






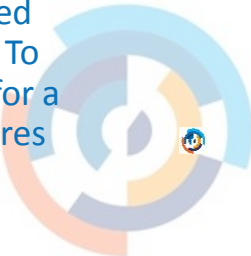


Important aspects to keep in mind

-  **Open Data is offered in many different formats enabling the data owner to choose how they prefer to retrieve the data.**
-  The format of an open dataset is the way in which the data is structured and made available for humans and machines.
-  Often enough the choice is limited and conditioned by the default formats used by the data provider. To maximise reuse of the data, it may be necessary for a publisher to use a number of formats and structures available across different platforms that suit the user's needs.
-  **Choosing the right format enables simpler management and reuse of the data.**
-  The formats can be grouped into five categories: Documents, Spreadsheets, Figures, Data to make a figure with and Compressed files.
-  Not all data formats are suitable for data analytics. There are data formats which are not adjustable. Others are easier to analyse, by machine, and so on.
-  All data formats need software to be read or adjusted. This varies from regular proprietary software, for example Microsoft Excel or Adobe Reader, to more advanced software to manipulate specific data, for example the statistics software R. Usually open source software can be found to read the various data formats.



Categorisation of the file formats of Open Data

Data format category	Example data formats	Extensions	Adjustable?
Documents	PDF, RDF, text document, Word document	PDF, RDF, txt, doc	PDF and RDF are not adjustable Text and word documents are adjustable
Spreadsheets	Microsoft Excel documents, XML spreadsheet, OpenDocument spreadsheet	csv, tsv, xlsx, xls, xml, ODS, JSON. Atom feed	All spreadsheet documents are adjustable
Figures	Pictures, online map viewer	jpeg, png, viz, invoke,	Figures are not adjustable
Data to make figures with	Publish data online (usually maps)	wms, wfs, wcs, shp	Usually the data is not adjustable
Compressed files	Compressed gzip file, compressed folder with multiple documents	GZ, zip	It depends on what is in the compressed file

The most common formats are csv, html and wms files.

Follow the eLearning module on Data Formats