



## 1.15 Exercise: Import data into iNZight Lite

This exercise will enable you to import data into iNZight Lite from a variety of sources. You will learn how to:

- Load data stored within iNZight Lite.
- Import a data file stored outside iNZight Lite. (Optional)
- Create and import your own data file. (Optional)

Follow these instructions to import data into iNZight Lite. If you have any difficulty completing this exercise, you may find the common questions on page 5 to be useful.

[The [getting started](#) video linked to the previous exercise supplemented and a new (40 sec) video about [finding the "Data to Insight" data sets](#) built into iNZight Lite covers this material.]

## Load data stored within iNZight Lite

Open iNZight Lite on your computer and go to **File > DataSet Examples**. Click on the drop down menu: **Select Data set category** and choose **Future-Learn**.

Select a data file, try **NHANES-1000**.

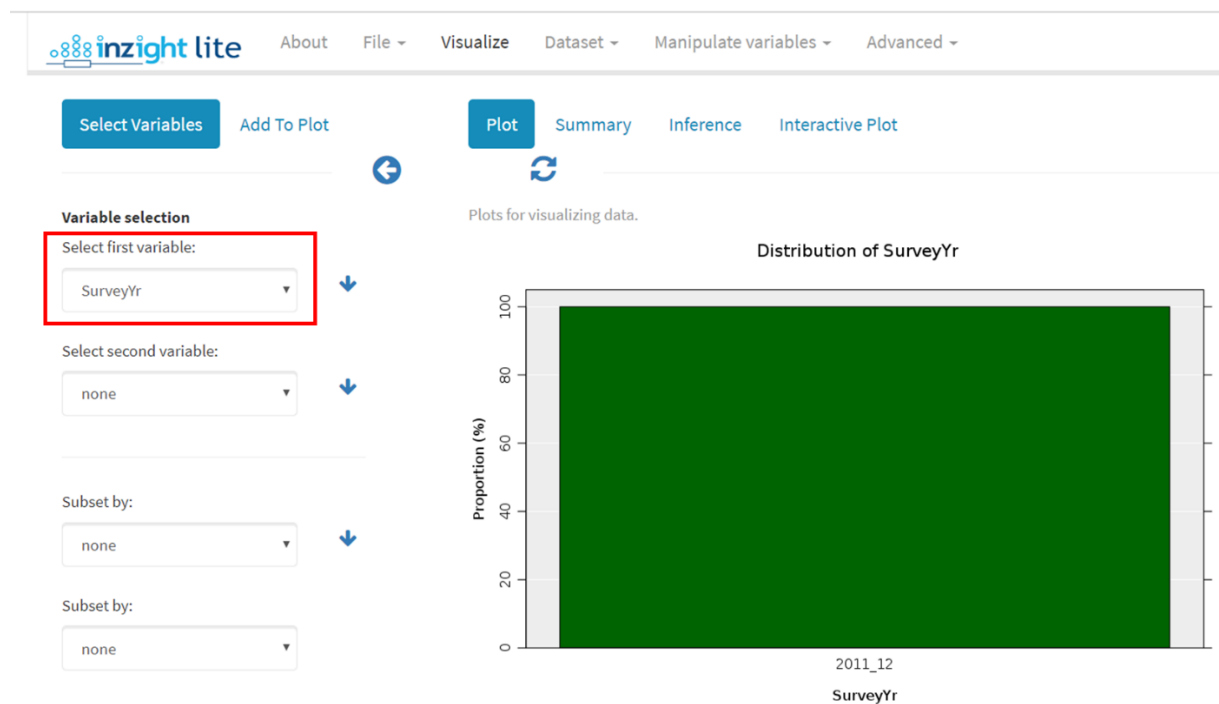
Then click on **Select Set** (*very important!*) to start using the data.

The screenshot shows the iNZight Lite web interface. The top navigation bar includes 'About', 'File', 'Visualize', 'Dataset', 'Manipulate variables', and 'Advanced'. The 'File' menu is open, showing 'Import Dataset', 'Export Dataset', 'Display Dataset', 'Remove Dataset', and 'DataSet Examples'. The 'DataSet Examples' dropdown is open, showing a 'Select Data set category' dropdown menu with 'Future-Learn' selected. Below this, a list of datasets is shown, with 'NHANES-1000' selected. A 'SELECT SET' button is highlighted at the bottom of the list. On the right side, the 'Selected data set: NHANES-1000' is displayed, along with the number of rows (1000) and columns (41). A list of column names is provided. Below this, the 'Data: NHANES-1000' section shows a table with 5 entries. The table has columns for SurveyYr, ID, Gender, Age, AgeDecade, Race3, Education, and MaritalStatus. The first row shows a female aged 5 in the 0-9 age decade, Black race, with no education listed. The second row shows a female aged 49 in the 40-49 age decade, White race, with CollegeGrad education and Married marital status. The third row shows a male aged 37 in the 30-39 age decade, White race, with SomeCollege education and Married marital status. The fourth row shows a female aged 37 in the 30-39 age decade, White race, with 9\_11thGrade education and Married marital status.

SurveyYr	ID	Gender	Age	AgeDecade	Race3	Education	Mar
2011_12	63549	female	5	0-9	Black		
2011_12	64871	female	49	40-49	White	CollegeGrad	Marri
2011_12	70551	male	37	30-39	White	SomeCollege	Marri
2011_12	66033	female	37	30-39	White	9_11thGrade	Marri

The data should be showing in the right hand side of your window. The top row shows the names of each variable and the data on that variable is in the column below. Scroll right to see all the variables. Click on the variable name to sort the data by that variable. Using the drop down at the top of the data you can view more data in each snapshot: 5, 30 or 50 at a time.

Now you have data, try to do something with it. Click on the **Visualize** tab at the top middle of the page. You should see this ...



Use **Select first variable** to select the variable **AgeDecade** and see what happens. Similarly select other variables as **first variable**.

Click inside the **Select first variable** box and use the up and down arrow keys and see what happens.

Select **Weight** as first variable. Now use **Select second variable** and choose some variables of interest to you. Then play with some other things. iNZight Lite is a program that really rewards just trying things out and seeing what happens.

To load a new data set for this course, just click on one of the other FutureLearn datasets on the left-hand side and click on **SELECT SET** again.

## PRACTICE

Load the **Gapminder-2008** data set into iNZight Lite. Scroll so that you can see more of the data. Try clicking on variable names to sort the data.

## Import a data file from outside iNZight Lite

Download the file [olympics100m.csv](#) and place it on your Desktop.

[This file contains data on the winning times for the 100 metre running race at the Olympics from 1896 to 2016.]

Now go:

- **File > Import Dataset**
- Press the **Browse** button and using the file browser that appears to navigate to **olympics100m** on your desktop and open the file. The data will get imported.
  - It is this easy because **olympics100m.csv** is in a file type format that iNZight Lite recognises (csv or "comma separated values")
- When you see it, click on the **Visualize** tab and explore this data as you did with the **NHANES-1000** data above.

## Create and import data (optional)

iNZight Lite expects **rectangular data** in which the first row of cells contains the **names of variables** and everything in the column beneath contains **data recorded on that variable**.

Below is some data for 10 people from the New Zealand **CensusAtSchool** database.

- **Open up a spreadsheet program** such as Excel and type in some of the data illustrated below (maybe the first 5 rows of the gender, age and height columns).

	A	B	C	D	E	F	G
1	gender	age	height	armspan	memory	reaction	region
2	female	12	165	173	44	0.328	Auckland Region
3	male	13	156	147	41	0.344	Otago Region
4	female	15	162	158	34	0.345	Bay of Plenty Region
5	female	13	176	176	31	0.387	Auckland Region
6	female	13	162	156	62	0.54	Wellington Region
7	male	12	144	144	59	3.597	Auckland Region
8	female	12	146	144	33	0.459	Auckland Region
9	male	11	153	151	75	0.948	Other North Island
10	female	15	163	160	47	0.544	Auckland Region
11	female	7	123	123	56	0.54	Other North Island
12							

- **Save your file on your Desktop:** Use **Save As**, type in some suitable name for your file and for **File Type** choose **Text (tab-delimited)** (it will then give your file .txt extension on its name).
- Now go **File > Import data...** and continue as above to import your new little data set into iNZight.

Both **csv** (comma-separated values) and **tab-delimited text** are common data formats that iNZight Lite automatically recognises from the filename extensions. It similarly imports **Excel** files (.xlsx) containing a single worksheet and a few other formats.

**Tab-delimited text** is preferable to **csv** if you are in continental Europe or any other country that uses "," as the decimal separator rather than "." (e.g. 0,57 instead of 0.57).

## Common questions

### *My data won't import at all*

Try opening your data file in Excel (or similar) and save it in **Text (tab-delimited)** format. When you go to import the data you should see a .txt extension on the end of its name..

### *My data still won't load*

Check your data set for spaces. If you are using csv files, ensure that your column names do not have commas in them.

Check that your numeric variables do not have "O" s instead of "0"s.

Try loading just the necessary rows and columns of your data by cutting and pasting them into another spreadsheet and saving as a new file.

### **Still have problems?**

iNZight's [FAQ page on importing data](#) may be useful.

Read the article called [Using your own data](#).