Table of Contents

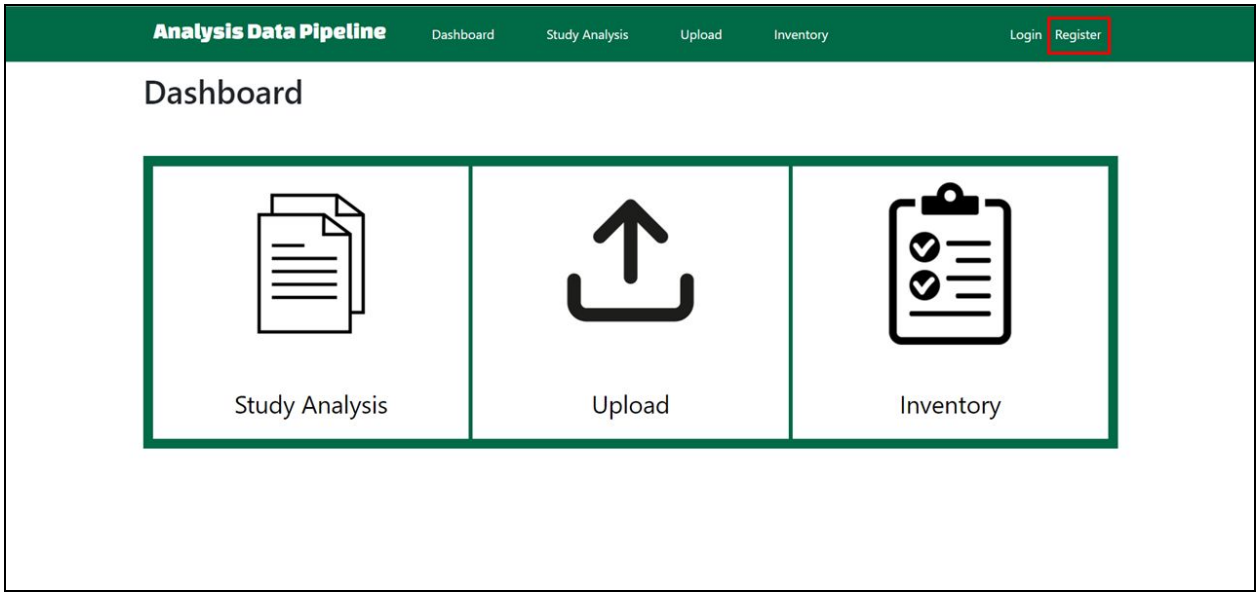
1. User System

This application allows users to create their own accounts. This is required for uploading study data, as the account is linked to the study so only verified users can upload to it. The study analysis and inventory analysis features do not require you to be logged in. However, you can only view studies that have a visibility of “Public”. Studies with the visibility of “Private” are only visible to their owners.

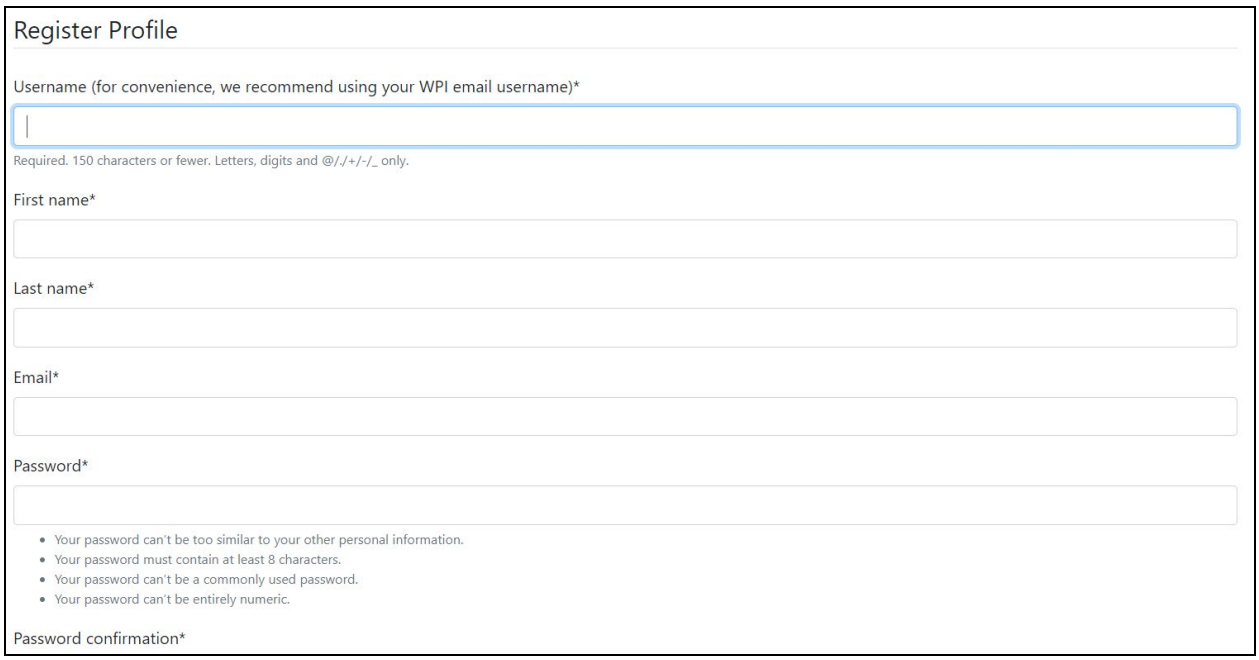
1.1 Register Profile and Login

1.1.1 Register

To register a profile click the “Register” text on the top right of the nav bar. Note this will only show on the nav bar if you are not logged in.

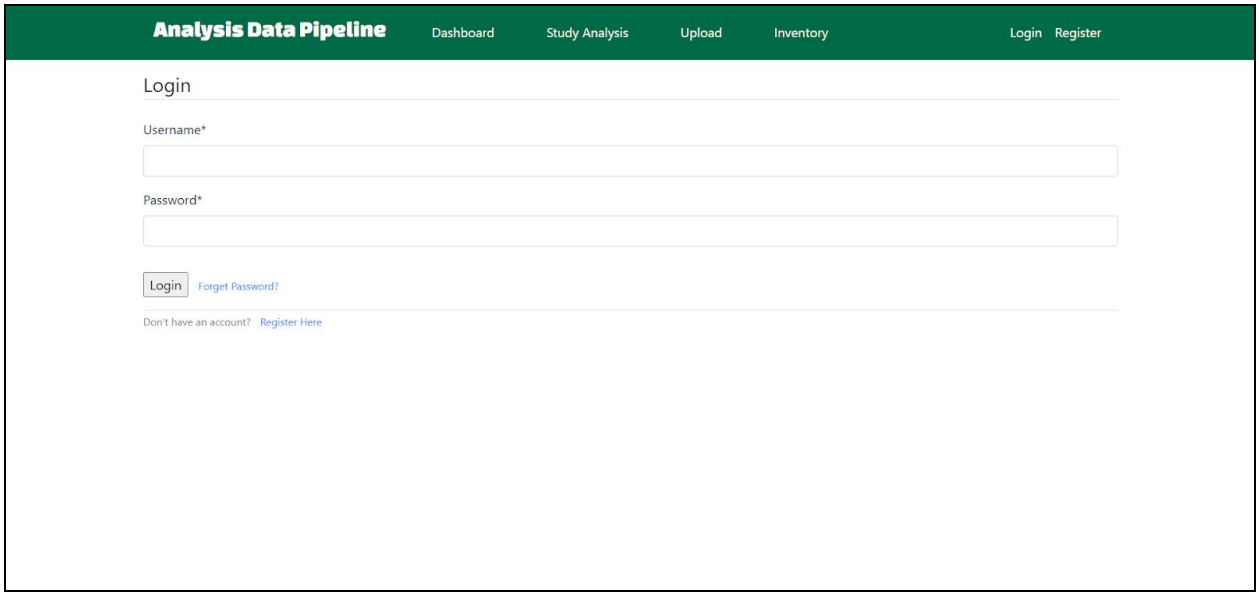


This will bring you to a new page with input fields for your account. Please enter a username, first name, last name, email address, and password. Then, click register to create your profile. Note that the password has specific requirements to ensure that it is secure. Profiles must have unique usernames and email addresses.



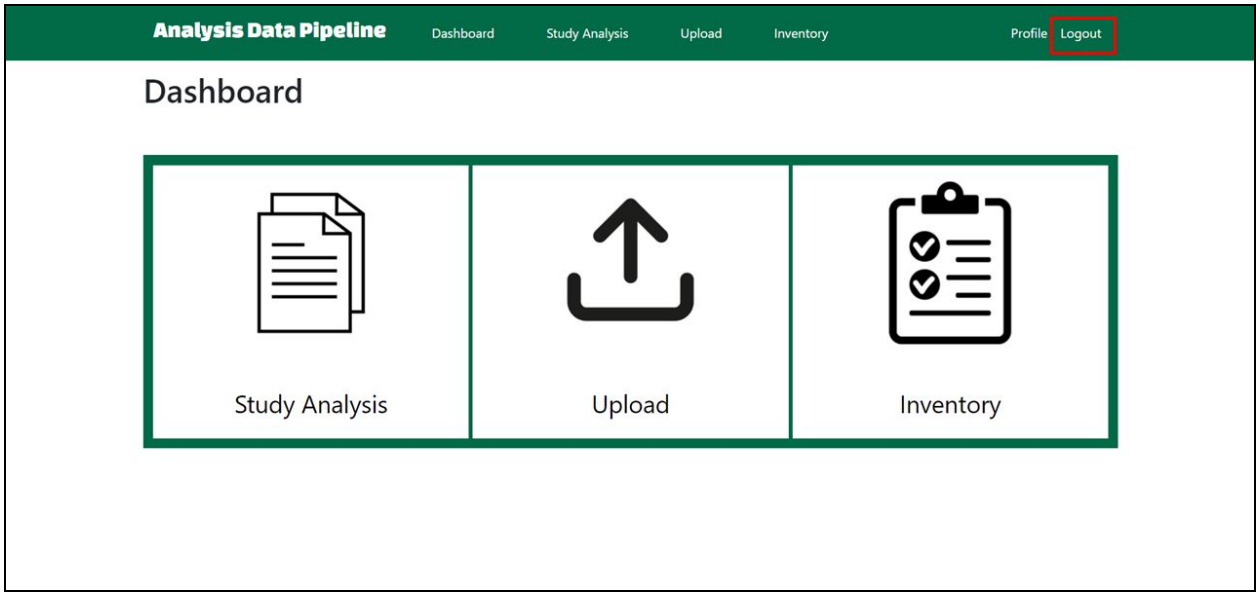
1.1.2 Login

Once the profile has been created, you will be prompted to log in. You can also reach the login screen by clicking the “Login” text on the nav bar.



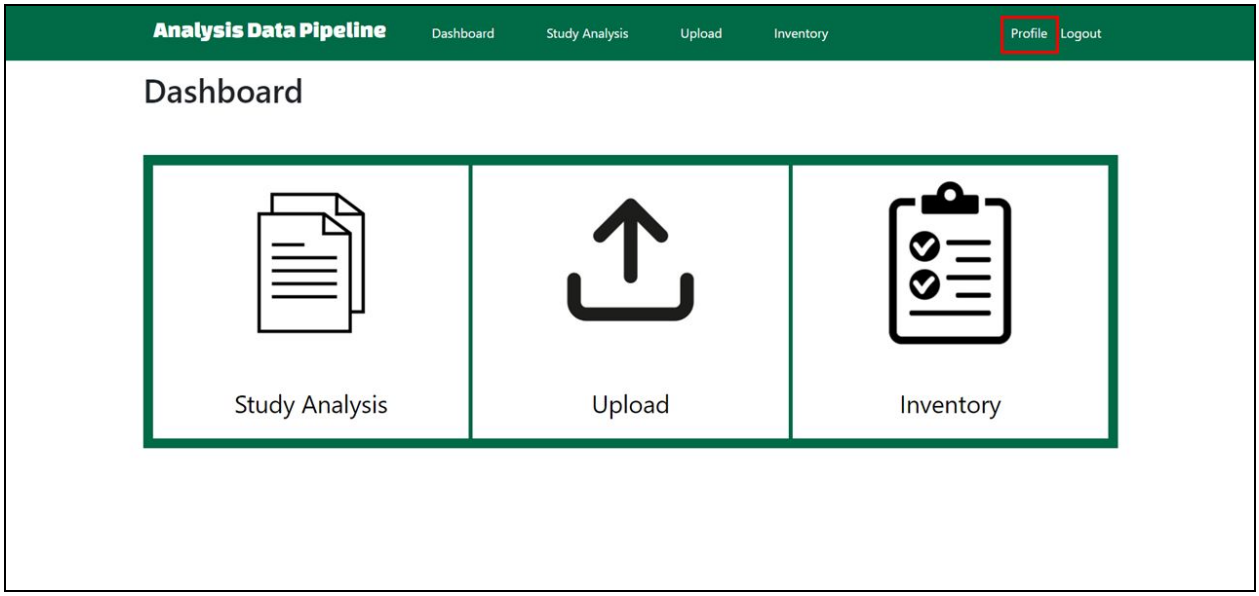
1.2 Logout

To log out, click the “Logout” text in the top right of the nav bar. Note that you need to be logged in to see this option.



1.3 Update Profile

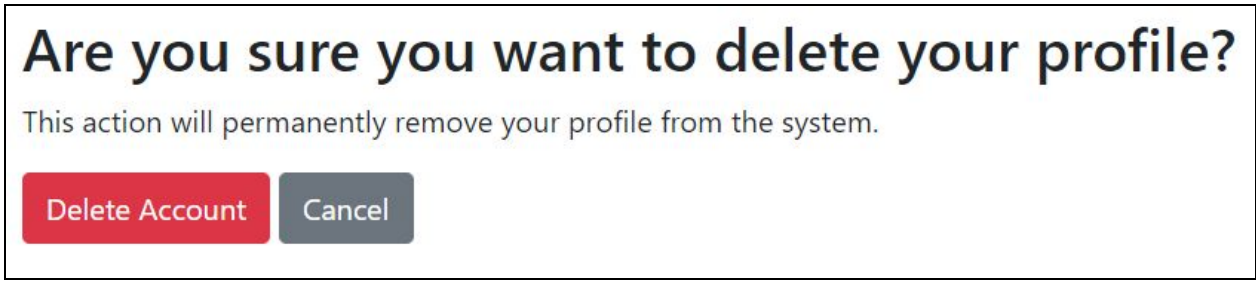
To view or update your profile information, click the “Profile” text at the top right of the nav bar.



This next page will show your profile information and allow you to update the information by clicking the “Update” button. If successful, a popup at the top of screen will confirm the update was successful.

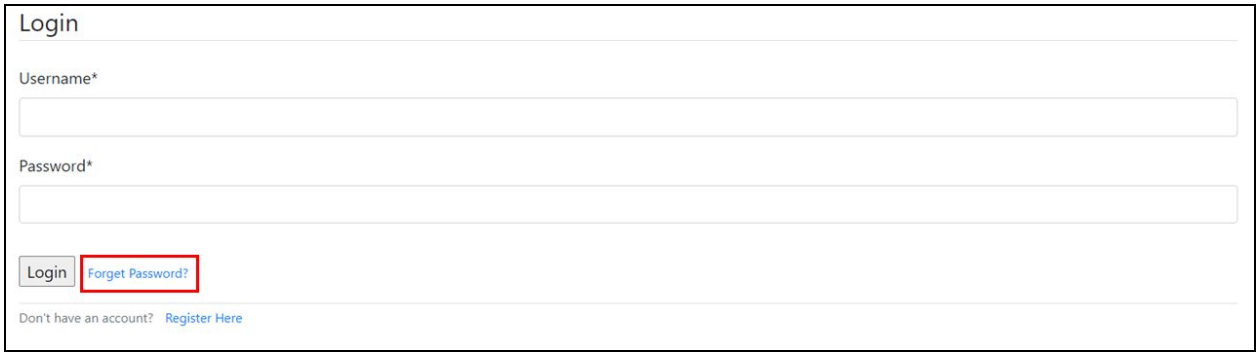
1.2.1 Delete Profile

Users can also delete their profile by clicking the “Delete Profile” button. This will take them to another screen to ensure they want to actually delete their profile. Clicking “Delete Account” will delete the account from the system and return the user to the login page.

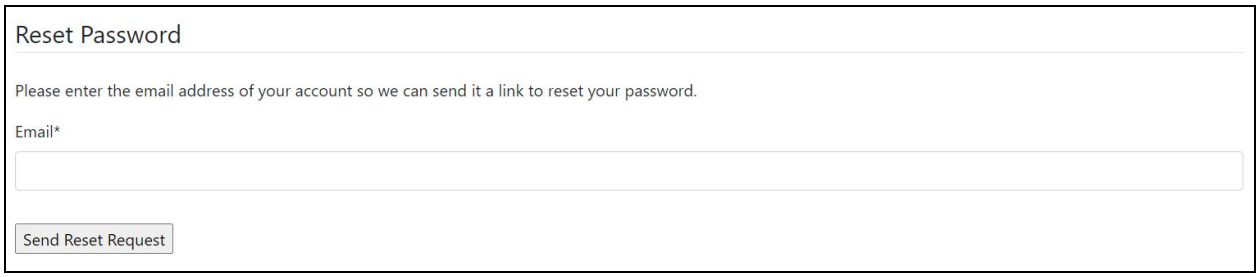


1.4 Reset Password

If you forgot your password or want to change it, go to the Login page and click “Forgot Password”. Note that you must be logged out to get to the login page.

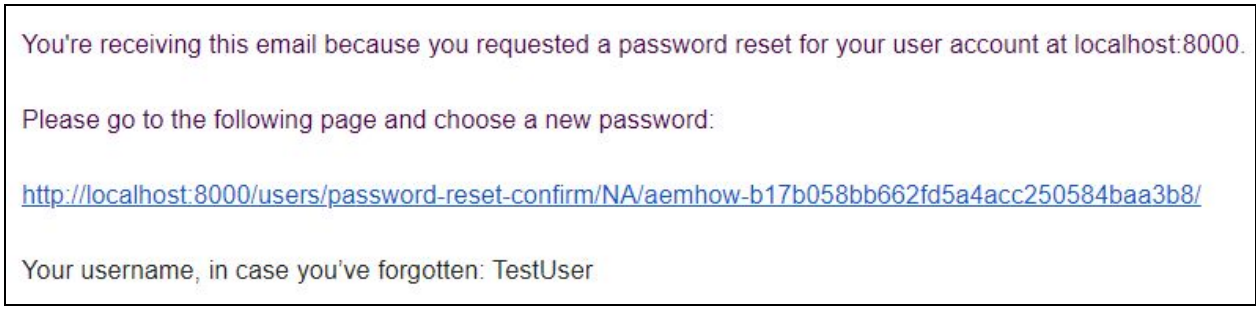


On the next page, enter the email address of your account and click “Send Reset Request”.



You will then be taken to the following screen. For security reasons, this will not confirm if there is an existing profile with that email.

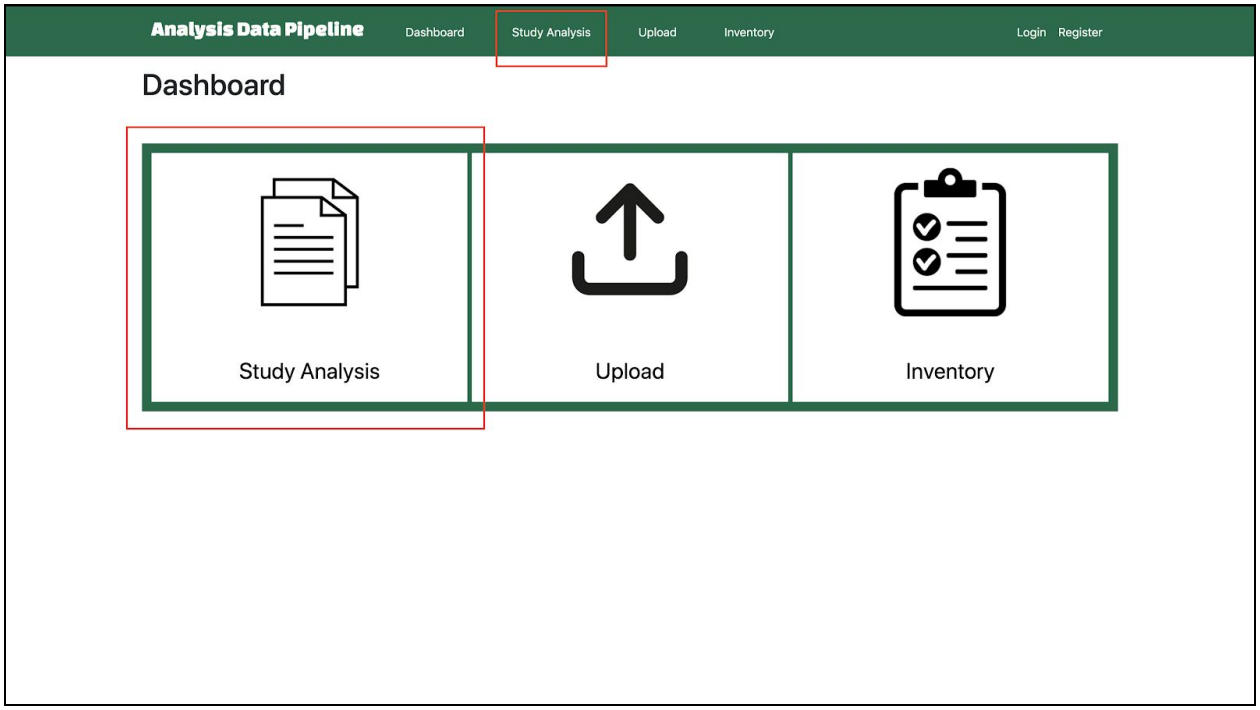
If there is a profile with the entered email address, it will receive an email from wpidatapipeline@gmail.com with the following format:



Clicking the link from the email will lead to the following page where the user can reset their password and then be redirected to the login page.

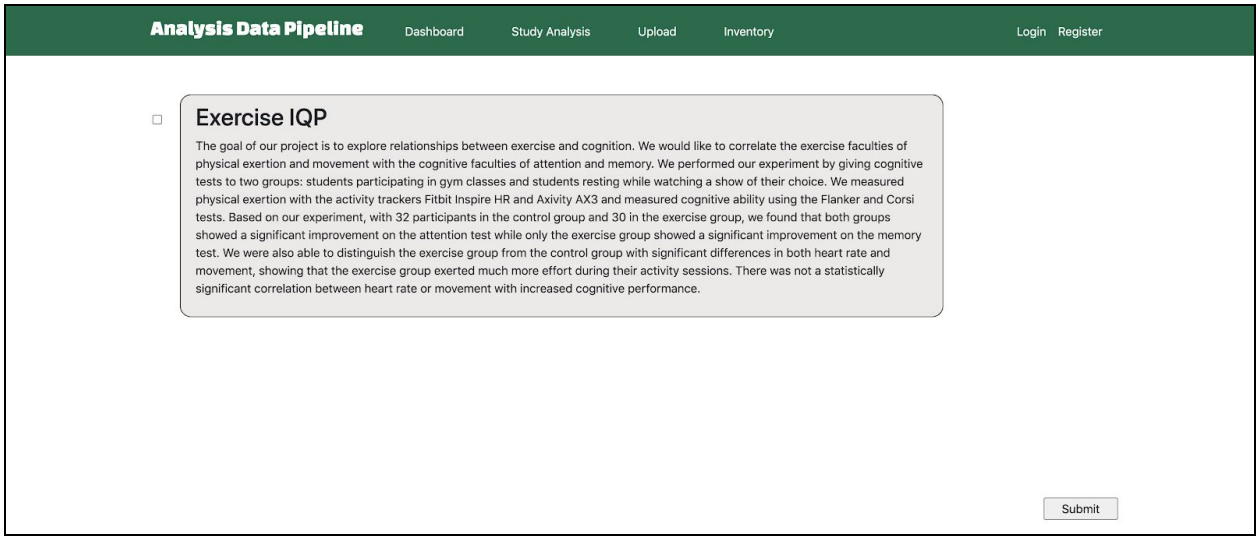
2. Study Analysis

The Study Analysis feature of the pipeline allows you to view and filter uploaded data to analyze and export for future use. You can access this feature by clicking “Study Analysis” on the Dashboard or the nav bar.



2.1. Selecting A Study

This page lists all of the studies with data in the pipeline, as well as the description of the study. Here, you may select one or more studies to analyze by clicking the checkboxes next to the study name. Once you have selected your study/studies, click “Submit”.



2.2 Selecting Data Category and Study Group

Next, you will be asked to select which data categories and study groups you would like to analyze.

Data Category: The type of data that was collected in the study. For example, a study exploring the relationship between exercise and memory would have several data categories, like heart rate and cognitive test results.

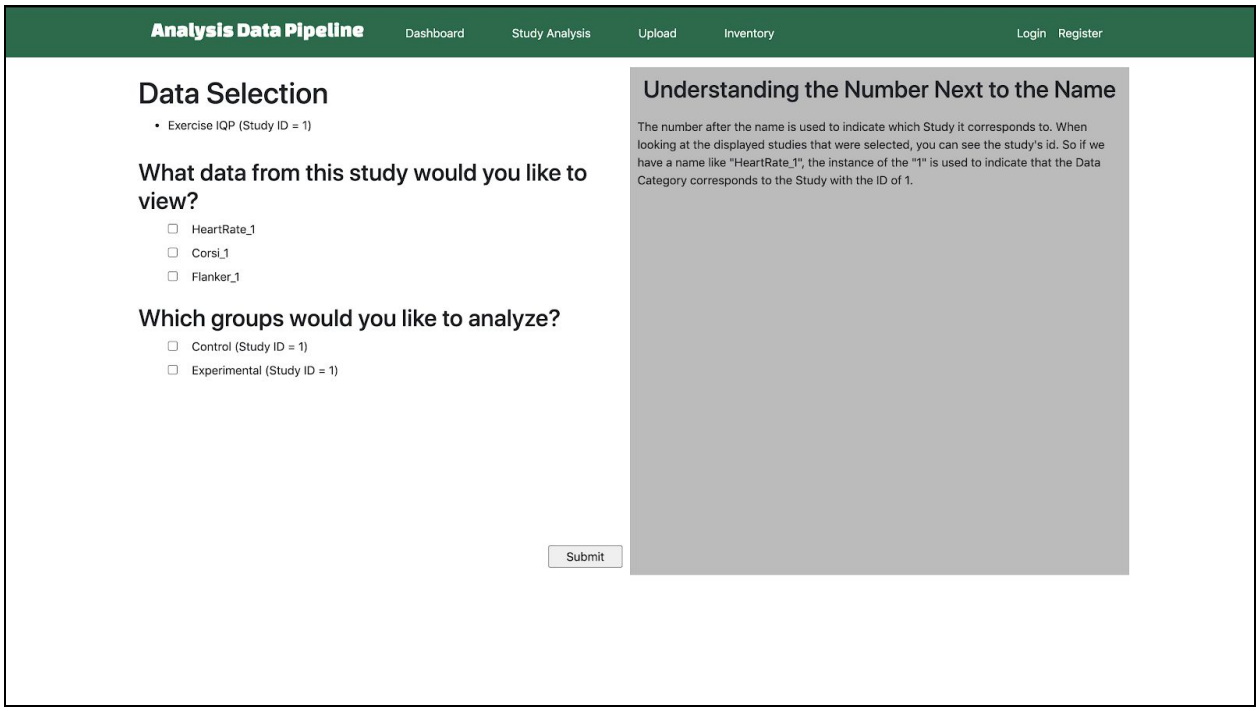
Study Group: The groups that subjects were split into for a study, such as control and experimental or several sample groups.

At the top of the page, you will see the study/studies that you chose, along with its study number. The study number uniquely identifies the study.

Below this, you will see all the data categories and study groups listed for the study/studies you selected. Next to each data category and study group is its corresponding study number. This is to aid you in identifying which data categories and study groups belong to which studies when doing cross study analysis.

You may select more than one data category, but you must select at least one. You may select more than one study group, but you must select at least one.

To select a data category or study group, click the checkbox next to its name. Once you have selected your data categories and study groups, click “Submit”.



2.3 Selecting Data Attributes and Filtering

On this page, you can select the specific attributes you would like to view and filter how you would like to view them.

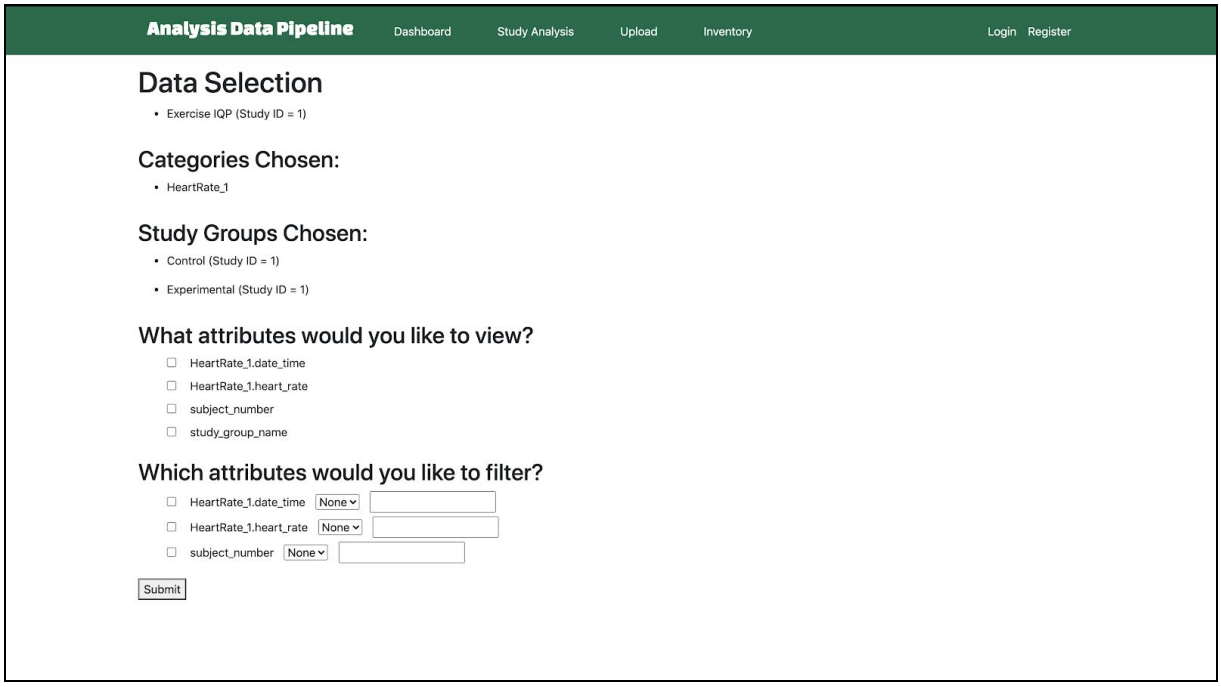
> **Attribute:** The specific fields collected for each data category. For example, a study collecting heart rate data (data category) would collect data for attributes like the heart rate in bpm, as well as what time that reading was taken. Often, these are the columns of the data table.

At the top of the page, you will see the study/studies, data categories, and study groups that you selected.

Below this, you will see a list of all the attributes for the data category/categories you selected, titled “What attributes would you like to view?” The attributes you select here will be shown as columns in the output table. You may select more than one data attribute, but you must select at least one. To select the attribute, click the checkbox next to its name.

In the section titles “Which attributes would you like to filter?”, each attribute is listed with a checkbox, a dropdown menu, and a text box. Here, you may filter the attributes based on their values. To filter the results based on the attribute, click the checkbox next to the attribute name, select a symbol from the dropdown, and enter the value that you would like to filter by into the text box. For example, setting heart rate to “=” in the dropdown box, and typing “106” into the text box will yield all the heart rate values that are equal to 106.

Note: You may filter by attributes that you did not select for viewing. Attributes selected for viewing and attributes selected for filtering are not dependent on each other.



2.4 Output Screen

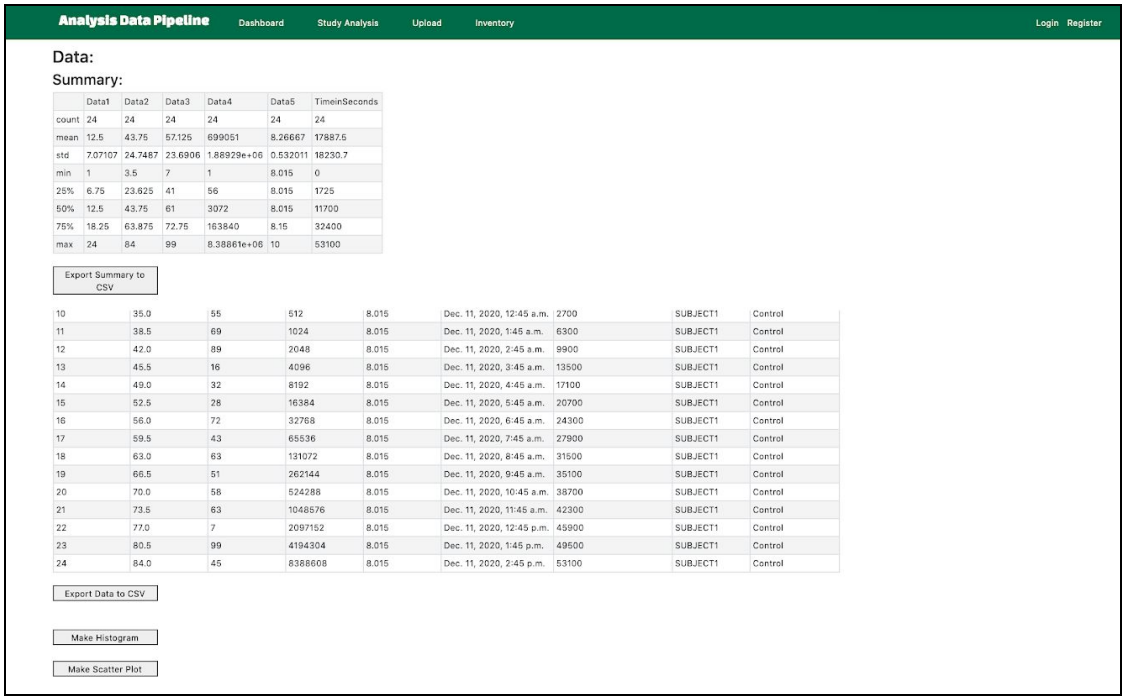
After performing a selection, two tables are shown. One is a statistical summary of the data and the other shows the selected data.

The data summary at the top of the page provides a list of all data columns with numerical values. The rows consist of eight descriptive metrics for each of these data columns.

The table below the summary contains all data that was selected from the previous process. It is formatted as a smaller box for formatting, but the user can scroll through the data to see it in its entirety.

Both tables can be exported to ".csv" files. The "Export data to CSV" buttons for each will be found under their respective tables, and clicking the button will start the ".csv" download.

To further analyze the data, select either option below to generate a histogram or a scatter plot.



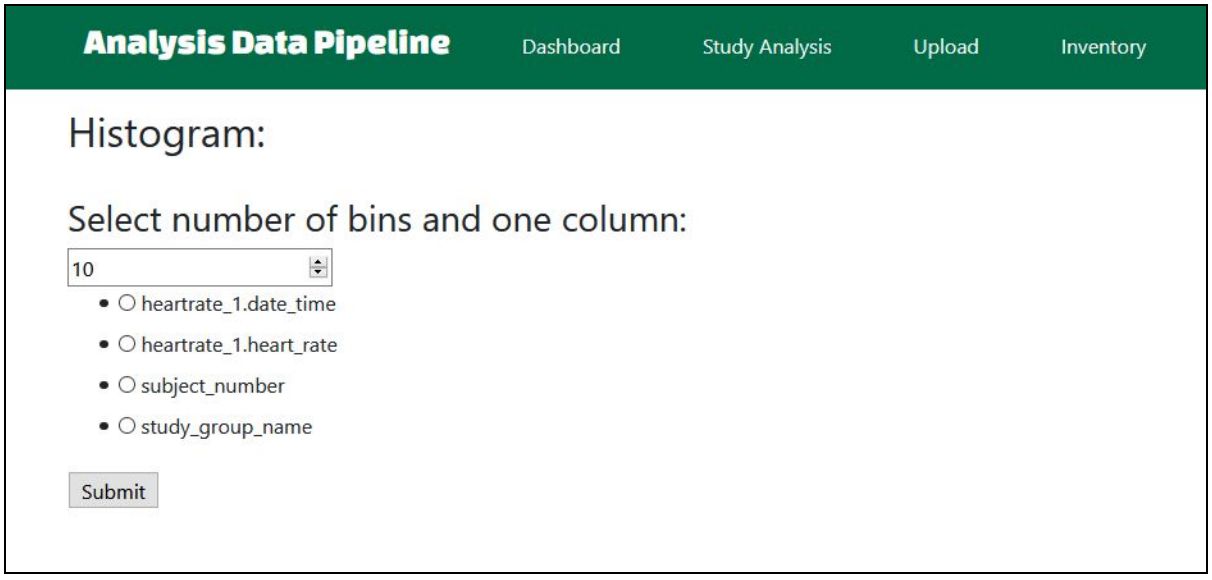
	Data1	Data2	Data3	Data4	Data5	TimeinSeconds
count	24	24	24	24	24	24
mean	12.5	43.75	57.125	699051	8.26667	17887.5
std	7.07107	24.7487	23.6906	1.88929e+06	0.532011	18230.7
min	1	3.5	7	1	8.015	0
25%	6.75	23.625	41	56	8.015	1725
50%	12.5	43.75	61	3072	8.015	11700
75%	18.25	63.875	72.75	163840	8.15	32400
max	24	84	99	8.38861e+06	10	53100

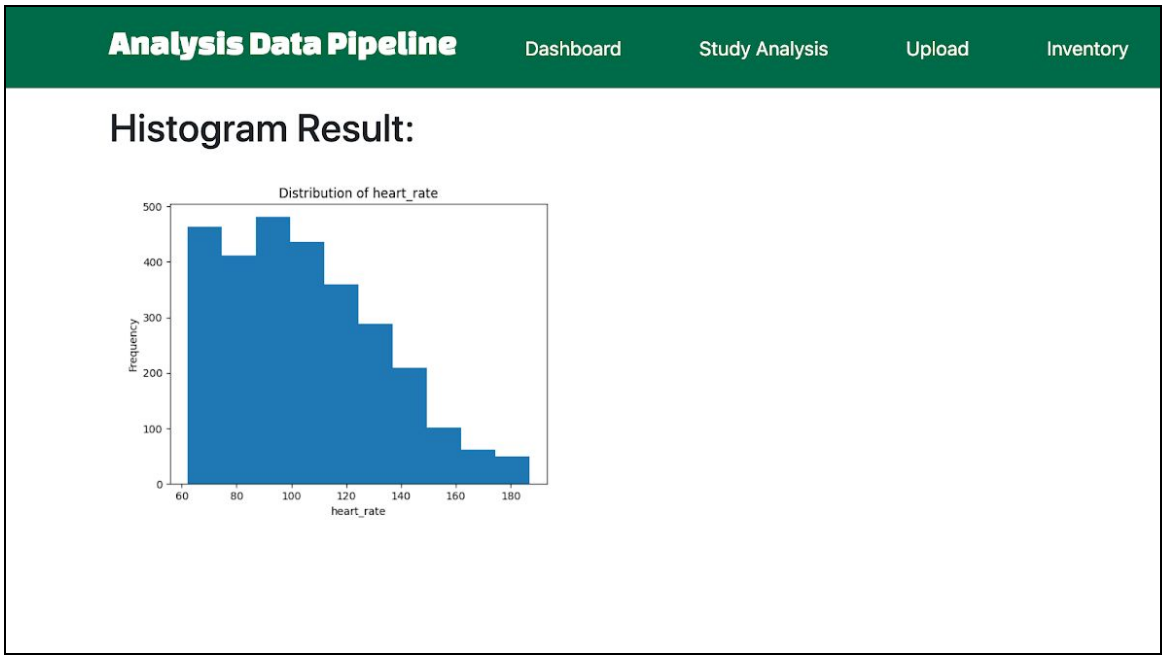
10	35.0	55	512	8.015	Dec. 11, 2020, 12:45 a.m.	2700	SUBJECT1	Control
11	38.5	69	1024	8.015	Dec. 11, 2020, 1:45 a.m.	6300	SUBJECT1	Control
12	42.0	89	2048	8.015	Dec. 11, 2020, 2:45 a.m.	9900	SUBJECT1	Control
13	45.5	16	4096	8.015	Dec. 11, 2020, 3:45 a.m.	13500	SUBJECT1	Control
14	49.0	32	8192	8.015	Dec. 11, 2020, 4:45 a.m.	17100	SUBJECT1	Control
15	52.5	28	16384	8.015	Dec. 11, 2020, 5:45 a.m.	20700	SUBJECT1	Control
16	56.0	72	32768	8.015	Dec. 11, 2020, 6:45 a.m.	24300	SUBJECT1	Control
17	59.5	43	65536	8.015	Dec. 11, 2020, 7:45 a.m.	27900	SUBJECT1	Control
18	63.0	63	131072	8.015	Dec. 11, 2020, 8:45 a.m.	31500	SUBJECT1	Control
19	66.5	51	262144	8.015	Dec. 11, 2020, 9:45 a.m.	35100	SUBJECT1	Control
20	70.0	58	524288	8.015	Dec. 11, 2020, 10:45 a.m.	38700	SUBJECT1	Control
21	73.5	63	1048576	8.015	Dec. 11, 2020, 11:45 a.m.	42300	SUBJECT1	Control
22	77.0	7	2097152	8.015	Dec. 11, 2020, 12:45 p.m.	45900	SUBJECT1	Control
23	80.5	99	4194304	8.015	Dec. 11, 2020, 1:45 p.m.	49500	SUBJECT1	Control
24	84.0	45	8388608	8.015	Dec. 11, 2020, 2:45 p.m.	53100	SUBJECT1	Control

2.4.1: Histogram

Clicking the “Make Histogram” button will bring the user to a new page with a selection prompt. From here, the user can select the amount of bins (equally spaced intervals for data points to fall into) as well as a single attribute from the selected data to be used in the histogram. The default number of bins is 10, but can go as low as 1 and infinitely high.

Clicking the submit button will redirect to a new page that displays the figure created from the selection prompt. For most Web browsers, right click and select 'Save as' to save the image to a local file.

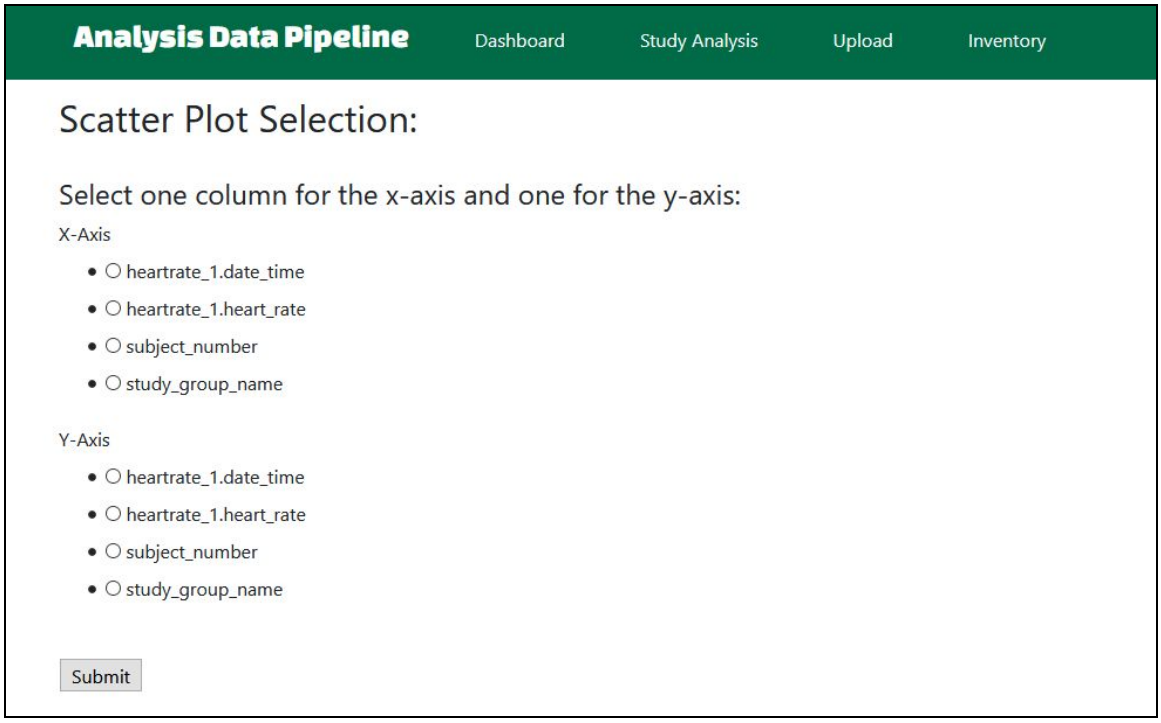


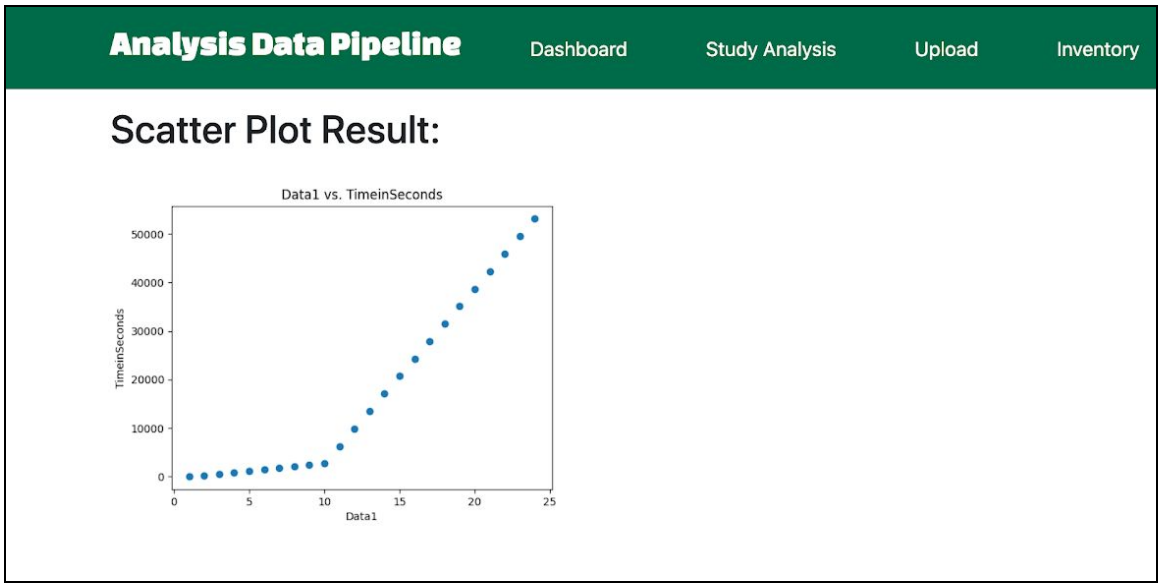


2.4.2: Scatter Plot

Clicking the “Make Scatter Plot” button will bring the user to a new page with a selection prompt. From here, the user will select one attribute from the “Select X-Column” boxes and one category from the “Select Y-Column” boxes.

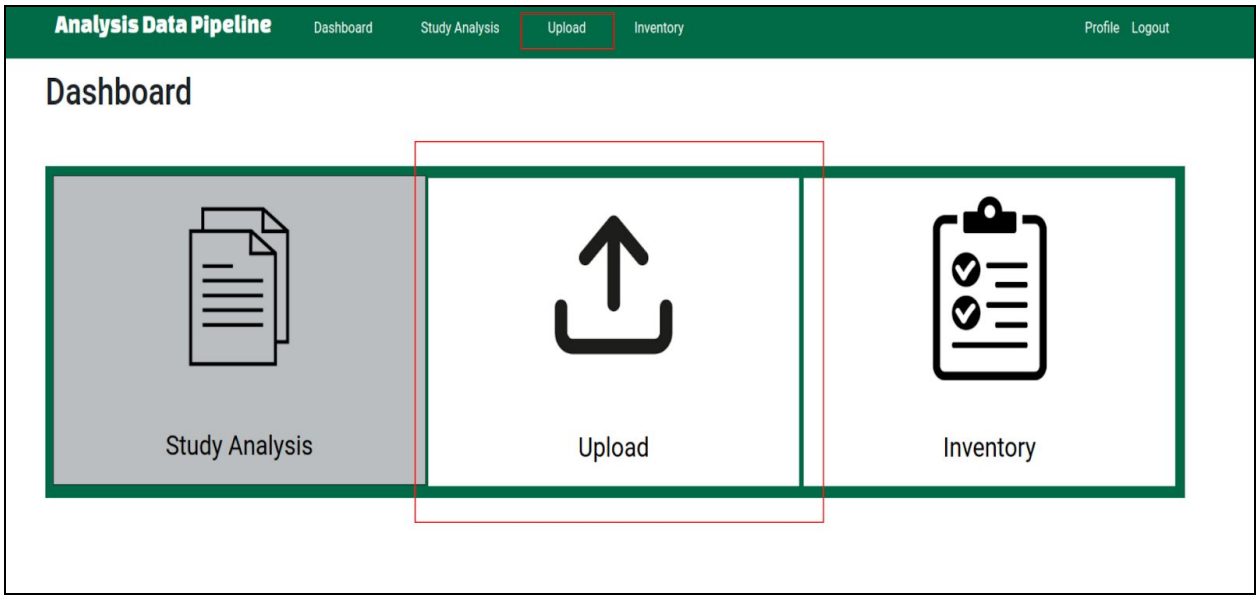
Clicking the submit button will redirect to a new page that displays the plot created with the previous selections. Selecting more than one category from either list will result in the figure output being the category selected highest in the list.





3. Upload

The Upload feature of the pipeline allows you to upload data to the application. You can access this feature by clicking “Upload” on the Dashboard or the navigation bar.



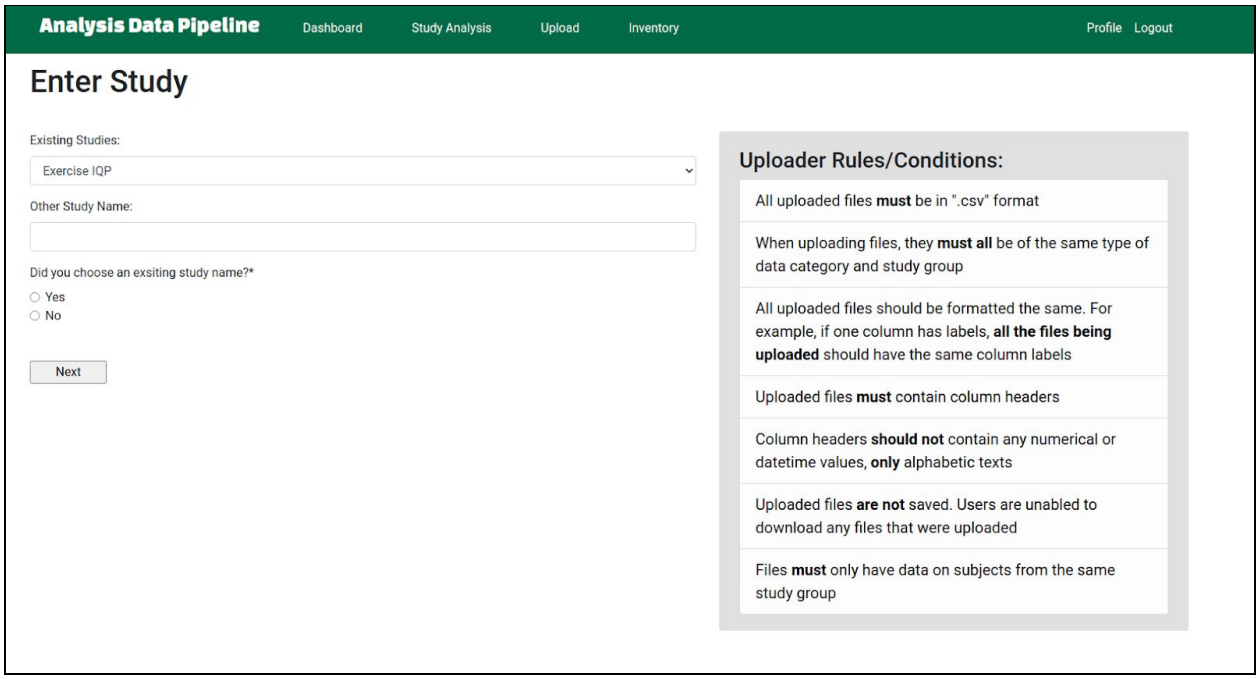
3.1: Study Name

On this page, the user is prompted to provide the name of the study that they want to upload data to. If there are currently existing studies within the database, the user will have the option to select one of those studies using the drop down menu. It is important that the user selects the ‘Yes’ option when asked if they chose an existing study name.

If there are no existing studies in the database, ‘Existing Studies’ drop down will not exist, and the user will be asked to enter in the name of the study.

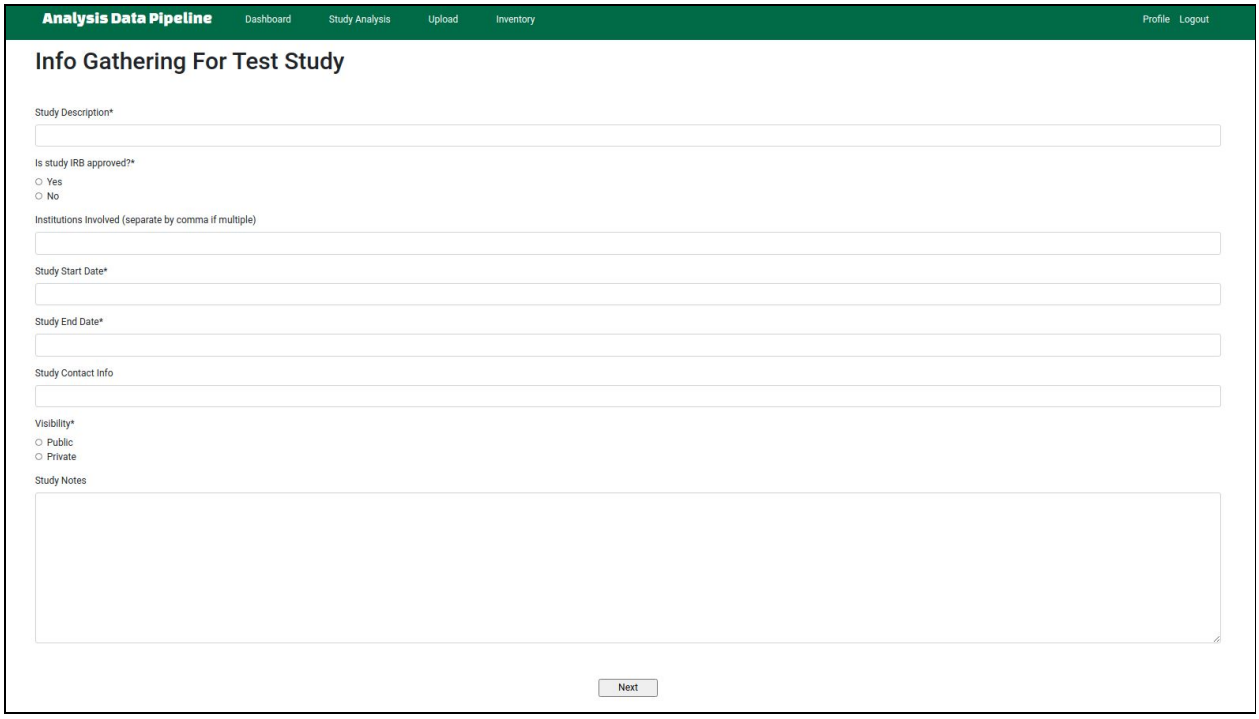
The page includes a small menu that displays the important uploader rules/conditions to follow in order to use this pipeline’s uploader feature. At the bottom of the menu, the user can download a pdf version of the rules/conditions.

By clicking on the ‘Next’ button, the user will be able to proceed forward in the uploader process.



3.1.1: Study Info

If the user reached this page, the system has detected that the user is adding a new study to the database. Because a new study is being added to the database, the user is required to provide metadata for that study so that other users of the system will be able to understand what this new study is.



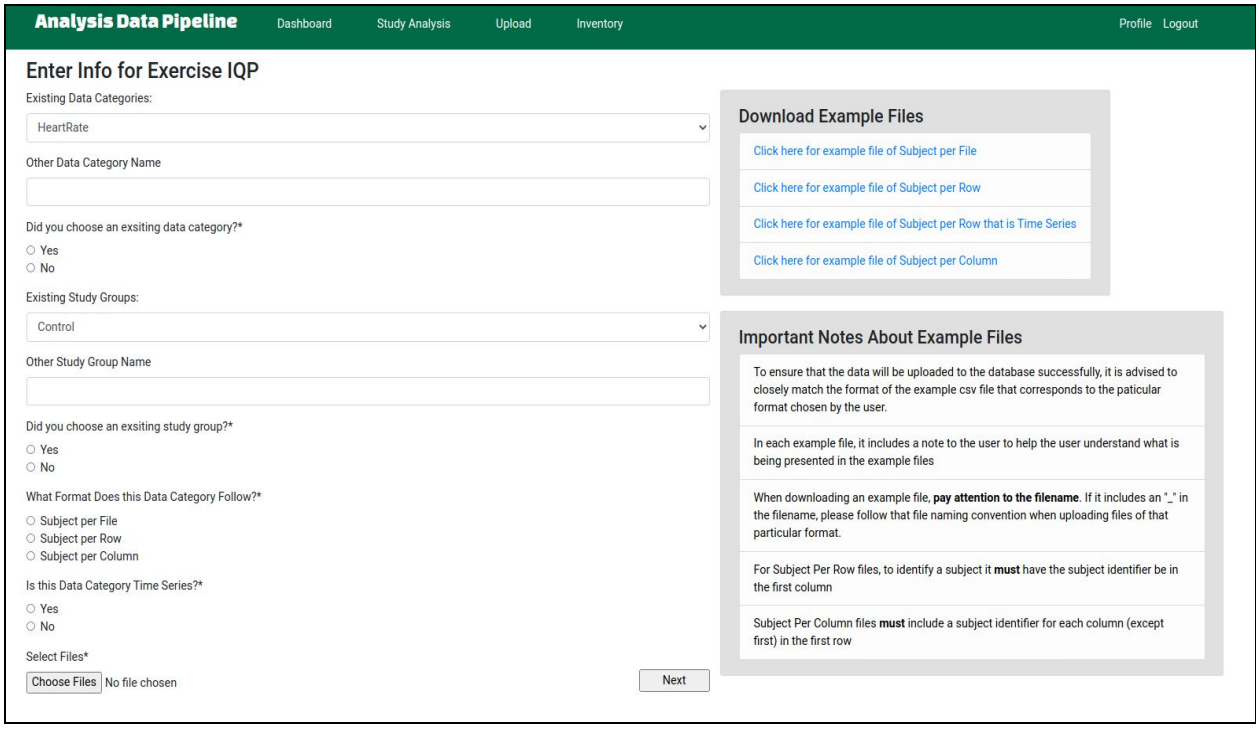
3.2: Info

Data Category: The type of data that was collected in the study. For example, a study exploring the relationship between exercise and memory would have several data categories, like heart rate and cognitive test results.

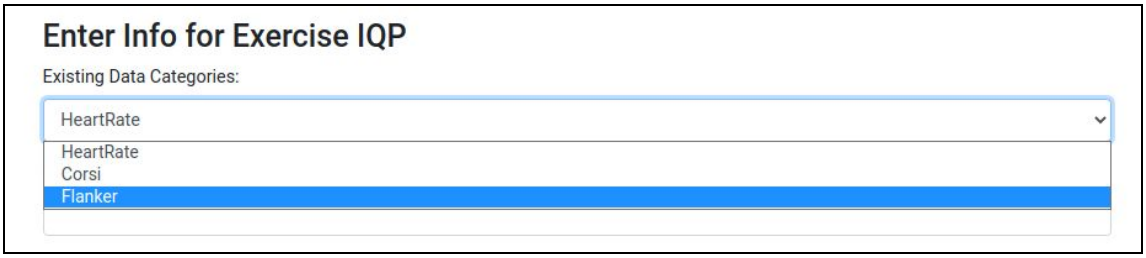
Study Group: The groups that subjects were split into for a study, such as control and experimental or several sample groups.

On this page, the user will provide information about which data category (if it exists) and study group (if it exists) they want to upload data to. If the option exists, the user should select ‘Yes’ if they chose an existing data category and/or an existing study group.

If there are existing data categories and/or study groups for the specified study, there will be a drop down box that will allow the user to select any of the existing data categories/study groups. In the image below, it indicates that HeartRate, Corsi, Flanker are an existing data category for the ‘Exercise IQP’ study.



The user is also prompted to provide information about the data they are uploading such as its format and if it is a time series data. Once all the information has been filled out, the user should upload the “.csv” files.

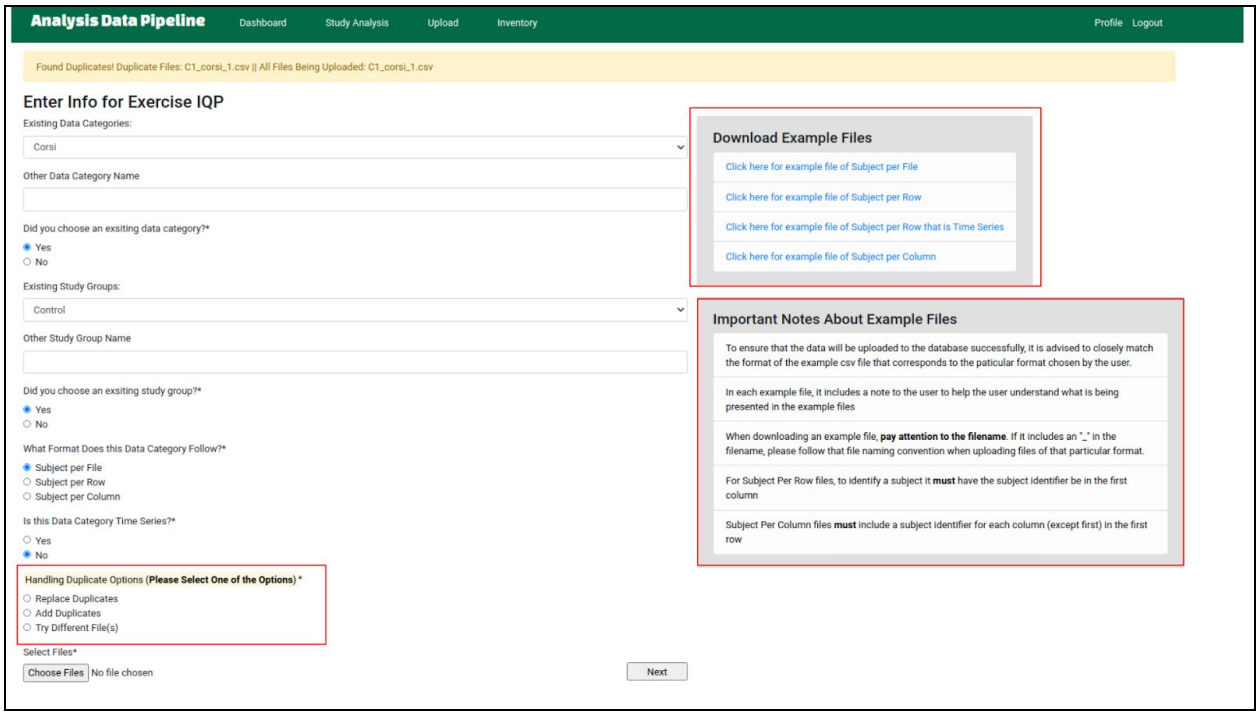


3.2.1: Info - Important Features to Lookout For

- Download Example Files
    - In case the user is unsure of how to structure their .csv file, they can download an example file of the corresponding format that they are interested in uploading

- Important Notes About Example Files
    - Important notes to help user understand how to format their uploaded files

- Duplicate File Handler
    - Upon detection of duplicate files, the page will include an additional prompt to ask the user on how they want to handle uploading duplicate files
        - ***Replace Duplicates***
            - This would remove all the data that was initially uploaded from the previous version of the duplicate file and replace them with the data from the current uploaded file
        - ***Add Duplicates***
            - This would simply just add the data without doing anything to the database. This potentially adds the risk of having duplicate data within the database table
        - ***Try Different File(s)***
            - This option allows the user to change the name of the detected duplicate file and have the user try uploading a newly named file. If the newly named file is a duplicate of an existing file, the user will simply be prompted again and repeat the process

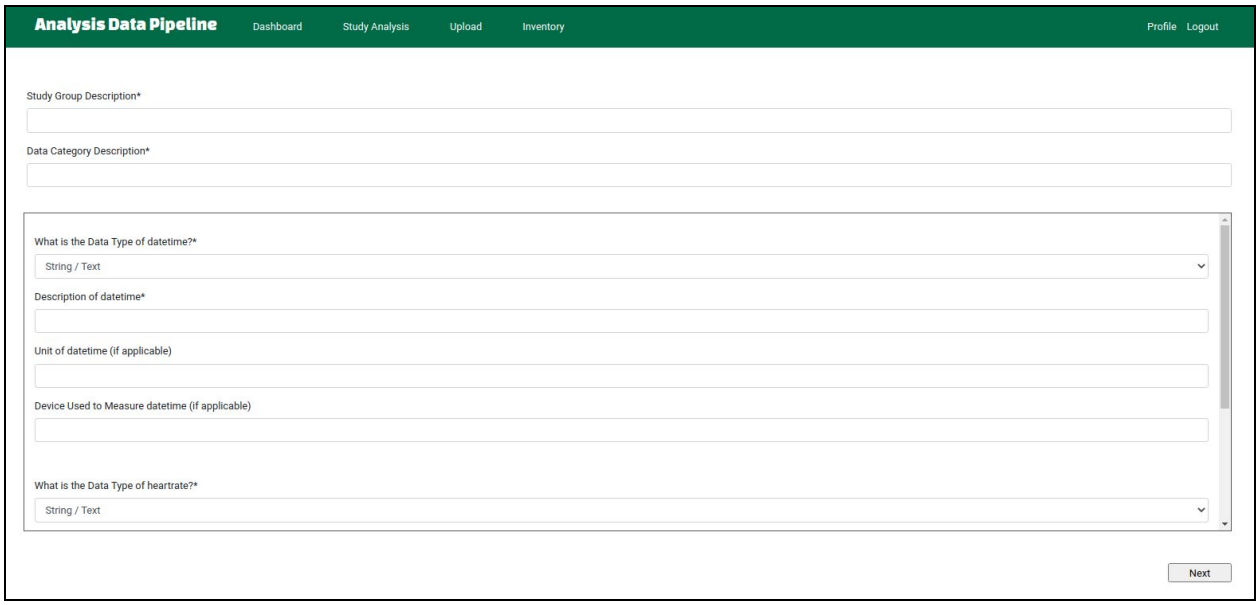


3.3: Extra Info

Attribute: The specific fields collected for each data category. For example, a study collecting heart rate data (data category) would collect data for attributes like the heart rate in bpm, as well as what time that reading was taken. Often, these are the columns of the data table.

If the user reached this page, it means that they are attempting to upload a new set of data to the database. This page asks the user all the necessary information it needs from the user based on the information that was provided from the Info page.

The purpose of this page is to acquire the metadata of the new data category to be added and/or study group name. Additionally, the user will be prompted to fill out the metadata for each of the detected column(s) / attribute(s) from the uploaded ".csv" file(s). By filling out the metadata, it helps other users of the system to be able to understand what other users have contributed to the system.

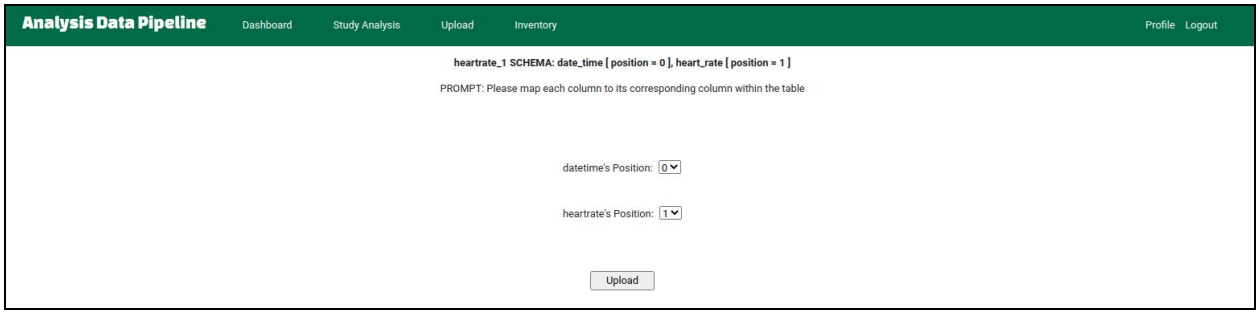


3.4: Final Prompt

On this page, the user is prompted to map the columns that were detected from their uploaded files to the corresponding columns of the table that the data will be uploading to. Because the column positioning of the uploaded .csv files will not always match the column order in the database tables, this feature provides the user the ability to simply remap their columns without requiring them to make their own alterations to their files.

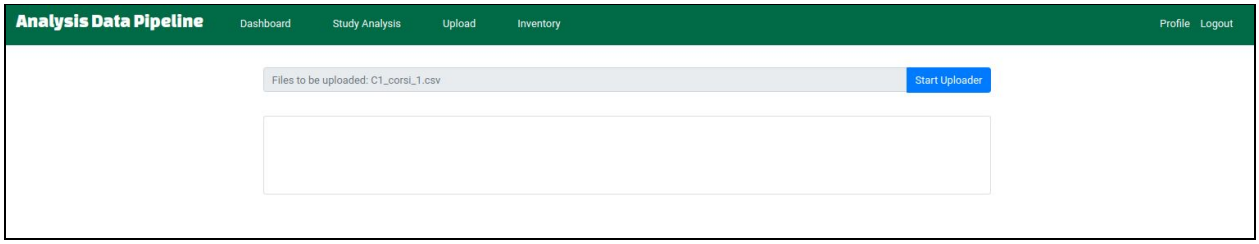
The user will use the drop down menu to select the correct positioning number. **Please note that the first column will have positioning of 0**.

Clicking the ‘Upload’ button, will direct the user to the uploader.

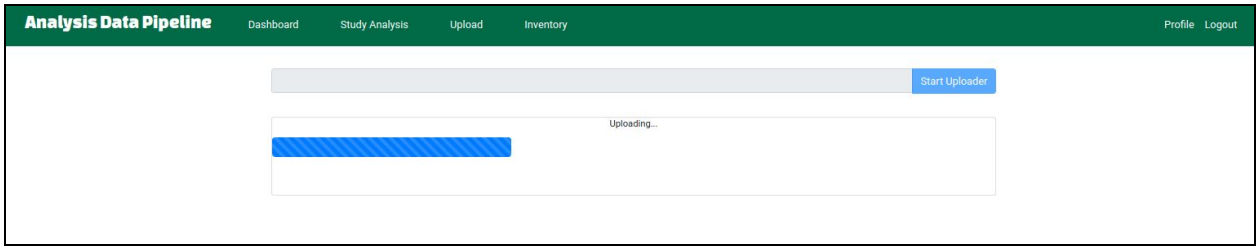


3.5: Uploader Page

At this page, this is where the uploading of the data to the database takes place. In the gray box, the user can view the files to be uploaded, and by pressing ‘Start Uploader’, it will begin the uploading process.

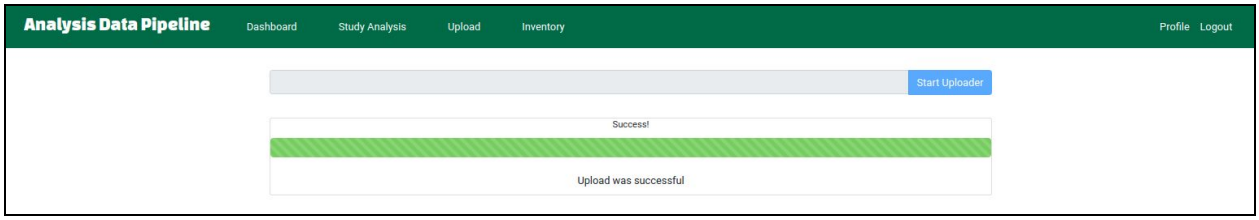


When the bar is blue, it indicates that the system is currently uploading the data. If it hangs, do not worry, it will eventually finish.



3.5.1: Success

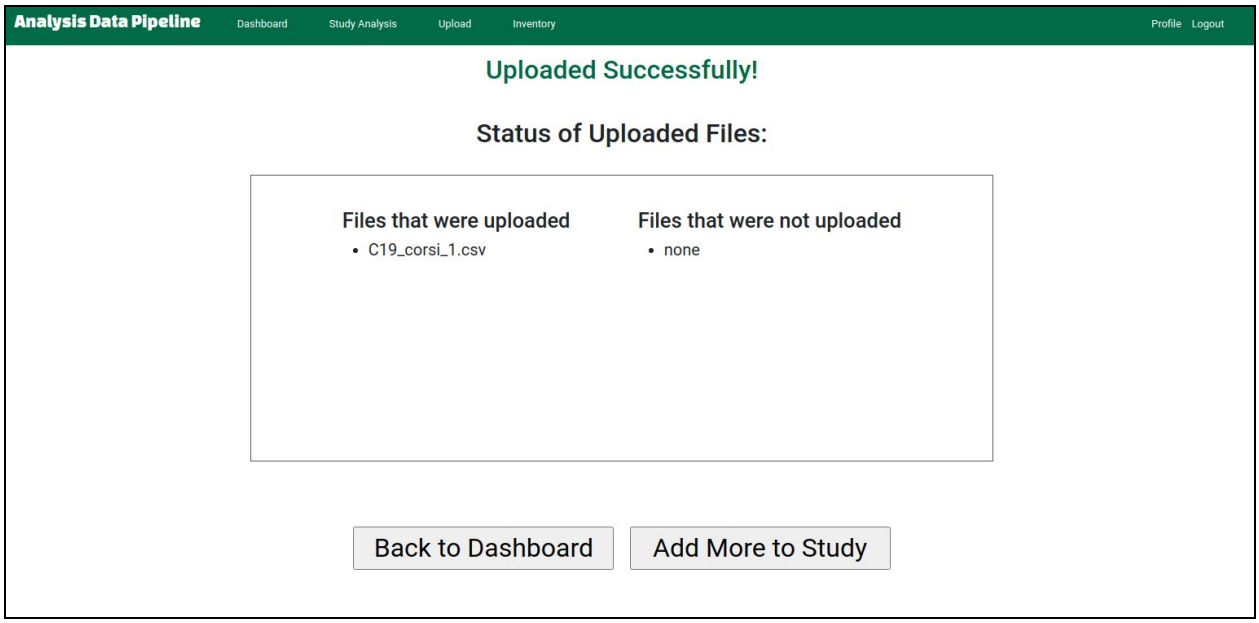
When the bar turns green from the uploader page, the green color indicates that the uploading process has completed and was successful. The user will soon be redirected to the successful upload page.



The user is brought to this page when all the data was uploaded successfully. All the files that were uploaded from the Info page would fall under the ‘Files that were uploaded’ column.

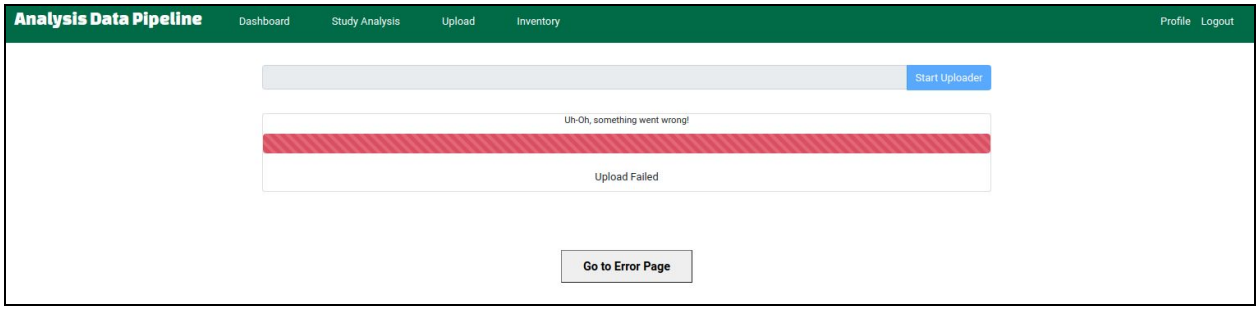
If the user clicks on the ‘Add More to Study’ button, the user will be redirected to the Info page where they can continue to upload more data to the same study.

If the user clicks on the ‘Back to Dashboard’ button, the user will be directed back to the Dashboard.



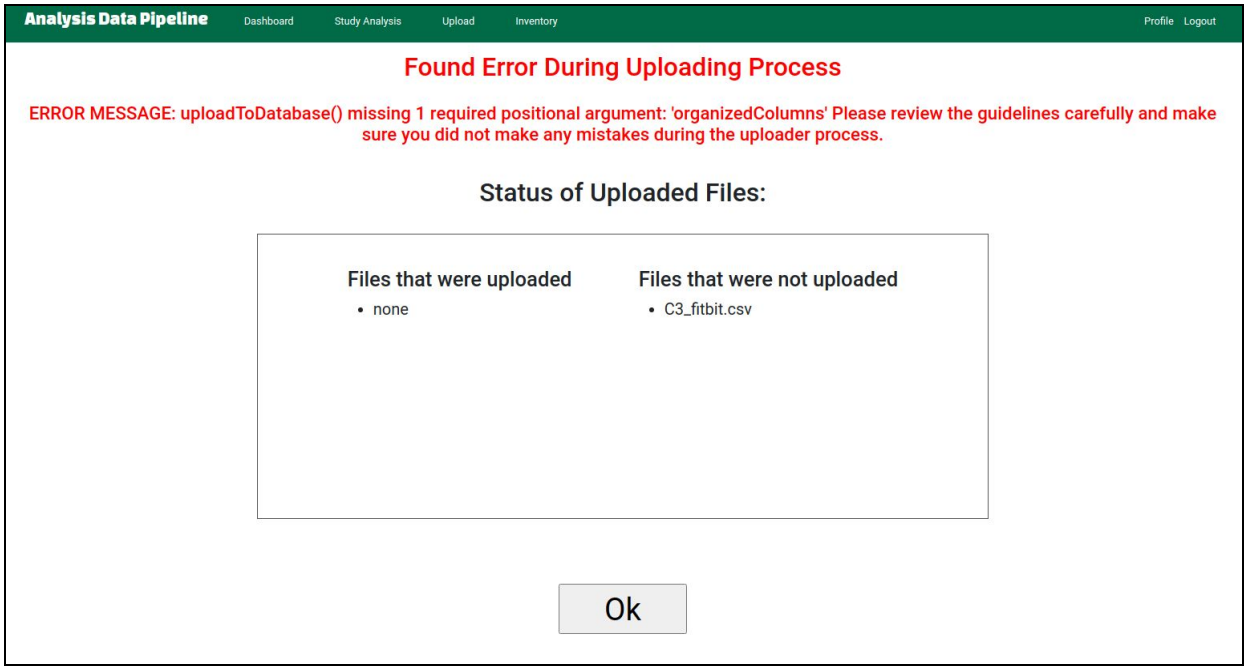
3.5.2: Error

When the bar turns red from the uploader page, it means an error has occurred during the uploading process. By clicking on the ‘Go to Error Page’ button, the user will see an error message and a list of all the files that were either uploaded or not uploaded successfully.



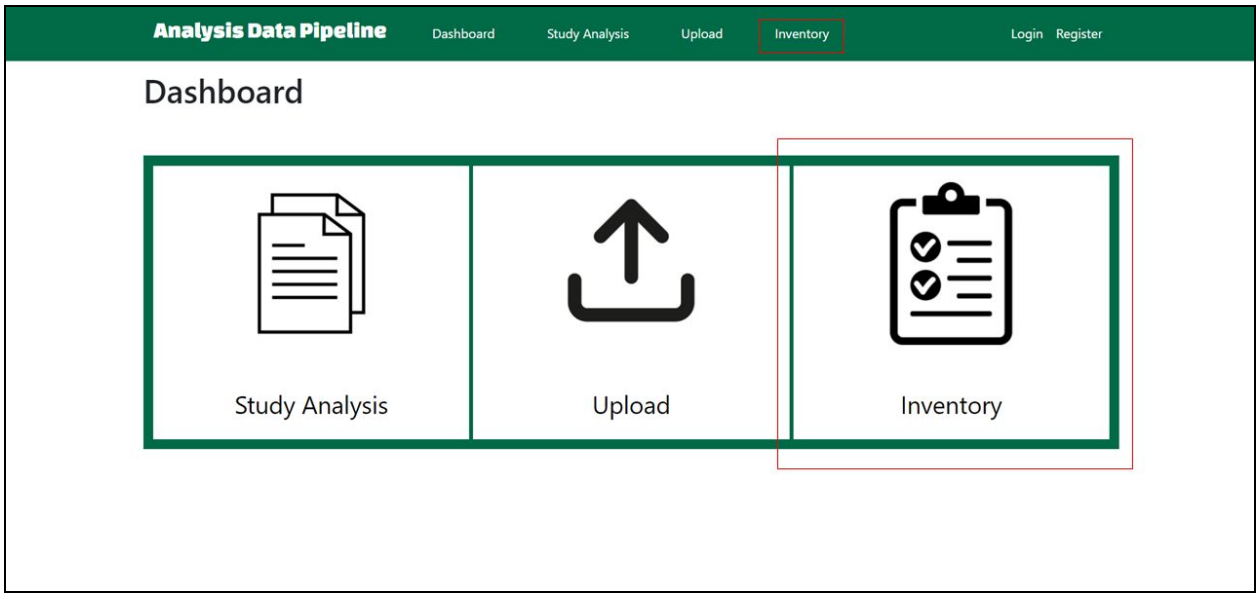
If there were any errors that occurred in the whole uploading process, the user will be redirected to this page, where they will be informed of the error that was found.

Similar to the Success page, this page also indicates which files were uploaded successfully to the database and which were not. By providing this information, the user will be able to learn which files they will have to try uploading again.



4. Inventory

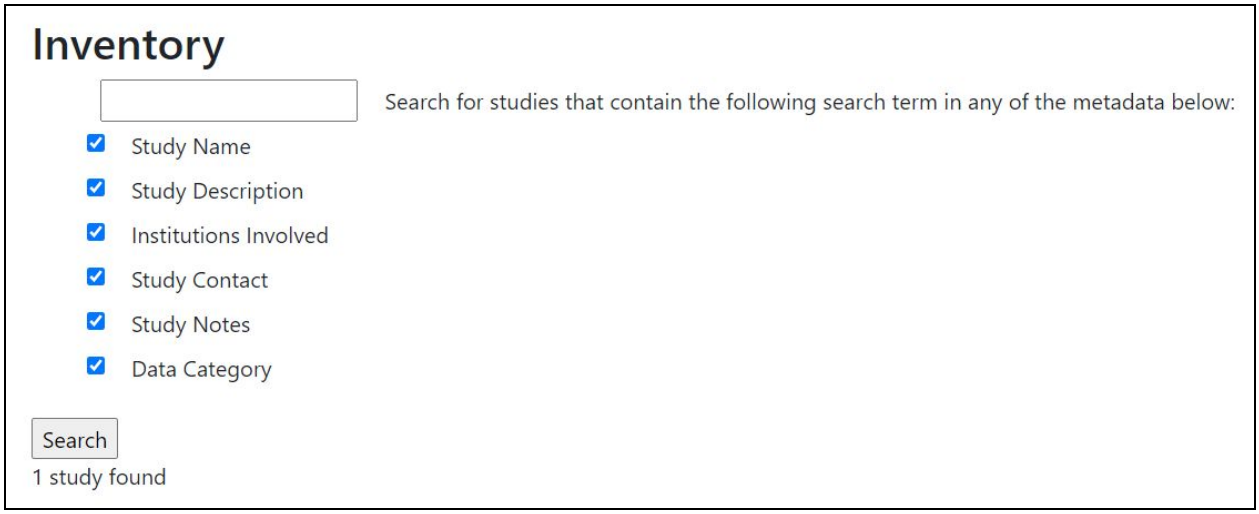
The Inventory feature of the pipeline allows you to view metadata about studies you have access to and find studies with specific data categories. You can access this feature by clicking “Inventory” on the Dashboard or nav bar.

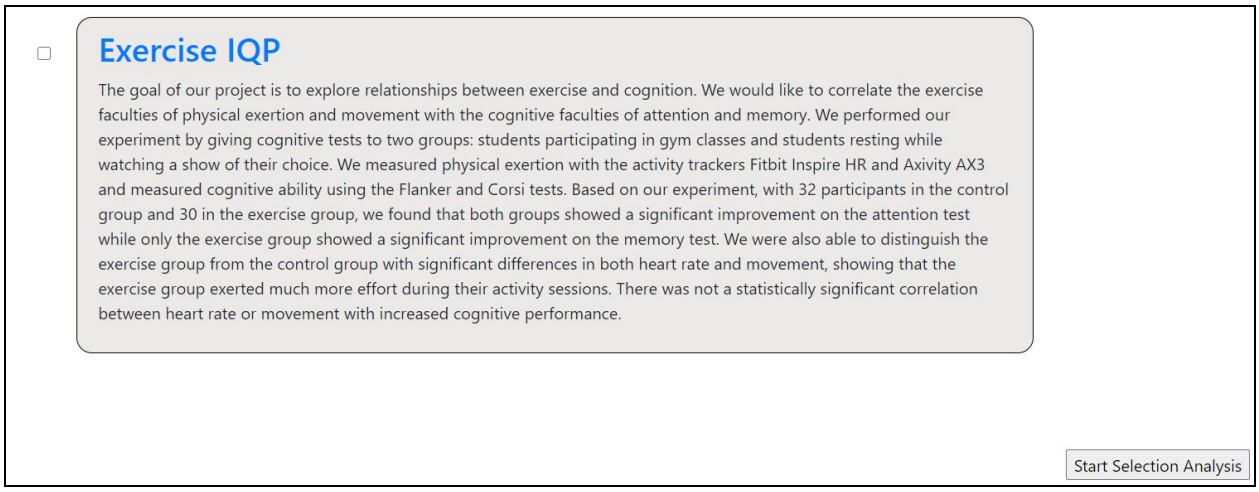


4.1 View Study List

Data Category: The type of data that was collected in the study. For example, a study exploring the relationship between exercise and memory would have several data categories, like heart rate and cognitive test results.

This view is similar to the study selection page. It shows the studies you have permission to view and allows you to select them using checkboxes and proceed into the selection analysis. However, this page also includes links for each study to allow you to view more details about them. Additionally, there is a search bar at the top of the page that filters studies based on if their metadata contains a given search term. This metadata includes the study title, study description, institutions involved, study contact, study notes, and data category within the study.



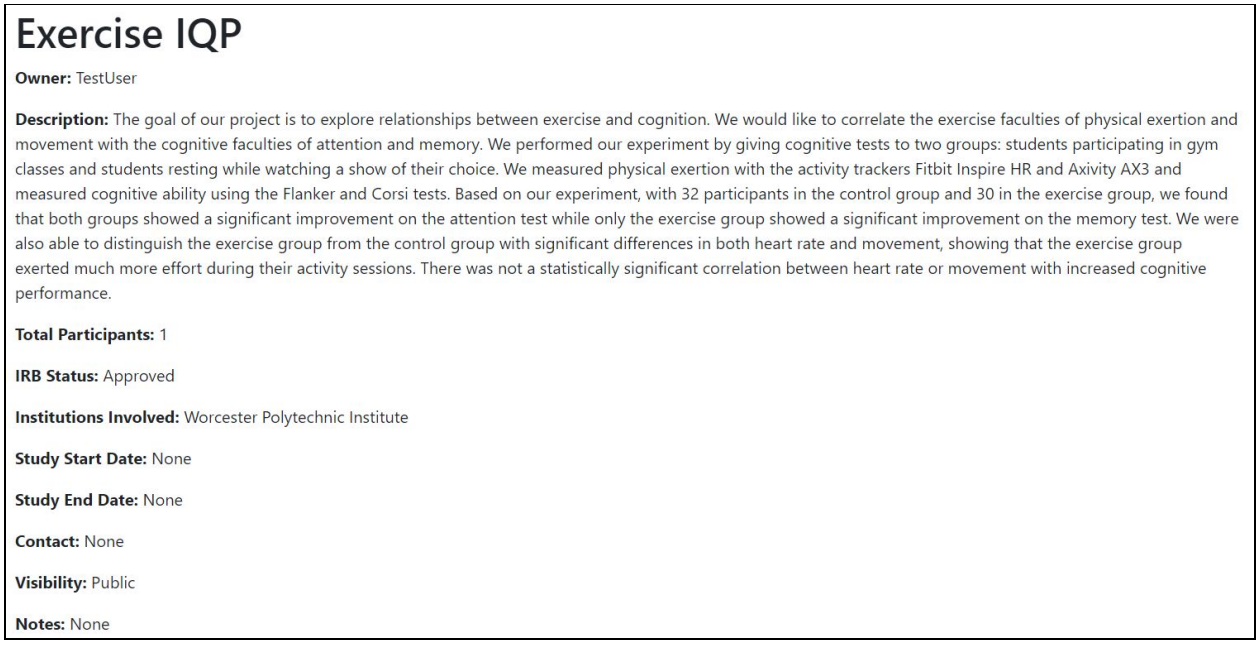


4.2 View Study Metadata

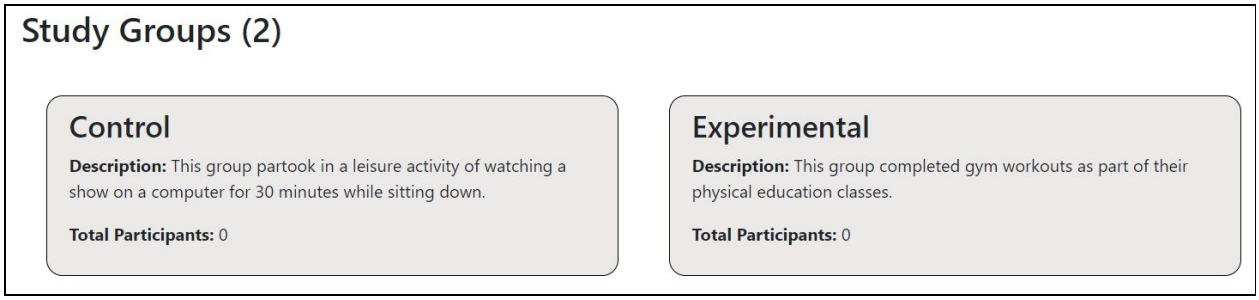
Study Group: The groups that subjects were split into for a study, such as control and experimental or several sample groups.

Attribute: The specific fields collected for each data category. For example, a study collecting heart rate data (data category) would collect data for attributes like the heart rate in bpm, as well as what time that reading was taken. Often, these are the columns of the data table.

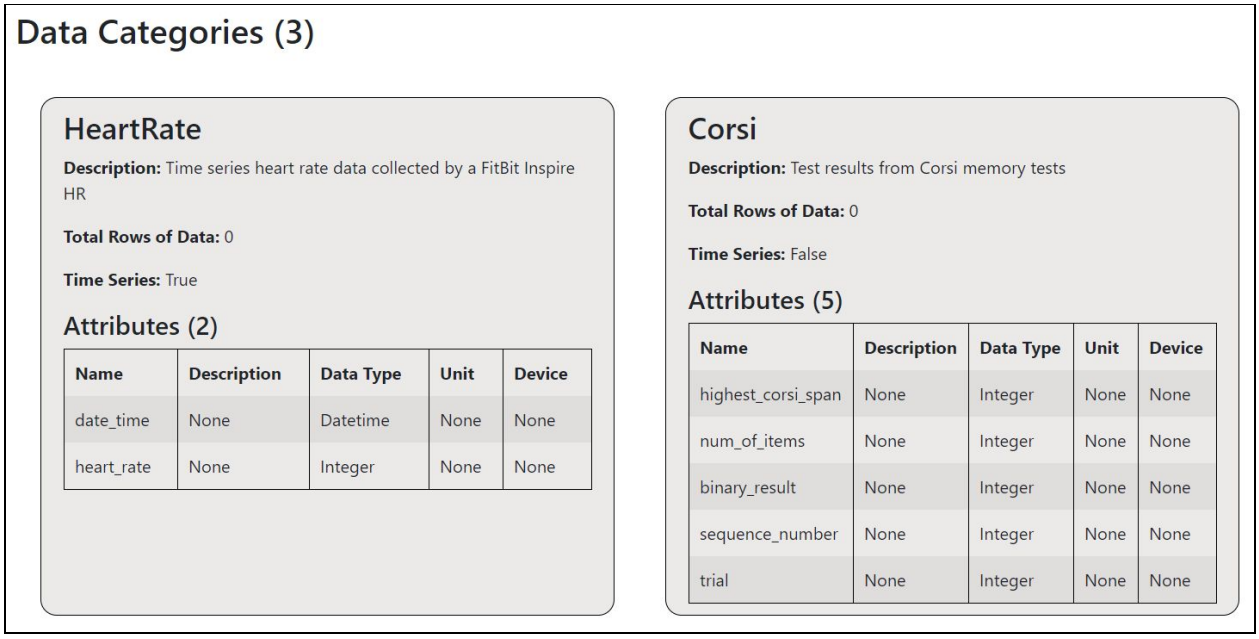
This page shows a variety of information for a specific study. This includes the study's description, total participants, IRB status, institutions involved, start date, end date, contact, visibility, and notes. The username of the user that owns that study is displayed beneath the study name.



The study groups of the study are displayed on the page as well. This shows the description and total participants for each of the groups.



The data categories of the study are displayed on the page as well. This shows the description, total rows of data, whether it is a time series, and a table of the column attributes for each data category.



Input Standards:

- Uploaded files **must** be in ".csv" format
- ~~Users can upload multiple of files at once~~
- Files **must** contain only one type of data category and study group
    - **Only** upload files from the same data category and study group
- Files **must** contain column headers
    - Column headers should not contain any numerical or datetime values, only strings
    - If file is organized in one-subject-per-column format, the file **must** contain row headers in addition to expected column headers
- For data categories that have one subject per file, the subject identifier **must** be in the beginning of the file name followed by an "_" with the rest of the file name
    - Ex: C1_heartrate.csv
- If the file is organized in a one-subject-per-row format and contains a subject identifier for each row, the identifier **must** be placed in the first column of the ".csv" file
- ".csv" files with the same data category **must** have the same columns. When uploading files at the same time, the order of the columns **must** be the same.
- Data categories and Study groups within a given study **must** have unique names
- ~~If entering in a new data categories, it is recommended to avoid adding spaces to the name~~
- Files **do not save**. If you leave the uploader in the middle of the uploader process, your files **Will be deleted**.

- File type
    - All uploaded files **must** be in ".csv" format
- Only contain one type of data category and study group
    - When uploading files, they **must all** be of the same type of data category and study group
- Column Headers
    - Files **must** contain column headers
        - Headers should not contain any numerical or datetime values, **only** strings
        - If subject-per-column format, the file **must** contain row headers in addition to having column headers
            - Row Header indicates the columns
            - Column Header indicates subject

        - Has some indicator value
- Subject identifier ← important
    - By row, must be in first column
    - By column, must be in first row
- Filename for one subject per file
- Files are not saved
    - No download option