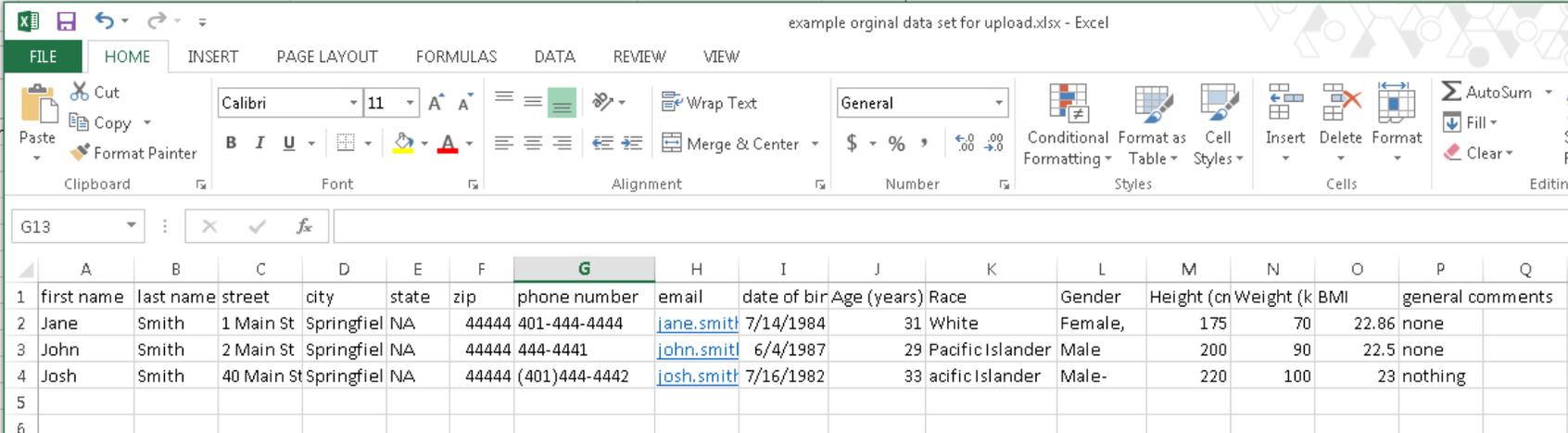


## Existing Database Upload:

Do you have an already existing database in excel that would like to upload to REDCap? We have developed some resource material to help you accomplish this. The below steps are meant for a user new to REDCap. Alternative approaches are possible using the data dictionary ([Penn State link](#))

Below is an example of a data set in excel with some common variables you might want to import. Some are categorical (female vs male), some are continuous (BMI, height), and some may require a special format (date of birth, email, phone number).



|   | A          | B         | C          | D           | E     | F     | G             | H  | I             | J           | K                | L       | M           | N           | O     | P                | Q |
|---|------------|-----------|------------|-------------|-------|-------|---------------|--|---------------|-------------|------------------|---------|-------------|-------------|-------|------------------|---|
| 1 | first name | last name | street     | city        | state | zip   | phone number  | email                                    | date of birth | Age (years) | Race             | Gender  | Height (cm) | Weight (kg) | BMI   | general comments |   |
| 2 | Jane       | Smith     | 1 Main St  | Springfield | NA    | 44444 | 401-444-4444  | <a href="mailto:jane.smit">jane.smit</a> | 7/14/1984     | 31          | White            | Female, | 175         | 70          | 22.86 | none             |   |
| 3 | John       | Smith     | 2 Main St  | Springfield | NA    | 44444 | 444-4441      | <a href="mailto:john.smit">john.smit</a> | 6/4/1987      | 29          | Pacific Islander | Male    | 200         | 90          | 22.5  | none             |   |
| 4 | Josh       | Smith     | 40 Main St | Springfield | NA    | 44444 | (401)444-4442 | <a href="mailto:josh.smit">josh.smit</a> | 7/16/1982     | 33          | acific Islander  | Male-   | 220         | 100         | 23    | nothing          |   |
| 5 |            |           |            |             |       |       |               |  |               |             |                  |         |             |             |       |                  |   |
| 6 |            |           |            |             |       |       |               |  |               |             |                  |         |             |             |       |                  |   |

- 1) First create a data form (also called instrument) in REDCap that lists the variables in the data set you are trying to import. Ideally, the form should list the variables in the same order of your data set. Some important points:
  - a. Variable names should be kept as simple as possible, while still being descriptive.
  - b. Variable names that are composed of more than one word must include an underscore between words (i.e. first name should appear first\_name).
  - c. Formats for export and import in REDCap are CSV (comma separated values), it is critical to avoid the use of commas, when creating your variables and database for upload.
  - d. **Do not include calculated fields as part of your form if you are trying to upload a database.**
  - e. Below are examples of how to create some common variable types.

### Edit Field ✕

You may add a new project field to this data collection instrument by completing the fields below and clicking the Save button at the bottom. When you add a new field, it will be added to the form on this page. For an overview of the different field types available, you may view the [Field Types video \(4 min\)](#).

**Field Type:** Text Box (Short Text, Number, Date/Time, ...)

**Field Label** [How to use Piping](#)

First Name

**Field Annotation** (optional) [Learn about Action Tags](#)

Explanatory notes - not displayed on any page ?

**Variable Name** (utilized during data export)

Enable auto naming of variable based upon its Field Label?

ONLY letters, numbers, and underscores

**Validation?** (optional) ---- None ----

- or -

**Enable searching within a biomedical ontology** ?

-- choose ontology to search --

**Required?\***  No  Yes

\* Prompt if field is blank

**Identifier?**  No  Yes

Does the field contain identifying information (e.g., name, SSN, address)?

**Custom Alignment** Right / Vertical (RV)

Align the position of the field on the page

**Field Note** (optional)

Small reminder text displayed underneath field

Save
Cancel

Be careful to indicate when a variable is an identifier (example name, address), this will help with data export and PHI requirements later.

### Edit Field ✕

You may add a new project field to this data collection instrument by completing the fields below and clicking the Save button at the bottom. When you add a new field, it will be added to the form on this page. For an overview of the different field types available, you may view the [Field Types video \(4 min\)](#).

**Field Type:** Multiple Choice - Radio Buttons (Single Answer) ▾

**Field Label** [How to use Piping](#)

Gender

**Choices (one choice per line)** [Copy existing choices](#)

0, Female  
1, Male

[How do I manually code the choices?](#)

**Field Annotation** (optional) [Learn about Action Tags](#)

Explanatory notes - not displayed on any page ?

**Variable Name** (utilized during data export)

sex  Enable auto naming of variable based upon its Field Label?

**Required?\***  No  Yes  
\* Prompt if field is blank

**Identifier?**  No  Yes  
Does the field contain identifying information (e.g., name, SSN, address)?

**Custom Alignment** Right / Vertical (RV) ▾  
Align the position of the field on the page

**Field Note** (optional)

Small reminder text displayed underneath field

**Save** **Cancel**

For categorical variables like gender, it is best to start the count at 0 and increment up with choices (ie, 0, 1, 2, 3) rather than starting with 1 (i.e. 1, 2, 3, 4). Starting at 0 will make analyzing data easier later.

### Edit Field ✕

You may add a new project field to this data collection instrument by completing the fields below and clicking the Save button at the bottom. When you add a new field, it will be added to the form on this page. For an overview of the different field types available, you may view the [Field Types video \(4 min\)](#).

**Field Type:** Text Box (Short Text, Number, Date/Time, ...)

**Field Label** [How to use Piping](#)

E-mail

**Field Annotation** (optional) [Learn about Action Tags](#)

Explanatory notes - not displayed on any page ?

**Variable Name** (utilized during data export)

email

ONLY letters, numbers, and underscores  Enable auto naming of variable based upon its Field Label?

**Validation?** (optional) Email

– or –

**Enable searching within a biomedical ontology** ?

-- choose ontology to search --

**Required?\***  No  Yes  
\* Prompt if field is blank

**Identifier?**  No  Yes  
Does the field contain identifying information (e.g., name, SSN, address)?

**Custom Alignment** Right / Vertical (RV)  
Align the position of the field on the page

**Field Note** (optional)   
Small reminder text displayed underneath field

Save
Cancel

When possible include validations (email, telephone, date...). This is especially helpful for quality control with large datasets. This helps verify you have the appropriate values in the appropriate places. If all your variable fields were simply text you could mistakenly upload email values to the gender category.

View equation

### Edit Field

You may add a new project field to this data collection instrument by completing the fields below and clicking the Save button at the bottom. When you add a new field, it will be added to the form on this page. For an overview of the different field types available, you may view the [Field Types video \(4 min\)](#).

**Field Type:** Text Box (Short Text, Number, Date/Time, ...)

**Field Label** [How to use Piping](#)

Weight (kilograms)

**Field Annotation** (optional) [Learn about Action Tags](#)

Explanatory notes - not displayed on any page ?

**Variable Name** (utilized during data export)

weight  Enable auto naming of variable based upon its Field Label?

ONLY letters, numbers, and underscores

**Validation?** (optional) Integer

**Minimum:** 35

**Maximum:** 200

– OR –

**Enable searching within a biomedical ontology** ?

-- choose ontology to search --

**Required?\***  No  Yes

\* Prompt if field is blank

**Identifier?**  No  Yes

Does the field contain identifying information (e.g., name, SSN, address)?

**Custom Alignment** Right / Vertical (RV)

Align the position of the field on the page

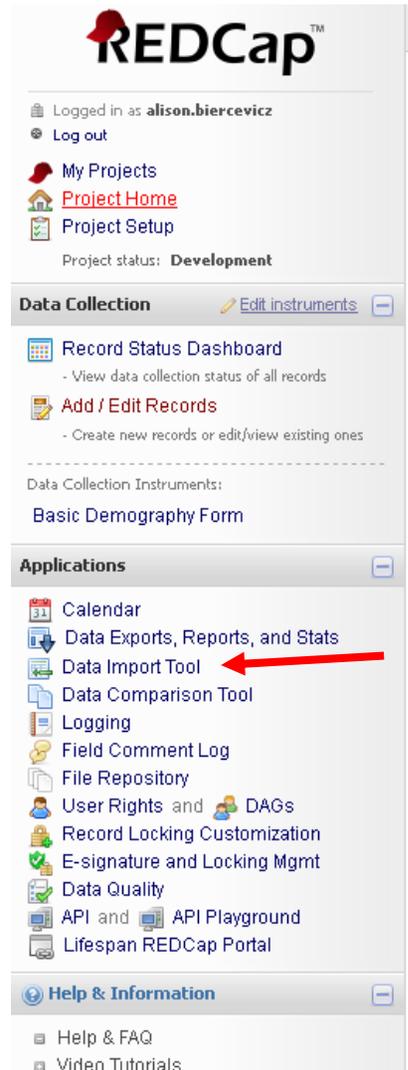
**Field Note** (optional)

Small reminder text displayed underneath field

**Save** **Cancel**

For continuous variables use the text box field type. It is good practice to include validation for numbers you have a good idea of an appropriate range, to avoid inputting incorrect data.

- 2) Once you have all of your variables entered in the form you can export your data import template to fill in your data. Under the Applications tool bar is Data Import Tool.



- 3) Now notice the difference between the data import template (top) and your original data set (bottom). The variable names are different. There is also a record\_id variable inserted in the first column and a demographics\_complete variable in the last column.

The top screenshot shows an Excel spreadsheet with the following data:

|   | A         | B          | C         | D       | E         | F     | G   | H    | I   | J      | K      | L        | M                     | N | O | P |
|---|-----------|------------|-----------|---------|-----------|-------|-----|------|-----|--------|--------|----------|-----------------------|---|---|---|
| 1 | record_id | first_name | last_name | address | telephone | email | dob | race | sex | height | weight | comments | demographics_complete |   |   |   |
| 2 |           |            |           |         |           |       |     |      |     |        |        |          |                       |   |   |   |

The bottom screenshot shows an Excel spreadsheet titled 'example original data set for upload.xlsx' with the following data:

|   | A          | B         | C          | D           | E     | F     | G             | H          | I             | J           | K                | L      | M           | N           | O     | P             |
|---|------------|-----------|------------|-------------|-------|-------|---------------|------------|---------------|-------------|------------------|--------|-------------|-------------|-------|---------------|
| 1 | first name | last name | street     | city        | state | zip   | phone number  | email      | date of birth | Age (years) | Race             | Gender | Height (cm) | Weight (kg) | BMI   | general color |
| 2 | Jane       | Smith     | 1 Main St  | Springfield | NA    | 44444 | 401-444-4444  | jane.smith | 7/14/1984     | 31          | White            | Female | 175         | 70          | 22.86 | none          |
| 3 | John       | Smith     | 2 Main St  | Springfield | NA    | 44444 | 444-4441      | john.smith | 6/4/1987      | 29          | Pacific Islander | Male   | 200         | 90          | 22.5  | none          |
| 4 | Josh       | Smith     | 40 Main St | Springfield | NA    | 44444 | (401)444-4442 | josh.smith | 7/16/1982     | 33          | Pacific Islander | Male   | 220         | 100         | 23    | nothing       |

You can now move your original data set to fill in the template for the appropriate variables in excel.

- 4) **IMPORTANT: Your formatted variables (i.e. date, phone number, email) and your categorical variables must exactly match the way you defined them in the form.**
- 5) For example in my original form I had address broken up into individual columns for street, city, state and zip. In the form I had originally defined address as a single text box. I will need to update the form to include this expanded form of the address to import the data correctly.
- 6) Now notice I have 3 different formats for phone number in my original dataset. These will have to be updated to be in the identical format defined in my form (US Format is the default).

| F    | G             | H                          | I             | J           | K                | L       | M      |
|------|---------------|----------------------------|---------------|-------------|------------------|---------|--------|
|      | phone number  | email                      | date of birth | Age (years) | Race             | Gender  | Height |
| 1444 | 401-444-4444  | <a href="#">jane.smith</a> | 7/14/1984     | 31          | White            | Female, |        |
| 1444 | 444-4441      | <a href="#">john.smith</a> | 6/4/1987      | 29          | Pacific Islander | Male    |        |
| 1444 | (401)444-4442 | <a href="#">josh.smith</a> | 7/16/1982     | 33          | Pacific Islander | Male-   |        |
|      |               |                            |               |             |                  |         |        |
|      |               |                            |               |             |                  |         |        |

7) **CATEGORICAL VARIABLES:** I also have race for one of my data points spelled incorrectly. Further, the way Pacific Island is defined is not identical to how it is defined in my form. If these were formatted variables (telephone and email) we would be concerned; however, how we treat categorical variables when importing is different than the formatted variables such as telephone and email. When importing categorical variables into REDCap the variable must reflect the number it was given when you defined it in the data form. I.E. Female will appear as "0", Male will appear as "1", White will appear as "4" and Pacific Islander will appear as "2" see the example below

|                    | J        | K   | L    | M   | N      | O      |     |
|--------------------|----------|-----|------|-----|--------|--------|-----|
|                    | dob      | age | race | sex | height | weight | bmi |
| <a href="#">it</a> | #####    | 31  | 4    | 0   | 175    | 70     |     |
| <a href="#">it</a> | 6/4/1987 | 29  | 2    | 1   | 200    | 90     |     |
| <a href="#">it</a> | #####    | 33  | 2    | 1   | 220    | 100    |     |
|                    |          |     |      |     |        |        |     |
|                    |          |     |      |     |        |        |     |

8) **REMINDER: Commas should not be included in any variable entries.** The format for the data import and export are CSV (comma separated values) and an extra comma in a data point will throw off the data upload.



Now you are ready for data upload. This is done in the same screen that the data template was downloaded from (Data Import Tool)

The screenshot displays the REDCap user interface. At the top, the REDCap logo is visible. Below it, the user is logged in as 'alison.biercevicz' with a 'Log out' link. Navigation links include 'My Projects', 'Project Home', and 'Project Setup'. The project status is 'Development'. The 'Data Collection' section includes 'Record Status Dashboard' and 'Add / Edit Records'. Under 'Data Collection Instruments', 'Basic Demography Form' is listed. The 'Applications' section contains a list of tools: 'Calendar', 'Data Exports, Reports, and Stats', 'Data Import Tool' (highlighted with a red arrow), 'Data Comparison Tool', 'Logging', 'Field Comment Log', 'File Repository', 'User Rights and DAGs', 'Record Locking Customization', 'E-signature and Locking Mgmt', 'Data Quality', 'API and API Playground', and 'Lifespan REDCap Portal'. The 'Help & Information' section at the bottom includes 'Help & FAQ' and 'Video Tutorials'.