The relationship between family income in the United States and child mortality - Detailed DMP

1. Data summary

State the purpose of the data collection/generation

The main purpose of the collection of the two data sets is to find the link and the relationship between them and build a result to take advantage of life if it is useful, we are working to improve it for the best and if it is not useful we work against it to get the best

Explain the relation to the objectives of the project

The resulting relationship is one of the most important objectives of the project. The relationship is to increase the family income in the United States with a decrease in the death rate of children aged 1 to 19 years from 1953 to 2015. In conclusion, this relationship can be confirmed in the future if household income increases. It can be applied to other countries

Specify the types and formats of data generated/collected

Source code :java.
Tables :CSV.
Text documents :PDF.
Structured Text :XML
Images :PNG.
The amount of data is 10 KB for the first data set It is in csv format .
The data quantity is 2.51 KB for the second data set It is in csv format .

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

Data will be available after the completion of the project directly for reasons of the author and will be usable by a third party but after the completion of the project and will be available for reuse for five years after completion .

Specify the origin of the data

I collected my raw data from two different sites, ," kaggle" and "data.gov", within the criteria and conditions for example licensing and being somewhat within one geographical area and within a similar time series between the two data sets.

State the expected size of the data (if known)

The amount of data is 2 KB for new data set It is in csv format .
The data quantity is 13 KB for the metadata
An image that shows the relationship between the data set and its size 144 KB

Outline the data utility: to whom will it be useful

Data is available at https://zenodo.org and the data ID is 10.1997 / rami.bar In csv format.
There will be no data blocking dates after completion of the project.
The code will be uploaded to my account on githup and open source

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

Outline the discoverability of data (metadata provision)

Metadata was created using an xml file that contains metadata metadata title, descriptor name, subject, date, simplified description, type, primary data links and other things. By entering data links, you have their own metadata and I have the right to use raw data for primary data by virtue of the licenses granted to the same data and were written under dcmi standards

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?

Yes Take advantage of Digital Object Identifiers The resulting data has been placed in https://zenodo.org and the data ID is 10.1997 / rami.bar .

Outline naming conventions used

zenodo.org

Outline the approach towards search keyword

They were given key words when they remembered going to the data directly from them including US family income and the death of children.

Outline the approach for clear versioning

Is to achieve the release of most of the objectives of the project during this release, including finding a relationship between the two data set and find the link between them and work on it

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

metadata contains metadata metadata title, descriptor name, subject, date, simplified description, type, primary data links and other things and were written under dcmi standards.

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

No data will be hidden from the characters and will be available directly with the finished project

Specify how the data will be made available

Data set and metadata will be available at zenodo.

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)?

The code will be open to my github account linked to the zenodo site. However, those who want the code will need to have a script editor to run the code in Java. To display the metadata you need an editor running xml to display and you need the Office program to view the documents.

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

Data and associated metadata are uploaded to zenodo.
And documents on my Google Drive account.
The code on my account is github.

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

By referring to the data administrator by sending an email or calling him because the contact information is located where the data is located

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.

We can apply the interoperability of the new dataset easily by a third user and software after the completion of the project because the methodology used in these data is known on a large scale and because it has the same format and has descriptive data in which all the details are written in standards and used DCMI standard

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?

All data used and stored in the project are either plain text, regular numbers or a diagram, easily understood and interacted with by users.

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

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

Data Licenses are a Creative Commons Attribution-ShareAlike License (CC-BY-SA) via this link
http://ufal.github.io/public-license-selector/ .

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

will be available after the completion of the project directly for reasons of the author and will be usable by a third party but after the completion of the project and will be available for reuse for five years after completion until the data is updated and rebuild the relationship again

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

The resulting data will be available after the completion of the project from the use and can not be reused by third parties again

Describe data quality assurance processes

The data was collected from the Internet from two different sites documented to them. Data on child mortality were the percentage of 100,000 children living in the United States. Income data was a year for all members of the community
Some operations were performed on child mortality data, and children were divided into several categories. The death toll was collected in these categories and collected in one category. A large category of years was excluded to fit the second data. Hashmap was created to store the years of data generated as a key and the number of child deaths as a value. Several columns of raw data that were not required in income and household income data were excluded from income from some years to fit years in child mortality data. HASHMAP was created and it took years as a key and annual income value as a value. A HASHMAP method was created for years from small to large and stored in TREEMAP. Exceptional columns were excluded from raw data and a graph line was created to represent the relationship between infant mortality and annual family income. With high household income, the infant mortality rate has declined. Child mortality can be derived from extracted data to build other relationships. The resulting data is stored in a .csv file and the raw data is attached to a special file within the source code folders. The Java language was used to represent these processes using the netbeans program and the code was pushed to Github.

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

The data will be available after completion of the project for five years

3. Allocation of resources

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

There is no economic cost because I made it on the zenodo site, which is free but it took me time and storage space.

Clearly identify responsibilities for data management in your project

The data will be managed during the project. The files will be stored periodically on the flash and will be uploaded to my account on githup. At the time of the emergency, the data will be retrieved and will be stored on my computer.

Describe costs and potential value of long term preservation

I do not think long-term preservation will cost

4. Data security

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

The data I work on is a little important data in the area of confidentiality, and if the delay in obtaining any pirated information is general information anyone can come from the Internet and the data will be controlled by the project manager. The project administrator's user name and password will be able to access and access the data, and it will be easy, and everyone will have the right to access the data without modification unless it has the power, with the help of zenodo

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

In my data, there is no moral or legal problem, there is no restriction on personal identity, and anyone interested in the data can look forward to it.

6. Other

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

No, there is not