

Big Data in Astronomy: Hands on with Large Surveys

AST 597 B

mjuric@astro.washington.edu

Today

- Administrivia
- Introduction to Large Surveys

About Me

- Mario Juric (mar-ee-oh you-rich)
- What I do:
 - Science derived from large surveys: Galactic structure, properties of the solar system
 - Astronomical algorithms and software research
 - Optical astronomer (=> some bias towards optical surveys)
- Hats I wear
 - Professor @ Astronomy
 - Sr. Fellow @ UW's eScience Institute
 - Data Management Project Scientist @ LSST

About Me

- Coordinates:
 - Office: C320
 - E-mail: mjuric@astro.washington.edu
- Best time to find me:
 - After class (except today!)
 - Fridays, 11:30am-12:30pm

About This Course

- The goal of this course is to prepare you for research with large survey data, teach you how to think about such data sets, and give you an overview of what is or soon will be available.
- We will:
 - Learn about how large observational data sets are changing astronomy
 - Introduce tools and techniques for working with large data sets (incl. access, analysis, and visualization)
 - Introduce and work hands-on with existing data sets (including SDSS, WISE, Kepler and others)
 - Learn about major upcoming surveys (including PanSTARRS, Gaia, and LSST)

Learning about Learning

- In survey astronomy (and related data-heavy disciplines), things change too quickly to be learned only once – it's a process that never ends.
- The major emphasis of this course will be on showing how to independently find information, evaluate, and use it.

Mailing List

- astr597b_wi16@uw.edu
- Some of you may be auditing
 - If you're not enrolled, e-mail me your e-mail address so I can add you to the class mailing list
 - mjuric@astro.washington.edu
- All announcements, etc., will go via that list

Lectures

- TTh, 1:30-2:50pm
 - Tuesday: Lecture
 - Thursday: Hands-on work
 - Will deviate from this pattern early on
- Lecture structure:
 - About an ~hour for introducing new materials, etc.
 - Leave ~20 minutes for open-ended discussion
 - **Interrupt and ask questions at any time!**

Course Materials

- I'll be adding most of what we need to the following repository on GitHub:

https://github.com/mjuric/astr597b_wi16

Smile, You're On Camera!

- Actually, it's mostly me who's on camera...
- Actually, it's mostly my slides that will be on camera...
- This is to help you review lectures later on, or to catch up in the (infrequent!) case you had to miss one.
- Viewing the recording later, via YouTube
 - I'll e-mail the link after class

Syllabus

- https://github.com/mjuric/astr597b_wi16
 - Still being finalized
- Three parts
 - Software / Data Analysis Tools (~ January)
 - Major Surveys (~ February)
 - Future Surveys and Data Challenges (~ March)

Grading, Homeworks, Etc.

- Two homeworks (50% of the grade):
 - All homeworks will be IPython notebooks, turned in via GitHub
 - HW #1, due end of January, focused on software tools
 - HW #2, due end of February, focused on survey data analysis
- Paper presentations (25% of the grade):
 - Present a paper on a survey / dataset of your choosing, making sure to cover:
 - The goals of the survey
 - The data it contains (and any caveats about it)
 - What it may be useful for
 - The best ways to obtain it
- Final project (25% of the grade):
 - More about this in a week or so ☺.

Questions?

- Large Survey and Why They're Different

Sur·vey

verb

/sər'vā/

1. (of a person or their eyes) look carefully and thoroughly at (someone or something), especially so as to appraise them.

"her green eyes surveyed him coolly"

synonyms: look at, look over, [observe](#), [view](#), [contemplate](#), [regard](#), gaze at, stare at, [eye](#); [More](#)

2. examine and record the area and features of (an area of land) so as to construct a map, plan, or description.

"he surveyed the coasts of New Zealand"

noun

/'sər,vā/

1. a general view, examination, or description of someone or something.

"the author provides a survey of the relevant literature"

synonyms: [study](#), [review](#), [consideration](#), [overview](#); [More](#)

2. an act of surveying an area of land.

"the flight involved a detailed aerial survey of military bases"



Translations, word origin, and more definitions

Hipparchus of Rhodes (180-125 BC)

Discovered the precession of the equinoxes.

Measured the length of the year to ~6 minutes.

In 129 BC, constructed one of the first star catalogs*, containing about 850 stars.



n.b.: also the one to blame for the magnitude system ...

Galileo Galilei (1564-1642)

Researched a variety of topics in physics, but called out here for the introduction of the *Galilean telescope*.

Galileo's telescope allowed us for the first time to *zoom in* on the cosmos, and study the individual objects in great detail.





Joseph von Fraunhofer (1787-1826)

Mounted a prism in front of an objective of a small telescope, and pointed it to the Sun.

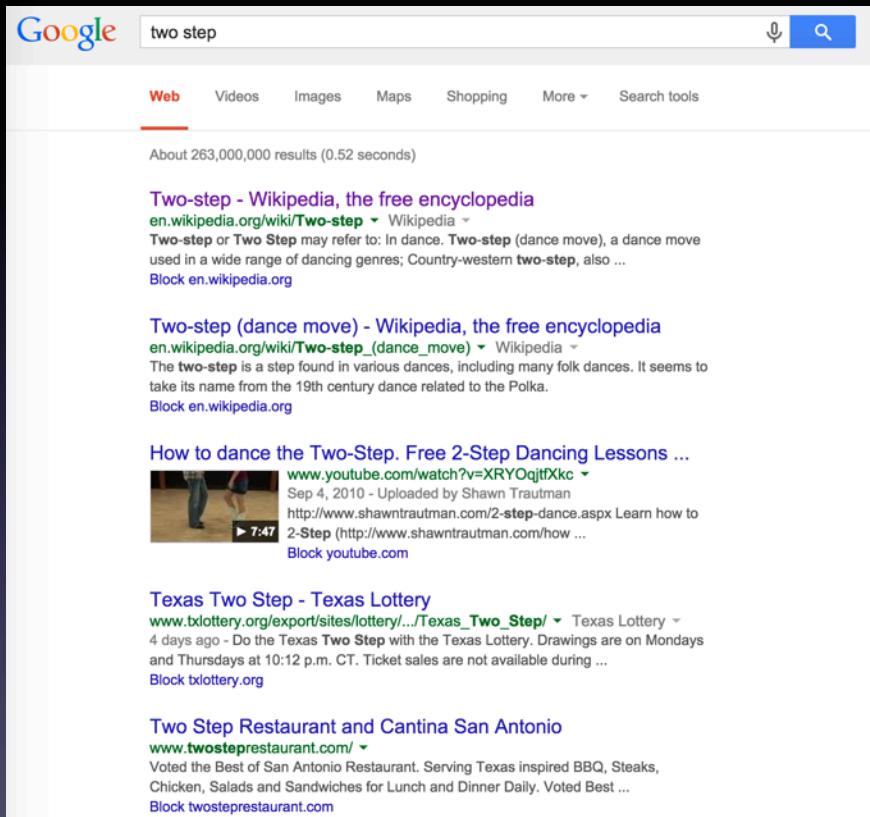
In 1859, Kirchhoff and Bunsen understood what Fraunhofer saw.

The birth of modern astrophysics!

The Astrophysics Two-Step

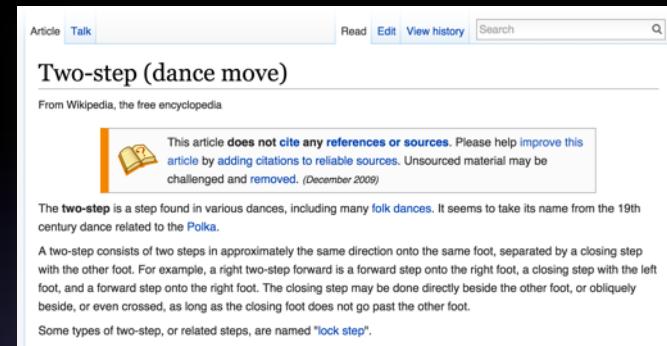
- Surveys
 - Construct catalogs and maps of objects in the sky. Focus on coarse classification and discovering targets for further follow-up.
- Large telescopes
 - Acquire detailed observations of a few representative objects. Understand the details of astrophysical processes that govern them, and extrapolate that understanding to the entire class.

Analogy: Google Search



Google search results for "two step". The search bar shows "two step". Below it, there are tabs for Web, Videos, Images, Maps, Shopping, More, and Search tools. The search results include:

- Two-step - Wikipedia, the free encyclopedia**
en.wikipedia.org/wiki/Two-step ▾ Wikipedia ▾
Two-step or Two Step may refer to: In dance. **Two-step (dance move)**, a dance move used in a wide range of dancing genres; Country-western **two-step**, also ...
Block en.wikipedia.org
- Two-step (dance move) - Wikipedia, the free encyclopedia**
[en.wikipedia.org/wiki/Two-step_\(dance_move\)](https://en.wikipedia.org/wiki/Two-step_(dance_move)) ▾ Wikipedia ▾
The **two-step** is a step found in various dances, including many folk dances. It seems to take its name from the 19th century dance related to the Polka.
Block en.wikipedia.org
- How to dance the Two-Step. Free 2-Step Dancing Lessons ...**
 www.youtube.com/watch?v=XRYOqjtXkc ▾
Sep 4, 2010 - Uploaded by Shawn Trautman
<http://www.shawntrautman.com/2-step-dance.aspx> Learn how to 2-Step (http://www.shawntrautman.com/how_...)
Block youtube.com
- Texas Two Step - Texas Lottery**
www.txdottery.org/export/sites/lottery/.../Texas_Two_Step/ ▾ Texas Lottery ▾
4 days ago - Do the Texas Two Step with the Texas Lottery. Drawings are on Mondays and Thursdays at 10:12 p.m. CT. Ticket sales are not available during ...
Block txdottery.org
- Two Step Restaurant and Cantina San Antonio**
www.twosteprestaurant.com/ ▾
Voted the Best of San Antonio Restaurant. Serving Texas inspired BBQ, Steaks, Chicken, Salads and Sandwiches for Lunch and Dinner Daily. Voted Best ...
Block twosteprestaurant.com



Wikipedia article on **Two-step (dance move)**. From Wikipedia, the free encyclopedia.

This article does not cite any references or sources. Please help improve this article by adding citations to reliable sources. Unsourced material may be challenged and removed. (December 2009)

The **two-step** is a step found in various dances, including many **folk dances**. It seems to take its name from the 19th century dance related to the **Polka**.

A two-step consists of two steps in approximately the same direction onto the same foot, separated by a closing step with the other foot. For example, a right two-step forward is a forward step onto the right foot, a closing step with the left foot, and a forward step onto the right foot. The closing step may be done directly beside the other foot, or obliquely beside, or even crossed, as long as the closing foot does not go past the other foot.

Some types of two-step, or related steps, are named "**lock step**".



YouTube video titled "How to dance the Two-Step. Free 2-Step Dancing Lessons w/Shawn Trautman". The video shows a man and a woman dancing the two-step on a wooden floor. The video player shows the progress bar at 0:10 / 7:46.

Google's index is a catalog of the Web. We use it to "zoom in" on individual entries to find out more.

How to dance the Two-Step. Free 2-Step Dancing Lessons w/Shawn Trautman



Subscribe

852,553

+ Add to Share More

719 139

21 ABP weather

Google weather

All Maps Apps News Books More Search tools

About 1,210,000,000 results (0.59 seconds)

Seattle, WA 98105

Tuesday 10:00 AM
Mostly Cloudy

 **4** °C | °F

Precipitation: 7%
Humidity: 88%
Wind: 6 km/h

Temperature Precipitation Wind



Time	Temperature (°C)
11 AM	4
2 PM	5
5 PM	6
8 PM	4
11 PM	4
2 AM	4
5 AM	3
8 AM	3

11 AM 2 PM 5 PM 8 PM 11 PM 2 AM 5 AM 8 AM

Day	Icon	Temp (°C)	Temp (°F)
Tue		5°	2°
Wed		7°	2°
Thu		5°	1°
Fri		4°	1°
Sat		6°	0°
Sun		5°	-1°
Mon		5°	2°
Tue		7°	3°

More on weather.com Feedback

[National and Local Weather Forecast, Hurricane, Radar and ...](#)
[www.weather.com/](#) The Weather Channel
The Weather Channel and **weather.com** provide a national and local **weather** forecast for cities, as well as **weather** radar, report and hurricane coverage.

But, it's more than just a catalog of pointers – more and more, Google itself collects, processes, indexes, visualizes, and serves the actual information we need.

More and more often, our “research” begins and ends with Google!

Modern Large Astronomical Surveys are the Googles of the Sky

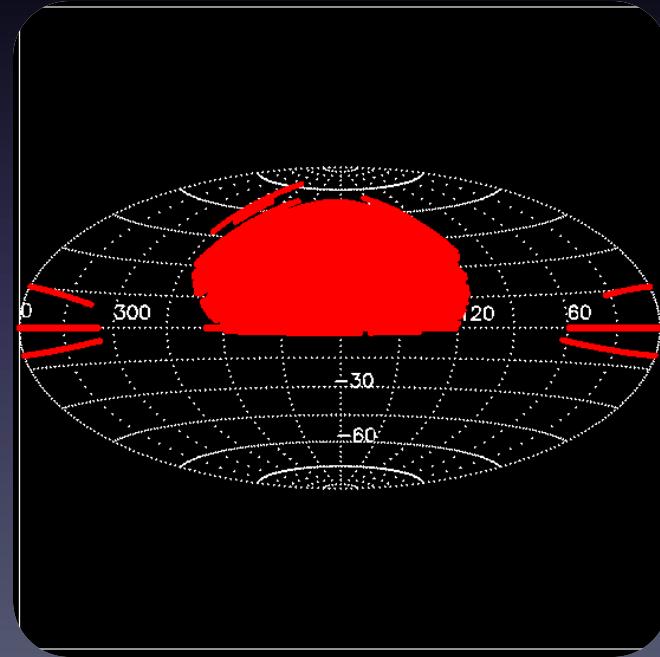
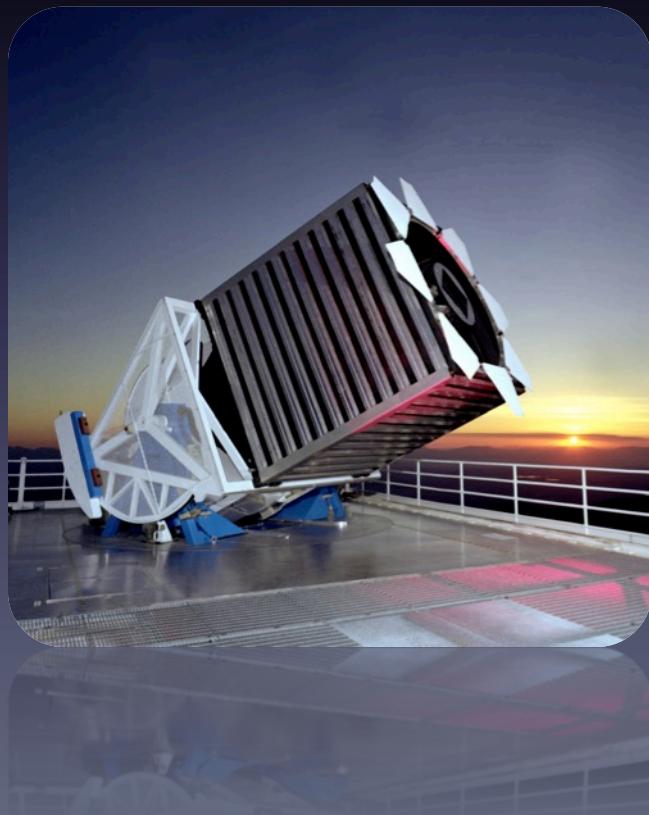
- There's a close parallel with large surveys in astronomy, in scale, quality, and richness of the collected information
 - Scale: We're entering the era when we can catalog the entire sky
 - Quality: Those catalogs will be as precise as the measurements taken with "pointed" observations (used to be ~5-10x worse)
 - Richness: Those catalogs contain not only positions and magnitudes, but also shapes, profiles, and temporal behavior of the objects.
- Quite often, the research begins and ends with the survey.
- This is what makes large surveys of today not just bigger, but different. They're more than just "finding charts".

Sloan Digital Sky Survey

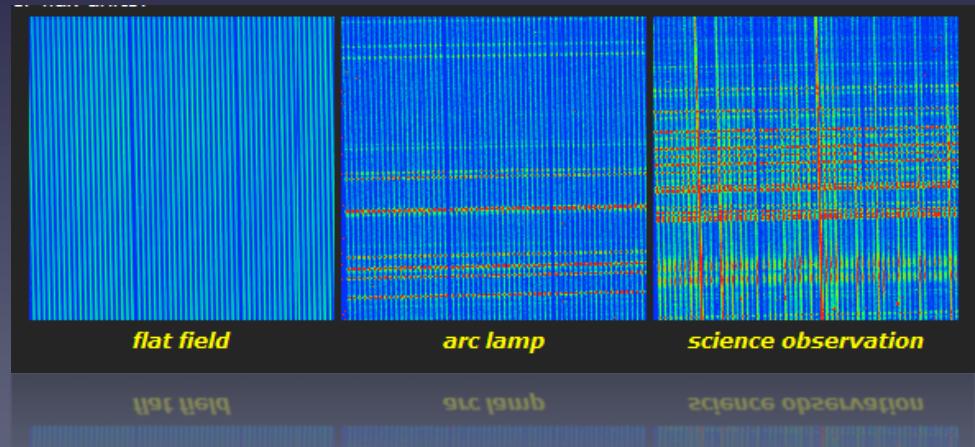
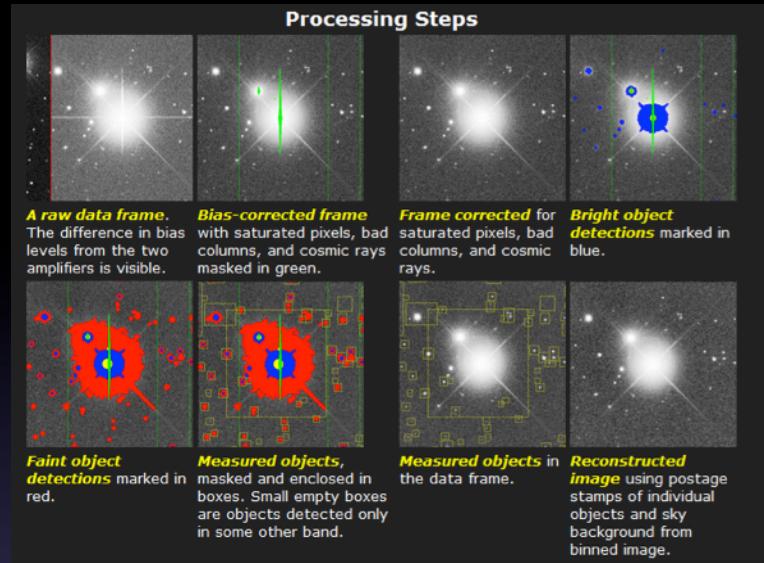
2.5m telescope 14,500 deg² 0.1" astrometry r<22.5 flux limit

5 band, 2%, photometry for >260M stars, 200M galaxies

Millions of spectra



Observing With SDSS



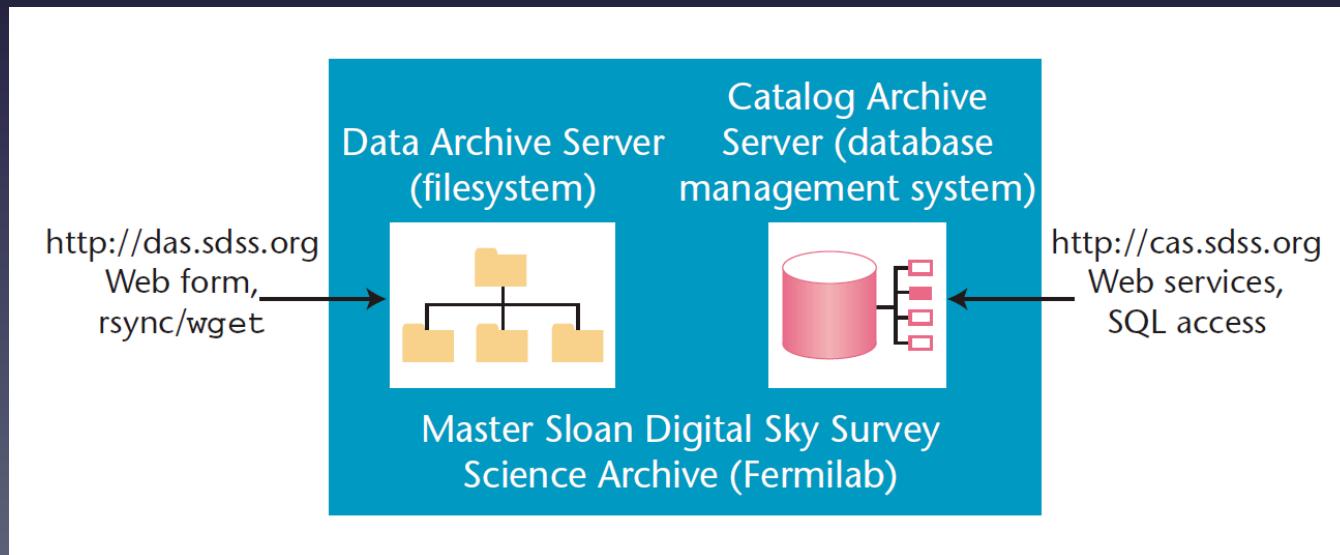
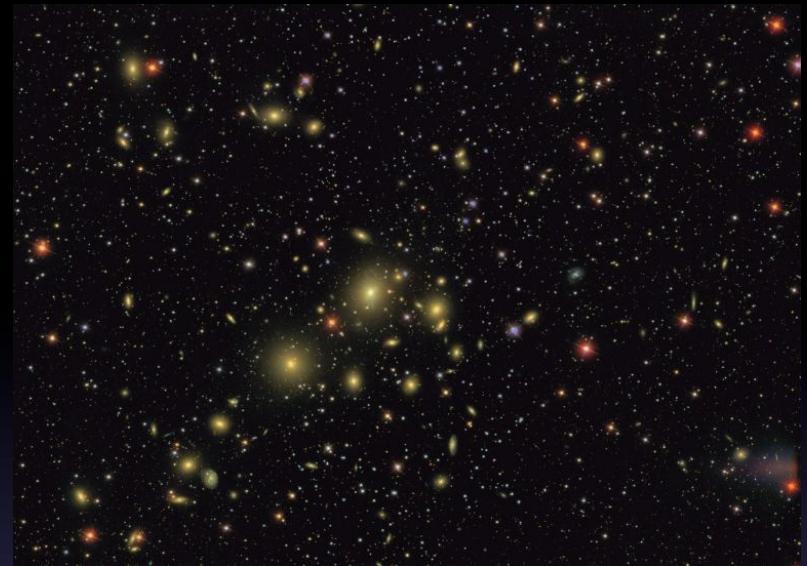
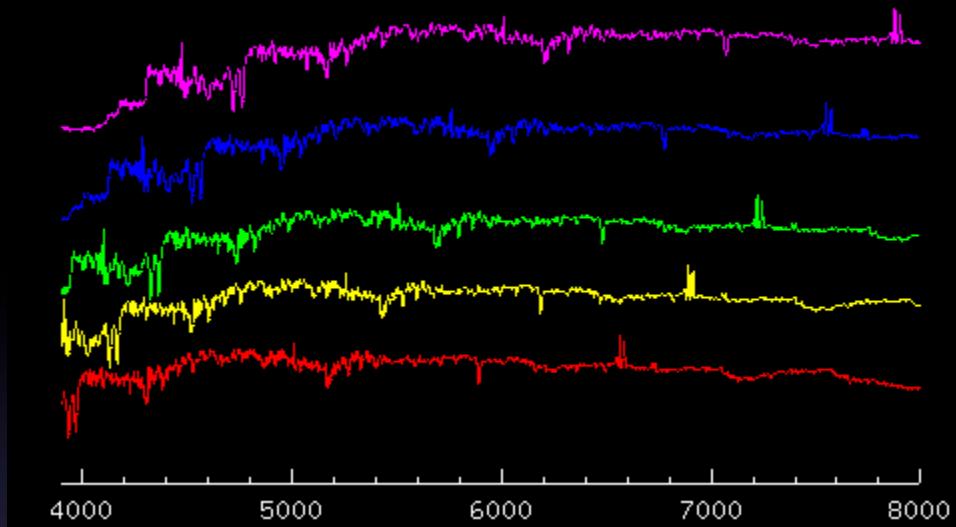


Table 1. Sloan Digital Sky Survey public data releases.

Release	Date	Size (catalogs)	Images (millions)	Spectra (thousands)	Distribution/mirrors
EDR ²	June 2001	200 Gbytes	14	54	Johns Hopkins Univ. (JHU), San Diego Supercomputing Center (SDSC), UK, Japan
DR1 ⁶	June 2003	1 Tbyte	53	186	JHU, SDSC, Canadian Astronomical Data Centre (CADC), Univ. of Pittsburgh, UK, Germany, Japan, India
DR2 ⁵	Mar. 2004	2 Tbytes	88	330	JHU, Univ. of Pittsburgh, SDSC, Germany
DR3 ³	Sept. 2004	3 Tbytes	141	478	JHU, UK, India
DR4 ⁴	June 2005	4 Tbytes	180	608	JHU, Germany, Hungary, Brazil
DR5 ⁶	June 2006	5 Tbytes	215	738	JHU, India, Russia, Hungary, Australia
DR6 ⁷	June 2007	6 Tbytes	287 M	1.27 M	JHU, India, Hungary
DR7 _a	June 2007	6 Tbytes	582 M	1.75 M	JHU, Canada, UK, Australia
DR7 _b	June 2008	6 Tbytes	512	1.38	JHU, Canada, UK, Australia
DR7 _c	June 2008	4 Tbytes	180	908	JHU, Germany, UK, Brazil

SDSS SkyServer DR7

http://cas.sdss.org/dr7/en/

Sloan Digital Sky Survey / SkyServer

SDSS

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

Welcome to the DR7 site!!!

This website presents data from the Sloan Digital Sky Survey, a project to make a map of a large part of the universe. We would like to show you the beauty of the universe, and share with you our excitement as we build the largest map in the history of the world.

News

The site hosts data from Data Release 7 (DR7). What's new in DR7, what's new on this site, and known problems. More...

For Astronomers

A separate branch of this website for professional astronomers (English) More...

SDSS is supported by

National Science Foundation (NSF)

NASA

Ministry of Education, Culture, Sports, Science and Technology (MEXT)

Powered by Microsoft

SkyServer Tools

- Famous places
- Get images
- Visual Tools
- Explore
- Search
- Object Cross-ID
- CasJobs

Science Projects

- Basic
- Advanced
- Challenges
- For Kids
- Games and Contests
- Teachers
- Links to other projects

Info Links

- About Astronomy
- About the SDSS
- About the SkyServer
- SDSS Data Release 7
- SDSS Project Website
- Open SkyQuery
- Images of RC3 Galaxies

Help

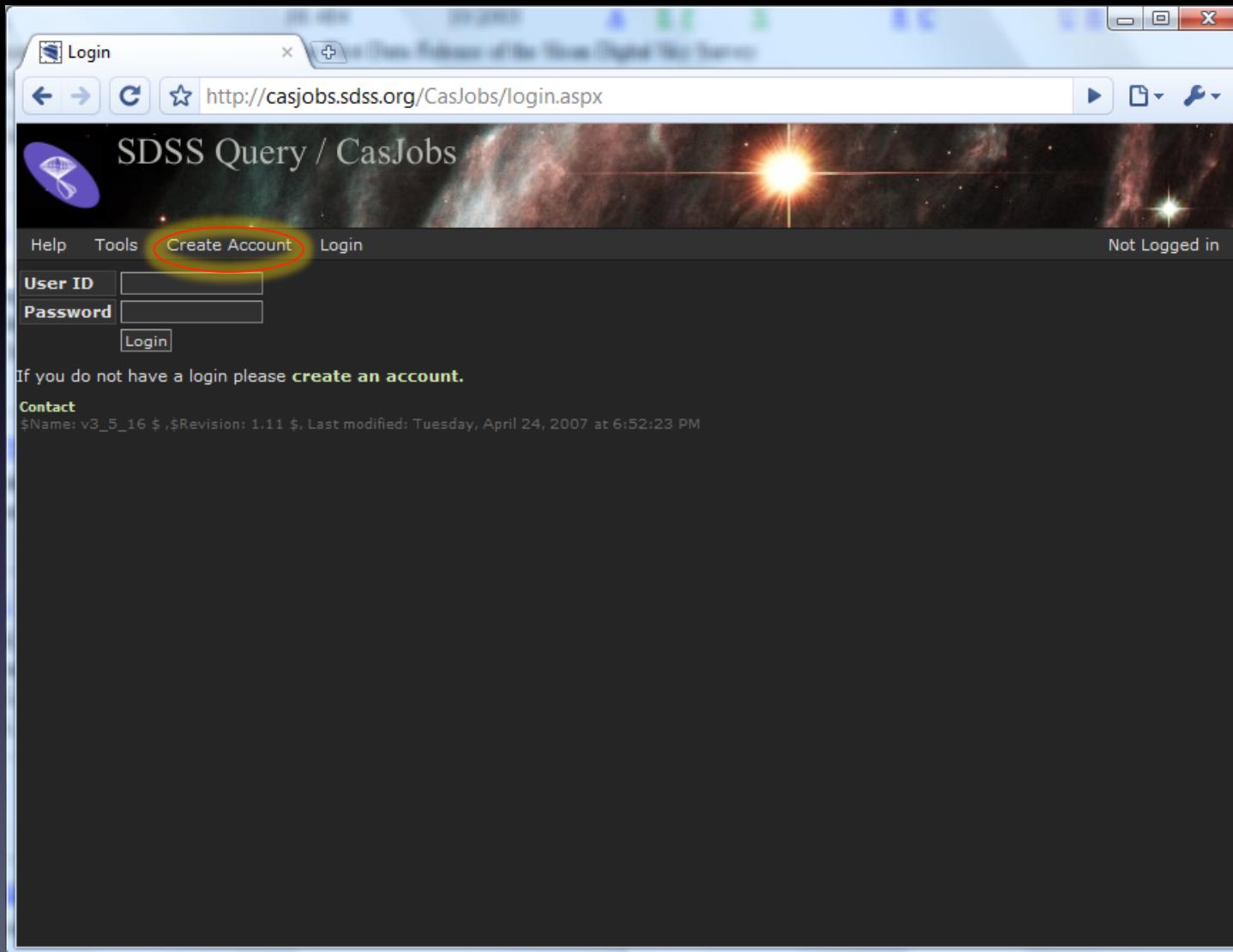
- Getting Started
- FAQ
- How To
- Glossary
- Schema Browser
- Sample SQL Queries
- Details of SDSS Data

Contact Us

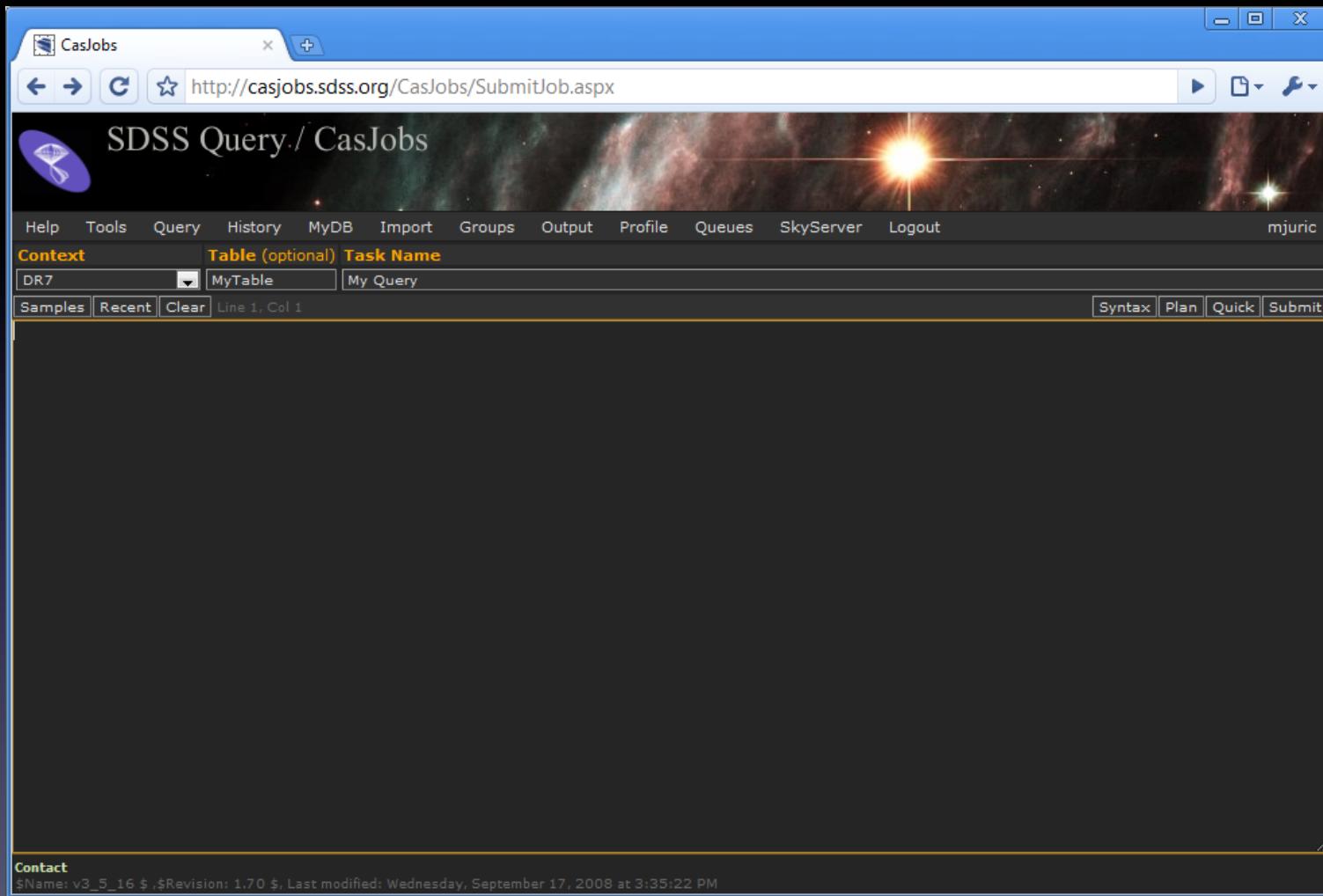
The contents for boundaries of the different reports

Report No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000
------------	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------

CasJobs: Running long and complex queries



SDSS CasJobs Service



CasJobs

http://casjobs.sdss.org/CasJobs/SubmitJob.aspx

SDSS Query./ CasJobs

Help Tools Query History MyDB Import Groups Output Profile Queues SkyServer Logout mjuric

Context Table (optional) Task Name

DR7 MyTable My Query

Samples Recent Clear Line 1, Col 1 Syntax Plan Quick Submit

Example:

Get the positions (in Galactic coordinates) and dereddened magnitudes of all stars within 1 degree of $\alpha = 341.6$, $\delta = 31.7$

Contact
\$Name: v3_5_16 \$,\$Revision: 1.70 \$, Last modified: Wednesday, September 17, 2008 at 3:35:22 PM

The screenshot shows a web browser window titled 'CasJobs' with the URL 'http://casjobs.sdss.org/CasJobs/SubmitJob.aspx'. The page header includes the SDSS logo and the title 'SDSS Query./ CasJobs'. A navigation bar at the top has links for Help, Tools, Query, History, MyDB, Import, Groups, Output, Profile, Queues, SkyServer, and Logout. A user name 'mjuric' is visible on the right. Below the header is a search form with fields for 'Context' (set to DR7), 'Table (optional)' (set to MyTable), and 'Task Name' (set to My Query). There are buttons for Samples, Recent, Clear, Line 1, Col 1, Syntax, Plan, Quick, and Submit. The main area contains the text 'Example:' and 'Get the positions (in Galactic coordinates) and dereddened magnitudes of all stars within 1 degree of $\alpha = 341.6$, $\delta = 31.7$ '. At the bottom, there is a 'Contact' section with revision information: '\$Name: v3_5_16 \$,\$Revision: 1.70 \$, Last modified: Wednesday, September 17, 2008 at 3:35:22 PM'.

An SQL Query

The screenshot shows a web-based SQL query interface titled "SDSS Query / CasJobs". The interface has a blue header bar with the title "CasJobs" and a back/forward, refresh, and search toolbar. Below the header is a URL bar showing "http://casjobs.sdss.org/CasJobs/SubmitJob.aspx". The main area is titled "SDSS Query / CasJobs" and features a logo of a telescope. A sidebar on the right contains a brief introduction to SQL.

SQL – Structured Query Language
A language for specifying and extracting subsets of data from large tabular databases.
Allows you to select only the data that satisfy a certain set of (possibly very complicated) restrictions that you specify

Context **Table (optional)** **Task Name**
DR7 MyTable My Query
Samples Recent Clear Line 18, Col 1

```
select

s.objid,
s.l, s.b,
s.extinction_r as Ar,
s.dered_u as u, s.dered_g as g, s.dered_r as r, s.dered_i as i, s.dered_z as z,
s.err_u, s.err_g, s.err_r, s.err_i, s.err_z,
s.flags

from
fGetNearbyObjEq(341.6, 31.7, 60) n, Star s

into
mydb.field1

where
n.objID=s.objID
```

Anatomy of an SQL Query

The screenshot shows the SDSS Query / CasJobs interface. In the top navigation bar, there is a logo, a search bar with the URL <http://casjobs.sdss.org/CasJobs/SubmitJob.aspx>, and a menu bar with items like Help, Tools, Query, History, MyDB, Import, Groups, and Output.

The main area is titled "SDSS Query / CasJobs". It has a toolbar with buttons for Samples, Recent, and Clear. Below the toolbar is a code editor containing an SQL query:

```
select s.objid,
       s.l, s.b,
       s.extinction_r as Ar,
       s.dered_u as u, s.dered_g as g, s.dered_r as r, s.dered_i as i, s.dered_z as z,
       s.err_u, s.err_g, s.err_r, s.err_i, s.err_z,
       s.flags
  from fGetNearbyObjEq(341.6, 31.7, 60) n, Star s
  into mydb.field1
 where n.objID=s.objID
```

Annotations explain the different parts of the query:

- SELECT** – select a subset of entries from the catalog, based on the criteria we will specify. (Annotation points to the word "select")
- Columns** – the columns we are interested in. Usually we don't want all (> 50!) available columns, but just some (e.g., id, position, extinction, magnitudes, errors, and flags). (Annotation points to the list of columns: s.objid, s.l, s.b, s.extinction_r as Ar, s.dered_u as u, s.dered_g as g, s.dered_r as r, s.dered_i as i, s.dered_z as z, s.err_u, s.err_g, s.err_r, s.err_i, s.err_z, s.flags)
- FROM clause** – the tables from which to fetch the columns above. Can be "virtual". (Annotation points to the "from" keyword and the function call fGetNearbyObjEq)
- INTO clause** – where to store the result. The results go to tables in your personal database. (Annotation points to the "into" keyword and the table name mydb.field1)
- WHERE clause** – the filter with which to filter the result. (Annotation points to the "where" keyword and the condition n.objID=s.objID)

Submitted Job

CasJobs 

http://casjobs.sdss.org/CasJobs/jobdetails.aspx?id=4928887

SDSS Query / CasJobs

Help Tools Query History MyDB Import Groups Output Profile Queues SkyServer Logout mjuric

'My Query' Details

This page will automatically refresh every 30 seconds

There are 0 jobs ahead of this one in the DR7 queue.

JobID	TaskName	Context	Queue	Submitted	Started	Finished	Status
4928887	My Query	DR7	600	7/28/2010 4:33:04 AM			Ready

Executed on
0 No Message

Query

```
select
s.objid,
s.l, s.b,
s.extinction_r as Ar,
s.dered_u as u, s.dered_g as g, s.dered_r as r, s.dered_i as i, s.dered_z as z,
s.err_u, s.err_g, s.err_r, s.err_i, s.err_z,
s.flags

from
fGetNearbyObjEq(341.6, 31.7, 60) n, Star s

into
mydb.field2

where
n.objID=s.objID
```

Contact
\$Name: v3_5_16 \$,\$Revision: 1.23 \$, Last modified: Wednesday, May 14, 2008 at 1:52:25 AM

Completed Job

The screenshot shows a web browser window titled "CasJobs" displaying the "SDSS Query / CasJobs" interface. The URL in the address bar is <http://casjobs.sdss.org/CasJobs/jobdetails.aspx?id=4928831&message=Details%20of%204928831>. The top navigation bar includes links for Help, Tools, Query, History, MyDB (which is highlighted with a yellow oval), Import, Groups, Output, Profile, Queues, SkyServer, and Logout. The user name "mjuric" is visible on the right.

The main content area is titled "'My Query' Details". It shows a table of job details:

JobID	TaskName	Context	Queue	Submitted	Started	Finished	Status
4928831	My Query	DR7	600	7/28/2010 4:03:05 AM	7/28/2010 4:03:13 AM	7/28/2010 4:03:43 AM	Finished

Below the table, it says "Executed on Rows DR7Best long 70454" and "Message: Query Complete".

The "Query" section contains the following SQL code:

```
select  
s.objid,  
s.l, s.b,  
s.extinction_r as Ar,  
s.dered_u as u, s.dered_g as g, s.dered_r as r, s.dered_i as i, s.dered_z as z,  
s.err_u, s.err_g, s.err_r, s.err_i, s.err_z,  
s.flags  
  
from  
fGetNearbyObjEq(341.6, 31.7, 60) n, Star s  
  
into  
mydb.field1  
  
where  
n.objID=s.objID
```

At the bottom left, there is a "Contact" link and a note: "\$Name: v3_5_16 \$,\$Revision: 1.23 \$, Last modified: Wednesday, May 14, 2008 at 1:52:25 AM\$". A white arrow points from the text "Query Complete" in the "Message" box towards the bottom right corner of the query results table.

MyDB miner - Google Search

http://casjobs.sdss.org/CasJobs/MyDB.aspx

SDSS Query / CasJobs

Help Tools Query History MyDB Import Groups Output Profile Queues SkyServer Logout mjuric

MyDB Local Only

Views

Tables

Functions

Procedures

Sort by... All selected...

Rows	kB	Name
70,454	9,160	field1

Mario Juric 's MyDB

9,624 kB of 2,000,000 kB used

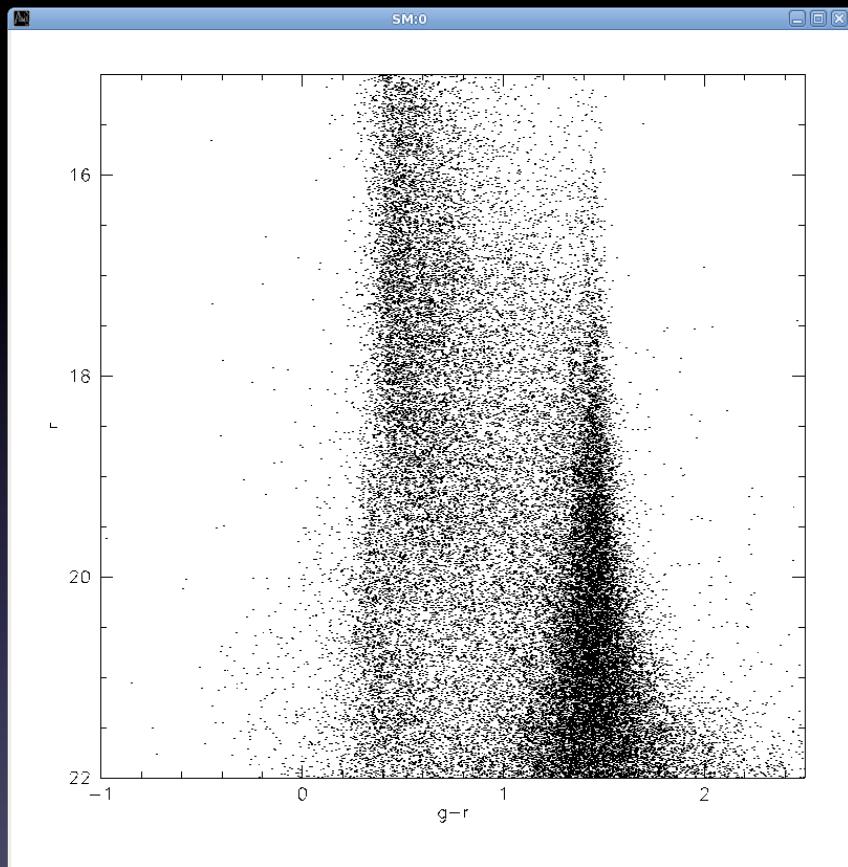
From this page you can get various information about the contents of both your MyDB and shared tables within your groups. Click the left table links to get information about a specific table, such as rows, columns or size. From the table pages you can also perform various table-specific tasks, such as:

- Download a table
- Manage your group tables
- Rename a table
- Drop a table

Sizes are approximations only.
Row counts are approximations only. For exact value run a count.
There's always some overhead, even empty MyDB's take up space.
Group tables do not count towards your MyDB size limit.

Contact
Name: v3_5_16 \$, Revision: 1.64 \$, Last modified: Tuesday, January 27, 2009 at 3:19:32 PM

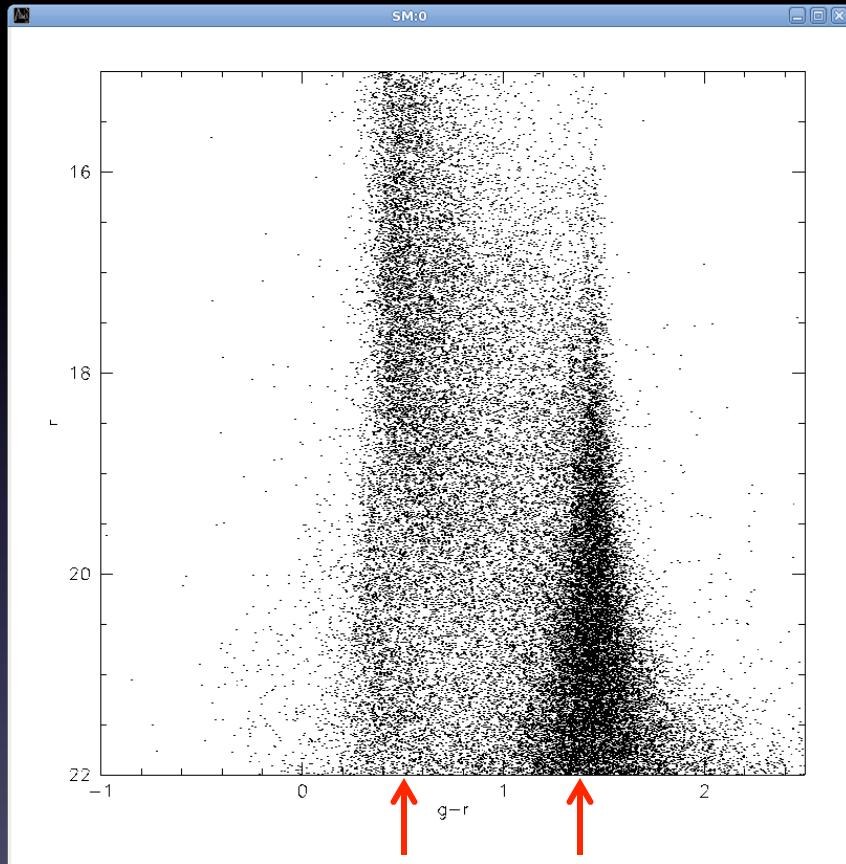
Basic Data Analysis: A Color-Magnitude Diagram (CMD)



```
Untitled - Notepad
File Edit Format View Help
data "field1.csv"
read <1 2 b 3 Ar 4 u 5 g 6 r 7 i 8 z 9>
set gr = g-r
lweight 2
expand 1.001
limits -1 2.5 22 15
ptype 0 0
erase
box
points gr r
xlabel g-r
ylabel r
```

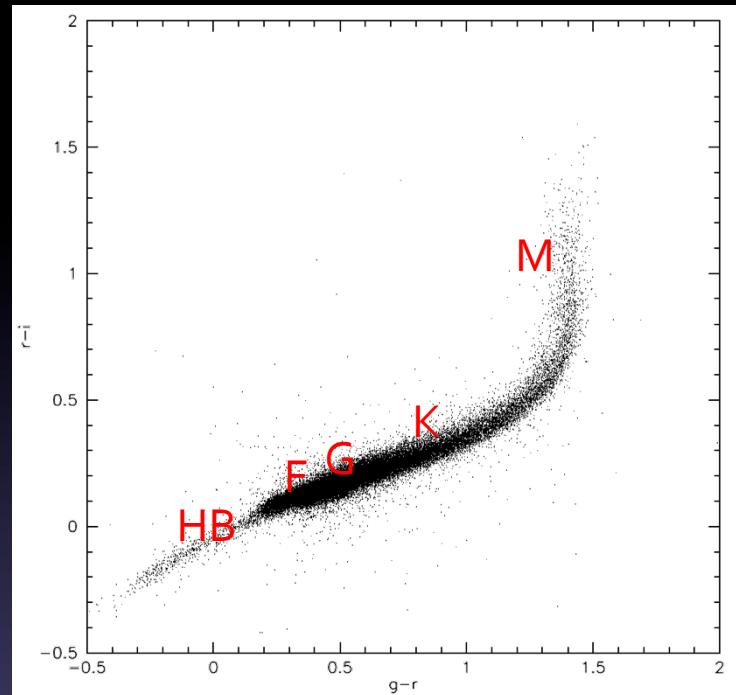
The screenshot shows a Windows Notepad window titled "Untitled - Notepad". The window contains Starmax command language (SM:0) code. The code reads data from a CSV file named "field1.csv" and performs several operations: it sets the color axis to "g-r", uses a line weight of 2, expands the data by a factor of 1.001, sets the limits for color (g-r) from -1 to 2.5 and magnitude (r) from 22 to 15, sets the point type to 0, erases the previous plot, draws a box around the data, and finally plots the data points with the color axis labeled "g-r" and the magnitude axis labeled "r".

Interpreting Stellar CMDs

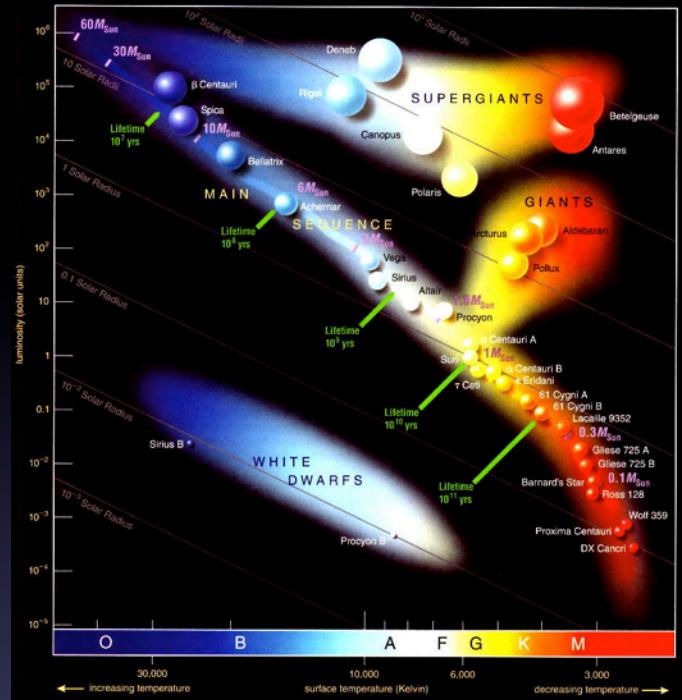
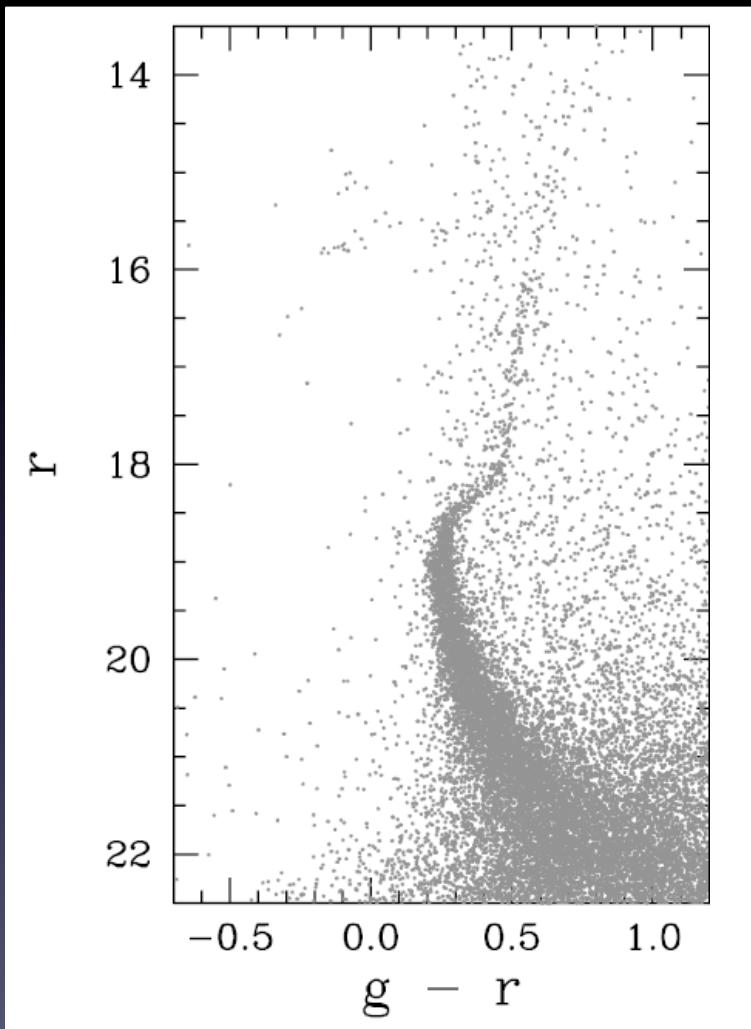


← More
Luminous

Less →
Luminous



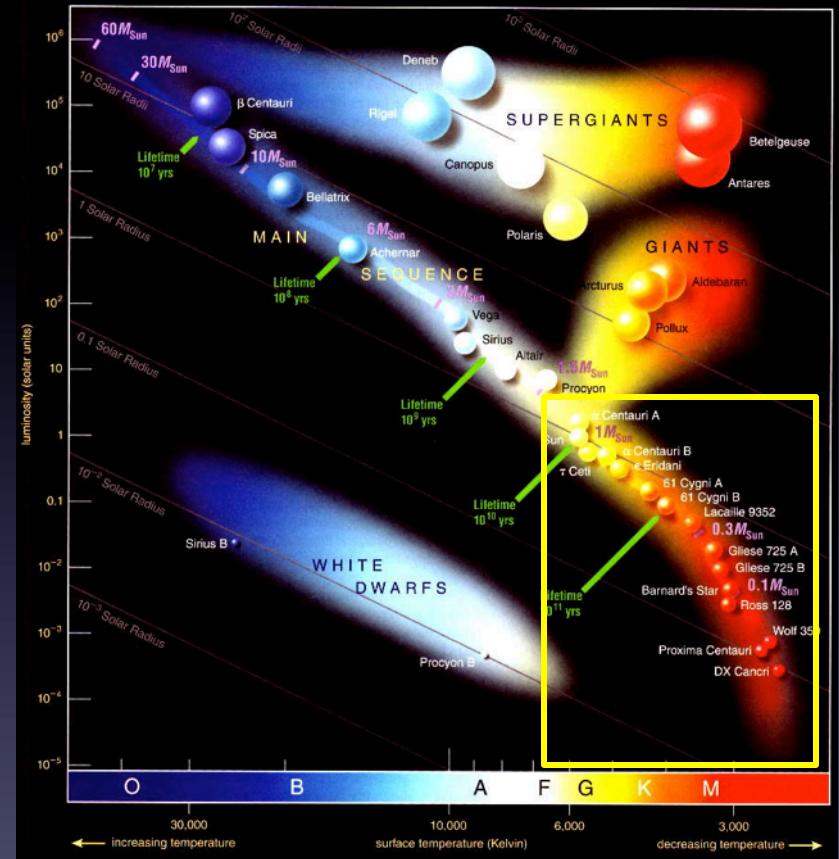
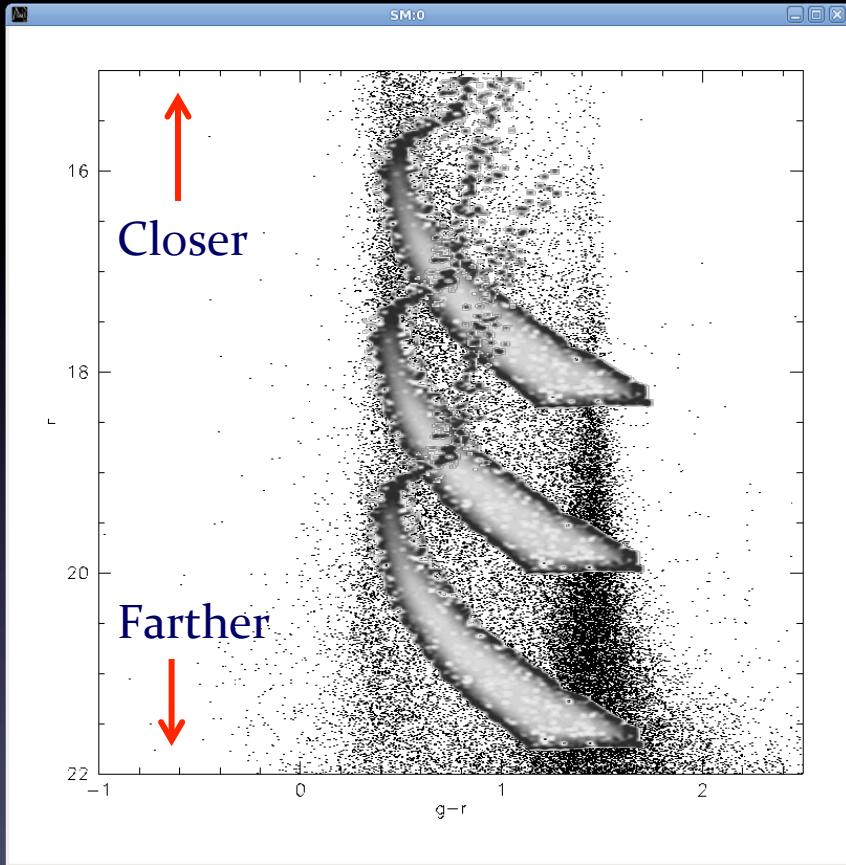
A CMD of a Globular Cluster



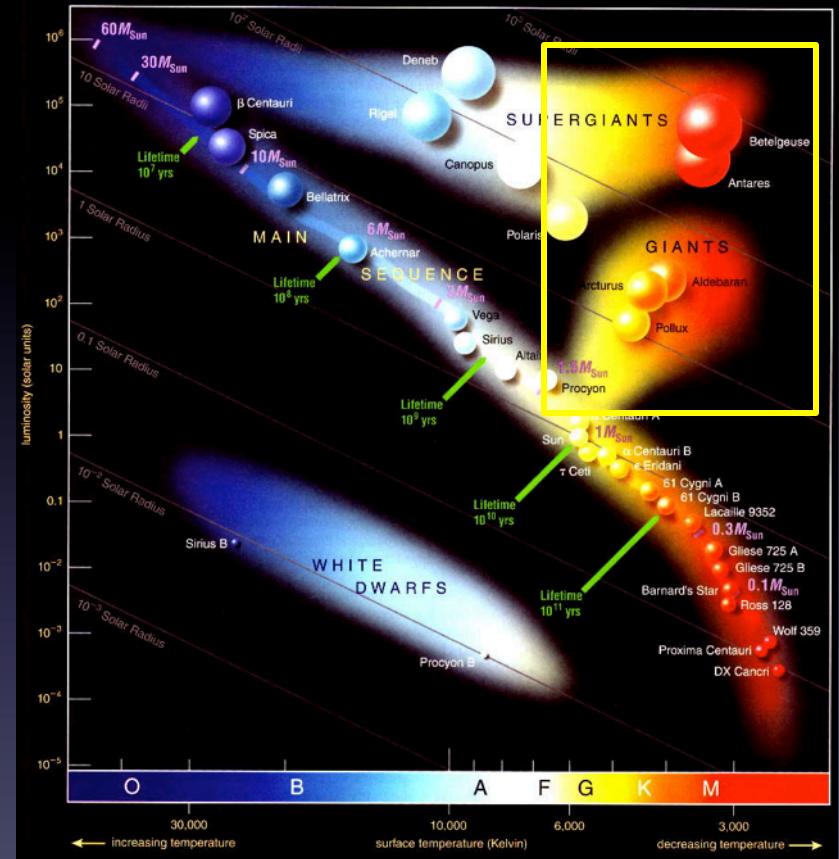
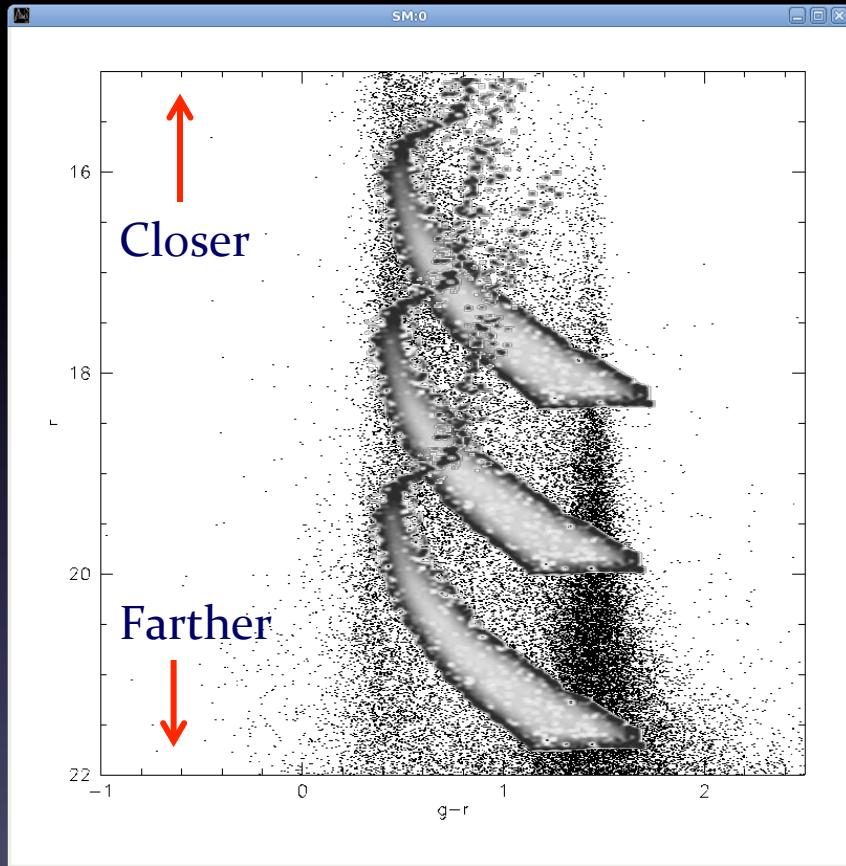
A globular cluster is a simple population
(same age, same metallicity), at the
same distance.

$$m = M + 5 \log_{10}(d/10\text{kpc})$$

Interpreting “Field” CMDs



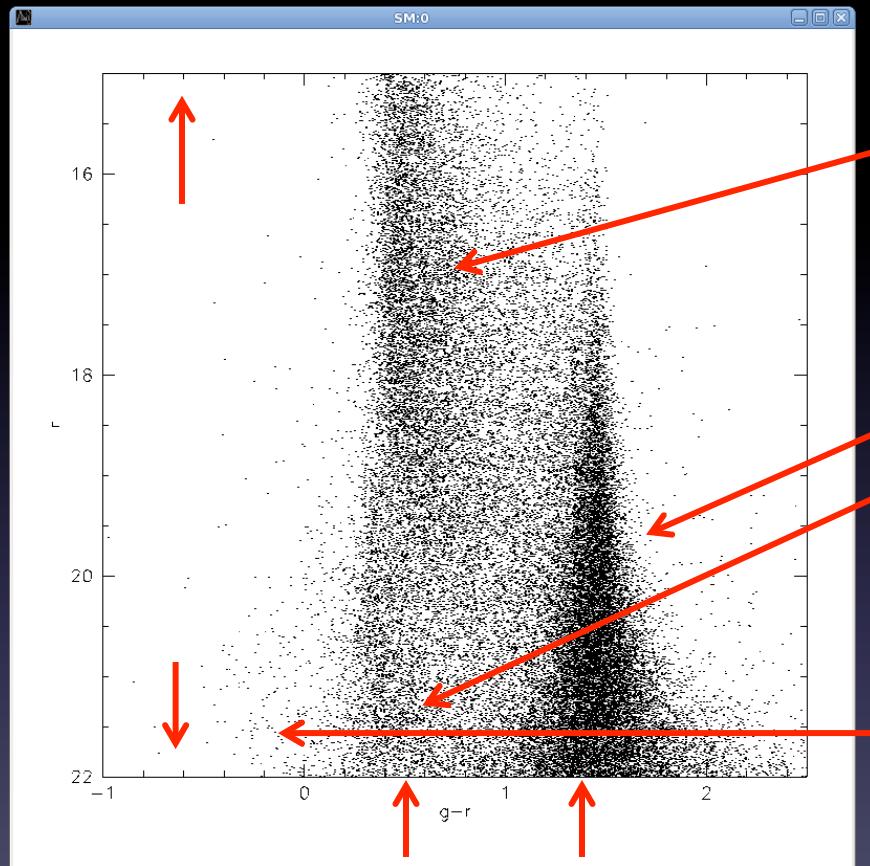
Interpreting “Field” CMDs



What about the giants? Yes, they are there, but significantly less numerous (<5%) than the main sequence stars.

The same applies for early type stars, white dwarfs, etc.

Basic Data Analysis: A Color-Magnitude Diagram



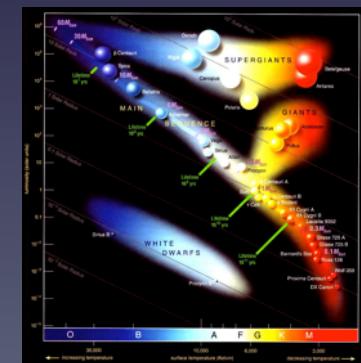
← More Luminous

Less →
Luminous

Disk G/F stars

Nearby Disk K/M stars
Halo stars

QSOs, RRLyrae,
WDs, errors...



CasJobs

http://casjobs.sdss.org/CasJobs/SubmitJob.aspx

SDSS Query./ CasJobs

Help Tools Query History MyDB Import Groups Output Profile Queues SkyServer Logout mjuric

Context Table (optional) Task Name

DR7 MyTable My Query

Samples Recent Clear Line 1, Col 1 Syntax Plan Quick Submit

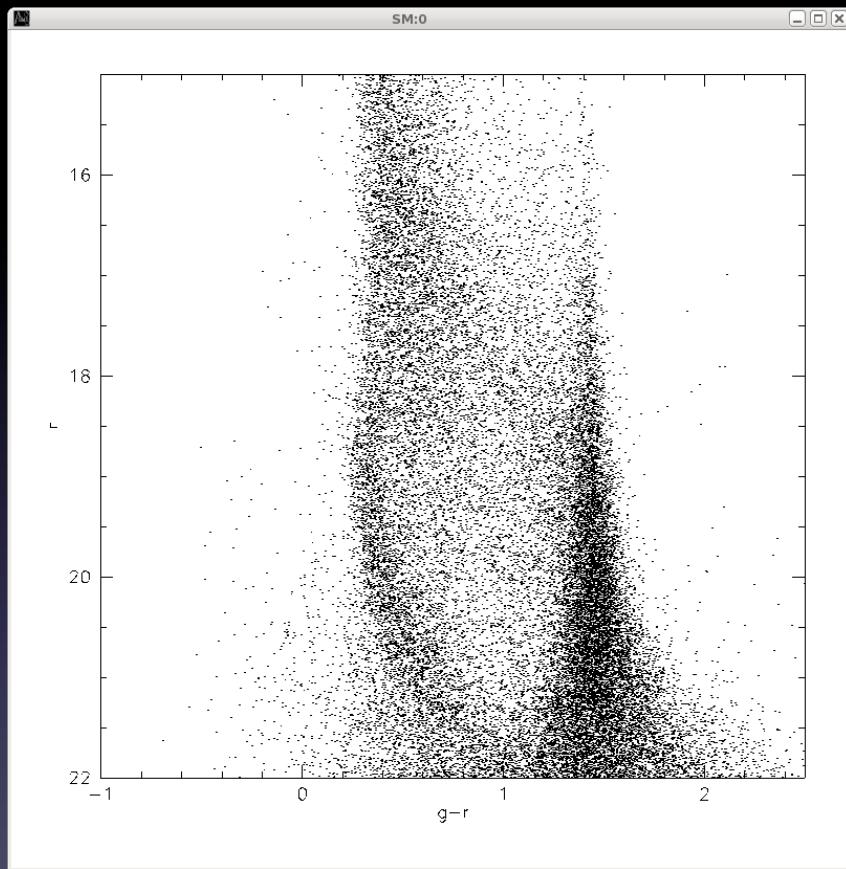
Get the positions (in Galactic coordinates) and dereddened magnitudes of all stars within 1 degree of $\alpha = 21^{\circ}16', \delta = 31^{\circ}7'$

$\alpha = 115.2, \delta = 32.7$

Contact
\$Name: v3_5_16 \$,\$Revision: 1.70 \$, Last modified: Wednesday, September 17, 2008 at 3:35:22 PM

The screenshot shows a web-based query interface for the SDSS. At the top, there's a header bar with the title 'CasJobs' and a URL 'http://casjobs.sdss.org/CasJobs/SubmitJob.aspx'. Below the header is a banner featuring a starry background with a prominent central star. The main menu includes 'Help', 'Tools', 'Query', 'History', 'MyDB', 'Import', 'Groups', 'Output', 'Profile', 'Queues', 'SkyServer', and 'Logout'. A user name 'mjuric' is visible on the right. The 'Query' section has a 'Context' dropdown set to 'DR7', a 'Table (optional)' dropdown set to 'MyTable', and a 'Task Name' input field set to 'My Query'. Below these are buttons for 'Samples', 'Recent', 'Clear', and 'Line 1, Col 1'. To the right of these buttons are links for 'Syntax', 'Plan', 'Quick', and 'Submit'. The main body of the page contains a large text area with the query instructions. At the bottom, there's a 'Contact' section with revision information.

Color-Magnitude Diagram

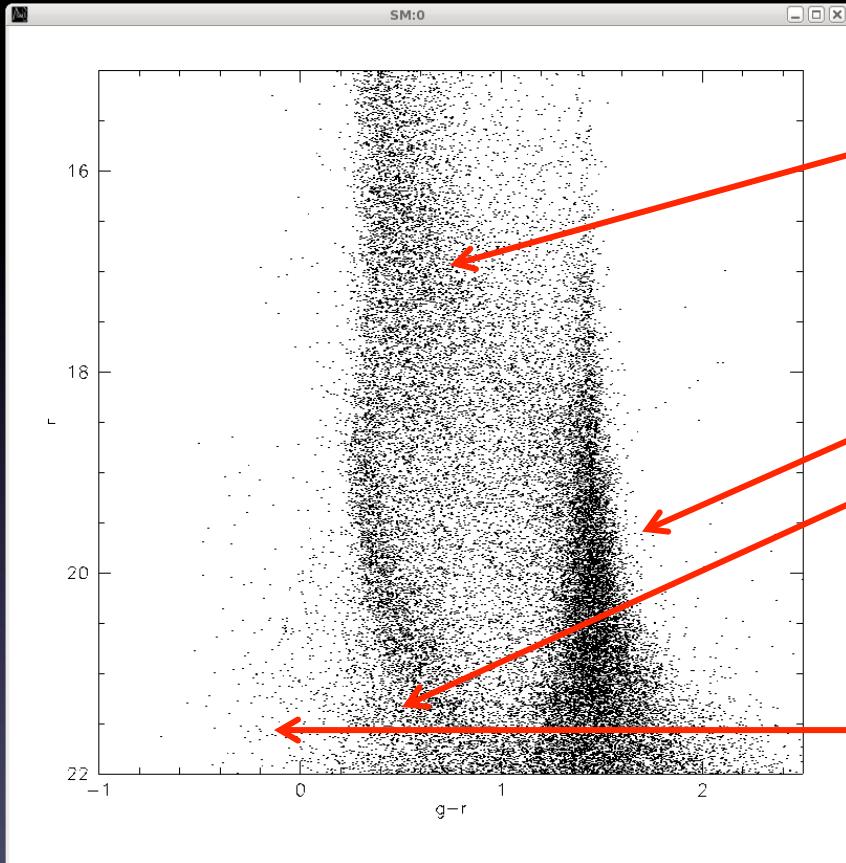


Untitled - Notepad

```
File Edit Format View Help
data "field2.csv"
read <1 2 b 3 Ar 4 u 5 g 6 r 7 i 8 z 9>

set gr = g-r
lweight 2
expand 1.001
llimits -1 2.5 22 15
ptype 0 0
erase
box
points gr r
xlabel g-r
ylabel r
```

Color-Magnitude Diagram



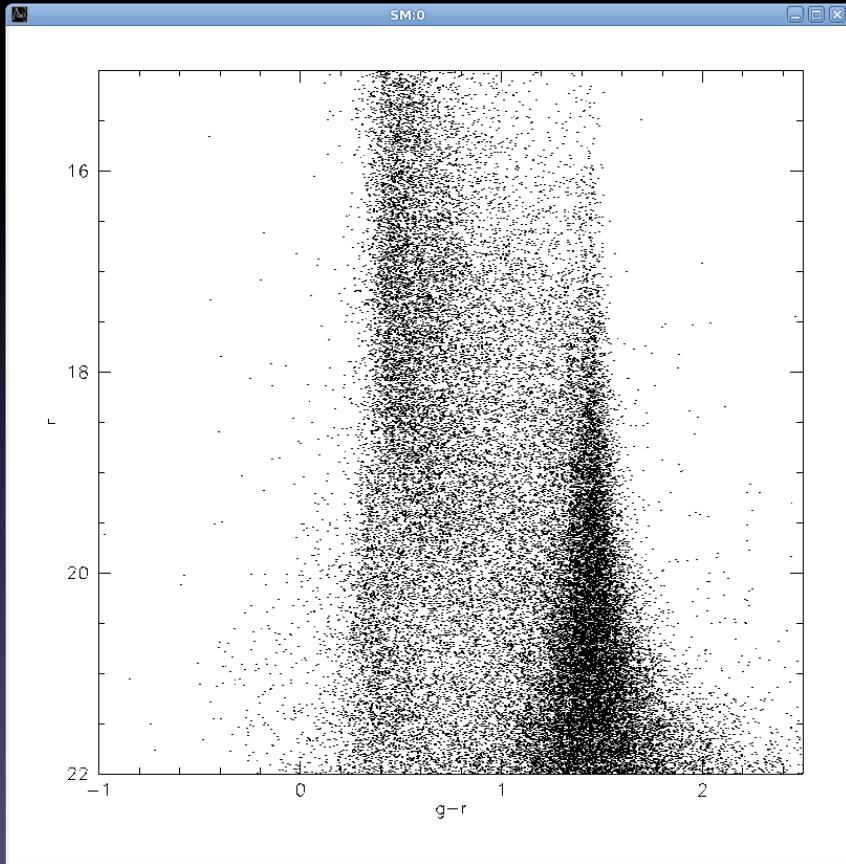
Disk G/F stars

Nearby Disk K/M stars
Halo stars

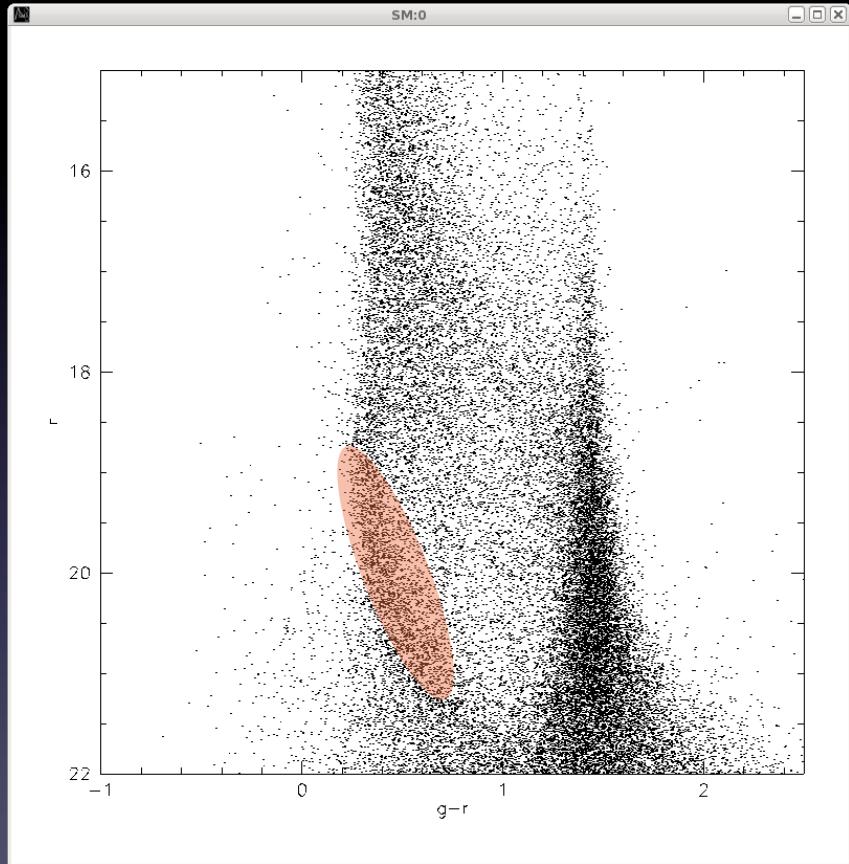
QSOs, RR Lyrae,
WDs, errors...

Notice Anything Different...?

Field #1



Field #2



Monoceros Stream (Newberg et al. 2002)

THE GHOST OF SAGITTARIUS AND LUMPS IN THE HALO OF THE MILKY WAY

HEIDI JO NEWBERG,^{1,2} BRIAN YANNY,^{1,3} CONNIE ROCKosi,⁴ EVA K. GREBEL,⁵ HANS-WALTER RIX,⁵ JON BRINKMANN,⁶ ISTVAN CSABAI,⁷ GREG HENNESSY,⁸ ROBERT B. HINDSLEY,⁸ RODRIGO IBATA,⁹ ZELJKO IVEZIĆ,¹⁰ DON LAMB,⁴ E. THOMAS NASH,³ MICHAEL ODENKIRCHEN,⁵ HEATHER A. RAVE,² D. P. SCHNEIDER,¹¹ J. ALLYN SMITH,¹² ANDREA STOLTE,⁵ AND DONALD G. YORK⁴

Received 2001 June 18; accepted 2001 December 5

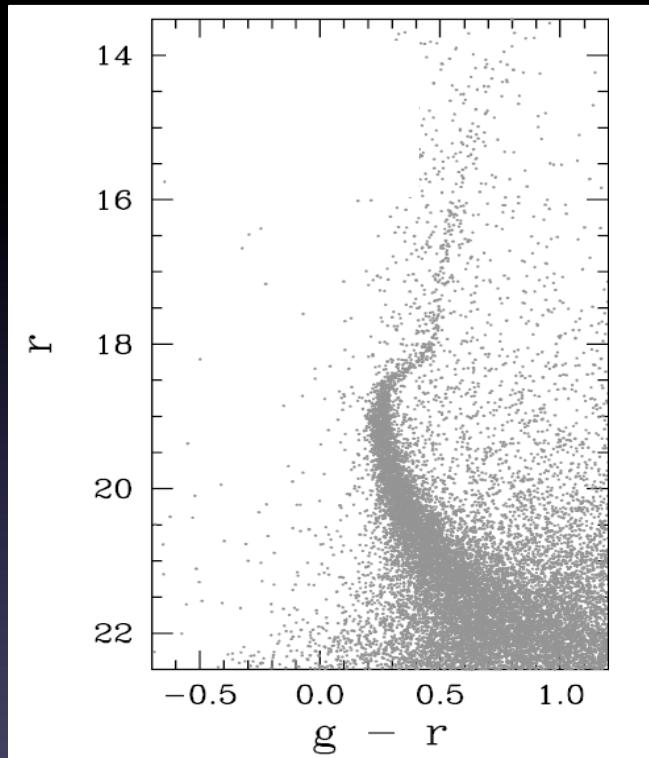
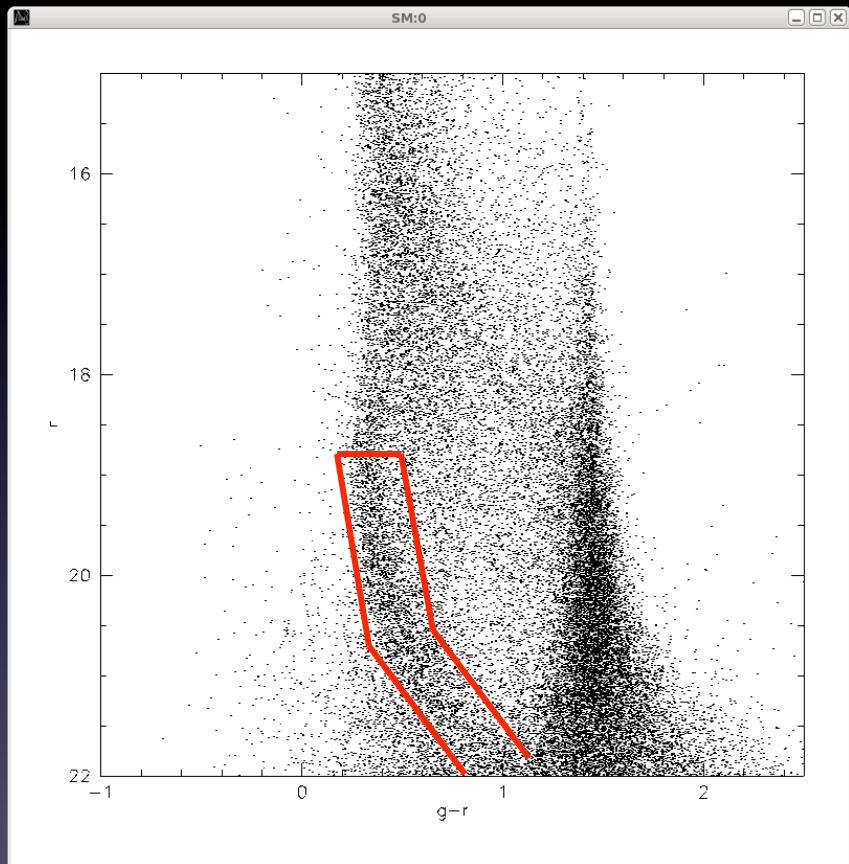
ABSTRACT

We identify new structures in the halo of the Milky Way from positions, colors, and magnitudes of five million stars detected in the Sloan Digital Sky Survey. Most of these stars are within $1^{\circ}26$ of the celestial equator. We present color-magnitude diagrams (CMDs) for stars in two previously discovered, tidally disrupted structures. The CMDs and turnoff colors are consistent with those of the Sagittarius dwarf galaxy, as had been predicted. In one direction, we are even able to detect a clump of red stars, similar to that of the Sagittarius dwarf, from stars spread across 110 deg^2 of sky. Focusing on stars with the colors of F turnoff objects, we identify at least five additional overdensities of stars. Four of these may be pieces of the same halo structure, which would cover a region of the sky at least 40° in diameter, at a distance of 11 kpc from the Sun (18 kpc from the center of the Galaxy). The turnoff is significantly bluer than that of thick-disk stars, yet the stars lie closer to the Galactic plane than a power-law spheroid predicts. We suggest two models to explain this new structure. One possibility is that this new structure could be a new dwarf satellite of the Milky Way, hidden in the Galactic plane and in the process of being tidally disrupted. The other possibility is that it could be part of a disklike distribution of stars which is metal-poor, with a scale height of approximately 2 kpc and a scale length of approximately 10 kpc. The fifth overdensity, which is 20 kpc away, is some distance from the Sagittarius dwarf streamer orbit and is not associated with any known Galactic structure. We have tentatively identified a sixth overdensity in the halo. If this sixth structure is instead part of a smooth distribution of halo stars (the spheroid), then the spheroid must be very flattened, with axial ratio $q = 0.5$. It is likely that there are many smaller streams of stars in the Galactic halo.

Subject headings: Galaxy: halo — Galaxy: structure

Learning about structures from CMDs

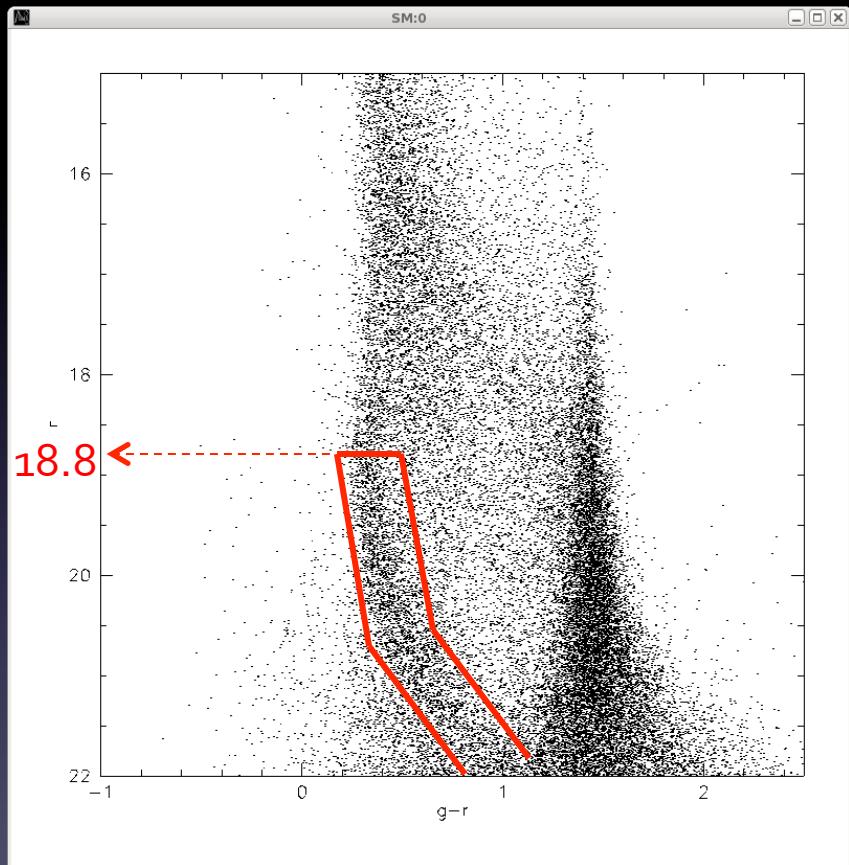
Monoceros stream CMD



⇒ Monoceros stream is a
localized structure

Learning about structures from CMDs

Monoceros stream CMD

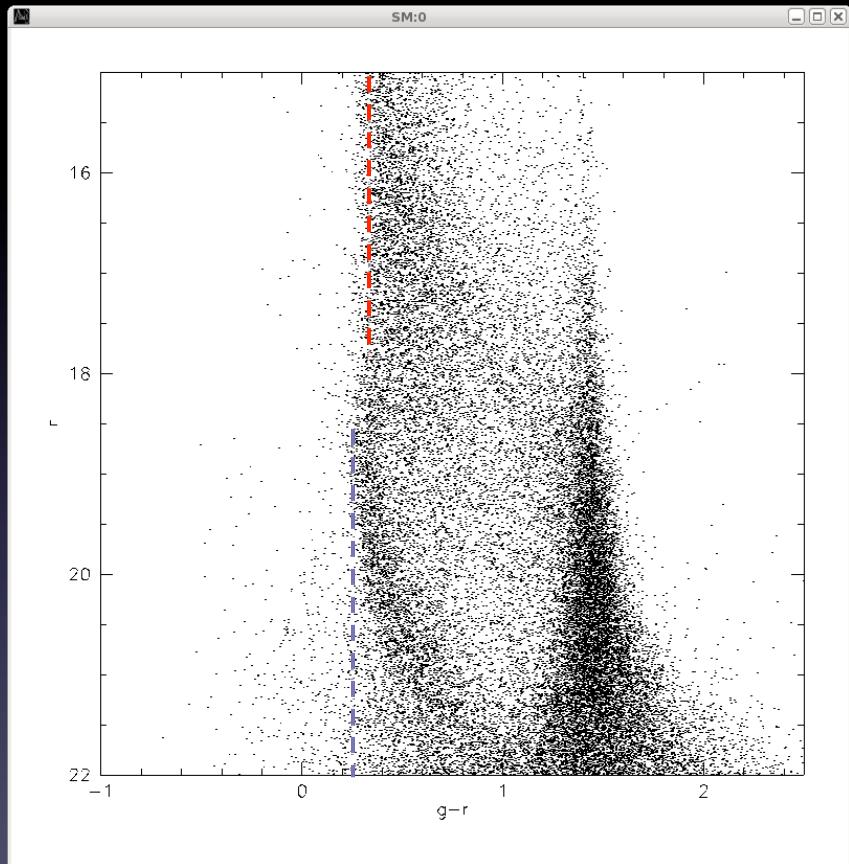


$$M_{MSTO} \approx 3.5$$
$$d = 10^{\frac{m-M}{5}+1}$$

⇒ *The distance to the Monoceros stream has to be around ~11.5kpc*

Learning about structures from CMDs

Monoceros stream CMD



Turn-off color of the Monoceros stream is bluer than that of the disk
⇒ Mon. stars are either more metal poor, or younger than disk stars

Monoceros Stream (Newberg et al. 2002)

THE GHOST OF SAGITTARIUS AND LUMPS IN THE HALO OF THE MILKY WAY

HEIDI JO NEWBERG,^{1,2} BRIAN YANNY,^{1,3} CONNIE ROCKosi,⁴ EVA K. GREBEL,⁵ HANS-WALTER RIX,⁵ JON BRINKMANN,⁶ ISTVAN CSABAI,⁷ GREG HENNESSY,⁸ ROBERT B. HINDSLEY,⁸ RODRIGO IBATA,⁹ ZELJKO IVEZIĆ,¹⁰ DON LAMB,⁴ E. THOMAS NASH,³ MICHAEL ODENKIRCHEN,⁵ HEATHER A. RAVE,² D. P. SCHNEIDER,¹¹ J. ALLYN SMITH,¹² ANDREA STOLTE,⁵ AND DONALD G. YORK⁴

Received 2001 June 18; accepted 2001 December 5

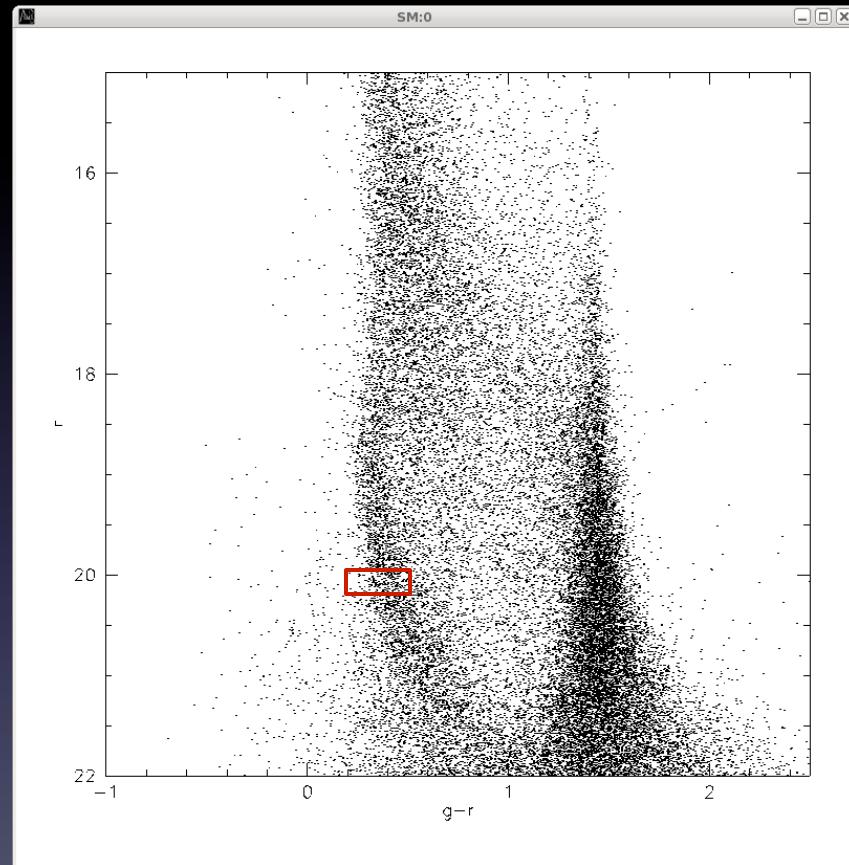
ABSTRACT

We identify new structures in the halo of the Milky Way from positions, colors, and magnitudes of five million stars detected in the Sloan Digital Sky Survey. Most of these stars are within $1^{\circ}26$ of the celestial equator. We present color-magnitude diagrams (CMDs) for stars in two previously discovered, tidally disrupted structures. The CMDs and turnoff colors are consistent with those of the Sagittarius dwarf galaxy, as had been predicted. In one direction, we are even able to detect a clump of red stars, similar to that of the Sagittarius dwarf, from stars spread across 110 deg^2 of sky. Focusing on stars with the colors of F turnoff objects, we identify at least five additional overdensities of stars. Four of these may be pieces of the same halo structure, which would cover a region of the sky at least 40° in diameter, at a distance of 11 kpc from the Sun (18 kpc from the center of the Galaxy). The turnoff is significantly bluer than that of thick-disk stars, yet the stars lie closer to the Galactic plane than a power-law spheroid predicts. We suggest two models to explain this new structure. One possibility is that this new structure could be a new dwarf satellite of the Milky Way, hidden in the Galactic plane and in the process of being tidally disrupted. The other possibility is that it could be part of a disklike distribution of stars which is metal-poor, with a scale height of approximately 2 kpc and a scale length of approximately 10 kpc. The fifth overdensity, which is 20 kpc away, is some distance from the Sagittarius dwarf streamer orbit and is not associated with any known Galactic structure. We have tentatively identified a sixth overdensity in the halo. If this sixth structure is instead part of a smooth distribution of halo stars (the spheroid), then the spheroid must be very flattened, with axial ratio $q = 0.5$. It is likely that there are many smaller streams of stars in the Galactic halo.

Subject headings: Galaxy: halo — Galaxy: structure

Hmmm... How do I do this for the whole sky?

Monoceros stream CMD



CasJobs X

http://casjobs.sdss.org/CasJobs/jobdetails.aspx?id=4930248

SDSS Query / CasJobs

Help Tools Query History MyDB Import Groups Output Profile Queues SkyServer Logout mjuric

'My Query' Details

[Resubmit Job](#)

JobID	TaskName	Context	Queue	Submitted	Started	Finished	Status
4930248	My Query	DR7	600	7/28/2010 8:04:02 AM	7/28/2010 8:04:10 AM	7/28/2010 8:09:21 AM	Finished

Executed on Rows Message

DR7Best long 94144 Query Complete

Query

```
select
    s.objid,
    s.l, s.b,
    s.ra, s.dec,
    s.extinction_r as Ar,
    s.dered_u as u, s.dered_g as g, s.dered_r as r, s.dered_i as i, s.dered_z as z,
    s.err_u, s.err_g, s.err_r, s.err_i, s.err_z,
    s.flags

from
    Star s

into
    mydb.northslice3

where
    b > 40
    and 0.2 < g - r and g - r < 0.4
    and 21 < r and r < 21.1
```

Contact
\$Name: v3_5_16 \$,\$Revision: 1.23 \$, Last modified: Wednesday, May 14, 2008 at 1:52:25 AM

MyDB

http://casjobs.sdss.org/CasJobs/MyDB.aspx

SDSS Query / CasJobs

Help Tools Query History MyDB Import Groups Output Profile Queues SkyServer Logout mjuric

Views

Tables

Functions

Procedures

Sort by... All selected...

Rows	kB	Name
94,144	13,768	northslice3

Mario Juric 's MyDB

14,232 kB of 2,000,000 kB used

From this page you can get various information about the contents of both your and shared tables within your groups. Click the left table links to get information about a specific table, such as rows, columns or size. From the table pages you can perform various table-specific tasks, such as:

- Download a table
- Manage your group tables
- Rename a table
- Drop a table

Sizes are approximations only.
Row counts are approximations only. For exact value run a count.
There's always some overhead, even empty MyDB's take up space.
Group tables do not count towards your MyDB size limit.

Contact
\$Name: v3_5_16 \$,\$Revision: 1.64 \$, Last modified: Tuesday, January 27, 2009 at 3:15:10 PM

MyDB

http://casjobs.sdss.org/CasJobs/MyDB.aspx

SDSS Query / CasJobs

mjuric

Views

Tables

Functions

Procedures

Sort by... All selected... Rows kB Name

94,144	13,768	northslice3
--------	--------	-------------

northslice3

Contains ~94,144 rows (~13,768 kB)

Notes Sample Job Plot BPlot Download Publish Neighbors Rename Drop

Table Schema type [size]

objid	I	b	ra	dec	Ar	u	g	r	g	i	z	Up
bigint [8]	float [8]	float [8]	float [8]	float [8]	float [8]	float [8]	float [8]	float [8]	float [8]	float [8]	float [8]	Up
Upper Limit	Up											
Lower Limit	Low											

X X X X X X X X
 Y Y Y Y Y Y Y Y

Plot Table

This produces non-interactive plots and is suitable for plotting a great numbers of objects. If you are only interested in plotting a few objects, you may want to try out the 'Quick' plot functionality, which can be reached by clicking 'Plot' on the query page after any 'Quick' query has completed.

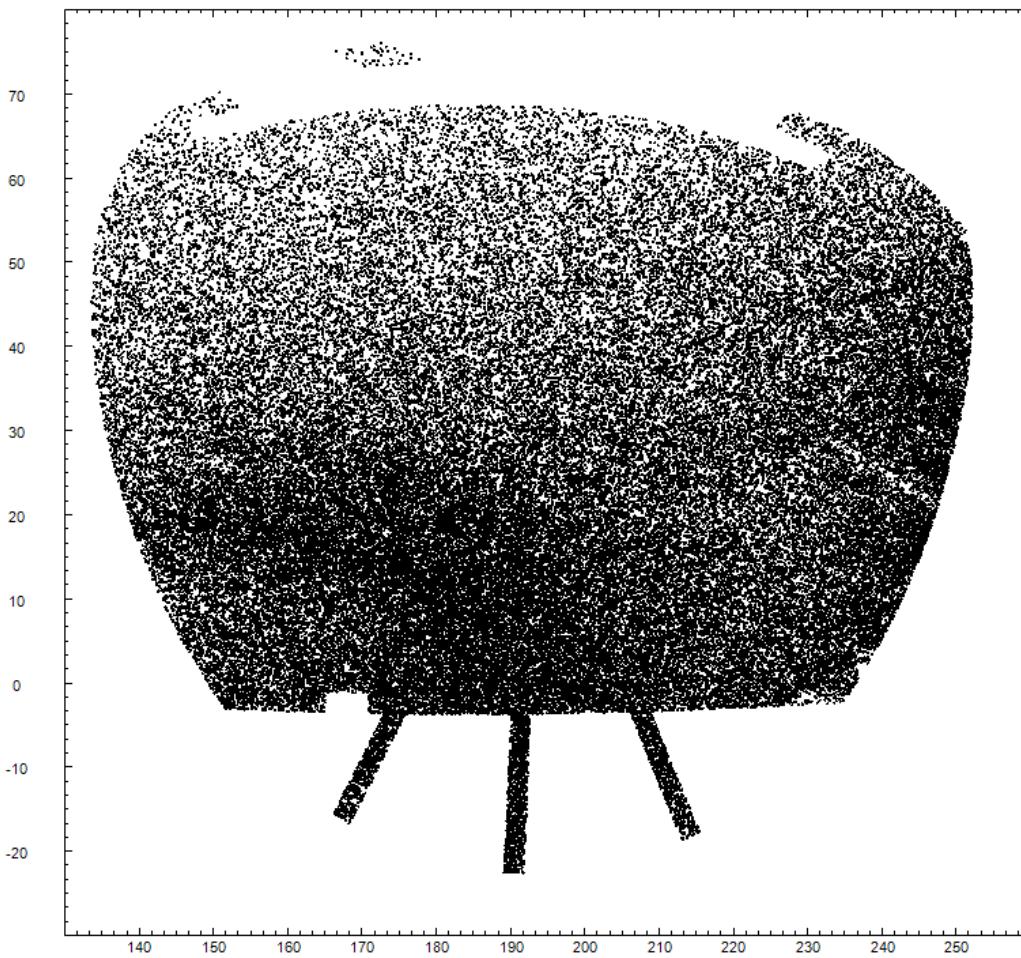
To plot a table, pick columns from above to represent values on the X and Y axis, then click 'Plot'. Only one 'X' and one 'Y' may be selected. A plotting output job with the parameters you've specified will be created and placed in queue.

You can also specify constraints on your plot by entering values into the limit fields. Any blank fields are considered unbounded. Constraints entered for columns that are not actively being plotted will still affect the plot.

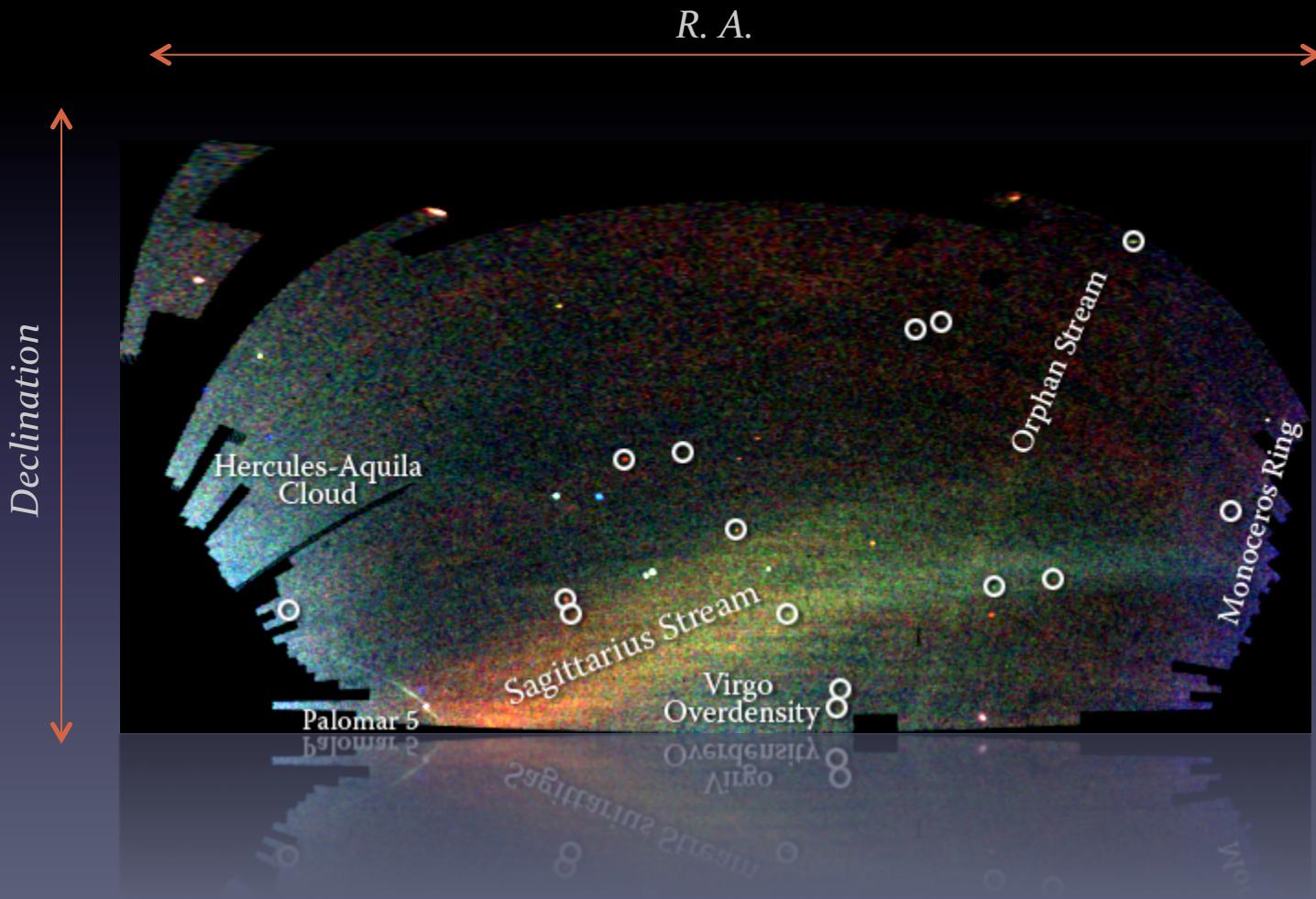
You may also manually specify your constraints by clicking 'Manually Enter Plot Constraints'.

Plot Manually Enter Plot Constraints

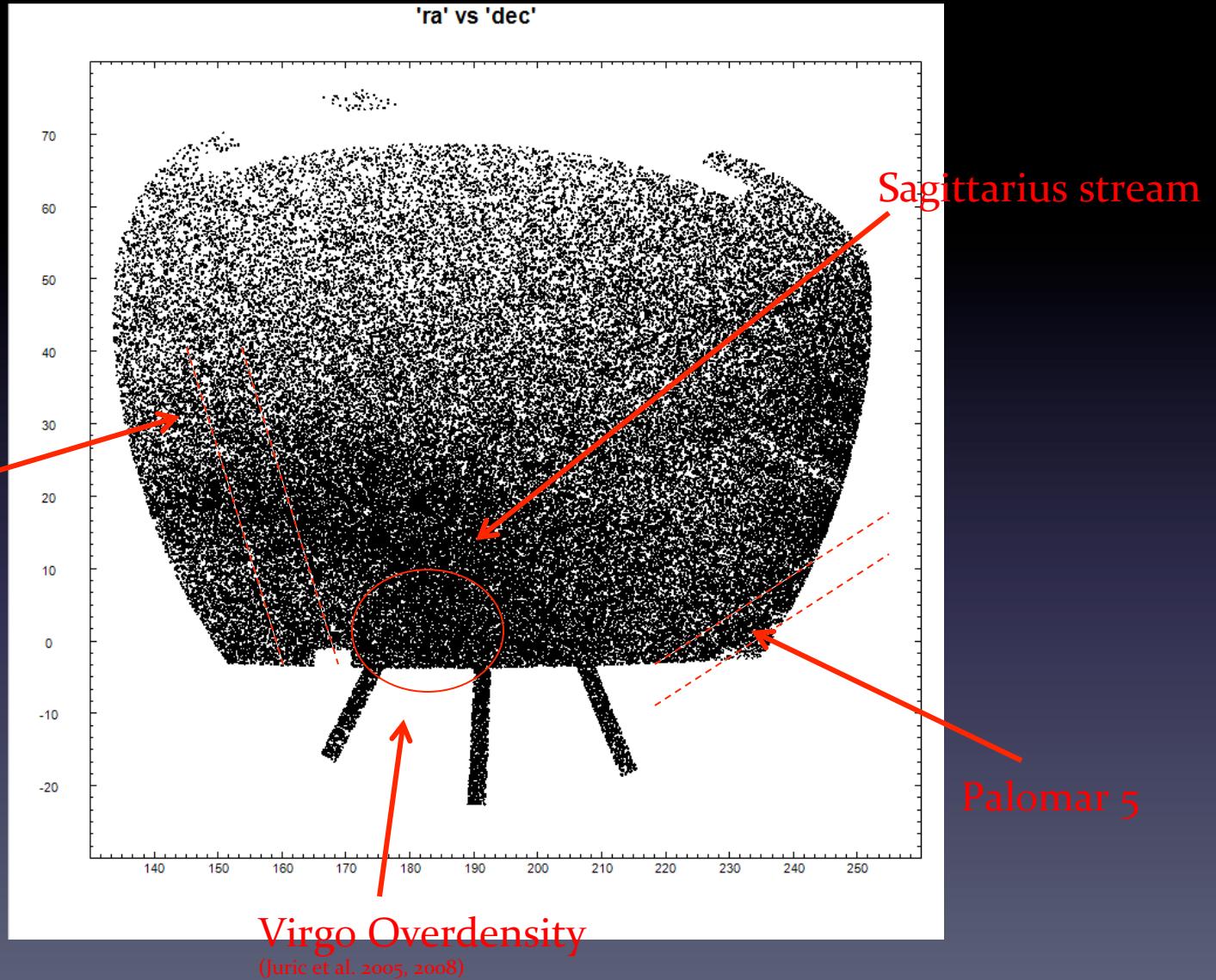
'ra' vs 'dec'



Field of Streams



Orphan Stream
(Belokurov et al. 2006)



Virgo Overdensity
(Juric et al. 2005, 2008)

- Let's summarize what we just did:
 - In less than 10 minutes, we “observed” two fields 120deg apart (~4 months), to 22nd magnitude, in 5 bands
 - We “reduced” and extinction-corrected the data (~70,000 stars/field) and analyzed it using color-magnitude plots
 - We discovered a main sequence feature in one of the streams at distance around 11.5kpc, likely metal poor
 - We then proceeded to discover three more streams over the whole sky
- 10 years ago, this would have been a multi-night observing proposal for top telescopes in the world. The challenge is obtaining the data.
- Today it’s an SQL query – *the data is already there, the challenge is to ask the right question!*

Astronomy in the Age of Large Surveys

- Traditionally, astronomy was a data-starved science. Our approach to research and our analysis methods were shaped by this environment. Surveys are altering it; data is becoming abundant and of unprecedented quality.

Sloan Digital Sky Survey

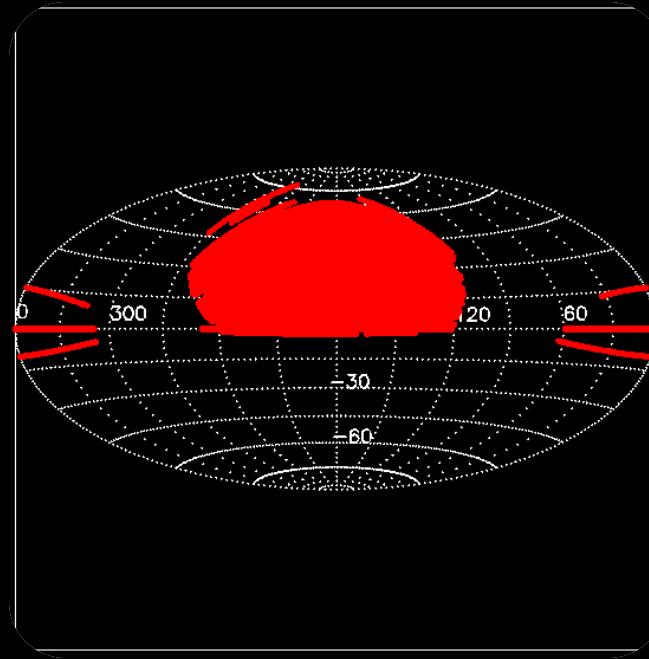
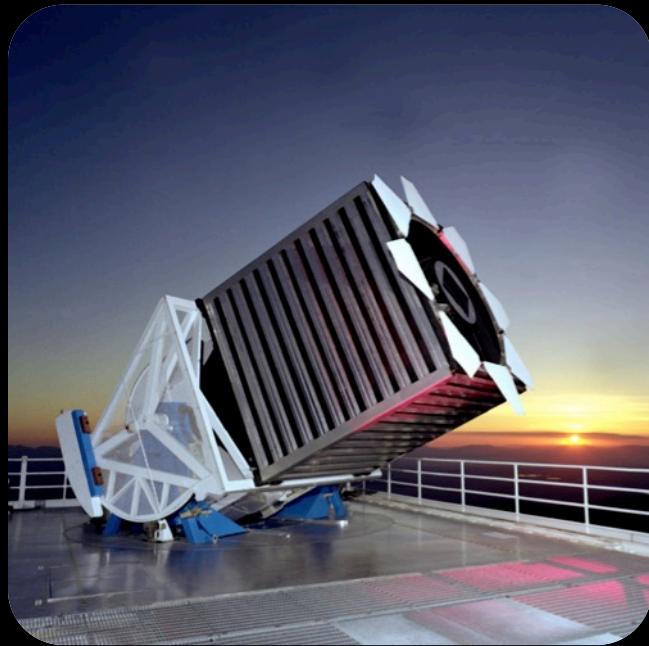
2.5m telescope

>14500 deg²

0.1" astrometry

r<22.5 flux limit

5 band, 2%, photometry for >460M objects
Millions of spectra



10 years of ops: ~10 TB of imaging

Panoramic Survey Telescope and Rapid Response System

1.8m telescope

30,000 deg²

50mas astrometry

r<23 flux limit

5 band, better than 1% photometry (goal)



~700 GB/night

LSST: A Deep, Wide, Fast, Optical Sky Survey



8.4m telescope

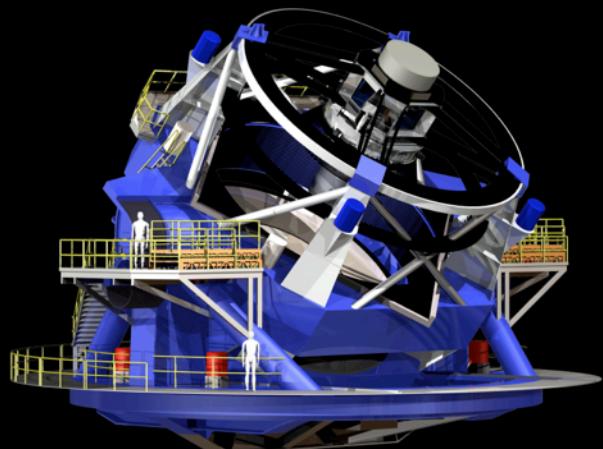
18,000+ deg²

10mas astrom.

r<24.5 (<27.5@10yr)

ugrizy

0.5-1% photometry

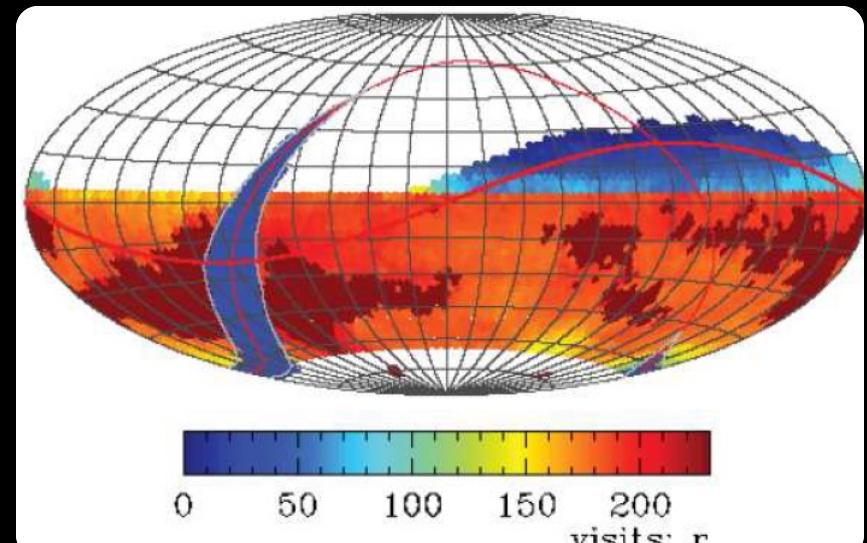


3.2Gpix camera

30sec exp/4sec rd

15TB/night

37 B objects



Imaging the visible sky, once every 3 days, for 10 years (825 revisits)

Low-Frequency Array for Radio Astronomy (LOFAR)



Estimated Data Volume

Data Type	Data Rate [MB/s]	Data Rate [GB/h]
Burst Mode	100	360
Monitoring Mode	0.03	0.1
Spectrometer Mode	0.34	1.2

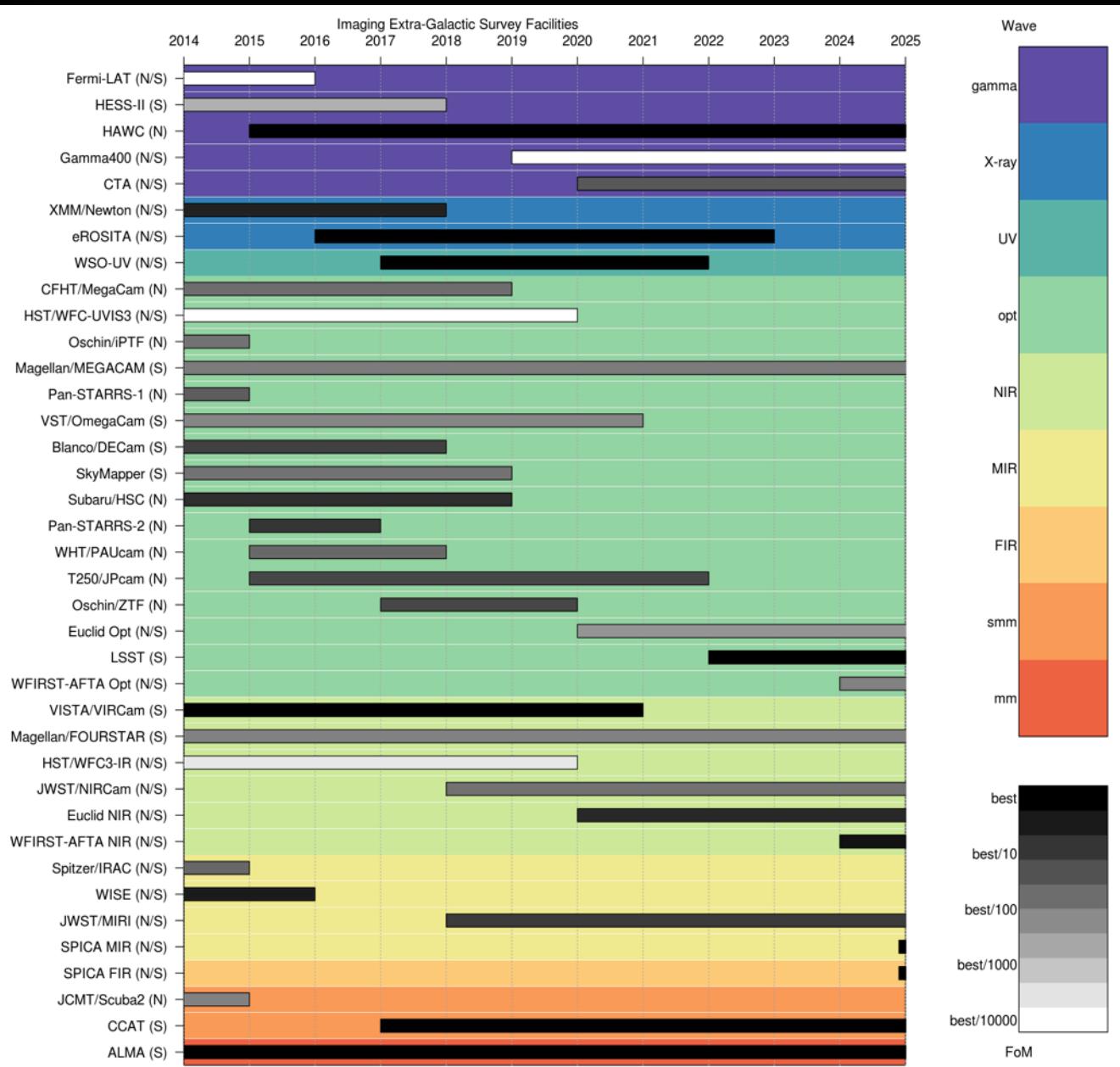


Square Kilometer Array



SPDO / Swinburne Astronomy Productions

10 PB / hour



Astronomy in the Age of Large Surveys

- Traditionally, astronomy was a data-starved science. Our approach to research and our analysis methods were shaped by this environment. Surveys are altering it; data is becoming abundant and of unprecedented quality.
- For optical, LSST will cap this transformation: it will deliver the positions, magnitudes and variability information for virtually *everything* in the southern sky to 24th-27th magnitude, with an order of magnitude better controlled systematics than current surveys.

LSST: A Quantitative Difference

Final Image Collection – All DRs	515 PB	All Data Releases Includes Virtual Data (475 PB)
Final Image Collection – DR11	114 PB	Data Release 11 (Year 10) Includes DR11 Virtual Data (88 PB) and all raw images from all years
Final Database	16 PB	Data Release 11 (Year 10) Includes Data, Indexes, and DB Swap
Final Disk Storage	375 PB	Archive and Base Sites
Final Tape Storage	121 PB	Single Site, Single Copy Only
Peak Number of Nodes	1750	Archive and Base Sites Compute and Database Nodes
Number of Alerts Generated	28 billion	Life of survey

- **Virtual Data is data that is dynamically recreated on-demand from provenance information**

LSST: A Qualitative Difference

9 hours on CFHT/MegaCam (Munoz et al. 2010)

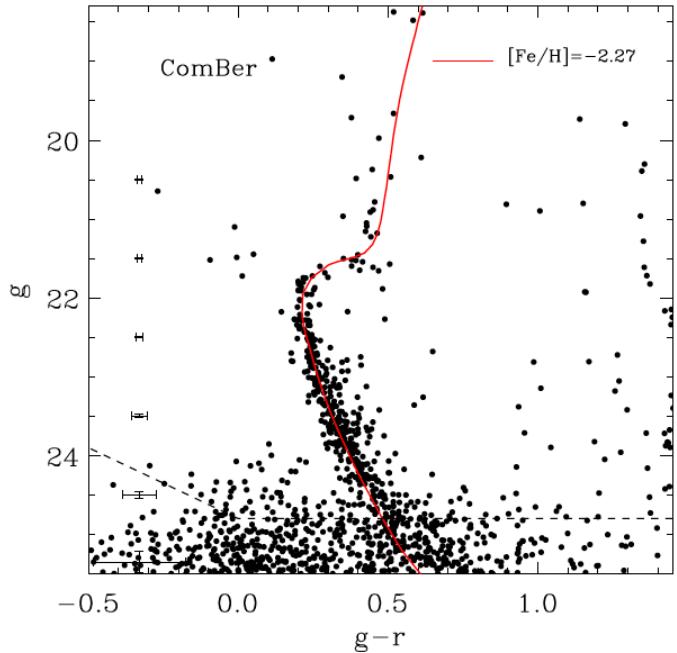
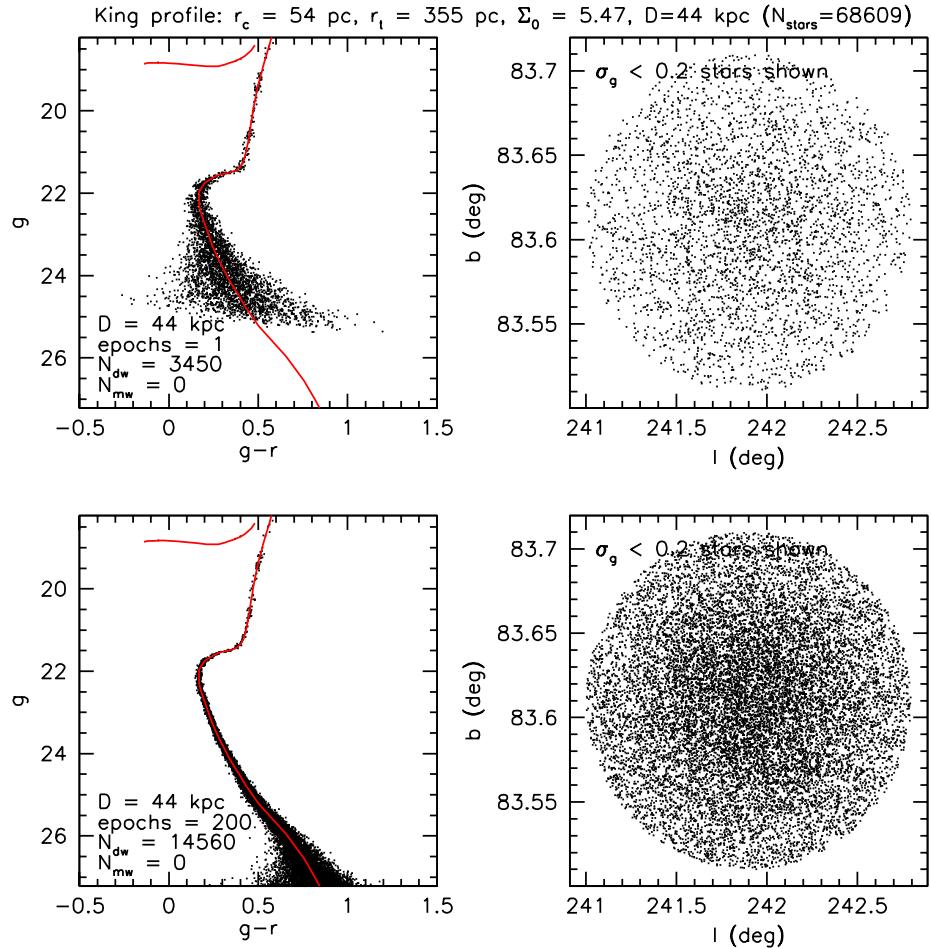


Figure 3. CMD for the inner region ($r < 6'$) of ComBer. The dashed lines mark the 90% completeness level after χ and sharp cuts have been applied to remove non-stellar objects. As it can be seen, our CFHT photometry reaches at least 3 mag below the main sequence turn-off of ComBer. We have complemented our photometry with SDSS data for $g > 20$. The error bars to the left were determined from the artificial star tests and represent the standard deviation of a Gaussian function fitted to the error distribution as a function of magnitude. A theoretical isochrone for a 13 Gyr old, $[Fe/H] = -2.27$ population is shown with a solid red line (from Girardi et al. 2004).

(A color version of this figure is available in the online journal.)

LSST, single-epoch and full-survey (simulation)



Astronomy in the Age of Large Surveys

- Traditionally, astronomy was a data-starved science. Our approach to research and our analysis methods were shaped by this environment. Surveys are altering it; data is becoming abundant and of unprecedented quality.
- For optical, LSST will cap this transformation: it will deliver the positions, magnitudes and variability information for virtually *everything* in the southern sky to 24th-27th magnitude, with an order of magnitude better controlled systematics than current surveys.
- **We're entering the age of abundance of high quality data. Success in research will depend on the ability to analyze and mine knowledge from that data.**

Next time ...

- Software Tools
 - Python
 - IPython
 - Git and github
 - Basic numerical and analysis libraries