6630 Final Project Process Book

Anthony Garcia, Stefan Kapetanovic
anthonygarcia918@hotmail.com
stefankapetanovic1@gmail.com

TABLE OF CONTENTS

1. 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

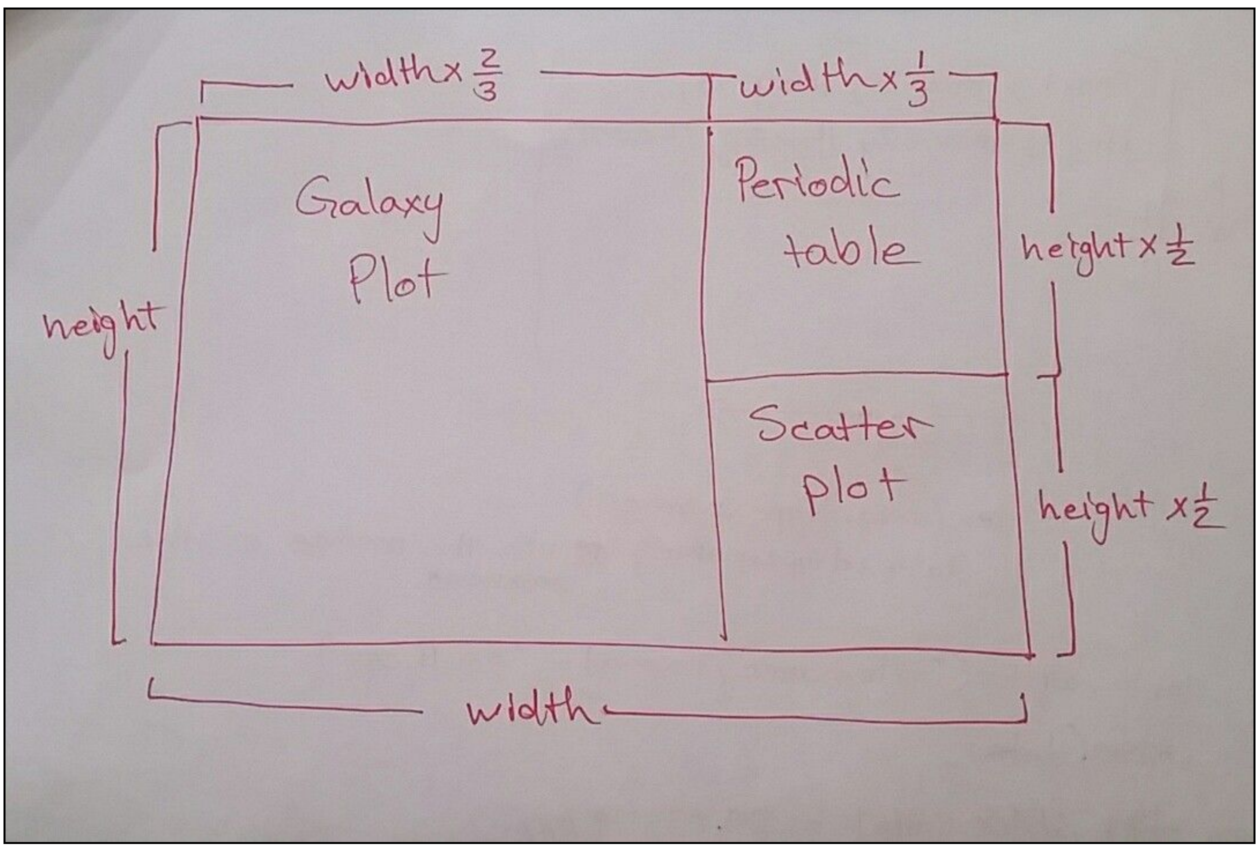
Overview and Motivation-

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 the Sun, etc. Providing unique and interactive celestial visuals with a informational backing is our primary focus in the project.

- **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. This visualization was directly sparked from interviews with Professors Gail Zasowski and Kyle Dawson. They conveyed to us that this project would be extremely useful in the professional field and we personally believe that it also interesting for beginners to view and learn about the Milky Ways' stars. 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 with was also a goal of ours as we generated our concept. We felt that a mixture of including concepts we knew we understood, with a new visualization would be a good project for the two of us to tackle.
- **Questions:** With a direct motive from astronomy professors stating that this would be a unique and informational project we decided to visually encompass the information of stars in the Milky Way Galaxy. Simply, 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 decision of how and what to display in order to satisfy visual constraints. Over the course of the semester, our project evolved to include the main components most researchers could use. Temperature, velocity, positional relations, chemical abundance (restricting the amount of chemicals given was saturated to about 25 elements), and chemical composition/correlation are all crucial components to the understanding on the MIlky Way galaxy stars. This evolution of questions answers the questions we are trying to answer in the most elegant fashion considering the data we acquired. New questions that arose during the analysis of data was how we were going to accurately display our visualizations. This was a design question that was crucial in how our website worked. We ultimately wanted a design where the user had minimal scrolling amongst all three visualizations.

Below is the beginning layout we aimed for:



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 includes: 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.

- **Data:** Our data stems from the APO Galactic Evolution Experiment (APOGEE) data set. APOGEE is using high-resolution, high signal-to-noise infrared spectroscopy to penetrate the dust that obscures significant fractions of the disk and bulge of our Galaxy. APOGEE surveys over 100,000 red giant stars across the full range of the Galactic bulge, bar, disk, and halo. Precise radial velocities, temperatures, sizes, star relation distances, and detailed chemical abundance "fingerprinting" will provide unprecedented insights into the dynamical structure and chemical history of the Galaxy. Our goal is to aggregate through this information as smoothly as possible to create informational but intriguing visualizations. Our data includes such a deep amount of information we knew we would have some data processing to do. Anthony had a great astronomical background and took this aspect on his own hands. Once we could work off the extremely accurate data we had to understand what tags and elements we could use. APOGEE data is open source data open to anyone interested in the Milky Way Galaxy so attaining the data wasn't too difficult, but organizing the data into a .csv excel spreadsheet allowed us to visually see the data and then split the aspects of the data that each of us was to focus on in the specific visualizations we assigned to each other. We had our professor advise the data capturing aspect in order to be positive our data did not lead us down a wrong path. When meeting with her we concluded we were ready to work on the APOGEE data quite quickly because it was very well constructed. Our only objective was to cluster, process, and average the star data to include within our visualizations.
- **Exploratory Data Analysis:** When it came to analyzing the data we were working with we had a good understanding of what we could do. Using ".csv" files to encompass our data we visually understood what tags we could use and what information could provide insights to our user. We knew that chemical abundances was a vital part we needed to included in our visuals. To do this we decided to deploy

the tile skills we acquired from our homework. Our scatter plot was also something we knew we could do but didn't want that to be the entire scope of our project. We agreed we wanted to work on something that we both had not experienced before. To do this we decided on making an overlay galaxy visualization that could encompass all of our stars. This process was not difficult and did not take us much time to deliberate. We decided that the galactic visual and periodic table were two visualizations that we wanted to include in our project as "Must-Haves." We also included the scatter plot as an "Optional" visualization we hoped to reach. By having a firm grasp of what our data represented we knew exactly how were were going to approach creating our visualizations. This took a lot of stress off of us in the design process initially. The difficulty would come directly from generating intriguing and informational visualizations. Designing the structure and layout of the visuals came into question later based off a user experience predicament rather than a data limitation.

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.

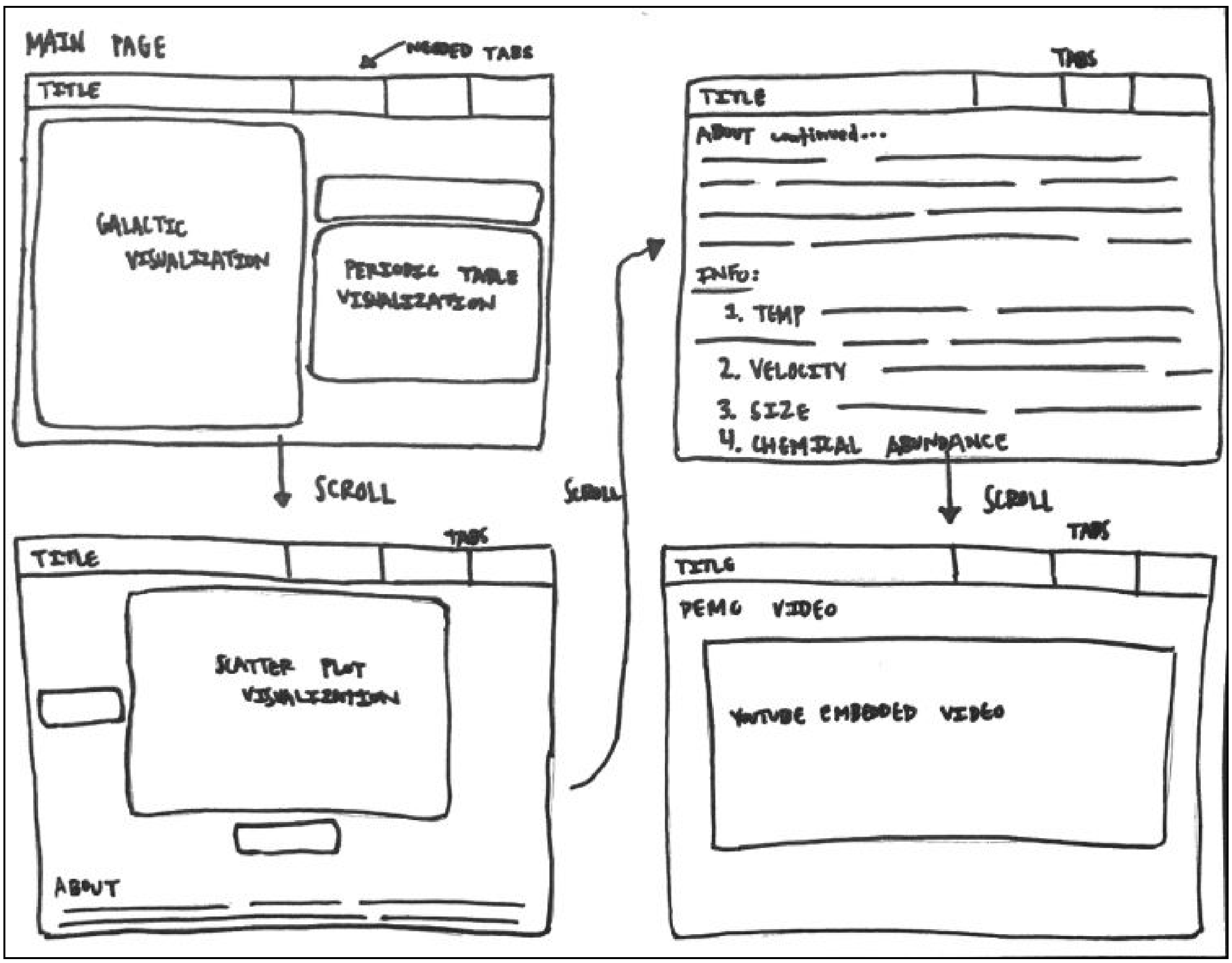
atomicNumber	symbol	Name	atomicMass	cpkHexColor	electronicConfiguration	electronegativity	atomicRadius	ionRadius	vanDelWa
1	H	Hydrogen	1.00794(4)	FFFFFF	1s1	2.2	37	120	120
2	He	Helium	4.002602(2)	D9FFFF	1s2	32	32	140	140
3	Li	Lithium	6.941(2)	CC80FF	[He] 2s1	0.98	134	76 (+1)	182
4	Be	Beryllium	9.012182(3)	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

Data Sources:
https://github.com/andrejewski/periodic-table
https://data.sdss.org/sas/dr14/apogee/spectro/redux/r8/stars/l31c/l31c.2/allStar-l31c.2.fits
https://data.sdss.org/datamodel/files/APOGEE_REDUX/APRED_VERS/APSTAR_VERS/ASPCAP_VERS/RESULTS_VERS/allStar.html
https://www.dropbox.com/s/jn9e5waryvr2f6f/apogee_distances-DR14_2017-04-26.fits?dl=0

2. Visualization Design Drafts-

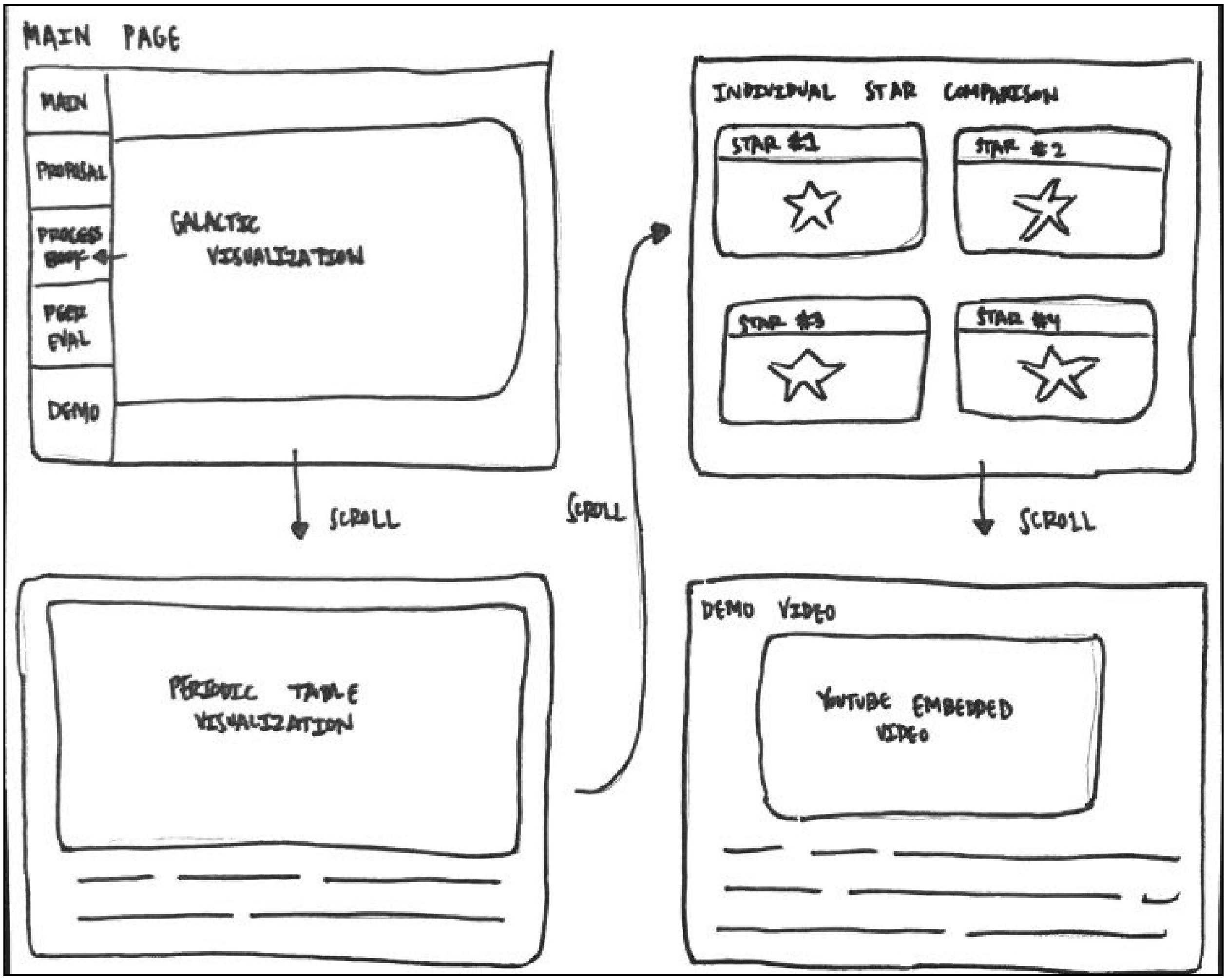
An overview of the designs are given below. Each design gives it pros and cons while explaining what is included into the final design. We include the layout of our visualizations but not the visualizations themselves because they will be constructed in the coming weeks. The general outlines of the three 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:

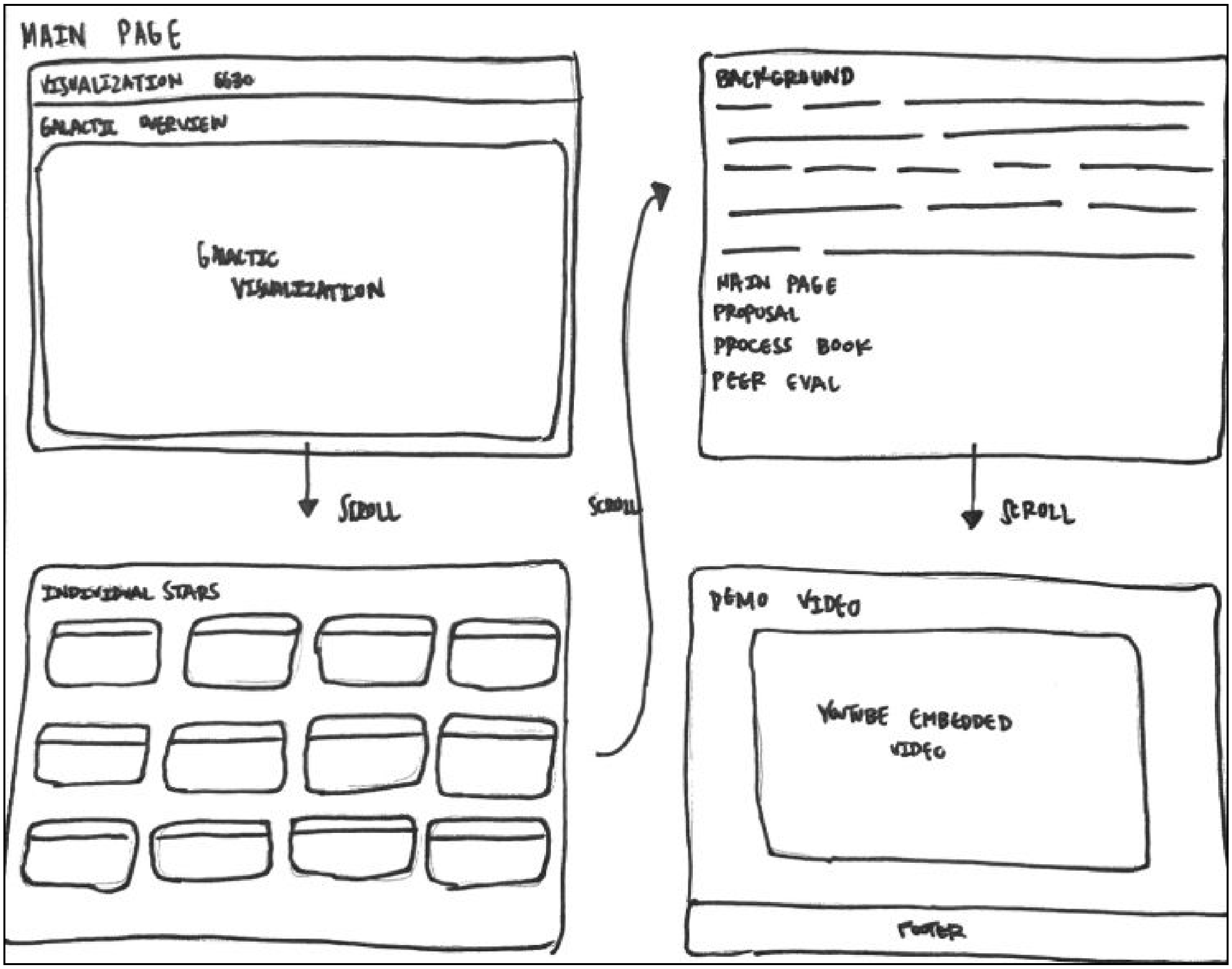


Our second design was interesting because it contained a collapsable side view that would direct you to the sections you wanted to view. We also had an interesting idea of star comparison but ultimately determined that we could get more interactivity and visual intrigue in a different way. This design 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.

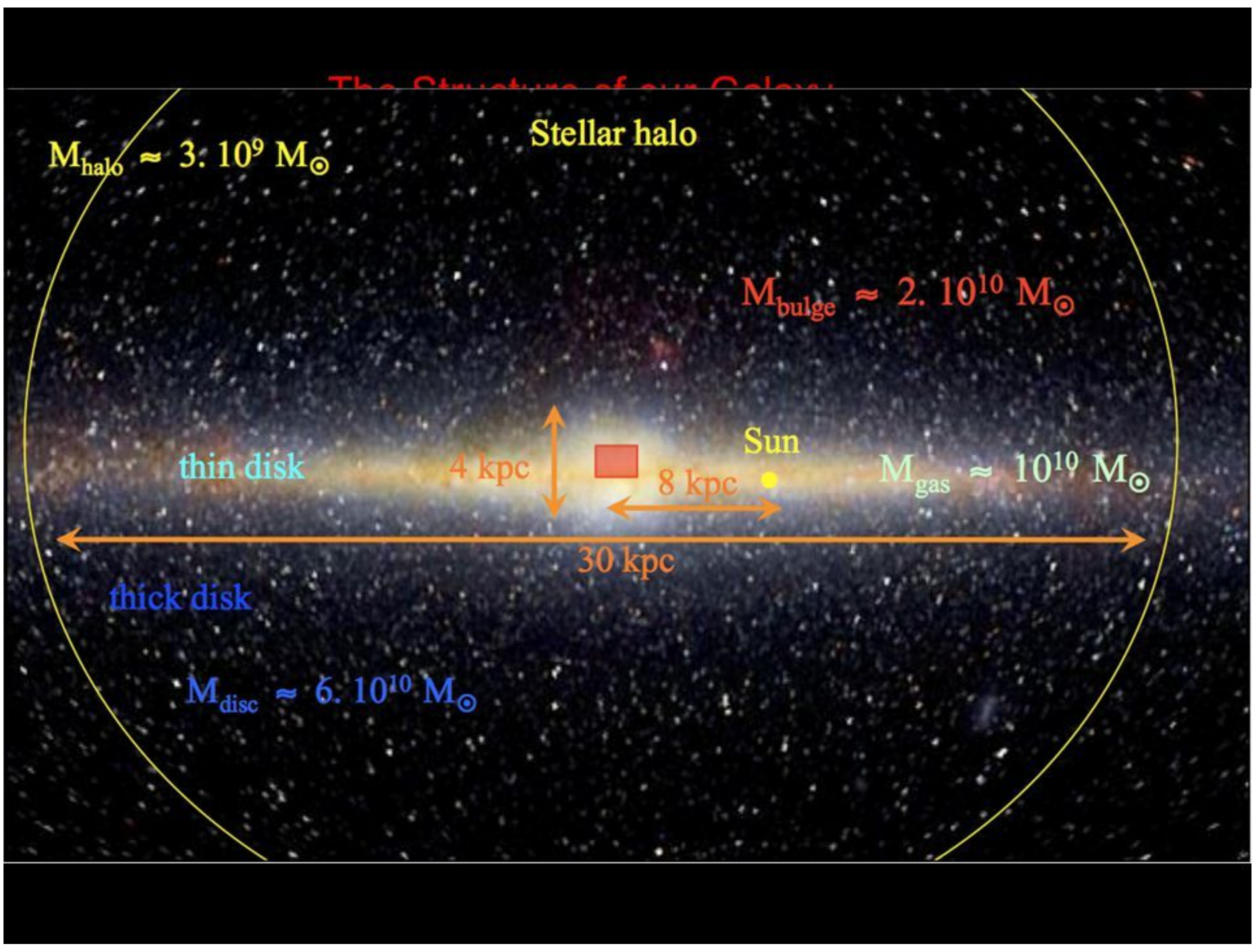
Afterwards, when grouping together we talked about our feedback and relayed to one another that our project is solid but fine tuning the details will elevate it to where we want it to be. We loved the feedback we got from Srija and thought it was very fair and more importantly very helpful. We know what 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.

3. 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.

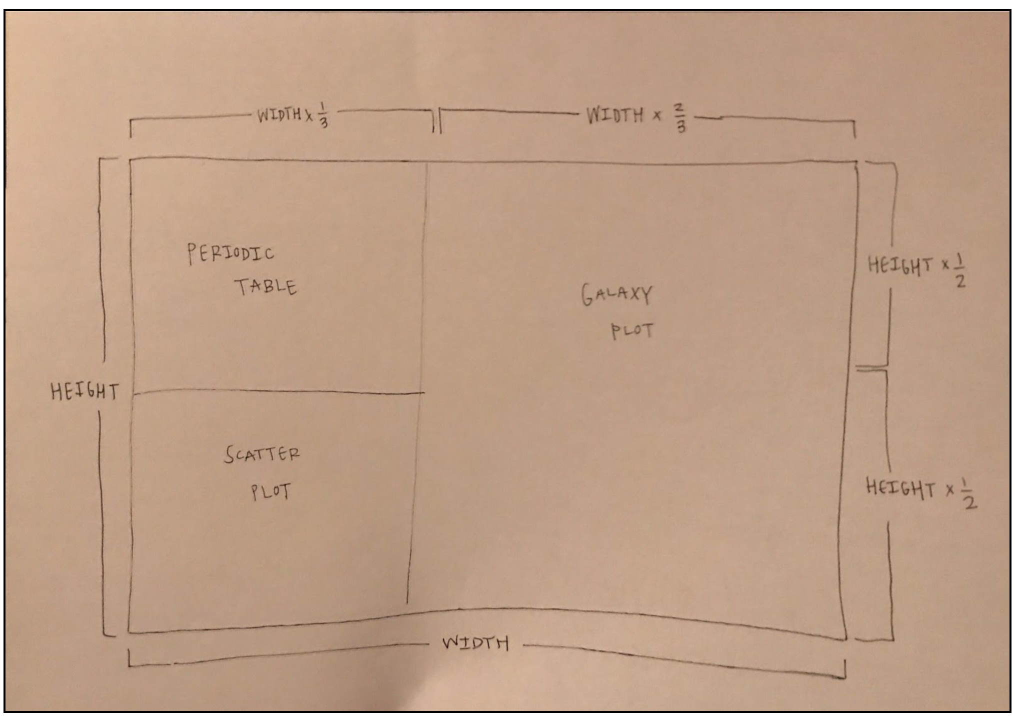
Must-Have Features-

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

Our final design ended up being delegated in a different manner. To eliminate scrolling as much and to generate visualizations that could be interactive and viewed on the same page.
We made sure to include all three of our visualizations in the following format:



The visualizations are described in detail below.

Galactic Visual-

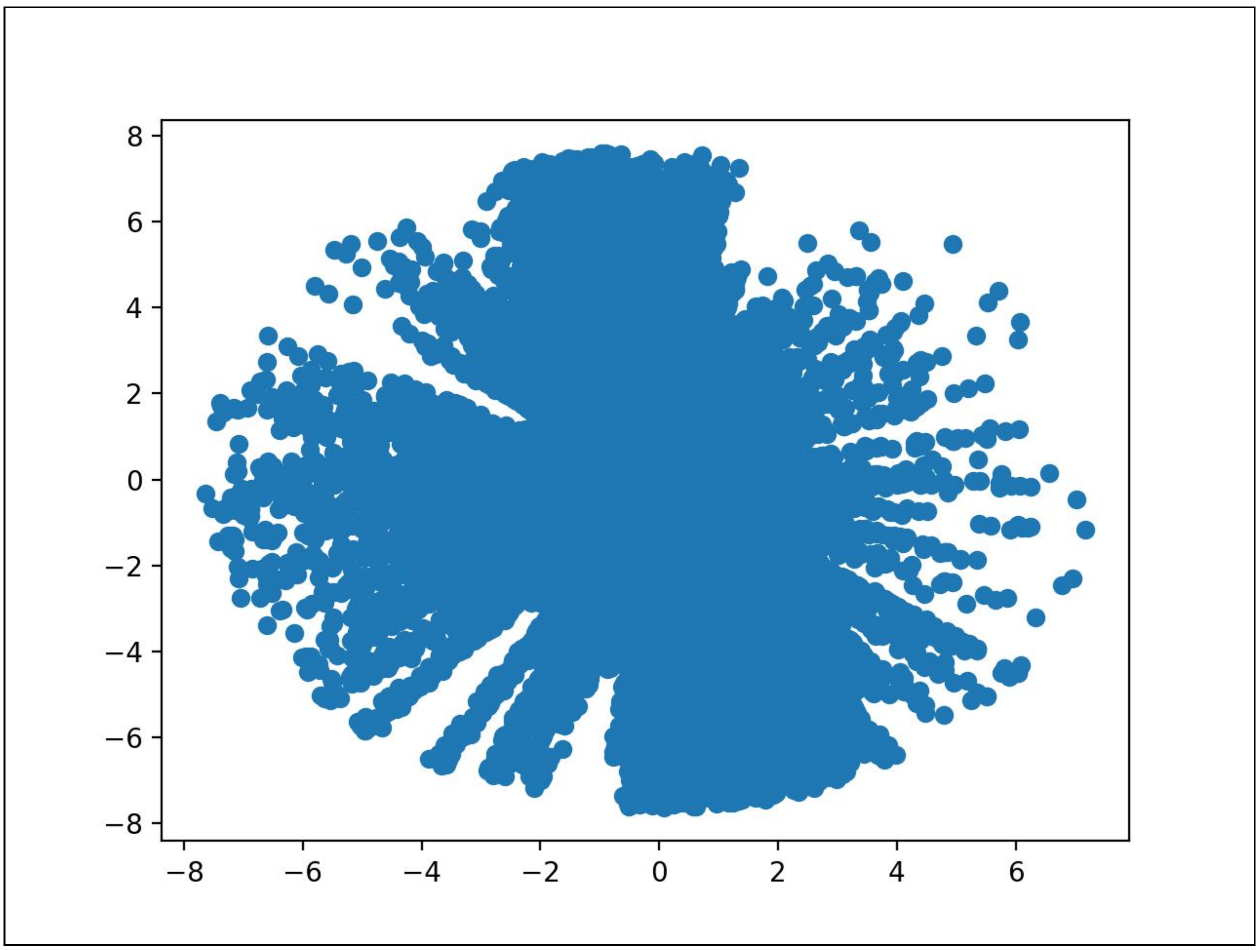
Our initial thoughts on the galactic visual included an overlay atop a still figure. The goal of this visualization is to interact directly with the scatter plot. In brief the scatter plot will display all of the stars in the range of our Milky Way Galaxy and organize the stars (each represented by a single dot) depending on the user's axis adjustment. We let the user select the X and Y axis to see correlations of what they desire. Once they do this they have the ability to use a brush to select another range of stars in the galaxy. This selects the stars, highlights them in the scatter plot but also shows there positional relation on the Galactic Visual. This lets the user see the accurate spatial relation of the stars they chose and correlated using specific terms. The Sloan Digital Sky Survey has created the most detailed three-dimensional maps of the Universe ever made, with deep multi-color images of one third of the sky, and spectra for more than three million astronomical objects. Learn and explore all phases and surveys—past, present, and future—of the SDSS.

Below is our starting Milky Way Galaxy image that we would use as a baseline for our first visual.

Upon gathering the data and analyzing the plots of critical data points for our project we saw numerous issues we needed to fix in order for the data to be useful at all.

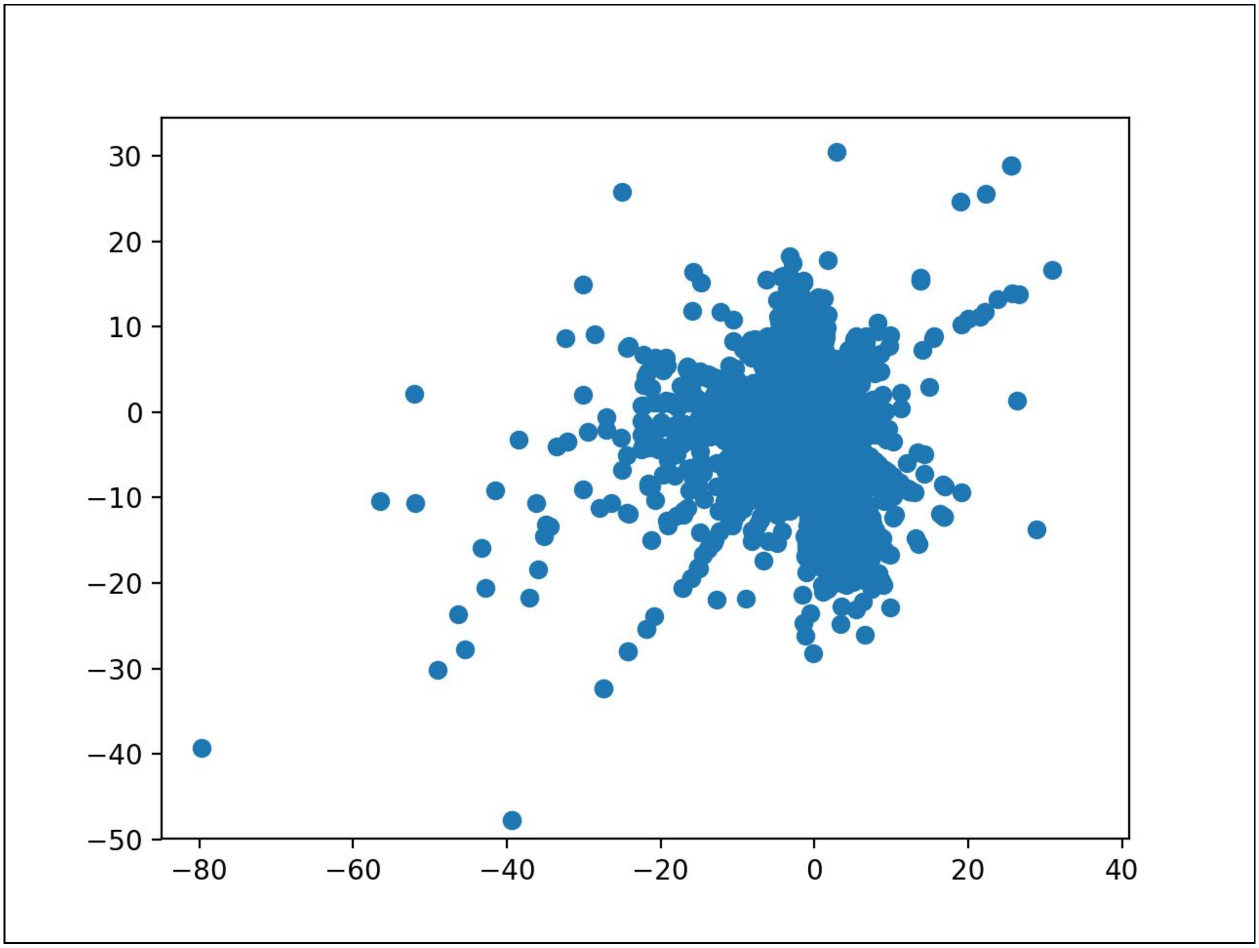
Some of the issues with the data collected below include but are not limited to:

- Overlapping data points
- Loading data took time and could need async tasks
- A spectrum that was out of our range in the Milky Way Galaxy
- Connecting this data to two other visuals without processing would cause an issue in formatting and scale
- Indistinguishable number of points



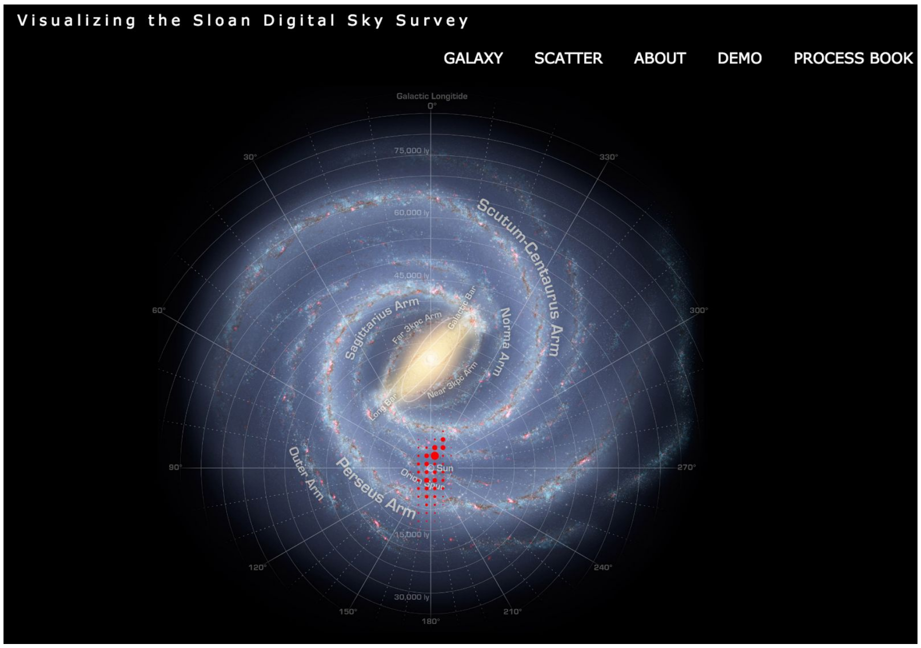
With this baseline of data, we decided to overhaul some major design concepts to improve our overall design. First, we decided to tackle the issue of the amount of data we gave the machine to avoid having to learn async tasks in JS. We did this by clustering and averaging our star data into individual cells of a grib. This would give us the approximate data of the stars within a close proximity while drastically improving our odds of decreasing our data plots to make the visuals comprehensible. While we did not deviate from our proposal we started to realize to make a visualization that was capable of helping a wide scope of people we needed to implement our data in a structured alternative. This would not only help the visualization but the functionality of the final product through a more intuitive and perceptive design. Our final galactic design was very similar to our original notion of what we wanted. The data we processed was the only issue we encountered and did not plan for. This step was crucial in generating a template for the galaxy interaction we planned to provide for our user.

Our first iteration back from the original data sample returned to look like the image below:

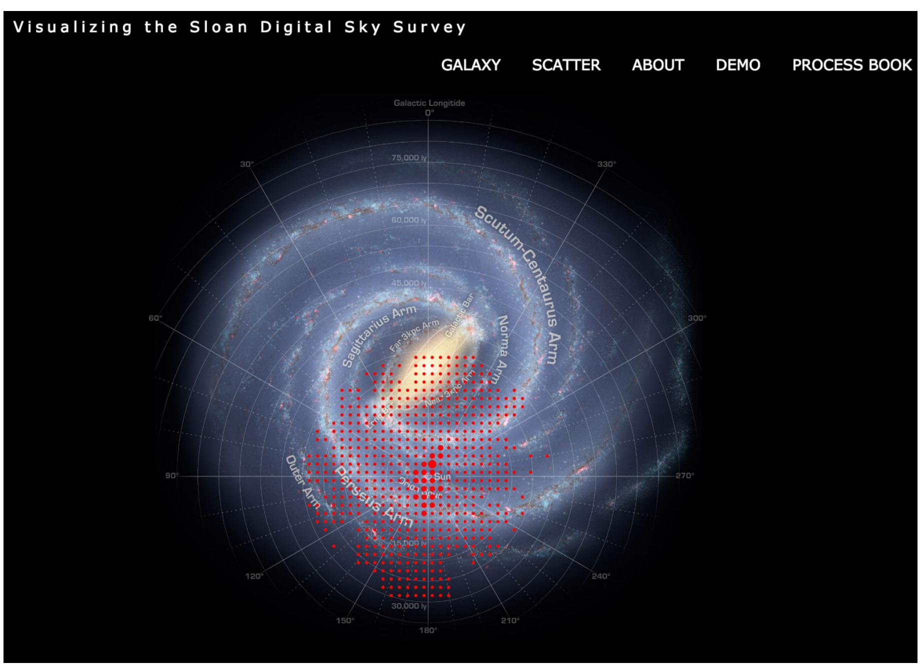


From this iteration we concluded that the data averaging had done a great deal in proving our project to render and compile faster. This took out our worry for having to learn async tasks. We still had a few concerns with overlapping and the radius of stars we were going to display on the galactic visual. Next we had to cluster the data in the exact regions they needed to be placed according to its positional relation to the sun in the Milky Way Galaxy. We encountered a design decision at this point where we had to understand how we wanted the scatter plot and visualizations to interact. We had two potential options that would help differentiate the selected stars. If you recall, the idea is to allow the periodic table to choose the color of the galaxy plot while the scatter plot would allow the user to find correlations they wanted and to select particular stars. These events all needed to occur simultaneously, in real time without delay, with animation, and visually intriguing. Using our skillset from the course we knew that since all of our data points were going to be displayed as circles the best differentiating feature we could use was color, fill, and size while encapsulating all of our ideas.

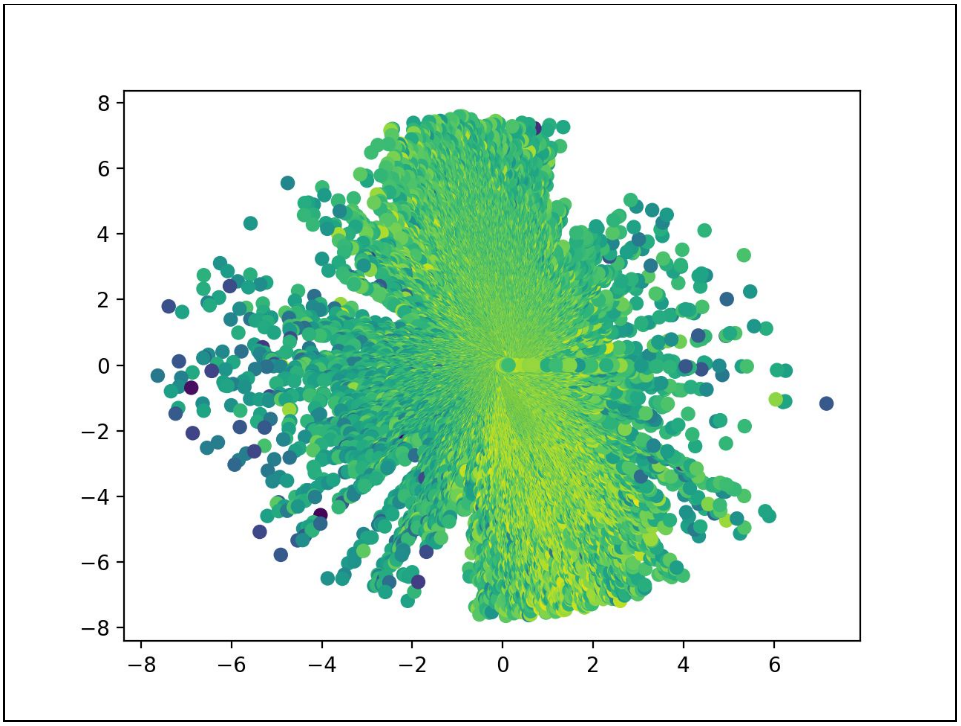
The two strategies considered are below:



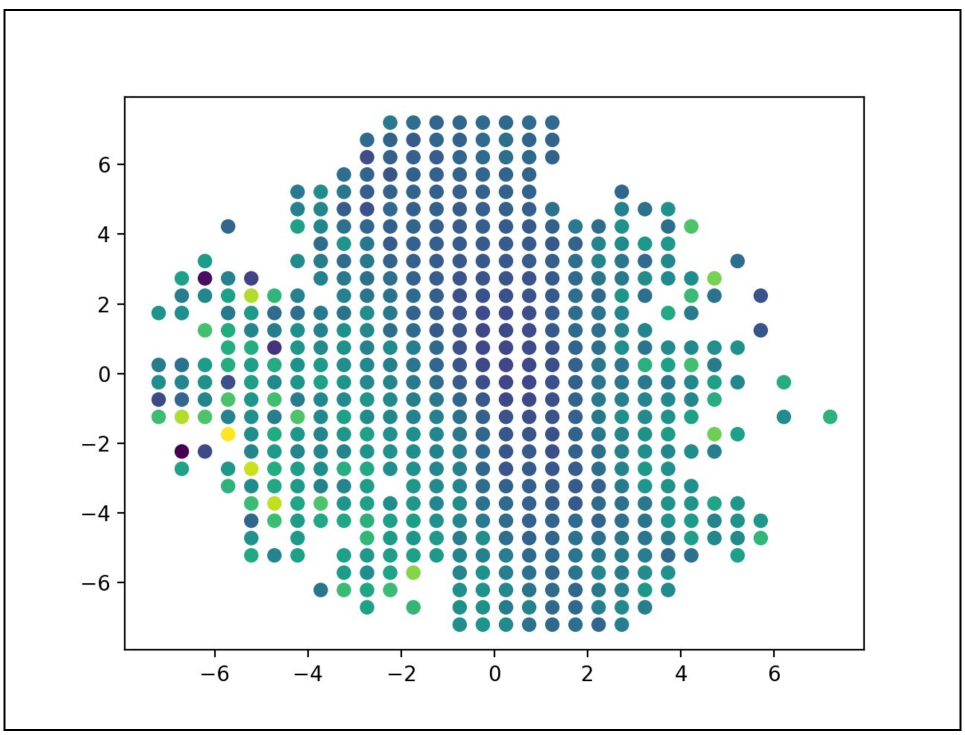
This first design choice didn't include the entire data we gathered but rather highlighted the specific stars chosen in the scatter plot visualization. We decided against not showing all the stars due to a possible confusion on the user's end. The visualization wasn't easily visible upon important change.



This new strategy shows all of our data points on the screen to begin. Upon the user interacting with the periodic table or the scatter plot the galactic visual will change in color or selection immediately.



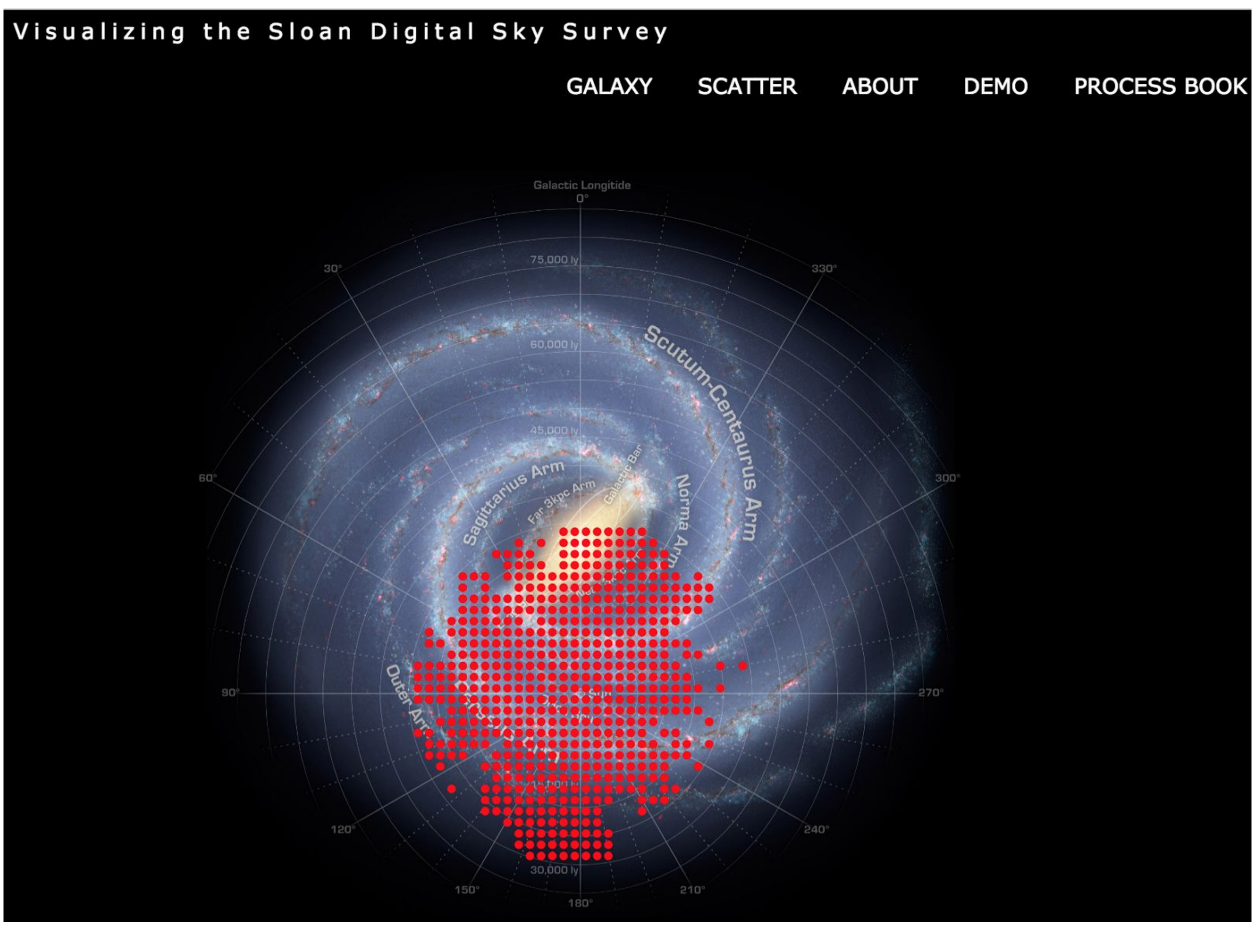
In the two examples on this page we wanted to show the relation of clustering/averaging and color to our original data sample. In the first image we see that color is separating data points but to an extent due to the amount of data. This period in our design evolution caused our design to change into a more organized grid.

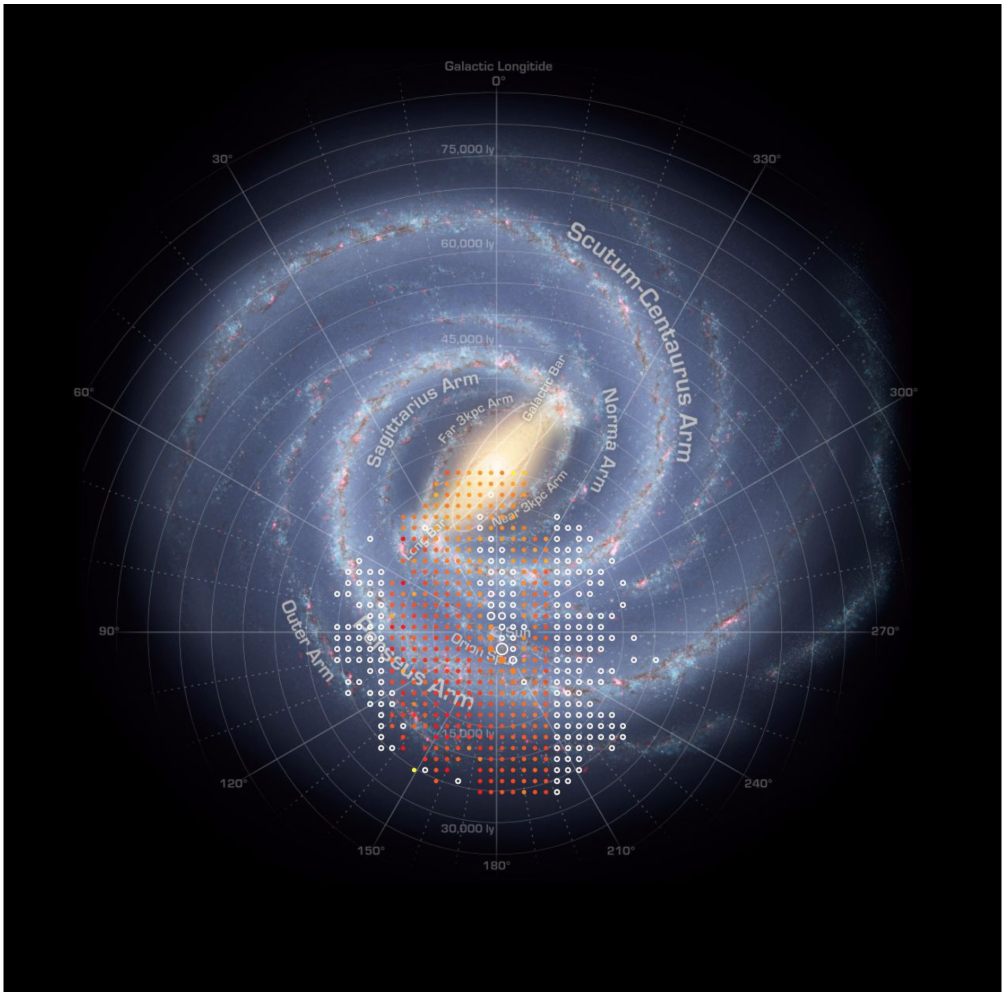


Our data now represented its true form. It was easily distinguishable while also employing color to separate chemical abundance depending on the color scheme and chemical chosen by the user.

This image shows our Galactic Visual nearly complete. As you can see the Milky Way Galaxy is the background image with an accurate spatial overlay of our clustered data of stars. With over 250,000 data points each with a star name, chemical abundance sheets, velocities, temperatures, distances in relation to other stars, if they have planets orbiting them as well as their sizes we had to cluster, average, and generate data points of all that data into one sun representing the group. Our points need minor resizing and the ability to interact with the scatter plot brush and the the periodic table chemical selection. But in terms of clarity and interpretation this design is progressively better than what our original data conveyed.

The image below shows all the star data points in the Milky Way Galaxy. While positionally accurate there is not interaction with the scatter plot and we had a design preference to make the radius of the points smaller in the next iteration. From that stage we want to have a tooltip show the star information as the user hovers over the clustered grid element. The data given will be averaged star data with the intent of showing correlations amongst whatever specific variables (temperature, position in relation to the sun, chemical abundance, etc.) the user chooses:



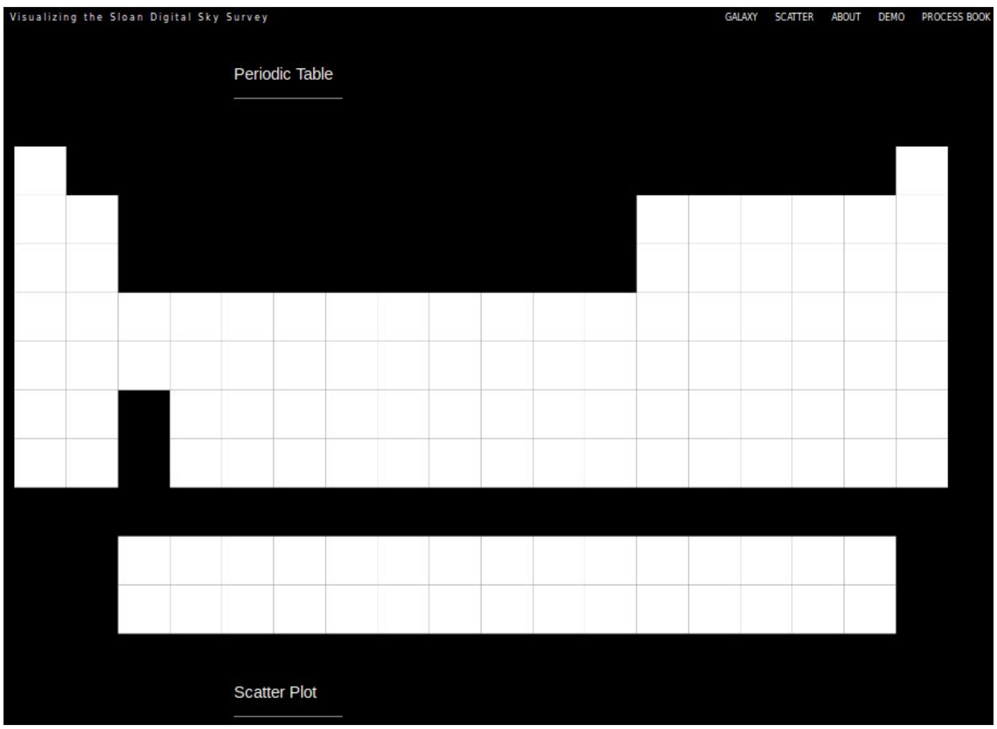


Our final design of the galactic visual came to fruition exactly how we wanted it to. We have our main star differentiators being color, fill, and size. The stars selected in the scatter plot are accurately displayed in the accurate positions. Our design evolved from learning how to interpret ouf data to making sure it was visually appealing. We did not want to compromise our need for aesthetics and give up the informational aspect our project could provide. As you can see we went through several stages in our design evolution of this particular visualization._At this stage our galactic visual encompasses all of the following:

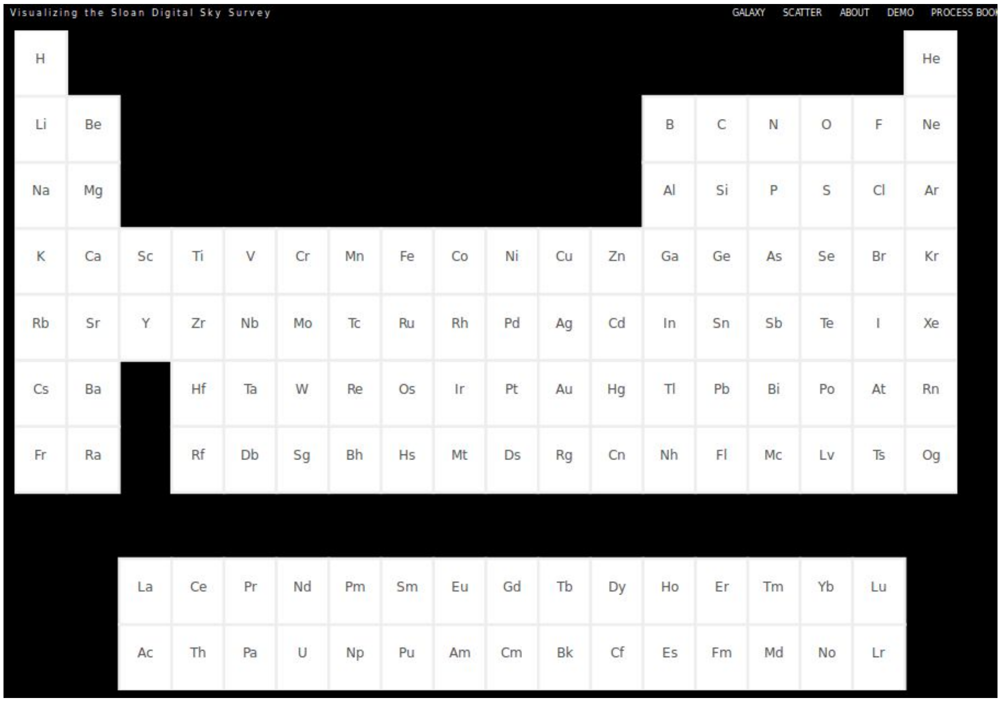
- Interactivity with the other visualizations (periodic table affects colors while the scatter plot affects selection with a brush)
- Interactivity with the user through a tool-tip hovering over star regions
- Color, fill, and size designation
- Fully clustered/averaged data
- Aesthetically pleasing visually that are intuitive and informational

Periodic Table Visual-

This visualization is vital to conveying that chemical abundance is stars is a attribute we considered in the data. This fact is vital to researches in the subject. The periodic table is very informative and pulls data directly from an our APOGEE online data. The goal of this visualization was to give an overview description for the stars. Understanding the composition of the elements and when they arose cosmically either in the big bang fusion, cosmic ray fission, merging neutron stars, exploding massive stars, dying low mass stars, and exploding white dwarfs allows the user to develop mental notes for particular chemicals they would like to look into further. The next general progression is to input the chemicals they're interested into the scatter plot axes to view correlations in position, temperature, averager metallicity, size, and velocity in the stars.



Our first iteration of the periodic table visualization is directly implemented from our knowledge in creating a tile chart in HW6. With that knowledge in hand we we quickly able to implement a periodic table that would serve its functionality exactly how we wanted. This visualization also was conceived and generated exactly how we intended it. As you can see this version only contains the sectioning partitions with no relevant information within it. As we progressed throughout the design we scaled, positioned and altered the design minimally throughout the design evolution. The hardest aspects to this visual were the color scaling and getting it to interact with the galactic visual properly.



As this visual versioned from infancy to final product we had some fun formatting our page. Since this visual included skills that we had already deployed in previous assignments its evolution was steady and consistent giving us extra time to make sure we implemented everything we wanted in this visual. At this stage, the periodic table was a stand alone SVg but slowly it developed a larger purpose. We included details like when the elements in the table were created.

As mentioned earlier, all of the elements were created from one of the following aspects:

- Big bang fusion
- Cosmic ray fission
- Merging neutron stars
- Exploding massive stars
- Dying low mass stars
- Exploding white dwarfs

We determined the best way to visually show this was to use color as our main differentiation. The colors of the tooltip interact with the colors on the periodic table. For example, if a chemical was mostly generated from a dying low mass star then the area within the periodic table representing the chemical will be full. The area consumed by a color directly responds to an event in which is was made.

We wanted to show this visually, our next challenge was making sure users understood the color and change references. Through our design evolution we debated whether this concept of colored area to the element region was enough to convey our notion of that chemical arose from the identical color action it corresponded to in the tool-tip. We decided the best way to do this was to create an “About” section below the visuals that made sure to incorporate all the information about all three of the visuals in order to use them to the best of their abilities. This is below the three primary visualizations and serves to bestow and misconceptions or ambiguity for users of the site. We believed this would be necessary for both the advanced viewers with a background in the data/Milky Way Galaxy as well as novice learners.

Below is the “About” section in our design. Although it is not necessary we decided to include this for those wanting to know the full ability of each visualization if it is not conveyed initially for whatever reason:

About

Galactic

Our initial thoughts on the galactic visual included an overlay atop a still figure. The goal of this visualization is to interact directly with the scatter plot. In brief the scatter plot will display all of the stars in the range of our Milky Way Galaxy and organize the stars (each represented by a single dot) depending on the user's axis adjustment. We let the user select the X and Y axis to see correlations of what they desire. Once they do this they have the ability to use a brush to select another range of stars in the galaxy. This selects the stars, highlights them in the scatter plot but also shows there positional relation on the Galactic Visual. This lets the user see the accurate spatial relation of the stars they chose and correlated using specific terms. The Sloan Digital Sky Survey has created the most detailed three-dimensional maps of the Universe ever made, with deep multi-color images of one third of the sky, and spectra for more than three million astronomical objects. Learn and explore all phases and surveys—past, present, and future—of the SDSS.

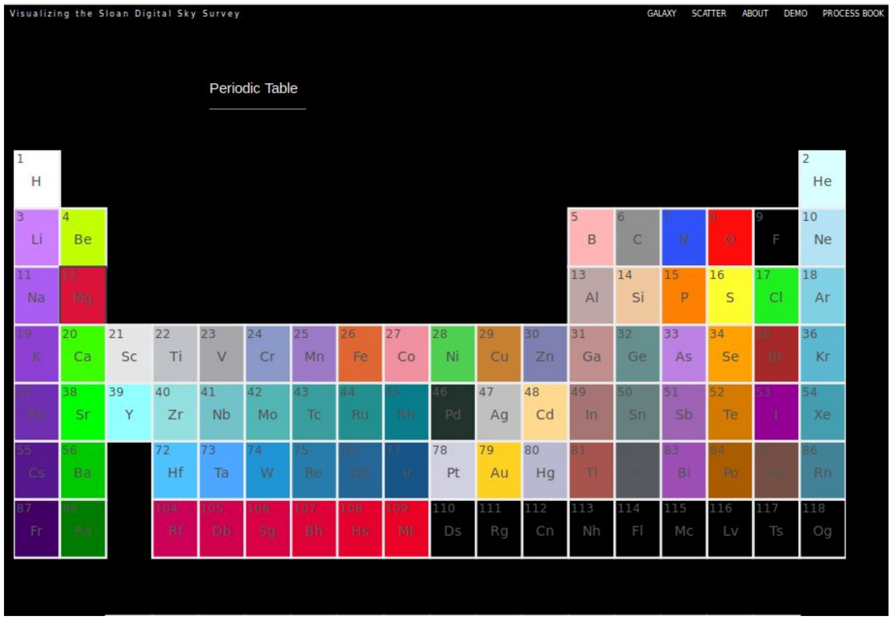
Periodic Table

The first visualization we began working on was the periodic table that will correlate to the chemical abundance in the stars. The periodic table is very informative and pulls data directly from an our APOGEE online data. The goal of this visualization was to give an overview description for the stars. Understanding the composition of the elements and when they arose cosmically either in the big bang fusion, cosmic ray fission, merging neutron stars, exploding massive stars, dying low mass stars, and exploding white dwarfs allows th user to develop mental notes for particular chemicals they would like to look into further. The next general progression is to input the chemicals they're interested into the scatter plot axes to view correlations in position, temperature, averager metallicity, size, and velocity in the stars. For this particular visualization we used data from https://github.com/andrejewski/periodic-table to make our perioudic table as professional as possible.

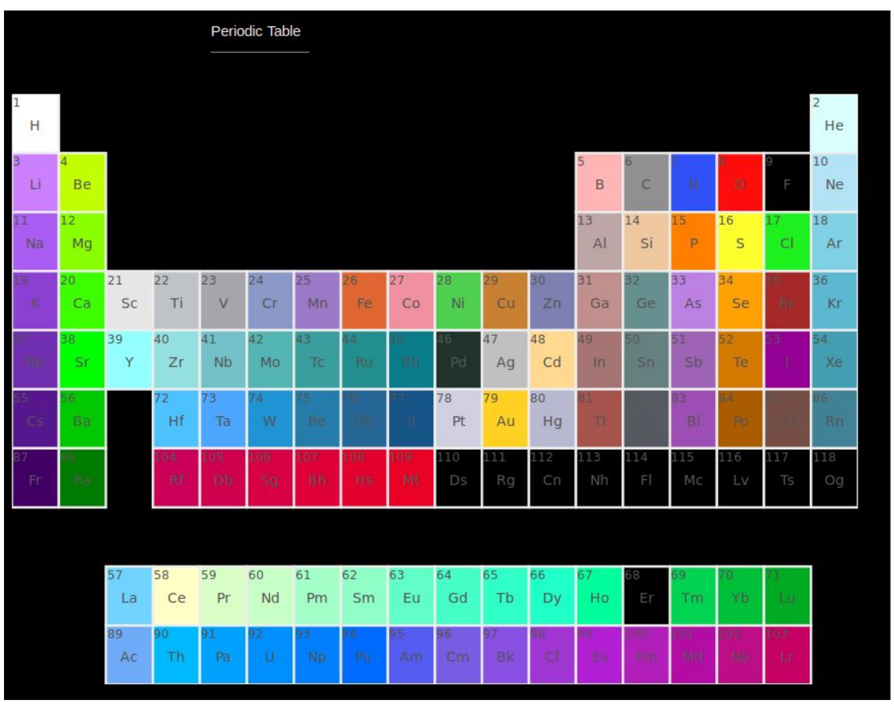
Scatter Plot-

The scatter plot ended up being the last visualization we implemented. It is an interactive visual that connects directly to the galactic visualization. The user can interact to develop a scatter plot of a correlation they would like (described further in the project process book) with the color they would like, then using the brush tool they can drag across the screen to select a particular region of stars. This selection will then be showed in the galactic visualization where those selected stars are positional in the Milky Way Galaxy. This gives the user the ability to connect stars together based off a variety of variables all dependent on their specific preferences. This personalization gives each user a new and unique informational experience to the website.

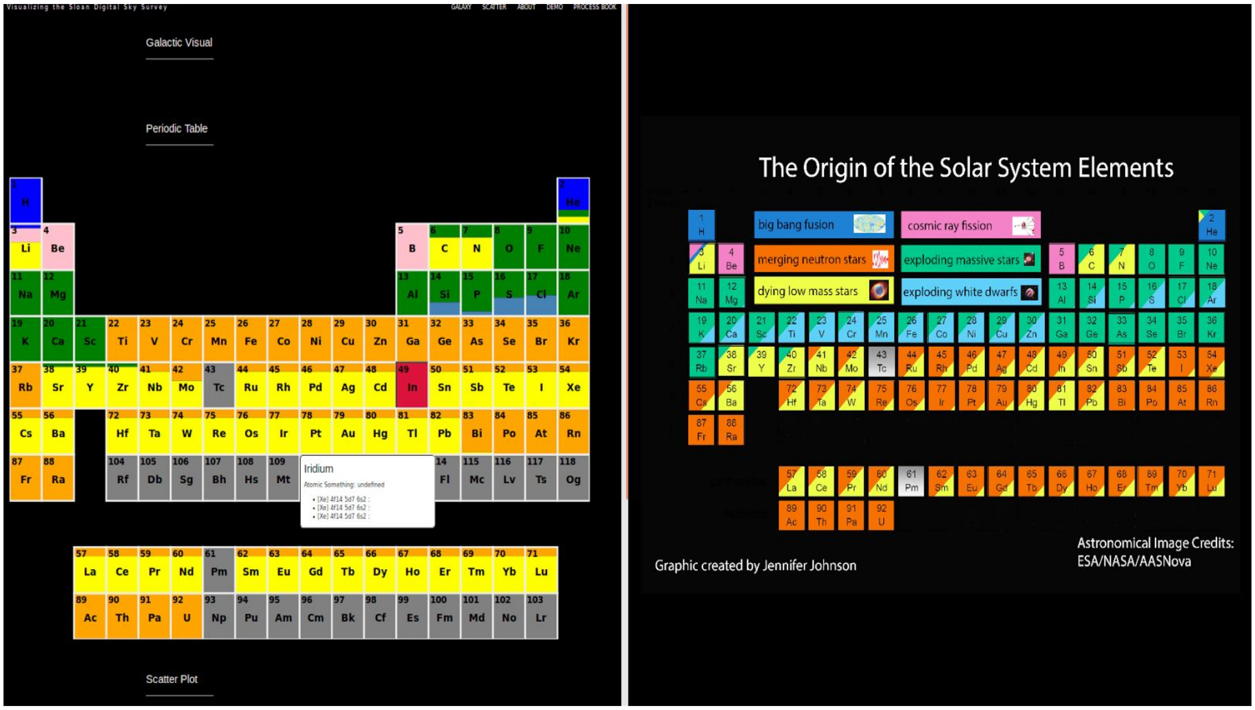
Color coding the table was vital in order for the user to have some visual understanding of the differences. The differences initially do not correlate with anything. But in other iterations we chose to make particular colors refer to being man made or created at the time of the big bang. Basically a 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.



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.

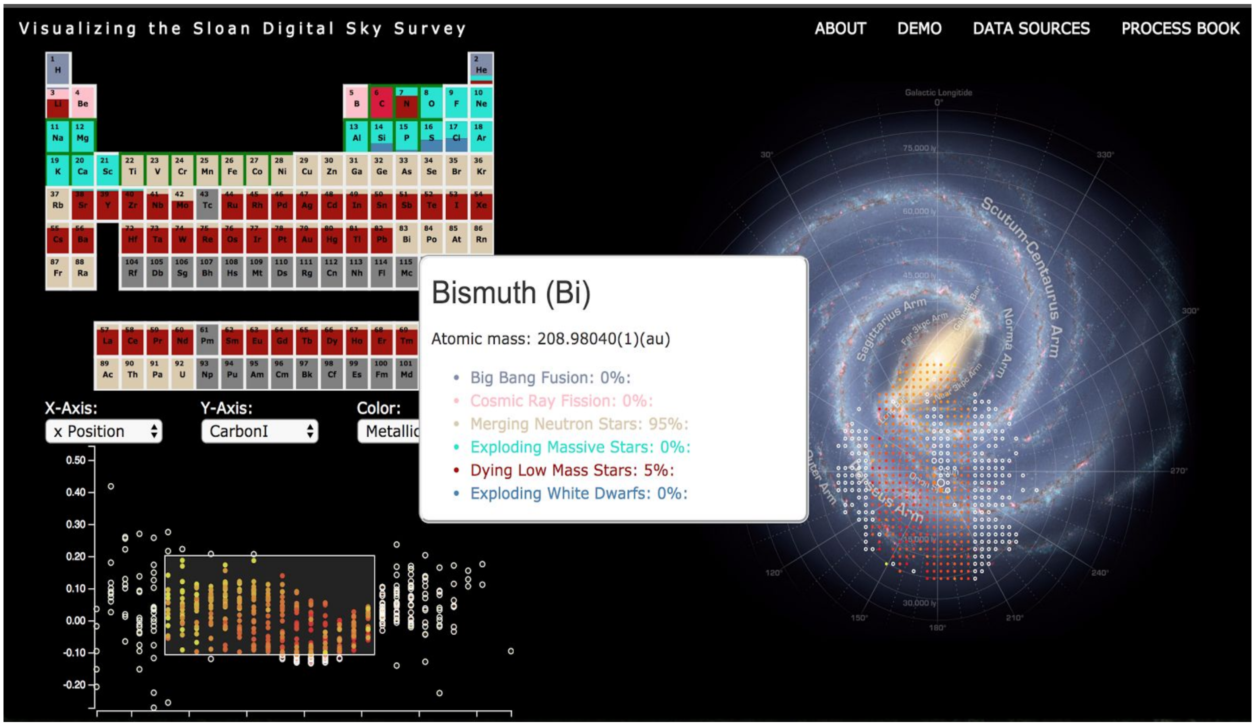


Lastly, we will implement the scatter plot and it will interact with the galactic visual just like the periodic table. The SDSS began regular survey operations in 2000, after a decade of design and construction.

It has progressed through several phases:

1. SDSS-I (2000-2005),
2. SDSS-II (2005-2008),
3. SDSS-III (2008-2014),
4. SDSS-IV (2014-).

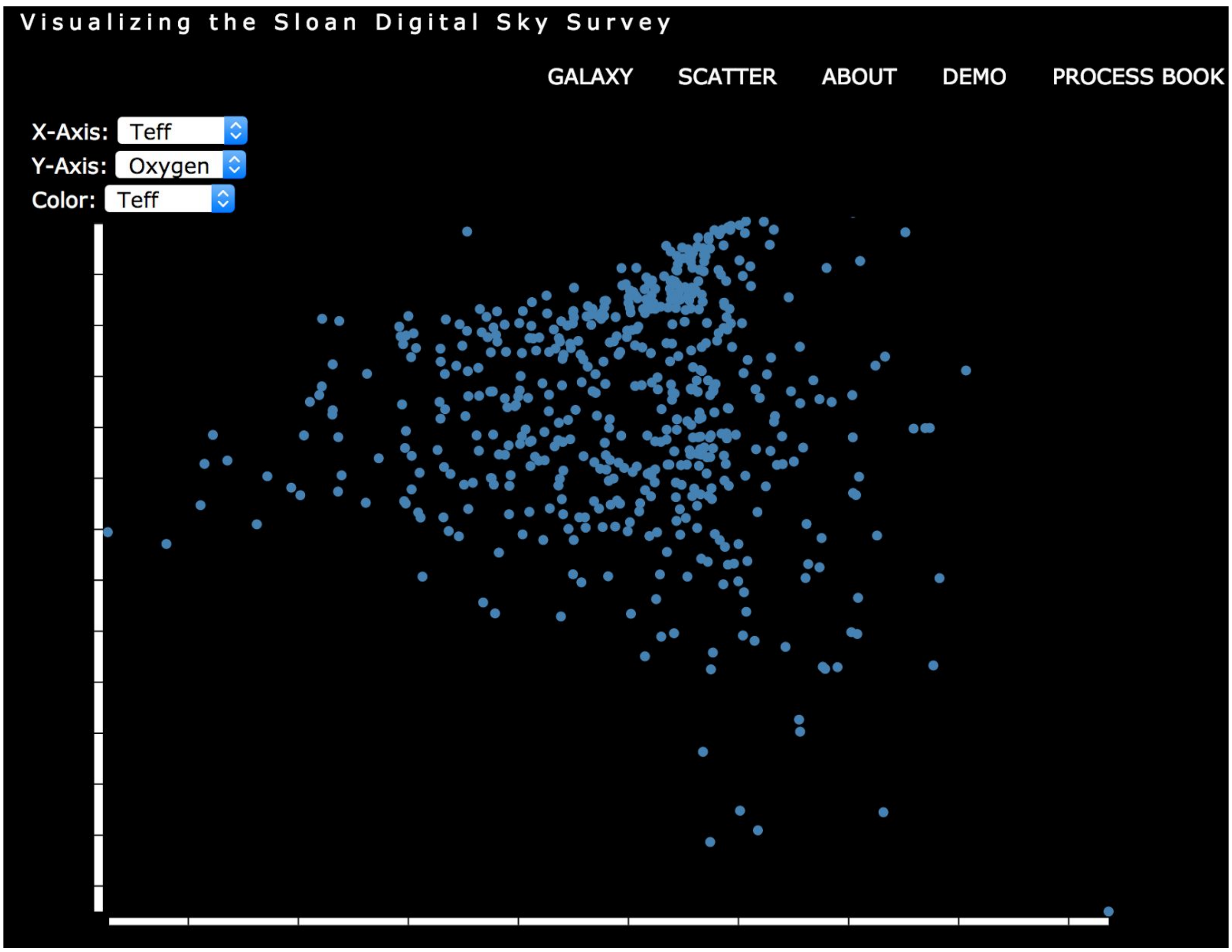
This allows us to make sure we can connect our galactic visual to our scatter plot in a variety of fields. Each of these phases has involved multiple surveys with interlocking science goals. The three surveys that comprise SDSS-IV are eBOSS, APOGEE-2, and MaNGA. We plan to correlate the two with a brush tool that the user can interact with.



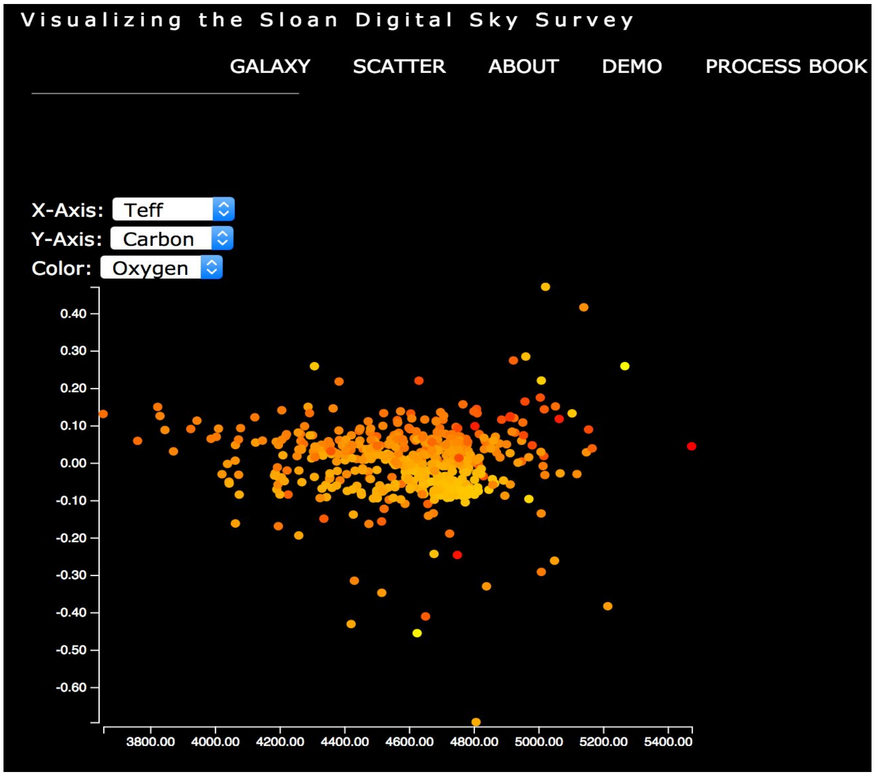
Our final periodic table looks like the one above. It displays the informational tool-tip and an ability to change the coloring and labeling of the entire visualization. This design evolution came from a rooted concept that we always stuck with our original designs. We did not deviate much from the proposal and learned that to communicate our idea in the best way possible we could add labeling. This tile chart is very clean and concise but we would've loved to generate a better area system to break apart our tiles in a cleaner fashion. We implemented the tiles as best as we could and like our current tiles but if we had more time or additional team members this is one thing we definitely would've liked to improve. The intent and functionality of all of our interactive visualizations give the user exactly what we wanted from the beginning of our design creations.

Scatter Plot Visual-

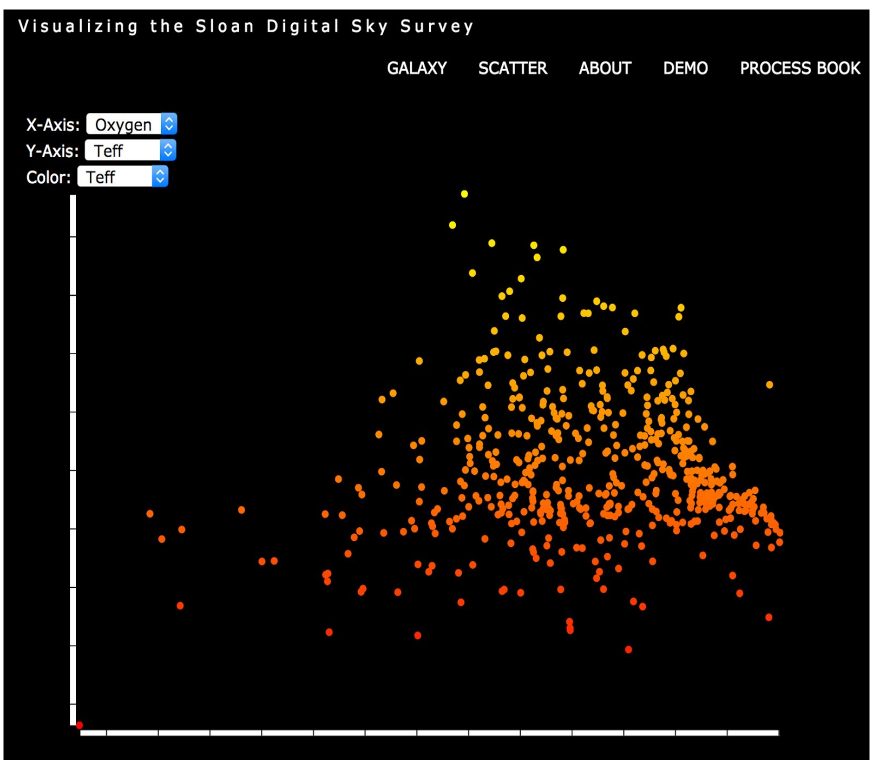
The scatter plot ended up being the last visualization we implemented. It is an interactive visual that connects directly to the galactic visualization. The user can interact to develop a scatter plot of a correlation they would like (described further in the project process book) with the color they would like, then using the brush tool they can drag across the screen to select a particular region of stars. This selection will then be showed in the galactic visualization where those selected stars are positional in the Milky Way Galaxy. This gives the user the ability to connect stars together based off a variety of variables all dependent on their specific preferences. This personalization gives each user a new and unique informational experience to the website.



Initially, our scatter plot was a great start. But we added quite a bit to it. We got a lot of our optional desired features and we knew we could make our scatter plot better than the above implementation. We unstacked our drop down menus to make them look better as we placed it in the lower left corner of the screen. We had some color scaling issues in this implementation and we changed that as our evolution progressed.



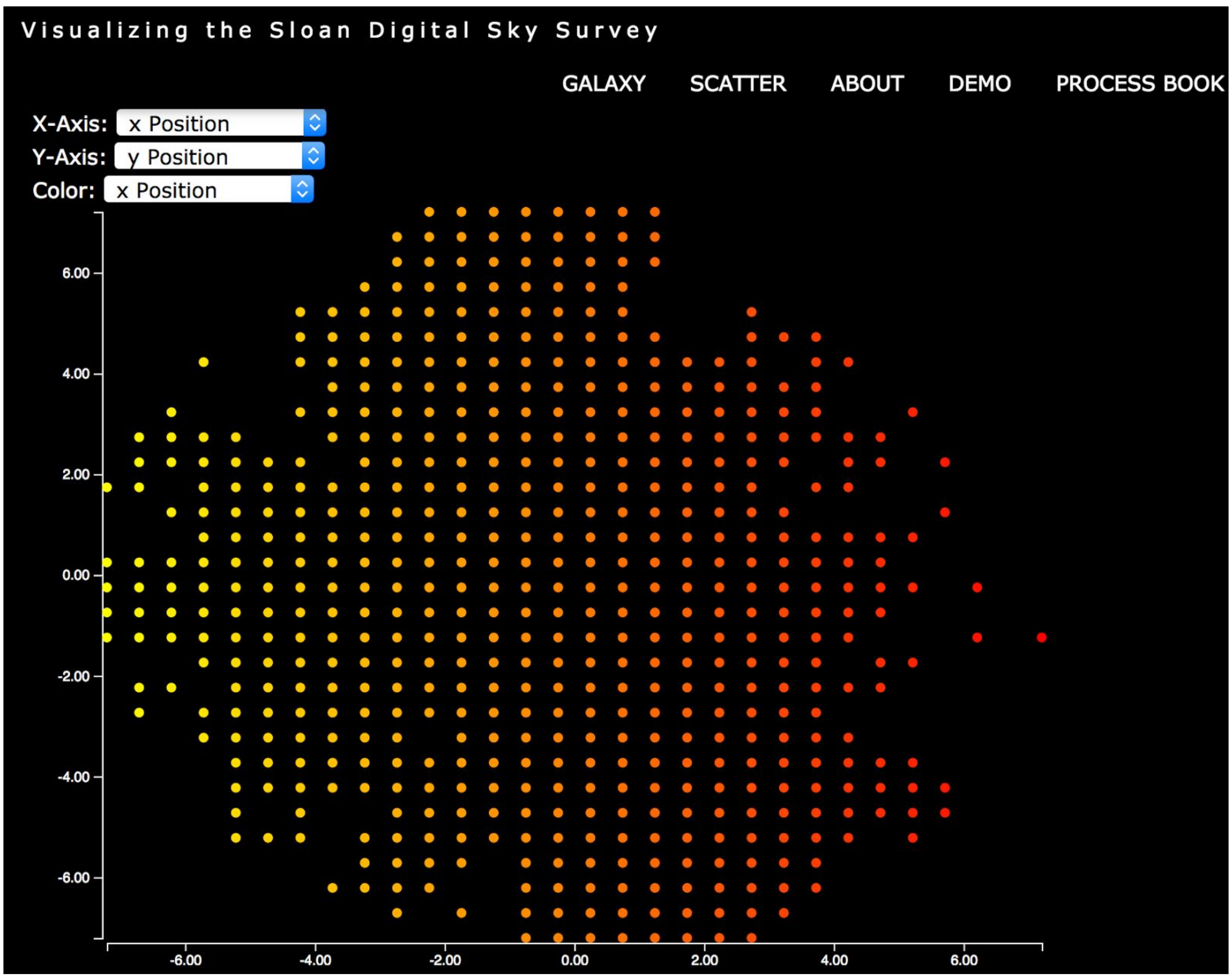
These two plots are to show the capability of the correlation applications of the scatter plot. This is an invaluable ability that we had designed from the proposal. We are incredibly proud of the possibilities we have created with or scatter plot and galactic visuals.

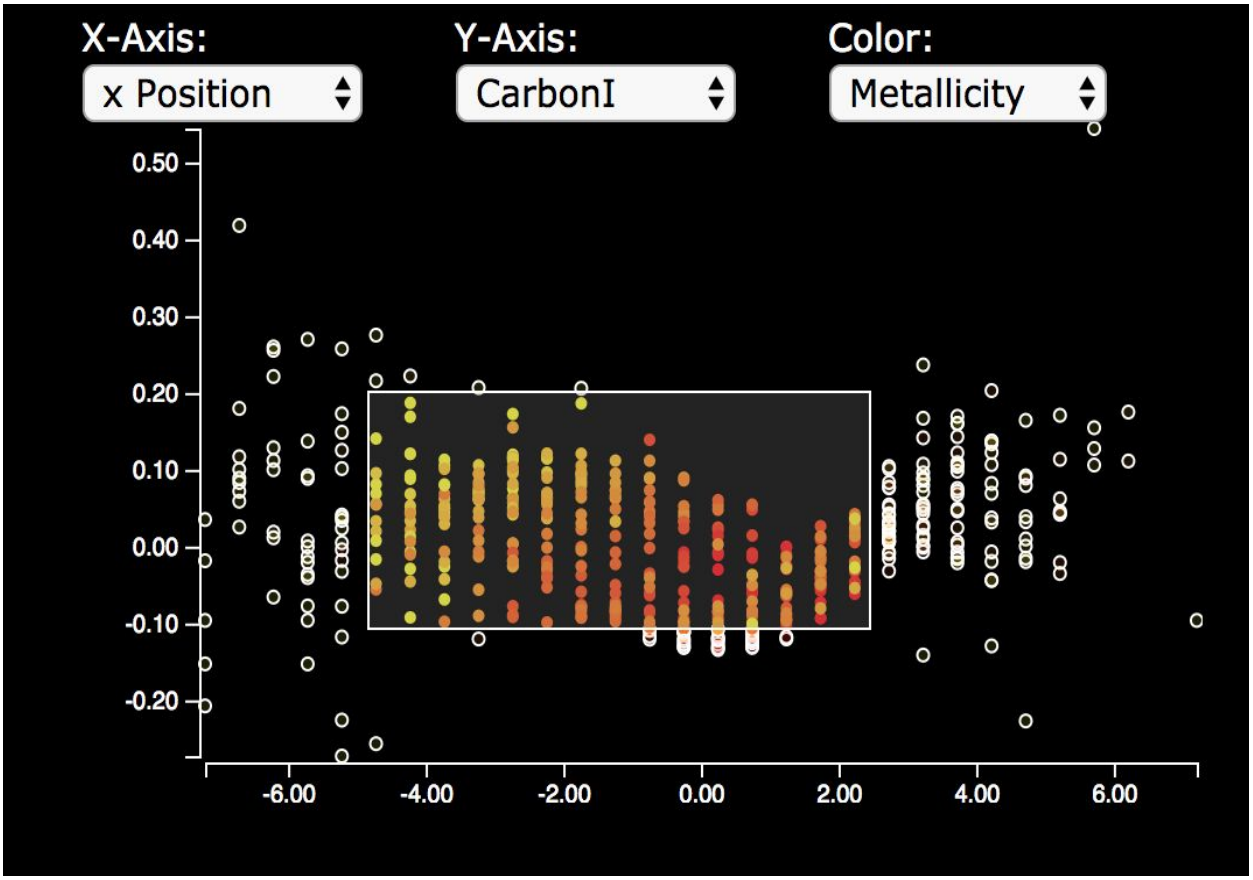


The galactic and scatter plot visualizations work flawlessly together. The coloring is better but we can make this even better to help the viewer experience a good informational system. Our implementation is extremely solid and includes nice D3 transitions as well. This design evolution has drastically improved from the infancy of the scatter plot.

Our final scatter plot that will be displayed is set with the X-Axis set at x position and the Y-Axis set at y position. The coloring is still to be fixed but the gradient is a much better look. In this format the stars directly match the galactic visualization layout. The scatter plot, while very simple shows the most out of our visualizations. Some of the information that can be gained from this visual includes:

- Brush selection allows precision star selection
- Correlations of any chemical abundance and or temperature is the most influential data shown from this plot
- Color/fill allows the user to have a visually intriguing layout
- The animation is the final touch on the display, it shows the extreme boundaries between stars as well as commonalities in the other case





Our final version of the scatter plot is shown above. The color, fill, bruch selection, axes, and overall implementation with the other galactic visual is complete. We always wanted to include the scatter plot within our visualization and we just had the time to do so. The scatter plot is one of our most proud elements in this assignment.

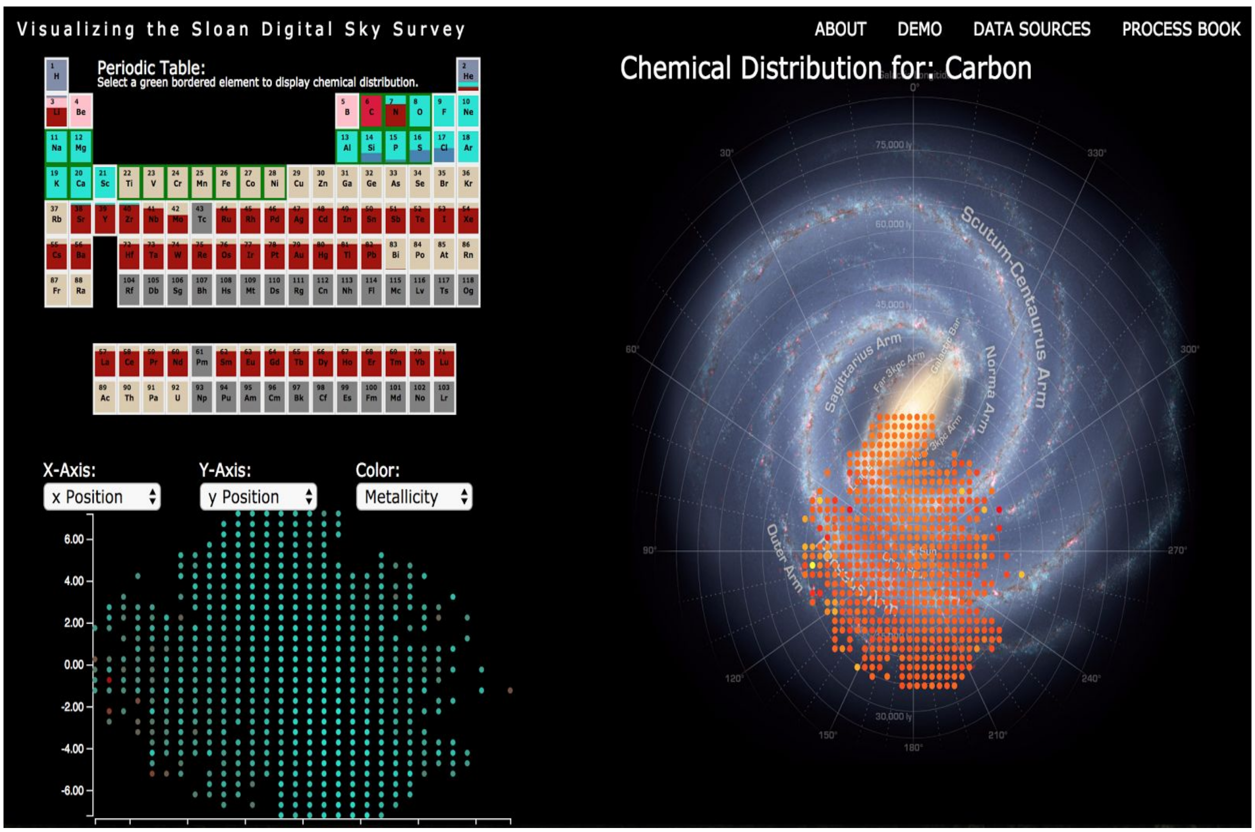
The key differences in this scatter plot from the other scatter plot implementations are:

- Horizontal drop down menus
- The brush tool is clearly shown highlighting specific stars
- Better axis labeling
- The gradient and D3 transitions are also new to this version

It is extremely informational while simultaneously being very visually appealing. If we were to harp on anything in this visual it would be the slight overlap in some correlations. Our design decision made the entire project better overall. The design principles we used started large. We wanted all of our visualizations to be seen on the webpage simultaneously so we put the scatter plot in the bottom left corner of the screen. We knew that color and transitions were crucial to the success of a plain scatter plot. While never deviating from the optional tag in the proposal we were able to generate a scatter plot with all of its functionality. We thought long and critically about our implementation and design to make sure that our scatter plot was great. We are very proud of the interaction we have within our visualizations and of how far we came.

Final Webpage Version-

Our main page of visualizations is shown below. It includes everything we wanted to include as well as the optional features we weren't sure we were going to be able to reach. We were very pleased with the work we put in. From visuals, to animations, to actual interactivity between the user and the visualizations, color scaling, solid HTML/CSS and powerful documentation we believe that our project has exceeded expectations. We look forward to having live users interact and work on the project to see what kind of interesting findings they encounter. We also plan to show professors Gail Zasowski and Kyle Dawson our progress since they helped us start a project that they believed could actually be beneficial to them and other researchers. Whether expert or novice we truly hope you enjoy our visualization project!



4. Evaluation-

Peer-to-Peer Evaluation-

Stefan: We learned a tremendous deal from our partnership and work loads. Anthony was a great partner to work alongside throughout this project. We were able to delegate visualizations, process book entries, and any other requirements seamlessly. We learned that our data had a lot of information but aggregating through the tables provided was a hassle and incomprehensible at times. Our goal was to make our visualizations tell as much of the story as possible. We answered questions like stellar chemical distribution and star pattern correlations throughout the Milky Way Galaxy. We believe we did a great job of showing the user a ton of information to the ser in a intuitive and aesthetic manner. If we were to change anything we would've liked to have a better color distribution among all the visuals so that they could directly coordinate in a more appealing way. This is not a huge problem, nevertheless it is something that can be improved on. Regardless, our visuals are very well implemented and function exactly to the specifications we desired.

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. I have never worked on a group project, but this project has taught me the importance of working as a group to accomplish a larger project. This project will be maintained and expanded on by me in my research, so I am excited to see what conclusions can be drawn from our visualization.

Conclusion-

That concludes our documentation for our visualizations. We hope that you enjoyed this project as much as we had fun making it. This process of creating a complicated visualization system was extremely challenging but also extremely satisfying. As all the pieces came together from our proposal we developed a project that tells a story while also being captivating. As a two person team we were able to accomplish a lot more than we expected of ourselves and we are proud of that. There were many constraints and obstacles to be dealt with, particularly in terms of feasibility and design. When there were many interactive components, efficiency and design were trade-offs. We made sure to use our skills used in class and our TA mentor's guidance to ensure our project was setup properly and functioning the best it could. Visualization is an important asset to have in the world we live in. We are both proud to have learned so much about data visualization and we know we will be able to apply it in later projects. The ability to capture someone's attention is vital and by taking the most precious commodity a person has, which is there time, we know how important it is to give the user a valuable experience in return. We both agree that we had a blast learning in this class and are proud of the final project that we made.

5. Acknowledgements

We'd like to thank the staff in our CS 6630 - Data Visualization class including the amazing Dr. Alexander Lex and our wonderful teaching assistants:

- Carolina Nobre, PhD student, Visualization / Computer Science
- Pranav Dommata, MS student, Visualization / Computer Science
- Dr. Trang Tran, PhD student, Computer Science

All the hard work you put into this course from filming lectures, to helping with our homeworks and projects, while being consistent with grading so quickly has not gone unnoticed.

We thank you!

THE UNIVERSITY OF UTAH