RELATION BETWEEN NUMBER OF UNEMPLOYMENT PEOPLE IN NEW YORK AND NUMBER OF DNA STRUCTURES RELEASED BY PROTEIN DATA BANK

A Data Management Plan created using DMPonline

Creator: IBAIDA TAHA

Affiliation: Other

Template: Digital Curation Centre

ORCID iD: https://orcid.org/0000-0001-6592-7016

Github: https://github.com/ibaidaTaha/DMP-of-Relation-between-Number-of-Unemployment-people-in-New-York-and-Number-of-DNA-Structures-

Doi: https://doi.org/10.6084/m9.figshare.6349478.v1

Project abstract:
This report presents the relation between Number of Unemployment people in New York and Number of DNA Structures Released by protein data bank from 2000 to 2018, It contains a data sets from two different open data repository.

Last modified: 01-06-2018

RELATION BETWEEN NUMBER OF UNEMPLOYMENT PEOPLE IN NEW YORK AND NUMBER OF DNA STRUCTURES RELEASED BY PROTEIN DATA BANK

DATA COLLECTION

What data will you collect or create?

Data that we used in our experiment was experimental observation and samples from the research done by protein data bank and US government's open data repository, This is the first file https://www.rcsb.org/stats/growth/dna, and has volume (1 KB) and the second is https://catalog.data.gov/dataset/local-area-unemployment-statistics-beginning-1976, and has a volume (3,5 KB), The excel spreadsheet will save as a comma separated value (.csv) file.

How will the data be collected or created?

We took a data sets from this two different open data repository:
- Unemployment data from the U.S. Government's repository (https://www.data.gov)
- DNA structures from Protein Data Bank (https://www.rcsb.org)

DOCUMENTATION AND METADATA

What documentation and metadata will accompany the data?

Data.gov follows the Project Open Data schema – a set of required fields (Title, Description, Tags, Last Update, Publisher, Contact Name, etc.) for every data set displayed on Data.gov.
All data has been worked in comma separated value (.csv) format with accompanying meta\data file in text (.json) format, and you can find this files on Github https://github.com/ibaidaTaha/DMP-of-Relation-between-Number-of-Unemployment-people-in-New-York-and-Number-of-DNA-Structures- .
Documentation, tutorials, and metadata for using both data and software will be archived with the relevant product, and the metadata schema used is the one used by the repository.

ETHICS AND LEGAL COMPLIANCE

How will you manage any ethical issues?

No ethical issues (no human subject data collected).
How will you manage copyright and Intellectual Property Rights (IPR) issues?

First, about copyright, the data we used from protein data bank and U.S. Government's repository is public and accessible for everyone.
Second, the data is owned by data bank and U.S. Government's repository and it's used for further non-commercial or commercial dissemination.
We used Excel tool to represent the relation between these two datasets.

STORAGE AND BACKUP

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

Digital copies will be stored on Figshare (https://figshare.com).

How will you manage access and security?

All data and metadata will be stored privately in the cloud on Figshare until publication, after which point it will be made open-access under a Creative Commons license, and citable in its own right. Data will be searchable on Figshare, and downloadable by any user.

SELECTION AND PRESERVATION

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

The final dataset will be transferred to the Figshare repository, which ensures sustainable archiving of the final research data.

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

For long-term archiving, there are a couple of options in place, but in this experiment, we used Fedora commons 4 repository to keep the data.

DATA SHARING

How will you share the data?

Potential data users for the data we will compile include people who interested in statistical and statistics and include anyone working with long-term data.

We used Fedora commons 4 to publish data, that's why we used CC-BY as a license.

Are any restrictions on data sharing required?

No restrictions on data necessary, or ethical or privacy issues. Data will be free to use under the expectation that it will be correctly attributed and cited using the Figshare DOI.

RESPONSIBILITIES AND RESOURCES

Who will be responsible for data management?

The team members will be primarily responsible for managing all information as relates to this project according to the FAIR principles.

What resources will you require to deliver your plan?

No resources needed.