



borealis The Canadian Dataverse Repository
Le dépôt Dataverse canadien



Odesi Deposit Guide (using Borealis Dataverse DDI 2.5)

Reference Guide

February, 2024

Odesi Markit! Program

Table of Contents

INTRODUCTION-----	3
1. OBTAINING AND PREPARING THE DATASET-----	3
a. Obtaining a Dataset:-----	3
b. Correcting and editing a Dataset that is already published in Odesi Dataverse:-----	4
c. Preparing the Dataset-----	4
2. CREATING AN ACCOUNT AND LOGIN PROCEDURES-----	4
a. Create an Account:-----	4
b. Logging into Borealis:-----	5
3. NAVIGATING THE ODESI DATAVERSE-----	5
4. ODESI DATAVERSE PERMISSIONS-----	6
5. CREATING A DATASET-----	6
a. Adding a New Dataset:-----	6
b. Create Citation Metadata Record:-----	7
c. Add Files:-----	7
6. ADDING METADATA AND EDITING A DATASET-----	9
a. Add and Edit Files:-----	9
b. Add Metadata:-----	10
c. Add Dataset Access Terms:-----	11
7. VARIABLE METADATA-----	12
a. Accessing the Data Curation Tool-----	12
If you are using a pop-up blocker it will prevent the Data Curation Tool from opening. Either disable your pop-up blocker or permit borealisdata.ca to send pop-ups and use redirects.-	12
b. Variable Groups-----	12
c. Variable Metadata Documentation-----	13
8. PUBLISHING A DATASET:-----	14
9. APPENDIX A: PREPARING AND WORKING WITH AN .SPSS SYNTAX FILE-----	14
1. Working with an .sps syntax file (creating an SPSS system file .sav)-----	15
2. Working with an SPSS system file (.sav) file.-----	17
10. APPENDIX B: QUICK MARK-UP FILES-----	17
11. APPENDIX C: DEPOSITING SYNTHETIC DATA-----	18
12. APPENDIX D: DEPOSITING RESTRICTED DATA-----	19
13. APPENDIX E: SPSS & FRENCH CHARACTERS-----	21

INTRODUCTION

This guide is centered on marking up and publishing Statistics Canada microdata files for [Odesi](#). In some cases, marking up other files - such as Public Opinion Polls or Surveys from non-Statistics Canada sources - will follow the same steps outlined in this guide; however, the data file will come from different sources and different DDI fields may be used for marking up public opinion polls or non-Statistics Canada files. Mark-up students should consult their supervisors for further information when marking up non-Statistics Canada files.

Detailed guidance on best practices for curating metadata in the [Odesi repository](#) is available in the [Odesi Metadata Best Practices Guide](#).

1. OBTAINING AND PREPARING THE DATASET

You will need to have local copies of the datasets - including microdata and aggregate data files - to work with in order to import to Borealis, add metadata, and publish the result in an Odesi Collection Dataverse. In most cases, the DLI contact for your institution will obtain local copies of the dataset, since they have the passwords required to access the Statistics Canada files. Mark-up students should consult their supervisors to retrieve the files from the appropriate source.

a. Obtaining a Dataset:

- i. Statistics Canada may already provide a copy of the data file and metadata. Check with your supervisor to obtain a copy from Statistics Canada's Data Liberation Initiative (EFT) or an equivalent site.
- ii. If Statistics Canada has not provided the data file, Mark-up students need to ask their supervisor to download the required data file.
- iii. Download all available datasets and organize them. These files may be available in the form of an ASCII text file (.txt or .dat) and a syntax file (.sps) etc., but will also include full system files in SPSS, SAS, STATA, and csv format. See Appendix A if you are working with an .sps syntax file.
- iv. While documentation in .xml format typically will not have been created at this early stage, check for any available metadata in any format to include or reuse.

- v. Documentation in the form of a codebook, user guide, and/or a questionnaire should be available along with the data files (.pdf format usually).

2. CREATING AN ACCOUNT AND LOGIN PROCEDURES

a. Create an Account:

- i. If you do not already have an account, navigate to the [Borealis homepage](#) and click the “Login” option in the top right corner of the page.

Note: If your institution is already configured to use Shibboleth login, you can simply sign in to create an account. Logging in with your institution’s Shibboleth login credentials will automatically create an account for you.

- ii. Click the “Sign Up” button and complete the “Account Information” form. The Affiliation field will auto-fill with the institution associated with your email address.

Account - Borealis

The screenshot shows the 'Account Information' form on the Borealis website. At the top, there is a tab labeled 'Account Information'. Below the tab, a message states: 'You can also create a Dataverse account with one of our other log in options.' The form contains four main fields: 'Email *', 'Affiliation *', 'Username *', and 'Password *'. Each field has a small question mark icon to its right. The 'Email *' field is a text input box. The 'Affiliation *' field is a greyed-out text input box. The 'Username *' field has a text input box and a description: 'Create a valid username of 2 to 60 characters in length containing letters (a-z), numbers (0-9), dashes (-), underscores (_), and periods (.).' The 'Password *' field has a text input box and a description: 'Your password must contain:' followed by two bullet points: '• At least 6 characters (passwords of at least 20 characters are exempt from all other requirements)' and '• At least 1 character from each of the following types: letter, numeral'.

- iii. After filling out the required account information, check the box at the bottom of the page indicating that you have read and accept the Borealis General Terms of Use, and then click the “Create Account” button.
- iv. Send your username to your supervisor. Your supervisor will add your username to the ODESI Admin Group which gives you special permissions to deposit data in the Odesi Dataverse.

Log In

Log in or sign up with your institutional account — more [information about account creation](#). Leaving your institution? Please contact [Borealis Support](#) for assistance.

Your Institution

[Allow me to type the name of my institution](#)

Institution not listed?

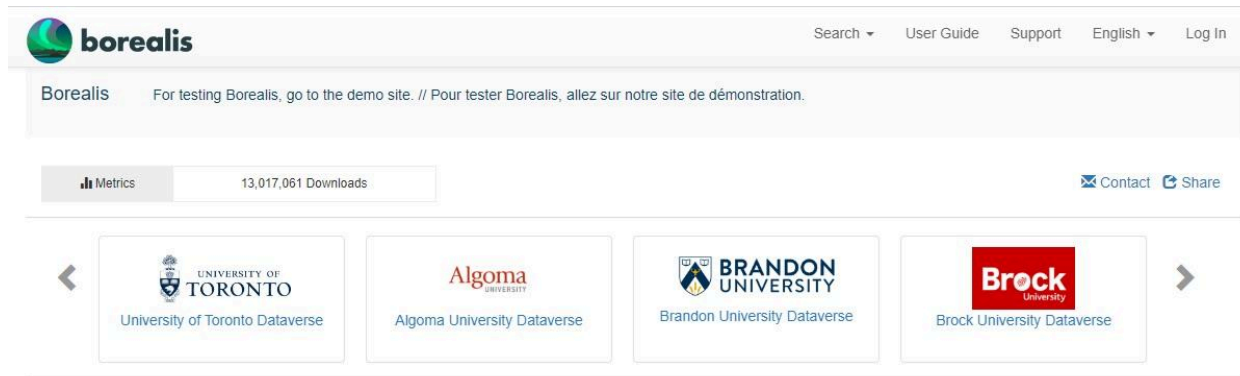
b. Logging into Borealis:

- i. After navigating to the [Borealis homepage](#), click the “Login” button in the top right corner of the page. You will have the option to login using your username/email or through your institutional login.
- ii. If you want to login through your institution, select your institution from the drop down window and click the “Continue” button. This will direct you to your institution’s login page. Login with your institutional credentials and you will land in your institution’s Dataverse.
- iii. To view all of the Dataverses in Borealis, click the Borealis logo in the top left corner to return to the homepage, and then click the “Explore Borealis” button OR click the “Borealis >” link located underneath the name of the Dataverse from within your current Dataverse.

3. NAVIGATING THE ODESI DATAVERSE

The [Odesi dataverse](#) contains thousands of datasets, including collections of Survey Microdata, Public Opinion Polls, Administrative Data, Summary Data, and relevant documentation. Users can search across data collections as well as filter, explore, tabulate, visualize, download, and subset datasets and variables.

- a. The Dataverses in Borealis correspond to the various institutions that publish in Borealis. They are accessible through a preview panel across the top of the page. **For Odesi publications, the collections dataverses exist within the [Odesi Dataverse](#).**



- b. When you visit a specific Dataverse, there will be a description of the collections that exist within that Dataverse. For the Odesi Dataverse, these collections include Public Use Microdata Files, Aggregate Data, etc. Determine which of these Odesi collections the Dataset you are working with belongs to and navigate to that collection’s Dataverse before creating the Dataset in Borealis. This will ensure that the Dataset is published under the proper collection.
- c. If you are interested in working with other collections in Borealis, please refer to the [Borealis User Guide](#) or contact the Institutional Collection owner.

4. ODESI DATAVERSE PERMISSIONS

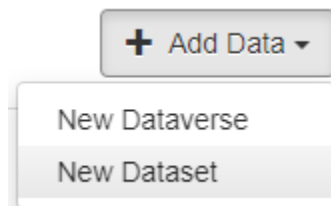
There are two permission groups in the Odesi Dataverse: ODESIAdmin and OdesiStudent. Members of the ODESIAdmin group are assigned an Admin role and have access to view and edit everything in the Odesi Dataverse. Members of the OdesiStudent group are assigned a Contributor Plus role, which allows them to create new Dataverses and Datasets, edit Dataset License + Terms, edit Dataset Permissions, and publish Datasets. All new students will be added to the OdesiStudent Group.

Note: For information concerning permissions for restricted data, see [Appendix E](#).

5. CREATING A DATASET

a. Adding a New Dataset:

- i. In the Odesi Dataverse, click the “+ Add Data” button and select “New Dataset” from the drop-down menu. If you do not see the “+ Add Data” button, it is because you do not have permission to deposit data in that Dataverse. Contact your supervisor to be granted deposit permissions. Do not create a new Dataverse unless directed to do so by your supervisor.



b. Create Citation Metadata Record:

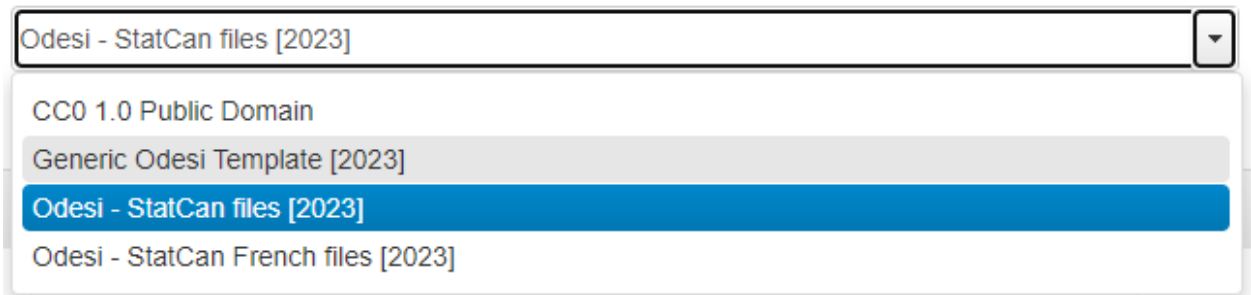
- i. Ensure you are creating the dataset in the appropriate Host Dataverse.

Host Dataverse 

Changing the host dataverse will clear any fields you may have entered data into.

Statistics Canada Public Use Microdata Files (PUMFs) / Collection de fichiers de microdonnées à c

- ii. For Odesi publications, there are two metadata templates that you can use: the Odesi - StatCan files template (for English and French) and the Generic Odesi template. The templates will provide instructions on required metadata fields, controlled vocabularies, and other Odesi-specific practices when creating metadata.



Odesi - StatCan files [2023]

CC0 1.0 Public Domain

Generic Odesi Template [2023]

Odesi - StatCan files [2023]

Odesi - StatCan French files [2023]

- iii. Required Fields: There are a number of Citation Metadata fields that must be filled in order to save a Dataset. These mandatory fields are indicated by an asterisk * and include:

- Title
- Author -> Name
- Contact -> Email
- Description -> Text
- Subject

There are additional required metadata fields that will be accessible once the Dataset has been saved. You will need to save the Dataset and then add the metadata to these fields following the steps in [Section 5: Adding Metadata and Editing a Dataset](#). For more information on required metadata fields, see the [Odesi Metadata Best Practices Guide](#).

- iv. Some metadata fields are repeatable, meaning they can record more than one value. Repeatable fields can be entered using the “+” button located on the right of the field. Refer to the [Odesi Metadata Best Practices Guide](#) for information about which fields are repeatable.

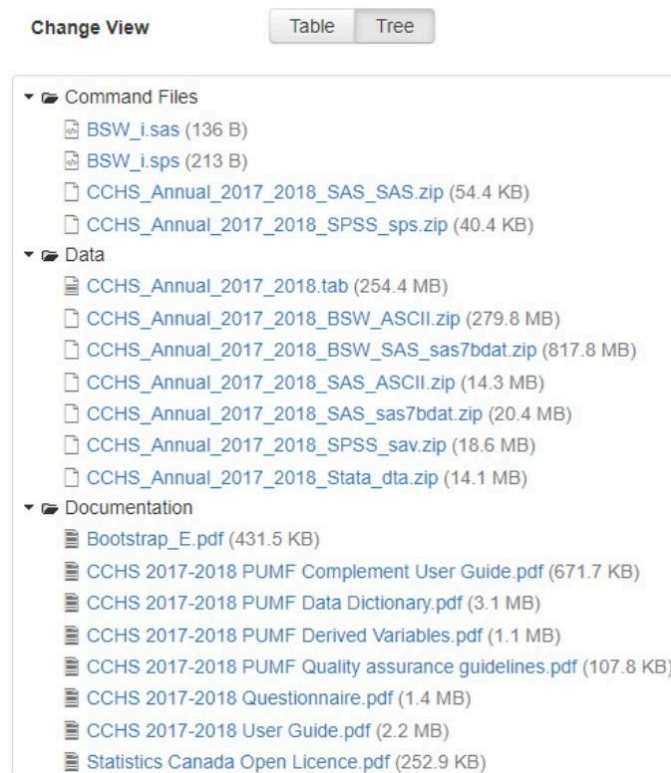
c. **Add Files:**

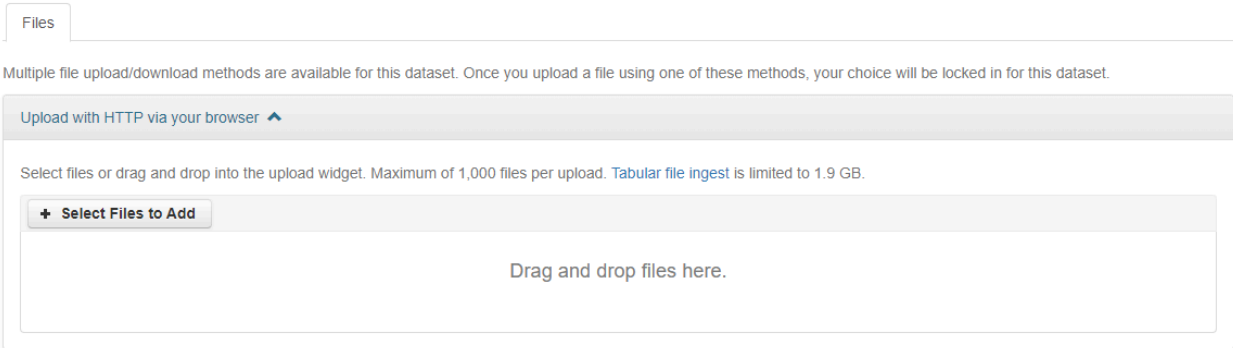
Before uploading, prepare a deposit package (.zip) that contains all available data and documentation files in different formats (e.g. SPSS, SAS, STATA, Excel, CSV, TXT, PDF). Dataverse ingests certain tabular data formats (e.g. SPSS, Excel, CSV, STATA, etc.) and creates a derivative .TAB file. The SPSS data file is always used to create the .TAB file and all other data formats that are not documentation should be zipped (.zip) to ensure that they are not converted to TAB files.

Note: When uploading all the files, please note a .zip package containing the SPSS, documentation and pre-zipped other data formats is preferred.

- i. Files can be uploaded either by clicking the “Select Files to Add” button or by dragging and dropping the files into the widget.
- ii. Files can also be added after the Dataset is saved.

Note: The maximum size of file that can be imported is approximately 3 Gigabytes, and 500 Megabytes for tabular files (.spss). Larger files can also be zipped and then uploaded.





- iii. When adding files, you should include all file formats of the data that are released by Statistics Canada but only make edits [\(Section 6\)](#) to the .spss file. If the release only includes an .spss file, consult your supervisor to determine if file conversion is necessary.
- iv. Drag and drop files into the widget OR click the “+ Select Files to Add” to open a file browser and choose the desired file(s). If you upload the wrong file, select the check box on the left, then click the “Edit” button and choose “Delete.”
- v. File Metadata: Make sure the file name is correct and add a file tag (either “Data” or “Documentation”) by clicking the three-dot menu, which distinguishes the file between Data files and Documentation files. Fill in the description field using the controlled vocabulary found in the [Odesi Metadata Best Practices Guide](#).



- vi. If you need to restrict access to the data file(s), select the checkbox to the left of the data file(s) then click the “Edit” button and select “Restrict.” Fill in the Terms of Access for Restricted Files and click the “Save Changes” button.

Note: See [Appendix E](#) for information on assigning access permissions.

- vii. Once all of the files are uploaded and tags and restrictions are set, click the “Save Changes” button in the bottom left corner of the page. After inputting the Citation Metadata and adding files, click the “Save Dataset” button. It may take a while for tabular

files to ingest. Always wait until the files have finished uploading before proceeding to add or edit metadata.

d. Correcting and editing a Dataset that is already published in Odesi Dataverse:

You will be able to edit the files and metadata by navigating to the Dataset and selecting an option from the “Edit Dataset” drop down menu.

If you do not have permission to edit a Dataset, contact your supervisor and they will adjust your permissions.

6. ADDING METADATA AND EDITING A DATASET

Once a Dataset has been saved, you will need to add additional Metadata as well as Terms and Conditions to the Dataset before publishing. These features can be accessed either by clicking the “Edit Dataset” drop-down button in the top right, or by navigating to the “Metadata” tab at the bottom of the page and clicking the “Add + Edit Metadata” button on the right.



a. Add and Edit Files:

Once the Dataset has been saved, you will be able to add and edit files. To do so, click the Edit Dataset drop down button and select “File (Upload)” OR Select the File tab and click the “+ Upload Files” button. Follow the instructions from [Section 4. c.](#) to complete the remaining steps.

b. Add Metadata:

When creating a metadata record, correct spelling and grammar are of the utmost importance. It may seem like you are repeating yourself when you enter metadata in multiple sections and that is because you are! Remember that the different fields into which you are entering information will appear in different parts of the finished document and must be complete on their own.

Detailed guidance and best practices for curating metadata, including the required fields and controlled vocabularies used for Odesi publications, are available in the [Odesi Metadata Best Practices Guide](#).

- i. Click the “Edit Dataset” drop-down button and select “Metadata”, or navigate to the Metadata tab and click the “Add + Edit Metadata” button.

Citation Metadata

Title * ?

Subtitle ?

Alternative Title ?

Alternative URL ?

Other ID ?

Agency ? Identifier ? +

Author * ?

Name * ? Affiliation ? +

Identifier Scheme ? Identifier ?

Contact * ?

Name ? Affiliation ? +

E-mail * ?

- ii. Following the [Odesi Metadata Best Practices Guide](#), input metadata into the appropriate fields. Required fields are marked with an asterisk *. Hovering over the question icon beside the name of each field will reveal a description of the field, but the instructions in the [Odesi Metadata Best Practices Guide](#) always take precedence over these descriptions.

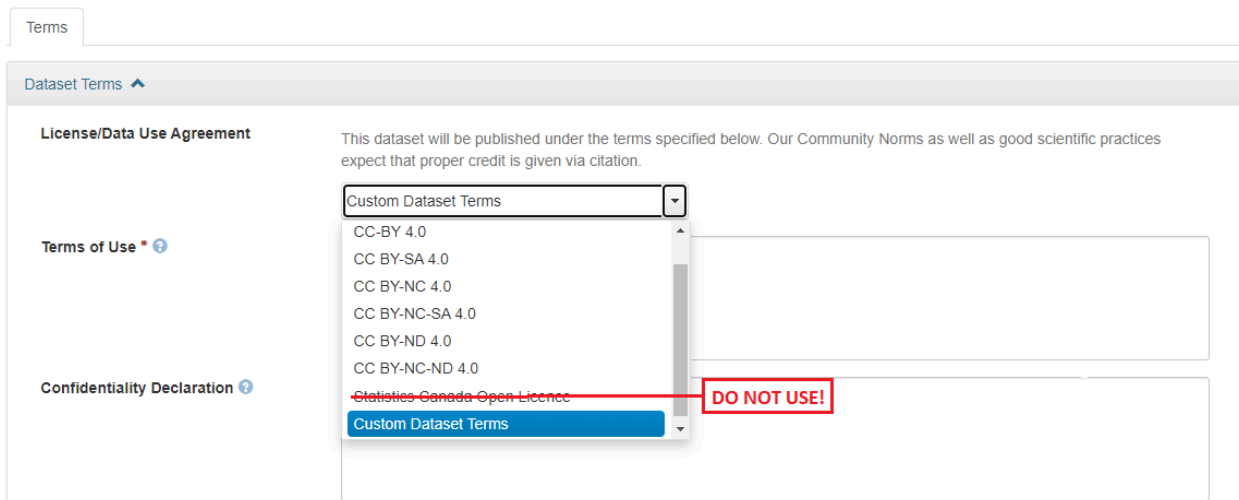
- Geospatial Metadata
- Social Science and Humanities Metadata
- Astronomy and Astrophysics Metadata
- Life Sciences Metadata
- Journal Metadata

- iii. The Citation Metadata drop-down menu will be open when you land on the page. There are additional drop down menus at the bottom of the page. The Geospatial Metadata and the Social Science and Humanities Metadata drop-down menus both contain required fields for Odesi documentation.
- iv. If a field is repeatable, use the “+” button to the right of the field to create a new input (e.g. text box or drop-down window) for that field.
- v. For guidance on filling out the metadata fields, see the [Odesi Metadata Best Practices Document](#).
- vi. Once all of the required fields are completed, click the “Save Changes” button.

c. Add Dataset Access Terms:

- i. Click the “Edit Dataset” drop-down button and select “Terms” OR navigate to the Terms tab and click the “Edit Terms Requirements” button.
- ii. In the Dataset Terms field, choose a License/Data Use Agreement. The default option is a Creative Commons CCO 1.0 Universal Public Domain Dedication.
- iii. For Odesi publications, select the Custom Dataset Terms template and fill in the Terms field with the appropriate license. Using the Odesi - StatCan files template will fill in the Terms of Use field with the Statistics Canada Open Licence.

CAUTION: Do **NOT** use the Statistics Canada Open Licence option in the License/Data Use Agreement drop-down menu (it does not map to DDI at this point).

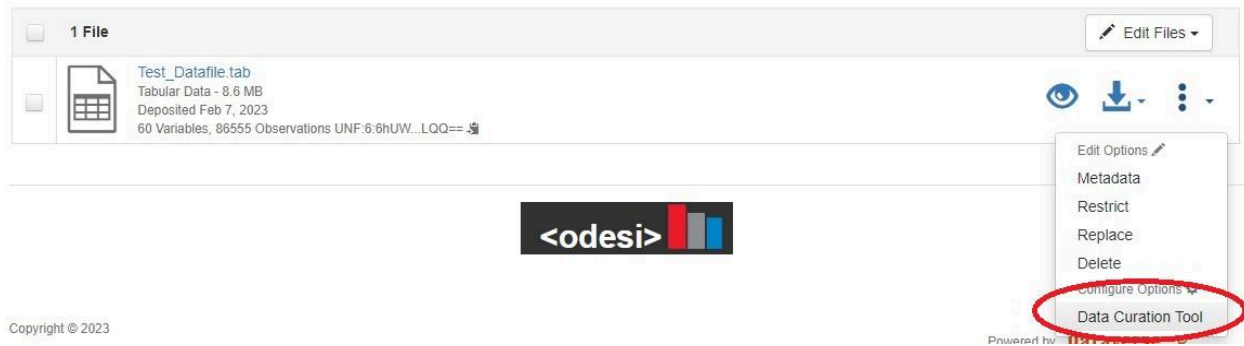


- iv. See [Appendix E](#) if you are publishing Restricted Data.

7. VARIABLE METADATA

a. Accessing the Data Curation Tool

Navigate to the file that you would like to edit. Click the three-dot drop-down and select “Data Curation Tool.” The Data Curation Tool will open in a new tab.



Note: If you are using a pop-up blocker it will prevent the Data Curation Tool from opening. You will need to either disable your pop-up blocker or permit borealisdata.ca to send pop-ups and use redirects.

Once you have opened a tabular data file in the Data Curation Tool, always check to ensure that the file is ready for public use. You will need to check for the following:

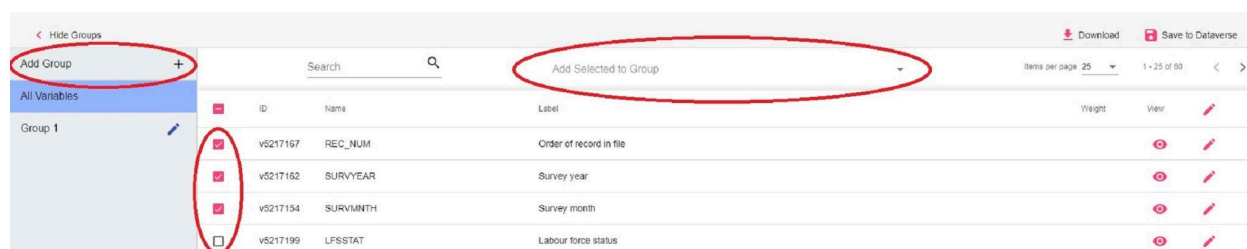
- an id variable;
- all the variable labels are correct and unique (avoid lengthy labels/repeats that read as a literal question);
- all the value labels are correct;
- all the missing values have been declared;
- all the recoded variables are correct; and
- the data are not weighted.

b. Variable Groups

Variables that may share a common subject, arise from the interpretation of a single question, or are linked by some other factor are arranged into Variable Groups. The Variable Groups can be found in the documentation provided by the data producer (codebook, data dictionary, etc).

For best practices on creating variable group metadata, see the [Odesi Metadata Best Practices Guide](#).

- i. In the Data Curation Tool, click the “Add Group +” button on the left-hand side of the page to create a variable group. The Variable Groups will be listed in the column on the left-hand side underneath the “All Variables” group. It is Odesi best practice to arrange the Variable Groups in the same order as they appear in the documentation provided by the data producer.
- ii. Once you have created and named the group, return to the list of All Variables and click the checkboxes next to the variables that belong to that group. Once you have selected all of the variables that belong to a particular Variable Group, scroll to the top of the page and select the Variable Group from the drop-down list in the centre of the page. The variables will be added to the Variable Group you select.
- iii. Follow the supporting documentation for the dataset to find information about naming the variable groups and to determine which variables belong to each group.



c. Variable Metadata Documentation

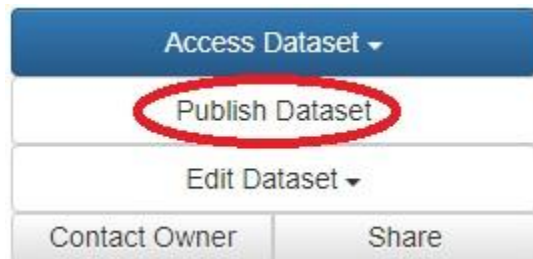
- i. Click the pencil icon to add and edit metadata for each variable.
- ii. A window will open with metadata fields for you to fill out and you will also be able to add the variable to a group.
- iii. If the variable you are working with is a weight variable, the “Is Weight” checkbox must be selected to indicate the variable’s status as a weight variable. All other variables will need to be weighted with that variable by selecting the weight variable from the “Weight Variable” drop-down menu.

Note: When weighting all other variables, make sure the weight variable is **NOT** selected, otherwise the weight indication will be removed.

- iv. When you have finished adding the variable metadata, click the “Update” button to apply the changes.
- v. Once you have finished marking up the Variable Metadata and adding the variables to groups, click the “Save to Dataverse” button to save your work.

8. **PUBLISHING A DATASET:**

Once you have completed all of the steps described above, the Dataset is ready to be published. Navigate to the Dataset and click the “Publish Dataset” button on the right side of the page.



If you have added Access Terms to the Dataset, a popup window will open for you to confirm that the Terms are correct before publishing. After confirming the Access Terms, click “Continue” to finish the publication process.

9. **APPENDIX A: PREPARING AND WORKING WITH AN .SPSS SYNTAX FILE**

Before marking up the file (entering the metadata into Borealis), always check to ensure that the file is ready for public use. You will need to check for the following:

- there is an id variable;

- all the variable labels are correct and unique (avoid lengthy labels/repeats that read as a literal question);
- all the value labels are correct;
- all the missing values have been declared;
- all the recoded variables are correct;
- any alphanumeric variables are truly alphanumeric; and
- the data are not weighted.

This guide assumes that you will perform these checks using SPSS. Generally, you will have either an SPSS system file (.sav format – obtained from Odesi or Statistics Canada), or you will have a combination of an ASCII text file (.txt or .dat) and a syntax file (.sps – obtained from Statistics Canada EFT site). The following will describe how to prepare the dataset in each case.

1. Working with an .sps syntax file (creating an SPSS system file .sav)

Open the syntax file (.sps file) by double-clicking on it. The .sps file is the program file which converts the raw data file (ASCII text file) to an SPSS system file (.sav).

Note: Statistics Canada will sometimes release the syntax file as reading cards or layout cards with each section of the syntax in its own file. These files usually have different extensions but can be opened with a text editor (e.g. Notepad++) and merged to create the syntax file which can then be saved as a .sps file. Reading/Layout cards extensions:

- .cde/.cdf – value labels in English/French
- .lay – record layout
- .lbe/.lbf – variable labels in English/French
- .mvs – missing values

a) Modify the SPSS command code

See sample SPSS syntax file (next page) for instructions on editing the syntax file. Tips on editing SPSS command code:

- all commands end with a period.
- all commands start in the first column; subsequent lines in the same command are indented by 1-2 columns.
- all lines of text are in quotation marks
- LRECL (Logical Record Length) can be determined by looking at the last position of the last variable on the raw data file, or the last number at the end of the DATA LIST FILE.

b) Run the file

SPSS will run the syntax file's code against the referenced raw data file.

- i) Save the file.
- ii) From the Menu bar, use the menu command: **Run** → **All** to run the code. When it is done, "SPSS processor is ready" will appear at the bottom of the screen.

c) Check results

- i) Check that the SPSS System File (.sav) was saved to the disk location you specified at the bottom of your SPSS program (.sav file will have been added to the working directory / folder).
- ii) From within the data editor screen (the one that looks like a spreadsheet), use the menu command: **Analyze** → **Descriptive Statistics** → **Frequencies**.
- iii) Choose an appropriate variable (e.g. Province, or some other categorized variable).
- iv) Run the analysis and compare your results (e.g. number of cases) with the number of cases indicated in the "Readme" file or Data Dictionary for the survey .
- v) The .sav file can now be imported into Borealis.

This document is a sample SPSS Syntax file

```

1  * Encoding: windows-1252
2  TITLE "General Social Survey, Cycle 27: Giving Volunteering and Participation"
3  Set Length=none width=80.
4  FILE HANDLE MICRO1 /NAME = 'H:\Surveys\GVP_PUMF_MAIN.TXT' /rform=fixed /rcli=5877.
5
6  DATA LIST FILE = MICRO1
7  /RECID          1-   5
8  /WGHT_PER      6-  15 (4)
9  /DH1GAGE      19-  20
10 /AGEGR10       21-  22
11 /WTBS_001     878- 887 (4)
12 /WTBS_002     888- 897 (4)
13
14 VARIABLE LABELS
15   RECID          "Record identification"
16   WGHT_PER       "Person weight"
17   DH1GAGE        "Age group of respondent (6 categories)"
18   AGEGR10        "Age group of respondent (groups of 10)"
19   WTBS_001       "Bootstrap weight # 1 for personal weight"
20
21 VALUE LABELS
22   /AGEGR10
23     01 "15 to 24 years"
24     02 "25 to 34 years"
25     03 "35 to 44 years"
26     04 "45 to 54 years"
27     05 "55 to 64 years"
28     06 "65 to 74 years"
29     07 "75 years and over"
30     96 "Valid skip"
31     97 "Don't know"
32     98 "Refusal"
33     99 "Not stated"
34   /BRTHCAN
35     1 "Born in Canada"
36     2 "Born outside Canada"
37     6 "Valid skip"
38     7 "Don't know"
39     8 "Refusal"
40     9 "Not stated"
41   /VD1DHRS VD1_04A VD1_04B VD1_04C VS_050
42     9996 "Valid skip"
43     9997 "Don't know"
44     9998 "Refusal"
45     9999 "Not stated"
46
47 MISSING VALUES
48   AGEGR10          (96 THRU 99)
49   AGEPRGR6         (96 THRU 99)
50   AM1DVG           (96 THRU 99)
51   VD1DT201        (9996 THRU 9999)
52   WGHT_PER         (99999.9996 THRU 99999.9999)
53   WTBS_001        (99999.9996 THRU 99999.9999)
54   WTBS_002        (99999.9996 THRU 99999.9999)
55
56 SAVE OUTFILE = 'H:\Surveys\GVP_PUMF_MAIN_E.SAV' / COMPRESSED.
57 EXECUTE.

```

Annotations:

- Title in quotes:** Points to the TITLE command.
- Sets the maximum page length and width for text output:** Points to the Set Length=none width=80. command.
- Same name:** Points to the FILE HANDLE command.
- Path to where the data file is:** Points to the FILE HANDLE command.
- Logical Record Length (Find by looking at the last position of the last variable on the raw data):** Points to the /rcli=5877. command.
- Numbers identify position of variables in raw data file. (1) - indicates number of decimal places to read. (A) - Indicates this variables is alphanumeric:** Points to the DATA LIST FILE command.
- Clean up variables if needed (e.g. spell words):** Points to the VARIABLE LABELS command.
- Make changes if necessary - i.e typos:** Points to the VALUE LABELS command.
- Always check to make sure all missing values have been declared. All Odesi files should have missing values declared.** Points to the MISSING VALUES command.
- Path to where data file is to be saved. Give dataset a meaningful name with survey acronym and year.** Points to the SAVE OUTFILE command.

2. Working with an SPSS system file (.sav) file.

If you are starting directly with a .sav file, you do not need to go through the process described in part one above.

- Open the .sav file in SPSS.
- Go through the checklist listed at the beginning of this Appendix.
- Make any required changes.
- Save the file.
- The .sav file can now be imported into Borealis.

10. APPENDIX B: QUICK MARK-UP FILES

When a data file is in high demand, a quick mark-up of the file is done to get it published within a few days of its release. Quick mark-up is not a preferred practice and it is the responsibility of the school that publishes the quick file to complete the mark-up.

a. ODESI Mark-up list

- i. If the survey is not listed, add it.
- ii. Status column – add “Quick file”.
- iii. Add the name of the school responsible for the quick mark-up to the “OCUL School Responsible” column.

b. Mark-up the file making sure that the following sections are completed:

- i. Document Description – all fields.
- ii. Study Description – as much of the metadata as possible.
 - Using the XML file from the previous version (if available) of a survey will make doing this section much faster.
- iii. Variable groups.
- iv. Other Materials.

c. Publish the quick file.

Note: If you are not able to complete the full mark-up, send an email to the MarkIt! Group (ocul-odesi@lists.uwaterloo.ca) to ask someone else to finish the mark-up.

d. Complete the markup of the file.

- i. Add question text, universe and notes.
- ii. Review the metadata to make sure that all fields are filled out.
- iii. Republish file.

e. Update the ODESI Markit list

- i. Status column: change from “Quick file” to “Completed”.

- ii. Add school name to the OCUL School Responsible column.
- iii. Add date completed to the Date Completed column.

11. APPENDIX C: DEPOSITING SYNTHETIC DATA

Note: Synthetic data files are not real data, that is, they are not the same as microdata or aggregate data files. Therefore, they require different mark-up procedures.

SPSS

- Open the dataset in SPSS.
- Eyeball the data to make sure that every variable has a label.

Dataverse

Create a Dataset and begin creating the metadata.

Follow the [Odesi Metadata Best Practices Guide](#) for entering the metadata **EXCEPT** for the following tags:

- **Title <titl>2.1.1.1**
 - All titles, subtitles, alternative titles start with “Synthetic: Title”
 - e.g. Synthetic: Aboriginal Children’s Survey, 2006 [Canada]
- **Alternative title <altTitl> 2.1.1.3 and 2.1.1.4**
 - Do **not** enter a parallel title.
- **Other ID <IDNo> 2.1.1.5**
 - Follow the formula given in the [Odesi Metadata Best Practices Guide](#) with the addition of the word “synthetic” at the beginning of the ID
 - Example 1: synthetic-cchs-E-2012-mental-health
 - Example 2: synthetic-nlscy-E-1998-1999-c-3-primary-file
- **Date of Collection <collDate>**
 - Do **not** enter the “Date of collection.”
- **Description <abstract> 2.2.2**
 - The following paragraph must be added to the Abstract. It should precede the description of the survey:

“Please note: This is a Synthetic data file, also known as a Dummy File – it is **NOT** real data. This synthetic data file should **NOT** be used for purposes other than to develop and test computer programs that are to be submitted by remote access. Each record in the synthetic file matches the format and content parameters of the real Statistics Canada Master File with which it is associated, but the data

themselves have been *'made up'*. They do **NOT** represent responses from real individuals and should **NOT** be used for actual analysis. These data are provided solely for the purpose of testing statistical package 'code' (e.g. SPSS syntax, SAS programs, etc.) in preparation for analysis using the associated Master File in a Research Data Centre, by Remote Job Submission, or by some other means of secure access. If statistical analysis 'code' works with the synthetic data, researchers can have some confidence that the same code will run successfully against the Master File data in the Research Data Centres."

- **Kind of data <dataKind> 2.2.3.10**
 - Synthetic data
- **Time Method <timeMeth> 2.3.1.2**
 - Do **NOT** enter the "Time Method"
- **Data Collector <dataCollector> 2.3.1.3**
 - Do **NOT** enter the "Data Collector"
- **Sampling Procedure <sampProc> 2.3.1.4**
 - Do **NOT** enter the "Sampling Procedure"
- **Collection mode <collMode> 2.3.1.6**
 - Do **NOT** enter the "Mode of Data Collection"
- **Weighting <weight> 2.3.1.12**
 - Do **NOT** enter the "*Weighting*"
- **Variable Groups <varGrp> 4.1**
 - Variable groups do **NOT** have to be added to the synthetic files

12. APPENDIX D: DEPOSITING RESTRICTED DATA

Access to collections and datasets are controlled through the Permissions function, which uses preset roles. Roles can be assigned at the collection, dataset, or file level. Once assigned, that role applies to all collections and datasets in the hierarchy below the location where the role was assigned. In other words, if a role is assigned at the collection level, that role is applicable for all datasets within that collection. Permissions for Odesi's restricted data collections are assigned at the Dataverse level and are therefore applied to all Datasets in the collection.

Note: Users need to have an Admin role in order to edit and change Dataverse permissions.

- i. After you have created a Dataverse, click the "Edit" drop-down button and select "Permissions".

Contact Share Publish Edit

TEST DV (University of Toronto)
 testdvrestric

- General Information
- Theme + Widgets
- Permissions**
- Groups
- Dataset Templates
- Dataset Guestbooks
- Featured Dataverses

Delete Dataverse

- ii. From here, you can set restrictions for who can add to your Dataverse and which users and groups have access to your Dataverse. By default, users have to request access for depositing in your Dataverse and can freely access files in your Dataverse.

Permissions ^ Current access configuration to your dataverse.

Select if all users or only certain users are able to add to this dataverse, by clicking the Edit Access button. Edit Access

Who can add to this dataverse?
Anyone adding to this dataverse needs to be given access

When a user adds a new dataset to this dataverse, which role should be automatically assigned to them on that dataset?
Contributor - Edit metadata, upload files, and edit files, edit Terms, Guestbook, Submit datasets for review

Users/Groups ^ All the users and groups that have access to your dataverse.

Assign Roles to Users/Groups

5 Users/Groups

User/Group Name (Affiliation) <a>^	ID <a>^	Role <a>^	Action
		Admin	<a>Remove Assigned Role
Dataverse Admin (OTHER)	@dataverseAdmin	Admin	<a>Remove Assigned Role
University of Toronto	&ip/UofTIP	File Downloader	<a>Remove Assigned Role
University of Toronto	&shib/4	Dataverse + Dataset Creator	<a>Remove Assigned Role
UofT ODESI Ezproxy IP	&ip/UofTEzproxyIP	File Downloader	<a>Remove Assigned Role

Roles ^ All the roles set up in your dataverse, that you can assign to users and groups.

- iii. The “Permissions” drop-down menu allows you to choose who can deposit data in your Dataverse. The “Users/Groups” drop down menu allows you to assign roles which give users and groups certain permissions for depositing and accessing data. A glossary of the various roles you can assign is available under the “Roles” drop-down menu.
- iv. For restricted ODESI publications, click the “Assign Roles to Users/Groups” button and assign the “File downloader” role to all the applicable institutions, making sure to select the appropriate access method.

13. APPENDIX E: SPSS & FRENCH CHARACTERS

When importing .sav files into Borealis for French-language studies, or with any special characters, you may experience issues where accented characters do not render properly. This problem started with SPSS 21 when SPSS began supporting Unicode in their files. The work-around for this problem is to turn off Unicode in SPSS. This change cannot be made if you have an open data file.

- a. Open a blank file in SPSS.
- b. Go to “Edit → Options”.
 - i. Click on the “Language” tab.
 - ii. In the “Character Encoding for Data and Syntax” box select “Locale’s writing system”. The Locale can be set for “English” or “OS writing system”.
 - iii. Click Apply, then OK.
- c. Resave your data file with this new setting before importing into Borealis.

