Data Management Plan

COVID-19 Gold Price Analysis

Contact person
- Ludwig Burtscher (e1526268@student.tuwien.ac.at, https://orcid.org/0000-0001-5453-8074)

Based on Common DSW Knowledge Model, 2.0.1 (dsw:root:2.0.1)

Generated on 06 Apr 2020

Data Management Plan created in Data Stewardship Wizard <https://ds-wizard.org>

Abstract

The current global Corona crisis not only has a huge impact on health systems around the world, but also on the economic situation. Gold is traditionally known to be a go-to asset in times of crisis and uncertainty and its price therefore usually increases when stock indices collapse. However, the Austrian newspaper Kurier recently published an article (https://kurier.at/wirtschaft/coronavirus-verliert-gold-seinen-ruf-als-krisenwaehrung/400787762) claiming that current market activity during the Corona crisis contradicts this theory. They also state that the gold price will stabilize again in the medium term as this also happened on other occasions.

This experiment aims at a data-driven approach to make validations of these assumptions easier. The development of the number of Corona cases and deaths in each country over time is compared to the corresponding gold prices. Since we have not yet overcome the current situation, validating assumptions about medium-term consequences on the gold price cannot be made now, but might be with future data.

Section A: Data Collection

1. What data will you collect or create?

Re-used datasets We will use the following already existing non-reference datasets:

- **COVID-19 cases and deaths**

  We already have a copy of this dataset.

- **Gold prices first quarter 2020**

  We already have a copy of this dataset.

Data formats and types We will be using the following data formats and types:

- **CSV**

  It is a standardized format. This is a suitable format for long-term archiving. We will have only a small amount of data stored in this format.

2. How will the data be collected or created?

There will be no instrument dataset in this project.

Storage and file conventions We will use a filesystem with files and folders with the following folder conventions:

- There will be a **folder for each sample/subject**. Each of those will use the following conventions: The name of a folder is equal to the name of the dataset in it.

Moreover, we have made appointments about naming the files.

We will not be storing data in an "object store" system.

We will not use a relational database system to store project data.

We will not use a graph database for data in the project.

We will not be storing data in a triple store.

Section B: Documentation and Meta-data

3. What documentation and meta-data will accompany the data?

List of data to be published is given in Section E, Question 9. This also includes information about catalogs where the data can be found. Information about data types used is given in Section A, Question 1.

We will be documenting the data with Dublin Core metadata.

Section C: Ethics and Legal Compliance

4. How will you manage any ethical issues?

Data we reuse

- **COVID-19 cases and deaths** does not need an extension of any consent because it is not personal data.
- **Gold prices first quarter 2020** does not need an extension of any consent because it is not personal data.

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

We will be working with the philosophy *as open as possible* for our data.

The data cannot become completely open immediately because the deadline for the Data Stewardship exercise has to be reached to prevent misuse.

Data that is not legally restrained will be released after a fixed time period, unconditionally.

For the reference and non-reference data sets that we reuse, conditions are as follows:

- **COVID-19 cases and deaths** – freely available with obligation to quote the source (e.g. CC-BY).
- **Gold prices first quarter 2020** – freely available for any use (public domain or CC0).

Section D: Storage and Backup

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

Data that project members themselves store are adequately backed up and traceable. Therefore data are protected against both equipment failure and human error. This is achieved by using two separate NAS-Systems at different locations.

7. How will you manage access and security?

Project members will not store data or software on computers in the lab or external hard drives connected to those computers.They can carry data with them on encrypted data carriers and password-protected laptops. Project members have been instructed about both generic and specific risks to the project.

The possible impact to the project or organization if information is lost is small because all data can be regenerated. The possible impact to the project or organization if information is leaked is small because it is publicly available on purpose. The risk of information vandalism in the project or organization is acceptably low.

We are not using any personal information.

Section E: Selection and Preservation

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

We plan to publish the following datasets:

- **Correlation of COVID-19 cases and the gold price** – This data set will be kept available as long as technically possible. – The metadata will be available even when the data no longer exists.

9. What is the longterm preservation plan for the dataset?

- **Correlation of COVID-19 cases and the gold price** will be stored in a domain-specific repository: Zenodo. The DOI of the dataset is 10.5281/zenodo.3741831.

None of the used repositories charge for their services.

Section F: Data Sharing

10. How will you share the data?

- **Correlation of COVID-19 cases and the gold price** – freely available for any use (public domain or CC0).

Information about used repositories (i.e. where will potential users find out about the data) is provided in Section E, Question 9.

Embargo on the data is described in Section C, Question 5, and Section F, Question 11.

11. Are any restrictions on data sharing required?

Ethical and legal restrictions are documented under Section C. We have used the Data Stewardship Wizard, which made us aware of options to minimize the restrictions.

Data that is not legally restrained will be released after a fixed time period. The embargo period will end on 20.04.2020.

Section G: Responsibilities and Resources

12. Who will be responsible for data management?

Ludwig Burtscher is responsible for implementing the DMP, and ensuring it is reviewed and revised.

13. What resources will you require to deliver your plan?

To execute the DMP, no additional specialist expertise is required.

We do not require any hardware or software in addition to what is usually available in the institute.

Charges applied by data repositories (if any) are mentioned already in Section E, Question 9.