Computer Assignment 1

Project A: Mass-metallicity relation

Deadline for the report: **13-10-2023 17:00h**
Upload your report in Assignments on Brightspace

The following tasks are meant to give you a familiarity with the world of astronomical databases and on their use to address relevant questions in galaxy evolution. The focus will be on SDSS data and the information about the dataset are available on classic.sdss.org/dr7/ .

This is based largely on a Lab developed by J. Brinchmann at Leiden Observatory, to whom we are very grateful.

1 CasJobs

A. Create a CasJobs account: go to https://skyserver.sdss.org/casjobs/ and click on "Login" and then "create a new account" (Fig. 1).

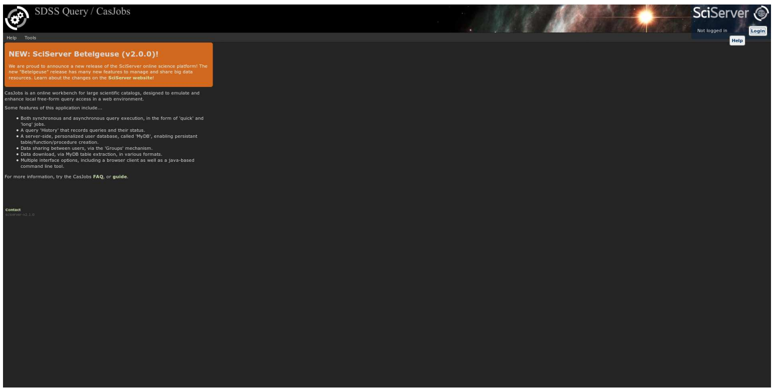


Figure 1: Home page of CasJobs

B. We will work with **Tables** and you can check the contents (and the syntax!) of every table in the **Schema Browser** (https://skyserver.sdss.org/CasJobs/SchemaBrowser.aspx). In particular we will work with the Table **SpecObjAll**, therefore before starting be sure to have read the description of this table.

C. Now you can create your **first query**.
We can start with a simple one: choose the first 100 objects in SpecObj table that is within 0.001 and z=0.04 **(be sure to put the context to dr7)**:

```
SELECT TOP 100 z, ra, dec
FROM SpecObj
WHERE abs(z-0.04)< 0.001
```

If you prefer you can run also the examples provided in CasJobs from the drop-down menu **Samples** (Fig. 2) for some tips on query syntax, but make sure that it does not take too much time. You can select a fixed number of objects with the command TOP, as done in the previous example.

D. While the first query is running click on the **History** link to be informed on its progresses.

E. If it has finished, you can go in **MyDB** and verify if there is a new table there. If you download a CSV or FITS file, then you will be able to explore the results.

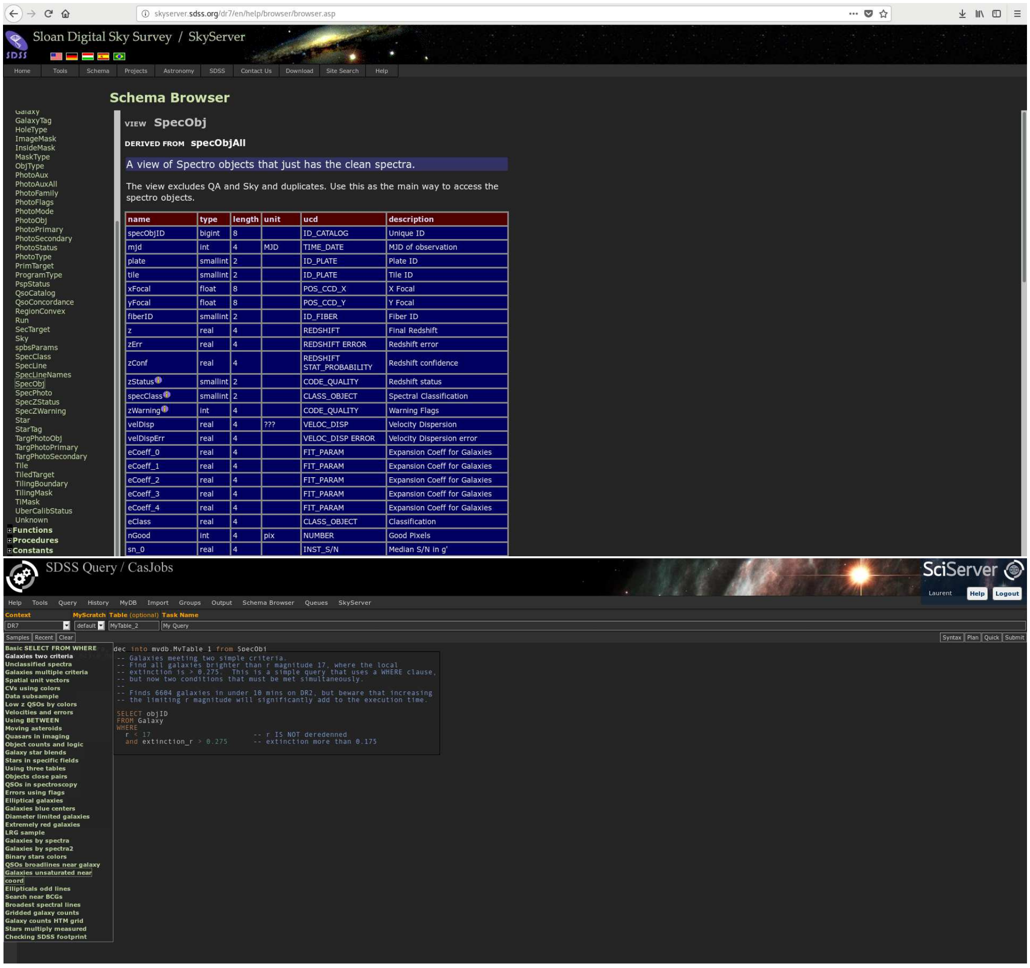
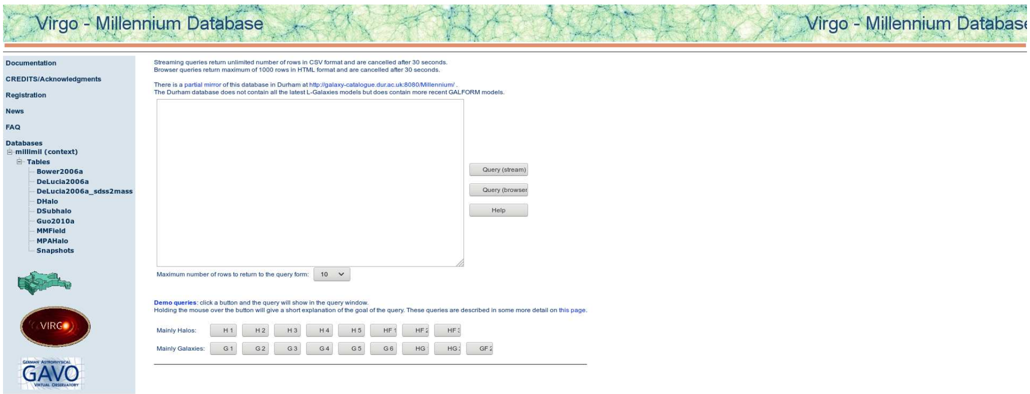


Figure 2: Contents of SpecObj (top); Samples of queries in CasJobs (bot).

2 SDSS data and Millennium simulation

You will need CasJobs (Sec.1) and the database interface to the **Millennium simulation** (http://gavo.mpa-garching.mpg.de/Millennium). The goal of this project is to compare the predictions from the Millennium simulation with the observations in the SDSS. We will focus on the quantitative trends.



In the SDSS CAS we are not given the stellar mass for the galaxies, so we need to use the luminosity of objects and we need the same in the Millennium simulation.

For the calculations you can use the Ned Wright's cosmology calculator (https://www.astro.ucla.edu/ wright/CosmoCalc.html). Note that you need to agree on a consistent cosmology (use $H_0 = 70$km/s/Mpc, $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$).

It is probably easier to work with absolute magnitudes rather than luminosities, so you need to recall that the absolute magnitude is given by $M = m - 25 - 5\log(d_L(z)/Mpc)$, where $d_L(z)$ is the luminosity distance at redshift z and m the apparent magnitude.

If you need it, you can assume that the absolute magnitude of the sun in the r-band is $M_{\odot,r} = 4.76$ and hence that the luminosity of an object in solar units is given by $L/L_\odot = 10^{-0.4(M_r - M_{\odot,r})}$.

Feel free to compare results between different models, but you should focus on DeLucia2006a and DeLucia2006a-sdss2mass tables.

2.1 Project A - Mass-metallicity relation

2.1.1 The theoretical mass-metallicity relation

Your theoretical work will be to extract information from the Millennium database to predict the relationship between r-band luminosity and the metallicity of the systems.

- You should use the DeLucia 2006 catalogue and extract out the mass in metals and stars and other quantities of interest. Select the model in the snapshot number that is closest to z=0.1.

Please refer to the **Help** section on the Millenium database website for a detailed instruction on how to write queries and an overview of query parameters, although the syntax is very similar to CAS.

In the observation part, you will provide abundances of oxygen relative to hydrogen. Therefore, you will need to transform the theoretical calculation into these units. In fact it is more confusing because in the observation section, you will use peculiar units where by actually calculating 12 + log O/H, where O/H is the oxygen abundance relative to the hydrogen abundance.

Remember first that the metallicity is the amount of metal divided by total mass. Secondly, the solar metallicity can be approximated by Z=0.02 for our work here (it is probably a bit lower) and the oxygen abundance of the sun can be taken to be 12 + Log O/H = 8.8 (it is probably somewhere between 8.65 and 8.9 but discussions continue - see Asplund et al. 2009). This should now allow you to compile a list of predicted absolute r-band magnitudes and oxygen abundance.

2.1.2 The observed mass-metallicity relation

To estimate metallicities observationally we will make use of a simple relation due to Pettini & Pagel (2004). They found that the oxygen abundance in nearby objects could be approximately found by taking the ratio of the flux in the nitrogen forbidden line at 6585 Å ([N II]6585) to that of Hα at 6565 Å:

$$12 + \log(O/H) = 8.9 + 0.57 \log\left(\frac{f_{[NII]}}{f_{H\alpha}}\right) \tag{1}$$

You will use this to calculate abundances and you also need the photometry of the objects to calculate their absolute magnitude. The goal here is for you to produce a diagram that plots absolute r-band magnitude on the x-axis and the oxygen abundance on the y-axis. Select your objects from the SDSS using CasJobs - SpecPhoto might be a good table to use. However to get the line fluxes you need to JOIN with the SpecLine table. Here is a simple example to get the Hβ line.

```
SELECT s.SpecObjID, s.z –we want the photometric objID
FROM SpecObj as s
JOIN SpecLine as L ON s.SpecObjID = L.specObjID –line L is detected in spectrum
WHERE –you could add a constraint that the spectral type is a galaxy
L.LineId =4863 –and the line L is the Hβ line
```

Your challenge is to modify this to get Hα, [N II] (their LineIDs corresponds to their central wavelength) and ensure that *the S/N is greater than 5 in each line*. You can either use the equivalent width or the height of the line as a measure of the flux for the abundance calculation. If you decide to do everything in SQL (which is easy), you might need to avoid division by zero - the best way to do this is to include a check in your calculation like:

```
L.ew/nullif(L.ewErr,0)>5
```

Make sure you also get the r-band magnitude and absolute magnitude in the r-band. Typical oxygen abundances are $12 + \log(O/H) = 8$ to 9 and typical absolute magnitudes probably between -15 and -25.

Get the theoretical predictions and compare to your results. How do they look? Are they the same? Different? Do they agree with each other? Why, or why not?

3 Suggestions

1. Give a reasonable name to your tables, so you can easily recognize every table. You can do it directly in the query using the command INTO.
2. In order to recognize the components of a table from the command, write the SQL keywords in uppercase (common practice in SQL).
3. Before submitting a query, check your query syntax using the botton SYNTAX.
4. Be sure that in your table there are some elements, otherwise something went wrong. You can download your table from MyDB page (Fig.3).

```
SELECT TOP 200 s.specObjID, s.z, L.ew
INTO mydb.Exercise1 FROM SpecPhoto AS s
JOIN SpecLine AS L ON s.specObjID = L.specObjID
WHERE L.LineID = 0000
AND L.lineID = 0000
AND L.ew > 0000
```

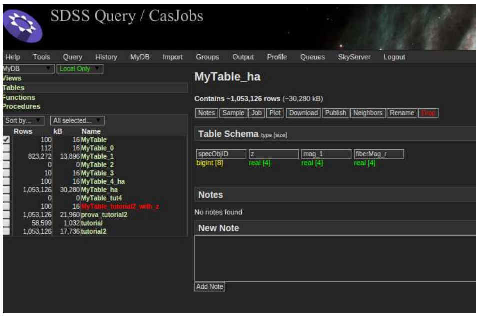


Figure 3: MyDB page: there are tables (left) without elements because of syntax errors.

4 The report

Deadline: 13-10-2023 17:00h
Upload your report in Assignments on Brightspace

Your report will be on the mass-metallicity relation (if you chose Project A) or SFR-luminosity (if you chose Project B). The report should be divided in the following sections:

- *Introduction*: here you describe the astrophysical background and the databases, with particular regard of the meaningful aspects for this project (max. 1 page);
- *Results*: here you discuss how you selected the data from SDSS and Millennium, how you used these to study the relations (plots!);
- *Discussion*: in this section there will be your discussion of the results. For example, is there a correlation between mass and metallicity in the subsample you have chosen? Does it have the same slope of the theoretical one? Is there any change if you select different type of galaxies (spirals, ellipticals etc..)? Is the scatter of the observed relation the same of the theoretical relation?

Please make your report 3-4 pages, including plots. The minimum fontsize should be 11 pt (smaller fonts are allowed for figure captions). You can add tables, figures and whatever you need to make your report complete, but do not exceed the page limit. The emphasis should be on how the results and how you interpret them. Do not forget to write your name and a title. Lastly, please include with your report the data you downloaded and a working example of the code you used for the assignment (e.g. a python script, jupyter notebook).

Computer Assignment 1

Project B: SFR-luminosity relation

Deadline for the report: **13-10-2023 17:00h**
Upload your report in Assignments on Brightspace

The following tasks are meant to give you a familiarity with the world of astronomical databases and on their use to address relevant questions in galaxy evolution. The focus will be on SDSS data and the information about the dataset are available on classic.sdss.org/dr7/ .

This is based largely on a Lab developed by J. Brinchmann at Leiden Observatory, to whom we are very grateful.

1 CasJobs

A. Create a CasJobs account: go to https://skyserver.sdss.org/casjobs/ and click on "Login" and then "create a new account" (Fig. 1).

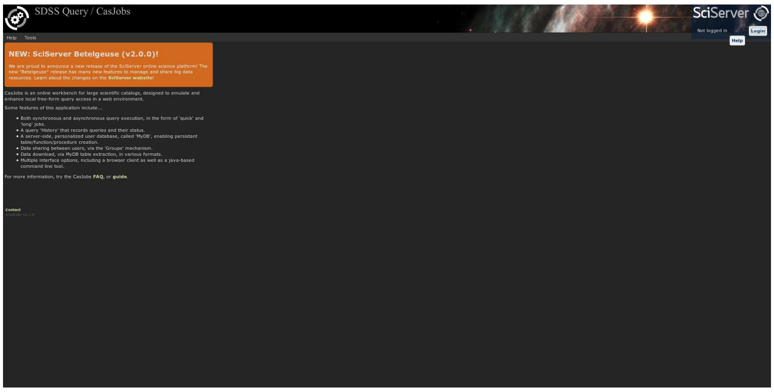


Figure 1: Home page of CasJobs

B. We will work with **Tables** and you can check the contents (and the syntax!) of every table in the **Schema Browser** (https://skyserver.sdss.org/CasJobs/SchemaBrowser.aspx). In particular we will work with the Table **SpecObjAll**, therefore before starting be sure to have read the description of this table.

C. Now you can create your **first query**.
We can start with a simple one: choose the first 100 objects in SpecObjAll table that is within 0.001 and z=0.04 **(be sure to put the context to dr7)**:

```
SELECT TOP 100 z, ra, dec
FROM SpecObj
WHERE abs(z-0.04)< 0.001
```

If you prefer you can run also the examples provided in CasJobs from the drop-down menu **Samples** (Fig. 2) for some tips on query syntax, but make sure that it does not take too much time. You can select a fixed number of objects with the command TOP, as done in the previous example.

D. While the first query is running click on the **History** link to be informed on its progresses.

E. If it has finished, you can go in **MyDB** and verify if there is a new table there. If you download a CSV or FITS file, then you will be able to explore the results.

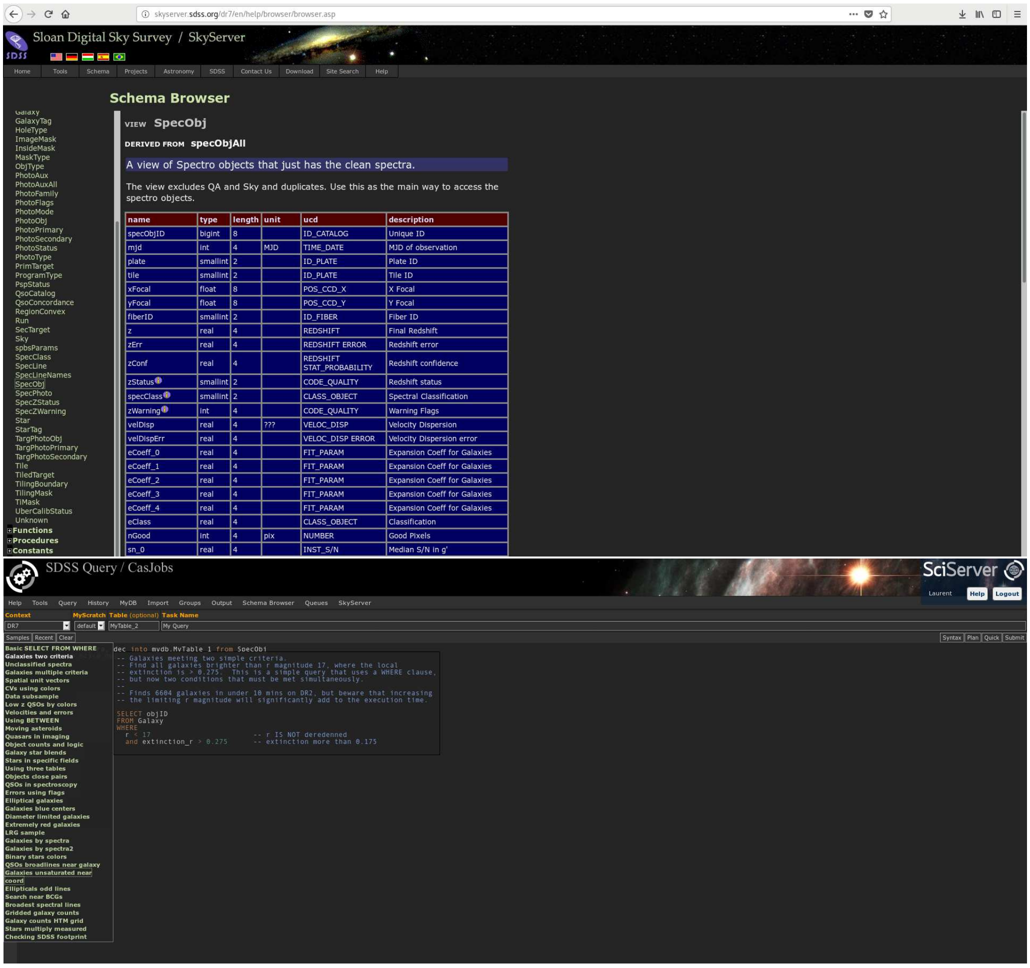
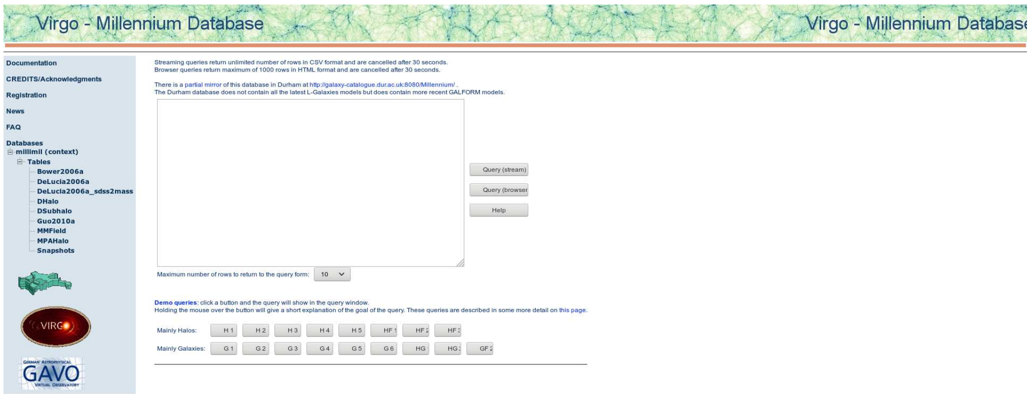


Figure 2: Contents of SpecObj (top); Samples of queries in CasJobs (bot).

2 SDSS data and Millennium simulation

You will need CasJobs (Sec.1) and the database interface to the **Millennium simulation** (http://gavo.mpa-garching.mpg.de/Millennium). The goal of this project is to compare the predictions from the Millennium simulation with the observations in the SDSS. We will focus on the quantitative trends.



In the SDSS CAS we are not given the stellar mass for the galaxies, so we need to use the luminosity of objects and we need the same in the Millennium simulation.

For the calculations you can use the Ned Wright's cosmology calculator (https://www.astro.ucla.edu/ wright/CosmoCalc.html). Note that you need to agree on a consistent cosmology (use $H_0 = 70$km/s/Mpc, $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$).

It is probably easier to work with absolute magnitudes rather than luminosities, so you need to recall that the absolute magnitude is given by $M = m - 25 - 5\log(d_L(z)/Mpc)$, where $d_L(z)$ is the luminosity distance at redshift z and m the apparent magnitude.

If you need it, you can assume that the absolute magnitude of the sun in the r-band is $M_{\odot,r} = 4.76$ and hence that the luminosity of an object in solar units is given by $L/L_\odot = 10^{-0.4(M_r - M_{\odot,r})}$.

Feel free to compare results between different models, but you should focus on DeLucia2006a and DeLucia2006a-sdss2mass tables.

2.1 Project B - SFR-Luminosity relation

2.1.1 The theoretical SFR vs L

You will create the theoretical predictions for the star formation rate (SFR) vs luminosity relation.

1. Assemble a catalogue from the Millennium simulation with various quantities, but at least the stellar mass, the SFR and the absolute magnitude in several bands. Select the model in the snapshot number that is closest to z=0.1.

2. Plot the log of the stellar mass against the absolute r-band magnitudes (is there a good correlation? Does it break down?) In Bell & De Jong (2001) you can find a popular prescription to estimate the M/L, which is $\log(M/L) = -0.565 + 1.132(B - V)$. Does it work for your case as well?

2.1.2 The observed SFR vs L

Your task here is to calculate the luminosity in the r-band, the SFR of a galaxy, and compare this to theoretical predictions. To do this we will want to use the luminosity of the Hα line. There is then a relationship between this and the star formation rate of the galaxy given by (Kennicutt 1998):

$$\mathrm{SFR}(M_\odot/yr) = 7.9 \times 10^{-42} L(H\alpha)[\mathrm{erg/s}], \tag{1}$$

where the IMF is assumed to be a Salpeter IMF. For our needs we need to divide this by 1.5 to approximately correct it to a Kroupa IMF. You need to write an SQL query that find the Hα luminosity for a source (rest-frame wavelength of Hα is 6565 Å). Note that the fCosmoDl function returns the luminosity distance and that the luminosity of a source at redshift z is given as flux $4\pi d_L^2$. Note also that the SDSS quotes fluxes in units of 10^{-17} erg/s/cm^2. You will need to join the SpecPhoto and SpecLine tables to get the line fluxes:

```
SELECT s.SpecObjID, s.z –we want the photometric objID
FROM SpecObj as s
JOIN SpecLine as L ON s.SpecObjID = L.specObjID –line L is detected in spectrum
WHERE –you could add a constraint that the spectral type is a galaxy
L.LineId =4863 –and the line L is the Hβ line
```

But be aware that in the SpecLine table they give the parameters of the Gaussian:

$$f(\lambda) = Ae^{-\frac{1}{2}\frac{(\lambda-\lambda_0)^2}{\sigma^2}} \tag{2}$$

where σ is the width of the line and A is the height of the line, while you need the flux (i.e. the integral of the line). Typical star formation rates vary from 0.01 to 100 solar masses per year, while the absolute magnitudes should lie in the range -15 to -25 or so.

When this is done you should obtain the theoretical predictions and compare your absolute magnitude in r-band versus $\log(SFR)$ against the theoretical predictions. How do they compare? Are they the same? Different? Do they agree with each other? Why, or why not?

3 Suggestions

1. Give a reasonable name to your tables, so you can easily recognize every table. You can do it directly in the query using the command INTO.
2. In order to recognize the components of a table from the command, write the SQL keywords in uppercase (common practice in SQL).
3. Before submitting a query, check your query syntax using the button SYNTAX.
4. Be sure that in your table there are some elements, otherwise something went wrong. You can download your table from MyDB page (Fig.3).

```
SELECT TOP 200 s.specObjID, s.z, L.ew
INTO mydb.Exercise1 FROM SpecPhoto AS s
JOIN SpecLine AS L ON s.specObjID = L.specObjID
WHERE L.LineID = 0000
AND L.lineID = 0000
AND L.ew > 0000
```

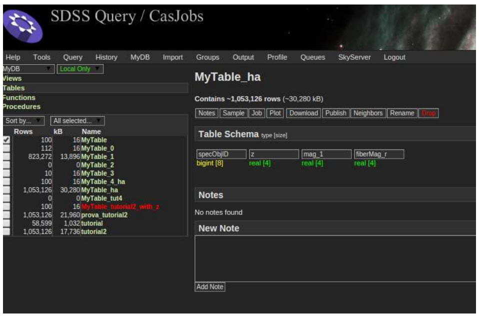


Figure 3: MyDB page: there are tables (left) without elements because of syntax errors.

4 The report

Deadline: 13-10-2023 17:00h
Upload your report in Assignments on Brightspace

Your report will be on the mass-metallicity relation (if you chose Project A) or SFR-luminosity (if you chose Project B). The report should be divided in the following sections:

- *Introduction*: here you describe the astrophysical background and the databases, with particular regard of the meaningful aspects for this project (max. 1 page);
- *Results*: here you discuss how you selected the data from SDSS and Millennium, how you used these to study the relations (plots!);
- *Discussion*: in this section there will be your discussion of the results. For example, is there a correlation between luminosity and star formation rate in the subsample you have chosen? Does it have the same slope of the theoretical one? Is there any change if you select different type of galaxies (spirals, ellipticals etc..)? Is the scatter of the observed relation the same of the theoretical relation?

Please make your report 3-4 pages, including plots. The minimum fontsize should be 11 pt (smaller fonts are allowed for figure captions). You can add tables, figures and whatever you need to make your report complete, but do not exceed the page limit. The emphasis should be on how the results and how you interpret them. Do not forget to write your name and a title. Lastly, please include with your report the data you downloaded and a working example of the code you used for the assignment (e.g. a python script, jupyter notebook).