

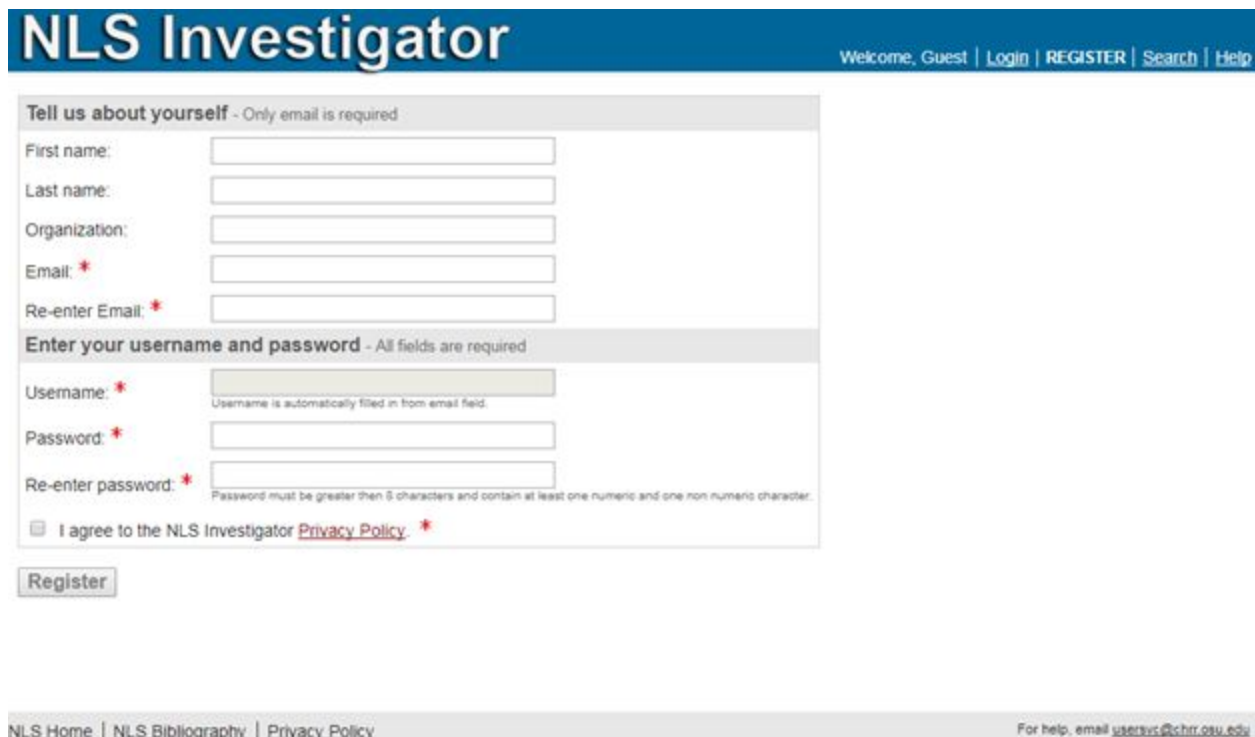
Instructions for Part A

This first part of the assignment we will work with a data set from the National Longitudinal Survey of Youth (NLSY). You may remember you used some data from the 1997 version of this survey in an earlier assignment. This time we will be working with data started in 1979 and continued through 2014. In this first part of the assignment we will:

1. Create an account to access the data
2. Download data files from the NLSY website on the subjects physical characteristics.
3. Download appropriate variable names from the NLSY website.

Create an account

1. Create an account on the NLS Investigator at
<https://www.nlsinfo.org/investigator/pages/register.jsp>



NLS Investigator Welcome, Guest | [Login](#) | [REGISTER](#) | [Search](#) | [Help](#)

Tell us about yourself - Only email is required

First name:

Last name:

Organization:

Email: *

Re-enter Email: *

Enter your username and password - All fields are required

Username: * Username is automatically filled in from email field.

Password: *

Re-enter password: * Password must be greater than 5 characters and contain at least one numeric and one non numeric character.

☐ I agree to the NLS Investigator [Privacy Policy](#). *

[NLS Home](#) | [NLS Bibliography](#) | [Privacy Policy](#) For help, email usersvc@chr.osu.edu

You will be directed to verify your account through the email address that you supplied. Follow the instructions to do so.

Download the data

2. Once you have verified your account, log in. After logging in, select the study NLSY79 (1979-2014)

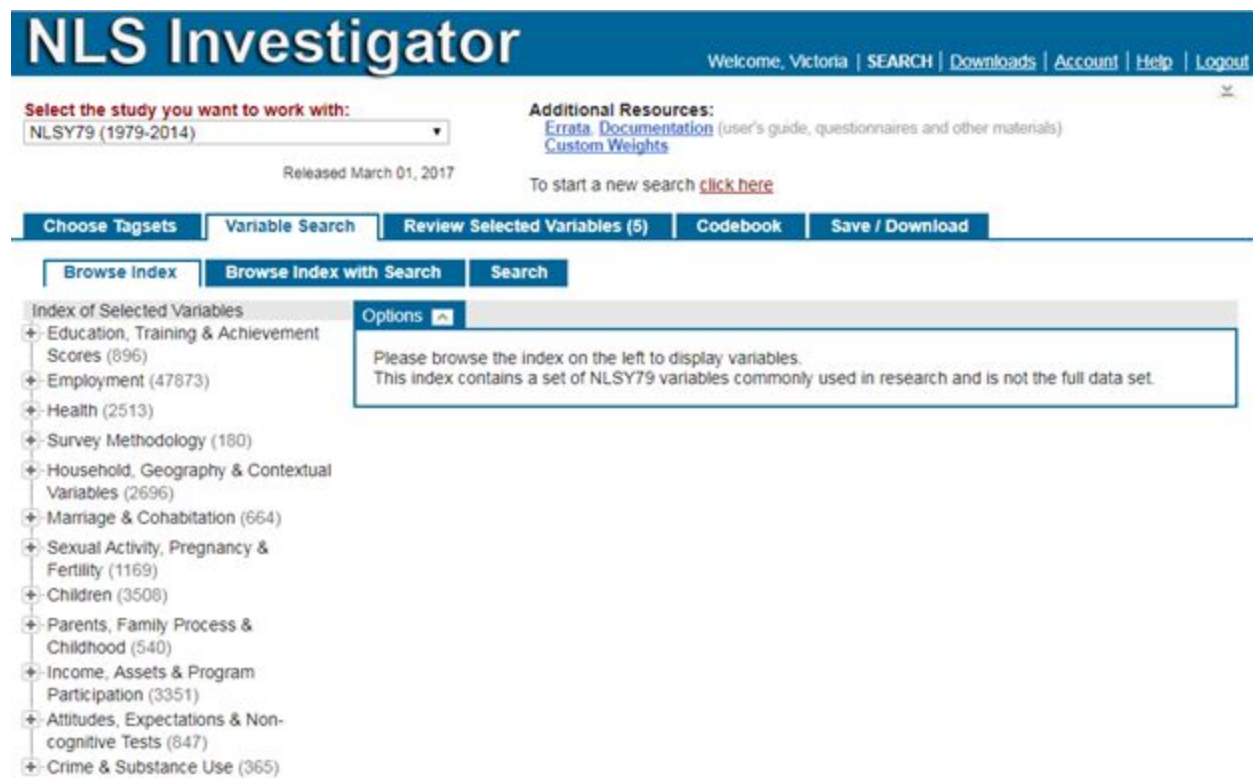


NLS Investigator Welcome, Victoria | [SEARCH](#) | [Downloads](#) | [Account](#) | [Help](#) | [Logout](#)

Select the study you want to work with:

- (Choose One)
- (Choose One)
- NLSY79 (1979-2014)**
- NLSY79 Child/Young Adults 1986-2014
- NLSY97 1997-2013 (rounds 1-16)
- Original Cohorts - Mature Women & Young Women
- Original Cohorts - Older Men & Young Men

3. In the search page that comes up, select the “Variable Search” tab, and then the “Browse Index” tab.



NLS Investigator Welcome, Victoria | [SEARCH](#) | [Downloads](#) | [Account](#) | [Help](#) | [Logout](#)

Select the study you want to work with: NLSY79 (1979-2014)

Released March 01, 2017

Additional Resources: [Errata](#), [Documentation](#) (user's guide, questionnaires and other materials), [Custom Weights](#)

To start a new search [click here](#)

Choose Tagsets | **Variable Search** | **Review Selected Variables (5)** | **Codebook** | **Save / Download**

Browse Index | **Browse Index with Search** | **Search**

Index of Selected Variables

- Education, Training & Achievement Scores (896)
- Employment (47873)
- Health (2513)
- Survey Methodology (180)
- Household, Geography & Contextual Variables (2696)
- Marriage & Cohabitation (664)
- Sexual Activity, Pregnancy & Fertility (1169)
- Children (3508)
- Parents, Family Process & Childhood (540)
- Income, Assets & Program Participation (3351)
- Attitudes, Expectations & Non-cognitive Tests (847)
- Crime & Substance Use (365)

Options

Please browse the index on the left to display variables.
This index contains a set of NLSY79 variables commonly used in research and is not the full data set.

4. In the expandable list of categories titled “Index of Selected Variables”, drill down on “Health”, then “Physical Characteristics”, revealing 3 categories, “Height”, “Weight” and “Eye & hair color”.

Choose Tagsets | **Variable Search**

Browse Index | **Browse Index w**

Index of Selected Variables

- + Education, Training & Achievement Scores (896)
- + Employment (47873)
- Health (2513)
 - Physical Characteristics (35)
 - Height (13)
 - Weight (20)
 - Eye & hair color (2)
 - + Health conditions (227)
 - + Mental health (51)
 - + Health care usage (252)
 - + Health Insurance (700)
 - + Health behaviors (256)
 - + Work-related health (892)
 - + Cognition (115)

- Select the “**height**” item. Under Options to the right, a list of variables with descriptions should appear. However, the list of variables reveals variables with multiple “Question Name” codes for different years. Early in the study, height was recorded as a single variable, while later in the study feet and inches were recorded as separate variables. We are largely interested in adult height, so to simplify the data, use the check boxes to just select all items with a Question Name Q11-10 A or Q11-10 B.

Choose Tagsets | Variable Search | Review Selected Variables (15) | Codebook | Save / Download

Browse Index | Browse Index with Search | Search

Index of Selected Variables

- Education, Training & Achievement Scores (896)
- Employment (47873)
- Health (2513)
 - Physical Characteristics (35)
 - Height (13)**
 - Weight (20)
 - Eye & hair color (2)
 - Health conditions (227)
 - Mental health (51)
 - Health care usage (252)
 - Health Insurance (700)
 - Health behaviors (256)
 - Work-related health (892)
 - Cognition (115)
- Survey Methodology (180)
- Household, Geography & Contextual Variables (2696)
- Marriage & Cohabitation (664)

Options Showing 13 of 13 filtered by All Variables 500 1

	RNUM	QUESTION NAME	VARIABLE TITLE	YEAR
1	<input type="checkbox"/> R0481600	HEALTH-HEIGHT	HEIGHT OF R	1981
2	<input type="checkbox"/> R0779800	HEALTH-HEIGHT	HEIGHT OF R	1982
3	<input type="checkbox"/> R1773900	HEALTH-HEIGHT	HEIGHT OF R	1985
4	<input checked="" type="checkbox"/> T0897400	Q11-10_A	R HEIGHT IN FEET	2006
5	<input checked="" type="checkbox"/> T0897500	Q11-10_B	R HEIGHT IN INCHES	2006
6	<input checked="" type="checkbox"/> T2053900	Q11-10_A	R HEIGHT IN FEET	2008
7	<input checked="" type="checkbox"/> T2054000	Q11-10_B	R HEIGHT IN INCHES	2008
8	<input checked="" type="checkbox"/> T3024800	Q11-10_A	R HEIGHT IN FEET	2010
9	<input checked="" type="checkbox"/> T3024900	Q11-10_B	R HEIGHT IN INCHES	2010
10	<input checked="" type="checkbox"/> T3955100	Q11-10_A	R HEIGHT IN FEET	2012
11	<input checked="" type="checkbox"/> T3955200	Q11-10_B	R HEIGHT IN INCHES	2012
12	<input checked="" type="checkbox"/> T4892800	Q11-10_A	R HEIGHT IN FEET	2014
13	<input checked="" type="checkbox"/> T4892900	Q11-10_B	R HEIGHT IN INCHES	2014

500 1

- Now select the “Save/Download” tab. This should reveal a tab “Save Tagset”; in the radio buttons that appear, select “Save to PC” and “By Rnum”, but **DO NOT** select “Save”

Choose Tagsets | Variable Search | Review Selected Variables (15) | Codebook | Save / Download

Save Tagset | Basic Download | Advanced Download | Manage Downloads

Choose where to save the tagset of your selected variables:

☒ Save to PC
☐ Save on our server

Tagset Type: ☒ By Rnum ☐ By Qname with Year

Filename:

Filename must only contain alpha, numeric, hyphen or underscore characters.

Tagsets saved on your PC or our server can be reloaded at a later time through the 'Choose Tagsets' tab.
Tagsets stored on our server with no activity may be deleted after 90 days.

7. Select the “Advanced Download” and set the options to look like this:

Save Tagset
Basic Download
Advanced Download
Manage Downloads

Customize your advanced download:

Create Download of Data

- ☒ Tagset (list of selected variables)
- ☐ SAS® control file (includes the datafile of selected variables)
- ☐ SPSS® control file (includes the datafile of selected variables)
- ☐ STATA® dictionary file of selected variables
- ☐ R® Source code (includes the datafile of selected variables)
- ☒ Codebook of selected variables
- ☐ Short Description File
- ☒ Comma-delimited datafile of selected variables (to be read in Excel, etc.)
Column headers -- Use ☒ Reference Number ☐ Question Name (does not guarantee uniqueness)

Create Frequency / Table

- ☐ Apply Universe Restrictors ([How to use Universe Restrictors](#))
- ☐ Notify me by email when download is complete.

Filename:

Filename must only contain alpha, numeric, hyphen or underscore characters.

Download status appears under 'Manage Downloads' tab.
Downloads may be deleted after 10 days of inactivity.

8. Enter an appropriate filename such as Project1data for your data and click “Download”.

9. After processing on the server, a listing with your file is presented and you should click the download hyperlink. This will download the data as a CSV file to an appropriate location on your local machine.

Choose Tagsets
Variable Search
Review Selected Variables (15)
Codebook
Save / Download

Save Tagset
Basic Download
Advanced Download
Manage Downloads

Download Status:

All downloads are available. Please click a download link below to begin downloading.

All Available Downloads:

	Date	Study	Name	Size	Download
1	2017-09-24 11:18:07	NLSY79	project1data	72.5K	download

Download appropriate variable names

10. In the tab set above, select “Review Selected Variables (15)”, revealing a list of 5 basic variables plus the 10 you added for height.

Choose Tagsets | Variable Search | Review Selected Variables (15) | Codebook | Save / Download

Options

500

	<input checked="" type="checkbox"/>	RNUM ▲	QUESTION NAME	VARIABLE TITLE	YEAR
1	<input checked="" type="checkbox"/>	A0002600	VERSION_R26	VERSION_R26_1	2014
2	<input checked="" type="checkbox"/>	R0000100	CASEID	IDENTIFICATION CODE	1979
3	<input checked="" type="checkbox"/>	R0173600	SAMPLE_ID	SAMPLE IDENTIFICATION CODE	1979
4	<input checked="" type="checkbox"/>	R0214700	SAMPLE_RACE	R'S RACIAL/ETHNIC COHORT FROM SCREENER	78SCRN
5	<input checked="" type="checkbox"/>	R0214800	SAMPLE_SEX	SEX OF R	1979
6	<input checked="" type="checkbox"/>	T0897400	Q11-10_A	R HEIGHT IN FEET	2006
7	<input checked="" type="checkbox"/>	T0897500	Q11-10_B	R HEIGHT IN INCHES	2006
8	<input checked="" type="checkbox"/>	T2053900	Q11-10_A	R HEIGHT IN FEET	2008
9	<input checked="" type="checkbox"/>	T2054000	Q11-10_B	R HEIGHT IN INCHES	2008
10	<input checked="" type="checkbox"/>	T3024800	Q11-10_A	R HEIGHT IN FEET	2010
11	<input checked="" type="checkbox"/>	T3024900	Q11-10_B	R HEIGHT IN INCHES	2010
12	<input checked="" type="checkbox"/>	T3955100	Q11-10_A	R HEIGHT IN FEET	2012
13	<input checked="" type="checkbox"/>	T3955200	Q11-10_B	R HEIGHT IN INCHES	2012
14	<input checked="" type="checkbox"/>	T4892800	Q11-10_A	R HEIGHT IN FEET	2014
15	<input checked="" type="checkbox"/>	T4892900	Q11-10_B	R HEIGHT IN INCHES	2014

11. Use your mouse to highlight or select the entries in the table, copy, and paste into a new Excel document. Save this to disk as a CSV file with an appropriate name such as “height_index.csv”. This file will be used to name the variables in R.