



Easing the Pain of Astronomical Database Access

<http://goo.gl/4X147>



A little history

- Started life as part of the Pan-STARRS project in Hawaii
- Pan-STARRS has a lot of data and a very complex schema
- Very difficult for the novice to retrieve anything meaningful
- Queries have a habit of submitting but never returning

<http://goo.gl/4X147>

The Pan-STARRS web interface

- Existing SDSS-like web interface
- Offers a query box and a basic query builder

but...

- Results need to be manually downloaded...
- ... then loaded into some desktop application for analysis

<http://goo.gl/4X147>

Objectives of VODb

- To make it easier to write SQL for astronomical databases
- To utilize existing VO software to analyze the results
- To provide the ability to work on data offline

<http://goo.gl/4X147>

VODb

Color-color all 330, 0.3

Pan-STARRS PS1_SAS_09

```

1 SELECT gMeanPSFMag - rMeanPSFMag, rMeanPSFMag - iMeanPSFMag
2 FROM Object
3 WHERE objID IN
4 (SELECT objID FROM dbo.fGetNearbyObjEq(330, 0.3, 0.8))
5 AND nDetections > 5
6 AND gMeanPSFMag != -999
7 AND rMeanPSFMag != -999
8 AND iMeanPSFMag != -999

```

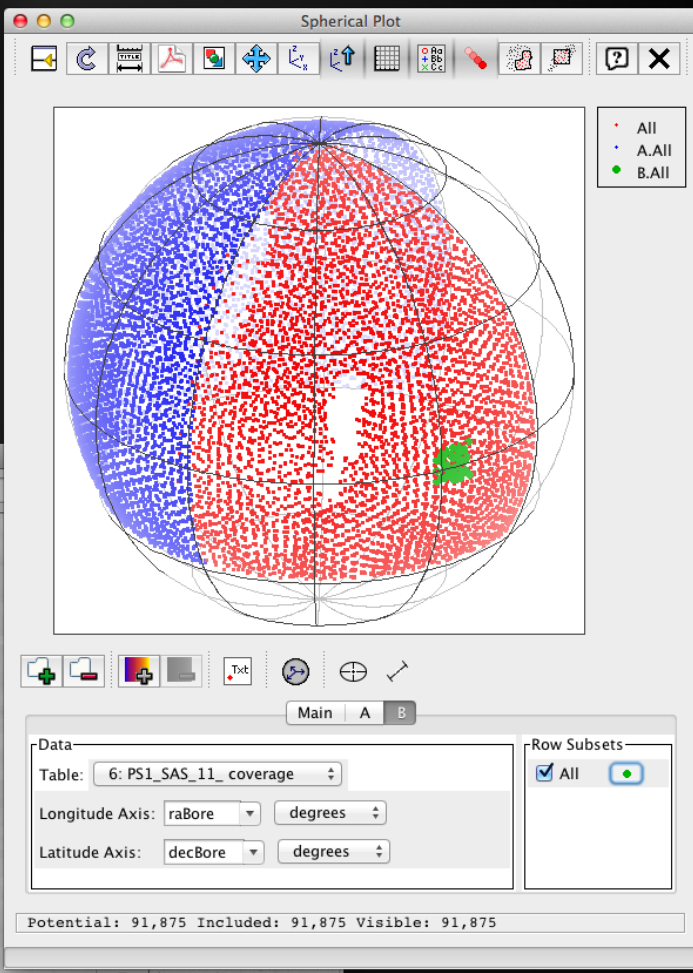
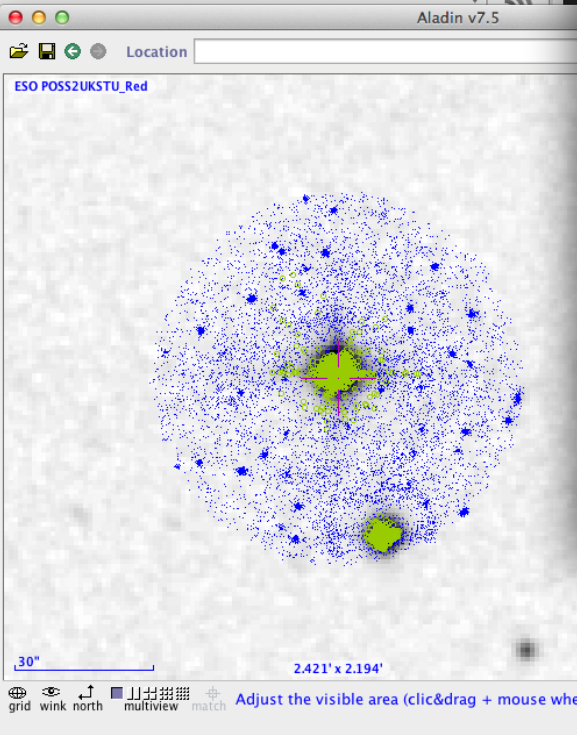
Edit Query View Diagnostics

00:09 : Complete

Send to all SAMP clients

Column1	Column2
-2.519	
1.239	
0.141	
0.45	
0.27	
1.604	
0.18	
0.752	
1.197	
1.114	
0.851	
48,706	

Derby SDSSDR8



Can't other VO tools do this?

Yes, but...

Tools like TOPCAT enable database querying as an *extra*.
VODb is a *dedicated* astronomical database access tool that
tries to apply the UNIX philosophy of

"Do one thing and do it well"

<http://goo.gl/4X147>

Some features

- A user-friendly 'double-click' desktop application
- SQL syntax highlighting
- Graphical query builder
- Local 'Derby' database to store results
- Astro date converter
- Astro coordinate converter
- Mask generator (e.g. for data quality flags)
- Query diagnostics display
- SAMP communication (pass data to TOPCAT etc)
- Query history (SQL, query time, rows returned etc)
- 'Plug-in' queries

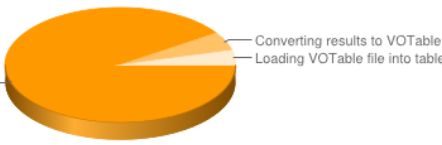
<http://goo.gl/4X147>

VODb

Color-color stars 330, 0.3 | All connections | coverage

Pan-STARRS | PS1_3PI

Query diagnostics. Total time: 00:04



Running query and downloading results

Converting results to VOTable

Loading VOTable file into table

Edit Query | View Diagnostics

00:04 : Complete

Send to all SAMP clients

frameID	raBore	decBore	
73406		266.093	46.74
73409		266.402	32.827
73411		266.42	37.843
73412		265.765	50.819
73413		266.386	42.276
73418		266.093	46.74
73421		266.42	37.843

Query Builder

Diagram...

DetectionCalib

- objID
- detectID
- ippObjID
- ippDetectID
- filterID
- surveyID
- ra
- dec
- raErr
- decErr
- calMag
- calMagErr
- calFlux
- calFluxErr
- dataRelease

Filter

- filterID
- filterType
- filterSpec

Survey

- surveyID
- name
- description

designer | syntax

Submit

Mask generator

Detection Flags

- Select/deselect all
- DEFAULT: Initial value: resets all bits
- PSFMODEL: Source fitted with a psf model (linear or non-linear)
- EXTMODEL: Source fitted with an extended-source model
- FITTED: Source fitted with non-linear model (PSF or EXT; good of
- FITFAIL: Fit (non-linear) failed (non-converge; off-edge; run t
- POORFIT: Fit succeeds; but low-SN; high-Chisq; or large (for
- PAIR: Source fitted with a double psf
- PSFSTAR: Source used to define PSF model

Binary

10101

Decimal

21

Hexadecimal

0x15

Cancel

Coordinate converter

RA HH:MM:SS.S

03:00:00.000

Dec +/-DD:MM:SS.S

+45:00:00.000

Degrees -90.0 -> 360.0

45.0

Cancel

Date converter

Date YYYY-MM-DD HH:MM:SS

1972-03-24 00:00:00

MJD 0.0 -> 60000.0

41400.1

JD 2400000.5 -> 2460000.5

24414C

Cancel

Why plug-in queries?

- To use a query builder requires a thorough understanding of the schema
- Sample queries are helpful, but cumbersome (copy-and-paste, edit)

<http://goo.gl/4X147>

What are plug-in queries?

- Parameterised sample queries
- User-friendly GUI generated for each query
- Parameters have constraints, eg $-90 \leq \text{decl} \leq 90$
- Defaults values enable quick successful queries
- Growing set of plug-ins maintained online
- Queries encoded in simple XML format...

Plug-in query XML

```
<query>
  <author>Millenium</author>
  <title>Tully-Fisher relation</title>
  <shortDescription>Find the Tully-Fisher relation...</shortDescription>
  <longDescription>Find the Tully-Fisher relation, Mag vs Vvir, for galaxies with a bulge/total mas
    <param>
      <name>PARAM_BULGE_MASS_RATIO</name>
      <description>Bulge/total mass ratio</description>
      <type>float</type>
      <max>1</max>
      <min>0</min>
      <default>0.1</default>
    </param>
    <param>
      <name>PARAM_SNAPNUM</name>
      <description>Snapnum (redshift)</description>
      <type>int</type>
      <max>63</max>
      <min>0</min>
      <default>41</default>
    </param>
  <sql>SELECT vVir, mag_b, mag_v, mag_i, mag_r, mag_k
  FROM millimil.DeLucia2006a
  WHERE (bulgeMass < PARAM_BULGE_MASS_RATIO*stellarMass OR bul
  AND snapnum = PARAM_SNAPNUM</sql>
</query>
```

Find the Tully-Fisher relation...

Bulge/total mass ratio 0.0 -> 1.0
0.1

Snapnum (redshift) 0.0 -> 63.0
41

Cancel OK

VODb

Color-color stars 330, 0.3

Pan-STARRS PS1_SAS_09

```

1 SELECT gMeanPSFMag - rMeanPSFMag, rMeanPSFMag - iMeanPSFMag
2 FROM Detection JOIN Object ON Detection.objID = Object.objID
3 WHERE Object.objID IN
4 (SELECT objID FROM dbo.fGetNearbyObjEq(330, 0.3, 0.8))
5 AND nDetections > 5
6 AND gMeanPSFMag != -999
7 AND rMeanPSFMag != -999

```

Color-color for just stars

Database: PS1_SAS_09

RA: 0.0 -> 360.0 degrees
Slider: 330

Dec: -90.0 -> 90.0 degrees
Slider: 0.3

Radius: 0.0 -> 90.0 degrees
Slider: 0.8

1st band (g, r, i ...): **p**

2nd band (g, r, i ...): r

3rd band (g, r, i ...): i

PSF likelihood >: 0.0 -> 1.0
Slider: 0.99

Buttons: Cancel, OK

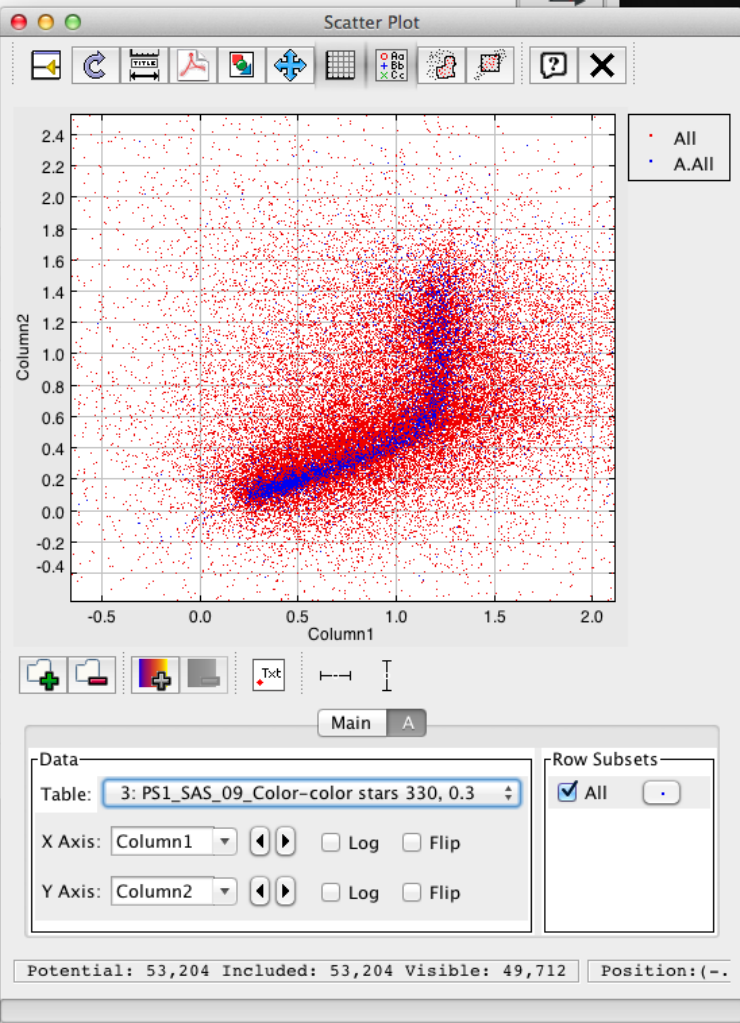
Edit Query View Diagram

00:25 : Comp

Column2
0.477
0.348
1.317
1.059
1.122
1.122
0.74
0.521
0.456
0.332
0.415

4,498 rows

Derby SDSSDR8 Pa



Learning by example

- It's difficult to write good SQL, easy to write bad SQL
- A badly placed JOIN can cost hours
- Plug-ins allow quick successful querying
- Allow users to adapt and enhance by example
- Help users improve SQL skills
- Help users and become familiar with schema

JDBC and TAP (and Pan-STARRS)

- Can connect to any JDBC-supported Db
- Can connect to any TAP-supported database
- Local Derby database means, for example, that Pan-STARRS data can be quickly compared to SDSS or 2MASS

<http://goo.gl/4X147>

VO compliant

- to plot your tables, pass them to TOPCAT
- to overlay catalogues on images, pass them to Aladin
- to perform statistical analysis, pass data to VOStat
- etc

<http://goo.gl/4X147>

Key principle: Abstraction

- Users don't need to worry about file formats
- Users don't need to know whether a database is TAP, JDBC or Casjobs
- Users are simply presented with a list of databases with meaningful names and descriptions
- e.g. Pan STARRS will likely never support TAP, but this should be irrelevant to end-users

Demo...

<http://goo.gl/4X147>

<http://goo.gl/4X147>