6630 Final Project Process Book

Anthony Garcia, Stefan Kapetanovic

anthonygarcia918@hotmail.com

stefankapetanovic1@gmail.com

TABLE OF CONTENTS

Basic Information-

- Title: Visualizing the Sloan Digital Sky Survey
- UID: u0805376 and u0871074
- Github Repository: https://github.com/stefankapetanovic/Visualizing-Galactic-Chemical-Distributions-with-SDSS

Overview and Motivation-

As a Physics major, Anthony Garcia has a background in working with researchers in astronomy. His background will be the basis of our gathered data, the approach we take in visualizing the data, and as a source of help/reference when necessary. Our primary motivation, is to encourage exploration of Sloan Digital Sky Survey (SDSS) data by the general public. Trends of chemical composition for stellar objects provide clues to galactic chemical evolution. By providing a means to explore a large set of stars in a visual way, these trends can be visualized. We want to display a very informational dataset of galactic coordinates in a pleasing and visually interactive way. This will allows us to identify and understand the chemical distribution trends found in the Milky Way Galaxy. We also would like to dedicate a portion to exploratory eye-candy for the general public. Exploratory results will heed star sizes, star temperatures, star distances relative to Earth, etc. Providing unique and interactive celestial visuals with a informational backing will be our primary focus.

- Related Work: From our research we concluded that no work out (to our knowledge resembles our visualizations). That was a primary motivation for us. Creating something new and useful for beginners and advanced astronomers alike. We knew from class work we could generate a table that resembled a periodic table as well as a scatter plot. The challenge of generating a visual that we had not had previous experience on was also a goal of ours as we generated our website concept.
- Questions: With a direct motive from astronomy professors stating that this would be a unique and informational project we decided to visually encompass the stars of the Milky Way Galaxy. Simplifying we are trying to answer the correlation questions involved with the stars of the Milky Way Galaxy. We knew from the beginning of acquiring our data we would have a variety of questions we could answer and we chose to develop a system that could answer as many questions as possible. Our data was so vast that at times it was overbearing so we had to make a decsion of how and what to display in order to satisfy visual constraints. Over the course of the semeter our project evolved to What questions are you trying to answer? How did these

questions evolve over the course of the project? What new questions did you consider in the course of your analysis?

Data and Data Processing-

Our SDSS DR14 data that has been crossmatched with GAIA observations. This data is formulated and structured to reflect the entire Milky Way Galaxy. It includes data on thousands of stars. Some of the data present is name, size, temperature, velocity, chemical abundance, etc. These data elements are what our projects basis will formulate around. WE will have to ensure that our data points do not start to contradict with the amount of visualizations we include. Cross Matching between identification numbers will be made between two different data sets will need to be performed. Derived data will be mostly positional information. Converting from one coordinate system (helio coordinates) to another (galactic central). We may also need to transform our chemical abundance values to be with respect to another element. For example, number of carbon atoms relative to the number of iron atoms may need to be converted to number of carbon atoms relative to nitrogen atoms.

Our data at a glance:

1	TEFF	LOGG	VMICRO	M_H	ALPHA_M	C
2	4830.54931640625	2.3597970008850098	1.327027678489685	-0.5674302577972412	0.11575902998447418	-0.03610187768936157
3	5104.53857421875	2.5360970497131348	1.6218290328979492	-0.680000901222229	0.24407699704170227	0.13198256492614746
4	4675.74658203125	2.1907029151916504	1.393060564994812	-0.9145915508270264	0.0712478756904602	-0.13521409034729004
5	4839.9736328125	1.9356029033660889	1.5959526300430298	-1.3047945499420166	0.2721533179283142	-0.330649733543396
6	5009.447265625	2.4841980934143066	1.612129807472229	-0.6034702062606812	0.0023045185953378677	-0.29206013679504395
7	3761.522216796875	0.7062593102455139	2.4370810985565186	-0.9575276374816895	0.24762605130672455	0.058451056480407715
8	5267.61376953125	2.2374765872955322	1.6867083311080933	-1.4991663694381714	0.3999782204627991	0.2583479881286621
9	5064.638671875	2.3741235733032227	1.6223657131195068	-0.6471663117408752	0.19101104140281677	0.11708474159240723
10	4807.6318359375	1.7507444620132446	0.8084999322891235	-1.9340882301330566	0.27778691053390503	-0.6936569213867188
11	4407.5498046875	1.816306471824646	1.9489020109176636	-0.4964139461517334	0.175821453332901	-0.030649960041046143
12	5050.17578125	2.2886271476745605	1.5789213180541992	-0.940174400806427	0.17884986102581024	-0.26201099157333374
13	4813.64404296875	2.5725135803222656	1.139188289642334	-0.6816353797912598	0.22070573270320892	0.0005702674388885498
14	4418.77294921875	1.8284544944763184	1.657219409942627	-0.4560634195804596	0.15269099175930023	0.03469541668891907
15	4653.27392578125	2.4687271118164062	1.260549783706665	-0.4243336021900177	0.19441260397434235	0.0374530553817749
16	4308.109375	2.3086113929748535	0.6511634588241577	-0.7940083742141724	-0.09594283998012543	0.25809890031814575
17	4981.2490234375	2.2854807376861572	1.7237330675125122	-0.7440061569213867	0.254031240940094	0.046630680561065674
18	4588.13818359375	1.814576506614685	1.4115450382232666	-0.6284843683242798	0.11248953640460968	-0.10999814420938492
19	4476.4326171875	1.640163779258728	1.6021398305892944	-1.0588468313217163	0.20495419204235077	-0.1636684536933899
20	4482.52294921875	1.949280023574829	1.2943490743637085	-0.6472985744476318	0.207508385181427	0.022253498435020447
21	4631.23876953125	2.0627846717834473	1.5104279518127441	-0.861946702003479	0.2858988046646118	0.2194281816482544
22	4473.56982421875	1.9627057313919067	1.3130508661270142	-0.8575751781463623	0.23898936808109283	0.026591777801513672

The above data set marks a subset of data used for coordinating the scatter plot axes. The different tags represent different measurements amongst the stars. In this set a few of the tags relate to temperature, x and y position, metallicity, velocity, etc. We deliberated to change the tags to more user friendly tags but decided that the tag names are best when represented directly from the APOGEE data.

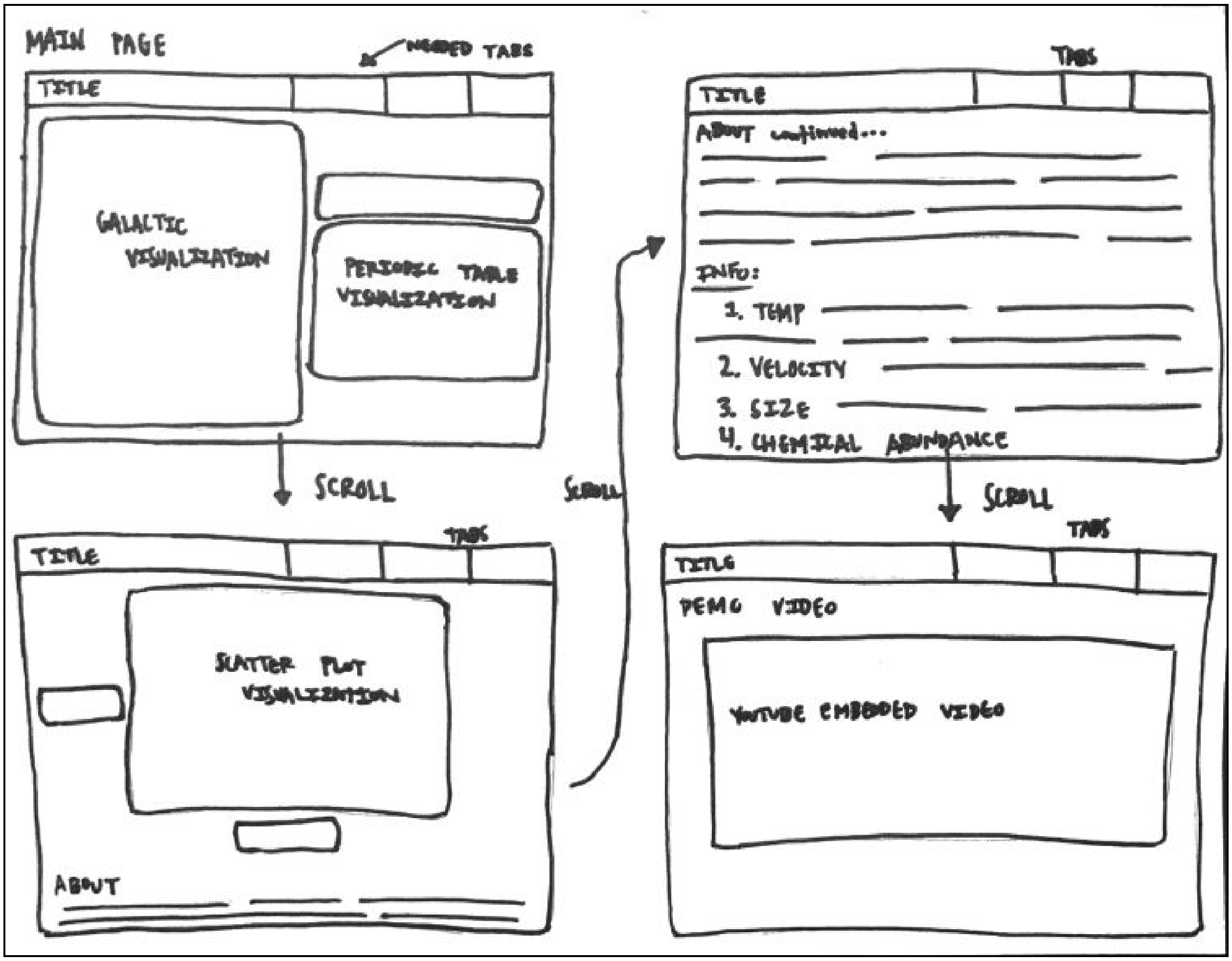
Below is data that contains all the necessary info required to generate our periodic table. All of the atomic quantities are present allowing us to use the tags to plot out chemicals in the exact spots they need to be.

atomicNumber	symbol	Name	atomicMass	cpkHexColor	electronicConfiguration	electronegativity	atomicRadius	ionRadius	vanDelWa
1	H	Hydrogen	1.00794(4)	FFFFFF	1s1	2.2	37	120	120
2	He	Helium	4.002602(2)	D9FFFF	1s2	32	32	140	140
3	Li	Lithium	6.941(2)	CC80FF	[He] 2s1	0.98	134	76 (+1)	182
4	Be	Beryllium	9.012182(3)	C2FF00	[He] 2s2	1.57	90	45 (+2)	900
5	B	Boron	10.811(7)	FFB5B5	[He] 2s2 2p1	2.04	82	27 (+3)	801
6	C	Carbon	12.0107(8)	909090	[He] 2s2 2p2	2.55	77	16 (+4)	170
7	N	Nitrogen	14.0067(2)	3050F8	[He] 2s2 2p3	3.04	75	146 (-3)	155
8	O	Oxygen	15.9994(3)	FF0D0D	[He] 2s2 2p4	3.44	73	140 (-2)	152
9	F	Fluorine	18.9984032(5)	9E+051	[He] 2s2 2p5	3.98	71	133 (-1)	147
10	Ne	Neon	20.1797(6)	B3E3F5	[He] 2s2 2p6	69	69	154	154
11	Na	Sodium	22.98976928(2)	AB5CF2	[Ne] 3s1	0.93	154	102 (+1)	227
12	Mg	Magnesium	24.3050(6)	8AFF00	[Ne] 3s2	1.31	130	72 (+2)	173
13	Al	Aluminum	26.9815386(8)	BFA6A6	[Ne] 3s2 3p1	1.61	118	53.5 (+3)	578
14	Si	Silicon	28.0855(3)	F0C8A0	[Ne] 3s2 3p2	1.9	111	40 (+4)	210
15	P	Phosphorus	30.973762(2)	FF8000	[Ne] 3s2 3p3	2.19	106	44 (+3)	180
16	S	Sulfur	32.065(5)	FFFF30	[Ne] 3s2 3p4	2.58	102	184 (-2)	180
17	Cl	Chlorine	35.453(2)	1FF01F	[Ne] 3s2 3p5	3.16	99	181 (-1)	175
18	Ar	Argon	39.948(1)	80D1E3	[Ne] 3s2 3p6	97	97	188	188
19	K	Potassium	39.0983(1)	8F40D4	[Ar] 4s1	0.82	196	138 (+1)	275
20	Ca	Calcium	40.078(4)	3DFF00	[Ar] 4s2	1	174	100 (+2)	590
21	Sc	Scandium	44.955912(6)	E6E6E6	[Ar] 3d1 4s2	1.36	144	74.5 (+3)	633

Visualization Design Drafts-

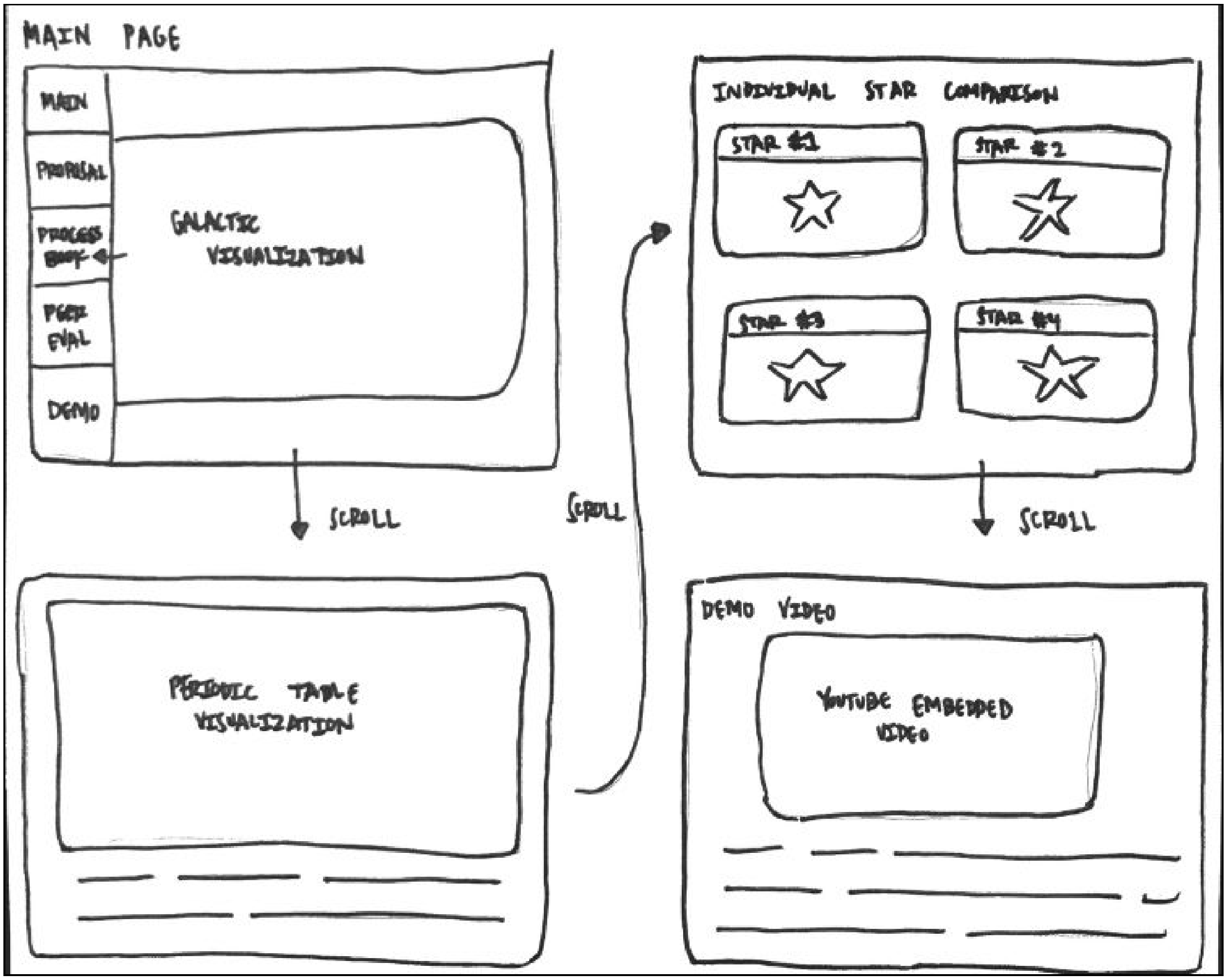
An overview of the designs are given below. Each design gives it pros and cons while explaining what is included into the final design. We include the layout of our visualizations but not the visualizations themselves because they will be constructed in the coming weeks. The general outlines of the three major visualizations we plan on certainly having include a galactic representation visual with stars and interactivity, the second is an interactive periodic table of elements, and the third is a scatter plot where the user can adjust axes.

Final Design:



Our final design envelopes the overarching theme of space. We believe it's an elegant and aesthetic approach to display celestial information. It is fully scrollable on one webpage but the tabs at the top directly and smoothly takes you to the section you want to view. We know we want at a minimum three different visualizations that are all connected. The major visualization on the main page will be directly interactive while the scatter plot will respond to the choices made above by the user. The major visualizations include a galactic overview and the periodic table of elements. All the designs incorporate the demo video at the end of the webpage so the user has the opportunity to go through the entire site and then watch a video about full interactivity.

Design 2:

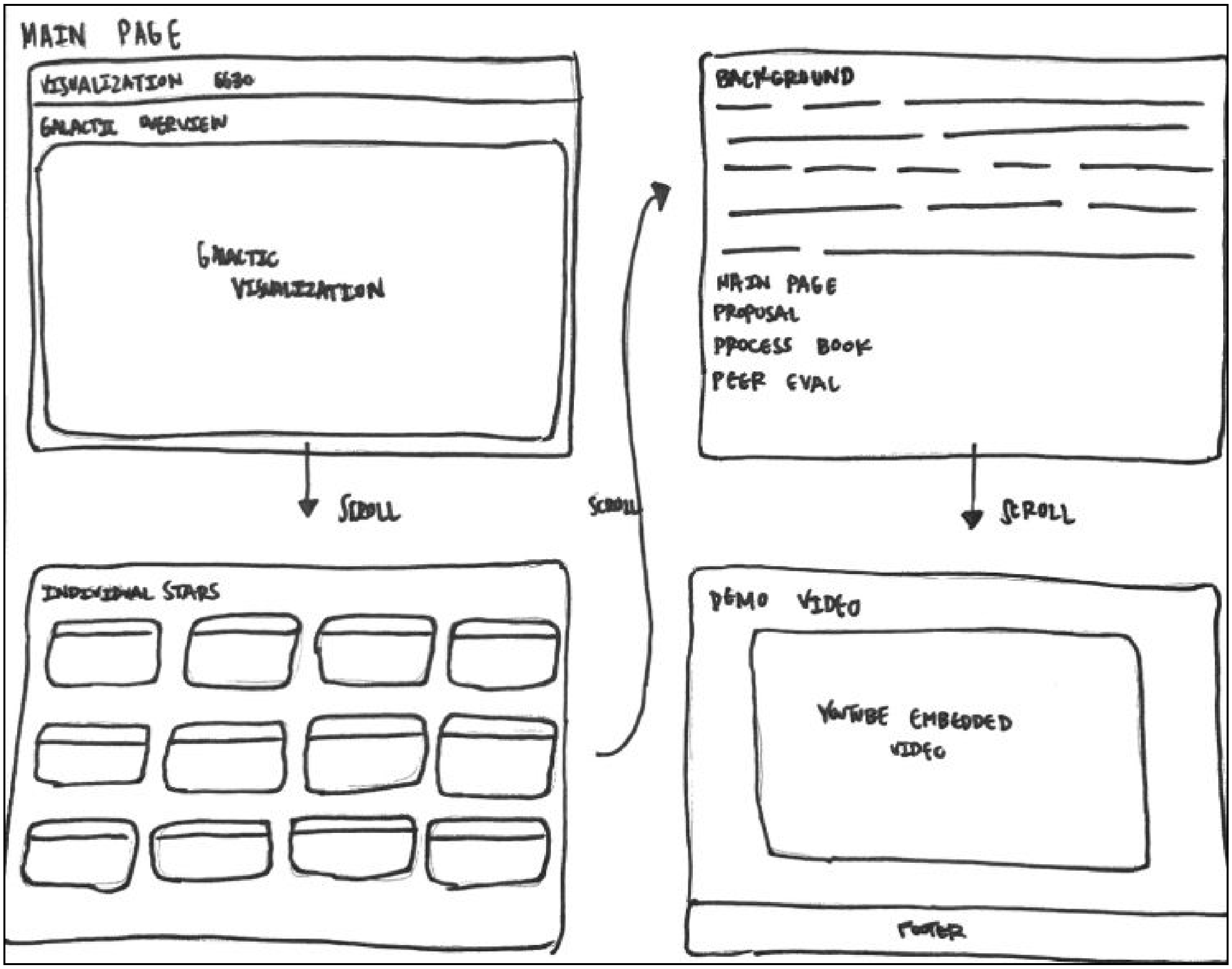


Our second design was interesting because it contained a collapsable side view that would direct you to the sections you wanted to view. We also had an interesting idea of star comparison but ultimately determined that we could get more interactivity and visual intrigue in a different way. This design separates the galactic visualization of the Milky Way and the periodic table. This will be critical decision made in the future:

- The layout of the three visualizations we plan to incorporate are vital since some visuals interact with one another.
- Ideally, we want to include all of our visualizations on the same page view. Minimizing the amount of scrolling for the user.

We decided we wanted the galactic and scatter plot to be interactive and change as the user played with the chemicals in the scatter plot and thus the stars would change. We will combine those visuals so they are easily and elegantly seen side by side if the monitor optimally allows for them.

Design 3:



This design provided our main visualization as the only major interactive piece. We wanted to balance our staple piece but also knew we'd need more visualizations to be a great project. Below the major visualizations were a series of stars and their entire scope demographics (i.e. name, size, velocity, temperature, etc.). This design ultimately had too much text throughout that we didn't like as a team. We wanted our visuals to convey everything they needed in the least amount of text. This will be a tricky design test that we will play around with as the project progresses.

Peer Feedback-

Anthony and I, Stefan Kapetanovic grouped with our peer Srija Adusumilli to evaluate our project thus far. WE got some very interesting feedback from Srija that we hadn't considered. That feedback is given below-

- A question that came up was if hovering over a star would elicit a tooltip in our galactic visual. We had planned on a tooltip but the information that we would include was still undecided. In this situation we got to see what Srija, a self proclaimed novice in astronomy, would want to see as a casual viewer. Her response was very interesting.

    - Basic star information like size, temperature, velocity,
    - Distance from planet Earth
    - Chemical abundance did not intrigue her particularly
- Srija was very impressed with the data we had already collected. In lots of aspects she thought we were ahead of the general student. She liked the story we were telling with our visualizations. She warned us to be careful with the color schemes and to keep the color of the periodic table visual we are planning consistent with the galactic visual.
- She also brought up a point that we had considered right before coming to class. The amount of data we are bringing in could be a worry. We may have to run asynchronous tasks to accommodate for a smooth functioning website.
- In terms of animation, Sirja said we may have to consider just a bit more in our scatter plot and tooltip. We completely agree we could incorporate more and will as we progress through our visuals.
- When it came to ask questions about displaying the amount of stars we wanted to get feedback on if we should cluster, average or just use smaller data samples. The risk of using all of our data which is over 250,000 stars is the possibility of overlapping stars that could generate unappealing visuals. To solve this we believe that if we took the average stars in a cluster and then displayed that information in our galactic visual we could solve two problems simultaneously. This would decrease the amount of data we would have to process while also clearing our visual by not having to account for every single stars data throughout the process.

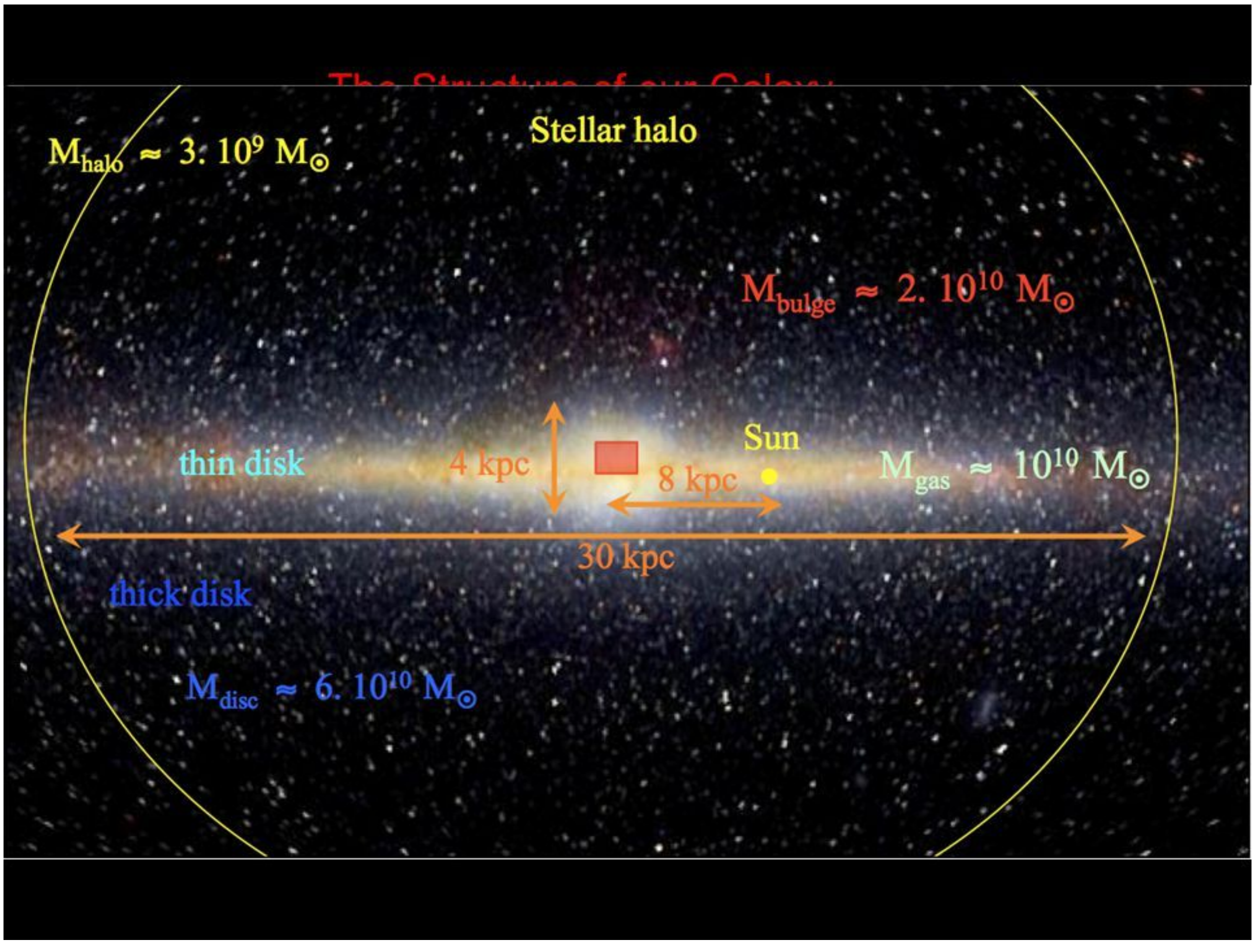
Afterwards, when grouping together we talked about our feedback and relayed to one another that our project is solid but fine tuning the details will elevate it to where we want it to be. We loved the feedback we got from Srija and thought it was very fair and more importantly very helpful. We know what we need to do in order to generate our visualizations in an elegant way. Averaging the star data prior to processing it will be a huge benefit. Clustering the stars into cells will also help us transcribe the data to the visual. We will make sure to use color schemes that help our project appearance and to display a broad enough amount of information where a novice and expert could find our project intriguing and fun to use.

Design Evolution-

When designing our visualizations we quickly realized that the amount of stars compared to the amount of real estate we had to work with on the monitor was overlapping. We needed some way of being able to show all the stars in the milky way galaxy effectively without causing an overlap in the visualization and to do this we chose to do quite a bit of data processing. We had to filter the amount of stars that we performed our clustering and data processing on. This became a spatial decision and the most practical method to encompass the most amount of stars without compromising visualization in the galactic view.

The filters we chose:

- Cartesian Z position: -2 to 2
- Radius of 30000 light years from the sun
- Abundance measurements that are flagged from the apogee that took the measurements



We decided deploy clustering to our advantage and to take an average of the star data we had accumulated. Clustering and then averaging the star data into cells allowed us to generalize enough information while continuing to provide a tremendous deal of detailed information for the user to enjoy. This will make our overlay of data on our galactic visual much more organized and the benefits greatly outweigh the necessity for every star to be displayed which our original notion entailed.

Must-Have Features-

1. Positional information related to chemical abundance

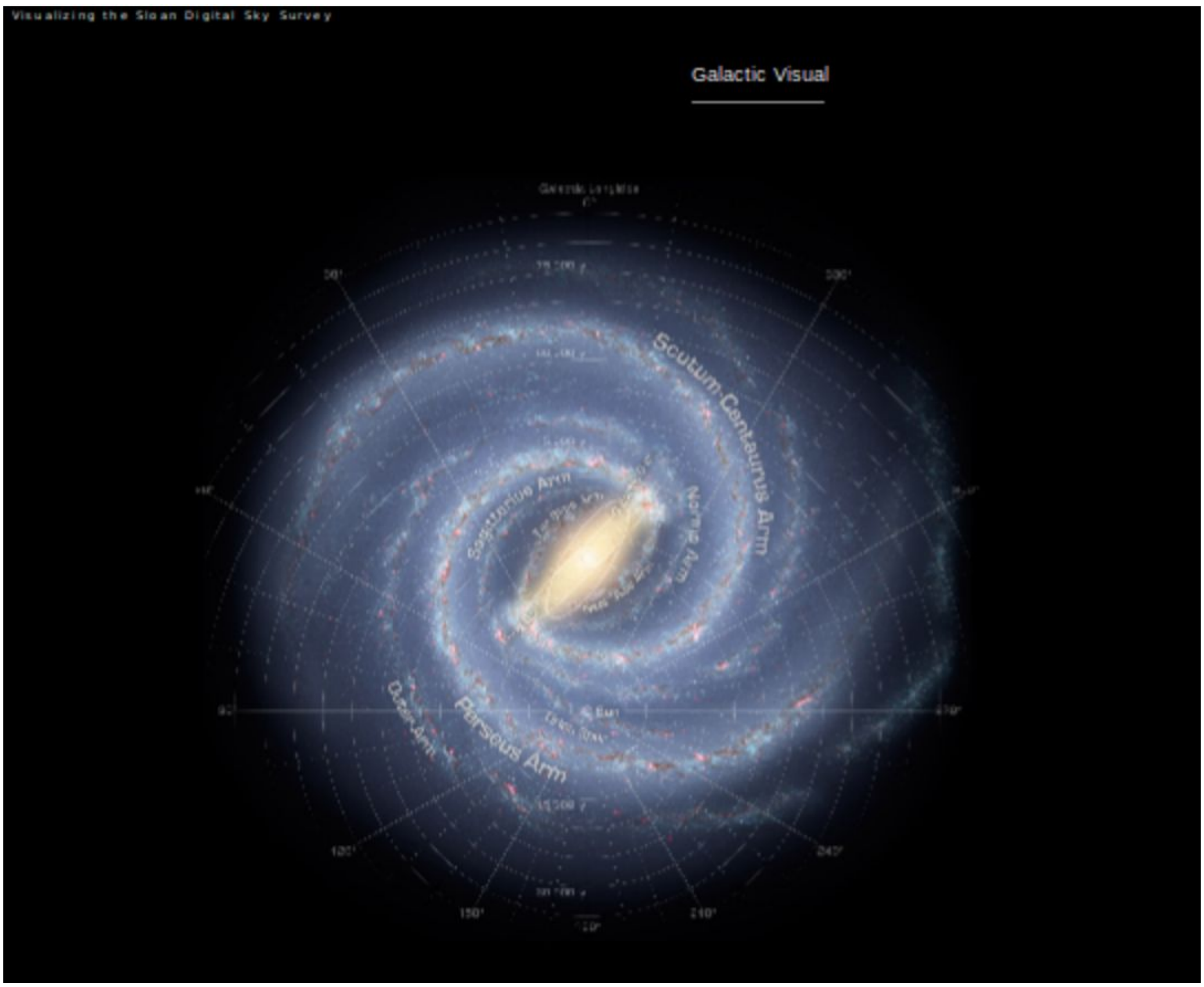
2. Interaction to change between elements and different quantities (ex. Temperature, mass, velocity, chemical abundance, etc.)
3. Interaction with the periodic table of elements and the galactic visualization
4. Scatter plot relation plot. With axis adjustments available

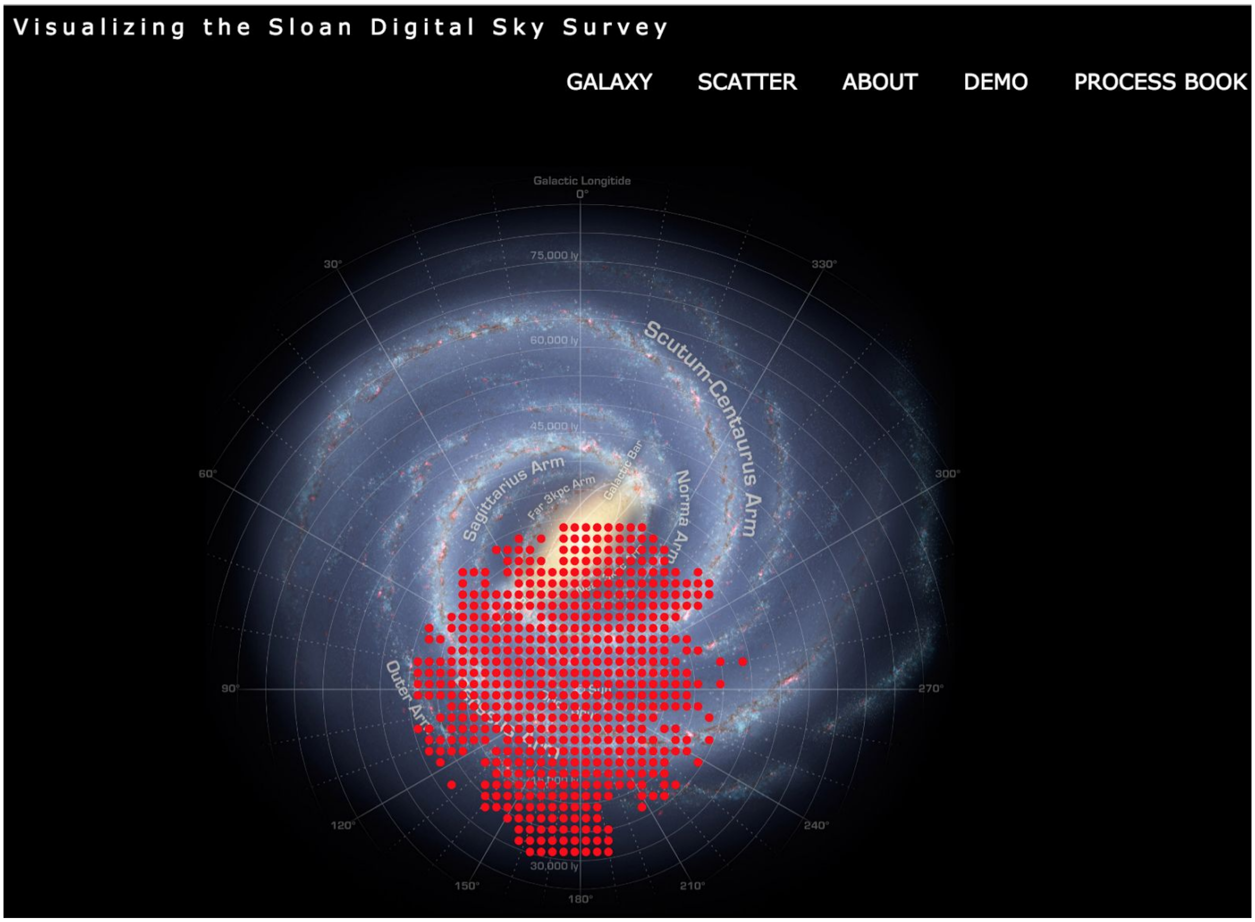
Optional-

1. Exoplanet cross reference. (Stars that are known to have planets orbiting them)
2. Visual background of a galaxy
3. Brush selection in scatter plot that will highlight in galactic plot

Galactic Visual-

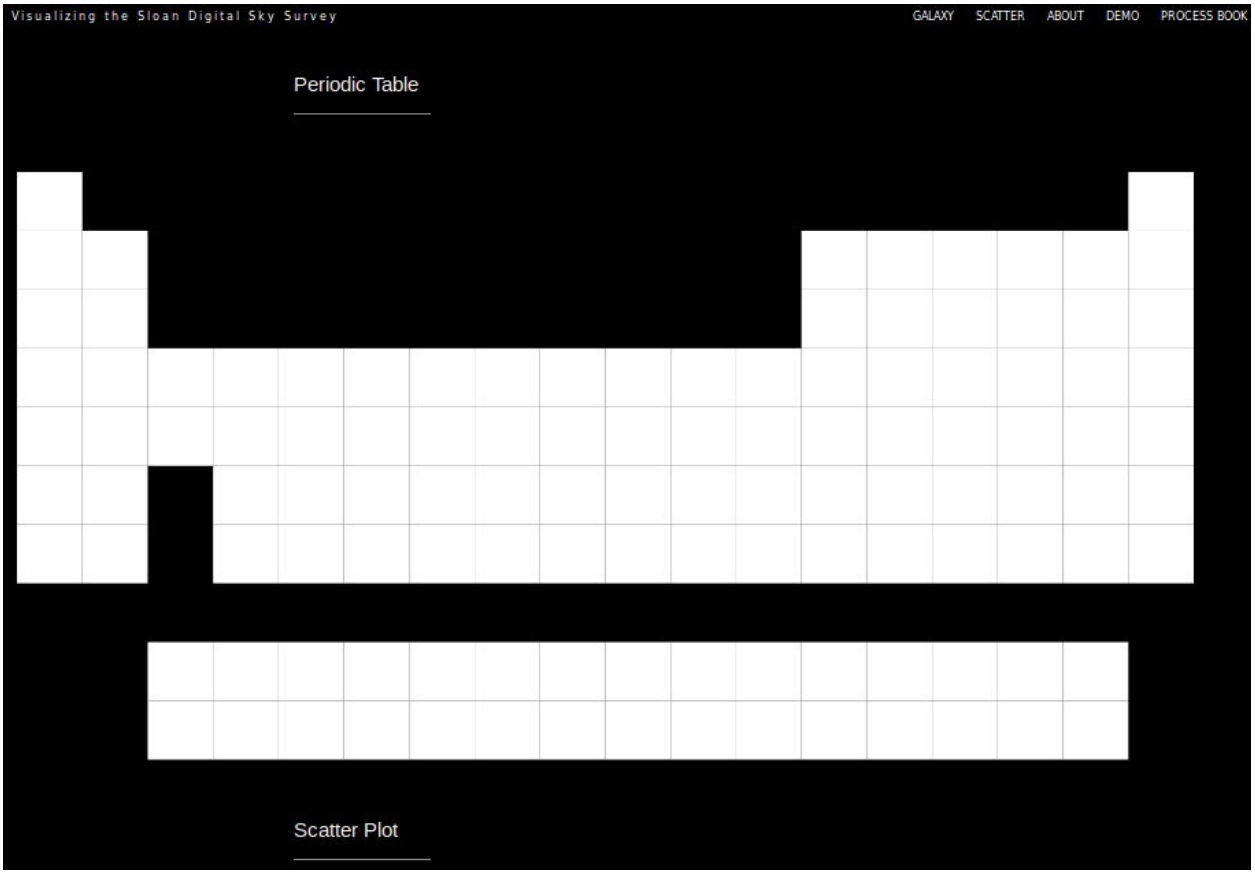
Our initial thoughts on the galactic visual included an overlay atop still figure.





Periodic Table Visual-

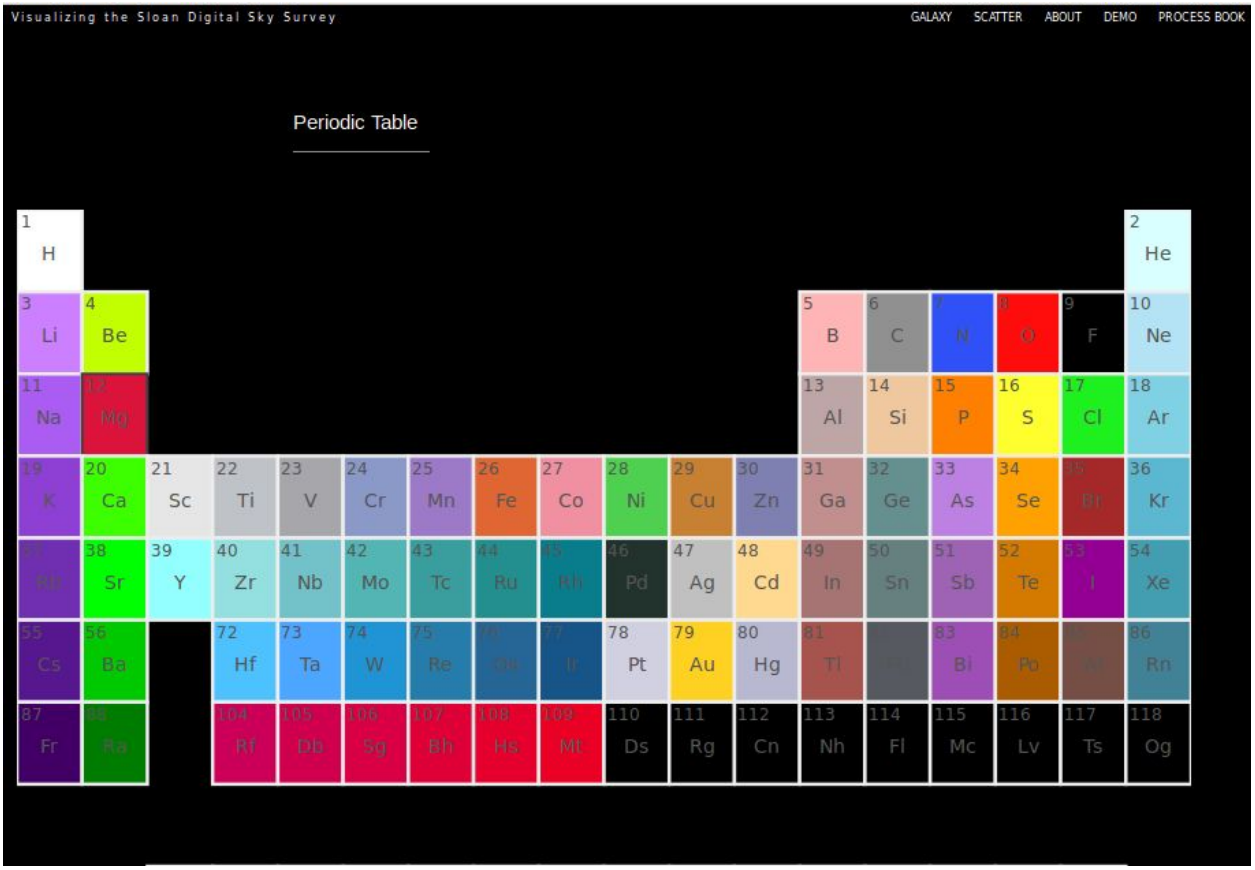
The first visualization we began working on was the periodic table that will correlate to the chemical abundance in the stars. The initial view directly below will display all the chemicals and let the user know what type of chemical they are (i.e. man made, etc.)



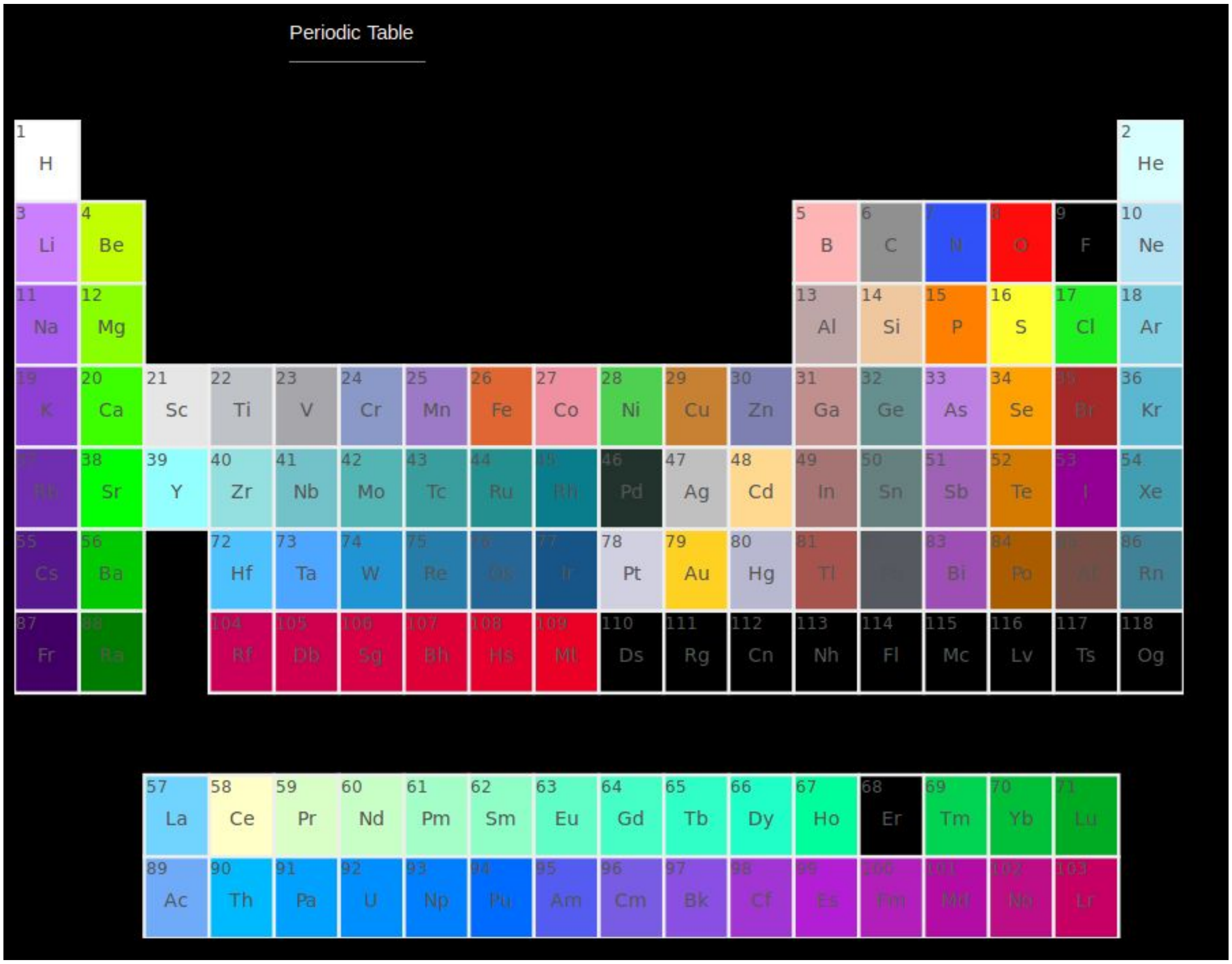
The next progression shown is a view that required our table to directly match or data cells. At this point we started thinking about the data and labeling the chemicals accordingly.

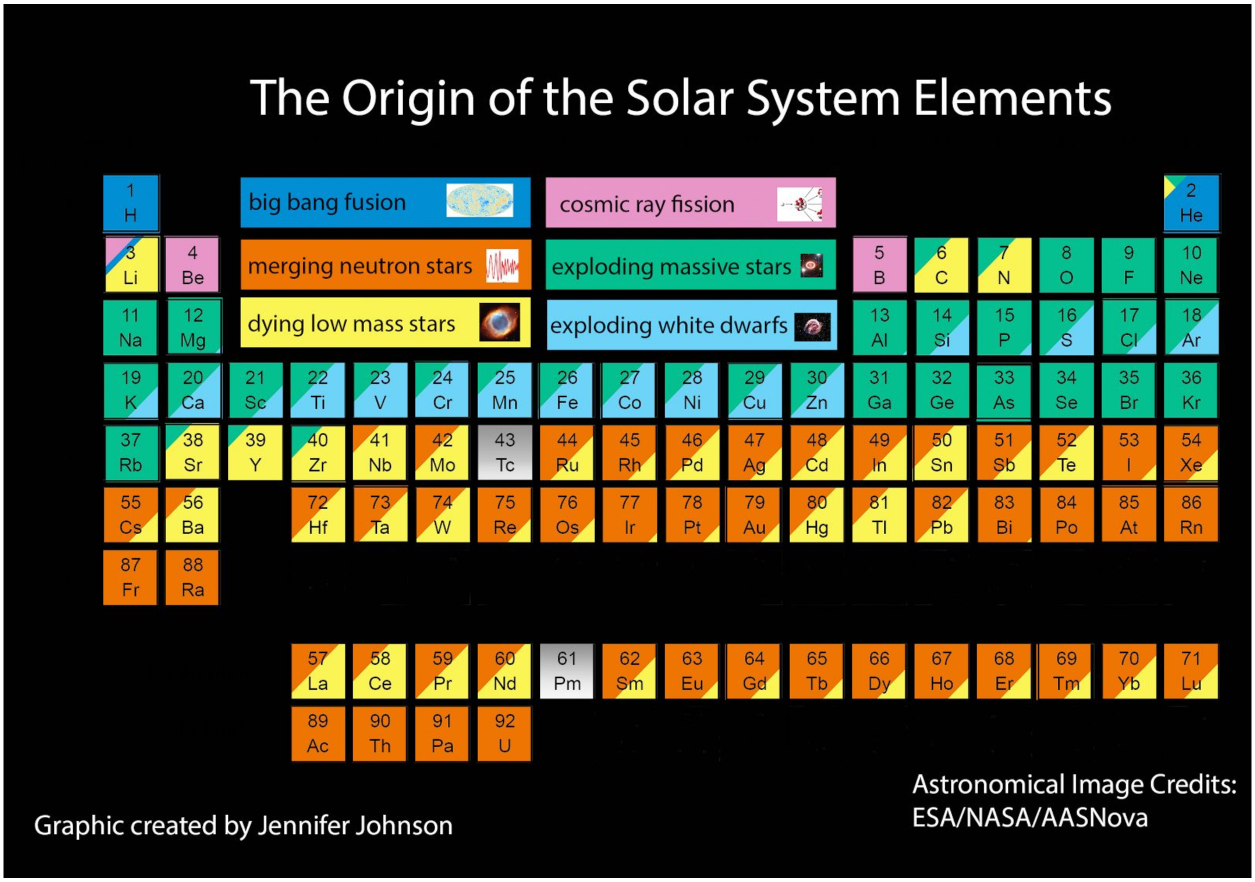

Visualizing the Sloan Digital Sky Survey
GALAXY
SCATTER
ABOUT
DEMO
H He
Li Be B C N O F Ne
Na Mg Al Si P S Cl Ar
K Ca Sc Ti V Cr Mn Fe Co Ni Cu Zn Ga Ge As Se Br Kr
Rb Sr Y Zr Nb Mo Tc Ru Rh Pd Ag Cd In Sn Sb Te I Xe
Cs Ba Hf Ta W Re Os Ir Pt Au Hg Tl Pb Bi Po At Rn
Fr Ra Rf Db Sg Bh Hs Mt Ds Rg Cn Nh Fl Mc Lv Ts Og
La Ce Pr Nd Pm Sm Eu Gd Tb Dy Ho Er Tm Yb Lu
Ac Th Pa U Np Pu Am Cm Bk Cf Es Fm Md No Lr

Color coding the table was vital in order for the user to have some visual understanding of the differences. The differences initially do not correlate with anything. But in other iterations we chose to make particular colors refer to being man made or created at the time of the big bang. Basically a color coded scale allowing the user to know what chemical is associated to which and thus they can relate that knowledge informatively when looking at the stars.

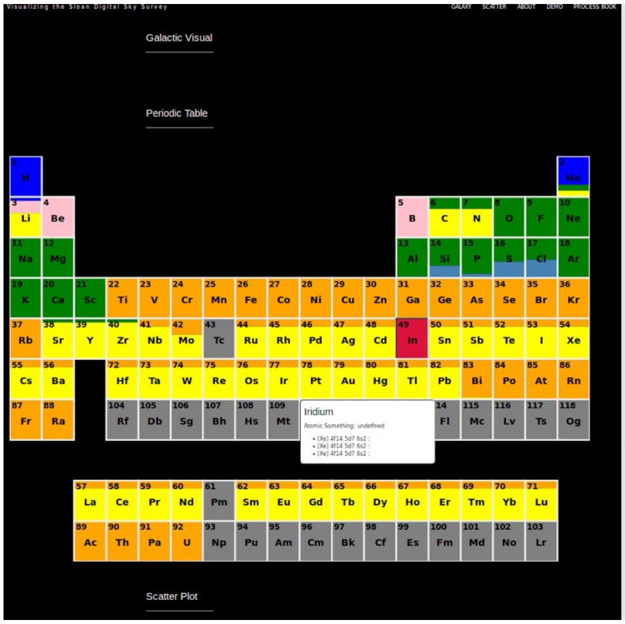


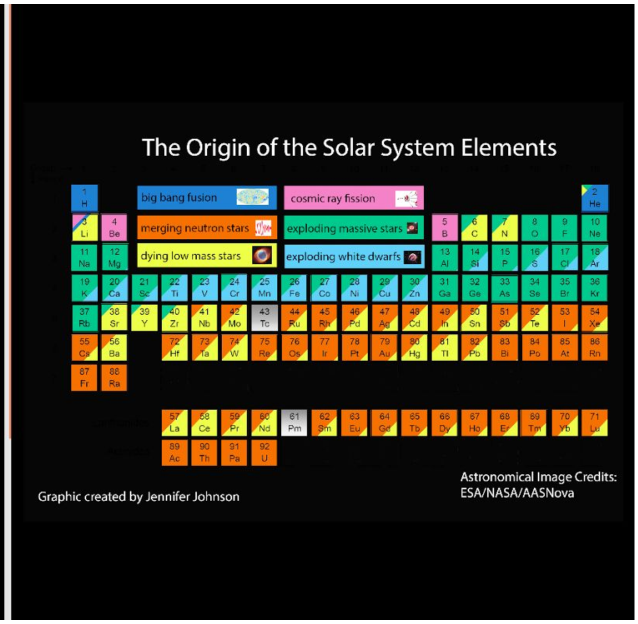
The addition of a few more chemicals was not as difficult as we originally thought. By adjusting the svg and the css we were able to and a column and row design below our initial periodic table and color code it using the same principles.



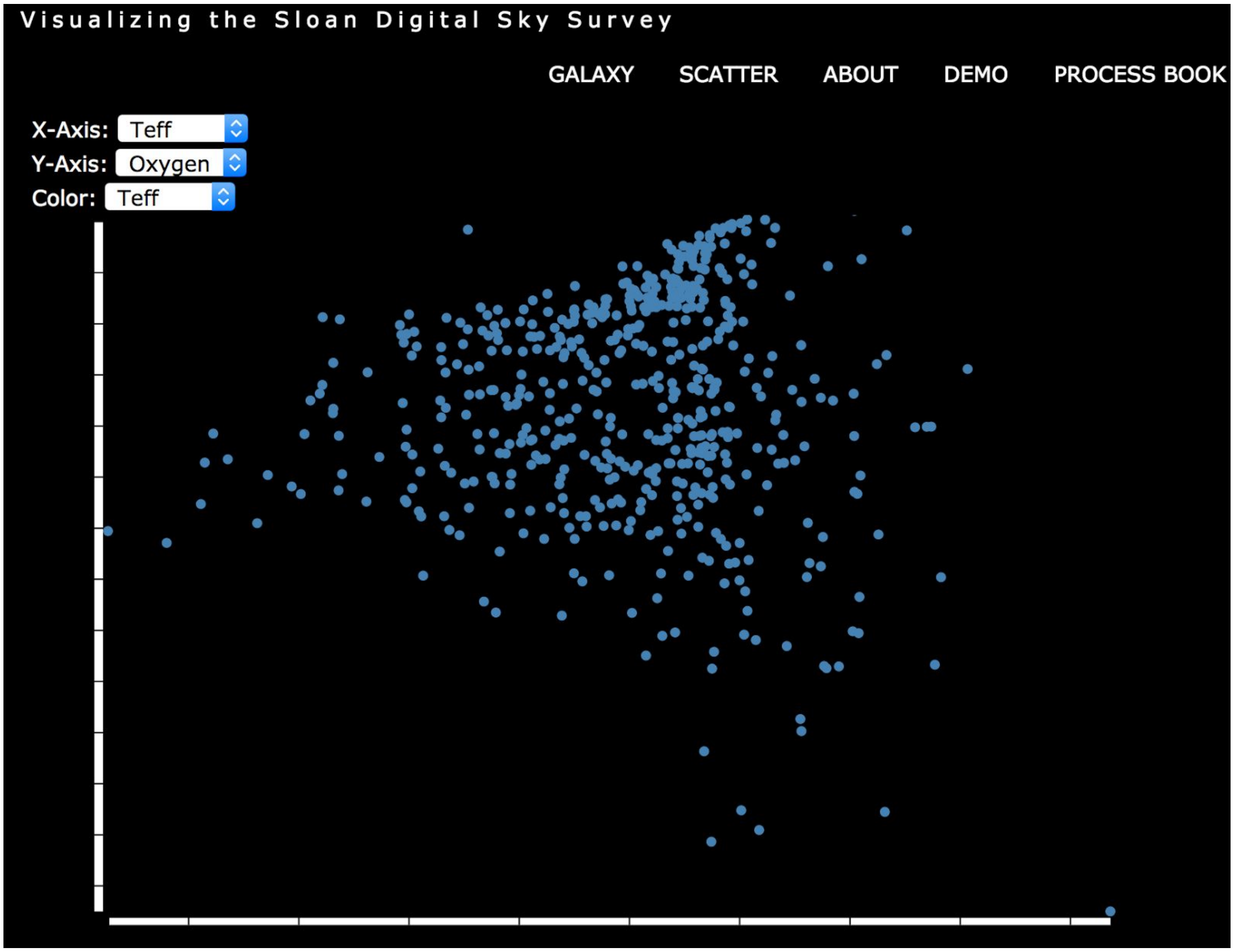

The Origin of the Solar System Elements
big bang fusion
cosmic ray fission
merging neutron stars
exploding massive stars
dying low mass stars
exploding white dwarfs
1 H
2 He
3 Li
4 Be
5 B
6 C
7 N
8 O
9 F
10 Ne
11 Na
12 Mg
13 Al
14 Si
15 P
16 S
17 Cl
18 Ar
19 K
20 Ca
21 Sc
22 Ti
23 V
24 Cr
25 Mn
26 Fe
27 Co
28 Ni
29 Cu
30 Zn
31 Ga
32 Ge
33 As
34 Se
35 Br
36 Kr
37 Rb
38 Sr
39 Y
40 Zr
41 Nb
42 Mo
43 Tc
44 Ru
45 Rh
46 Pd
47 Ag
48 Cd
49 In
50 Sn
51 Sb
52 Te
53 I
54 Xe
55 Cs
56 Ba
72 Hf
73 Ta
74 W
75 Re
76 Os
77 Ir
78 Pt
79 Au
80 Hg
81 Tl
82 Pb
83 Bi
84 Po
85 At
86 Rn
87 Fr
88 Ra
57 La
58 Ce
59 Pr
60 Nd
61 Pm
62 Sm
63 Eu
64 Gd
65 Tb
66 Dy
67 Ho
68 Er
69 Tm
70 Yb
71 Lu
89 Ac
90 Th
91 Pa
92 U
Graphic created by Jennifer Johnson
Astronomical Image Credits:
ESA/NASA/AASNova

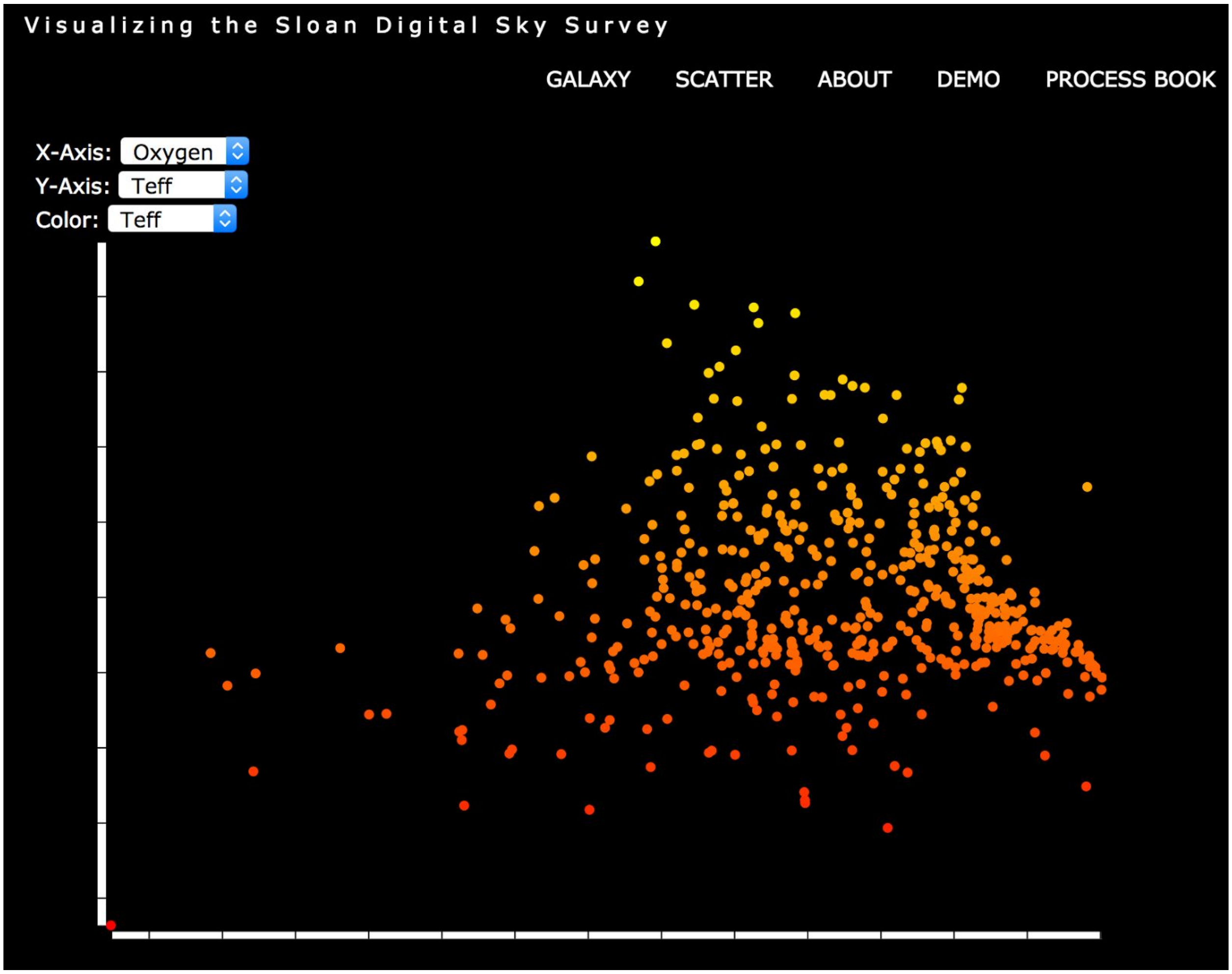
Below is a side by side representation of our merely finalized periodic table in comparison to a graphic we used as reference created by Jennifer Johnson. As you can see our coloring is a bit more custom pertaining to our specific liking and generates a tooltip as hovering over the element occurs with a mouse. We are quite pleased with what we have accomplished thus far.

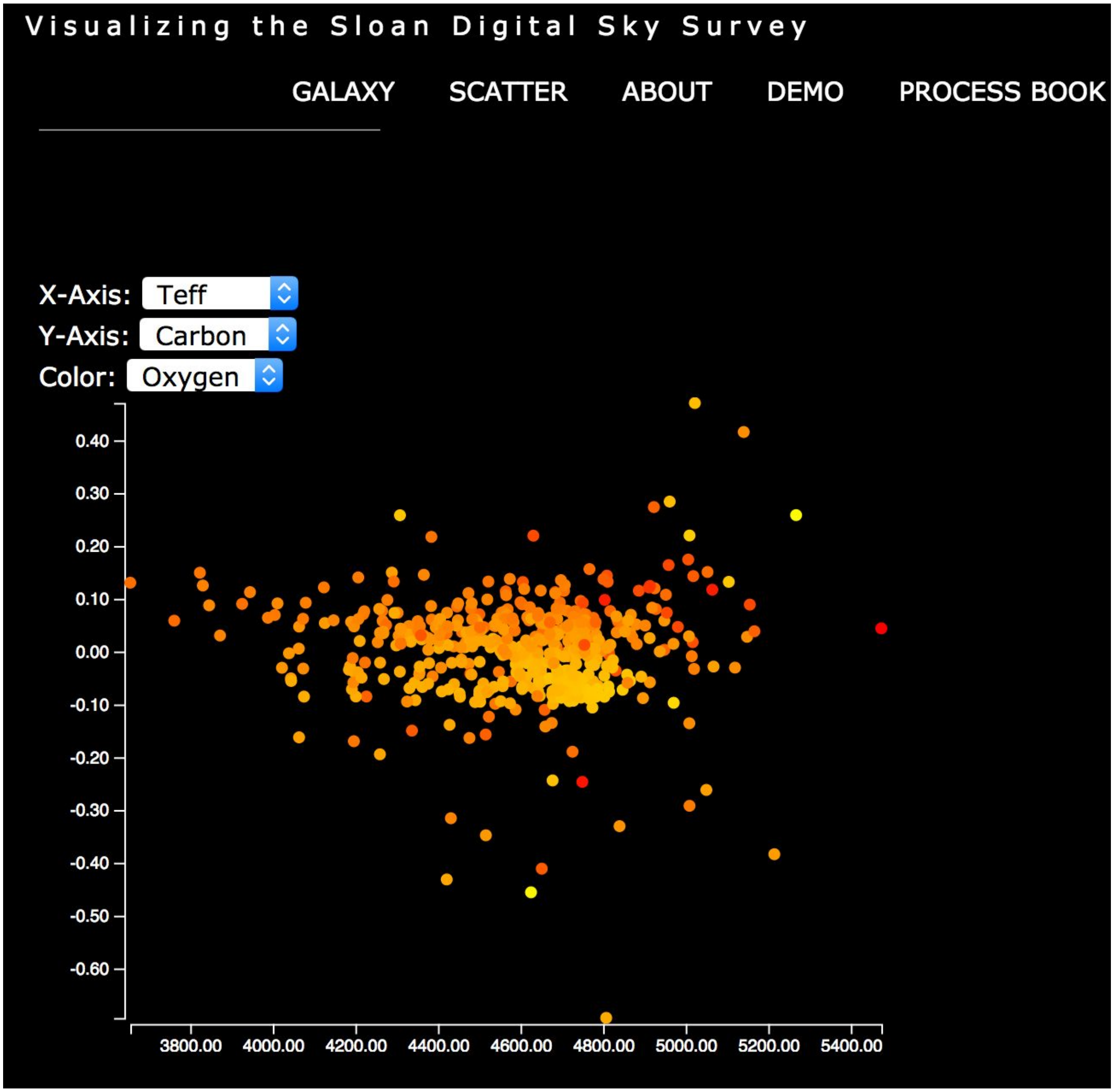


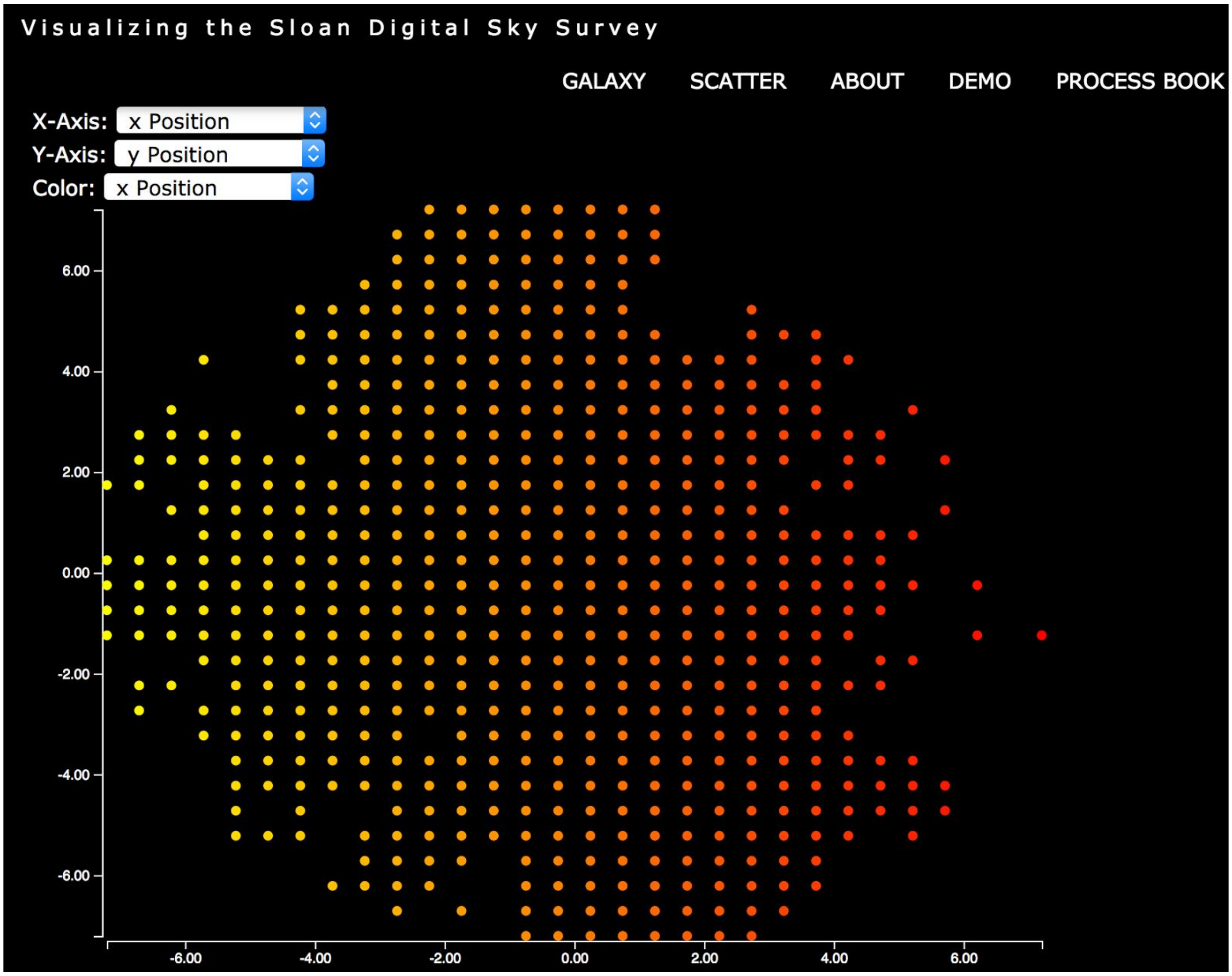


Scatter Plot Progression:


Visualizing the Sloan Digital Sky Survey
GALAXY
SCATTER
ABOUT
DEMO
PROCESS BOOK
X-Axis: Teff
Y-Axis: Oxygen
Color: Teff


Visualizing the Sloan Digital Sky Survey
GALAXY
SCATTER
ABOUT
DEMO
PROCESS BOOK
X-Axis: Oxygen
Y-Axis: Teff
Color: Teff


Visualizing the Sloan Digital Sky Survey
GALAXY
SCATTER
ABOUT
DEMO
PROCESS BOOK
X-Axis: Teff
Y-Axis: Carbon
Color: Oxygen
0.40
0.30
0.20
0.10
0.00
-0.10
-0.20
-0.30
-0.40
-0.50
-0.60
3800.00
4000.00
4200.00
4400.00
4600.00
4800.00
5000.00
5200.00
5400.00



Evaluation-

Stefan: Thus far I think Anthony and I are a great match. We have effectively split our workload and we can focus on the project. We meet bi-weekly which always lets us catch up if we need to.

Anthony: Both of us want to make our project as good as it can be and that helps tremendously. Splitting the visualizations is crucial and trusting your teammate to complete certains aspects has not been a problem.