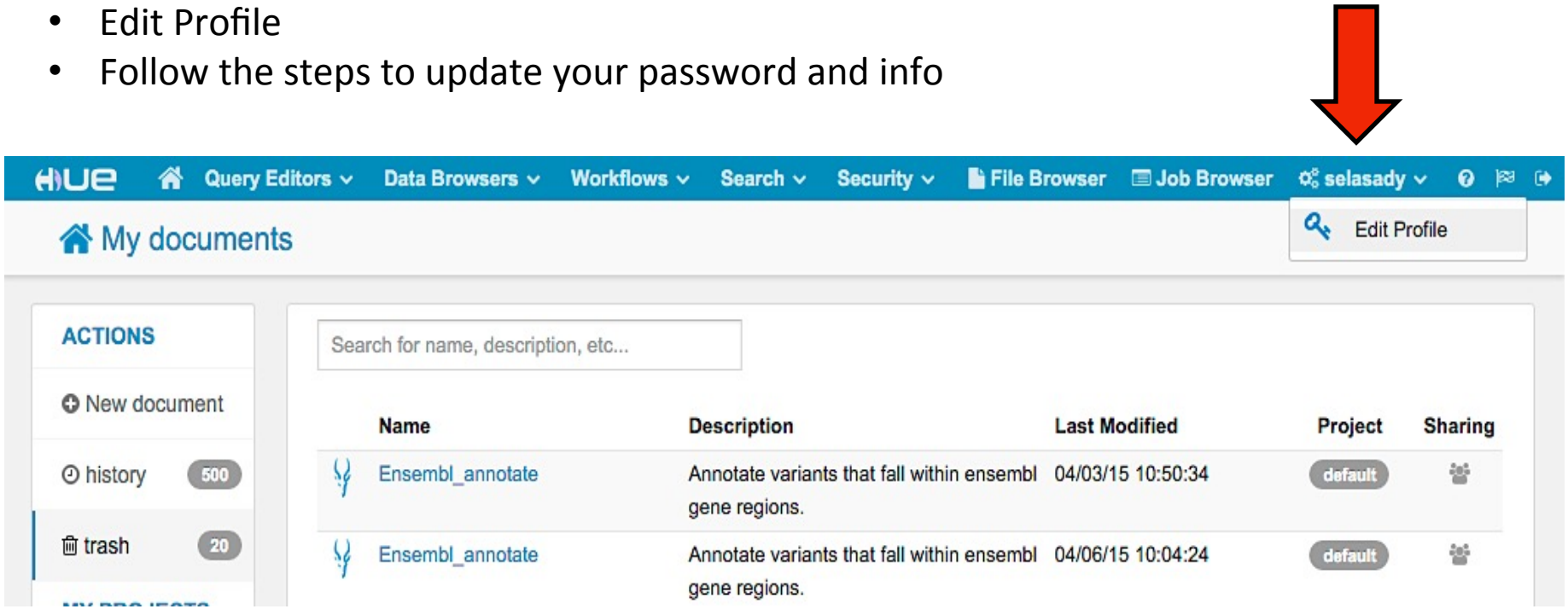
Using the Hue Web Interface

Summer Elasady

Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Updating Account Info

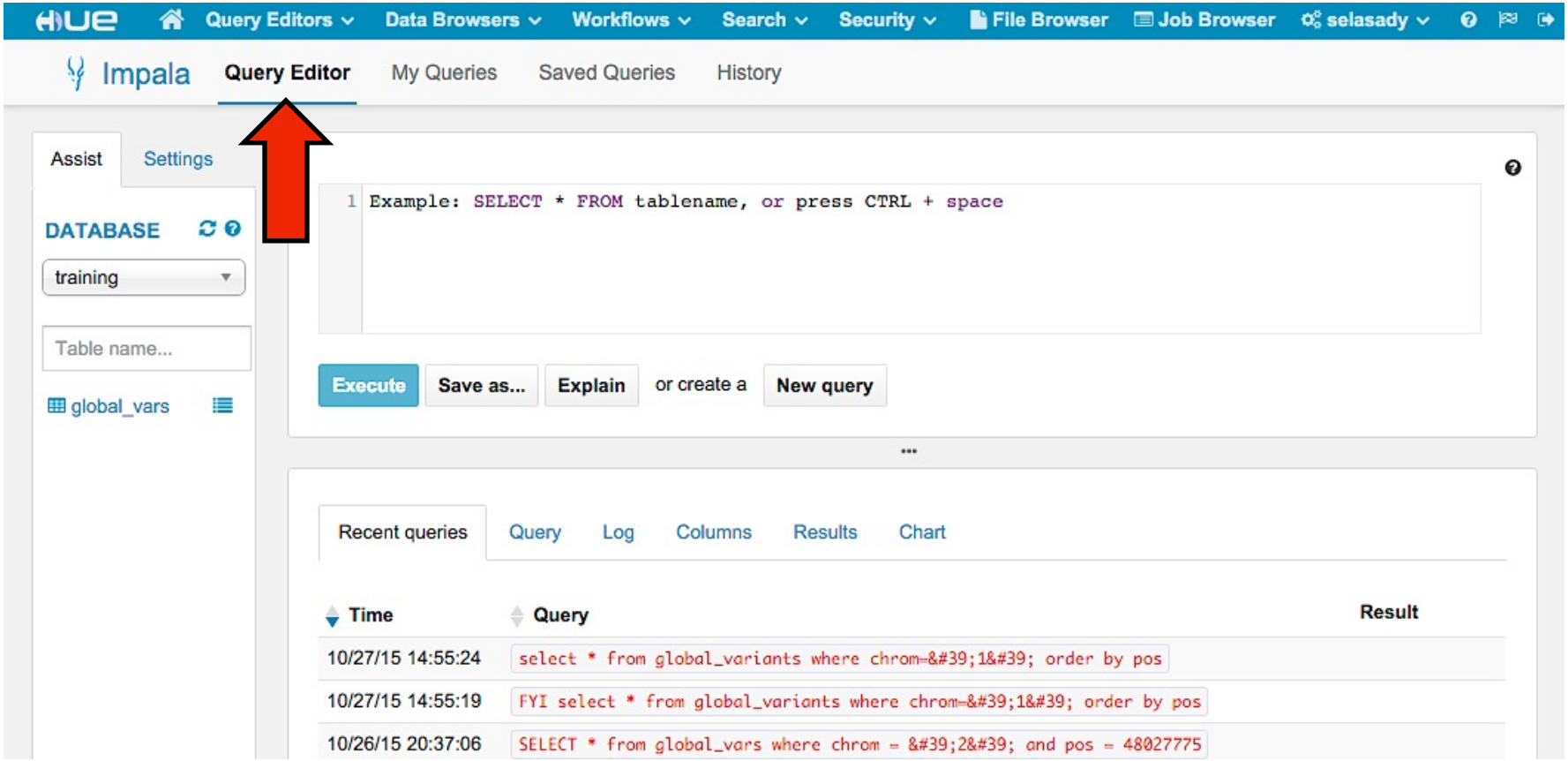
- Click on your user name
- Edit Profile
- Follow the steps to update your password and info



Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Query Editor- Impala

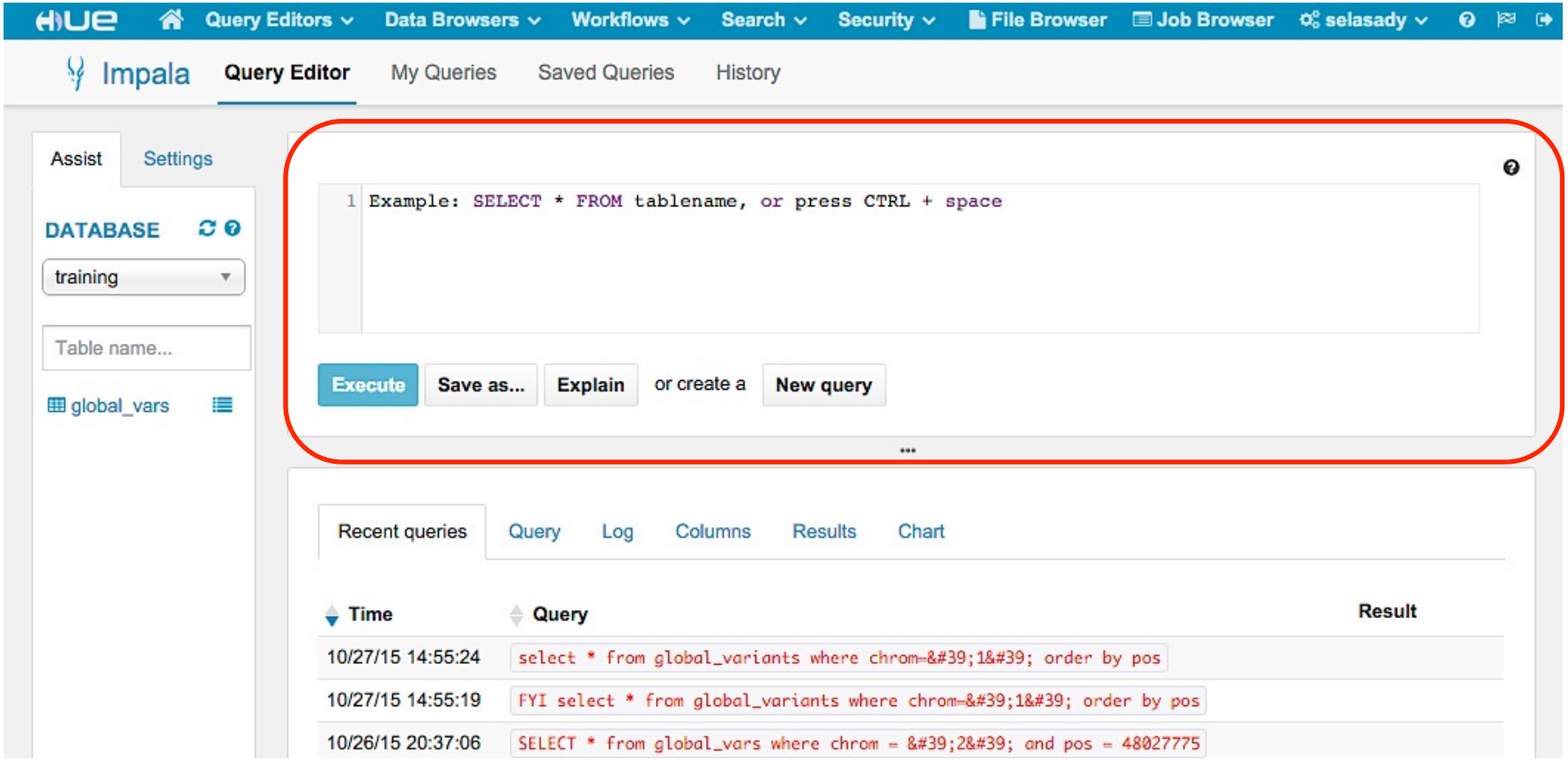
- Click on Query Editor
- Click on impala



Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Query Editor- Impala

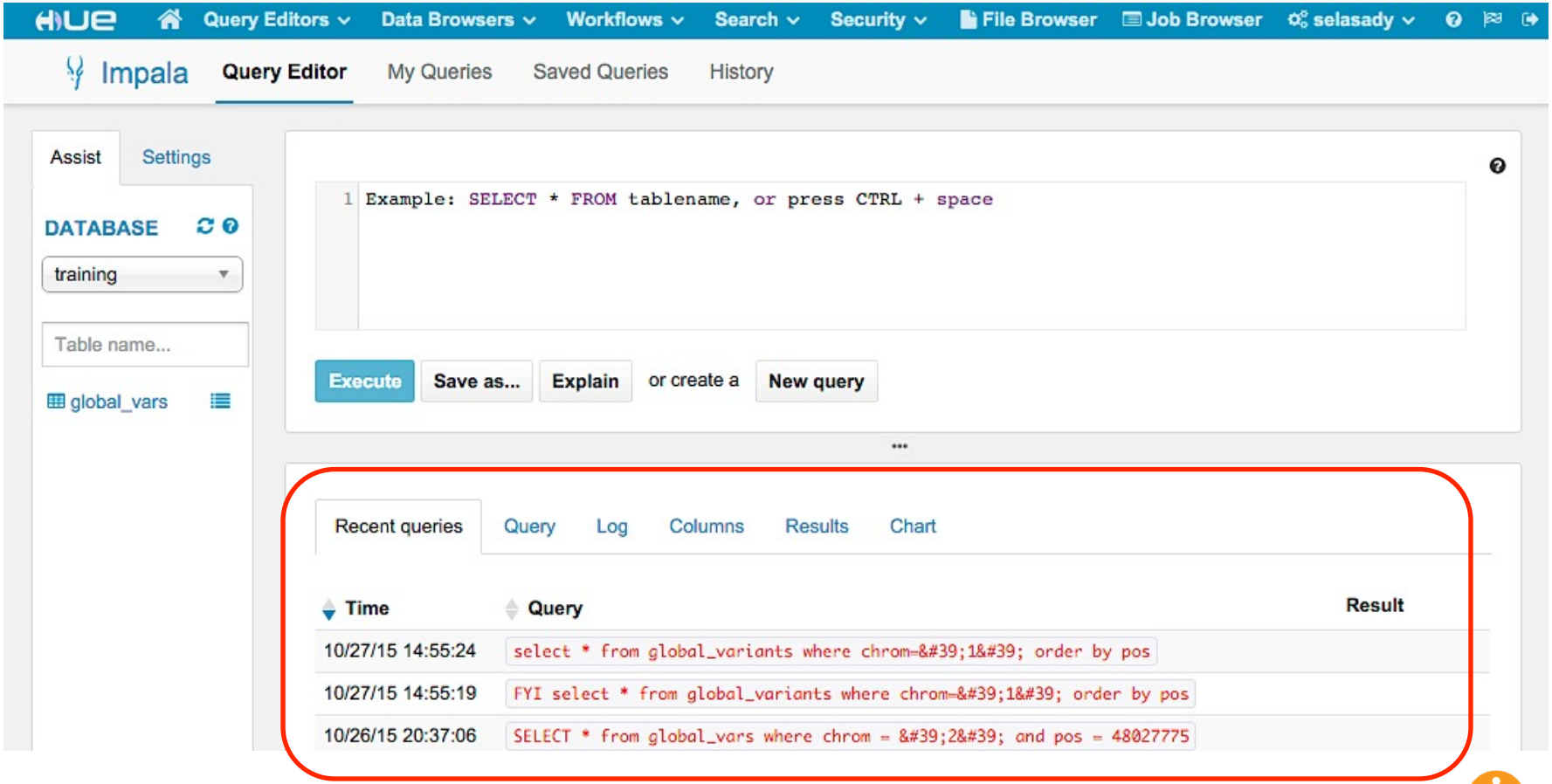
- Query window where you write and execute queries



Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Query Editor- Impala

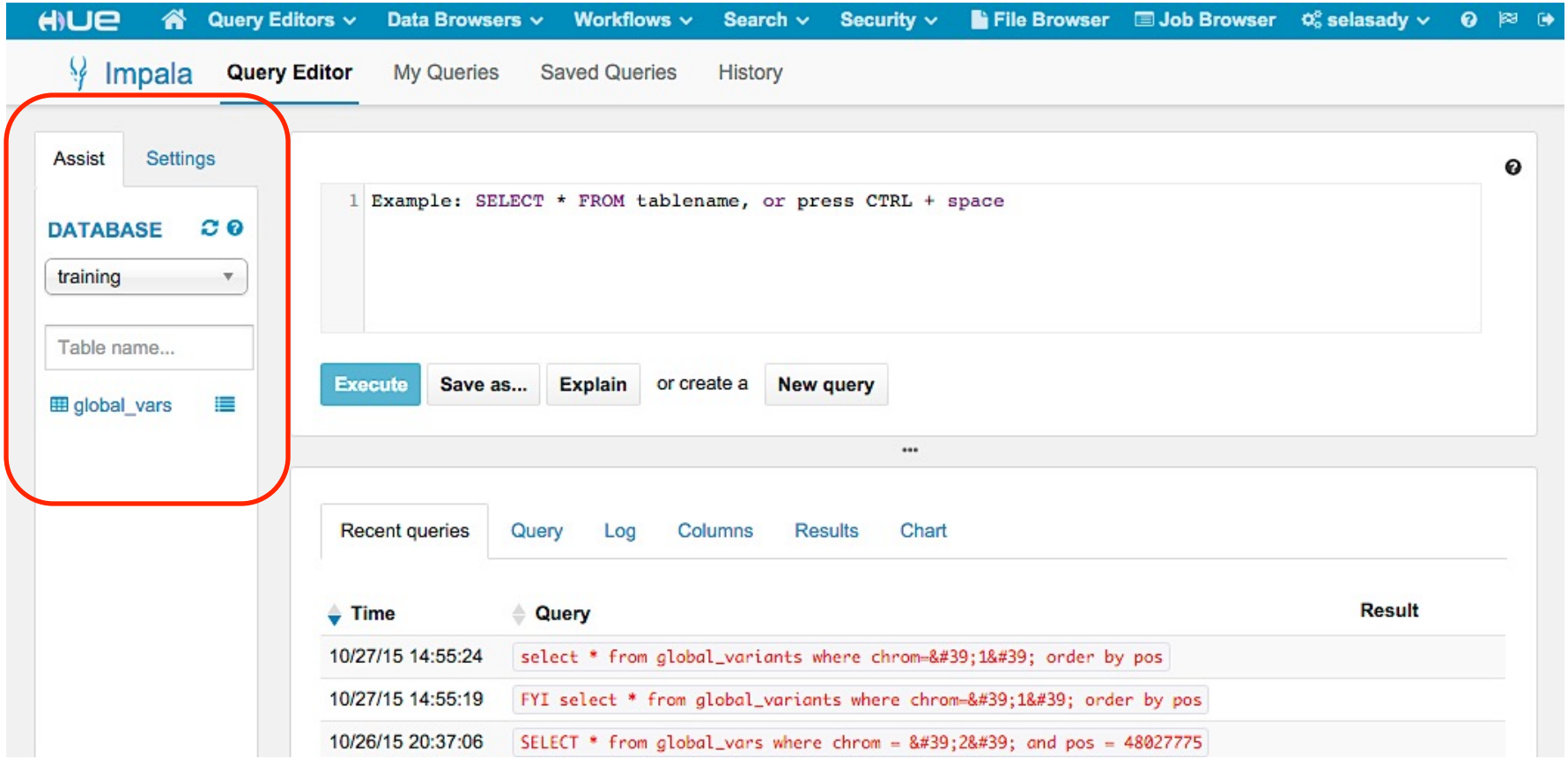
- View recent queries and query results



Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Selecting and Viewing Database

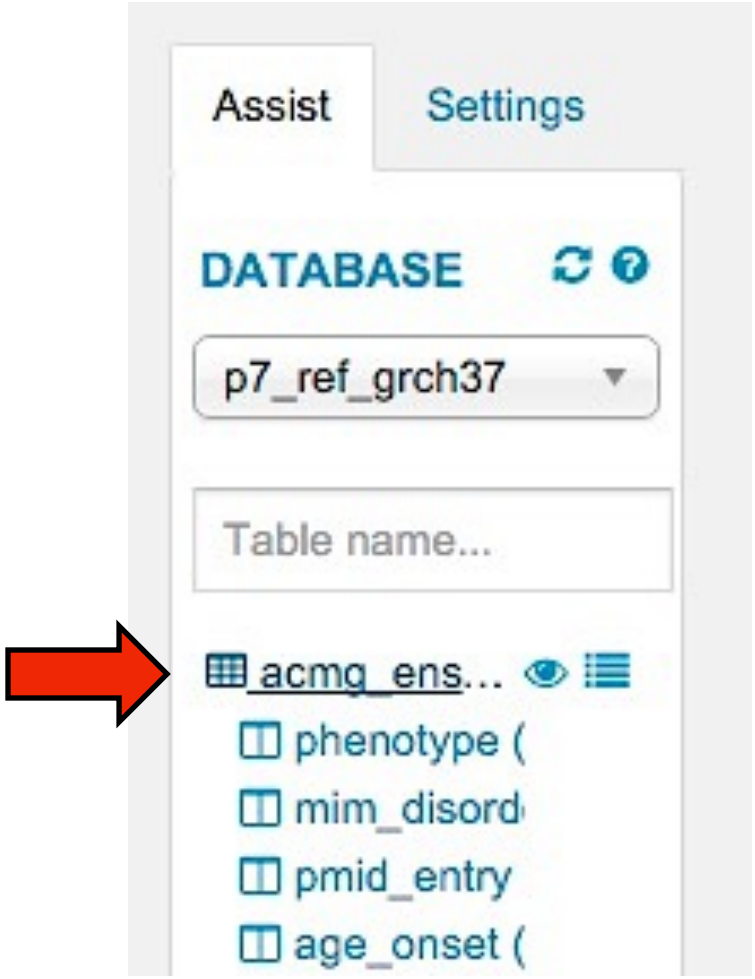
- Click on the drop-down menu to view and select databases



Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

View Database Information

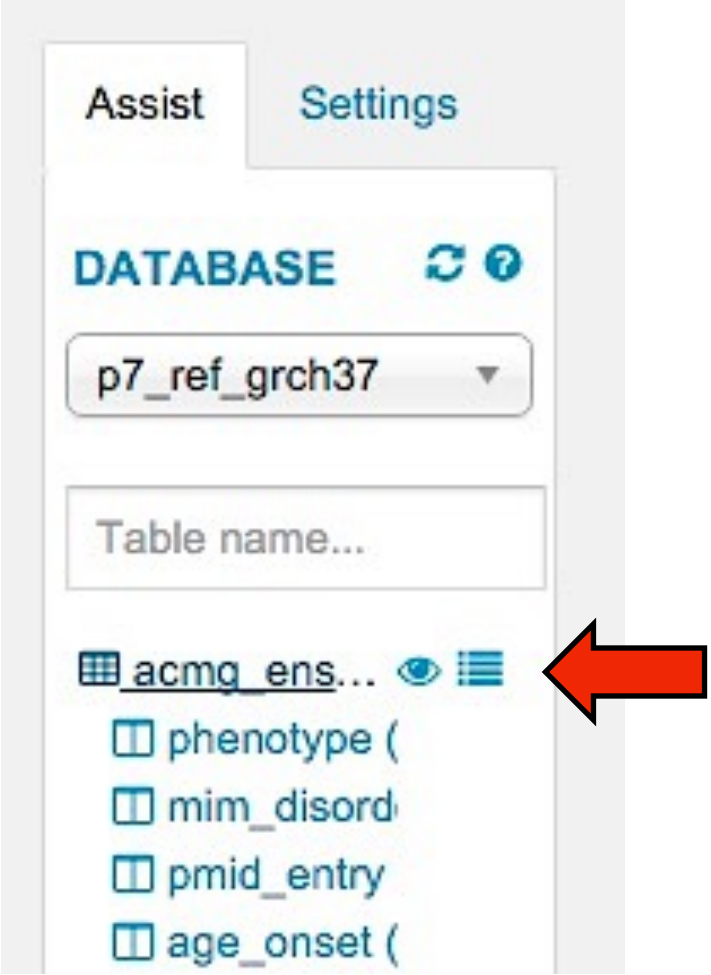
- Click on the database icon to view column names and data types



Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

View Database Information

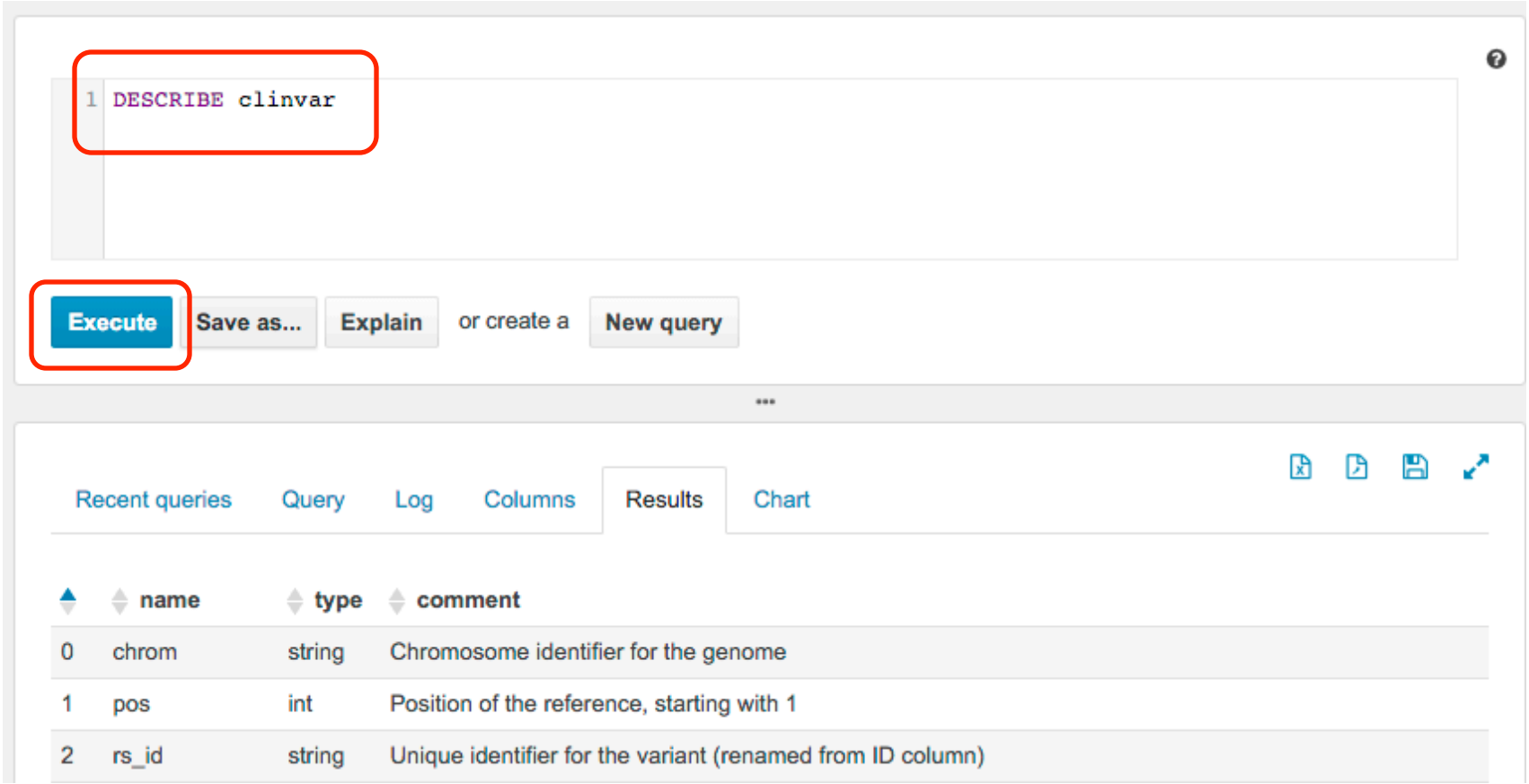
- Click on the eye to view information and a data sample in a new window
- Click on the list icon to view this information in the same window



Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Describe Database Columns

- Type DESECRIBE table_name and click on Execute
- View descriptions of what is in each column



	name	type	comment
0	chrom	string	Chromosome identifier for the genome
1	pos	int	Position of the reference, starting with 1
2	rs_id	string	Unique identifier for the variant (renamed from ID column)

Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Create a Query

- Select the p7_ref_grch37 database
- Enter the query shown
- Click on Explain to review the query plan

```
SELECT chrom, pos, ref, alt, gene
FROM clinvar
WHERE chrom = '1'
GROUP BY chrom, pos, ref, alt, gene
LIMIT 50
```

Execute | Save as... | Explain | or create a | New query

Recent queries | Query | Log | Columns | Results | Chart

	name	type	comment
0	chrom	string	Chromosome identifier for the genome
1	pos	int	Position of the reference, starting with 1
2	rs_id	string	Unique identifier for the variant (renamed from ID column)

Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Review Explain Plan and Execute Query

- Read in numerical order
- Look for ways to optimize
- Identify bottlenecks
- Click Execute to run the query

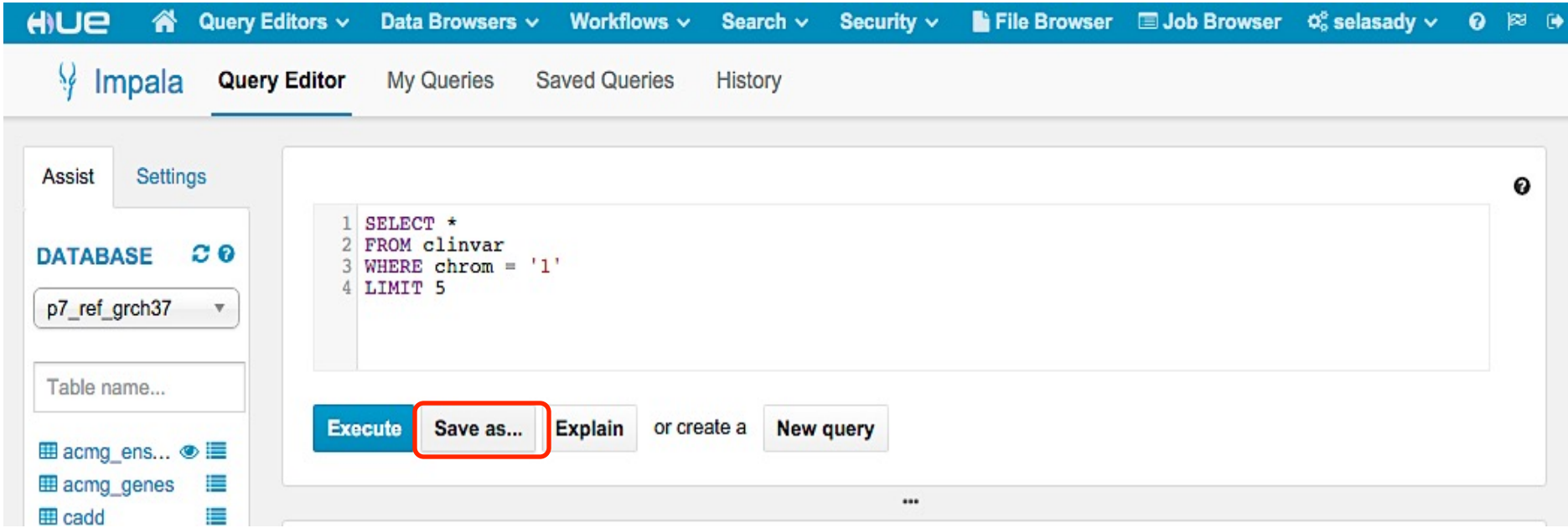
Recent queries | Query | Explanation

```
Estimated Per-Host Requirements: Memory=90.00MB VCores=2

04:EXCHANGE [UNPARTITIONED]
|  limit: 50
|
03:AGGREGATE [FINALIZE]
|  group by: chrom, pos, ref, alt, gene
|  limit: 50
|
02:EXCHANGE [HASH(chrom,pos,ref,alt,gene)]
|
01:AGGREGATE
|  group by: chrom, pos, ref, alt, gene
|
00:SCAN HDFS [p7_ref_grch37.clinvar]
   partitions=1/1 files=1 size=5.41MB
   predicates: chrom = '1'
```

Save Queries

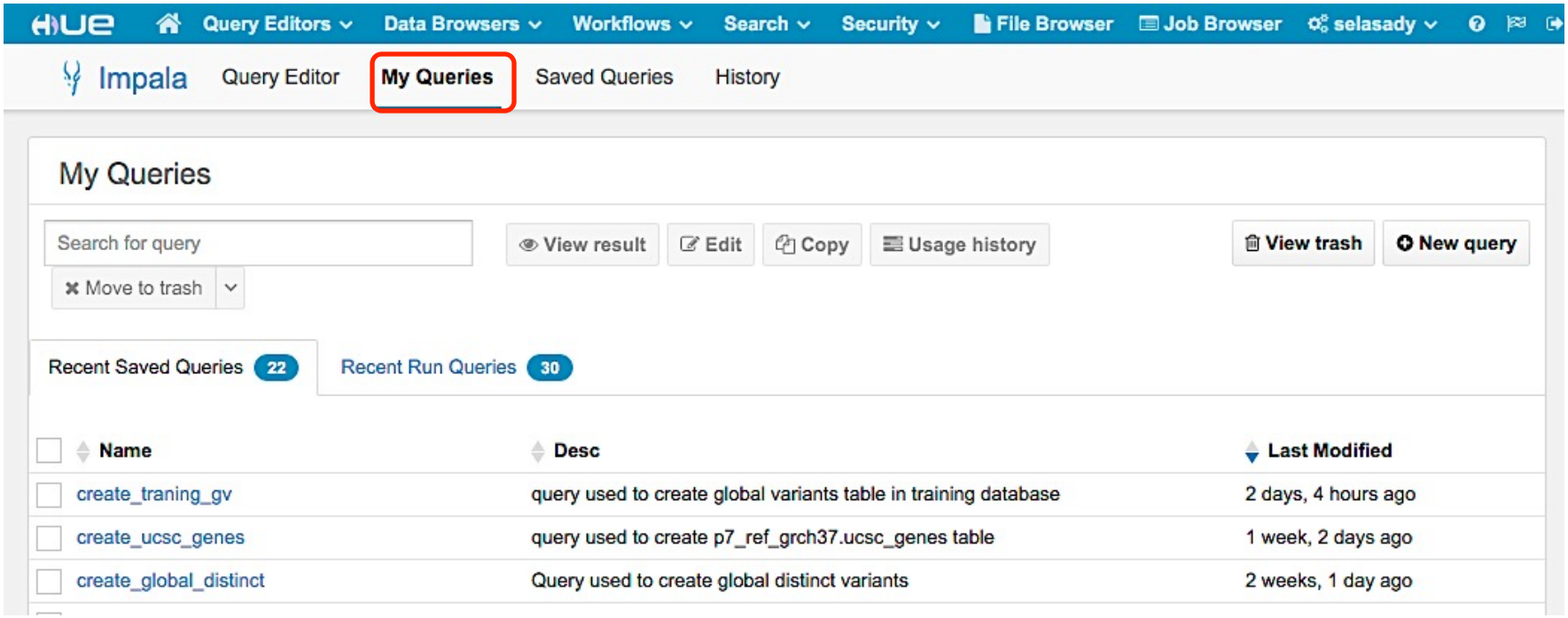
- After executing a query click on Save as
- Give the query a name and description



Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Run Saved Queries

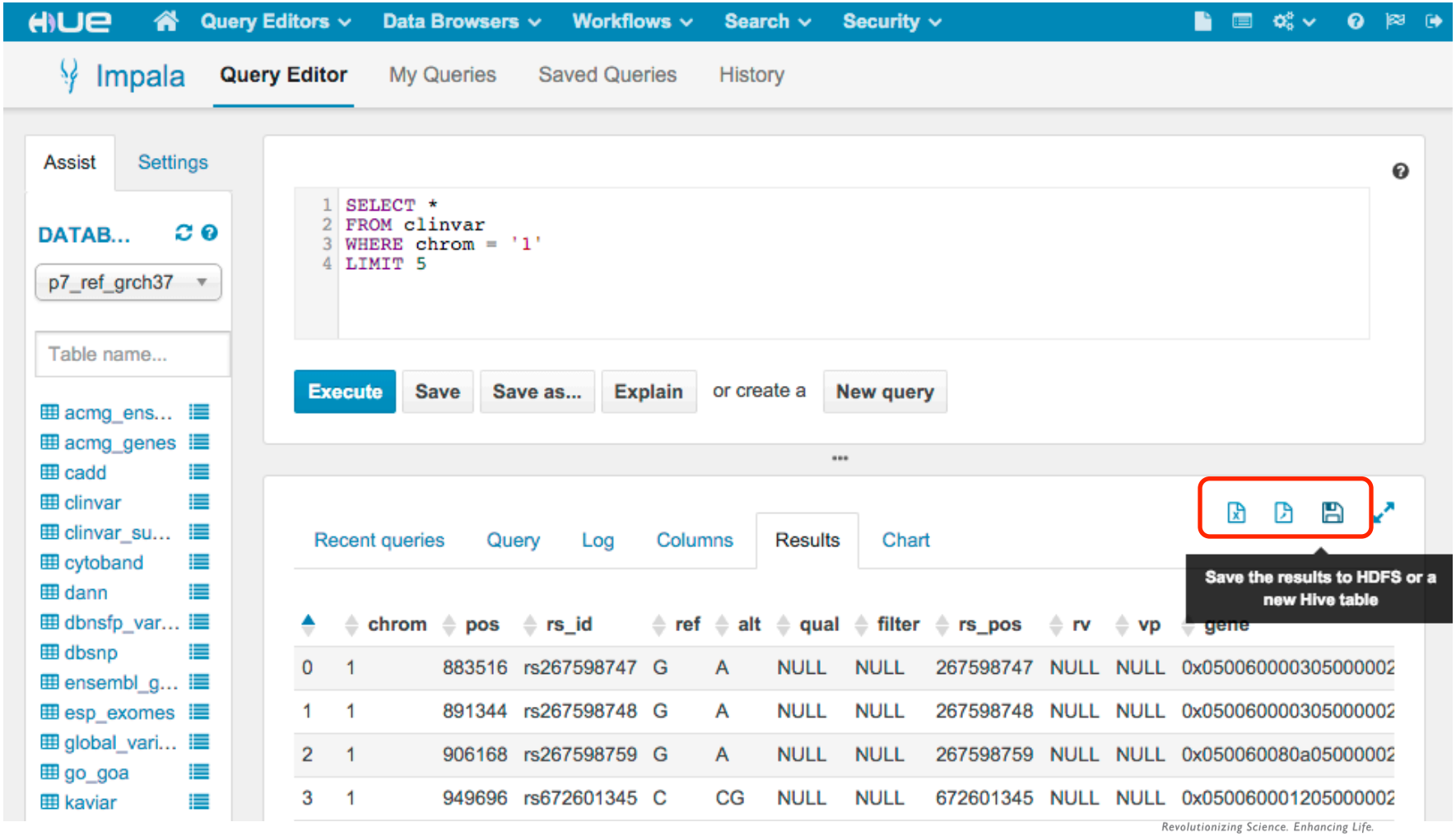
- Click on My Queries to view your saved queries
- Click on the query name to run it in the impala query editor



Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

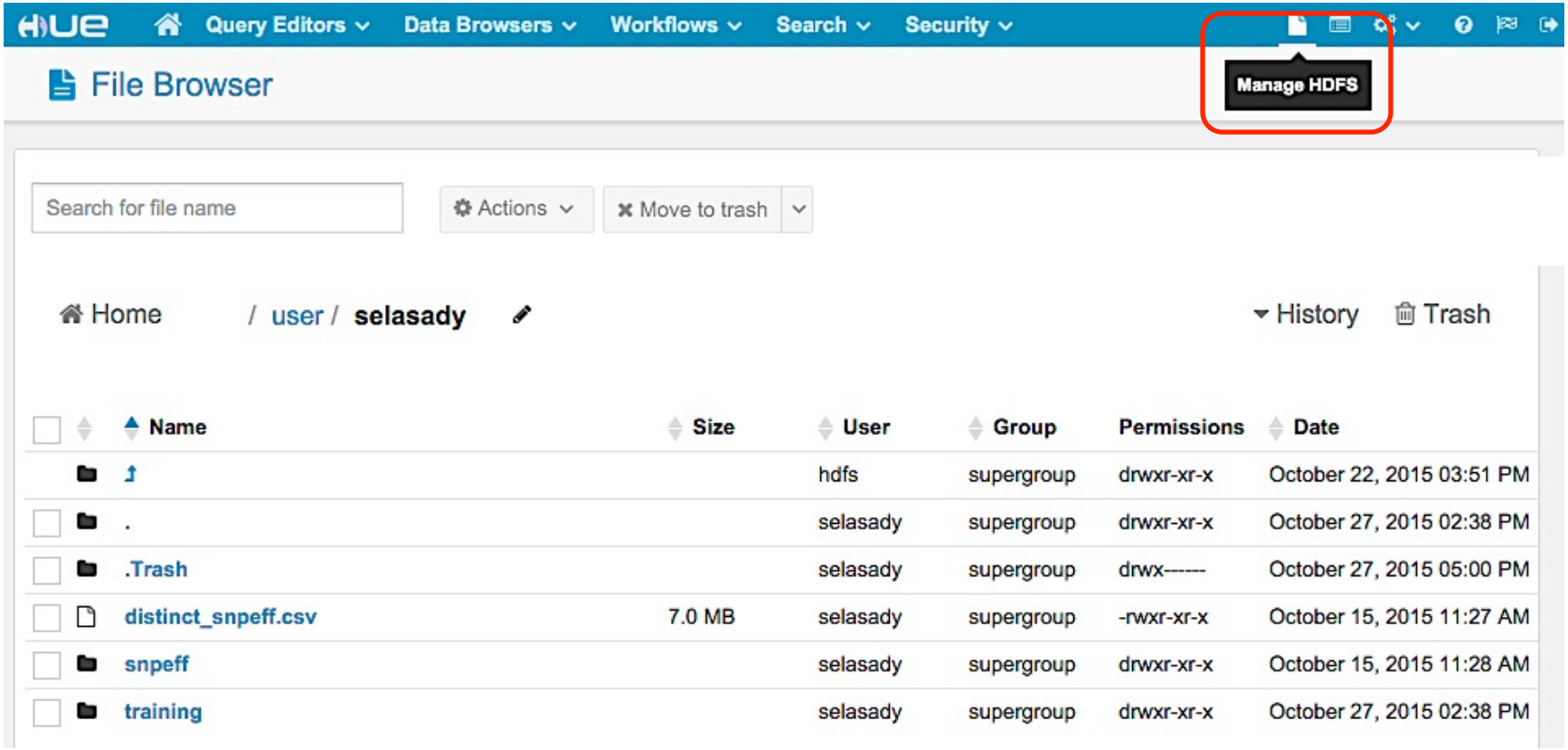
Save Results

- Click on one of the icons shown in the red box to save your results as a table on impala, or to download as a csv or excel file



Working with HDFS

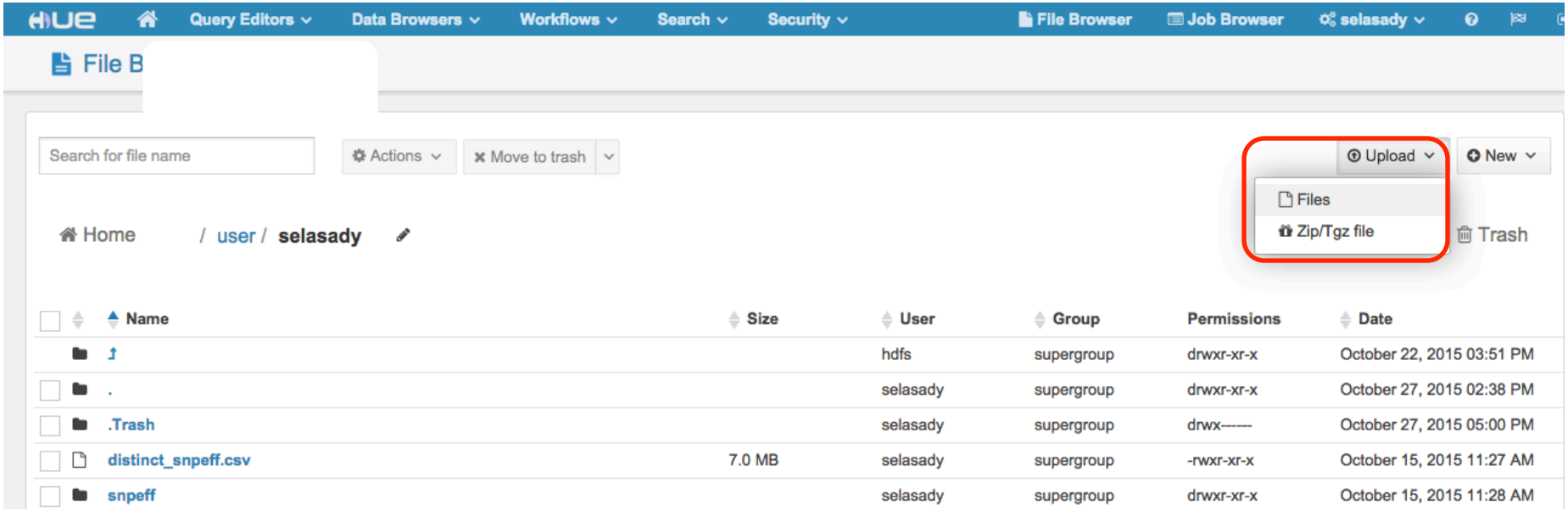
- Click on the file icon
- Click on manage HDFS



Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Uploading a TSV/CSV to HDFS

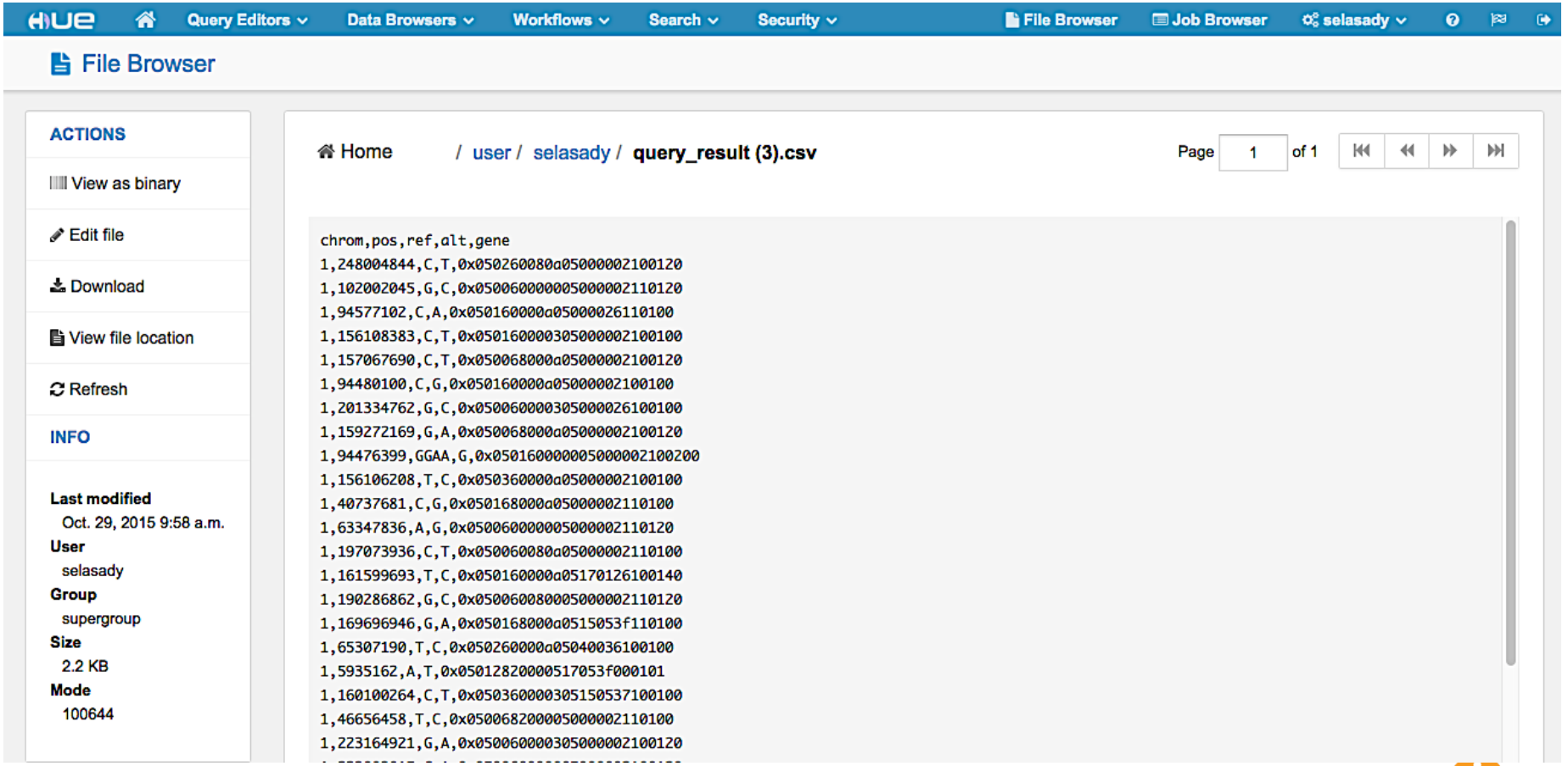
- Click on the file icon
- Click on manage HDFS



Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Viewing uploaded files

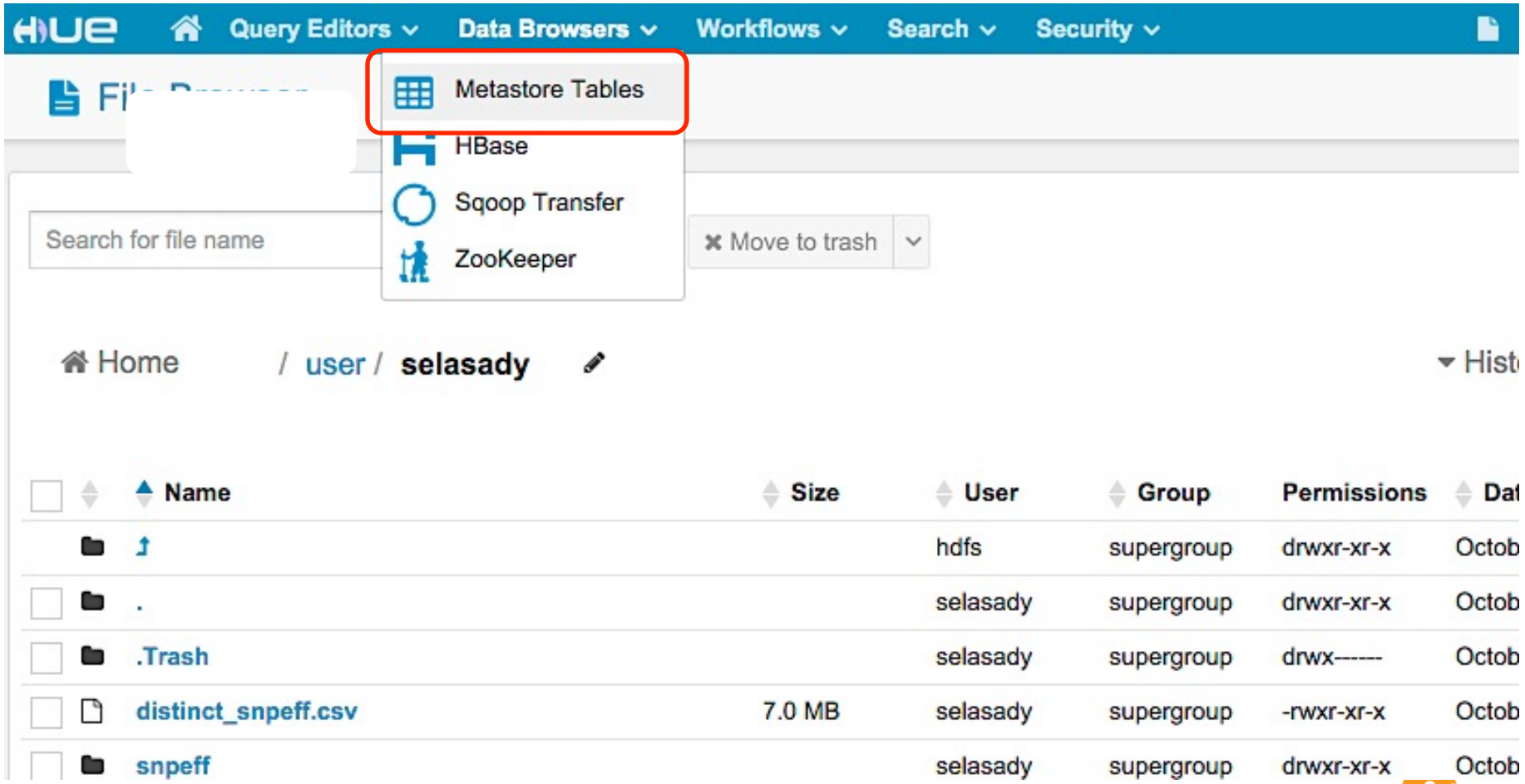
- Once a file is uploaded to HDFS click on that file
- Make sure everything looks correct before converting to a table



Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Converting an HDFS file to a table

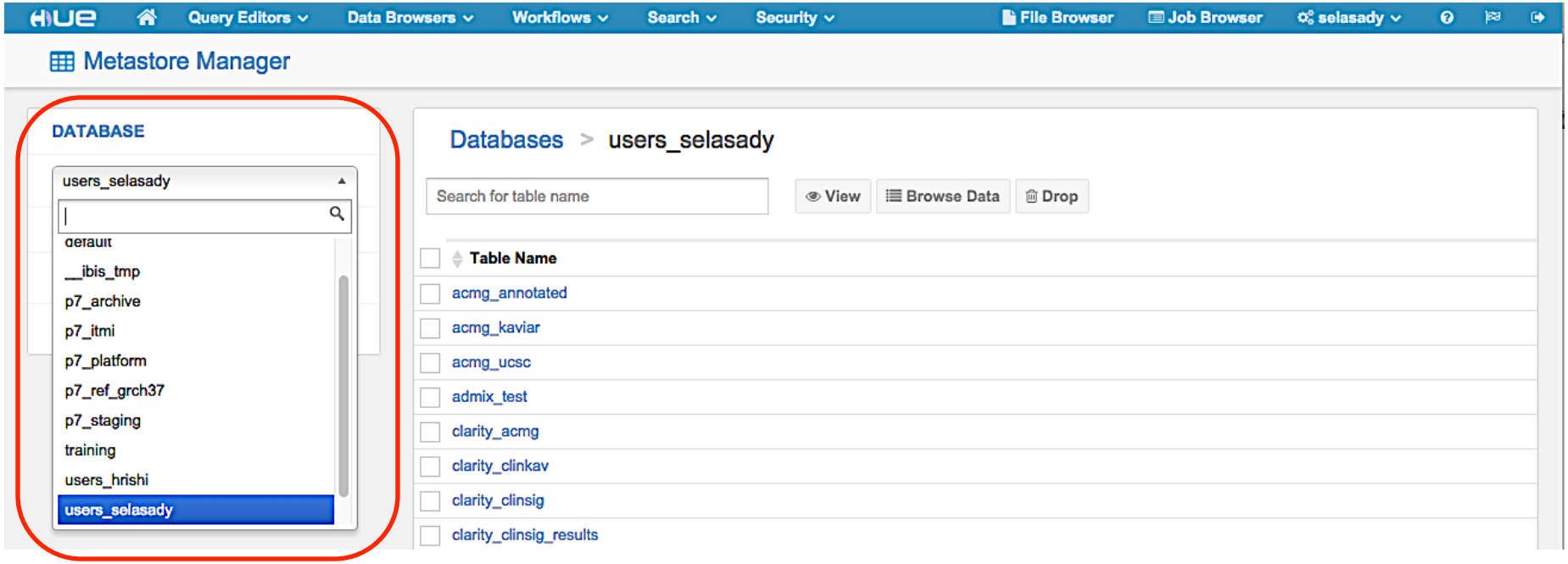
- Once a file has been uploaded to HDFS, click on Data Browses
- Click on Metastore Tables



Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

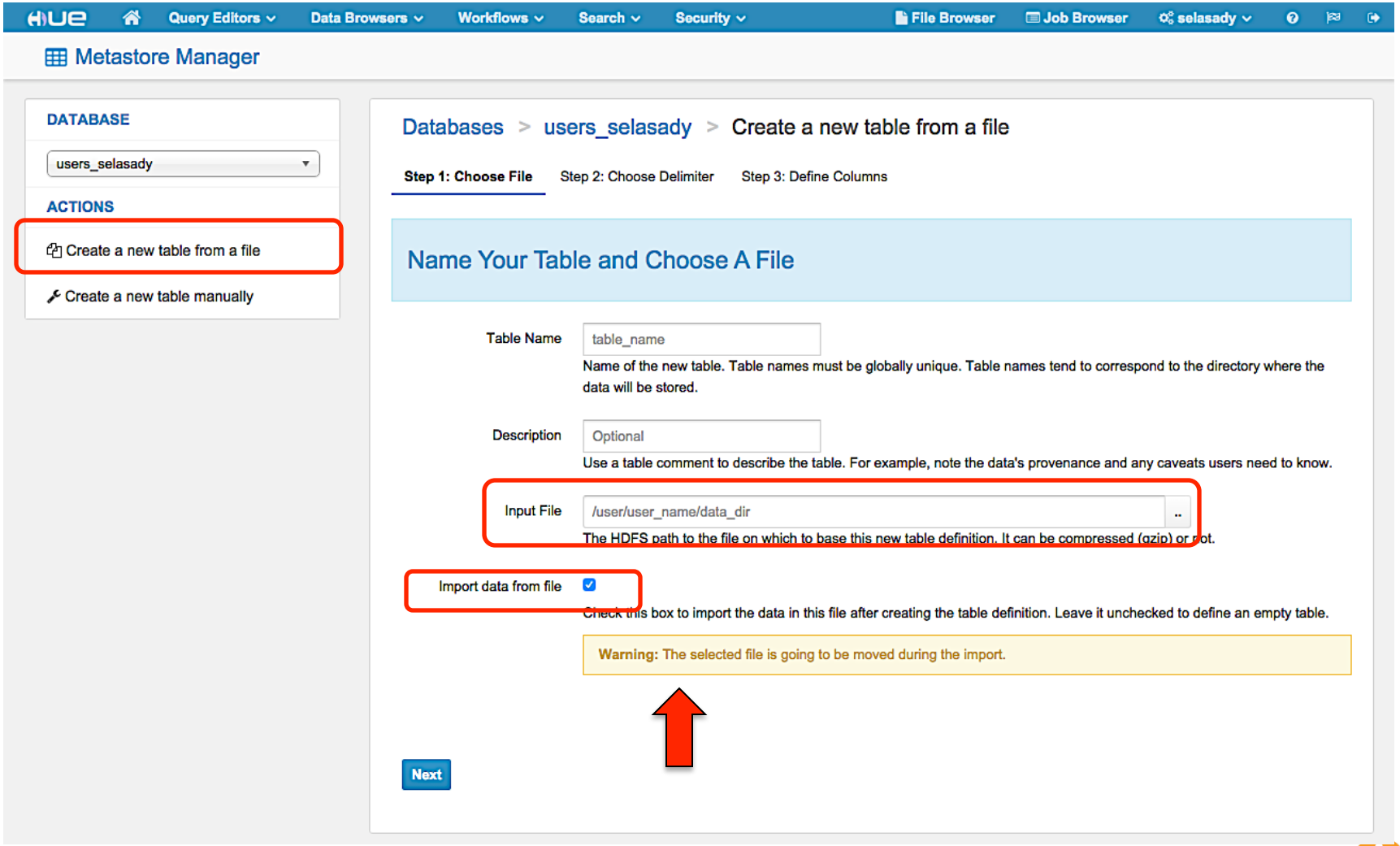
Select a database to place the table

- Click on the file icon
- Click on manage HDFS



Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

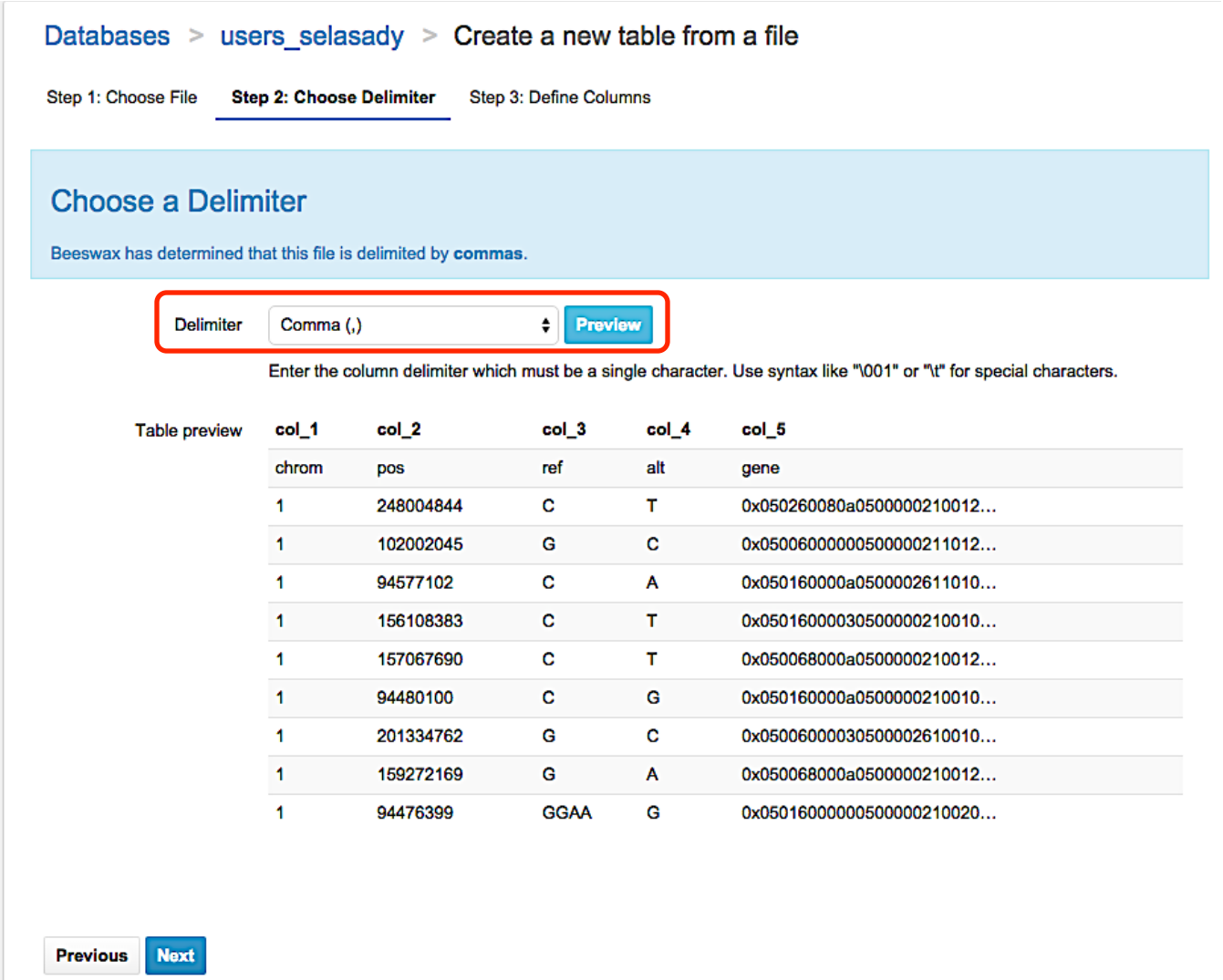
Create a new table from a file



Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Choose your delimiter

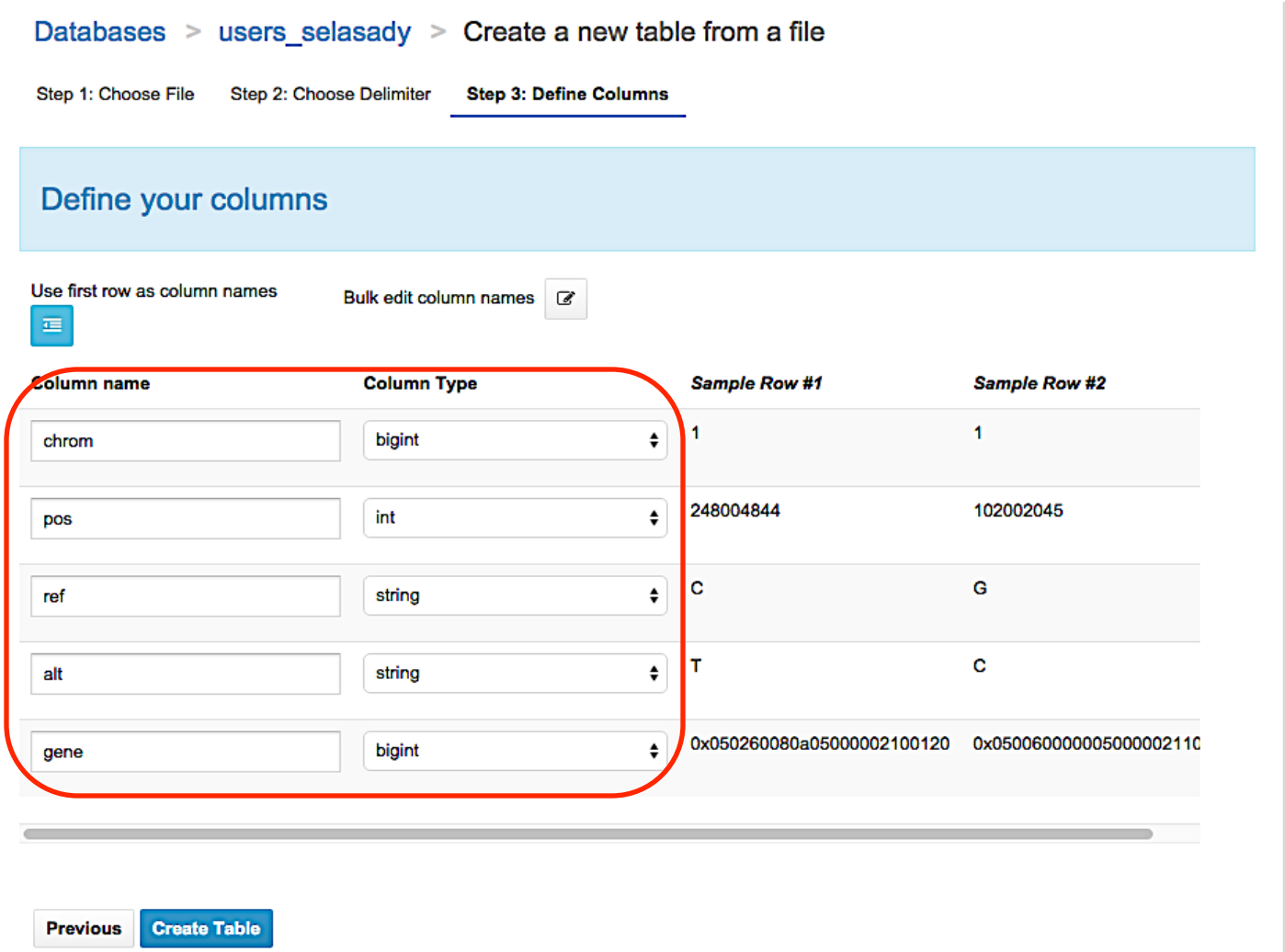
- Choose how your file is delimited



Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Define your columns

- Check that the proper data type is selected for each column



Institute for Systems Biology
Revolutionizing Science. Enhancing Life.