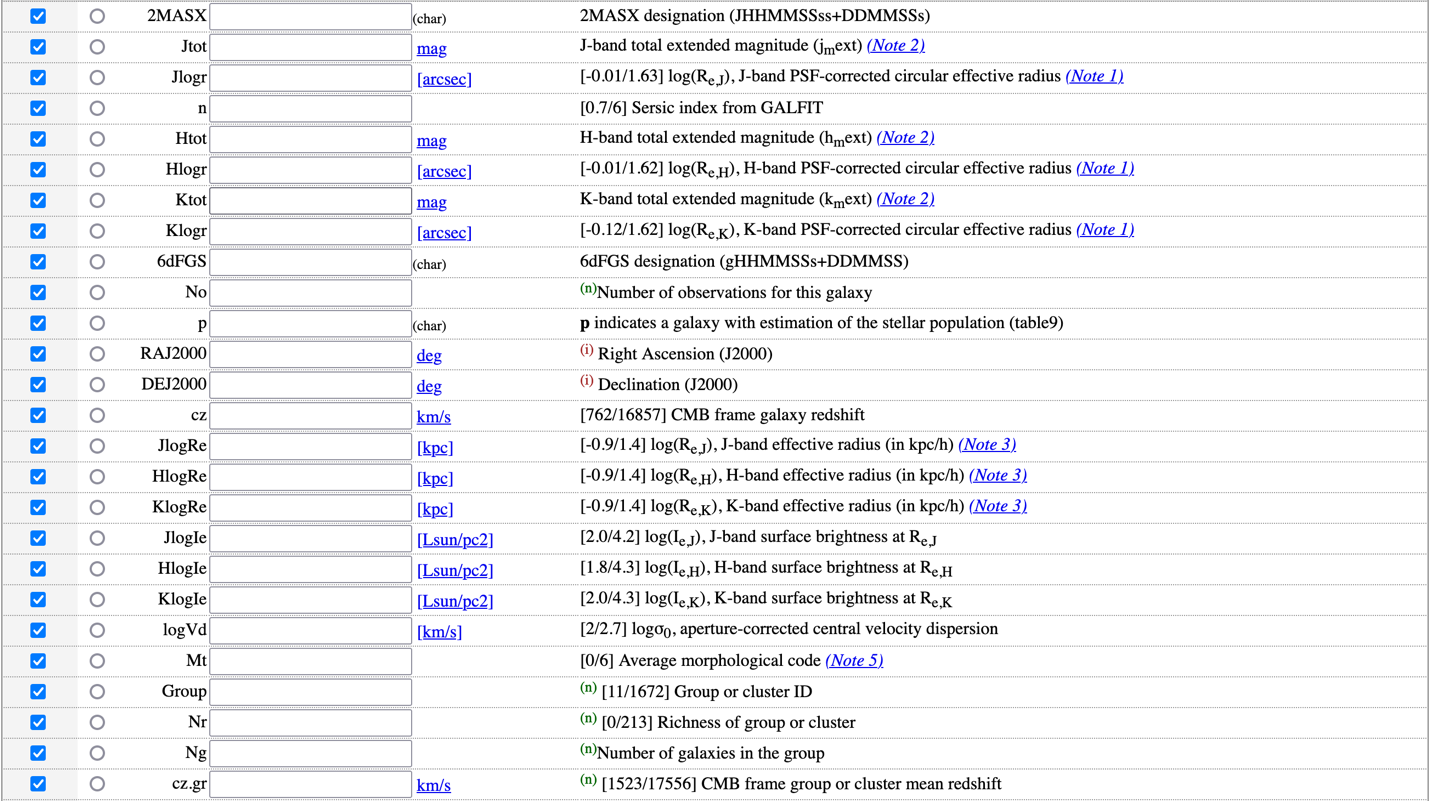
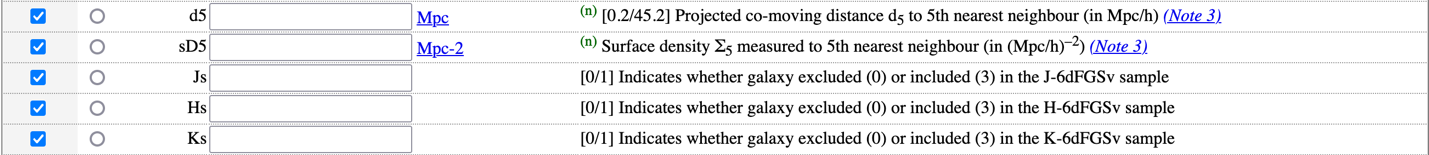
**WHAT I HAVE BEEN DOING**

1. **Obtaining the raw spectroscopy/velocity dispersions data**
2. 6dFGS (there are two data sources, I have not inspected whether they are different):
   * + 1. First source is what I actually used throughout my thesis, Campbell’s table on Vizier obtained [here](https://vizier.cds.unistra.fr/viz-bin/VizieR-3?-source=+J%2FMNRAS%2F443%2F1231%2Ftable2&-from=nav&-nav=cat%3AJ%2FMNRAS%2F443%2F1231%26tab%3A%7BJ%2FMNRAS%2F443%2F1231%2FFPsample%7D%26key%3Asource%3DJ%2FMNRAS%2F443%2F1231%2Ftable2%26HTTPPRM%3A%26) . I use two tables, first is FPsample table (FP quantities) with the following schema:





second table is table2 (velocity dispersion) with the following schema:

A screenshot of a computer

Description automatically generated

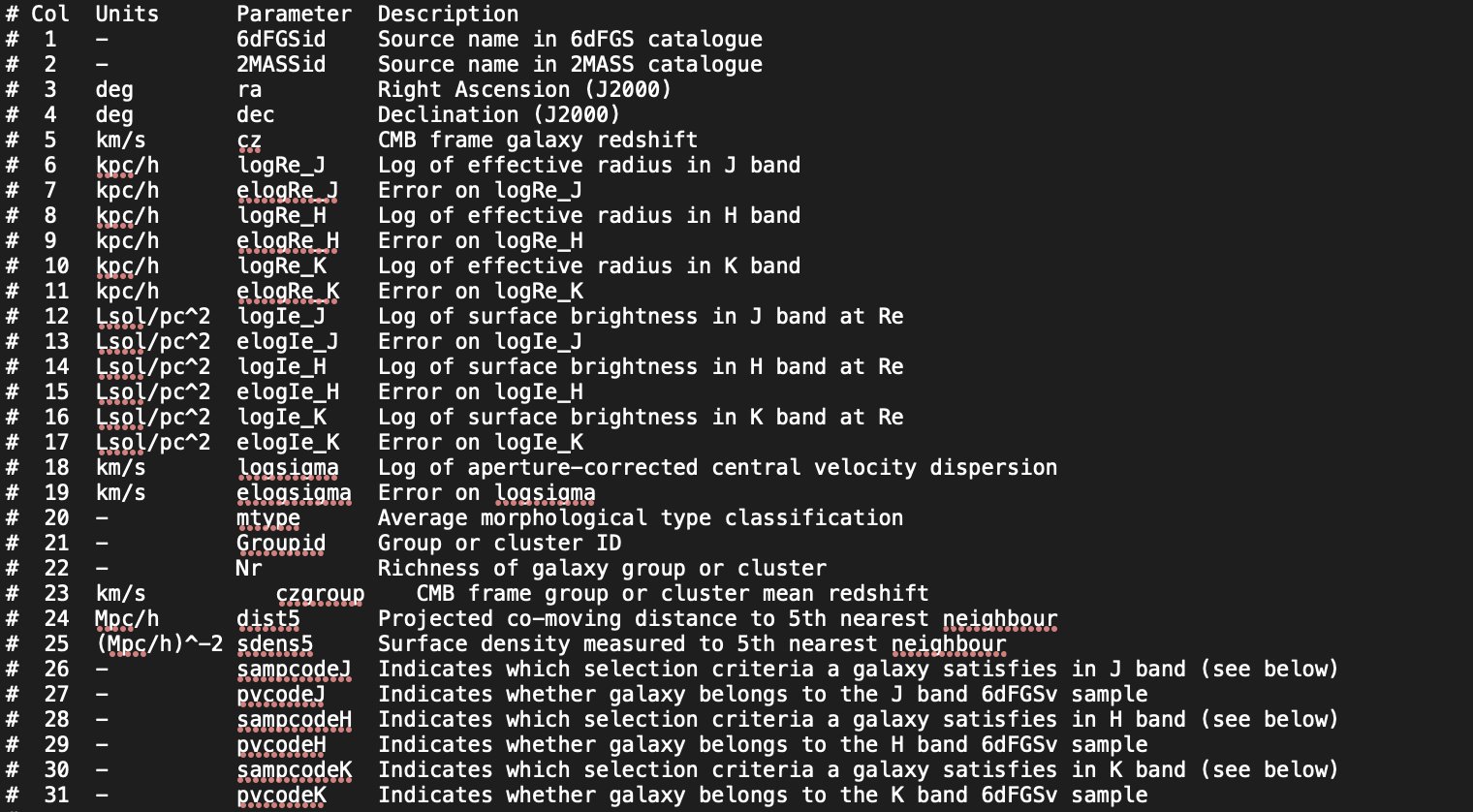
* + - 1. Second source is from Campbell et al. (2014) on MNRAS obtained [here](https://academic.oup.com/mnras/article/443/2/1231/1060810) (supplementary data section). The files are campbell\_table2.ascii (velocity dispersion data), campbell\_table4.ascii (NIR photometry) and campbell\_table8.ascii (derived FP data).
         1. table2 schema (unsure if the descriptions are true):

|  |  |  |  |
| --- | --- | --- | --- |
| column name | type | Unit | description |
| sdf | string |  | 6dFGS id |
| tmass | string |  | 2MASS id |
| mjd | float |  | Modified julian date of the observation |
| Z\_helio | float | Redshift | Heliocentric redshift |
| s2n | float |  | Signal to noise ratio of velocity dispersion measurement |
| Cor\_sigma | float |  | Cross-correlation R parameter? |
| Dex\_error | float | [km/s] | Error of log\_Vd? |

* + - * 1. table4 schema

|  |  |  |  |
| --- | --- | --- | --- |
| column name | type | Unit | description |
| MASS\_name | string |  | 2MASS id |
| J\_m\_ext | string | Mag | Extrapolated J magnitude |
| Fwhm\_j | float | Arcsecond? | Modified julian date of the observation |
| Delta\_r\_j | float | Arcsecond? | PSF correction? |
| Log\_re\_j | float | [arcsecond]? | Log PSF-corrected effective radius |
| nser | float |  | Sersic index |
| H\_m\_ext | float | Mag | Extrapolated H magnitude |
| Fwhm\_h | float | Arcsecond? |  |
| Delta\_r\_h | float | Arcsecond? |  |
| Log\_re\_h | float | [arcsecond]? |  |
| K\_m\_ext | float | Mag | Extrapolated K magnitude |
| Fwhm\_k | float | Arcsecond? |  |
| Delta\_r\_k | float | Arcsecond? |  |
| Log\_re\_k | float | [arcsecond]? |  |

* + - * 1. table8 schema



1. SDSS: obtained from SDSS DR14 via the following CasJobs query:

SELECT s.specobjid, p.objID, s.ra, s.dec, s.plate, s.instrument, s.mjd, s.fiberid, p.devMag\_u, p.devRad\_u, devAB\_u, p.devMag\_r, p.devRad\_r, devAB\_r, s.z,s.zErr,s.veldisp,s.veldispErr, em.sigmaStars, em.sigmaStarsErr

into mydb.SDSS\_spectro

From SpecObjAll as s

JOIN emissionLinesPort em ON (em.specObjID = s.specobjid)

JOIN PhotoObjAll p ON (p.specObjID = s.SpecObjID)

WHERE

(s.sdssPrimary = 1)

AND (s.z <= 0.1)

1. LAMOST: the file lamost\_DR7\_VDcat\_20200825.fits obtained from Khaled. Schema:

A screenshot of a computer error

Description automatically generated

1. **Obtaining supplementary data**

* Two supplementary data required: John’s radii measurements and Tempel et al. SDSS DR8 groups and clusters data.
* John provided colours table (data/raw/r\_e\_jrl/colours.ascii), radii measurements (data/raw/r\_e\_jrl/jhk\_r\_e.csv), and LAMOST ETG list (data/raw/r\_e\_jrl/lamost\_good\_pv\_list.csv)
  + Colours table (I didn’t use these, so I don’t know what most of the columns are):

|  |  |  |  |
| --- | --- | --- | --- |
| column name | type | Unit | description |
| tmass | string |  | 2MASS ID |
| Ra | float | deg |  |
| Dec | float | deg |  |
| J\_ext | float |  |  |
| J\_ext\_error | float |  |  |
| G\_r\_ext | float |  |  |
| GAIA3\_B\_R | float |  |  |
| GAIA3\_B\_R\_error | float |  |  |
| PS1\_g\_r\_5 | float |  |  |
| PS1\_g\_r\_5\_error | float |  |  |
| SM3\_g\_r\_5 | float |  |  |
| SM3\_g\_r\_5\_error | float |  |  |
| J\_K\_5 | float |  |  |
| J\_K\_5\_err | float |  |  |
| W2\_W3 | float |  |  |
| W2\_W3\_err | float |  |  |

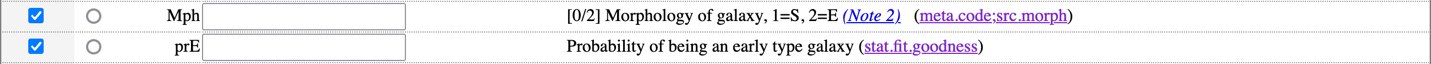
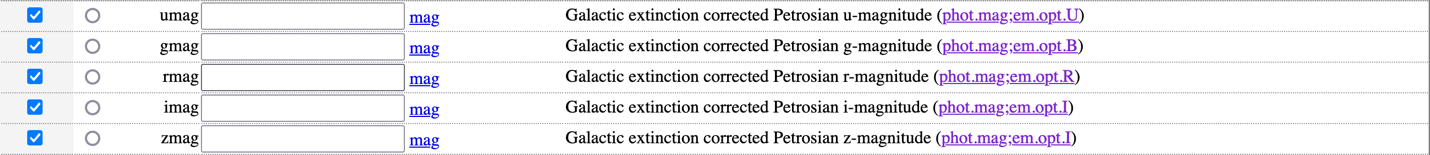
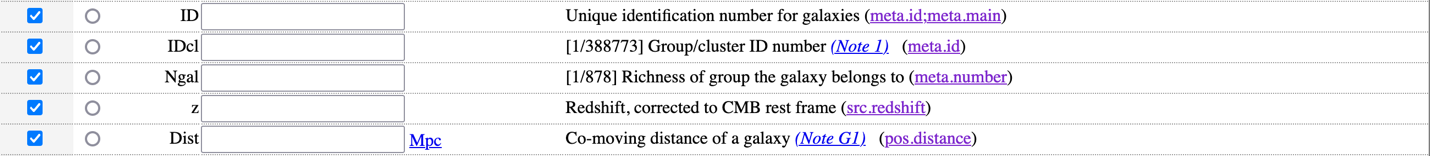
* + Radii measurements

|  |  |  |  |
| --- | --- | --- | --- |
| column name | type | Unit | description |
| tmass | string |  | 2MASS ID |
| Log\_r\_h\_app\_j |  |  |  |
| Log\_r\_h\_smodel\_j |  |  |  |
| Log\_r\_h\_model\_j |  |  |  |
| Red\_chi\_j |  |  |  |
| Galfit\_ser\_j |  |  |  |
| Fwhm\_j |  |  |  |
| Fit\_ok\_j |  |  |  |
| Log\_r\_h\_app\_h |  |  |  |
| Log\_r\_h\_smodel\_h |  |  |  |
| Log\_r\_h\_model\_h |  |  |  |
| Red\_chi\_h |  |  |  |
| Galfit\_ser\_h |  |  |  |
| Fwhm\_h |  |  |  |
| Fit\_ok\_h |  |  |  |
| Log\_r\_h\_app\_k |  |  |  |
| Log\_r\_h\_smodel\_k |  |  |  |
| Log\_r\_h\_model\_k |  |  |  |
| Red\_chi\_k |  |  |  |
| Galfit\_ser\_k |  |  |  |
| Fwhm\_k |  |  |  |
| Fit\_ok\_k |  |  |  |

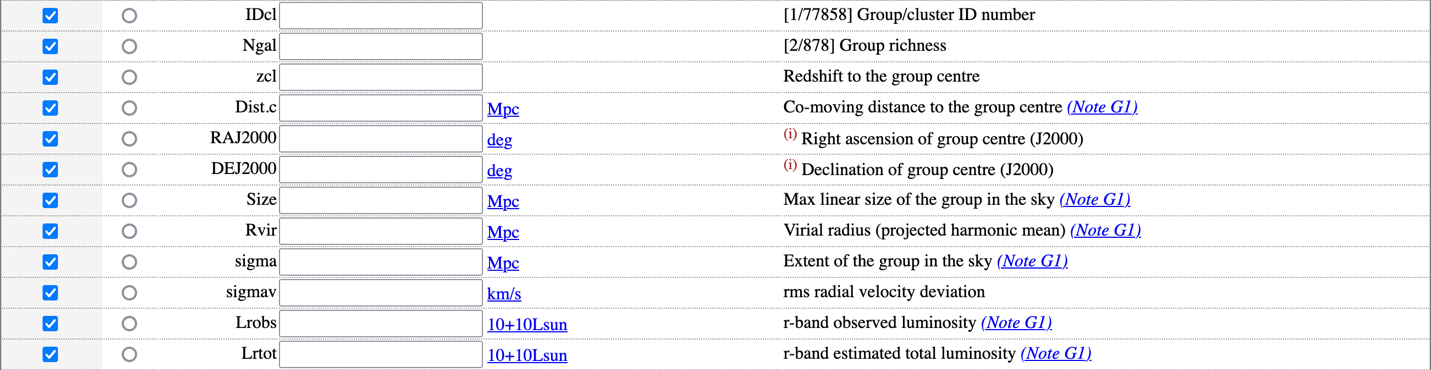
* + lamost\_good\_pv\_list.csv

|  |  |  |  |
| --- | --- | --- | --- |
| column name | type | Unit | description |
| Col1 | string |  | 2MASS ID |
| 2MASS | string |  | 2MASS ID |

* I downloaded Tempel et al. SDSS DR8 groups and clusters from vizier ([here](https://vizier.cds.unistra.fr/viz-bin/VizieR-3?-source=J/A%2bA/540/A106/dr8gal)). It contains two tables:
  + dr8gal 🡪 individual galaxy data. I filtered Ngal > 1 to exclude groups with only 1 member

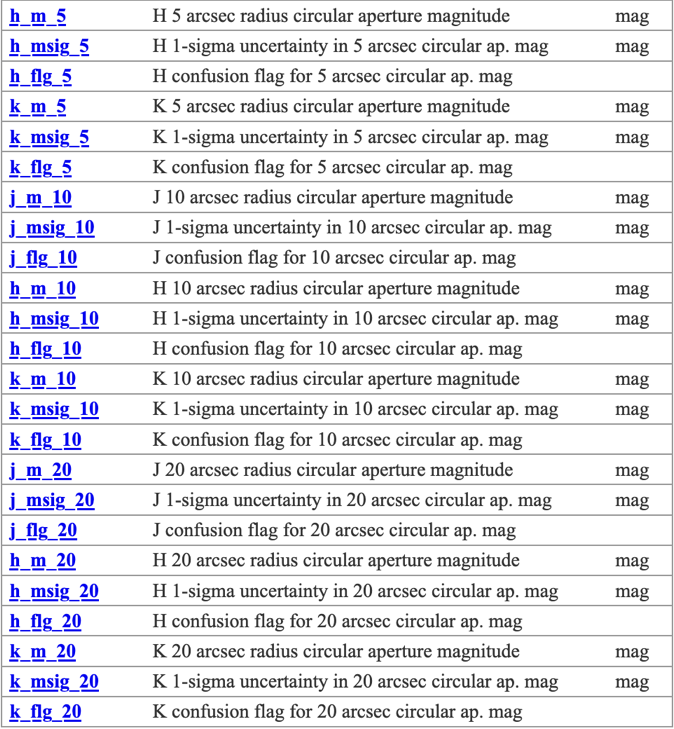
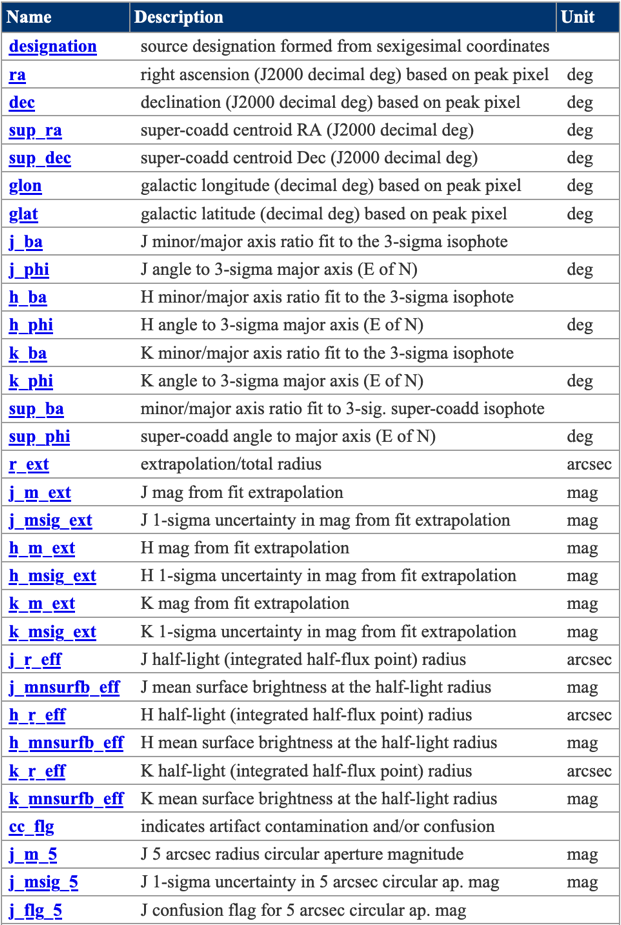
****

* + dr8gr 🡪 groups and clusters data

****

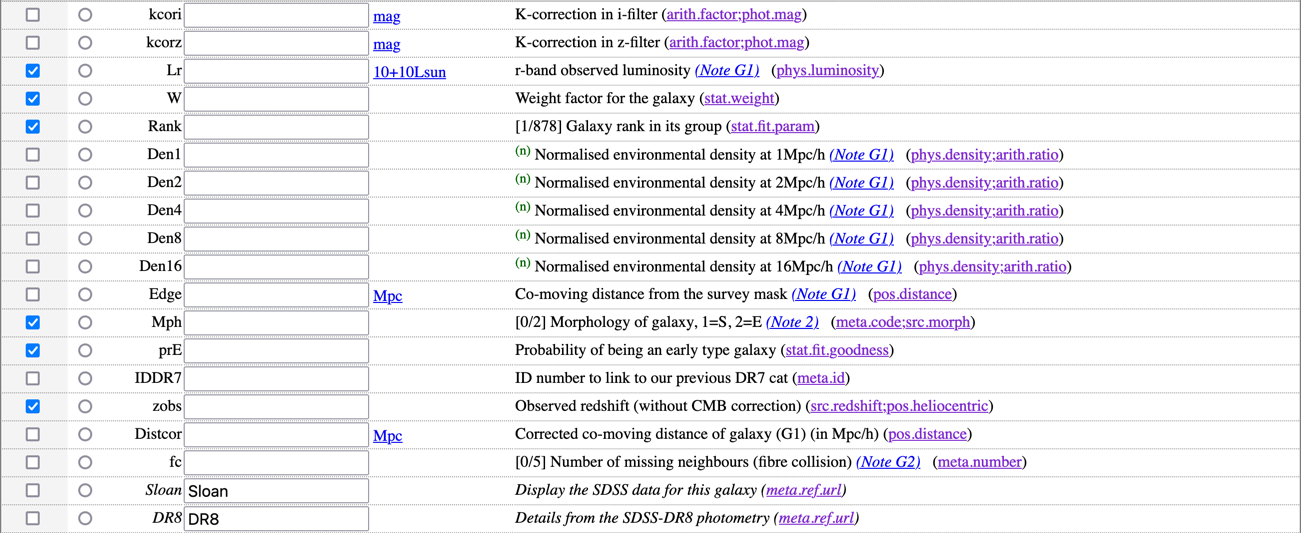
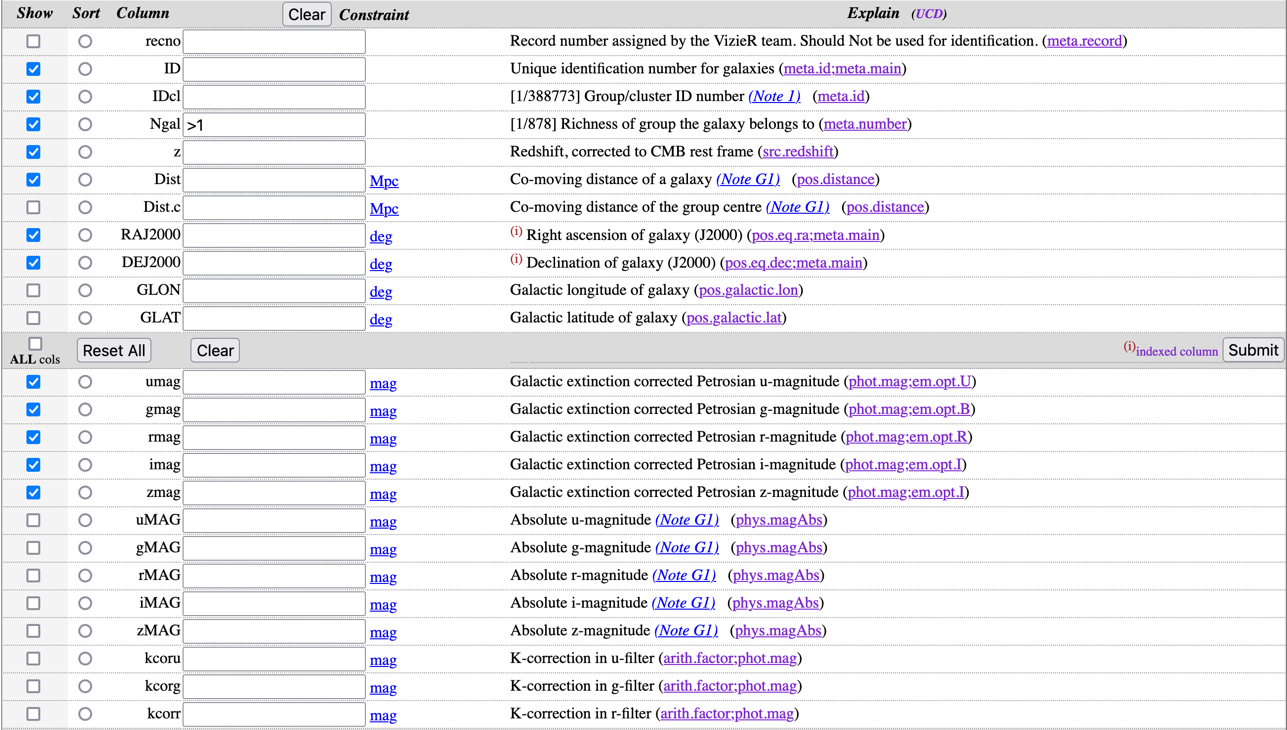
1. **Obtaining the sky coordinates from all the raw galaxies**

* I use the get\_coordinates.py script to fetch the (ra, dec) and store them to data/preprocessed/sky\_coord/<survey>.ascii file (IPAC format).
* I had to multiply 6dFGS ra by 15 to get the ra in degrees (it was given in hour), while for SDSS and LAMOST I fetch the ra and dec directly.
* To set the cone search radius, I first queried 6dFGS galaxies to 2MASS XSC. I compared each input galaxy’s ra dec to its 2MASS counterpart’s ra dec. I found that the maximum separation was 2.22”.
* I set the cone search radius to be 2.5 arcsecond (based on the argument above but more relaxed) and checked the ‘One to One Match’ box (basically doing a left join) so that I can simply merge the resulting dataframes later. These are the fields that I queried:



I saved the results at data/raw/2mass/<survey>\_tmass.csv. Also backed up the files in Google Drive (cannot push them to GitHub as they are too large).

1. Merge the original spectroscopy data with all others in the following order:
   1. 2MASS (from the response above): this file has the same number of rows as the original spectroscopy data, and I checked that their orders are the same, so I merged them simply by using their indices. Then I dropped all rows where ‘designation’ is null (no 2MASS counterpart).
   2. John’s radii measurements: merge them by using the 2MASS id as primary key. Actually the number of galaxies are different (this time SDSS sample has much more), and not sure why.
   3. For SDSS and LAMOST: cluster measurements from Tempel et al. (SDSS DR8 groups and clusters) done in two steps:
      1. First, I performed sky crossmatch for the individual galaxies
      2. Second, I join the individual galaxy with the group and cluster table to obtain the cluster redshift, member, and richness
      3. Galaxy table:



* + 1. Group table (Groups with IDcl>77858 have only 1 galaxy.):

