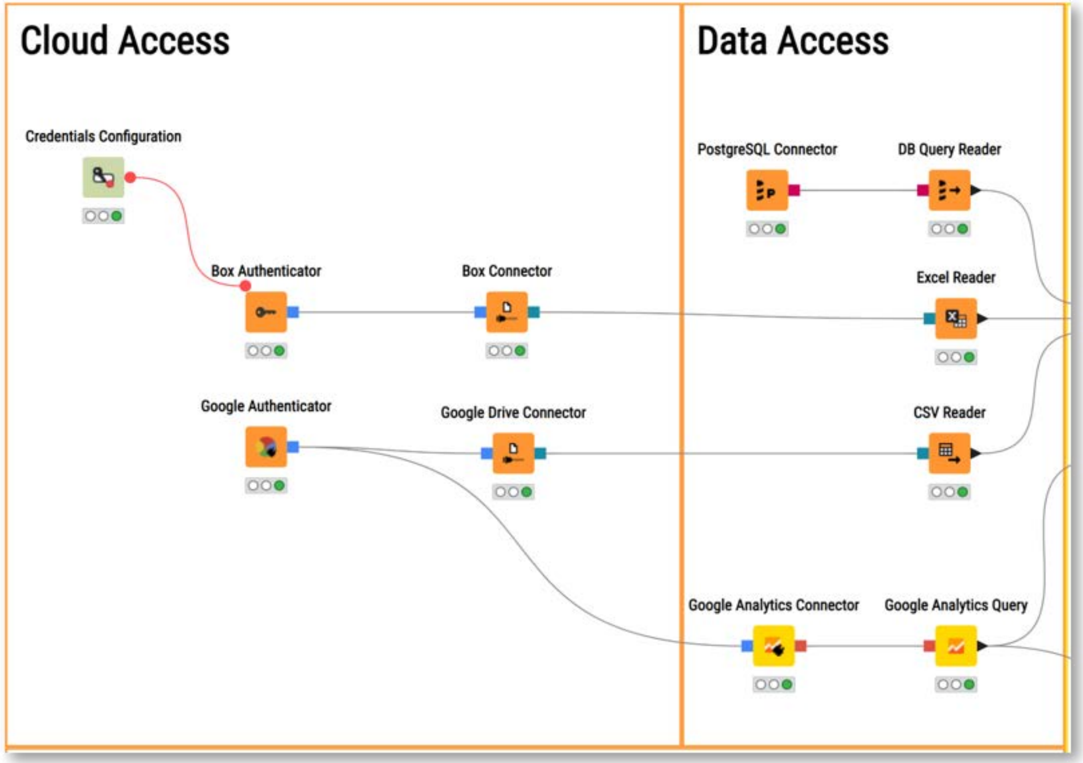
Data Cleaning and Transformation

Data Blending and Export

Analytics: aggregation, visualisation, reporting

Advanced Analytics

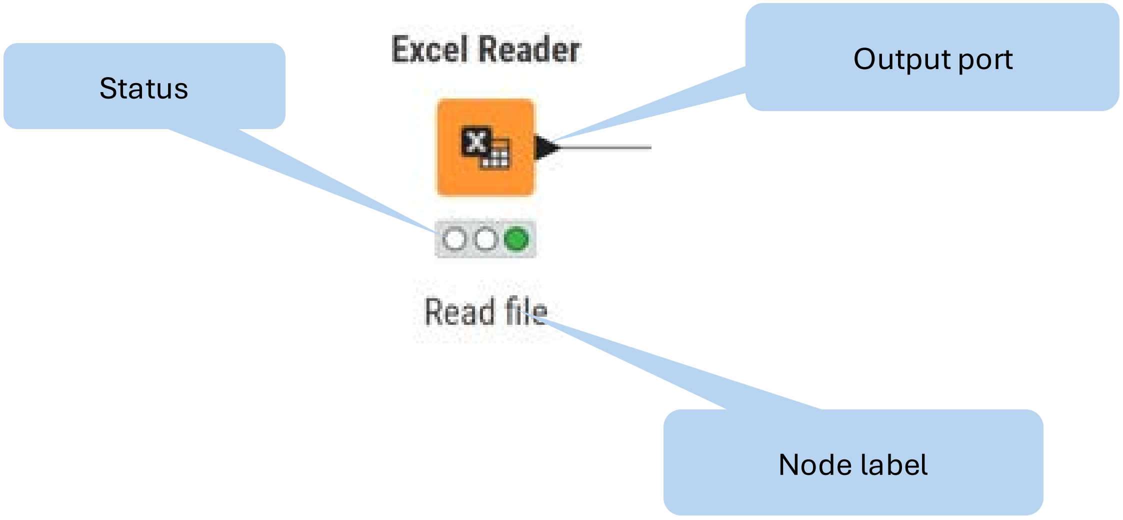
Data Access in KNIME Analytics Platform



- Files
  - CSV, txt, Excel, Word, PDF
  - SAS, SPSS
  - XML, JSON, PMML
  - Images, texts, networks
- Databases
  - MySQL, PostgreSQL, Oracle
  - Theobald
  - any JDBC (DB2, MS SQL Server)
  - Amazon DynamoDB
- Other
  - Twitter, Google
  - Amazon S3, Azure Blob Store
  - Sharepoint, Salesforce
  - Kafka
  - REST, Web services

Data Access in KNIME Analytics Platform

- Typically characterised by:
- Orange color
- By default no input ports, 1-2 output ports
- Many nodes for many data formats
- Support reading from different File Systems



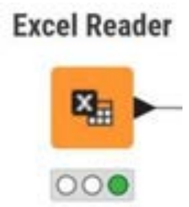
Status

Output port

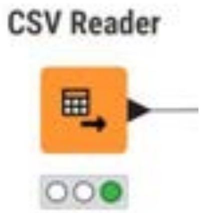
Node label

File types in Use Case

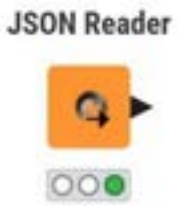
- Excel Files



- CSV Files

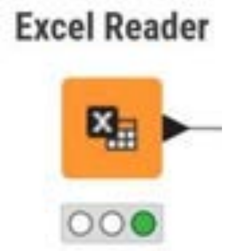


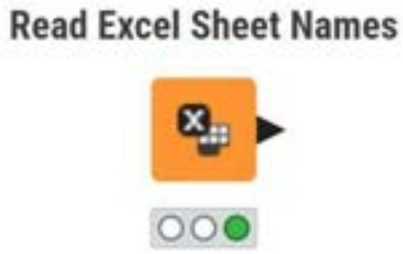
- JSON Files



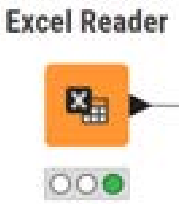
Read Excel files

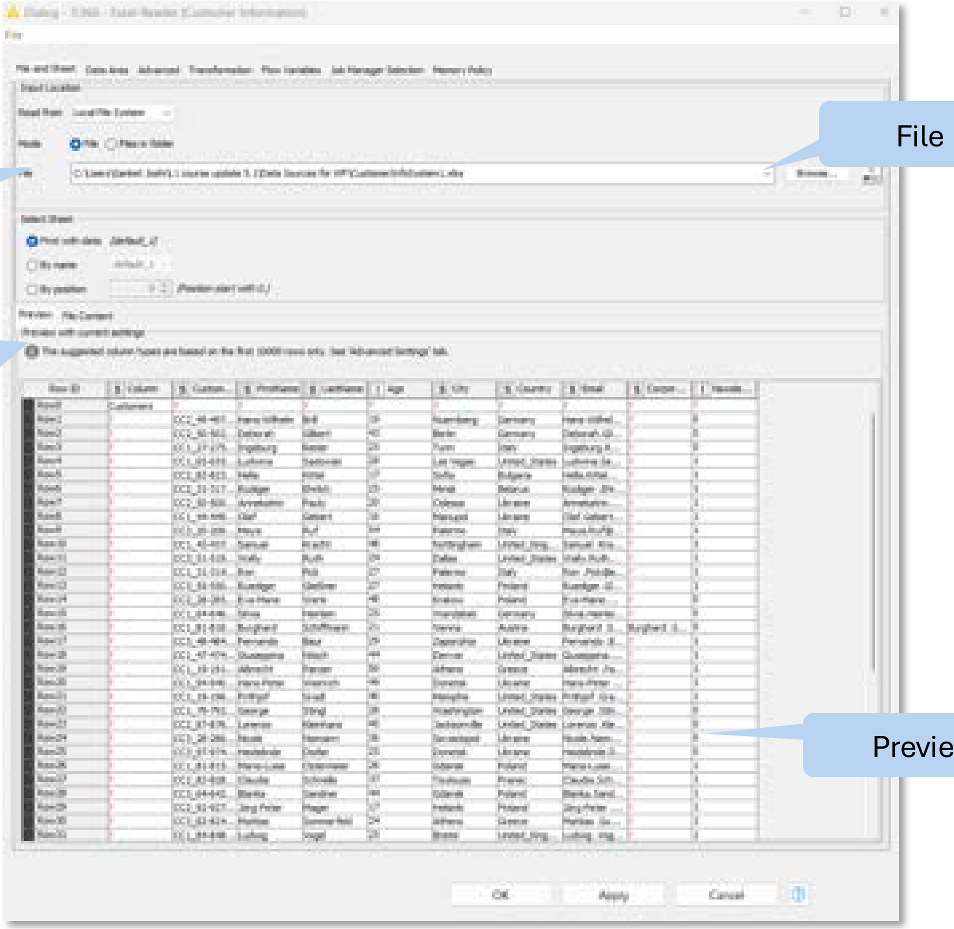
- Excel Reader - reads .xls and .xlsx file from Microsoft Excel
- Supports reading from multiple sheets





Excel Reader - Configuration





File and Sheet | Data Area | Advanced | Transformation | Flow Variables | Job Manager Selection | Memory Policy

Input Location

Read from: Local File System

Mode: ● File ○ Files in folder

Select Sheet

● First with data (default_1)

○ By name

○ By position (Position start with 0)

Preview | File Content

Preview with current settings

The suggested column types are based on the first 10000 rows only. See 'Advanced Settings' tab.

File system

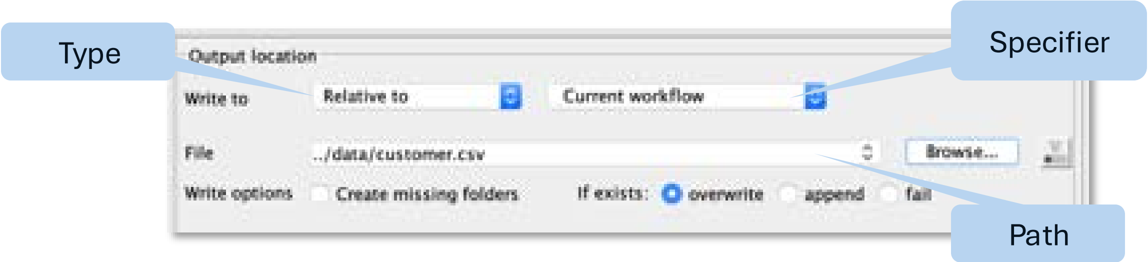
File path

Sheet specific settings

Preview

Common Settings: File Path

- A path consists of three parts:
  - **Type**: Specifies the file system type - e.g., local, relative, mountpoint, custom URL or connected
  - **Specifier**: Optional string with additional file system specific information - e.g. relative to which location (knime.workflow, LOCAL mountpoint…)
  - **Path**: Specifies the location within the file system



Output location

Write to: Relative to | Current workflow

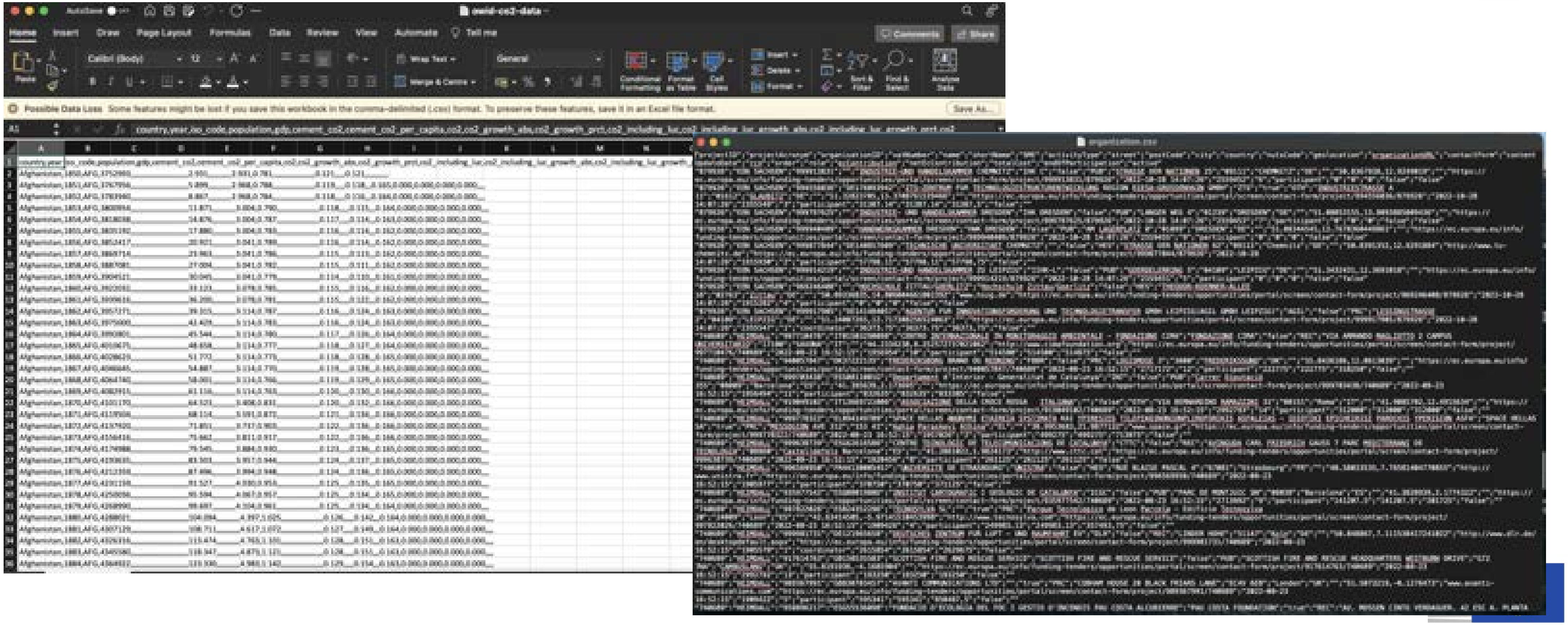
File: ../data/customer.csv [Browse...]

Write options: □ Create missing folders   If exists: ● overwrite ○ append ○ fail

- Examples:
  - (LOCAL, , C:\Users\username\Desktop)
  - (RELATIVE, knime.workflow, file1.csv)
  - (MOUNTPOINT, MOUNTPOINT_NAME, /path/to/file1.csv)
  - (CONNECTED, amazon-s3:eu-west-1, /mybucket/file1.csv)

Read CSV files

- CSV – Comma-separated Values

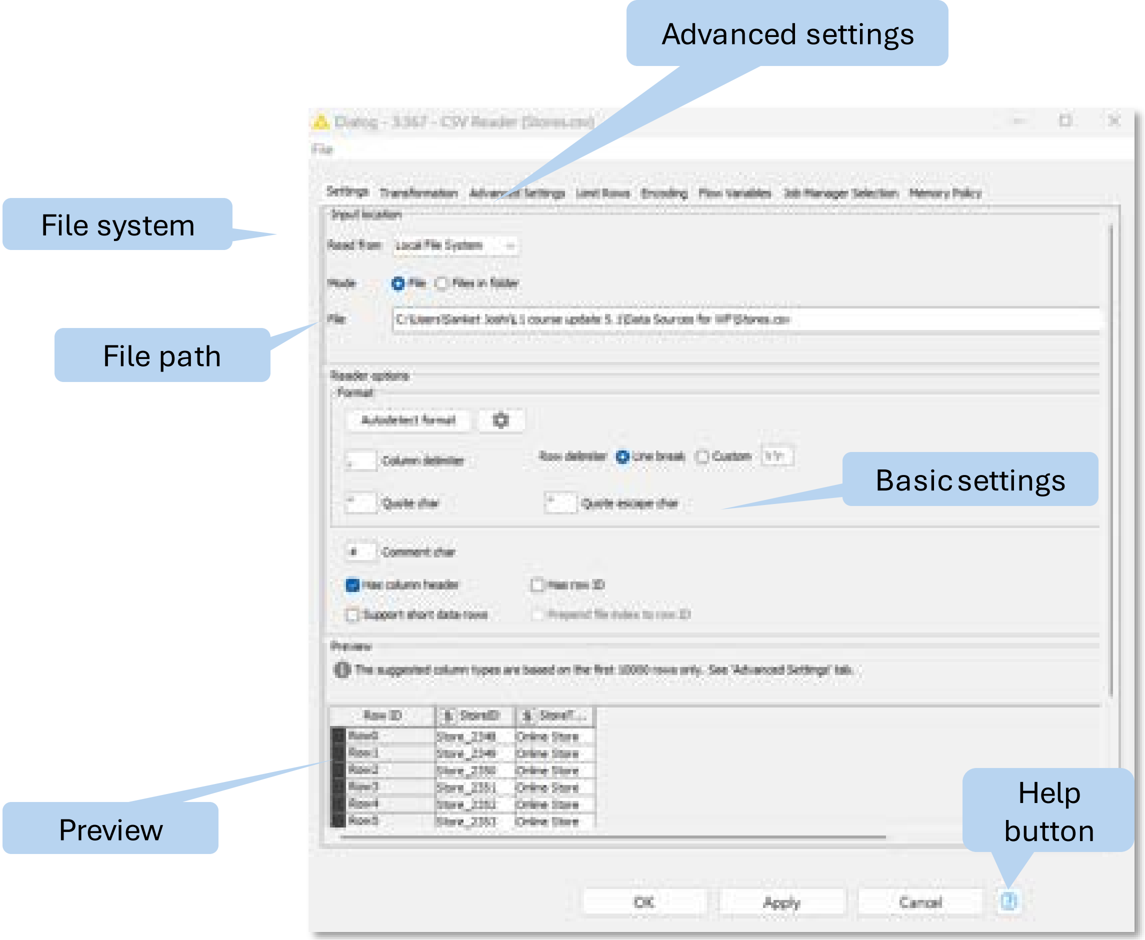


Possible Data Loss Some features might be lost if you save this workbook in the comma-delimited (.csv) format. To preserve these features, save it in an Excel file format.

```
country,year,iso_code,population,gdp,cement_co2,cement_co2_per_capita,co2,co2_growth_abs,co2_growth_prct,co2_including_luc,co2_including_luc_growth_abs,co2_including_luc_growth_prct,co2
```

Read CSV files

- Reads either one or multiple .csv and .txt files
- Further tabs to
  - Select columns
  - Limit the rows
  - Handle quotes
  - Select encoding


Advanced settings

File system

File path

Basic settings

Preview

Help button

Dialog - 3:367 - CSV Reader (Stores.csv)

File

Input location

Read from: Local File System

Mode: ● File ○ Files in folder

File: C:\Users\Sanket Joshi\L1 course update 5.1\Data Sources for WP\Stores.csv

Reader options

Format

, Column delimiter

Row delimiter ● Line break ○ Custom \r\n

" Quote char

" Quote escape char

# Comment char

☑ Has column header

☐ Has row ID

☐ Support short data rows

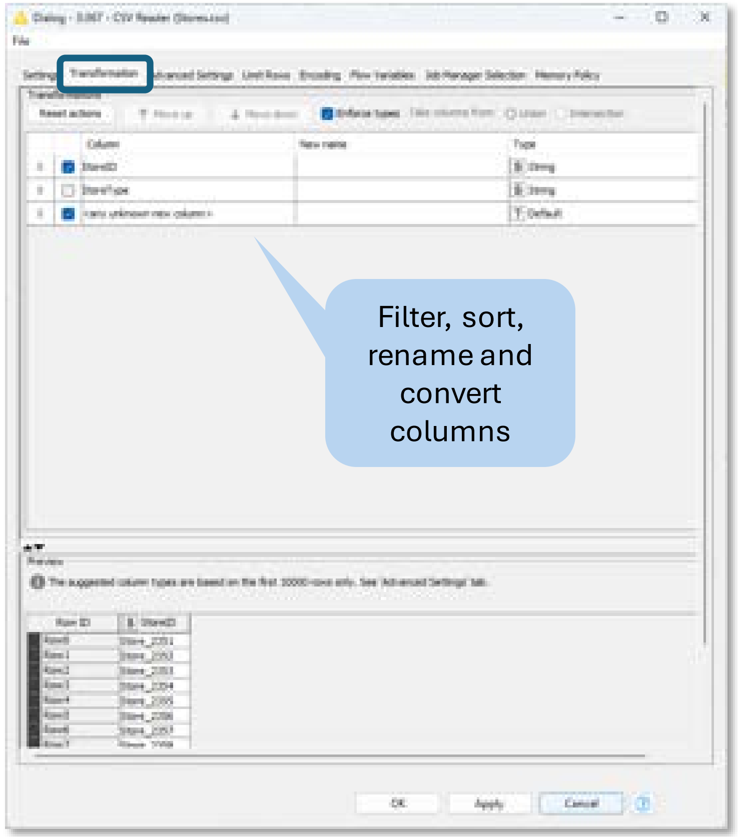
☐ Prepend file index to row ID

Preview

The suggested column types are based on the first 10000 rows only. See "Advanced Settings" tab.

Row ID	S StoreID	S StoreT...
Row0	Store_2348	Online Store
Row1	Store_2349	Online Store
Row2	Store_2350	Online Store
Row3	Store_2351	Online Store
Row4	Store_2352	Online Store
Row5	Store_2353	Online Store

Read CSV files



Dialog - 3:867 - CSV Reader (Stores.csv)

File

Settings | Transformation | Advanced Settings | Limit Rows | Encoding | Flow Variables | Job Manager Selection | Memory Policy

Transformations

Reset actions | Move up | Move down | ☑ Enforce types | Take columns from: ○ Union ○ Intersection

		Column	New name	Type
1	☑	StoreID		S String
2	□	StoreType		S String
3	☑	<any unknown new column>		T Default

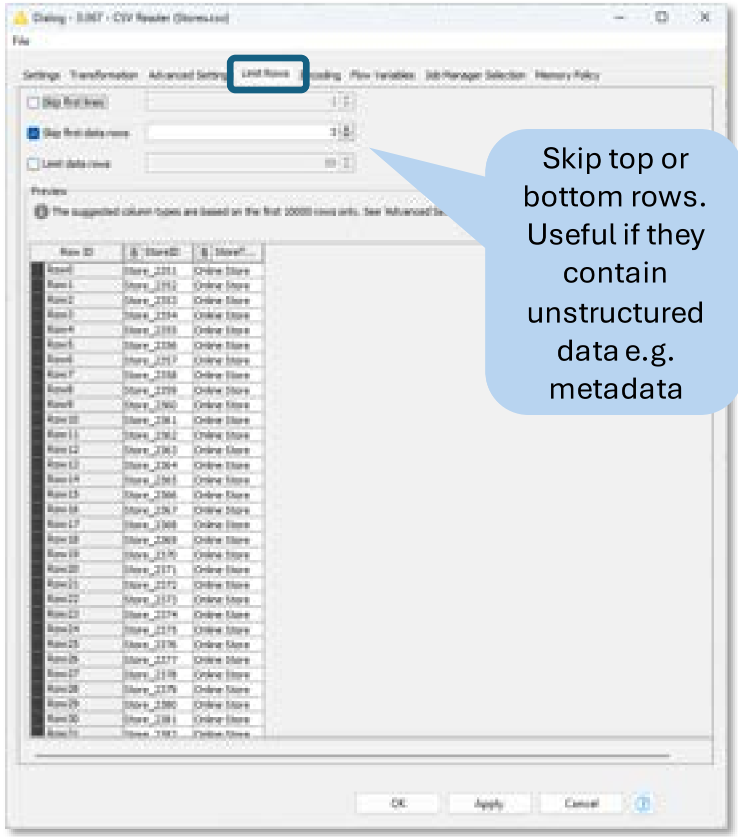
Preview

ⓘ The suggested column types are based on the first 10000 rows only. See 'Advanced Settings' tab.

Row ID	S StoreID
Row0	Store_2351
Row1	Store_2352
Row2	Store_2353
Row3	Store_2354
Row4	Store_2355
Row5	Store_2356
Row6	Store_2357
Row7	Store_2358

OK | Apply | Cancel

Filter, sort, rename and convert columns



Dialog - 3:867 - CSV Reader (Stores.csv)

File

Settings | Transformation | Advanced Settings | Limit Rows | Encoding | Flow Variables | Job Manager Selection | Memory Policy

□ Skip first lines 1

☑ Skip first data rows 1

□ Limit data rows 50

Preview

ⓘ The suggested column types are based on the first 10000 rows only. See 'Advanced Se

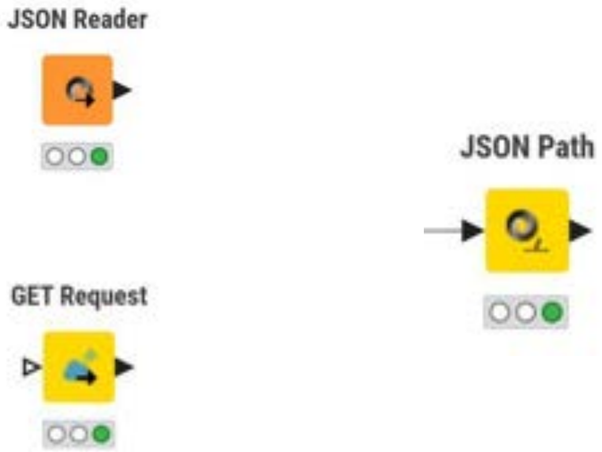
Row ID	S StoreID	S StoreT...
Row0	Store_2351	Online Store
Row1	Store_2352	Online Store
Row2	Store_2353	Online Store
Row3	Store_2354	Online Store
Row4	Store_2355	Online Store
Row5	Store_2356	Online Store
Row6	Store_2357	Online Store
Row7	Store_2358	Online Store
Row8	Store_2359	Online Store
Row9	Store_2360	Online Store
Row10	Store_2361	Online Store
Row11	Store_2362	Online Store
Row12	Store_2363	Online Store
Row13	Store_2364	Online Store
Row14	Store_2365	Online Store
Row15	Store_2366	Online Store
Row16	Store_2367	Online Store
Row17	Store_2368	Online Store
Row18	Store_2369	Online Store
Row19	Store_2370	Online Store
Row20	Store_2371	Online Store
Row21	Store_2372	Online Store
Row22	Store_2373	Online Store
Row23	Store_2374	Online Store
Row24	Store_2375	Online Store
Row25	Store_2376	Online Store
Row26	Store_2377	Online Store
Row27	Store_2378	Online Store
Row28	Store_2379	Online Store
Row29	Store_2380	Online Store
Row30	Store_2381	Online Store
Row31	Store_2382	Online Store

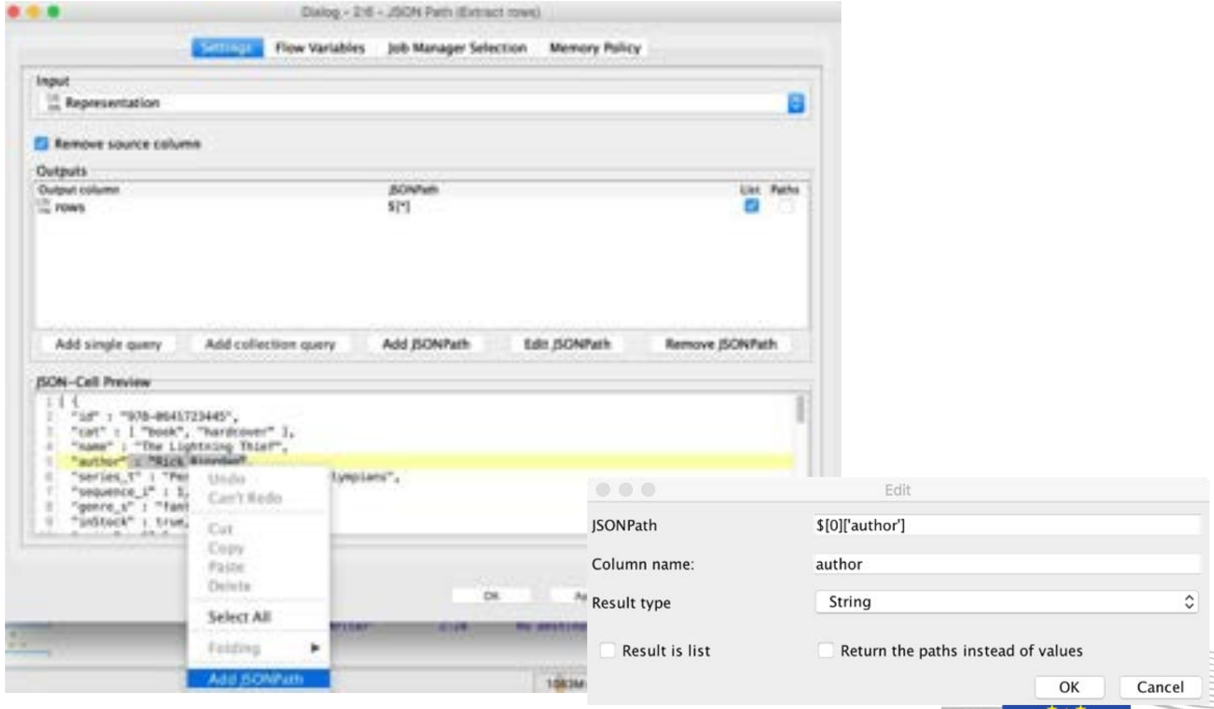
OK | Apply | Cancel

Skip top or bottom rows. Useful if they contain unstructured data e.g. metadata

Read JSON format

- Use the JSON Reader (or GET Request) node to get a JSON cell
- Use the JSON Path node to query the JSON file and extract parameters
  - Editor window simplifies construction of JSON queries by auto-generating them





Input: Representation

☑ Remove source column

Outputs

Output column	JSONPath	List	Paths
rows	$[*]	☑	□

JSON-Cell Preview

```
[ {
  "id" : "978-0641723445",
  "cat" : [ "book", "hardcover" ],
  "name" : "The Lightning Thief",
  "author" : "Rick Riordan",
  "series_t" : "Per[illegible]lympians",
  "sequence_i" : 1,
  "genre_s" : "fant[illegible]
  "inStock" : true,
```

JSONPath: $[0]['author']

Column name: author

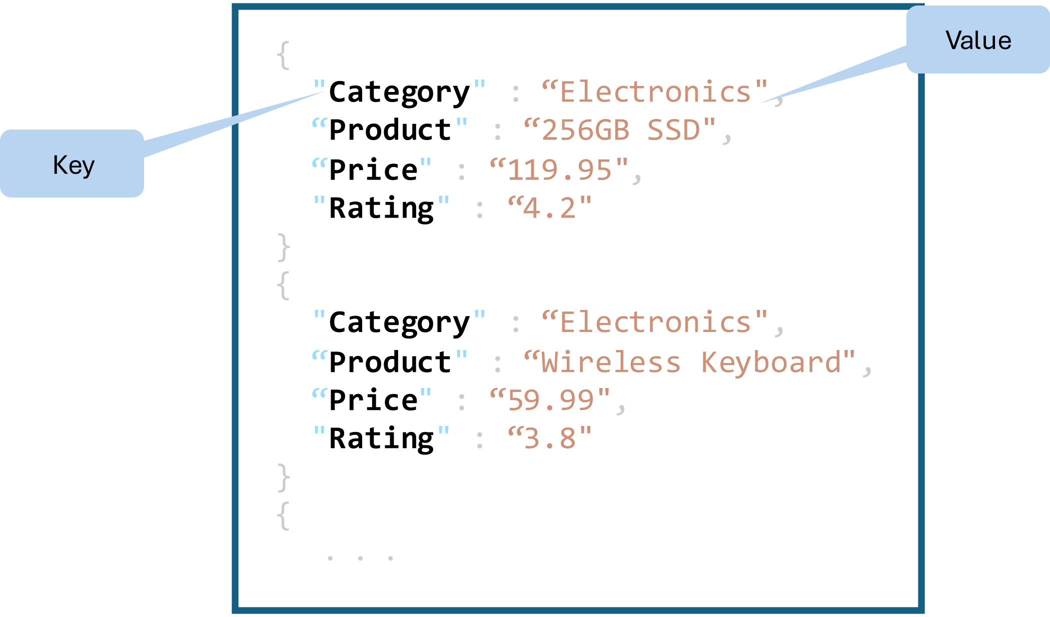
Result type: String

□ Result is list

□ Return the paths instead of values

JSON Objects

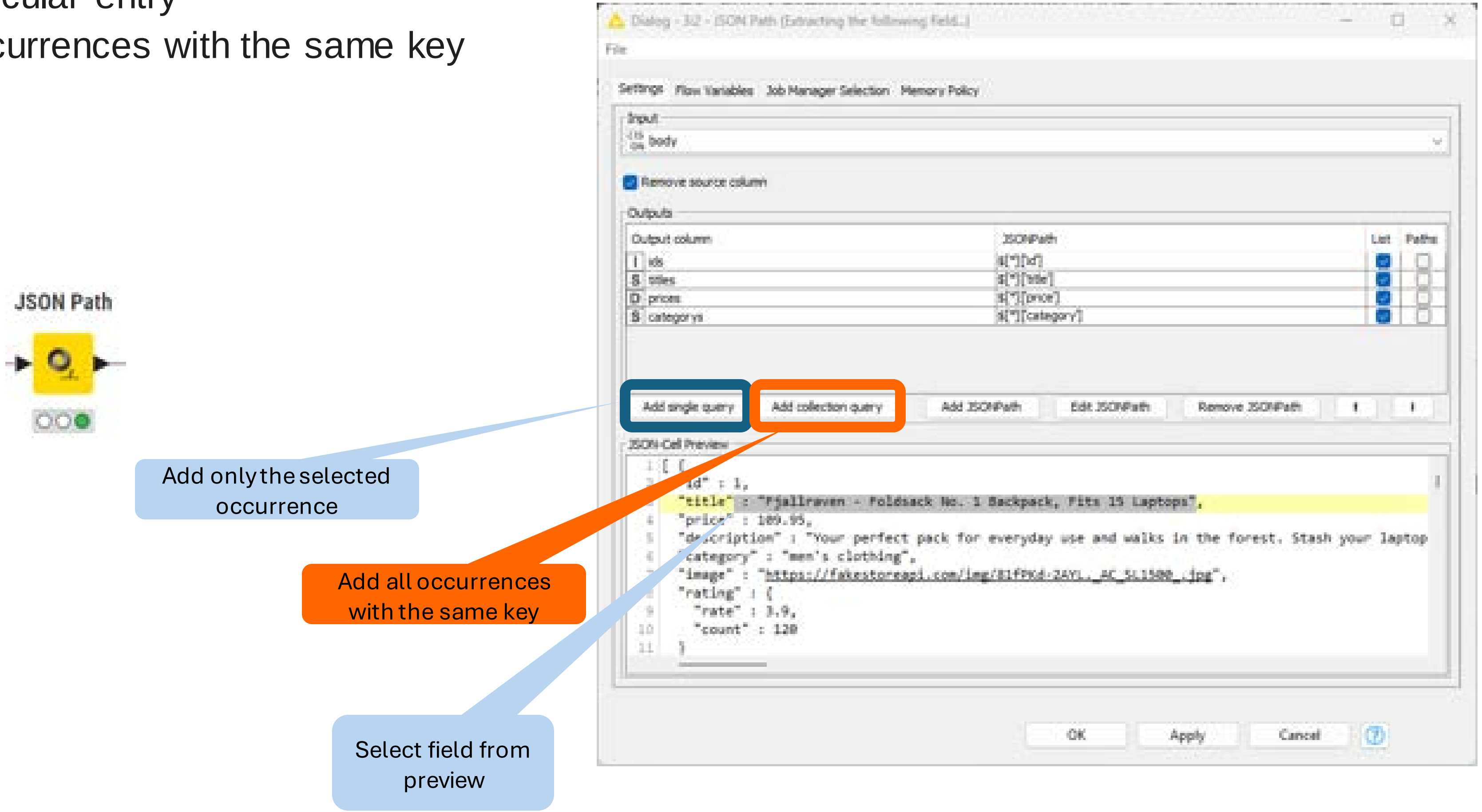
- Stores data in a key-value pair format (e.g., “Category”: “Electronics”)
- Supports nested structures
    - JSON objects inside JSON objects



```
{
  "Category" : “Electronics",
  “Product" : “256GB SSD",
  “Price" : “119.95",
  "Rating" : “4.2"
}
{
  "Category" : “Electronics",
  “Product" : “Wireless Keyboard",
  “Price" : “59.99",
  "Rating" : “3.8"
}
{
  . . .
```

JSON Path

- Extracts fields of interest from JSON
  - A particular entry
  - All occurrences with the same key



Dialog - 3:2 - JSON Path (Extracting the following field...)

Input: body

☑ Remove source column

Outputs

Output column	JSONPath	List	Paths
I ids	$[*]['id']	☑	□
S titles	$[*]['title']	☑	□
D prices	$[*]['price']	☑	□
S categorys	$[*]['category']	☑	□

JSON-Cell Preview

```
[ {
  "id" : 1,
  "title" : "Fjallraven - Foldsack No. 1 Backpack, Fits 15 Laptops",
  "price" : 109.95,
  "description" : "Your perfect pack for everyday use and walks in the forest. Stash your laptop
  "category" : "men's clothing",
  "image" : "https://fakestoreapi.com/img/81fPKd-2AYL._AC_SL1500_.jpg",
  "rating" : {
    "rate" : 3.9,
    "count" : 120
  }
```

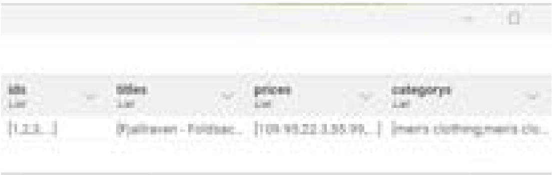
Add only the selected occurrence

Add all occurrences with the same key

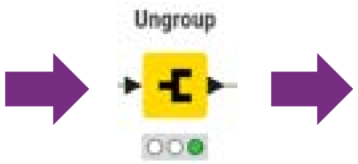
Select field from preview

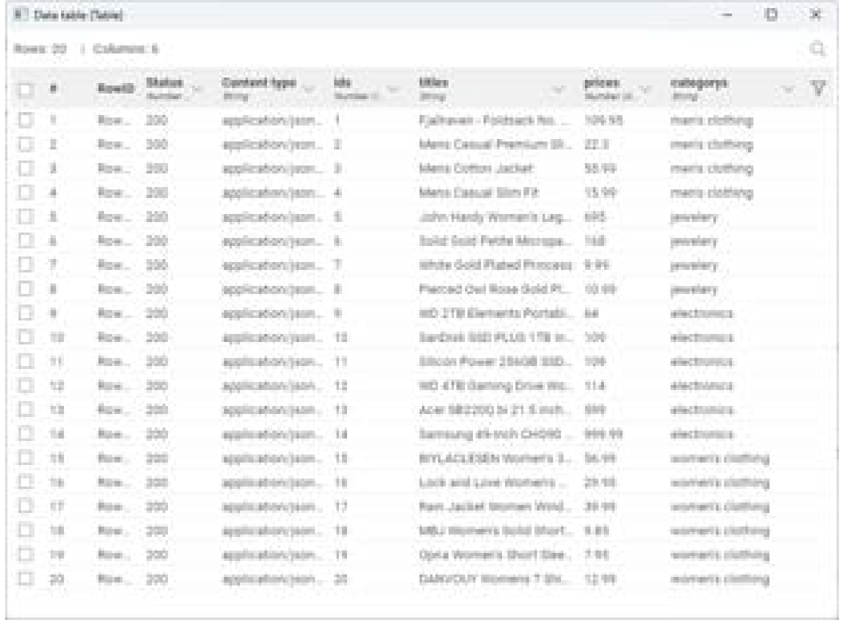
Ungroup

- Values for a JSON Path node are returned as collections
  - Multiple values in a cell
- Ungroup node converts collection cells to rows of data



ids List	titles List	prices List	categorys List
[1,2,3,...]	[Fjallraven - Foldsac...	[109.95,22.3,55.99,...]	[men's clothing,men's clo...





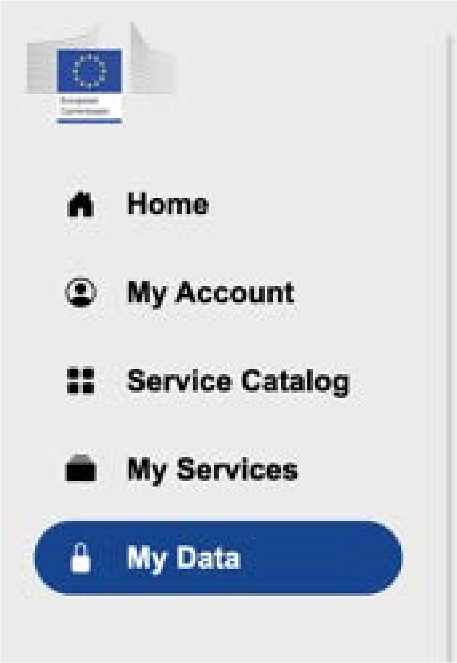
Data table (Table)

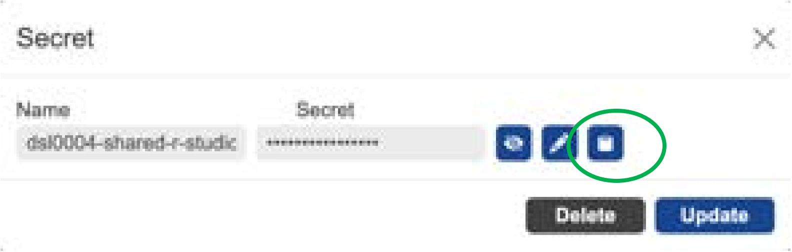
Rows: 20 | Columns: 6

□	#	RowID	Status Number ...	Content type String	ids Number (i...	titles String	prices Number (d...	categorys String
□	1	Row...	200	application/json...	1	Fjallraven - Foldsack No. ...	109.95	men's clothing
□	2	Row...	200	application/json...	2	Mens Casual Premium Sl...	22.3	men's clothing
□	3	Row...	200	application/json...	3	Mens Cotton Jacket	55.99	men's clothing
□	4	Row...	200	application/json...	4	Mens Casual Slim Fit	15.99	men's clothing
□	5	Row...	200	application/json...	5	John Hardy Women's Leg...	695	jewelery
□	6	Row...	200	application/json...	6	Solid Gold Petite Micropa...	168	jewelery
□	7	Row...	200	application/json...	7	White Gold Plated Princess	9.99	jewelery
□	8	Row...	200	application/json...	8	Pierced Owl Rose Gold Pl...	10.99	jewelery
□	9	Row...	200	application/json...	9	WD 2TB Elements Portabl...	64	electronics
□	10	Row...	200	application/json...	10	SanDisk SSD PLUS 1TB In...	109	electronics
□	11	Row...	200	application/json...	11	Silicon Power 256GB SSD...	109	electronics
□	12	Row...	200	application/json...	12	WD 4TB Gaming Drive Wo...	114	electronics
□	13	Row...	200	application/json...	13	Acer SB220Q bi 21.5 inch...	599	electronics
□	14	Row...	200	application/json...	14	Samsung 49-Inch CHG90 ...	999.99	electronics
□	15	Row...	200	application/json...	15	BIYLACLESEN Women's 3...	56.99	women's clothing
□	16	Row...	200	application/json...	16	Lock and Love Women's ...	29.95	women's clothing
□	17	Row...	200	application/json...	17	Rain Jacket Women Wind...	39.99	women's clothing
□	18	Row...	200	application/json...	18	MBJ Women's Solid Short...	9.85	women's clothing
□	19	Row...	200	application/json...	19	Opna Women's Short Slee...	7.95	women's clothing
□	20	Row...	200	application/json...	20	DANVOUY Womens T Shi...	12.99	women's clothing

KNIME Analytic Platform

- Sign in with your credentials stored in "My Data"



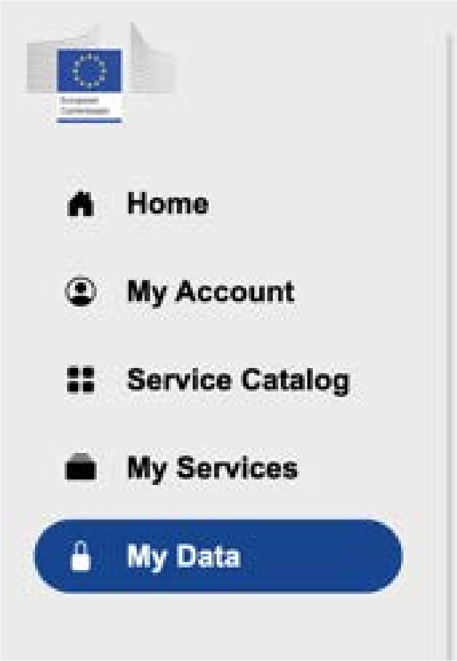

Secret

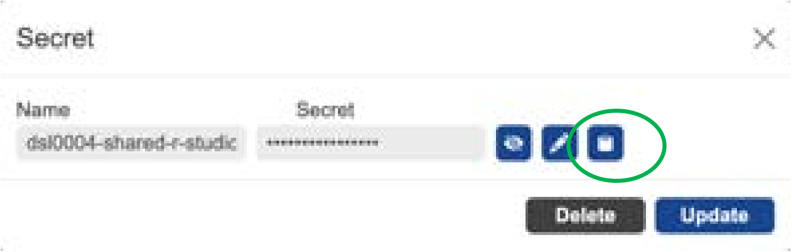
Name: dsl0004-shared-r-studic

Secret: ••••••••••••••••••

R-studio

- Sign in with your credentials stored in "My Data"




Secret

Name: dsl0004-shared-r-studic

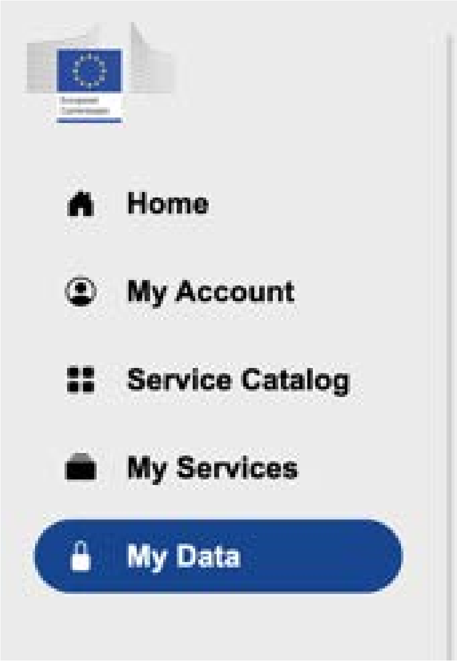
Secret: •••••••••••••••••

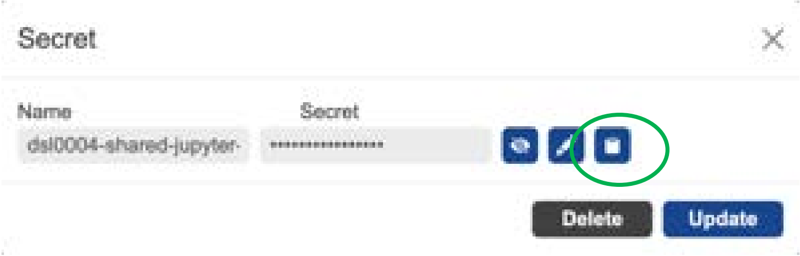
Use Case

Some colleagues from Zoi are **more advanced** and know how to code in R-studio, they want to know how to use this tool within the BDTI framework to integrate and collaborate with the rest of the team.

Jupyter lab - Python

- Sign in with your credentials stored in "My Data"





Secret

Name: dsl0004-shared-jupyter-

Secret: ******************

Use Case

Some colleagues from Zoi are **more advanced** and know how to code in Python, they want to know how to use this tool within the BDTI framework to integrate and collaborate with the rest of the team.

Summary

- **Now you should be able to:**

- Access BDTI framework

- Navigate BDTI Service Catalog

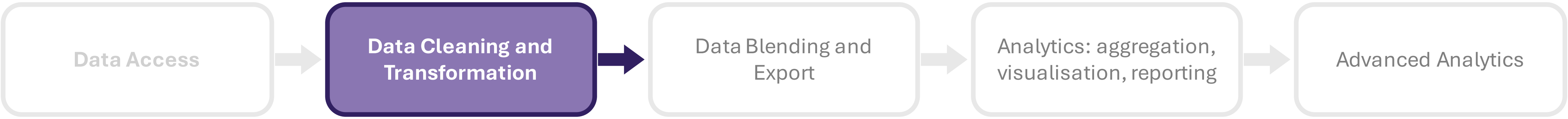
- Launch various data analytic tools
  - KNIME Analytics Platform, R-studio, Jupyter Notebook.

- Read different data file types into the analytic tools:
  - CSV, Excel, JSON.

- Complete basic data exploration in relation to the use case

Next session: 16/02/24

11:00 AM - 12:15 PM CEST



Data Access

Data Cleaning and Transformation

Data Blending and Export

Analytics: aggregation, visualisation, reporting

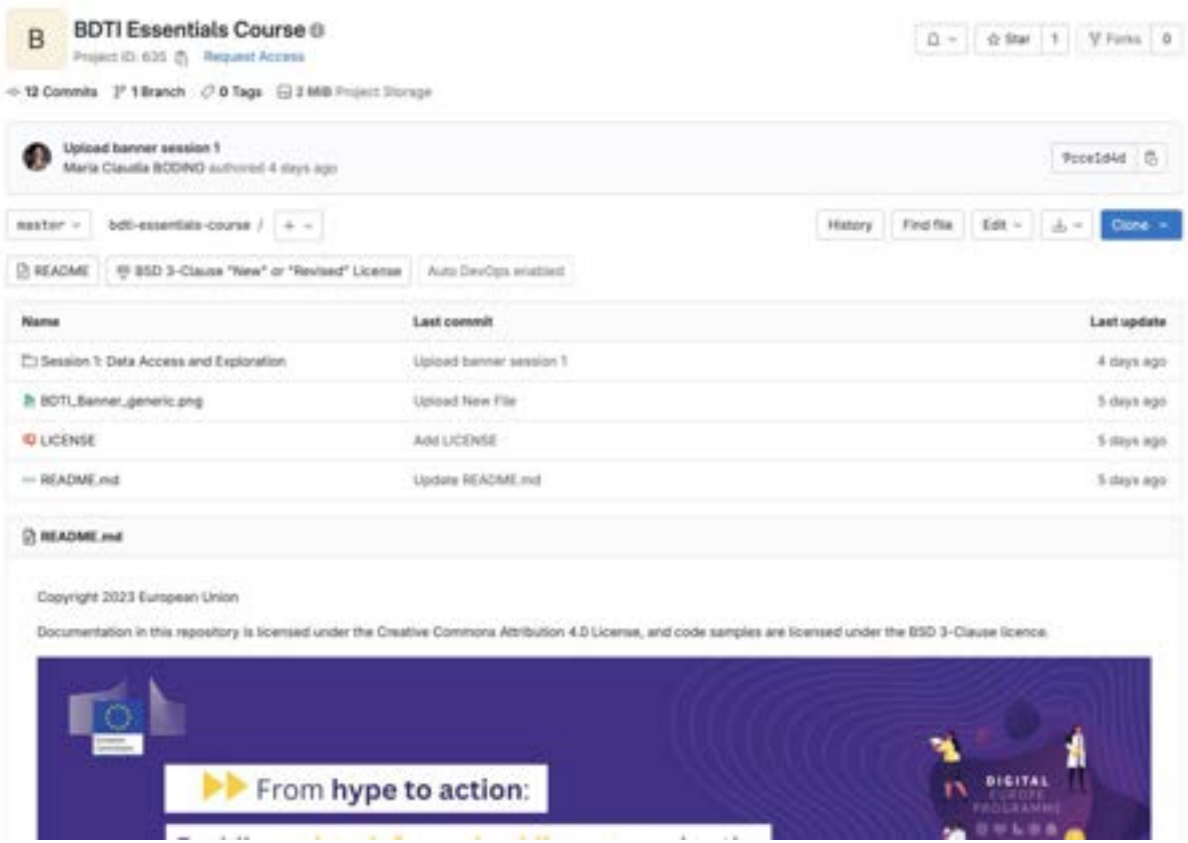
Advanced Analytics

https://big-data-test-infrastructure.ec.europa.eu/resources/courses-and-training/bdti-essentials-course_en

Practice, practice and practice!

- Slides, hands-on exercises, and solutions will be provided on the webinar Gitlab repository

https://code.europa.eu/bdti/bdti-essentials-course



BDTI Essentials Course

Project ID: 635

12 Commits · 1 Branch · 0 Tags · 2 MiB Project Storage

Upload banner session 1
Maria Claudia BODINO authored 4 days ago

9cce1d4d

master · bdti-essentials-course /

README · BSD 3-Clause "New" or "Revised" License · Auto DevOps enabled

Name	Last commit	Last update
Session 1: Data Access and Exploration	Upload banner session 1	4 days ago
BDTI_Banner_generic.png	Upload New File	5 days ago
LICENSE	Add LICENSE	5 days ago
README.md	Update README.md	5 days ago

README.md

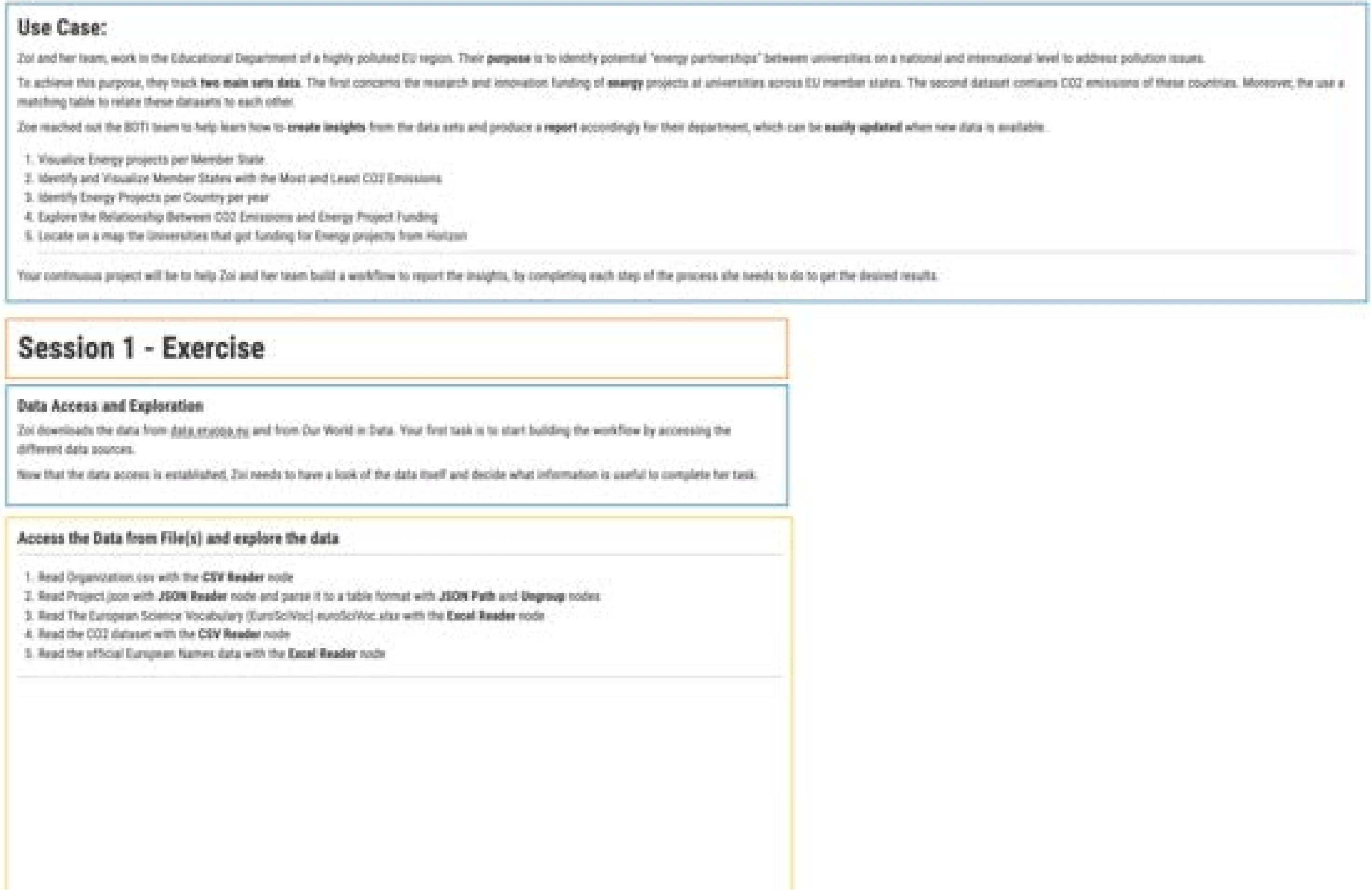
Copyright 2023 European Union

Documentation in this repository is licensed under the Creative Commons Attribution 4.0 License, and code samples are licensed under the BSD 3-Clause licence.

Practice, practice and practice!

- Slides, hands-on exercises, and solutions will be provided on the webinar Gitlab repository

https://code.europa.eu/bdti/bdti-essentials-course



Use Case:

Zoi and her team, work in the Educational Department of a highly polluted EU region. Their **purpose** is to identify potential "energy partnerships" between universities on a national and international level to address pollution issues.

To achieve this purpose, they track **two main sets data**. The first concerns the research and innovation funding of **energy** projects at universities across EU member states. The second dataset contains CO2 emissions of these countries. Moreover, the use a matching table to relate these datasets to each other.

Zoe reached out the BDTI team to help learn how to **create insights** from the data sets and produce a **report** accordingly for their department, which can be **easily updated** when new data is available.

1. Visualize Energy projects per Member State
2. Identify and Visualize Member States with the Most and Least CO2 Emissions
3. Identify Energy Projects per Country per year
4. Explore the Relationship Between CO2 Emissions and Energy Project Funding
5. Locate on a map the Universities that got funding for Energy projects from Horizon

Your continuous project will be to help Zoi and her team build a workflow to report the insights, by completing each step of the process she needs to do to get the desired results.

Session 1 - Exercise

Data Access and Exploration

Zoi downloads the data from data.europa.eu and from Our World in Data. Your first task is to start building the workflow by accessing the different data sources.

Now that the data access is established, Zoi needs to have a look of the data itself and decide what information is useful to complete her task.

Access the Data from File(s) and explore the data

1. Read Organization.csv with the **CSV Reader** node
2. Read Project.json with **JSON Reader** node and parse it to a table format with **JSON Path** and **Ungroup** nodes
3. Read The European Science Vocabulary (EuroSciVoc) euroSciVoc.xlsx with the **Excel Reader** node
4. Read the CO2 dataset with the **CSV Reader** node
5. Read the official European Names data with the **Excel Reader** node

Please check all the material in our open source repo:
https://code.europa.eu/bdti/bdti-essentials-course

Q&A

Your BDTI journey starts here

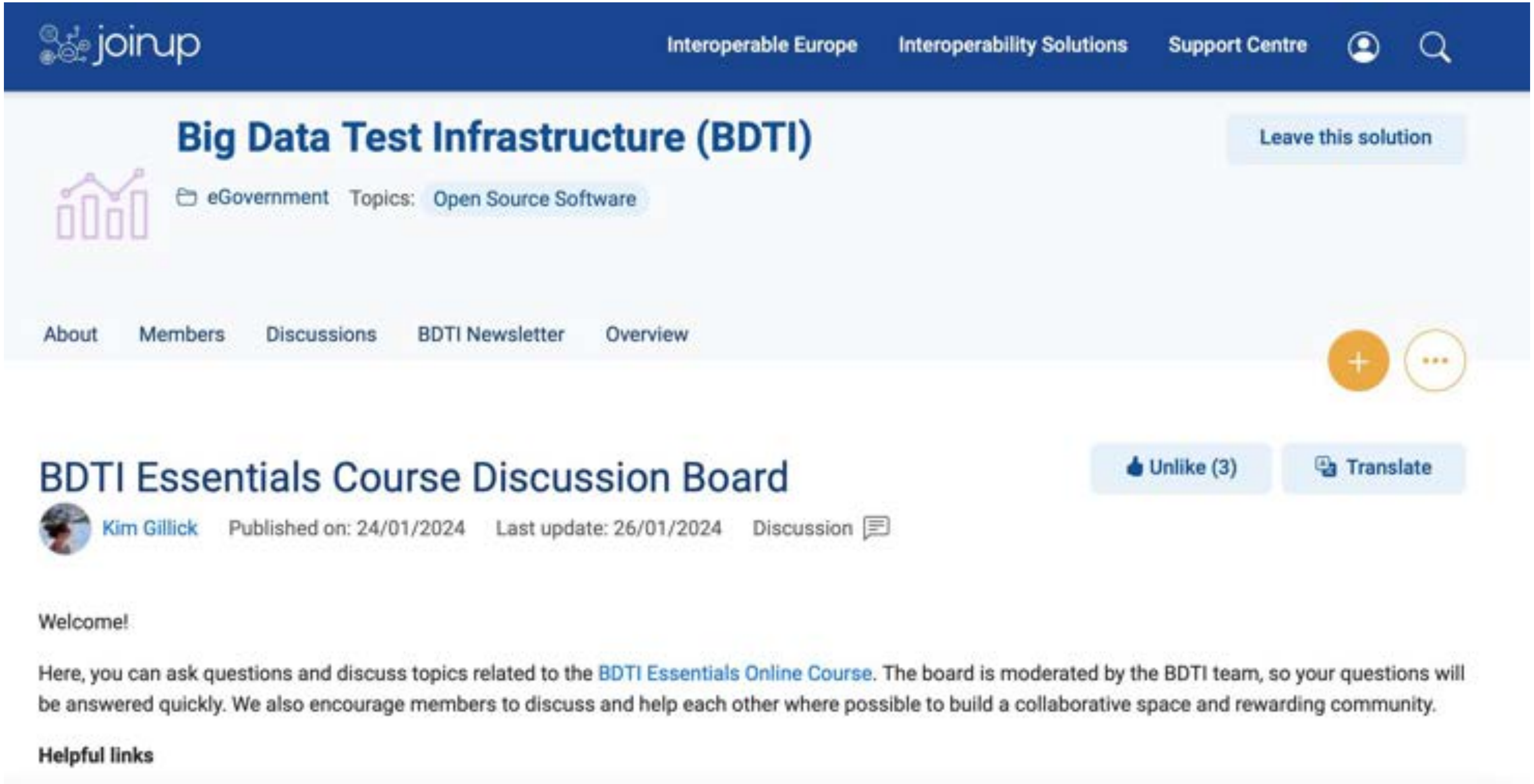
Congrats! You are on the first step to data-driven innovation.

- BDTI Essentials Course
- Brainstorm your data project
- Apply for BDTI Pilot-light process

Apply for BDTI: https://big-data-test-infrastructure.ec.europa.eu/apply-bdti_en

Course discussion board

DIGITAL EUROPE PROGRAMME



Big Data Test Infrastructure (BDTI)

eGovernment Topics: Open Source Software

BDTI Essentials Course Discussion Board

Kim Gillick Published on: 24/01/2024 Last update: 26/01/2024 Discussion

Welcome!

Here, you can ask questions and discuss topics related to the BDTI Essentials Online Course. The board is moderated by the BDTI team, so your questions will be answered quickly. We also encourage members to discuss and help each other where possible to build a collaborative space and rewarding community.

Helpful links

European Commission

Get in touch and follow the BDTI activities

Are you working for a public administration in need of infrastructure for data analytics?

EC-BDTI-PILOTS@ec.europa.eu

Visit BDTI's website

Subscribe to BDTI's newsletter

Subscribe to BDTI's Joinup

Register for BDTI Essentials online course

https://big-data-test-infrastructure.ec.europa.eu/

Thanks for attending, see you in 2 weeks!

Are you working for a public administration interested in testing a data analytics playground?

Schedule a 1-to-1 with the BDTI team: EC-BDTI-PILOTS@ec.europa.eu

Don’t miss the next session.
Register at the QR code below.

Questions? Comments? Ideas?
Join the dedicated discussion board.

https://big-data-test-infrastructure.ec.europa.eu/