Plan Overview

A Data Management Plan created using DMPonline

Title: Comparison of divorce rate between Austria and the EU

Creator: Martin Pichler

Principal Investigator: Martin Pichler

Data Manager: Martin Pichler

Affiliation: Other

Template: DCC Template

ORCID iD: 0000-0001-5305-9063

Project abstract:

Comparison on how the number of divorces between 1990 and 2015 correlate between Austria and the EU.

ID: 39550

Last modified: 19-04-2019

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Comparison of divorce rate between Austria and the EU

Data Collection

What data will you collect or create?

Input Data
A .csv and a .tsv file are saved in the project repository in the folder *data/raw*.

1. Ehescheidungen (Statistik Austria)
   Source: data.gv.at (Accessed on April 19 2019)
   UUID: 2d8ad82c-4730-3354-9971-9406f2ccf72c
   File Location: data/raw/OGD_scheidung_ext_SCH_1.csv
   File Size: 1KB
2. Divorces by duration of marriage (Eurostat)
   Source: data.europa.eu (Accessed on April 19 2019)
   Identifier: demo_ndivdur
3. File Location: data/raw/demo_ndivdur.tsv
   File Size: 297KB

Produced Data
The project produces 3 .png images that are stored in the *results* folder. The images are used for visual analysis of the correlation between the divorce rate of Austria and the EU.

1. scatter.png (~15KB)
2. time_corr.png (~15KB)
3. time_change.png (~30KB)

How will the data be collected or created?

Input Data
The input data files are downloaded and stored alongside the source code in the repository in the folder *data/raw*. Filenames are kept as provided by the data sources.

Produced Data
The experiment is conducted using a single jupyter notebook *divorce_analysis.ipynb*. Running the jupyter notebook will create three image files in the *results* folder. Data manipulation of the input data is only done in memory and no input data is altered during the execution of the jupyter notebook.

A readme.md is included in the repository describing how the experiment can be run either natively or with docker.

Documentation and Metadata

What documentation and metadata will accompany the data?

A description of both input datasets can be found in *divorce_analysis.ipynb.* No additional metadata is stored.

Ethics and Legal Compliance

How will you manage any ethical issues?

The data used for this experiment does not have any personalized data and therefore no ethical issues need to be managed.

How will you manage copyright and Intellectual Property Rights (IPR) issues?

Both datasets are openly accessible as *Open Data* under either the Creative Commons Attribution License 3.0 licence or have no licence applied.

Storage and Backup

How will the data be stored and backed up during the research?

The project and data are managed as a git repository and as such are stored as a public repository on Github (https://github.com/martinpichler/data_stewardship_ex1). This allows also to track changes in the dataset and can be used as a long term storage solution.

How will you manage access and security?

The project is publicly accessible on Github. Changes to the repository have to be accepted by the owner or an administrator. Since no confidential information or data is stored no additional security measures have to be set.

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

While both input data files are accessible by either the Austrian or European open data portals, the original input data files are stored long term within the git repository.

What is the long-term preservation plan for the dataset?

The project will be stored long-term on Github (https://github.com/martinpichler/data_stewardship_ex1). Since Github is a free service, storing the data long-term does not bear any additional costs.

Data Sharing

How will you share the data?

The data and code will be publicly available on Github (https://github.com/martinpichler/data_stewardship_ex1) under the MIT licence.

Are any restrictions on data sharing required?

There are no restrictions necessary.

Responsibilities and Resources

Who will be responsible for data management?

Project creator Martin Pichler is responsible for creating and maintaining the DMP. He will also be responsible for managing merge requests for the repository.

What resources will you require to deliver your plan?

No additional resources needed.