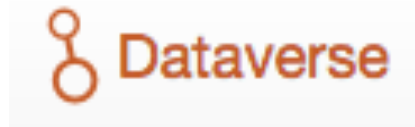**Brief Overview of the**

This semester we will be working with data sets shared by community partners, primarily the City of Boston, and curated by the Boston Area Research Initiative (BARI). They are hosted on BARI's Boston Data Library, which is supported by the Dataverse, a platform built by the Institute for Quantitative Social Science at Harvard University. This document will guide you through the use of the Dataverse, with a focus on the steps necessary to: Create an account; Access data; and Edit and upload datasets and files. (Note that the instructions here are specifically for the activities of our class. If you would like to learn more about using the Dataverse, please visit: http://guides.dataverse.org/en/4.1/user/).

Create an account

Go to https://dataverse.harvard.edu/

Click on the 'Sign Up' button on the top right of the screen and then follow directions.

Access Original Dataset Files

The Dataverse actually contains many sub-Dataverses, which themselves contain additional sub-Dataverses and Datasets. Datasets then contain a collection of files, including those that we would traditionally call "datasets," like an Excel spreadsheet, as well as spatial shapefiles, documentation, and associated R syntax.

The data that we will use in this course are all contained in a Dataverse called "SPPUA 5262 Dataverse." This Dataverse consists of multiple Datasets, each one containing the data and associated documentation needed for a particular project (e.g., Tax Assessor's Data: Cross-Sectional). You will receive a notification e-mail once you have been granted access to the Dataset specific to your group, at which time its contents will become visible through My Data:

Click on your name in the top right corner, then click on 'My Data.' All shared Dataverses and Datasets should be visible.

You will note that your Dataset, like all others used in this class, contains a single data file, documentation file, and R syntax. It is possible, however, to have a variety of configurations, including a Dataset with many related data files, one or more documentation files, etc.

**Edit and Upload New Dataset Files**

As part of the Midterm and Final Projects, collaborative teams will upload updated versions of data and documentation into the Dataverse, based on the work that they have done with the data.

When you are ready to upload new files to your group's dataset you need to do the following:
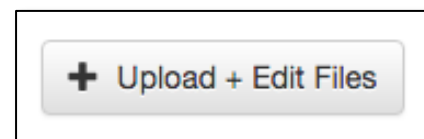
1. Click on the SPPUA 5262 Dataverse.
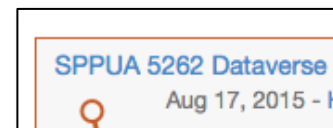
2. Click on your group's dataset.
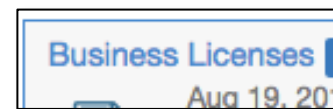
3. Select 'Upload + Edit Files' on the right side.

How to Edit Data in the Dataverse?

Go to your dataset page and click on the "Edit Dataset" button. There, depending on your level of permissions for this dataset, you will have the following options:

- Files (Upload or Edit Data): to add or edit files in this dataset.
- Metadata: to add/edit metadata including additional metadata than was not previously available during Dataset Creation.
- Delete Dataset (only available before your Dataset is published)

You can also directly select the Files, Metadata, Terms or Versions tabs found below the dataset summary information to specifically edit any of those parts of your dataset. You will be unlikely to use most of these options apart from Files.

How to Add Tabular (spreadsheet) Data to a Dataverse?

1. Follow directions above, then scroll down to the "Files" section and click on "Select Files to Add" to add all the relevant files to your Dataset. **Tip:** You can drag and drop or select multiple files at a time from your desktop, directly into the upload widget. Your files will appear below the "Select Files to Add" button where you can add a description and tags (via the "Edit Tag" button) for each file.
2. Click the "Add Dataset" button when you are done. Your unpublished dataset is now created.

How to Add Spatial Data?

Later in the course we will be learning about geospatial (or mappable) data in the form of shapefiles. These can be uploaded into the Dataverse and further explored and manipulated through its integration with WorldMap, a geospatial data visualization and analysis tool developed by the Center for Geographic Analysis at Harvard University. Once you publish your dataset with your shape files, you will be able to visualize your data using the "Map Data" button in Dataverse. *Note*: In order to map your data file, a copy will be sent to Harvard's WorldMap platform. You have the ability to delete any maps, and associated data, from the Harvard WorldMap platform, at any time.