Cryptocurrency Litecoin (LTC): Risk and Dependence Assessment - Detailed DMP

1. Data summary

State the purpose of the data collection/generation

Produced Data:
This Project produces aggregated dataset in CSV format (Filesize ~450 KB) that contains excess returns of assets and the riskfree rate, and a histogram of these in PNG format (Filesize ~100KB).

Explain the relation to the objectives of the project

The goal of this project is to identify the risk of LTC. Is there a systematic risk? Is it changing over time? Or are newer Methods like Conditional Value at Risk more appropriate for measuring risk?

Specify the types and formats of data generated/collected

Title: Assets
Description: Assets from YahooFinance (2019)
Type: Dataset
Format: csv
Source: Input
Title: LTC
Description: ltc (2019)
Type: Dataset
Format: csv
Source: Input
Title: RF
Description: dayli riskfree rate (FamaFrench) (2019)
Type: Dataset
Format: csv
Source: Input
Title: excess_return
Description: excess returns of all assets
Type: Dataset
Format: csv
Source: output
Title: Histogram
Description: Histogram of all excess returns
Type: Image
Format: png
Source: output

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

Assets (Yahoo Finance), LTC (coinmetrics) , RF (Fama French) are re-used for this project.

Specify the origin of the data

Assets was downloaded from https://finance.yahoo.com/ .
LTC was downloaded from https://coinmetrics.io/ .
RF was downloaded from https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html .
excess_return was produced during research.
Histogram was produced during research.

State the expected size of the data (if known)

Assets has a size of 265KB.
LTC has a size of 432KB.
RF has a size of 40KB.
excess_return has a size of 450KB.
Histogram has a size of 100KB.

Outline the data utility: to whom will it be useful

Other Researchers, who want to estimate the risk for LTC.

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

Outline the discoverability of data (metadata provision)

The metadata is provided in "metadata.xml".

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?

At the end of the project a snapshot of the remaining data will be exported and stored in the
research data repository Zenodo obtaining a DOI. (10.5283/zenodo.2358372)

Outline naming conventions used

The input files are saved with their original file names.
Other files, code and data use naming conventions commonly used with python. (https://visualgit.readthedocs.io/en/latest/pages/naming_conventions.html)

Outline the approach towards search keyword

Not implemented.

Outline the approach for clear versioning

The project team will systematically use Git as the source code management system for software development. Git will allow the collaborate development of the software providing effective distributed version control. It is easy to use and yet powerful and efficient to handle even large projects.
A copy of each Git repository will be uploaded ("pushed") to a publicly available Git server such as GitHub or Bitbucket. They are both well-established online servers which support distributed source code development, management, and revision control. They enable worldwide collaboration between developers and also provide some additional facilities to work on documentation and to track issues. GitHub provides paid and free service plans. Free service plans can have any number of public, open-access repositories with unlimited collaborators. Private, non-public repositories require a paid service plan while Bitbucket allows private repositories to be shared with up to five collaborators.

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

The metadata is provided in the folder documentation/metadata.xml. For example, author, format, sources are created.

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

Created using DMPonline. Last modified 01 July 2019

2 of 5

Assets is open access.
LTC is open access.
RF is open access.
excess_return is open access.
Histogram is open access.

Specify how the data will be made available

The Data is available on the open Git repository (https://github.com/Sn3llius/Cryptocurrency-Litecoin-LTC-Risk-and-Dependence-Assessment).
Additionally, at the end of the project a copy/snapshot of all relevant data as well as other
data categories (publications, presentations, etc.) will be made available to open data
repositories such as 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 data will be made available online through the WWW service. Thus, in order to discover,
or download the data only a common browser will be required from the user. The format
of downloadable files will be (solely or additionally) in an open standard such as CSV, XLSX, R and PDF.
Tools to access these files will be freely available in the internet.

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

The Data is available on the open Git repository (https://github.com/Sn3llius/Cryptocurrency-Litecoin-LTC-Risk-and-Dependence-Assessment).
Additionally, at the end of the project a copy/snapshot of all relevant data as well as other
data categories (publications, presentations, etc.) will be made available to open data
repositories such as Zenodo.

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

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.

Datasets produced by the project will be interoperable allowing data exchange and reuse. The project will be well documented and unless otherwise specified
openly accessible. The project will follow established European metadata vocabularies, standards and methodologies.

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?

Standard vocabulary will be used, no mapping required.

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

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

The project is licenced under the MIT licence (https://opensource.org/licences/MIT) and open for re-use by anyone.

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

Already open 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

All data produced and used in this project are available by third parties or will be made available after the end of the project.

Describe data quality assurance processes

Some aspects that are considered, are the following:

- missing values
- wrong data type
- useing right data (e.g. maturity of Treasury Bills)
- incomplete or erroneous results

Such validation checks, are allready done. So the results of this project are trustworthy.

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

All open data will remain re-usable for at least one years after the end of the project and maintained by the project teammembers.

3. Allocation of resources

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

Since the data is stored in a public Github repository, no additional cost has to be coverd.

Clearly identify responsibilities for data management in your project

Christoph Wimmer
christoph.wimmer@s.wu.ac.at
https://orcid.org/0000-0001-5430-1543

Describe costs and potential value of long term preservation

A snapshot of all relevant project data will be stored in the research data repository Zenodo upon termination of the project. This dataset will receive a DOI and will be preserved for the foreseeable future. There will be no huge potential value of long term preservation.

4. Data security

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

All research data underpinning publications will be made available for verification and re-use unless there are justified reasons for keeping specific datasets confidential. The main elements when considering confidentiality of datasets are:
Protection of intellectual property regarding new processes, products and technologies where the data could be used to derive sensitive information
Commercial agreements as part of the procurements of components or materials that might foresee the confidentiality of data
Personal data that might have been collected in the project where sharing them is not allowed by the national and European legislation.
At the end of the project, a copy/snapshot of all significant data will be stored in the project's repository at Zenodo. The infrastructure of Zenodo has been developed and is supported by CERN which guarantees long term data safety and availability.

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

No ethical issues exist.

6. Other

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

No additional procedures are used.

Created using DMPonline. Last modified 01 July 2019

5 of 5