



Sharing SDSS Data with the World

Ani Thakar
Jordan Raddick

*Center for Astrophysical Sciences,
The Johns Hopkins University*

Outline

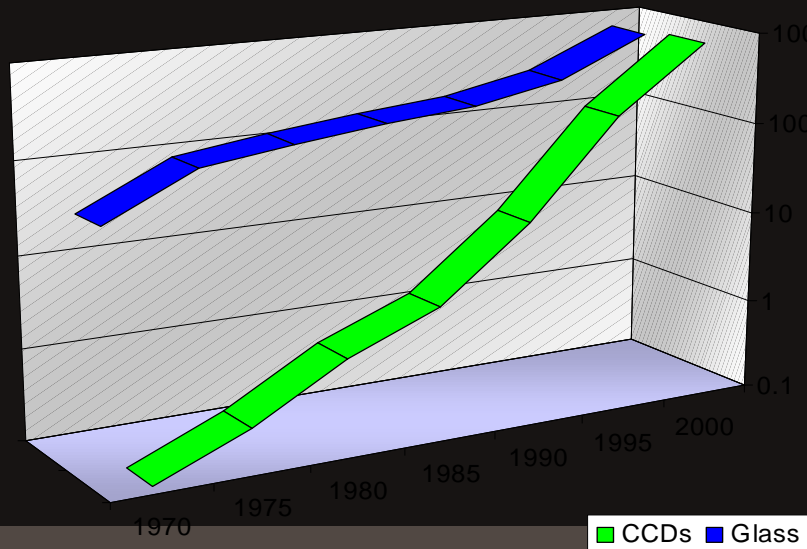


- Sharing Astronomy Data
- SDSS Overview
- Data Access Tools
- Data to Astronomers
- Data to the Public

Astronomy Data



- In the “old days” astronomers took photos.
- Starting in the 1960’s they began to digitize.
- New instruments are digital (100s of GB/nite)
- Detectors are following Moore’s law.
- Data avalanche: double every 2 years

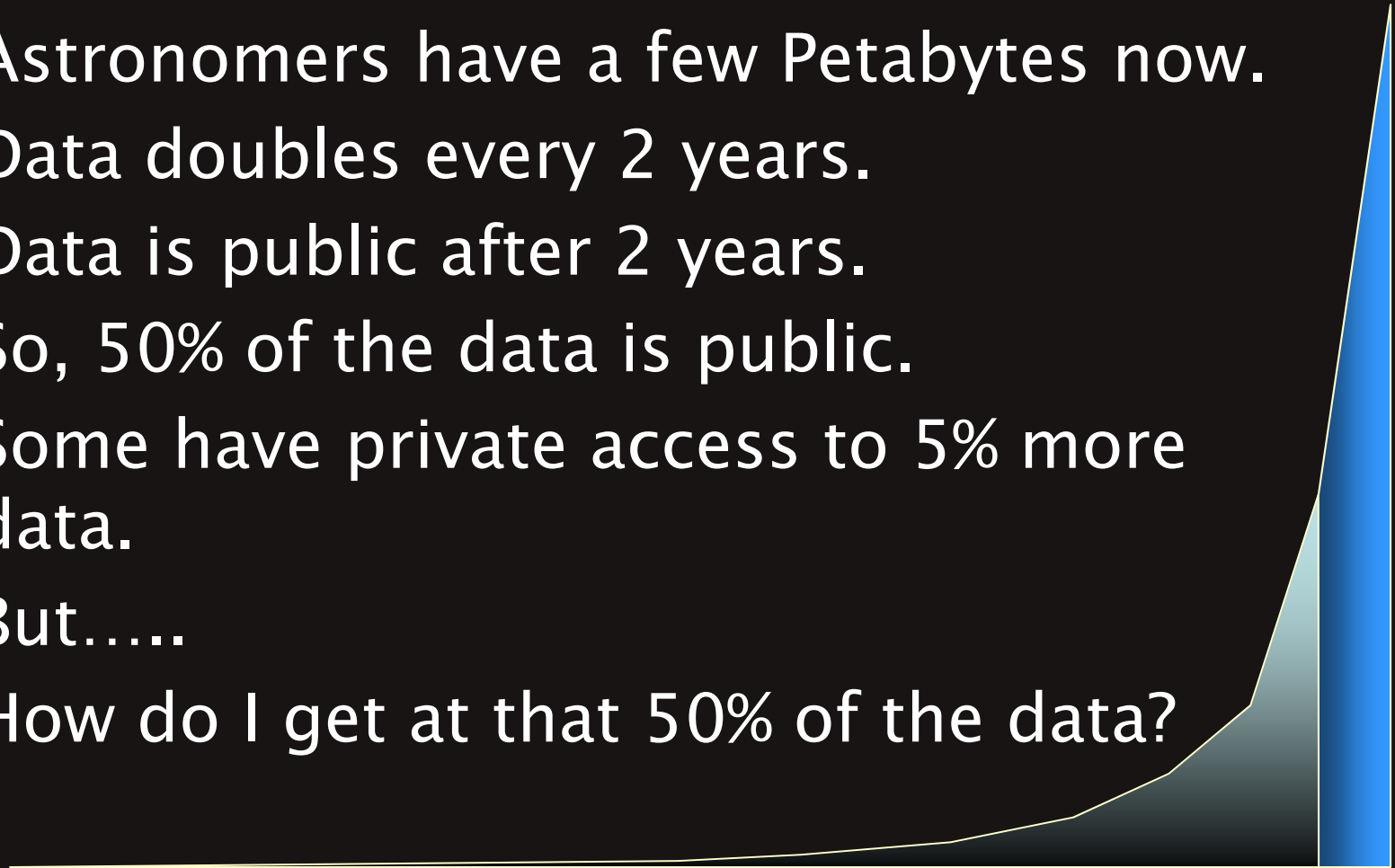


Total area of 3m+ telescopes in the world in m², total number of CCD pixels in Megapix, as a function of time. Growth over 25 years is a factor of 30 in glass, 3000 in pixels.

Astronomy Data



- Astronomers have a few Petabytes now.
- Data doubles every 2 years.
- Data is public after 2 years.
- So, 50% of the data is public.
- Some have private access to 5% more data.
- But.....
- How do I get at that 50% of the data?



New Astronomy– Different!

- Data “Avalanche”
 - The flood of Terabytes of data is already happening, whether we like it or not
 - Present techniques do not scale well
 - Drinking from the “Fire Hose”!
- Systematic data exploration
 - Will have a central role
 - Statistical analysis of “typical” objects
 - Automated search for the “rare” events
- Digital archives of the sky

Data Intensive Science



- Data avalanche in astronomy and other sciences
- Need new paradigm to deal with it
- Old file-based solutions do not cut it
- Old programming models don't work
- Scientists need to learn new tricks

Data Intensive Science Tools

- Databases instead of files (DBMSs)
 - Data integrity ensured
 - Optimized data access (DB indices)
 - Ability to define methods on data
 - Do the science *inside* the database!
 - Bring the analysis to the data, not vice-versa
- Scalable (parallel) data access
- High-speed transport protocols
 - Move the data rapidly when necessary
- Asynchronous Web services
 - Web browsers cannot handle data volumes

Virtual Observatory



- Premise: Most data is (or could be) online
- So, the Internet is the world's best telescope:
 - It has data on every part of the sky
 - In every measured spectral band: optical, x-ray, radio..
 - As deep as the best instruments (2 years ago).
 - It is up when you are up.
The observing conditions are always great (no working at night, no clouds no moons no..).
 - It's a smart telescope:
links objects and data to literature on them.
- **World-Wide Telescope coming soon!**

The Sloan Digital Sky Survey

Apache Point Obs., NM



SDSS Sky Coverage



www.sdss.org

SDSS Camera with 54 CCD chips



- Digital map in 5 spectral bands covering $\frac{1}{4}$ of the sky.
- Will obtain **40 TB** of raw pixel data.
- Photometric catalog with more than **200 million** objects.
- Spectra of \sim **1 million** objects.
- Data Release 6 – DR6: 350 M images, 1.1M spectra.

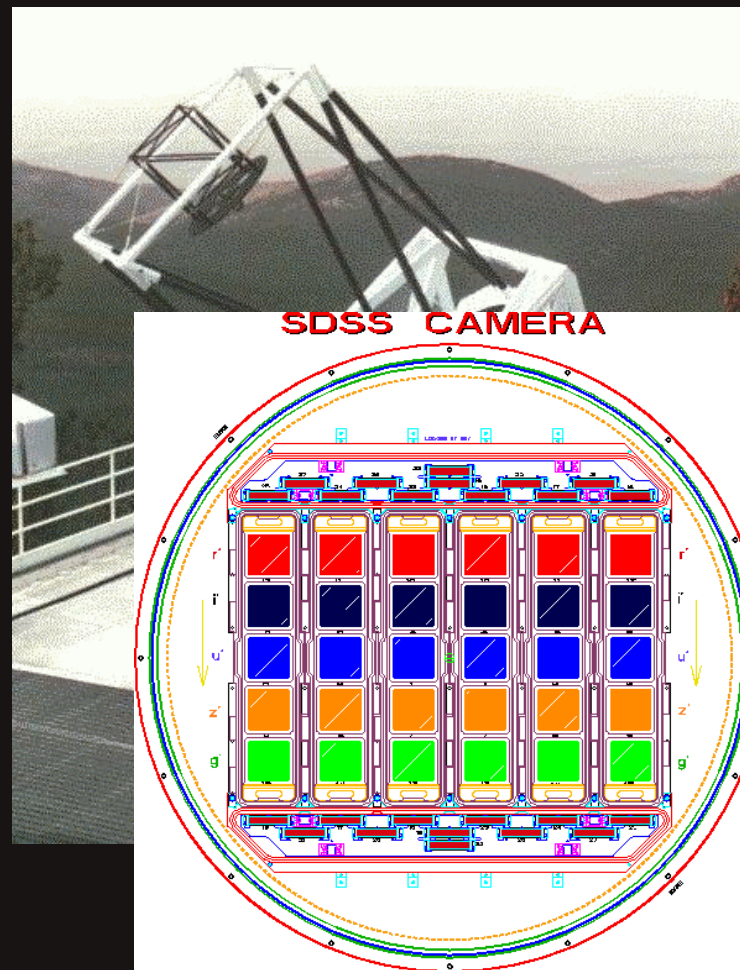
Next release ...

Data Release 7 – DR7: 400 M images, 1.2 M spectra.

Ground-breaking Instruments

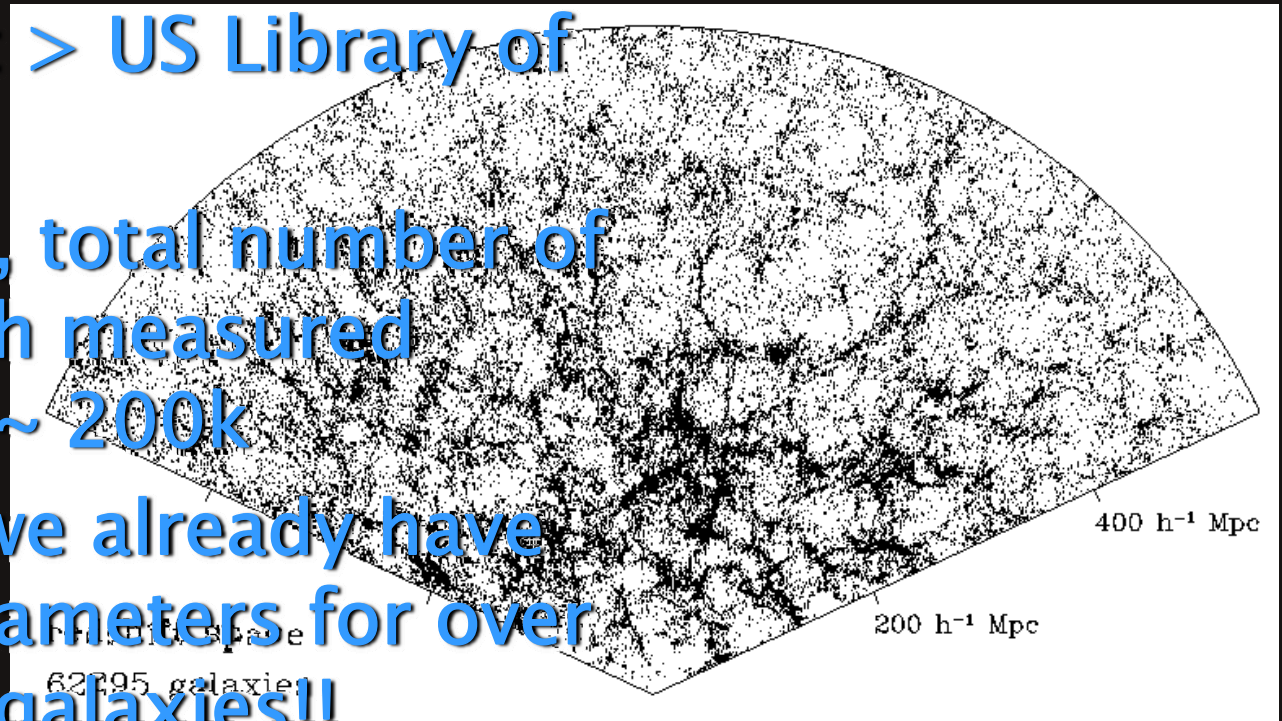


Multi-Fiber SpectroGraph (JHU)



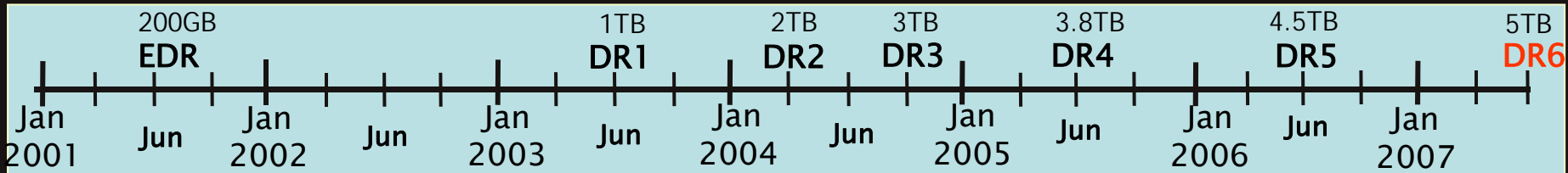
An Ambitious Survey

- Info content > US Library of Congress
- Before SDSS, total number of galaxies with measured parameters ~ 200k
- With SDSS, we already have detailed parameters for over 200 Million galaxies!!



A thousand-fold increase in the amount of data!

SDSS Data Releases



<i>Rel</i>	<i>Date</i>	<i>CAS Size</i>	<i>Images</i>	<i>Spectra</i>	<i>Sq Deg</i>	<i>CAS Mirrors</i>
DR6	6/29/07	5 TB	300 M	885k	8520	Hungary, India
DR5	6/28/06	4.5 TB	240 M	740k	8000	JHU, Portsmouth, STScI, Hungary, Moscow
DR4	6/29/05	3.8 TB	180 M	608k	6670	JHU, India
DR3	9/27/04	3 TB	141M	478k	5200	JHU, India, Portsmouth
DR2	4/15/04	2 TB	88M	330k	3324	JHU, UPitt, SDSC, Germany
DR1	6/15/03	1 TB	53M	186k	2099	JHU, SDSC, CDAC, UPitt UK, Germany, Japan, India
EDR	6/06/01	200 GB	14M	54k	462	JHU, SDSC, UK (ROE), Japan

SDSS Data Overview

Data Archive Server (DAS)

FITS files (raw data)

Images, spectra, corrected frames, atlas images, binned images, masks

Online form-based access

~~Binary and wget file retrieval~~

Catalog Archive Server (CAS)

Science parameters extracted to catalogs
Stuffed into relational DBMS (SQL Server)

Heavily indexed, optimized

Online access via SkyServer

Several levels of access, query tools

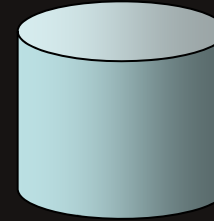
Data volume	catalogs (DAS, fits format)	2 TB			
	catalogs (CAS, SQL database)	4 TB			
Average wavelengths and magnitude limits	u	g	r	i	z
	1924.1	1708.3	1492.8	1241.1	1021.8

SDSS
Data
Release

The CAS Databases



SDSS Pipelines



- Processed data is stuffed into a commercial relational DBMS
 - Microsoft SQL Server 2000
- Allows fast exploration and analysis – Data Mining
- Two versions of the sky: Best and Target
 - Target is version of sky on which spectroscopic targets were chosen
 - Best is latest, greatest processing of the data
 - 2 DBs for each release, e.g. BestDR3 and TargDR3
- Heavily indexed to speed up access – HTM + DB Indices
- Short queries can run interactively.
- Long queries (> 1 hr) require a custom Batch Query System.

SkyServer Web Access

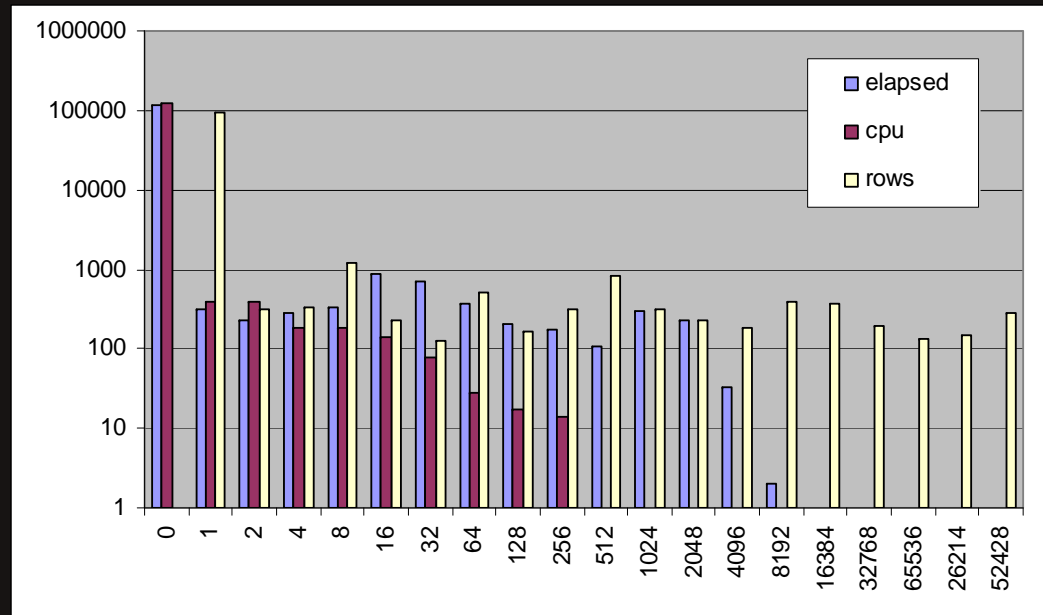


- Supports several levels of SQL access
 - Novice/casual users
 - Radial (cone) and Rectangular search
 - Intermediate/astro users
 - Imaging and Spectro Form Query
 - Expert
 - Free-form SQL, Object Crossid (upload RA/dec list)
 - CasJobs workbench environment (MyDB)
- Visual tools
 - ImageCutout service
 - Finding Chart, Navigate/browse images, image lists
 - Explore tool
 - Detailed info for each image object, including spectrum
- Downloadable interfaces
 - Emacs, Java tool (sdssQA), sqlcl (command-line)

<http://cas.sdss.org/>, <http://skyserver.sdss.org/>

CAS Workload Management

- Query execution time follows power law
- Vast majority of queries under a minute
- Separate short and long queries
- Execute long queries in batch mode



- Short and long queues (short=1 min, long=8hrs)
- Strictly limit time of query in a queue
- Provide user workspace DB – MyDB
 - Reduce network traffic from repeat and intermediate results
 - Allow sharing of query results between user groups

CasJobs and MyDB



- Batch Query Workbench for SDSS CAS
- Queries are queued for batch execution
 - Load balancing – queues on multiple servers
 - Limit of 2 simultaneous queries per server
- Short (1 minute) queue for immediate mode
 - Query aborted after 1 minute
- MyDB personal database
 - 1 GB (more on demand) SQL DB for each user
 - Long queries write to MyDB table by default
 - User can extract output (download) when ready
 - Ability to share MyDB tables with other users via group visibility

MyDB Features

- Tables
- Views
- Functions
- Procedures
- Extract
- Publish
- Create
- Drop
- Rename

MyDB - Mozilla Firefox

File Edit View History Bookmarks Tools Help

SDSS Query / CasJobs

Help Tools Query History MyDB Import Groups Output Profile Admin SkyServer Logout thakar

DR6

Views
Tables
Functions
Procedures

Sort by... No actions...

fBoundaryCircle
fBoundaryField
fBoundaryPrimary
fBoundaryPrim
fBoundaryRun
fBoundarySegment
fBoundaryStripe
fBoundaryTiGeom
fBoundaryTiPrimary
fCamcol
fCoordsFromEq
fCoordType
fCoordTypeN
fDatediffSec
fDistanceArcMinEq
fDistanceArcMinXYZ
fDMS
fDMSbase
fDocColumns
fDocColumnsWithRank
fDocFunctionParams
fEnum
fEqFromMuNu
fEtaFromEq
fEtaToNormal
fFiber
fField
fFieldMask
fFieldMaskN
fFieldQuality
fFieldQualityN
fFirstFieldBit
fFootprintEq
fFramesStatus
fFramesStatusN
fGetDiagChecksum
fGetJpegObjects
fGetLat
fGetLon
fGetLonLat
fGetNearbyFrameEq
fGetNearbyObjAllEq
fGetNearbyObjAllXYZ
fGetNearbyObjEq
fGetNearbyObjXYZ
fGetNearbyTiledTargetsEq
fGetNearestFrameEq
fGetNearestFrameIdEq
fGetNearestObjAllEq
fGetNearestObjEq
fGetNearestObjIdAllEq
fGetNearestObjIdEq
fGetNearestObjIdEqMode
fGetNearestObjIdEqType
fGetNearestObjXYZ
fGetObjectsEq
fGetObjectsMaskEq
fGetObjFromRect
fGetObjFromRectEq

fEtaFromEq

Notes Source

```
CREATE FUNCTION fEtaFromEq(@ra float,@dec float)
--/H Returns eta from ra,dec
-----
--/D This is the generic service function for SDSS specific
--/D coordinate conversions. It derives all the coordinates
--/D from the ra,dec, based on the definition of the primary
--/D areas of the stripes. For the Southern stripes, it returns
--/D a mu value shifted by 360 degrees, in order to have a
--/D non-negative and monotonic quantity.
-----
RETURNS float
AS BEGIN
DECLARE
@cx float, @cy float, @cz float,
@lambda float, @eta float,
@stripeEta float;
SET @cx = cos(radians(@dec)) * cos(radians(@ra-95.0));
SET @cy = cos(radians(@dec)) * sin(radians(@ra-95.0));
SET @cz = sin(radians(@dec));
SET @lambda = -degrees(asin(@cx));
IF (@cy = 0.0 and @cz = 0.0)
SET @cy = 1e-16;
SET @eta = degrees(atan2(@cz,@cy))-32.5;
SET @stripeEta = @eta;
IF @lambda < -180.0 SET @lambda = @lambda+360.0;
IF @lambda >= 180.0 SET @lambda = @lambda-360.0;
IF ABS(@lambda) > 90.0
BEGIN
SET @lambda = 180.0-@lambda;
SET @eta = @eta+180.0;
END
IF @lambda < -180.0 SET @lambda = @lambda+360.0;
IF @lambda >= 180.0 SET @lambda = @lambda-360.0;
IF @eta < 0.0 SET @eta = @eta+360.0;
IF @eta >= 360.0 SET @eta = @eta-360.0;
IF ABS(@lambda) = 90.0 SET @eta = 0.0;
IF @eta < -180.0 SET @eta = @eta+360.0;
IF @eta >= 180.0 SET @eta = @eta-360.0;
IF @eta > 57.5 -- 90.0-32.5
BEGIN
SET @eta = @eta-180.0;
END
RETURN @eta
END
```

Find: ingen

Next Previous Highlight all Match case

A. Thakar & J. Raddick,



Job Management

- Asynchronous
- Separate query and output jobs
- BatchAdmin DB keeps track of session, jobs, servers, queues and user privileges
- Job History page allows user to monitor, cancel, and resubmit jobs

SDSS Query / CasJobs

Home Help Tools Query History MyDB Import Groups Output Profile Admin Logout womullan

Refresh this page to get latest info

Status Target Submitted Name Like Apply

ANY ANY This Week

	Name	Query	Target	Submitted	Time(h:m:s)	Rows	Status
Info	My Query	select * from servers	BatchAdmin/1	1/21/2005 10:29:42 AM	0:0:0	0	Finished
Info	My Query	SELECT top 10 dbo.fSDSS(objId) as ID	DR4/500	1/20/2005 5:17:36 PM	0:0:0	10	Finished
Info	My Query	SELECT top 10 dbo.fSDSS(objId) as ID	DR4/500	1/20/2005 5:07:31 PM	0:0:1	10	Finished
Info	My Query	drop table mytable_3	MyDB/1	1/20/2005 5:06:56 PM	0:0:0	0	Failed
Info	My Query	drop table mytable_2	MyDB/500	1/20/2005 5:06:29 PM	0:0:0	0	Finished
Info	My Query	drop table mytable_1	MyDB/1	1/20/2005 5:06:25 PM	0:0:0	0	Finished
Info	My Query	drop table mytable_0	MyDB/1	1/20/2005 5:06:22 PM	0:0:0	0	Finished
Info	My Query	drop table mytable	MyDB/1	1/20/2005 5:06:18 PM	0:0:0	0	Finished
Info	My Query	select * from users where userid like 'k	BatchAdmin/1	1/20/2005 4:39:11 PM	0:0:0	0	Finished
Info	My Query	drop table mydb.mytable_2	MyDB/1	1/19/2005 9:34:43 AM	0:0:0	0	Failed
Info	My Query	drop table mydb.mytable_2	DR2/1	1/19/2005 9:34:33 AM	0:0:1	0	Failed

William O'Mullane
Last Modified :Thursday, January 20, 2005 at 5:04:35 PM, \$Name: v2_3_5 \$, \$Revision: 1.1.1.1 \$

Statistical Analysis



- Programmability in schema
 - Analysis stored right in the database
- Workbench environment
 - Can handle arbitrary computational queries
 - Results can be saved server-side in MyDB
 - Iterative analysis possible
- Provide capability to download results to their own machines for further analysis
 - Not trivial with TBs of data!!

Usage Logging

You have only two things to fear – failure and success.

– SQL Q

– IP don

• Analyze

• Logs ha

I don't know what the secret to success is, but the secret to failure is to try to please everyone.

• Invalua

designi

Sloan Digital Sky Survey / SkyServer

Web Site Traffic

Yearly, monthly, daily traffic (UTC GMT time). Counts are summarized as hits to the entire site and to the English, Japanese, German and the Projects sub-branches. There are other service branches, which are outside the language subtrees, so the sum of the columns do not add up to the hits.

The traffic on the education project weblog can be analyzed with the [SQL Search Page](#) or the [JHU CasJobs site](#).

Jim Gray

date	hits	English	German	Hungarian	Japanese	Spanish	Portuguese	Project	SkyServer	SkyService	SQL
Total	389904207	161360135	2500825	3087160	8214784	5194918	2383204	16690410	269485344	110128505	47327220
Yearly											
2008	12031850	6370389	147560	187027	2345	210277	144196	415144	11320712	711138	2010554
2007	152774705	51132700	1078198	1259260	503041	1954341	1437677	3353938	105568340	36908129	18517716
2006	93409785	32575824	814770	1019162	1842619	1816751	801329	4427810	59689360	33720425	9221372
2005	65905884	27830753	105305	547170	1288526	1104252	2	3514000	35478680	30427204	8259434
2004	33644996	22376283	75744	74527	1673737	106803	0	2254356	27492819	6152177	8123053
2003	18808934	11715659	90648	14	2242370	2494	0	1735133	17108133	1701229	1081253
2002	9738538	7004531	110644	0	456051	0	0	737353	9238570	506918	113838
2001	3589515	2853946	77956	0	36095	0	0	252676	3588730	1285	0
2008/01	10304859	5522483	122402	159766	2033	179190	124038	353719	9720954	583905	1816508
2007/12	1843313	933339	3666	9336	1006	7028	1878	95418	77542	3755	2857538
2007/10	12664999	4028830	90808	76034	302	170665	994	262206	7964516	4760383	1125176
2007/09	16407885	6720199	103798	109882	2294	212532	127527	253978	10675294	5732591	4594575
2007/08	22138674	7267105	94069	104728	2004	164843	102203	196688	14376446	7762228	2071399
2007/07	28150281	5878623	128803	152514	2832	180013	134549	222024	11549965	6302080	1527121
2007/06	7065624	2541830	89115	117134	5577	177821	107085	212652	6190157	875467	1042223
2007/05	9564368	2889552	84068	115126	20221	163114	146257	318766	7303322	2261046	341225
2007/04	9693840	2306460	147	147	147	142243	356591	8898274	795566	293970	
2007/03	7687967	2522715	115	115	115	146646	331270	7082859	605108	681688	
Daily (most recent)											
2008/02/7	105221	56358	505	658	38	1643	1151	5058	89115	16106	19957
2008/02/6	375266	188959	5785	4091	42	5773	4992	15300	352274	22992	24925
2008/02/5	51274	19011	596	646	32	700	0	1954	30333	20941	33748
2008/02/4	279834	135310	4478	5566	47	6521	3679	11467	262954	16880	21166
2008/02/3	285481	135022	4334	6545	60	5704	3186	10208	268096	17385	19216

Bill Cosby



Data to the Public: from Access to Learning

Jordan Raddick, Ani Thakar, Alex Szalay
Johns Hopkins University

Robert Sparks
National Optical Astronomy Observatory

Jim Gray
Microsoft Research

Countless others

Why data to the public?



- Commitment to data sharing
- Free choice – people can look at any star or galaxy
- **Inquiry** learning known to be effective
 - Learners answer a question themselves, with their own design

Audiences for data



- Formal education
 - Intro college students and teachers
 - K–12 students and teachers (science and more)
- Amateur astronomers
- Museum visitors
- Media
- General public
- **Access** must be tailored to each audience

Formal Ed: SkyServer

The screenshot shows the SDSS SkyServer DR5 website in a Mozilla Firefox browser window. The browser's address bar displays the URL <http://cas.sdss.org/dr5/en/>. The website header features the Sloan Digital Sky Survey logo and navigation links for Home, Tools, Schema, Projects, Astronomy, SDSS, Contact Us, Download, Site Search, and Help. The main content area is divided into several sections: a welcome message for DR5, a 'News' section highlighting the Data Release 5 (DR5), a 'For Astronomers' section with a link to a separate branch for professional astronomers, and four columns of links for SkyServer Tools, Science Projects, Info Links, and Help. The 'SkyServer Tools' column includes links for Famous places, Get images, Visual Tools, Explore, Search, Object upload, and CasJobs. The 'Science Projects' column includes links for Basic, Advanced, Challenges, For Kids, Games and Contests, Teachers, and Links to other projects. The 'Info Links' column includes links for About Astronomy, About the SDSS, About the SkyServer, SDSS Data Release 5, SDSS Project Website, Open SkyQuery, and Images of RC3 Galaxies. The 'Help' column includes links for Getting Started, FAQ, How To, Glossary, Schema Browser, Sample SQL Queries, and Details of SDSS Data. On the right side, there is a 'SDSS is supported by' section with logos for the Sloan Foundation, NSF, NASA, and MEXT. Below this is a 'Powered by Microsoft' logo and links for Site Traffic and Privacy Policy. At the bottom of the page, there is a 'Contact Us' section and a 'Done' status bar.

Browse Data

SDSS DR5 Navigate Tool - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://cas.sdss.org/dr5/en/tools/chart/navi.asp?ra=156.910563&dec=60.837925

Customize Links Free Hotmail Welcome to the I Do ... Windows Marketplace Windows Media Windows

DR5
SDSS
Home | Help | Chart | List | Exp |

Parameters

ra	202.48208	deg
dec	47.23139	deg
opt		

Get Image

Drawing options

- Grid
- Label
- PhotoObjs
- SpecObjs
- Targets
- Outline
- BoundingBox
- Fields
- Masks
- Plates
- InvertImage

Selected object

ra	202.63273
dec	47.38324
type	STAR
u	24.63
g	15.90
r	16.58
i	16.08
z	17.60

Explore
Recenter
Add to notes
Show notes

Click to open Sky Maps ?
To see Sky Maps, install the latest [Flash](#) and [Shockwave](#) players

Explore Data

Quick Look tool: SkyServer - Mozilla Firefox

File Edit View History Bookmarks Tools Help del_jcio.us

http://cas.sdss.org/dr6/en/tools/quicklook/quickobj.asp

Customize Links SDSS SkyServer DR4 Babelfish Subscribe...

skype US Search for: Google Web Search US +1

SDSS DR6

Summary
More Data
Search by
ObjID
Ra,dec
5-part SDSS
Plate-MJD-Fiber
Notes
Add to Notes
Show Notes
Finding Chart
Print
Help
Tutorial
Examples

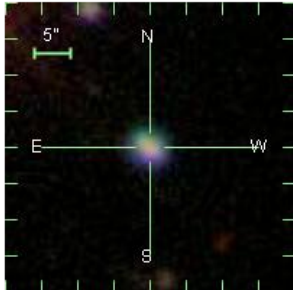
Summary data for: SDSS J113459.47+002509.1

Position Data (How do I find it?)

Object ID (objID):	Right ascension (ra):	Declination (dec):
588848900446814264	173.74781796	0.41921316

Image Data (What does it look like?)

Preview image (click to go to Navigate tool)



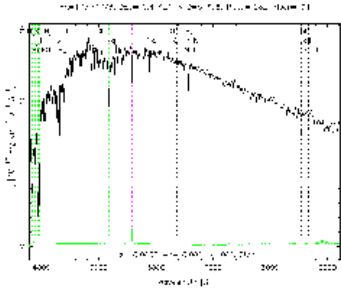
Object Type (type): GALAXY

Magnitudes:

Ultraviolet (u):	19.55 ± 0.03
Green (g):	18.04 ± 0.01
Red (r):	17.55 ± 0.01
Infrared - 7600 Å (i):	17.33 ± 0.01
Infrared - 9100 Å (z):	17.19 ± 0.02

Spectrum Data (What does its spectrum look like?)

Preview spectrum (click for a larger version)



Spectral classification (specClass): STAR

Redshift Data:

Redshift (z):	2.09106E-4
---------------	------------

[Get spectrum as CSV](#)

Done

Search for Data

SkyServer DR5 Search Form - Mozilla Firefox

http://cas.sdss.org/dr5/en/tools/search/form/form.asp

Sloan Digital Sky Survey / SkyServer

Home Tools Schema Projects Astronomy SDSS Contact Us Download Site Search Help

SkyServer Search Form

Mouseover the ? links for quick help, or click for the User Guide

Show me ? in the region ?

with magnitudes ?

<input type="text"/>	<	u	<	<input type="text"/>
<input type="text"/>	<	g	<	<input type="text"/>
<input type="text"/>	<	r	<	<input type="text"/>
<input type="text"/>	<	i	<	<input type="text"/>
<input type="text"/>	<	z	<	<input type="text"/>

and colors ?

<input type="text"/>	<	u-g	<	<input type="text"/>
<input type="text"/>	<	g-r	<	<input type="text"/>
<input type="text"/>	<	r-i	<	<input type="text"/>
<input type="text"/>	<	i-z	<	<input type="text"/>
<input type="text"/>	<	u-r	<	<input type="text"/>

Remember that the magnitude scale is backward!
For objects brighter than 18 in g, use g < 18.

for ?

Please return: ?

Num. of Objects: ?	<input type="radio"/> Count Only	<input type="radio"/> 10	<input type="radio"/> 100	<input type="radio"/> 1000	<input type="radio"/> All
Image Data: ?	<input checked="" type="checkbox"/> object IDs	<input checked="" type="checkbox"/> RA and Dec	<input type="checkbox"/> magnitudes		

Done

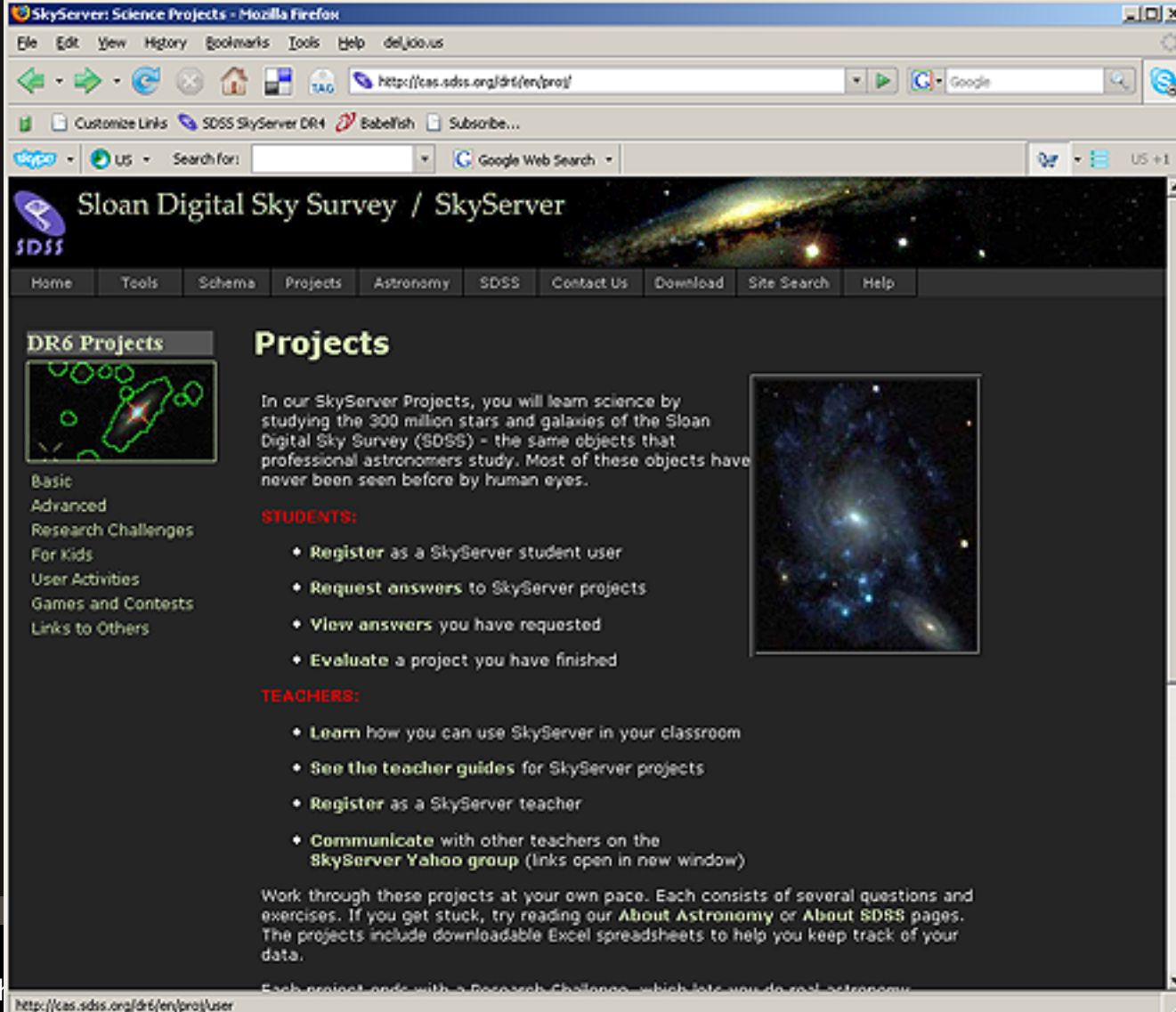
From Access to Learning



- Not enough just to give public access to data
- How are they using it?
- Are they **learning** anything from it?
 - Do they understand what they're doing and why?*

*Understanding = long-term memory, transfer to new situations (*How People Learn*)

Formal Education: Projects



Sloan Digital Sky Survey / SkyServer

Home Tools Schema Projects Astronomy SDSS Contact Us Download Site Search Help

DR6 Projects

- Basic
- Advanced
- Research Challenges
- For Kids
- User Activities
- Games and Contests
- Links to Others

Projects

In our SkyServer Projects, you will learn science by studying the 300 million stars and galaxies of the Sloan Digital Sky Survey (SDSS) – the same objects that professional astronomers study. Most of these objects have never been seen before by human eyes.

STUDENTS:

- Register as a SkyServer student user
- Request answers to SkyServer projects
- View answers you have requested
- Evaluate a project you have finished

TEACHERS:

- Learn how you can use SkyServer in your classroom
- See the teacher guides for SkyServer projects
- Register as a SkyServer teacher
- Communicate with other teachers on the SkyServer Yahoo group (links open in new window)

Work through these projects at your own pace. Each consists of several questions and exercises. If you get stuck, try reading our [About Astronomy](#) or [About SDSS](#) pages. The projects include downloadable Excel spreadsheets to help you keep track of your data.

Each project ends with a Research Challenge, which lets you do real astronomy.

<http://cas.sdss.org/dr6/en/proj/user>

SkyServer projects



- Three levels
 - Kids (K–8)
 - Basic (high school or Astro 101)
 - Advanced (skilled and motivated students)
- Research challenges
 - Independent research (open inquiry) with data
- Activities created by users

SkyServer projects

Types of Stars - Your Classifications - Mozilla Firefox


File Edit View Go Bookmarks Tools Help

http://cas.sdss.org/dr5/en/proj/basic/spectraltypes/studentclasses.asp

Customize Links Free Hotmail Welcome to the I Do ... Windows Marketplace Windows Media Windows


Home Tools Schema Projects Astronomy SDSS Contact Us Download Site Search Help

DR5 Projects



- Basic
- Scavenger Hunt
- The Universe
- Asteroids
- **Types of Stars**
- Color
- Galaxies
- Advanced
- Challenges
- For Kids
- User Activities
- Games and Contests
- Links to Others

Types of Stars



Your Classifications

Spectral Lines

Done

Classifying Stars

Pretend that you are an astronomer living in the early 1900s, before the OBAFGKM spectral types were developed. You are one of the first astronomers who ever looked at spectra of stars, and it is up to you to invent a way to classify them.

The table below shows a list of the first stars you are trying to classify. Spectra in the SDSS are sorted by Plate and Fiber number. Click the fiber numbers below to go to the Object Explorer entry for each star. When you click the first number, the tool will open in a new window; when you click another number, the entry for the new star will appear in the same Object Explorer window.

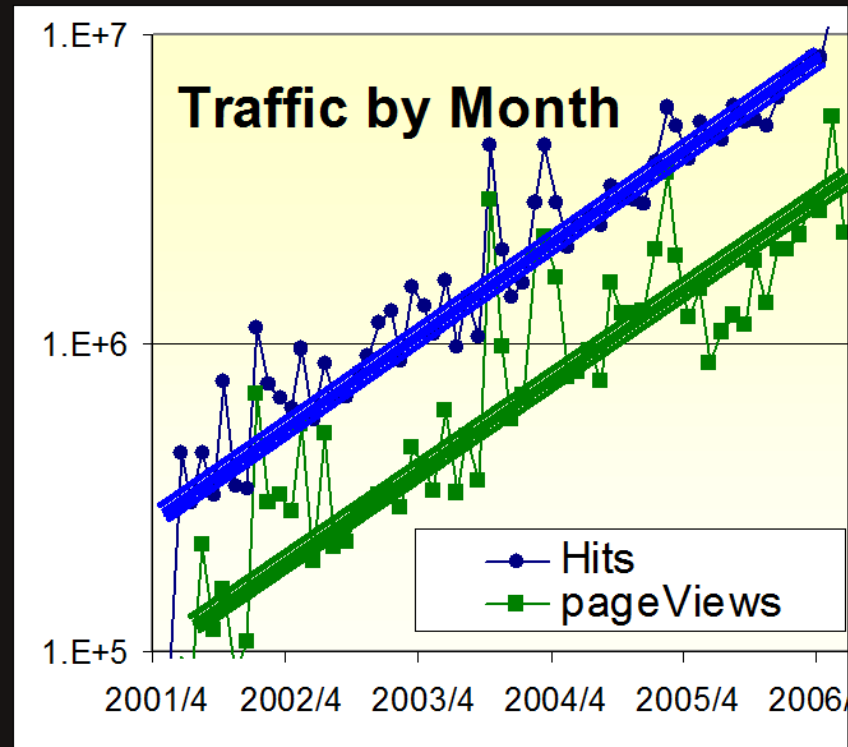
Once you have a star loaded into the Object Explorer, scroll down in the left-hand frame and click "Spectrum." You will see a full-sized picture of the star's spectrum. If you click "Print," you can print out the spectrum.

<i>Plate</i>	<i>Fiber</i>	<i>Plate</i>	<i>Fiber</i>
266/51630	21	266/51630	173
266/51630	275	266/51630	314
266/51630	365	266/51630	513
273/51597	2	498/51984	538
273/51597	157	273/51597	245
273/51597	589	281/51614	3
281/51614	4	281/51614	133

Look at the spectra of these 14 stars, and divide the spectra into several groups. There is no set number of groups you should have, and the groups do not have to have equal numbers of stars (because not all types of stars are equally common). If you find a

Use of educational projects

- 2 of top 20 IPs to hit entire site are K-12 districts
 - Orlando, FL & Los Angeles



Effectiveness of projects



- Not just who's using it, but how effectively
- How much are students really learning?
- Results are...

Effectiveness of projects



- (education research is really, really hard)
- This is a top priority in the future

Lessons Learned



- Users appreciate trust implied by giving them real data
- Public tools also used by astronomers
 - But astronomers want too many options for public
 - “Scope creep” is deadly
- SkyServer projects are good for high-achieving students with knowledgeable teachers
 - But it confuses others

Lessons Learned



- Instructional design is easy to do, hard to do right
 - Teacher guides must be very detailed
- Huge audience for citizen science
 - Rare chance to show “science in action”
 - But research problem must be authentic

Decisions to Make



- (for SkyServer and for all of us)
- Who is our audience in formal education?
 - Top 2% of students?
 - People getting only one exposure to science?
 - Some of both?
- What interactive tools can we make for museums?
 - You have 30 seconds to get attention, teach

Contact information



Contact Information

Jordan Raddick

410-516-8889

raddick@jhu.edu

Citizen science: Galaxy Zoo

- Background: galaxies come in different shapes



spiral



elliptical

- Classifying hard for computers, easy for people
- So get the public to help!

Galaxy Zoo



Galaxy Analysis | Galaxy Zoo - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.galaxyzoo.org/GalaxyAnalysis.aspx

SDSS SkyServer DR6 | SDSS DR6 Navigate ... | SkyServer DR6 Search... | Home | Galaxy Zoo | Data in Education | PocketVO Applet - St... | AltaVista - Babel Fish...

GALAXY ZOO.org

Hi **astromoose** | Home | The Science | How to Take Part | Galaxy Analysis | Forum | Press & News | FAQ | Links | Contact Us | Logout | Profile

- Galaxy Tutorial
- Galaxy Analysis
- Current Statistics
- Show My Galaxies

Galaxy Analysis

Welcome to Galaxy Zoo's view of the Universe. If you're here you should already have seen the [Tutorial](#), but feel free to go and remind yourself. There's no need to agonise for too long over any one image, just make your best guess in each case.

Galaxy Ref:
588023670777577606

Choose the Galaxy Profile by clicking the buttons below

CLOCK ANTI EDGE ON / SPIRAL GALAXY

ELLIPTICAL GALAXY

STAR / DON'T KNOW MERGERS

Done

Citizen science learning



- Even people who love science often don't understand peer review, proposals, etc.
- Want to use Galaxy Zoo to show **process** of science
- As we write science papers, we share results with volunteers

Galaxy Zoo blog

Galaxy Zoo Blog - Do galaxies care where they live? - Mozilla Firefox

File Edit View History Bookmarks Tools Help del.icio.us

http://www.galaxyzoo.org/2008/01/21/morphology_environment_1/

Customize Links SDSS SkyServer DR4 Sabelish Subscribe...

US Search for: Google Web Search US +1

ABOUT

GALAXY ZOO.org

Galaxy Zoo Blog

The official blog of the Galaxy Zoo project

JAN 21 **DO GALAXIES CARE WHERE THEY LIVE?**

Does where we live make a difference to the kind of person we are? This is a question that has been addressed many times by social scientists, and certainly with more refined thought than the following example, but it will serve our purposes.

Consider one person, Victor, living in a small countryside village, and another, Claire, who lives in the centre of a city. The nearest shops to Victor are many miles away. When he has a sudden biscuit craving and opens the cupboard to find, to his horror, that his wife finished off the last packet the previous evening, it is a great effort for him to travel to the shops to get another. Claire, on the other hand, has merely to stroll to the corner of her road to satisfy her craving for something crunchy. However, while Claire often finds herself nipping out for a packet of biscuits, Victor rarely has the need. He always makes sure he buys plenty of biscuits on his regular weekly shopping trip, and there is always the packet hidden at the back of the other cupboard that his wife hasn't noticed. Victor is very organised, while Claire clearly isn't, at least when it comes to biscuits. Does this have anything to do with where they live?

Of course, biscuit buying habits, although important, aren't the only thing one can say about an individual. Each person is complex and unique, imperfectly describable even by a very large number of personality traits. However, there are simple and obvious ways of crudely dividing up the population. Although we have so far confined ourselves to biscuits, the chances are that Victor is generally more organised than Claire. Perhaps there is a way of dividing people into groups by how organised they are. I've no idea, but there are small

SEARCH

CATEGORIES

- > Anze
- > Gill
- > Chris
- > Jordan
- > Kate
- > Kevin
- > Steven

ARCHIVES

- > February 2008
- > January 2008

Done

A.

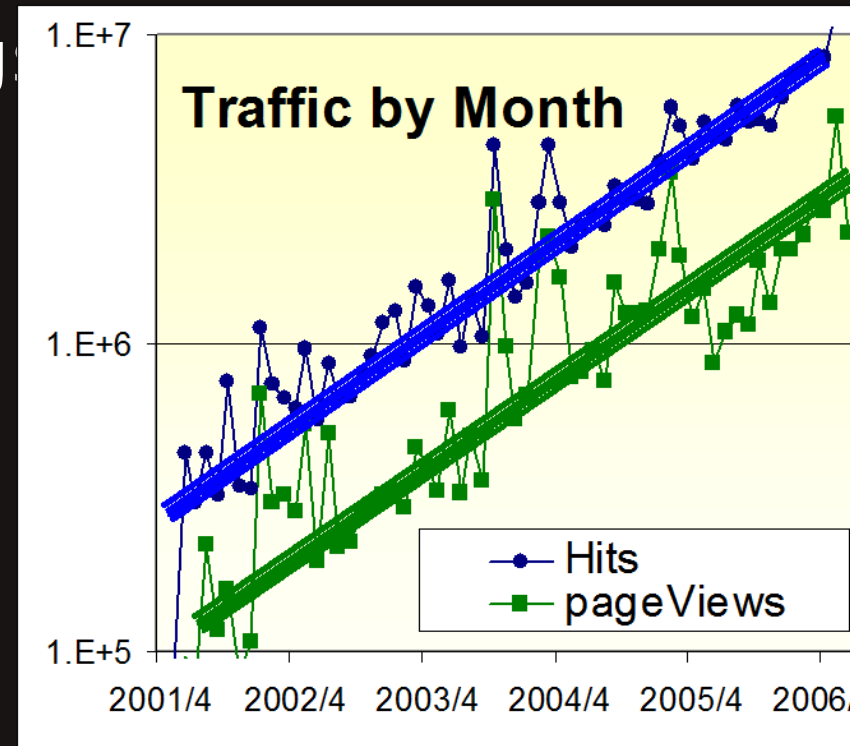
Use of projects



- What projects are users hitting? When?
- (Warning! Warning!): Not yet cleaned data
- Most hits on Tues, Thurs; fewest on weekend
- Looked at project hits in April 2006
 - School in session
 - But, hit counts to projects are *underestimates*, because people have to get data too

Use of educational projects

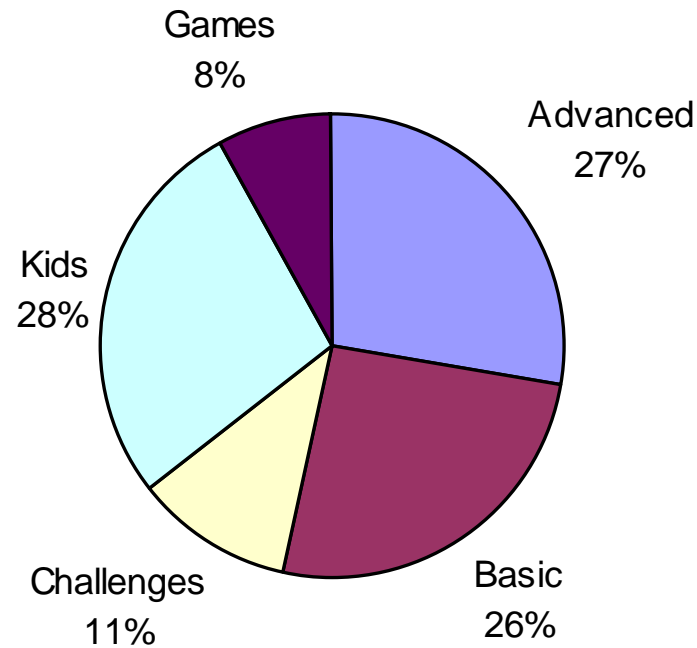
- Count hits with hit log
- Cleaned data
 - Removed webcrawler hits
- SQL analysis
- Hits & page views doubled each year
- 2 of top 20 hitting groups (by IP) are school districts
 - Orlando, FL & Los Angeles



Use of projects



Hits/projects by type (total hits / # projects)



Future Needs



- Back up all design decisions with education research
 - Now working with JHU School of Education
- **Study**: does use of real data make a difference for student understanding of X?
 - X = astronomy content, process of science, attitude toward science
 - Ed.D. dissertation in waiting to happen
- Need a community of people interested in this