

Exercise for Functional Task 1:

Describing a Data Set at the Variable Level

Overview:

For this exercise, students should form groups of 2-3 people, at least one of whom is comfortable working with XML either in a text editor or an XML editor. Using a pre-prepared template, students will be asked to describe a simple ASCII file using DDI 3.2. This will involve describing a Universe, Codes, Categories, Variables, and the logical and physical content of the data file.

Scenario:

You have a data set containing information about a set of corporate financial accounts, and you want to describe it at the variable level using DDI 3.2. A set of standard statistical/theoretical concepts is used for this type of data, and these have already been documented. The data is reported on a daily basis, and concerns corporations whose business relates to the environment (agriculture, forestry, and fisheries). The study is being conducted in Luxembourg, and all corporations and banks are domestic.

These concepts are:

Account ID – The identifier of the account used by the corporation’s financial institution.

Financial Institution – The legal name of the corporation’s financial institution which holds the account.

Corporation Name – The legal name of the corporation holding the account.

Business Sector – The business sector of the corporation, according to a simple classification (the only values are “Forestry”, “Fisheries”, and “Agriculture”).

Observation Time – The date of the day on which the financial standing of the corporation was recorded, at the official close of business.

Financial Holdings - The value of the corporate account, expressed as a total measured in Euros.

Daily Change – The relative change in financial holdings from the previous business day, measured at the end of business and expressed in Euros.

For convenience, a screen-shot of the data file is shown below.

	A	B	C	D	E	F	G	H
1	Case ID	Account Number	Bank Name	Corporation	Sector	Time	Account Balance	Change
2	1	A1234	Luxembourg Capitol Bank AG	Luxemburg Lumber Supplies AG	FOR	2013-09-09	1334.698	245.22
3	2	A4789	Prime Investments AG	Whole Foods Luxembourg AG	AGR	2013-09-09	256789.45	6523.78
4	3	D5442	Mosel Finacial AG	Seafood National AG	FSH	2013-09-09	6237.23	-345.21
5	4	J7982	Kirchberg Banking AG	Aquatic Foods AG	FSH	2013-09-09	98704.34	8976.34
6								
7								

Notes on the XML Template

You will need to add an “id” for all elements with a `<r:URN></r:URN>` in the template. The URN content in the template has been filled in EXCEPT for the ID of the object. This is what you will see in the template:

```
<r:URN>urn:ddi:de.dagstuhl:[ID]:1</r:URN>
```

All you replace is the [ID] with the “id” of the element. For example, where id equals VAR_1

```
<r:URN>urn:ddi:de.dagstuhl:VAR_1:1</r:URN>
```

When making references in DDI 3.2, you must insert the value of the “id” of the element being referenced into the `<r:URN>urn:ddi:de.dagstuhl:[ID]:1</r:URN>` of the referring element. For example you wish to reference a Concept with the id of CON_A:

```
<r:URN>urn:ddi:de.dagstuhl:CON_A:1</r:URN>
```

A reference also identifies the TypeOfObject which for this exercise will be already completed in the template.

Step 1:

Open the DDI XML template in a text editor or XML editor (Exercise1DDITemplate.xml)

Step 2:

Open the data file in Excel or a text editor (Exercise1Data.csv)

Step 3:

Identify the Universe (that is, the population being studied) and enter it into the file. Complete all needed fields in the DDI XML document.

Step 4:

Describe the logical record in the DDI XML file.

Step 5:

Examine the Variables to determine the representation of each one. If necessary, create Categories and Codes in the DDI XML document. Record any other representations on a piece of paper, for use in describing the Variables themselves.

Step 6:

Examine the Variables and describe them in the DDI XML file. Each Variable should have a name, label, description, a reference to the relevant concept, and a representation. The different types of Variables are provided in the template (Text, Case ID, Date-Time, Numeric, Coded) – these may need to be cut-and-pasted if there are more than one of a single type of Variable.

Step 7:

Look at the set of Variables and the data file and create a Physical Structure and a Record Layout in the DDI XML file.

Step 8:

Create a Physical Instance in the DDI XML file.