Frequently Asked Questions

More Technical Questions

Answers

1. What is the Sloan Digital Sky Survey? Top

The Sloan Digital Sky Survey (SDSS) was one of the most ambitious scientific projects of all time. Its primary science goal was to make a high-quality three-dimensional map of the universe. The survey used a specially-built 2.5-meter telescope in New Mexico and a CCD camera to take images of about one-quarter of the night sky. It entire dataset is now available online to astronomers and the general public.

The SDSS has taken images of more than 300 million objects. The survey has already completed a preliminary map of the universe, and has made many scientific discoveries from our Solar System to the edge of the Universe. The SDSS's discoveries are detailed in the Science section of the sdss.org website (link opens in a new window).

2. How can I get SDSS data? Top

SDSS data is available from this Data Release 7 website.

3. What is the Catalog Archive Server (CAS)? What is the Data Archive Server (DAS)? Top

SDSS data can be accessed in two ways. It can be viewed as catalog data (measured parameters for sky objects) or as FITS images and data files. Catalog data, along with preview images and spectra, are available from the Catalog Archive Server. FITS images and data are available from the Data Archive Server.

4. What is SkyServer? Top

SkyServer is a web site where students and members of the general public can get data from the Sloan Digital Sky Survey (SDSS). SkyServer shares many data access tools with the Catalog Archive Server, but also includes some tools designed specifically for the public. SkyServer's Projects are excellent resources for astronomy educators who want to include real data in their courses.

5. Where in the sky do the data come from? Top

The SDSS takes data in long, narrow "stripes." See the Sky Coverage page (link opens in a new window) for maps and tables that show where in they sky the current data come from. The CAS Navigate tool (new window) also has an interactive sky globe that shows where the SDSS has data. To see if a specific area has been seen by the SDSS, enter its coordinates into the Finding Chart (new window).

6. How can I search for data? Top

When you search for data in the SDSS, you are going through the SDSS database and looking for objects that match criteria you choose. For simple searches of photometric data, use the Imaging Query Form (new window). For simple searches of spectroscopic data, use the Spectroscopic Query Form (new window). For more complex searches, use Structured Query Language (SQL). See the CAS's guide on Searching for Datato learn more about SQL. To see thumbnail results of objects that meet your criteria, use the Image List tool (new window).

For large, complex queries that will take a long time to run, use CasJobs (new window), the SDSS's batch query interface.

The CAS has many other tools too. See Getting Started (new window) for more information on all the tools.

7. What help is available? Top

This site has an extensive Help section, including a Glossary of SDSS terms and a set of Tutorials, which are guides for doing common research tasks with SDSS data. Each tool also has its own Help section.

8. I am a teacher. How can I use the data in my classes? Top

SkyServer's Projects use SDSS data to teach topics in astronomy and other sciences, using guided and open inquiry. With our Projects, you and your students can learn about spectra and colors of stars, galaxy types, the history of the universe, and much more.

You are welcome to use and adapt any of our projects in your classes, free of charge. For more information on what you can do with SkyServer in the classroom, see our Teacher FAQ.

9. How can I see if the SDSS has an image of my favorite object? Top

Find the coordinates of the object using a name resolver like SIMBAD (link opens in a new window) or NED(new window).

Then, go to the CAS's Navigate tool (new window) and enter the object's coordinates. You can enter the coordinates as decimal degrees or as sexagesimal in the format HH:MM:SS and (+/-)DD:MM:SS. Click "Get Image" to see the object, and click on the object for its SDSS data. See the Help link in the Navigate tool for more information.

10. How can I match a list of objects to see what the SDSS knows about them? Top

Go to the CAS's Imaging Upload tool (opens in new window). Paste your list of objects, or upload a file containing data with the last two columns as (ra, dec) in decimal degrees. Click Submit. The next page will show only those objects that appear in the SDSS, with SDSS Object IDs that link to the Explore tool.

To see a thumbnail SDSS image of each matching object, use the Image List tool. Enter your list in the textbox on the upper left and click "Get Image". Click on one of the thumbnails to go to that position in the Navigatetool, or on one of the object names to go to that object's Explore tool entry.

11. Why doesn't the SDSS have data for well-known visible stars (Sirius, Vega, etc.)? Top

The SDSS has a very sensitive camera. Stars that you can see with your unaided eyes are a little too bright for the SDSS's camera, so they show up as washed out. The SDSS still gets an image of those stars (for example, here is an image of Pollux from SkyServer - link opens in a new window), but their images are unreliable, and the SDSS gets no catalog data.

12. Why are some bright stars classified as galaxies? Top

The SDSS distinguishes between stars and galaxies based on their shapes: single points of light are stars, and fuzzy patches of light are galaxies. Some stars are bright enough that their light washes out the camera, so to the SDSS's camera, they look like fuzzy disks instead of single points of light. Their appearance fools the SDSS's software into classifying them as galaxies.

13. What does the long SDSS ID number mean, and how do I work with it? Top

The SDSS needs a way to uniquely identify every object in the database, so it generates ID numbers. The ID numbers are code numbers that include information about how the object was observed.

One very important point when working with SkyServer is that the object IDs are so long that they get cut off in Excel, and show up with 000 as the last three digits. This means you won't be able to find your objects anymore! To get around this problem, see this workaround.

More Technical Answers

1. What interfaces are available to SDSS catalog data? Top

The Catalog Archive Server (CAS) is the database that contains the SDSS's catalog data. There are multiple interfaces to the data, including (all links open in a new window):

1. Catalog Archive Server: a web-based, synchronous interface for browsing and searching
2. SkyServer: a special interface to the CAS designed for the general public and students
3. CasJobs: a batch (asynchronous) system for querying the database and storing results
4. sqlcl: a command-line interface
5. Emacs buffer: an Emacs interface for running queries

Alternatively, you can retrieve SDSS FITS images with the Data Archive Server.

2. How do I convert from the SDSS's ugriz magnitudes to UBVRI magnitudes? Top

The SDSS measures magnitudes through *ugriz* filters, which give *ugriz* magnitudes. These magnitudes can be converted into UBVRI magnitudes using a set of transformations described on the Algorithms page of this site.

3. What is the difference between Target and Best? Top

Because the survey imaging and processing was improved over the course of the survey, the SDSS often obtained improved photometric measurements of objects AFTER they have been chosen ("targeted") for spectroscopy. However, it is important to keep a record of the photometric measurements at the time objects were targeted. We therefore maintain two versions of the photometric catalog:

1. **Target:**

   Known as TARGDR7 in the DR7 CAS, this database contains the photometric catalogs AS THEY WERE WHEN OBJECTS WERE CHOSEN FOR SPECTROSCOPY. This database contains the union of all target chunks. It may cover a slightly different area than Best, blended objects may be deblendeddifferently, image quality may be worse, and photometric calibration may be less accurate. However, if you want to see what the SDSS thought the magnitudes and other properties of an object were when it was chosen for spectroscopy, this is the place to look. Note that this database DOES NOT contain links from the photometric objects to the spectroscopy (you can always get the Target data for spectroscopic objects using the TargObjID field in the `SpecObjAll` table), nor does it contain the tiling information. This is because the Target database is intended to be a snapshot of the survey before any spectroscopy is done.

2. **Best:**

   Known as BESTDR7 in the DR7 CAS, this database contains the latest, best versions of the imaging data, processed with the latest version of the photometric processing software, and with the most recent understanding of the photometric calibration applied. For any science based on object photometry, you will want to use the Best data. In addition, only the Best database contains all of the spectroscopy and tiling information.

4. What is the difference between *SpecObj* and *SpecObjAll*? What does*sciencePrimary* mean? Top

The `SpecObjAll` table contains ALL spectroscopic objects, regardless of their status in the survey. Queries on this table can produce unusual or undesired objects. Thus, we have created the `SpecObj` view, which contains data for ONLY those fibers defined as SciencePrimary. To be defined as SciencePrimary, an object *must*meet all of the following criteria:

1. It was targeted in the target skyVersion
2. The plate on which the spectrum was taken is the primary observation of that tile
3. The plate was a main survey plate (not part of the Southern survey or a special project)
4. The objType is *not* QA, SKY, or SPECTROPHOTO_STD (these object types are repeatedly observed)
5. The fiber was mapped correctly (this is the `zWarning` check)

As a result, some plates may have many (or even all) of their fibers excluded from SpecObj. Some instances where this occurs are:

- Fibers and plates that were targeted but are outside the official survey boundaries. These objects cannot be matched to a *PhotoPrimary* in the target) version of the sky so they do not make it into`SpecObj`. Plates 344-346, 348, and 364 (tileRun 6) have 0 SpecObjs. Other plates from tile runs 4 and 6 have reduced numbers of objects, especially 315 and 342. This could affect plates 266-315, 363, 361 from tileRun 4 and plates 342-348, 364 from tileRun 6.
- Fibers where objType = 'SKY' do not show up in the `SpecObj` view. Some plates (for undetermined reasons) have large numbers of sky fibers: Plate 417 has 214 sky fibers, and they all seem to be on one half of the plate; Plate 595 has 91 SKYs and Plate 359 has 84 SKYs.
- Fibers where objType = 'QA' also do not show up in the `SpecObj` view. Below is a table of the plates with more than 100 QA fibers and the number of such fibers:

Plate	# of QA fibers
483	174
471	136
500	125
470	123
418	120
550	108

- **Note:** The definition of **SciencePrimary** relies purely on spectroscopic and geometric considerations. There are objects in `SpecObj` which *do not* have a corresponding **Best** photometric object.

5. What are the differences between *PhotoObj*, *PhotoTag*, and *PhotoObjAll*? Top

PhotoObjAll is a table in the Best and Target databases which contains **all** of the measured photometric quantities for **all** of the imaging objects. Because we measure hundreds of parameters for each of 340 million objects, this is a very large table, and queries can take very long to run.

In an effort to speed up queries, we have created a table with only a subset of the parameters that are requested most often (a "thin table"). This table is called *PhotoTag*. If you have a query that uses and returns only values stored in *PhotoTag*, it will execute much faster than if you used *PhotoObjAll*.

In addition, we have created a view of *PhotoObjAll* that contains only those objects that are Primary orSecondary. This view is called *PhotoObj*. Because this view effectively contains fewer objects than PhotoObjAll (but all the measured quantities for these objects), queries will execute faster.

Given the above, you should consider:

0. Querying from *PhotoTag* if it contains everything you are looking for
1. Querying from *PhotoObj* otherwise, UNLESS you are interested in data for objects which are *neither PRIMARY nor SECONDARY*. In that case, you will need to use *PhotoObjAll*.
2. Importantly, the "shorthand" quantities *u,g,r,i,z* do not exist in the *PhotoTag* table (because we want to keep it as thin as possible). Instead, you must use ModelMag_[ugriz], which is indexed to make queries faster. **HOWEVER**, in *PhotoObjAll* and its views, only the *u,g,r,i,z* are indexed, and *not* the ModelMags!

Because PhotoTag has many fewer parameters, larger portions of it can be cached, improving performance. We have found that for almost all queries which contain parameters fully in PhotoTag, it is faster. If you were looking for objects that had been detected multiple times, the fastest approach would be to perform a join on*PhotoTag* with itself, requiring that one object be Primary and the other Secondary.

6. How do I get photometry for spectroscopic objects? What is the *SpecPhotoAll*table? Top

The *SpecPhotoAll* table is a precomputed join between the Best *PhotoObjAll* and *SpecObjAll* tables. It includes the most requested parameters from these two tables, as well as a few pieces of information abouttiling. It also includes the TargetObjID, which allows user to retrieve the Target version of the photometry.

Note that all spectro-photo matches are not included in SpecPhotoAll, since there are additional JOIN conditions imposed on tiling and targeting information. The actual SQL fragment from the SpecPhotoAll construction is shown below to indicate all the JOINs involved:

```
FROM SpecObjAll s
    JOIN TileAll w ON s.tile=w.tile
    LEFT OUTER JOIN TilingInfo t ON s.targetid=t.targetid and
```

```
w.tileRun=t.tileRun
JOIN TargetInfo i on s.targetObjid=i.targetObjid
LEFT OUTER JOIN PhotoObjAll p WITH (nolock) ON
s.bestObjid=p.objid
LEFT OUTER JOIN PhotoTag q ON p.objid=q.objid
```

SpecPhotoAll is very useful for viewing and comparing objects' photometric and spectroscopic properties.

7. What is the difference between *SpecPhoto* and *SpecPhotoAll*? Top

As described above, *SpecPhotoAll* is a precomputed join between the Best *PhotoObjAll* and *SpecObjAll*tables. This includes non-science objects, and a variety of objects many users will not be interested in. The*SpecPhoto* view includes only those pairs where the SpecObj is a sciencePrimary (see the definition above), and the BEST PhotoObj is a PRIMARY object.

8. Why do z and zErr (in SpecObj) have different numerical precisions? Top

Internally, these numbers are stored to their full precision as they come out of the spectroscopic pipeline. When you perform a query, they have some default string format applied that cuts them to what you see. But you can use SQL's str() function to change the string format to whatever you like.

To get z to 6 decimals, for example, change your query to 'select str(z,8,6) as z' instead of just z, and analogously for zErr. This applies the function str() to the values in column z and returns the result with column label z (without the "as", the result of a function has no column label). The str(col,length,dec) function takes the numerical value in 'col' and formats it as a string of length 'length' and with 'dec' significant digits. In other words, str(z,8,6) is the SQL equivalent to the C function printf("%8.6f",z). str() rounds the result to the number of decimals you request.

9. How do I change the default precision of values in the output of my query? Top

Use the **str(*column*,*n*,*d*)** SQL construct (where *n* is the total number of digits and *d* is the number of decimal places) to set the precision of the column that your query requests. SkyServer returns values with a default precision that is set for each data type, and this may not be enough precision for some types of science. See the Selected neighbors in run or the Uniform Quasar Sample sample queries (both open in new windows) for examples of how to use STR.

10. What is the difference between *specClass* and *objType* for spectroscopic objects, and which one should I use? Top

The *objType* parameter in *SpecObj* and other tables is set when the objects are targeted for spectroscopy, when the spectroscopic plates are prepared. The *specClass* parameter is set by the spectroscopic pipeline after the spectrum is observed. For science, you should use the specClass attribute. The objType field is included for studies of the targeting algorithm.

11. Why does SDSS use the long (64-bit) objID fields, and what is the composition of the PhotoObj objID and SpecObj specObjID fields? Top

The 64-bit ID fields are required as primary keys (unique identifiers) in the SDSS database tables. They are used to uniquely identify each record in the database indices (link opens in a new window) for enhanced performance. Each of them are bit-encoded with information about the observational origin, i.e., the run,rerun, camera column, etc. for photometric data, and the plate, MJD, fiberID etc. for spectroscopic objects. Please see the entry for SDSS ObjID Encoding (new window) in the Algorithms page.

12. I want to mirror the SDSS archive - how can I get a copy of *all* the data? Top

A copy of the current publicly-available SDSS data release is available from UIC (University of Illinois at Chicago) for worldwide distribution over fast links. Please see the SkyServer support site at skyserver.org(new window) for further details on how to host a mirror site and where to get the data. Click on the **SDSS Mirrors** link on that site.

13. Where can I get a copy of the HTM (Hierarchical Triangular Mesh) spatial index library? Top

Freely downloadable copies of the HTM library (in C++, Java and now C#) are available at the SkyServer support site at skyserver.org (new window). Click on the **HTM** link on that site.