**Werkcollege from 23 September**
**Astronomial databases - SDSS CasJobs**
**Solutions**

*Originally created by Jarle Brinchmann, Leiden Observatory for IAC 2010 Modern Data Mining and Virtual Observatory*

# 1 Astronomical Databases and Catalogs

## 1.1 a) Bright Star Catalogue Supplemen:

- Selection criteria: V-band between 6.5 and 7.1, Complete all-sky
- Spectral coverage: 3500-9000 A (U,B,V,R,I, but not complete in U,B,R,I)
- Spectral resolution: about 4 (imaging)
- Spatial coverage: all sky
- Spatial resolution: irrelevant (point sources)

## 1.2 b) Pickles 1985:

- Selection criteria: Some bright stars, not complete
- Spectral coverage: 3600-10000A, spectral resolution: 10-17A
- Spatial coverage: all sky
- Spatial resolution: irrelevant

## 1.3 c) James et al. 2004:

- Selection criteria: Galaxies in the UGC (B-band < 14.5, Northern Hemisphere), with v<3000 km/s, spirals or irregulars, Diameter between 1.7ánd 6́(see paper). Complete.
- Spectral Coverage: H$\alpha$ (6561A).
- Spectral resolution: R=6561/50=130
- Spatial resolution: $\approx$ 1-2;̈ Field of View 10'x10'

## 1.4 d) Ebeling et al. 1998:

- Selection criteria: X-ray selected (band from 0.1-2.4 keV), Redshift < 0.3, flux > $4.4\,10^{-12} erg\,cm^{-2}\,s^{-1}$, complete
- Spectral coverage: $\approx$ 1-2 keV, resolution $\approx$ 1
- Spatial coverage: —b—>20, Northern Hemisphere
- Spatial resolution: <5 arcsec

## 2 CasJobs

Create a CasJobs account - run one simple search and download the result. This is not a terribly challenging task but it is very useful to be familiar with CasJobs - and at first glance it can be a bit bewildering. Note that the schema for the SDSS database is available on `http://cas.sdss.org/dr7/en`

To set up an account or later log in: `casjobs.sdss.org/CasJobs`

a)

b)

c)

d) There is no TOP in MySQL syntaxis. Nevertheless this feature can be introduced by using LIMIT:

```
SELECT z, ra, dec
FROM SpecObj
WHERE abs(z-0.04) < 0.001 LIMIT 10
```

In general case - LIMIT a,b - to select records with numbers from a to b

## 3 Stellar abundance distribution

a)

b) 471

```
select COUNT(*)
FROM sppparams
Where
fehaerr > 0
AND
feha < -3
```

c) Looking at the schema for the sppparams table we find the l & b columns to correspond to Galactic longitude and latitude. We can then create an SQL query to find this, an example would be:

```
select floor(b/10 + 0.5)*10 as latitude, COUNT(*) as num
FROM sppparams
Where feha < -3
GROUP BY floor(b/10+0.5)*10
ORDER BY latitude
```

d) Find the fraction of metal poor stars with latitude. Here we do two searches and join them. It would also be possible to combine these by first outputting them to MyDB and then joining afterwards, but this sorts it out more quickly:

```
select a.latitude, a.num as NlowFe, b.numall as Nall, cast(a.num AS float)/cast(b.numall AS float
FROM (
select floor(b/10 + 0.5)*10 as latitude, COUNT(*) as num
from sppparams
Where
fehaerr > 0
AND
feha < -3
GROUP BY floor(b/10+0.5)*10
) as a
```

```
JOIN
(
select floor(b/10 + 0.5)*10 as latitude, COUNT(*) as numall
FROM sppparams
WHERE
fehaerr > 0
GROUP BY floor(b/10+0.5)*10
) as b ON
a.latitude=b.latitude
ORDER BY a.latitude
```

The data can be saved in MyDB and visualized with any available software (export as VOTable and use topcat, for example).

# 4 Bright high-redshift objects

a)
```
Select TOP 10 name = str(plate, 4, 0)+'-'+str(mjd,5,0)+'-'+str(fiberid,3,0), ra, dec
From SpecPhoto Where
petroMag_R < 14
AND
petroMagErr_R < 0.1
ORDER BY -z
```

You can copy the result of the query into SDSS ImageList application (cut-out service).


Figure 1: Bright high-redshift objects cut-outs

b) The challenge here is to figure out what parameters to use to be able to reject poor data. Looking through the Schema for the SpecPhoto table you will quickly see the zConf parameter [the logical start is to search for quality, confidence, uncertainty, error.

In addition it would also help to use image flags to select objects with good photometry but that slows down the search a lot. So let us use the zConf parameter & zErr as well.

If you use zConf ¿ 0.97 then you still get some poor matches, so a better choice is to require zConf ¿ 0.99 and zErr ¡ 500/3e9 (ie. error less than 500 km/s). So the final query (with the redshift added to the name of the object) is:

```
Select TOP 10 name = str(plate, 4, 0)+'-'+str(mjd,5,0)+'-'+ltrim(str(fiberid,3,0))+'-z='+str(z,5,
From SpecPhoto
```

```
Where
petroMag_R < 14
AND
petroMagErr_R < 0.1
AND
zconf > 0.99
AND
zerr < 500/3e5
ORDER BY -z
```

and the images of the objects are on Figure 2.

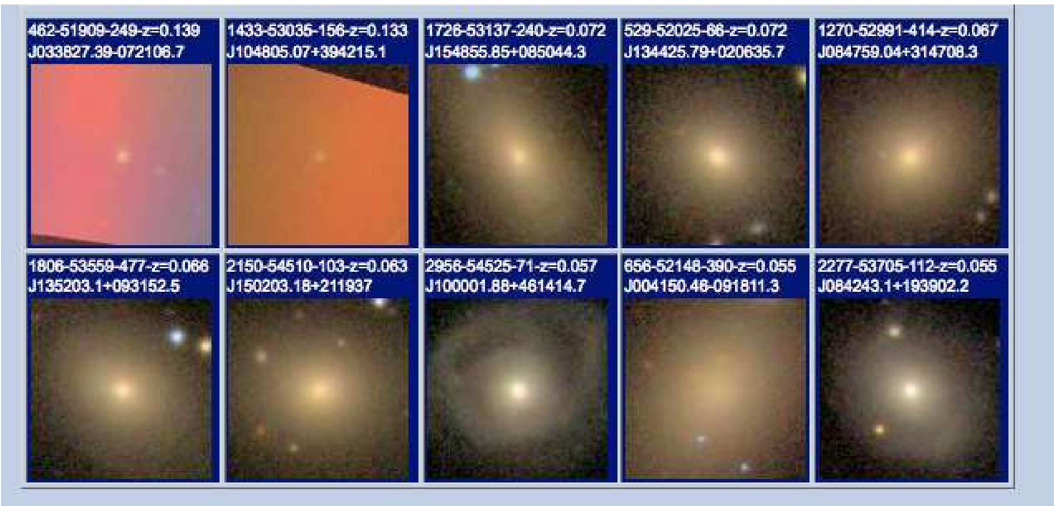


Figure 2: Bright high-redshift objects cut-outs with subselection

In general this makes sense - elliptical galaxies are the most massive objects in the Universe and that is what most of thee objects are. However the first two matches warrant a justification - they are genuine distant objects, but where the photometry is badly affected by scattered light.

Interestingly this is a problem that is inherently hard - if you repeat the same but with a magnitude cut of 15, you will find a large population of high redshift galaxies that peek through a foreground object and hence the photometry is for the foreground object but the redshift is for the background one.

c) dbo.fGetNearestObjIdEq(RA, DE,radus)

# 5 Planetary nebulae

a) Catalogue of Galactic Planetary Nebulae. Search for these on Vizier and you can here choose from several - either Catalogue of Galactic Planetary Nebulae (Kohoutek, 2001) or Strasbourg- ESO Catalogue of Galactic Planetary Nebulae (Acker+, 1992).

b) The creation of a new catalogue for upload would not strictly be necessary because you could just download the necessary columns from Vizier. However it is useful to have the extra information for later.

c) You match to the closest SDSS object by:

```
SELECT p._RAJ2000 as ra, p._DEJ2000 as dec,
dbo.fGetNearestObjIdEq(p._RAJ2000, p._DEJ2000,3.0) as objId
FROM
MyDB.PNe as p
WHERE
dbo.fGetNearestObjIdEq(p._RAJ2000, p._DEJ2000,3.0) IS NOT NULL
```

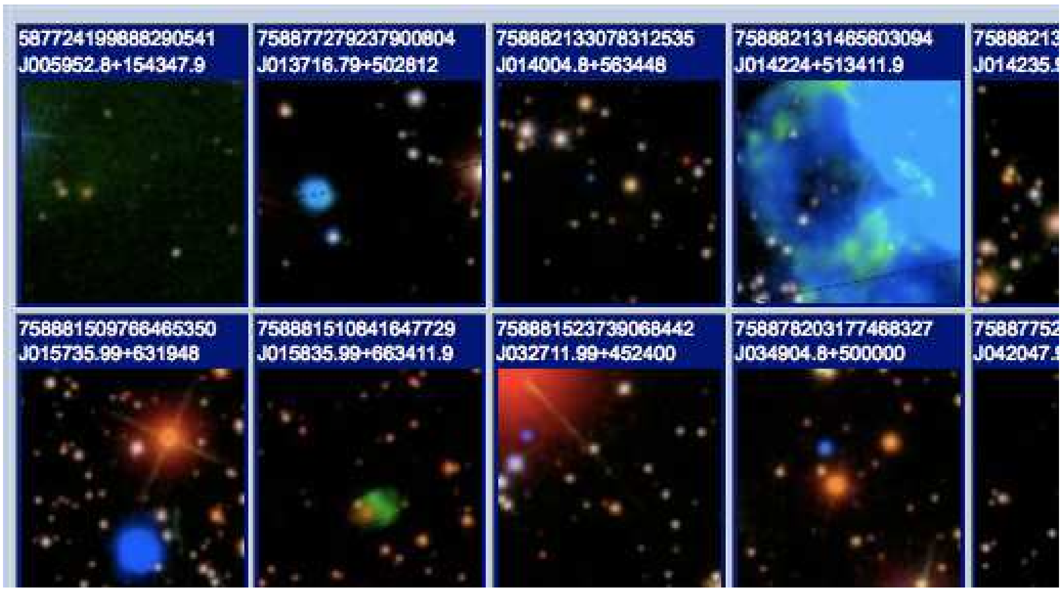


Figure 3: Planetary Nebulae

d) Match to get photometry - this is easy because the previous search stored the object ID so we can use this to JOIN with the PhotoPrimary table:

```
SELECT p.ra, p.dec, pp.u, pp.g, pp.r, pp.i, pp.z
FROM
MyDB.PNMatch as p
JOIN
PhotoPrimary as pp
ON
p.objid = pp.objid
```

e) Finally to create images you can carry out a simple query:

```
Select objID, ra, dec
FROM MyDB..PNMatch2
```

to get the positions and names. Copy the results and paste into the ImageList application and you might get something like on Figure 3

# 6 Additional tasks

## 6.1 Problem 1 - Finding asteroids

a) The pixel size is given to be 0.394 arcsec and we expect to see shifts as small as 0.0394 arcsec over a time of 72s. Keeping in mind that there are 3600 arcseconds in a degree and 24*3600 seconds in a day we get:

$$v = \frac{\Delta\, x\, [arcsec]}{\Delta\, t\, [s]} = 24 \frac{\Delta\, x\, [deg]}{\Delta\, t\, [day]}$$

or, when putting in numbers, 0.013 degrees/day.

b) Finding columns giving velocities is easily done by searching on the schema page. The rowv and colv columns contain the necessary velocity information

As was informed later, the settings suggested in the problem set were a bit too aggressive, thus a simple query that does work is:

```
select objID, u, g, r, i, z,
sqrt( power(rowv,2) + power(colv, 2) ) as velocity into
mydb.AsteroidsGood from PhotoObj
where
r < 20
and
rowv >= 0 and colv >=0
AND
(power(rowv,2) + power(colv, 2)) between 0.017 AND 0.25
```

c) The total number of matches from this was 174124. If you just look at a plot of u versus g you see plenty of objects with bad photometry. If you select only the good photometry ones, for instance using a criterion: $g > 10$ & $g < 22$ & $r > 10$ & $r < 22$ & $i > 10$ & $i < 22$, the subset will have 124,770 objects and you can use this to create colour-magnitude diagrams.

## 6.2 Problem 2 - Exoplanets

a)

b) Uploading a table to CasJobs can be iffy because the interface does not deal well with empty columns. Thus the simplest approach is to create a copy of the exoplanet table including only the positions of the star and some information available for all objeects. You can created a file with the following columns: `Pl_Name`, `Pl_Mass`, `St_Name`, ra, dec.

The problem set asks to match this to the Star table. This is entirely possible but it is more challenging than was intended. Thus let us first look at what was intended, namely to make use of the fGetNearestObjIdEq function which finds the objId of the closest object in the SDSS PhotoPrimary that is within some specified distance of your object.

```
SELECT e.ra as ra, e.dec as dec, e.Pl_Name as Name,
dbo.fGetNearestObjIdEq(e.ra, e.dec, 10.0) as objId
FROM MyDB.Exoplanets as e
WHERE
 e.ra IS NOT NULL
AND
e.dec IS NOT NULL
AND
dbo.fGetNearestObjIdEq(e.ra, e.dec, 10.0) IS NOT NULL
```

This works well and is fast with a total of 102 matches. The first ten are shown as images on Figure 4.

Now this does get you stars and it does tell you that the stars are saturated (overexposed) and hence not useful for doing science and you see that . So there is not really much need to match to the Star table. [the 5 repeated matches on the bottom row are five exposures of the central star in a extra-solar system with 5 planets - 55 Cnc.

If you want to match to the Star table, how would you do that? The simplest is to use the previous query to save a table in your MyDB. This table now contains the objID of the closest object and we now want to find the closest star to this - we can do this using the Neighbours table in CasJobs:

```
SELECT e.ra as ra, e.dec as dec, e.Name as name, e.objId, n.objID, n.type
FROM MyDB.ExoMatch3 as e
JOIN Neighbors as n
ON e.objID = n.objiD
WHERE
n.neighborObjId =(select top 1 nn.neighborObjId
from Neighbors as nn
join StarTag as s ON nn.neighborObjId = s.objID
-- Notice that the Match catalog comes here too!
```

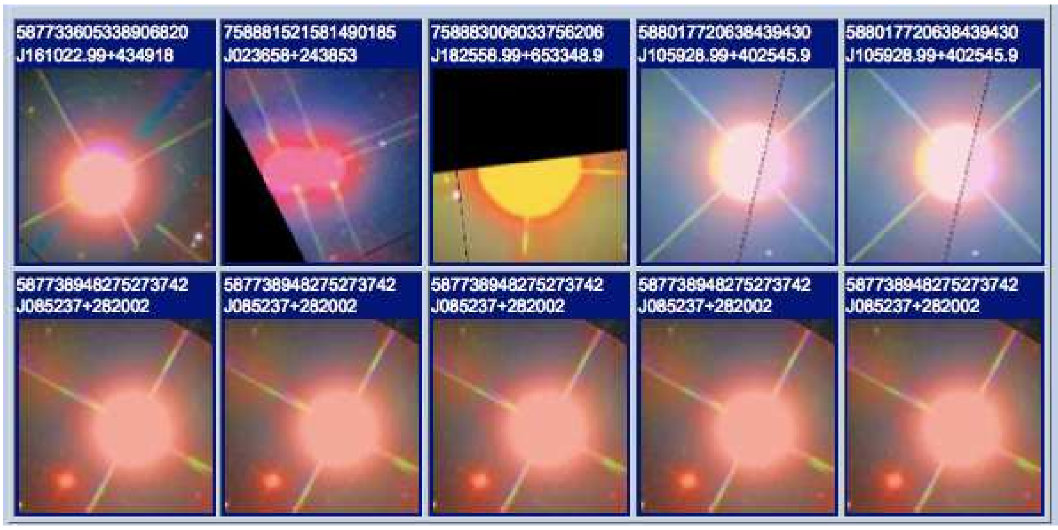


Figure 4: Exoplanets

```
where nn.objId = e.objId
order by nn.distance)
```

That achieved we order according to distance (closest first) and return the first neighborObjId to the wherestatement where it all started.

It reduced the list to 93 matches.

c) To find a table on the habitability of exo-planets it is best to use Vizier. I found a catalogue Habitability of known exoplanetary systems (Jones+, 2006) which satisfies the requirement. Load this catalog into CasJobs and match it with existing table by coordinates as in the previous solutions.

d)