



FUSION REGISTRY COMMUNITY EDITION DATA FORMATS

FUSION REGISTRY COMMUNITY EDITION
VERSION 9

Data Formats

This guide documents the supported Data Formats for import and export in the Fusion Registry

1 Contents

2	Overview	2
3	SDMX Data Formats	3
3.1	Generic Data	3
3.2	Structure Specific / Compact Data	3
3.3	SDMX-EDI	4

2 Overview

This document provides an overview of all data formats supported as import by the Fusion Registry

For SDMX data formats, additional documentation is available from sdmx.org.

The Fusion Registry includes a data samples directory in the distribution, which contains the same dataset described for each of the formats contained in this document.

It is important to note that this document discusses the advantages of some of the formats with regards to the level of validation that can be achieved by an XML Schema (XSD), the Fusion Registry however uses its own data validation processes to give a deeper level of validation than an XML schema can perform. The Fusion Registry data validation processes are the same regardless of data import format, and therefore the notes in this document about schema validation are for information only. The chosen data import format has no impact on the level of validation that will be performed on the imported dataset.

3 SDMX Data Formats

3.1 Generic Data

Generic Data is in all versions of the SDMX Standard, each version is supported by the Fusion Registry including GenericTimeSeriesData which was introduced in version 2.1 of SDMX.

A Generic Dataset is defined by the SDMX Schemas (xsd files) distributed in section 3B of the standard. A Generic Dataset contains the same XML nodes and attributes regardless of the data it is transmitting, hence the name 'generic'. Below is an example of a series in a Generic Dataset.

The SDMX Schema can validate a Generic Dataset but it cannot be used to verify any values in the reported dataset.

```
<generic:Series>
  <generic:SeriesKey>
    <generic:Value id="GEO" value="AT"/>
    <generic:Value id="SEX" value="T"/>
    <generic:Value id="HST" value="NPRV"/>
    <generic:Value id="LMS" value="DIV"/>
    <generic:Value id="CAS" value="POP"/>
    <generic:Value id="POB" value="TOTAL"/>
    <generic:Value id="COC" value="TOTAL"/>
    <generic:Value id="AGE" value="TOTAL"/>
    <generic:Value id="FREQ" value="A"/>
  </generic:SeriesKey>
  <generic:Attributes>
    <generic:Value id="TIME_FORMAT" value="P1Y"/>
  </generic:Attributes>
  <generic:Obs>
    <generic:ObsDimension value="2011"/>
    <generic:ObsValue value="10488.0"/>
    <generic:Attributes>
      <generic:Value id="OBS_STATUS" value="I"/>
      <generic:Value id="OBS_NOTE" value="LMS: Including persons whose same-sex registered partnership was legally dissolved."/>
    </generic:Attributes>
  </generic:Obs>
</generic:Series>
```

Figure 1 showing a single series for a Generic Dataset

3.2 Structure Specific / Compact Data

Structure Specific Data is a SDMX v2.1 re-branding of the term Compact Data which existed in versions 1.0 and 2.0 of the SDMX specification. The Fusion Registry supports each version including StructureSpecificTimeSeriesData which was introduced in version 2.1 of the SDMX specification.

Section 3B of the SDMX distribution includes the abstract schema (XSD) definition for a Structure Specific dataset, but the concrete implementation is dependent on what Provision Agreement, Dataflow, or Data Structure Definition the dataset is for.

The Fusion Registry is able to generate a Structure Specific XSD on request from the web services page.

As the generated Schema is specific to the structural metadata held in the Fusion Registry, the level of validation offered by the schema is far greater than with a Generic dataset. A Structure Specific dataset is also more terse than a Generic dataset, as shown in the example below.

```

<Series GEO="AT" SEX="T" HST="NPRV" LMS="DIV" CAS="POP" POB="TOTAL" COC="TOTAL" AGE="TOTAL" FREQ="A" TIME_FORMAT="P1Y">
<Obs TIME_PERIOD="2011" OBS_VALUE="10488.0" OBS_STATUS="I" OBS_NOTE="LMS: Including persons whose same-sex registered partnership wa
</Series>
<Series GEO="AT" SEX="T" HST="NPRV" LMS="MAR" CAS="POP" POB="TOTAL" COC="TOTAL" AGE="TOTAL" FREQ="A" TIME_FORMAT="P1Y">
<Obs TIME_PERIOD="2011" OBS_VALUE="17811.0" OBS_STATUS="I" OBS_NOTE="LMS: Including persons in a same-sex registered partnership."/>
</Series>

```

Figure 2 showing an example of two series in Structure Specific / Compact format

3.3 SDMX-EDI

SDMX-EDI is a 1990s syntax that uses the UN/EDIFACT syntax. It is an extremely terse data format that supports time series only. The format is still popular for reporting financial data in the Central Banking community. However, unless you have a specific requirement to use this format then it is not recommended as the syntax is largely superseded by 21st century syntaxes such as XML and JSON.

For more information on the structure and syntax of SDMX-EDI, please visit sdmx.org.

```

ARR++AT:T:NPRV:DIV:POP:TOTAL:TOTAL:TOTAL:A:2011:602:10488.0:I'
ARR++AT:T:NPRV:MAR:POP:TOTAL:TOTAL:TOTAL:A:2011:602:17811.0:I'

```

Figure 3 showing an example of two series represented in SDMX-EDI format