

**Tertiary Education
Commission**

Te Amorangi Mātauranga Matua



Guidance for CSV files

Additional Guidance v.02
Released 15 September 2023

Document purpose

This document is for Tertiary Education Organisations (TEOs) delivering provider-based learning reporting via the Single Data return (SDR). It provides guidance about Comma-Separated Values (CSV) files to be used alongside the Provider-based Data Specification.

CSV files are a simple file format used to store tabular data, like a spreadsheet or a database.

Guidance

1. Structure of a CSV File

A CSV file is essentially a list of rows, where each row contains values separated by commas. Think of it as a table where each row is a new line in the file, and each value within that row is separated by a comma.

2. Column Titles (Headers)

The first row of a CSV file often contains the column titles, also known as headers. These titles describe the data in each column. For example:

```
Provider Code,Student Identification Code,Gender,Name ID Code  
1234,5678901234,M,LEE A  
2345,3567892345,F,VAN M
```

In this example, "Provider Code", "Student Identification Code", "Gender", and "Name ID Code" are the column titles.

3. Handling Commas in Data

Since commas are used to separate values, you might wonder what happens if the data itself contains a comma. In such cases, the data is enclosed in double quotes.

For example, if someone's Name ID Code is "Alice, Jr", the CSV would look like:

```
Gender,Name ID Code  
F,"JR, A"
```

4. Handling Numbers

Numbers in CSV files are written as they are, without any commas for thousands. For instance, one thousand is written as 1000, not 1,000.

5. Columns with Multiple Entries

Sometimes, a column might have multiple entries for a single row. We will handle this using a common method separating these entries with a semicolon.

For example, if Alice has 3 Ethnicities, the CSV might look like:

```
Name ID Code,Ethnicity  
"JR, A",111;121;129
```

6. File names for the 5 SDR files

- Learner file: LEAR[Providercode].csv
- Course Register: REGT[Providercode].csv
- Course Enrolment File: ENRL[Providercode].csv
- Course Completion File: COURCOMP[Providercode].csv
- Qualification Completion File: QUALCOMP[Providercode].csv

7. File size

File size should not exceed a certain size (please note that discussions are ongoing with the development team as we have observed course enrolment file size for certain universities exceeding 47MB in production).

While partial file uploads are not permitted, you can resubmit, and the latest submission will be considered.

8. Tips for Working with CSV Files

- Software: You can open and edit CSV files with many software programs, including Microsoft Excel, Google Sheets, and plain text editors like Notepad.
- Encoding: CSV files are usually saved in a standard character encoding like UTF-8. This ensures that special characters (like accents in names) are displayed correctly.
- Encoding Issues: Characters like accented letters or symbols from non-Latin scripts need to be saved in a compatible encoding (like UTF-8). If not, they might appear as gibberish or cause errors when the CSV is opened in software that doesn't support that encoding.
- Consistency: Ensure that every row has the same number of columns. If a particular data point is missing, the comma should still be there to indicate an empty space.
- Special Characters in Headers: While headers typically contain simple alphanumeric descriptions, they can also contain spaces, numbers, and other special characters. It's a good practice to keep headers simple and without special characters to avoid confusion, but if needed, they can be enclosed in double quotes.

Published by the Tertiary Education Commission
Te Amorangi Mātauranga Matua

National Office
44 The Terrace
PO Box 27048
Wellington, New Zealand

June 2023

Authors: The Tertiary Education Commission

Every effort is made to provide accurate and factual content. The TEC, however, cannot accept responsibility for any inadvertent errors or omissions that may occur.



This work is licensed under the Creative Commons Attribution 4.0 International licence. You are free to copy, distribute, and adapt the work, as long as you attribute the work to the Tertiary Education Commission and abide by the other licence terms. Please note you may not use any departmental or governmental emblem, logo, or coat of arms in any way that infringes any provision of the Flags, Emblems, and Names Protection Act 1981.