
Data stewardship: Climate and energy resources

A Data Management Plan created using DMPonline

Creator: L I

Affiliation: Other

Template: European Commission (Horizon 2020)

ORCID iD: <https://orcid.org/0000-0001-7683-3402>

Project abstract:

This experiment is used to show the correlation between the yearly climate values and the used energy resources, in the years 1992-2016. The idea is that the climate has an impact on the amount of energy needed. The input data for this experiments is taken for the climate values from data.gv and for the energy resources from Statistik Austria. DOI: 10.5281/zenodo.2648558

Last modified: 22-04-2019

Data stewardship: Climate and energy resources - Final review DMP

1. Data summary

State the purpose of the data collection/generation

This experiment is used to show the correlation between the yearly climate values and the used energy resources, in the years 1992-2016. The idea is that the climate has an impact on the amount of energy needed.

Explain the relation to the objectives of the project

The data loaded and transformed is needed to determine the correlation as specified in the experiment goal.

Specify the types and formats of data generated/collected

The data format of the data generated is csv. This has been chosen since all data is tabular data and the input is also of type csv.

Specify if existing data is being re-used (if any)

Climate data (csv) from data.gv (Original 22 attributes, 4 used):

- Year (1992-2016)
- Ice days (numeric, # days)
- Summer days (numeric, # days)
- Mean temperature (numeric, degree celsius)

Energy resources data (csv) from Statistik Austria (4 attributes):

- Year (1992-2016)
- Renewables (numeric, Tera Joul)
- Oil (numeric, Tera Joul)
- Gas (numeric, Tera Joul)

Specify the origin of the data

The data is from data.gv (<https://www.data.gv.at/katalog/dataset/e6d471ec-f983-4e9a-b435-957cfc1bb48c>) and Statistik Austria (<http://statcube.at/statcube/opendatabase?id=deebilanz>).

State the expected size of the data (if known)

2kB

Outline the data utility: to whom will it be useful

Data will be provided via a DOI, stored on a public repository. The data might be interesting to researches looking to see the connection between climate and energy resource data.

DOI: [10.5281/zenodo.2648500](https://doi.org/10.5281/zenodo.2648500)

2.1 Making data findable, including provisions for metadata [FAIR data]

Outline the discoverability of data (metadata provision)

All datasets that were used and generated by the experiment (output.csv) are stored on a Zenodo repository, which is an open-access repository, that is based on a public Github repository.

This yields in a DOI citation (Digital Object Identifier).

Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?

The data is stored on a Github repository, making use of version control (git). Which is referenced by a DOI.

Outline naming conventions used

The input data is stored as provided. The output is always named as output.

Outline the approach towards search keyword

Metatags were used for keywords in the github repository.

Outline the approach for clear versioning

The Github versioning was used (Git based VCS).

Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how

Metadata describing the type and origin of data.

2.2 Making data openly accessible [FAIR data]

Specify which data will be made openly available? If some data is kept closed provide rationale for doing so

All datasets used and generated are made public available.

Specify how the data will be made available

The data is stored as csv in the public repository and can thus be downloaded and manipulated freely.

Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?

A browser to access the repository. To act on the source code, R has to be installed. To run the code a description is provided in the Readme.

Specify where the data and associated metadata, documentation and code are deposited

All the data, code and documentation is stored on a public Github repository.

Specify how access will be provided in case there are any restrictions

There are no restrictions.

2.3 Making data interoperable [FAIR data]

Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate

interoperability.

Since the standard file format 'csv' has been used, the data can be used by a variety of tools (editors, scripts, and so on).

Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?

A description of the data used and generated is provided in the repository.

2.4 Increase data re-use (through clarifying licenses) [FAIR data]

Specify how the data will be licenced to permit the widest reuse possible

MIT License

Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed

The data is already made available.

Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why

It can be freely reused.

Describe data quality assurance processes

The original data was not modified, just transformed and reduced to the minimum needed.

Specify the length of time for which the data will remain re-usable

There is no time limit planned.

3. Allocation of resources

Estimate the costs for making your data FAIR. Describe how you intend to cover these costs

There are no costs associated with a public github repository at the time of writing.

Clearly identify responsibilities for data management in your project

All data was gathered, stored and shared by one person, the dmp holder.

Describe costs and potential value of long term preservation

There are no costs associated.

4. Data security

Address data recovery as well as secure storage and transfer of sensitive data

The data is both stored at Zenodo and Github which handle the data security.

5. Ethical aspects

To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former

The input and output data of this project are completely anonymized. There is no unethical act.

6. Other

Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)

There are no other procedures.