Influences, Deaths, and Mediums of Renowned Artists

Process Book

Nick Lord-Ender-Laing, Nishita Kharche, & Eisen Montalvo

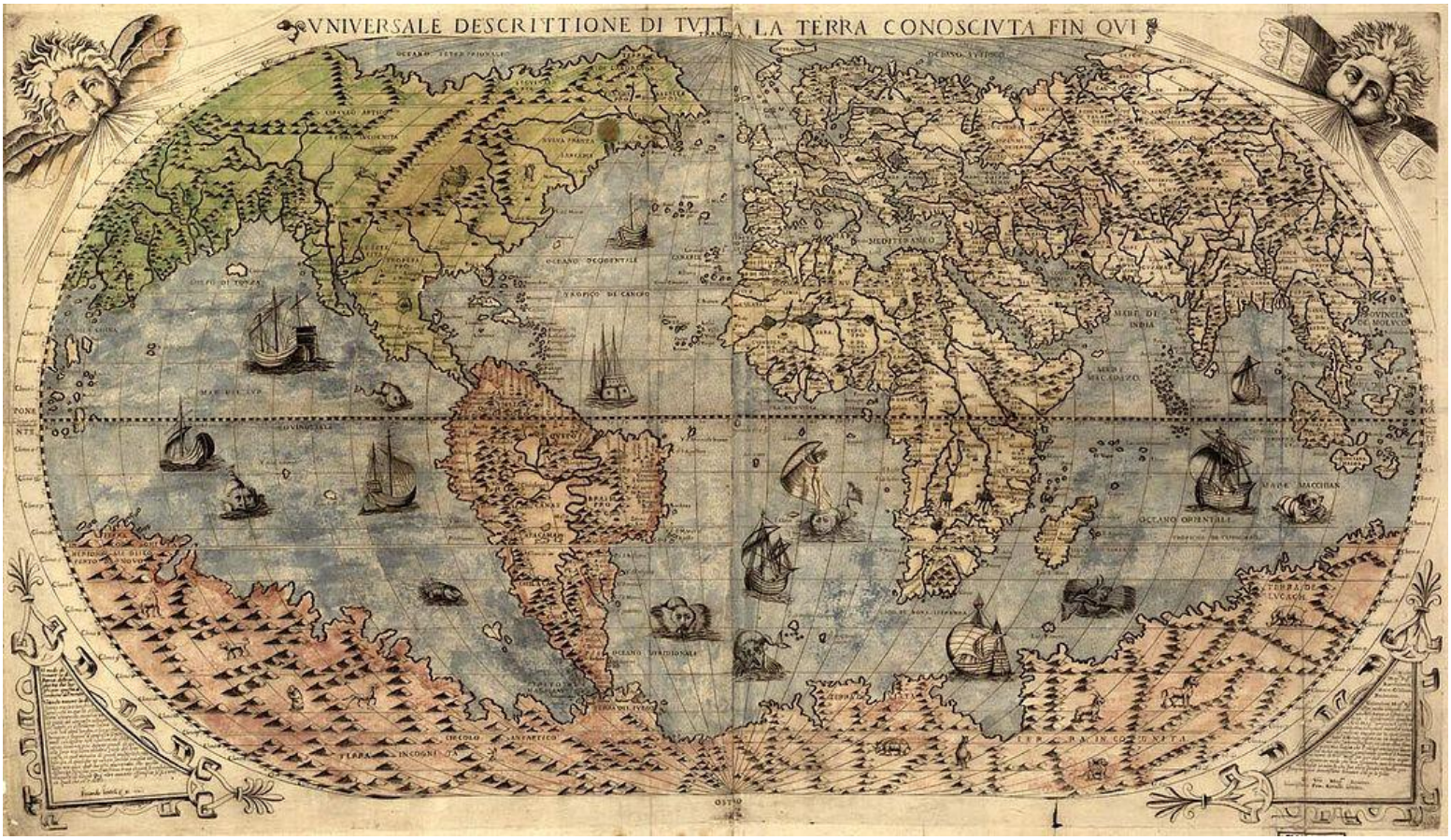


Table of Contents

- Overview & Motivation
  - Inciting Story and High-Level Motivations
  - Concrete Technical Deliverables: Influences
  - Concrete Technical Deliverables: Deaths & Mediums
  - If We Have the Time
- Related Work
  - Initial Inspirations
  - Inspirations While Revising
- Questions
  - Core Questions That We Are Trying to Answer
  - Development of Existing and Creation of New Questions
- Data
  - From Book to Table
  - Punching Up the Data
  - Outside of the Tables
  - The Map and the Old Timeline
  - The Birth of the Network View
  - The Ribbons Meet the Artists
  - Briefly Exploring the Adjacency Matrix
  - Minding the Gap
- Design Evolution
  - Phase 1: Early Stages
  - Phase 2: Post-Proposal: Reworking What's There
  - Phase 2: Post-Proposal: Brand New Parts
  - Phase 3: Final Designs: Big Picture
  - Phase 3: Final Designs: Traditional Tools
- Implementation
  - First Tab: Influences
  - Second Tab: Deaths
  - Third Tab: Mediums
- Evaluation
  - What Did We Learn About Our Data?
  - How Did We Answer Our Questions?
  - How Well Does Everything Work?
  - How Can We Improve?

Overview & Motivation

Inciting Story and High-Level Motivations

When brainstorming an idea from scratch for this project, we contemplated three very different ones. The first two were fine, but we spent only a few minutes speculating on fun and interesting things we could do with each. The third idea, by comparison, totally and effortlessly blindsided us. We spent easily over ten minutes just unloading exciting idea after exciting idea. By the end of the session, it went without saying that this third idea was the one we were going with, and that is the idea that we report on now.

This concept being: a data visualization that conveys relationships between each person in a set of historic artists. More specifically, we want to convey *influential* relationships. In the realm of art, it is not uncommon for an artist to have inspired the work of other artists. What makes this link so interesting to us is that they can extend from the influencer to another person hundreds of years later and on an entirely different continent (see fig. 1).

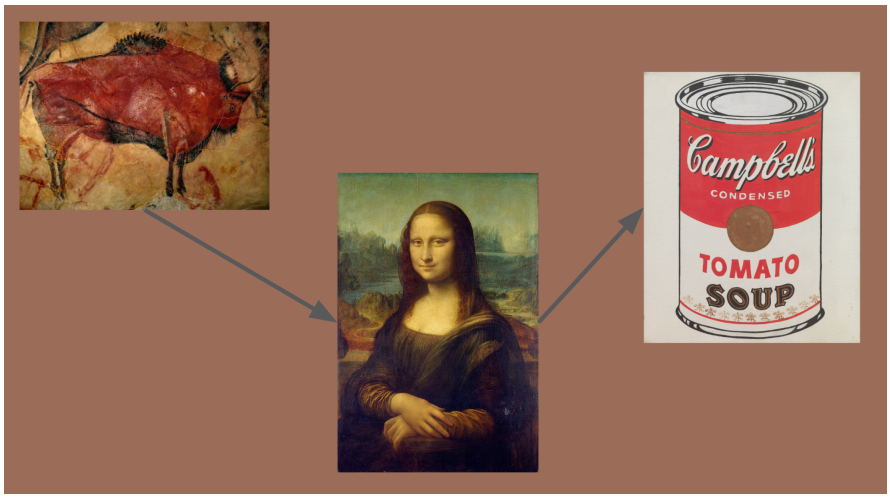


Figure 1. *An extreme influential link between the work of a cave painter and da Vinci, and da Vinci and Andy Warhol (this is not an actual link; it is sheerly for the sake of a high-level example)*

We not only want to share these fascinating invisible connections, but we could not help but think that they are begging to be made visible. Further, no one has taken it upon themselves to make this type of

rich information accessible and centralized for artists exclusively. This is plainly evident by one of our more concrete inspirations: art books. A member of the team had a particular interest in books that touch on these inspirational throughlines, and the best way to centralize and compare those connections was to flip through the pages and hope you remembered the last link you read about. A digestible data visualization that extracts the inspirations from those books seemed like a great remedy for both these wants of ours to convey this interesting type of data, and the gaps that we could not help but notice after conceiving the idea.

In a way, our overarching motivation is education. While one can say that about any visualization, we want to fully embrace a more traditional educational tone. The subject of historic artists lends itself naturally to that, and we want a visualization that would not look out-of-place as a museum display. The level of polish and presentation that we would need in order to match that is also a big goal for us. We really want to make it interactive too. We do not want something that is comparable to an unimpressive static map that you would glance at and then walk away from. We want to effortlessly convey to a user that there are more interesting things about our data than the default view that they can find out about by digging around. Our want for this virtual digging is important, and thus it informed many, if not all of our design decisions.

Concrete Technical Deliverables: Influences

With our broader motivations laying the abstract groundwork for our project, we next decided on some solid more technical goals that we want to accomplish by the end. These goals changed over time (see *Design Evolution* for further details), and so these represent the final version of our deliverables. Without these parts implemented and with interaction amongst them, we would consider our project a failure. The first one that we knew we must have is a world map. We repeatedly asked ourselves “do we really need a map?” and while its exact purpose changed, its presence as a goal never went away. Plotting these artists on that map is crucially important because it takes our data and grounds it in the

real world. Another way of grounding this data set visually is to include a component that shows a picture of each artist. We really want to convey that these are people who had their own lives (not just numbers), and a portrait is a simple and effective way of doing that. More specifically, we want a view with a standard network visualization showing images of the artists on each node, and edges showing their influences. We would then filter the map to *only* plot those artists selected from the network view (as just one example of our intended interactivity). This way, we can have accurate geographic spacing without worrying about clutter on the map. Avoiding clutter by dividing this into two views goes a long way in presenting something that looks good, to put it plainly.

On the other hand, a more blunt view is an equally important goal for us. We want to cater to those who like a less glossy type of visualization, and given our relationship-driven data, an adjacency matrix was the perfect choice. So for our influence data, we must implement an adjacency matrix, a map view, and a network visualization.

Concrete Technical Deliverables: Deaths & Mediums

At the same time, we want to make the most of our data. When we learned more about artists for this project, we latched onto how interesting both their means of death and artistic mediums were. These make for natural categories, and thus we feel that visualizing this, in addition to their influences, would make our final product richer.

To represent this data, we want to use both a more traditional visualization method, and a less orthodox one that looks more aesthetically pleasing. Our data set is related to master artists after all, and so we want a colorful, and visually striking creative view that does this type of data justice in combination with a more familiar component.

We want to use this traditional and unorthodox combination twice. Once to represent our death data, and again to represent our medium data. The more traditional half of these representations did not

matter that much to us at first, but we thought of a D3 chord visualization for the more creative half (see later sections for specifics) (see fig. 2).

Figure 2. *An example of how similar a D3 chord chart is to a painter's palette. This shows how naturally this chart lends itself to this kind of data in one way.*

Overall then, we would consider our project a failure if we do not have this scalable chord visualization that is capable of representing artists' deaths and mediums in two distinct views respectively. It would also be a failure if both views did not show our chord component hand-in-hand with a more traditional chart that fits the view's respective data (so mediums or deaths). This traditional and chord combination became very important, and proved its efficacy nicely in the latter stages of our work.

If We Have the Time

Lastly, we can briefly touch on our optional goals. These are things that would be nice to have, but we would not consider the project a failure if they are not present in the final product. We want our visualization web page to be perfectly legible on a browser regardless of the window size. While there can be some awkward ratios for any given window, not to mention vertically-oriented displays like phones, we want to accommodate for as many possibilities as we can. At the same time, we know that most of our users will be viewing this on a computer with at least HD resolution and horizontal orientation. So while we developed with this visual scalability in mind, we would not consider our project

a failure if we do not optimize for the extreme minority of users that use vertically-oriented computer monitors, or other unusual means of viewing like that. There will also likely be edge cases in which some things might be minorly obscured.

Another stretch goal is incorporating some text-based information on the artists. If we have the time, we would like to somehow incorporate links to the Wikipedia articles for each of our artists, and a user could go to the page if they want to learn more. More generally, this ties into storytelling. While a lot of explicit data visualization storytelling would be nice, it is largely icing on the cake, so to speak. If there is the time for it, we would want to accomplish something like the storytelling component of the sixth homework assignment for this class (see fig. 3).

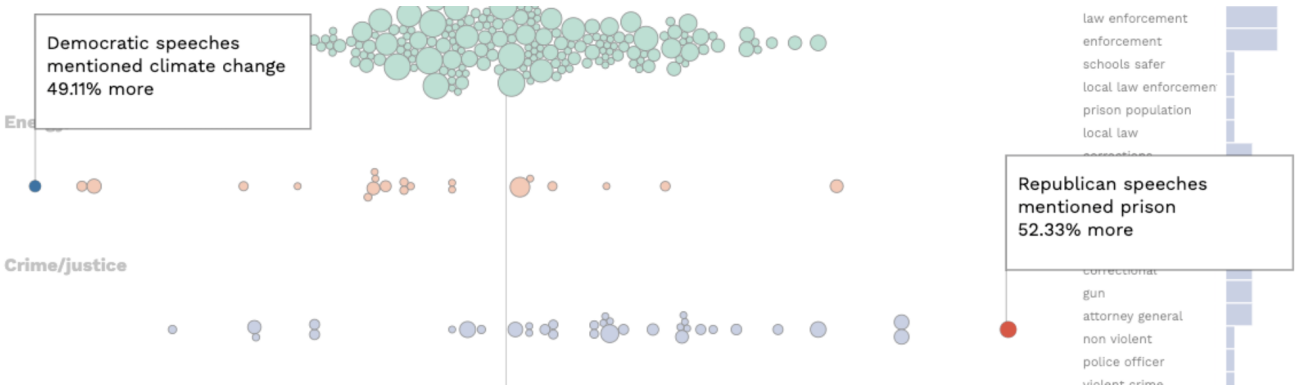


Figure 3. *The kind of minor storytelling we would want to accomplish if we have the time*

Related Work

Initial Inspirations

During our initial planning phase, our foremost inspiration for this project as a whole is our general interest in the history of artists and how they relate to each other via inspirations. We discussed how this would lend itself naturally to a type of network visualization (in relation to our latter interest) that could rely on a world map in order to ground it. More specifically, the books, *Art: The Definitive Visual Guide* and *Artists: Their Lives And Works* were our main source of inspiration. The stories of the people detailed in those books were the most fascinating parts, and we want to convey aspects of those stories in a visual way. Our focus on the stories of these people led us naturally to a technological source of inspiration, the screen-based displays you often find in museums (see fig. 4).

Figure 4. *An example of an interactive museum display that inspired us (see: https://prodisplay.com/market-sectors/public/museums-galleries/)*

We want our data visualization to convey a fairly neutral tone that encourages exploration and learning. We want to make something similar to this, but achievable on a web browser.

Inspirations While Revising

After we discussed our proposal with our assigned TA, we had a few takeaways concerning the direction of our project. One of them was that we should be visualizing more interesting data using more traditional methods. We then brainstormed some interesting traits about the artists in our data (namely, art medium, and means of death), and thought about representing those using a parallel coordinate visualization (see fig. 5).

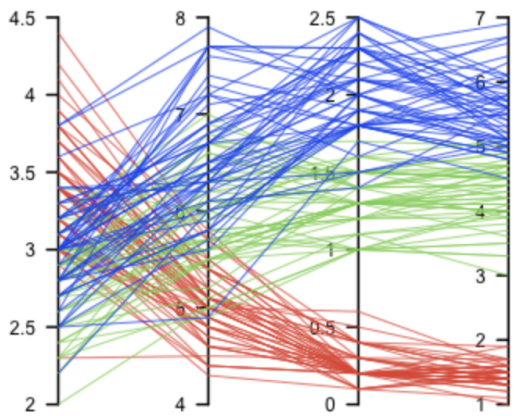


Figure 5. *An example of what initially made our gears turn in relation to potential additions (see Wikipedia article on parallel coordinates)*

However, we decided that this kind of visualization would not be very interesting because it is unlikely that art medium has any correlation with means of death (for example). We then remembered a somewhat similar visualization technique that would be both more aesthetically pleasing and appropriate for our type of data: chord visualization.

In particular, we really like the famous *The Lord of the Rings* chord visualization (see fig. 6). We were inspired to use this visualization's look and functionality to convey those tangential traits of our artists. In combination with dropdowns that can change both the grouping of our artists and the data attribute they map correspond to (medium, means of death, etc.), we can have a dynamic version of this visualization (for more details, see *Design Evolution* for further details).

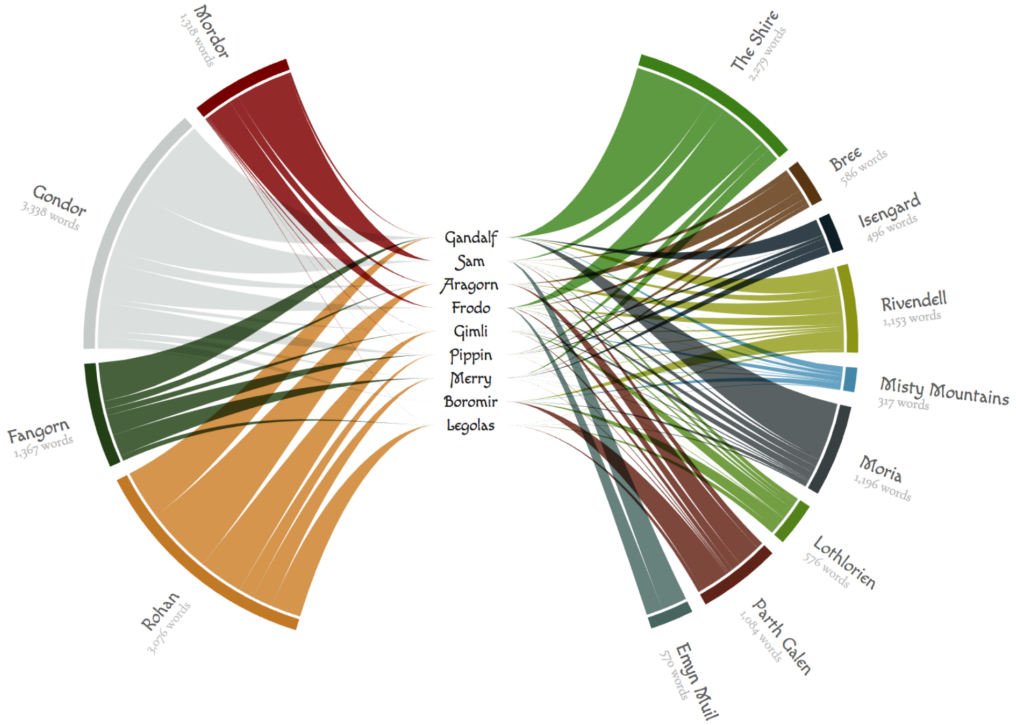
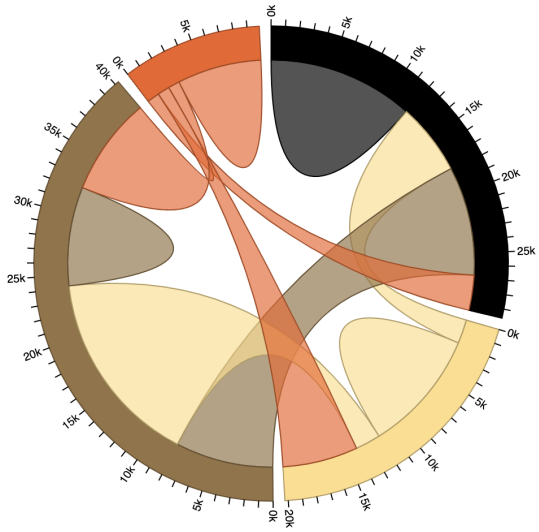


Figure 6. (see https://www.datasketch.es/project/lotr)

In addition to this example, we also took large inspiration from the in-class discussion on chord visualizations, and the accompanying code (see fig. 7). However, we took even more direct inspiration from the freely available chord and ribbon code on the official Github repository for D3. We followed the footsteps of the creator of figure 6, and decided to take the general methodology in the D3 code, and tweak it to look closer to what we want (see *Implementation* for further details).




```
<!DOCTYPE html>
<meta charset="utf-8">
<style>
    body {
        font: 10px sans-serif;
    }

    .group-tick line {
        stroke: #000;
    }

    .ribbons {
        fill-opacity: 0.67;
    }
</style>
<svg width="500" height="500"></svg>
<script src="https://d3js.org/d3.v7.min.js"></script>
<script>
    {
        // Our input data
        let matrix = [
            [11975, 5871, 8916, 2868],
            [1951, 10048, 2060, 6171],
            [8010, 16145, 8090, 8045],
            [1013, 990, 940, 6907]
        ];

        let svg = d3.select("svg"),
            width = svg.attr("width"),
            height = svg.attr("height"),
            outerRadius = Math.min(width, height) * 0.5 - 40,
            innerRadius = outerRadius - 30;
```

Figure 7. *The class code and example of a chord visualization that helped get the ball rolling*

Another part of the course that inspired us was the lecture on world map projections. Originally, we used an equirectangular projection to get some code off the ground quickly, but after learning about other projections from that lecture, we changed it to the one that we learned is the most well-rounded: Winkel tripel (*Citation to lecture slides perhaps*). You can see this difference in our early code in figure 8.

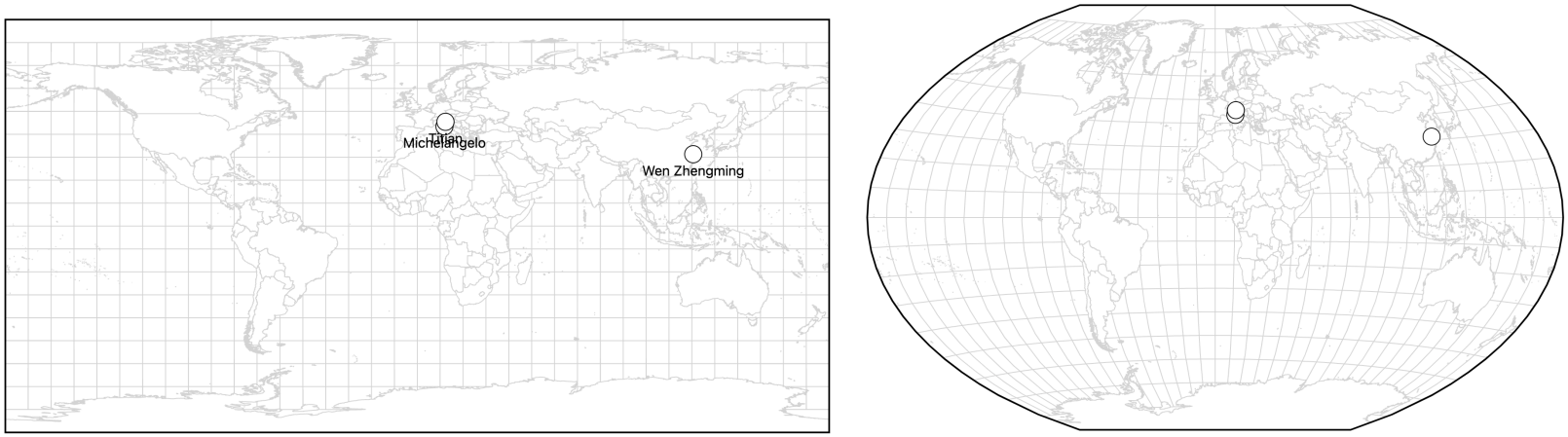


Figure 8. *A comparison showing the improvement of Winkel tripel over our original projection*

A short while after that lecture, we took a lot of notes on week 12's lecture, *Visualizing Networks and Trees*, and week 13's lecture, *Visualizing Multivariate Networks*. These two, and especially the latter one, was just what we needed to think of better network visualizations that we could add to our work. The type of data in the class examples seemed very similar to our influence data, and we arrived on an

adjacency matrix as a good simple view to add. In particular, we took inspiration from Micha Stubb's observable implementation of one in D3 (see fig. 9).

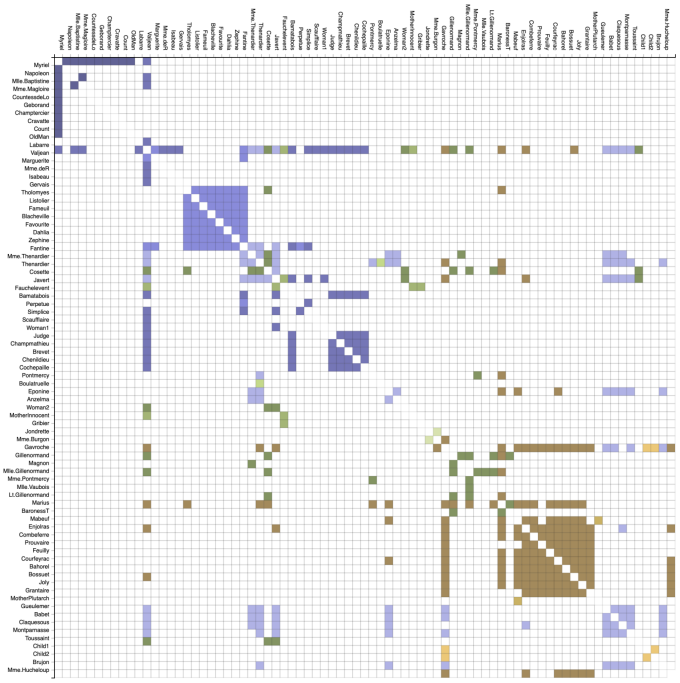


Figure 9. *An adjacency matrix that shows how valuable sorting can be in emphasizing visual trends (see https://bl.ocks.org/micahstubbs/7f360cc66abfa28b400b96bc75b8984e)*

We also took direct inspiration from our hands-on experiences with the homework assignments for this class. In particular, homework 6's usage of a D3 physics simulation gave us the idea to put one in our project's code. In particular, we used one for the rendering of our artist nodes in our traditional node and edge network view (see fig. 10).

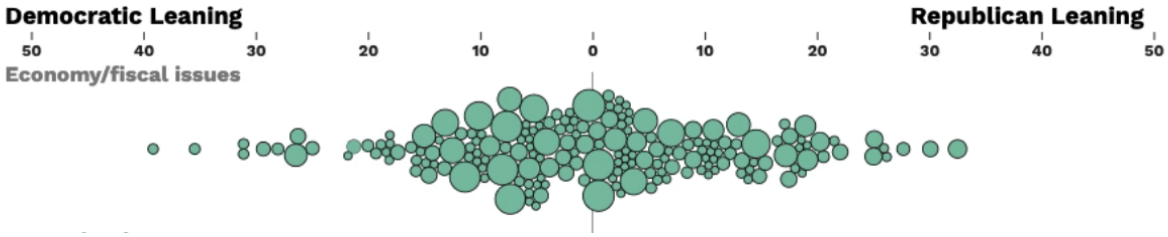


Figure 10. *The physics simulation from homework 6 that gave us the idea*

Another assignment that gave us some good inspiration was homework 4 (see fig. 10).

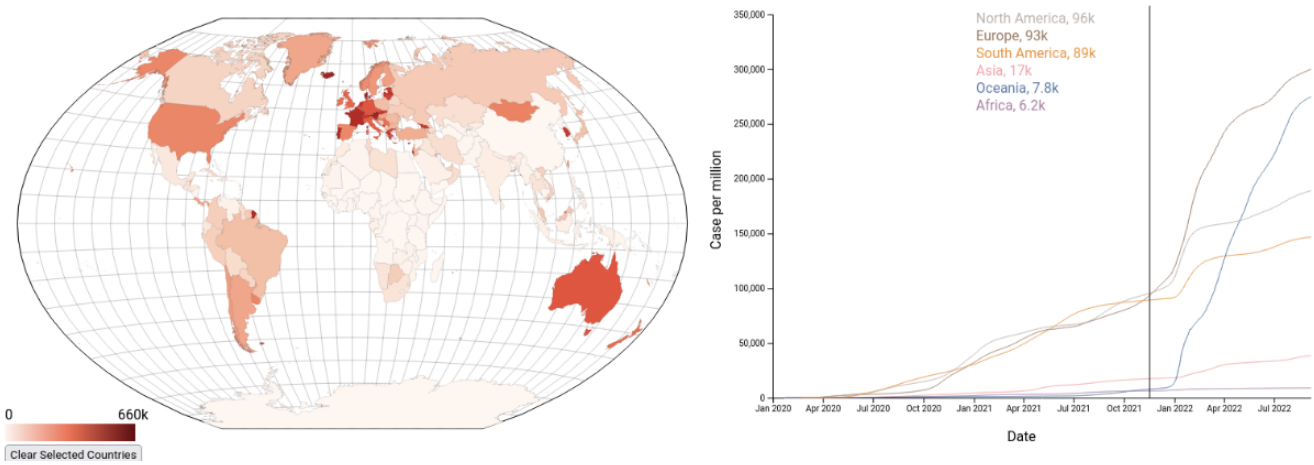


Figure 10. *Homework 4 from this class that includes the precision line on the right*

Specifically, the text that showed which lines were at which exact numbers was exactly what we ended up doing for our stacked area chart component (more on that in later sections).

Lastly, while this next bit of inspiration came to our attention late in this course, it was eerily similar to our project's premise, and thus it could not help but inspire us with some concrete ideas (see fig. 11).

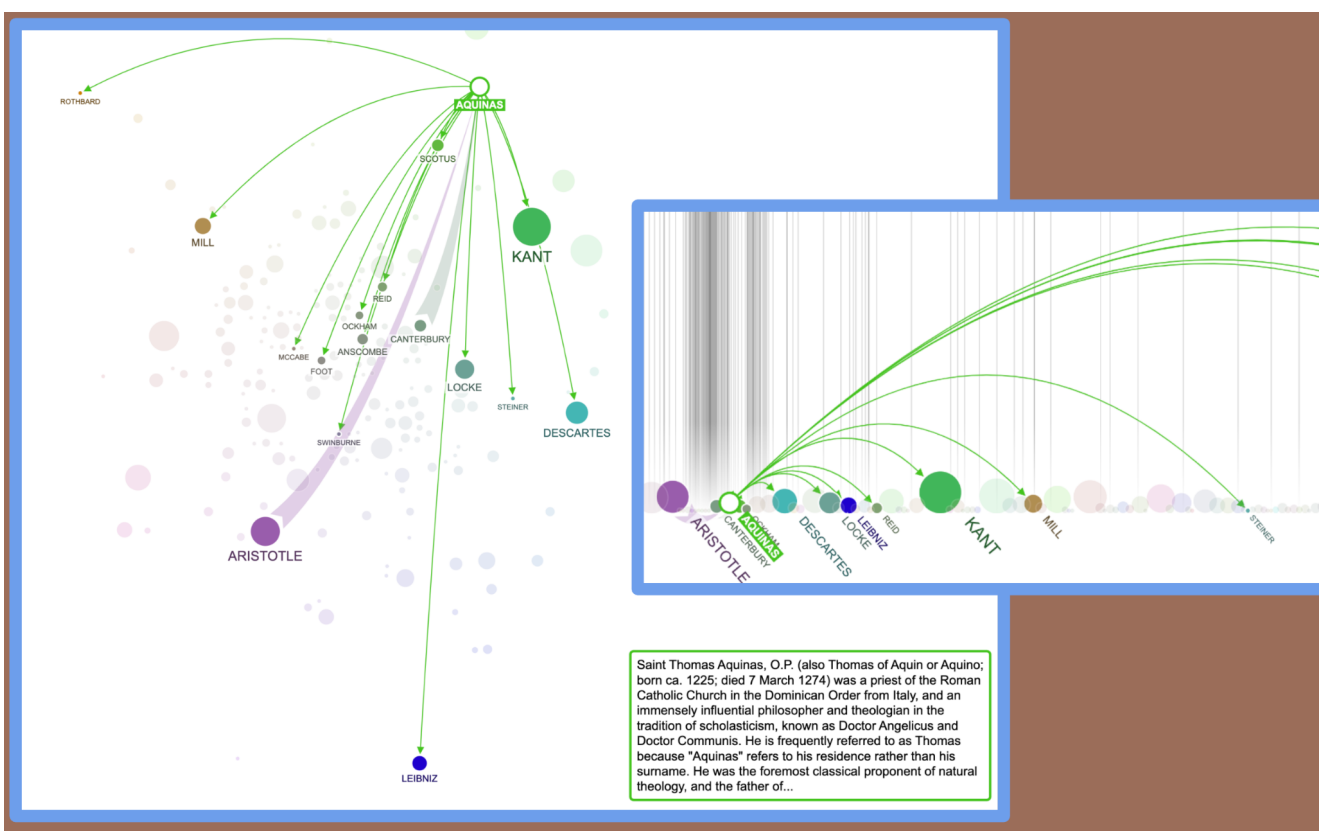


Figure 11. *The impressive edge map visualization from class*

Introduced to us in class during the lecture on filtering and aggregation, Marian Dörk's edge map visualization had an uncanny resemblance to what we wanted to do. This inspired us to do a few things despite already having most of our code off the ground. Firstly, this nudged us towards using multiple views like she did instead of rendering everything all at once. The transition between figure 11's edge map and timeline views made us realize that we have this option. Secondly, this visualization reminded us a bit of storytelling. While still an optional goal for us, we kicked around some ideas in relation to displaying some informative text similar to figure 11's approach. However, we think that this tool would have had even more influence on us if we found out about it earlier in the process.

Questions

Core Questions That We Are Trying to Answer

The most important questions that we want to answer visually are those in relation to the nature of inspirations between artists across all of history. A good way to contextualize this is to look at an example artist like Diego Velázquez. The first couple of questions we have about Velázquez are,

> ***"Which notable artists were major inspirations for Velázquez?" and "which notable artists did Velázquez inspire similarly?"***

Within these major questions, we then have more detailed questions concerning the nature of the inspiration both ways. For example,

> ***"Did Velázquez inspire people while he was alive, or did his work become more inspirational long after his death? If so, how long after?"***

are examples of the nuance of those deeper questions we are aiming to answer.

Meanwhile, there are more cut and dry questions that we want to answer. Still using Velázquez as an example, we would like to know,

> ***"When did Velázquez die? When was he born? Where was he born? When did he die? Where did he die? Who were his contemporaries?"***

While these are all simple facts that one can find easily on Wikipedia, one of our goals is to centralize the blunt answers to these questions visually. All together, the answers to these questions build a more comprehensive picture of the people in our data.

Development of Existing and Creation of New Questions

Of our initial questions, the ones that changed the most clearly are those related to influences between artists. Initially, we simply wanted to answer the question,

"Who influenced whom?"

That question became more layered, and after we finished our project proposal, we wanted to answer the more nuanced questions that fall under that main question more overtly. For example,

"What is the range of this particular influence relationship temporally?"

is a question we now want to clearly answer visually rather than simply providing the dates through text and letting the audience figure it out.

Lastly, after expanding our data to include the artists' mediums and ways in which they died, we were able to provide answers to some new questions. In relation to deaths, several visualization components that we developed can answer questions like,

"How many artists died in this particular way? Of the artists that died this way, how are they dispersed temporally? Did more artists die of a certain way in this century compared to the last? Is there a trend involving age for certain types of deaths in the data? etc."

Our chord visualization and a more traditional component made for death data specifically answer these questions through their interactions with each other, and their visual strengths. Meanwhile, our chord component and a different traditional chart that we were developing will answer these questions in relation to artistic mediums:

"How many artists in a given century used a particular medium? Was there a predominant medium for a given century? When was this particular medium at its peak in popularity? etc."

Answering these questions will both establish context, and convey interesting trends (see *Implementation* & *Design Evolution* for more details). These developments are key, and we believe that they will make our audience take away more information, which is one of our main intentions.

Data

From Book to Table

While our code that directly renders our visualization took a while to get right, a process that took a comparable amount of time was collecting our data. While there is plenty of data on historically significant artists, we found no data set that contains influences amongst them. So there was not a data set out there that we could use directly or even derive data from. We created our own data set by curating information from many sources *manually*. Our invaluable sources throughout this process were primarily books, and then we used Wikipedia to fill in any missing simplistic information like locations of birth and death. The books we referred to are: *Art: The Definitive Visual Guide* and *Artists: Their Lives And Works*, (see fig. 12 for covers) and the most important data we derived from there is the aforementioned influential links.

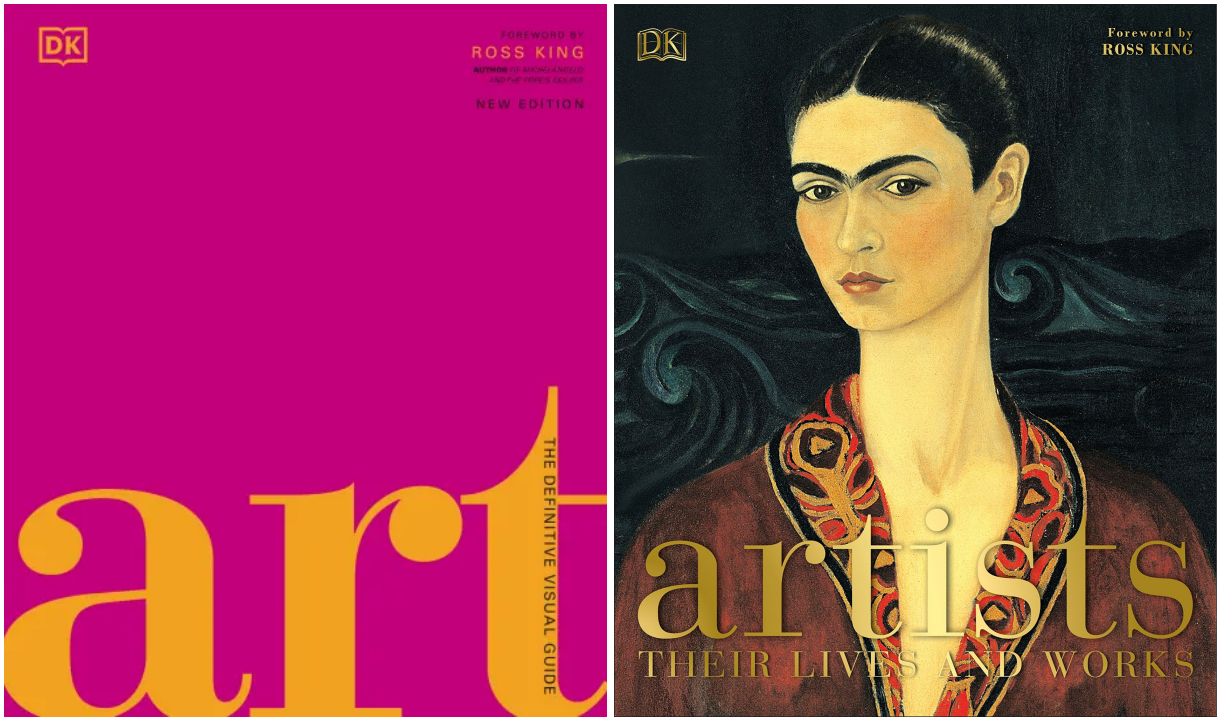
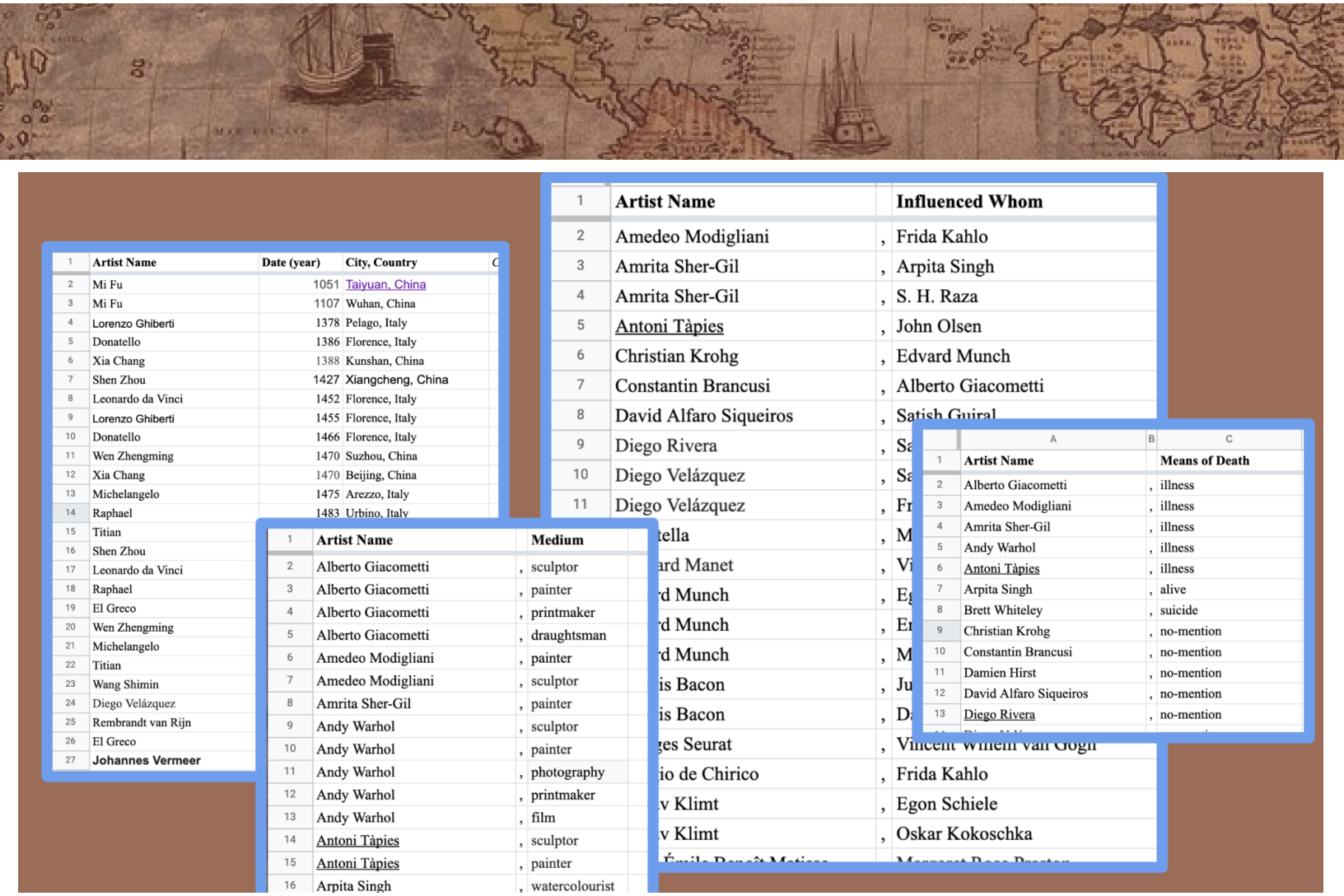


Figure 12. *Visual references for the books that we relied on*

Our data set is fairly small, around 100 artists, but the information on each artist is detailed, and spread across multiple tables. We created those tables on Google Sheets (see fig. 13).



1	Artist Name	Date (year)	City, Country
2	Mi Fu	1051	Taiyuan, China
3	Mi Fu	1107	Wuhan, China
4	Lorenzo Ghiberti	1378	Pelago, Italy
5	Donatello	1386	Florence, Italy
6	Xia Chang	1388	Kunshan, China
7	Shen Zhou	1427	Xiangcheng, China
8	Leonardo da Vinci	1452	Florence, Italy
9	Lorenzo Ghiberti	1455	Florence, Italy
10	Donatello	1466	Florence, Italy
11	Wen Zhengming	1470	Suzhou, China
12	Xia Chang	1470	Beijing, China
13	Michelangelo	1475	Arezzo, Italy
14	Raphael	1483	Urbino, Italy
15	Titian		
16	Shen Zhou		
17	Leonardo da Vinci		
18	Raphael		
19	El Greco		
20	Wen Zhengming		
21	Michelangelo		
22	Titian		
23	Wang Shimin		
24	Diego Velázquez		
25	Rembrandt van Rijn		
26	El Greco		
27	Johannes Vermeer		

1	Artist Name	Influenced Whom
2	Amedeo Modigliani	Frida Kahlo
3	Amrita Sher-Gil	Arpita Singh
4	Amrita Sher-Gil	S. H. Raza
5	Antoni Tàpies	John Olsen
6	Christian Krohg	Edvard Munch
7	Constantin Brancusi	Alberto Giacometti
8	David Alfaro Siqueiros	Satish Guiral
9	Diego Rivera	Sa
10	Diego Velázquez	Sa
11	Diego Velázquez	Fr
	tella	M
	ard Manet	Vi
	rd Munch	Eg
	rd Munch	Er
	rd Munch	M
	is Bacon	Ju
	is Bacon	Da
	ges Seurat	Vincent Willem van Gogh
	io de Chirico	Frida Kahlo
	v Klimt	Egon Schiele
	v Klimt	Oskar Kokoschka

1	Artist Name	Medium
2	Alberto Giacometti	sculptor
3	Alberto Giacometti	painter
4	Alberto Giacometti	printmaker
5	Alberto Giacometti	draughtsman
6	Amedeo Modigliani	painter
7	Amedeo Modigliani	sculptor
8	Amrita Sher-Gil	painter
9	Andy Warhol	sculptor
10	Andy Warhol	painter
11	Andy Warhol	photography
12	Andy Warhol	printmaker
13	Andy Warhol	film
14	Antoni Tàpies	sculptor
15	Antoni Tàpies	painter
16	Arpita Singh	watercolourist

	A	B	C
1	Artist Name		Means of Death
2	Alberto Giacometti		illness
3	Amedeo Modigliani		illness
4	Amrita Sher-Gil		illness
5	Andy Warhol		illness
6	Antoni Tàpies		illness
7	Arpita Singh		alive
8	Brett Whiteley		suicide
9	Christian Krohg		no-mention
10	Constantin Brancusi		no-mention
11	Damien Hirst		no-mention
12	David Alfaro Siqueiros		no-mention
13	Diego Rivera		no-mention

Figure 13. *High level overview of our initial data acquisition process*

We manually filled a table that contains the artists' mediums of art, a table that has their means of death, another with the influence-based connections, and one with locations and dates associated with their births and deaths. Then, the key for accessing an artist's data across different tables is simply their name.

This method had the great advantage of keeping our data just as we wanted it to be. It would be pointless in a visualization like ours to include artists with no influence link in the set, and thus we kept a standard that all of our artists would have to have at least one link. This also allowed us to get very familiar with the data, as we would even personally discuss particular characters among the set.

Punching Up the Data

This method was difficult in some areas, and a big example involves a library we use for converting a city and country to latitude and longitude. If an artist's location did not fit that format, we

would have issues. For example, Marcel Duchamp's birthplace is Blainville-Crevon, France. This is a commune, and so we checked Google Maps to see which city it roughly maps to. If the closest city was not clear on Google Maps, we then had to reference a more detailed map of France (see fig. 14).

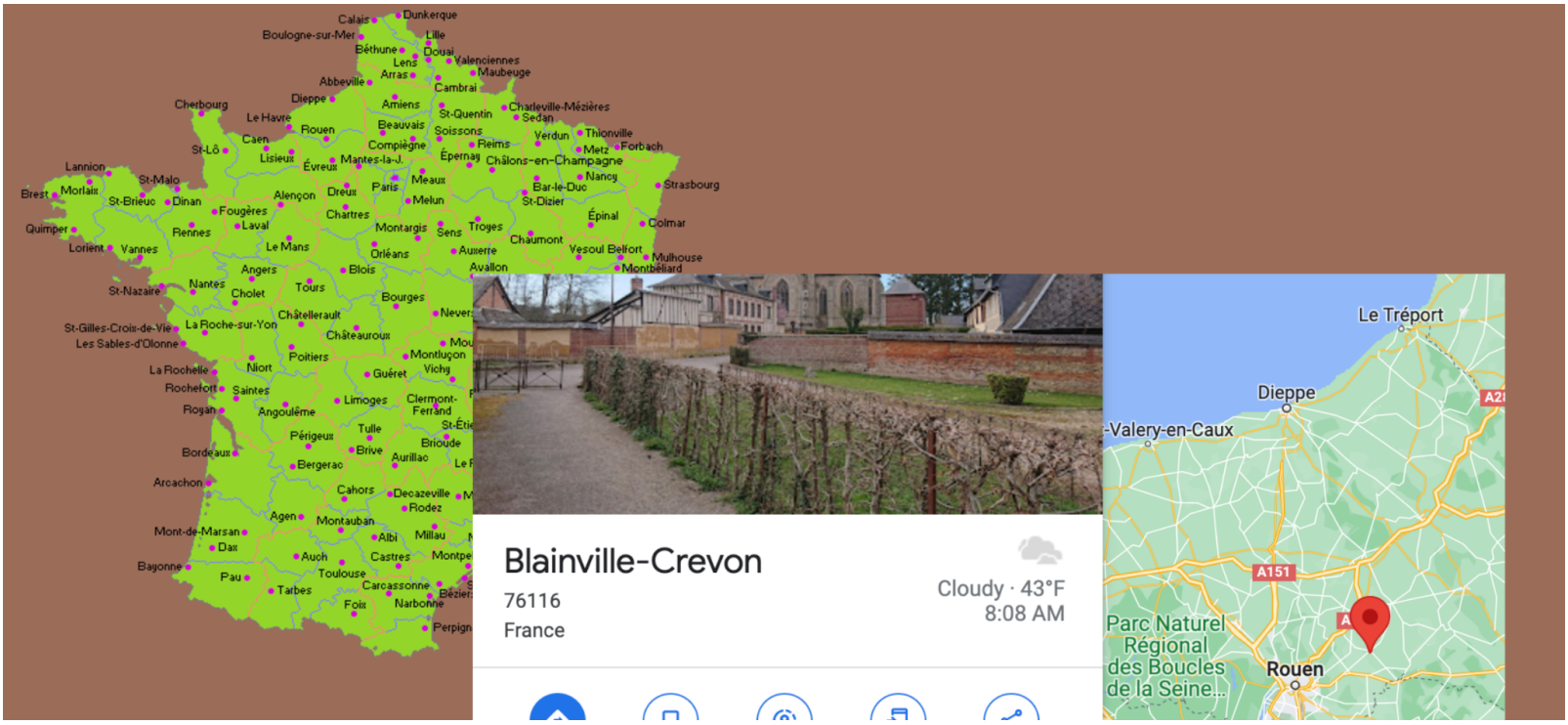


Figure 14. *The type of map that we would use and an example town on Google Maps that caused the workflow*

While examples like these did not pop up often, they appeared enough to be noteworthy, and we had to use similar processes if something like that happened in a different country. We also cleaned up our data in some minor ways to make it more code-friendly. An example was getting rid of any and all accent marks on our artists' names.

Also, there were a small handful of times when we wanted to augment our data to have at least one artist from each continent or at least one from each century. In these few times, we would simply take to various articles online on those very topics.

Outside of the Tables

After gathering all of our data into tables, there were some elements and processes that did not fit cleanly into our Google Sheets workflow. Firstly, it is worth noting that we considered adopting a

database approach for this project, but because of the size of our data, we decided to use local .json files. This also removes the perhaps needless task of maintaining a database properly. We used scripts to convert our .csv files into JSON, and then stored them in a simple folder. One of the biggest changes to the data in this JSON form involves the aforementioned library, node-geocoder, that we used to convert our locations into latitude and longitude that we could then plot on a map.

A chunk of our data that we also keep local is the portrait images for each artist in our set. Rather than scraping images from Wikipedia of varying dimensions, we manually gathered portrait images for all roughly 100 of our artists. We then used a clever Adobe Photoshop workflow to convert them into circular thumbnail images (see fig. 15).

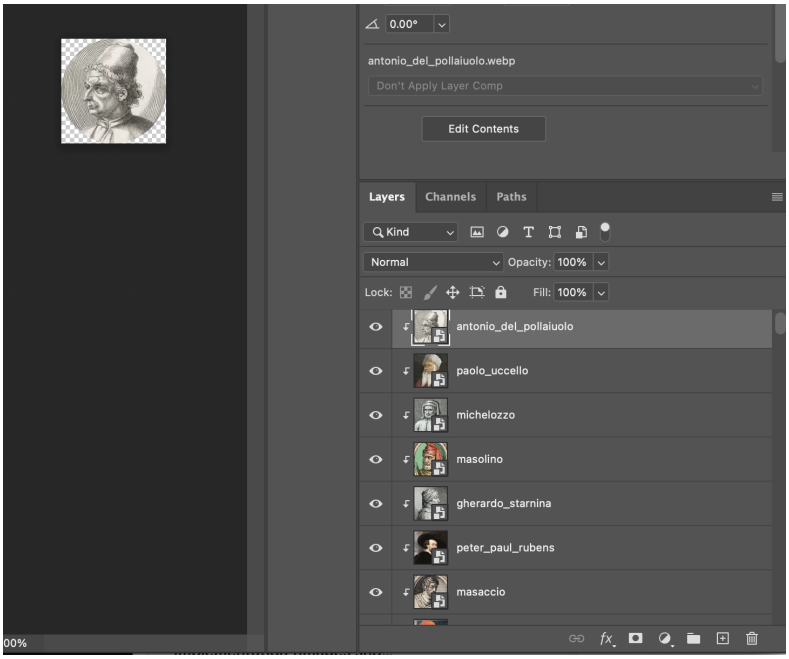


Figure 15. *Our clever Adobe Photoshop workflow using layers effectively for masking*

These images are all of the same size, and that size is small enough so that they do not take up too much space. We also augmented our JSON-ified table containing means of death for each artist to include another attribute. That being the local file path to that artist's thumbnail image. Overall, the data collection and adjustment processes were extremely hands-on, but we find this appropriate given the nature of our data.

Exploratory Data Analysis

The Map and the Old Timeline

The first way in which we visually explored our data was plotting the appropriate part of it onto an early version of our map component. After seeing the data this way (see fig. 16), a particularly important improvement became obvious.

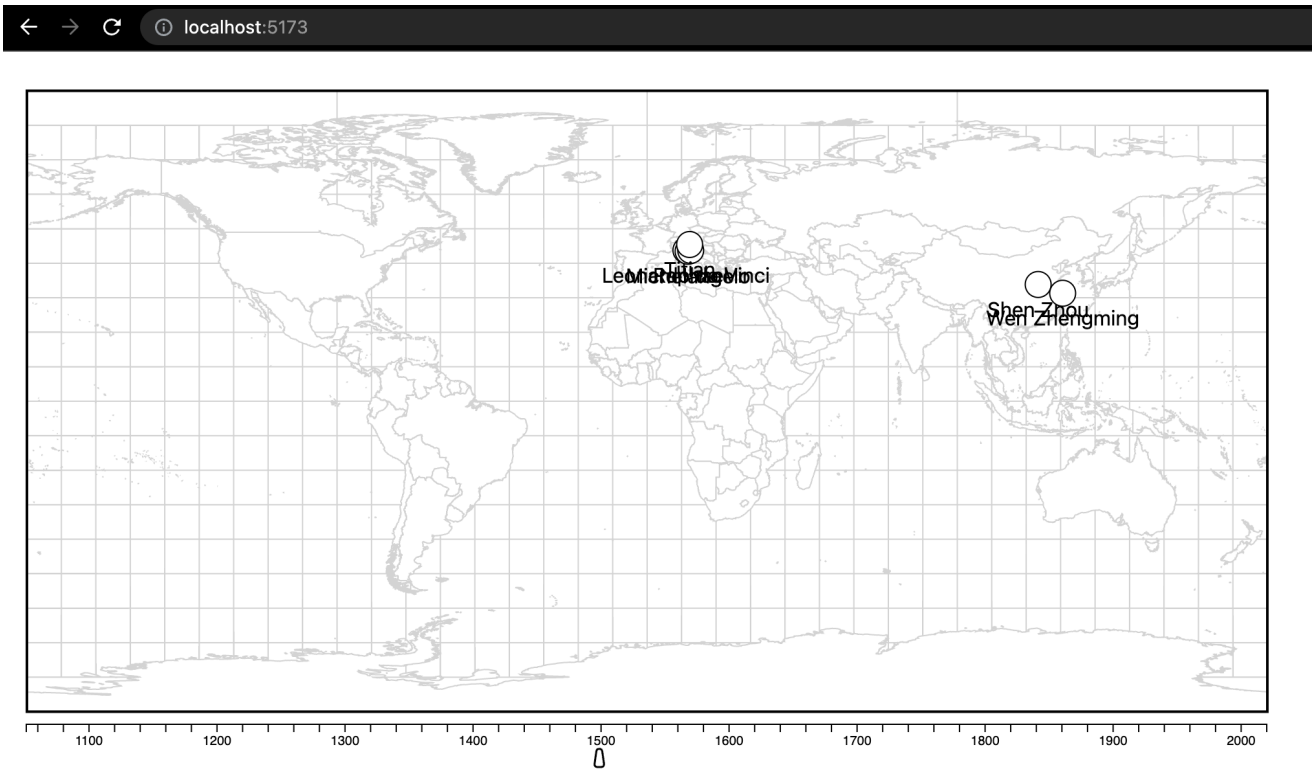


Figure 16. *An example of bad geographic clutter*

While we are no longer filtering the map view based on a selected date on a timeline, the general problem of clustering nodes led us to looking into D3 force physics simulations to space these markers well. The fact that the artists' nodes cluster in particular regions around particular timelines also led us to the greater decision of abandoning the timeline component as a whole, and filtering the map view in a different way (see *Implementation*).

Something else that became apparent through our exploration in relation to figure 16 is that there are geographic gaps in our data. We had little to no artists from Africa, South America, Canada, and parts of east Asia. To make the most of our screen real estate we both augmented the data to fill in

these gaps, and made the map only one of four views instead of our main component (that idea being a remnant from some of our earliest plans).

Speaking of gaps, we also have some temporal gaps in our data. This problem exists mainly between the oldest artist in our set, who lived around the 11th century, and the late 14th century. As a minimum remedy, we wanted to have at least one artist per century. At the same time, the realization that it would take a lot more data to pad out every slight temporal gap in our data led us even further away from our originally planned timeline component.

The Birth of the Network View

Something else that our map-related analysis led us to is our current network view. We looked at these apparent clusters in our data, and we could not really get rid of them due to how interesting the links are. We thought, *if only we could space our nodes optimally without worrying about location-related accuracy*. To attain just that, we decided after this analysis to create a more traditional node network view (see fig. 17).

Figure 17. *An example of how neat and tidy the network looks over having everything on the map*

We developed a way of using this new view to filter the map view in a way that prevents geographic or temporal uncomfortable clustering (see later sections for more details). Additionally, something that figure 17's presentation showed us is that there are some obvious islands in our network. These island groups could take priority over other nodes that are not as significant. For example, if we need to get rid

of some nodes for whatever reason, or replace some with others for the sake of geographic or temporal diversity, we would acknowledge that these islands are probably not good candidates for deletion.

The Ribbons Meet the Artists

Our chord and chart underwent some experiments with fake data while getting it off the ground in its very early stages. The crux of this component are the groupings and subgroupings of all of the data (the arcs), and the mapping of those subgroups onto their attributes (the ribbons show this). When working with just the arcs, we replicated a realistic number of groups, 10, and experimented with an odd number of them as well (see fig 18).

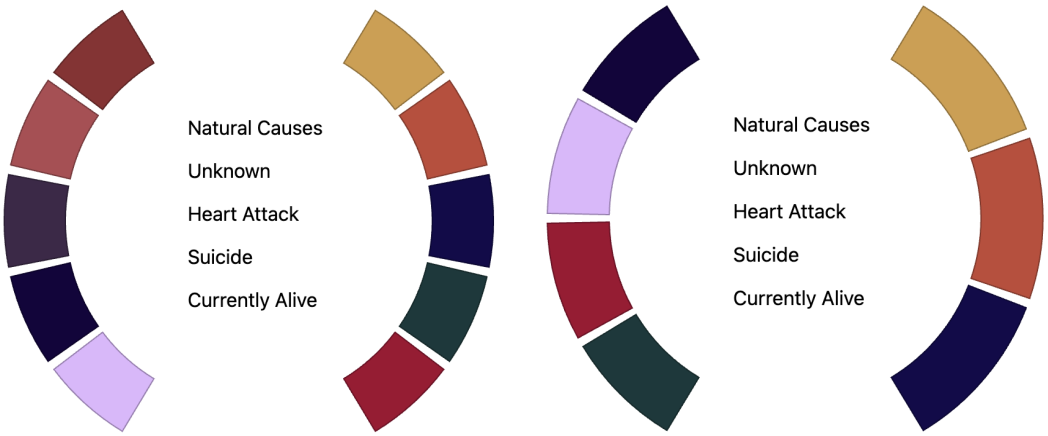


Figure 18. *Two examples of our early chord visualization on n=10 (left) and n=7 (right) data groupings.*

Through this, we saw that we likely would not have to worry about running out of space for this unorthodox component, but wiring real data to it made us make a decision for the better.

When wiring real data initially, we realized that a uniform sizing for each arc would be a potential missed opportunity in regards to encoding. We decided that we should encode the number of artists in a grouping by the size of the arc, even if this meant complicating the program. After accounting for both this and wiring real data, one can see how it looked in figure 19 when combined with our ribbon generating method.

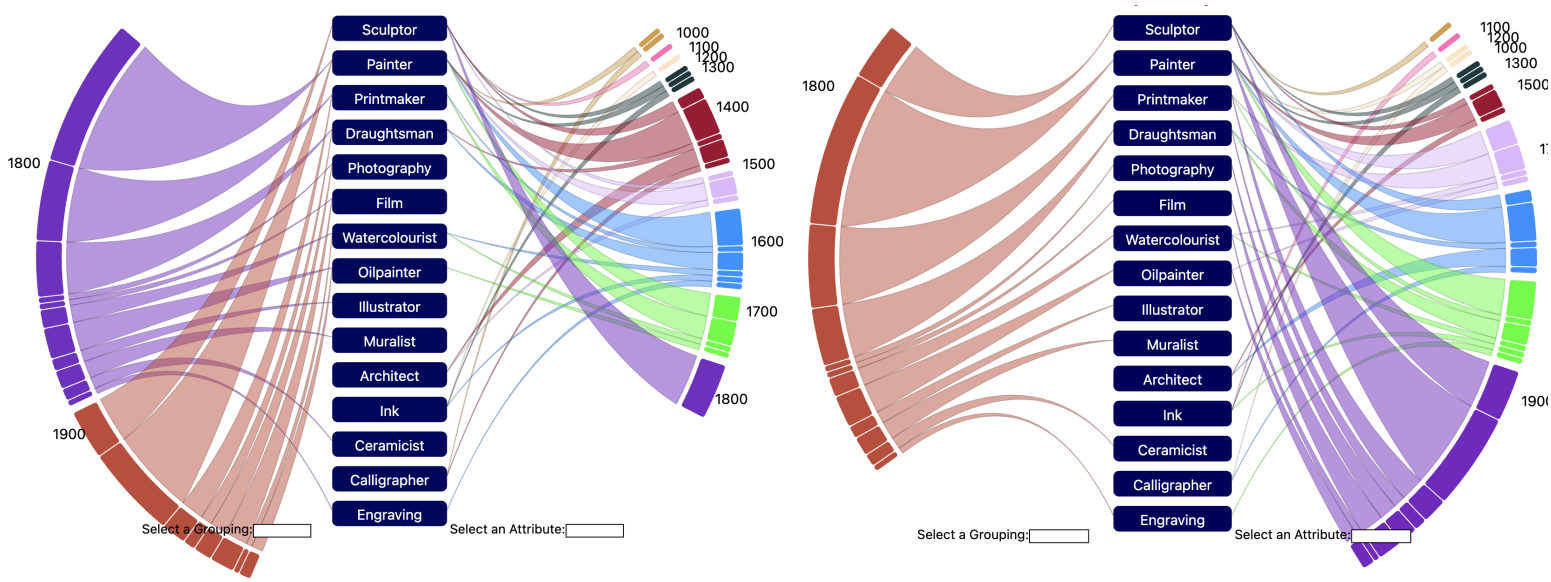


Figure 19. *Sorted vs unsorted chord visualizations with real data*

After analyzing figure 19's left image, we realized that to make the chart look better, we needed to sort the ribbons by the size of their groups. This led us to figure 19's right image. When tweaking the code for that version, we did not want to separate a group's ribbons and chords from each other. However, we decided in the end that maintaining symmetry was more important than keeping our groups close together at all costs. This is because we use both color and text as an encoding for that grouping anyway, and this happy medium is still sensical (see fig. 20).

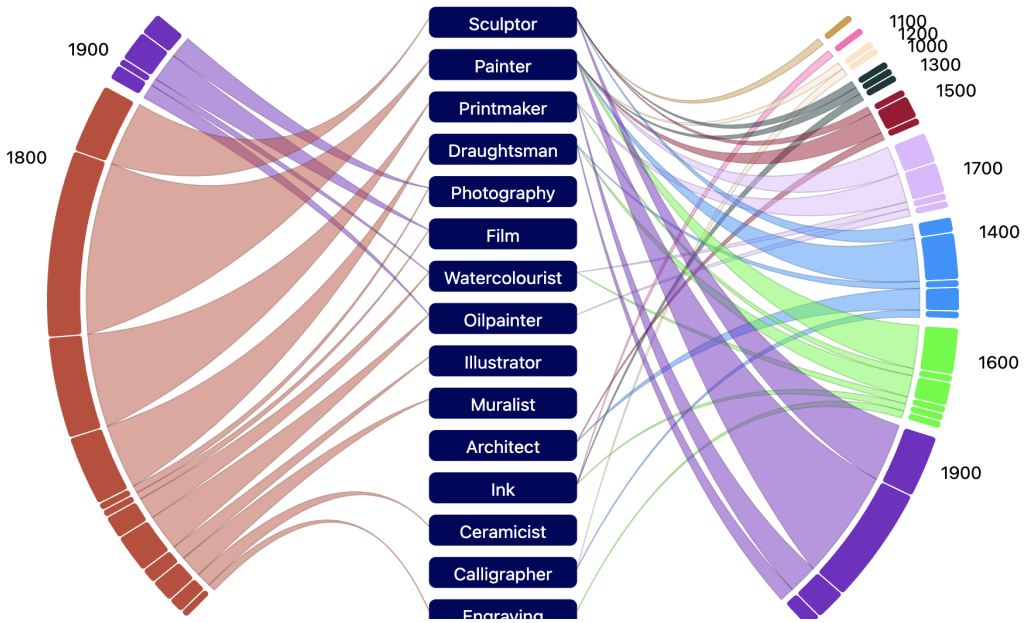


Figure 20. *Our balanced solution*

Briefly Exploring the Adjacency Matrix

Compared to our other final components, our adjacency matrix had the least amount of growth from exploratory data analysis. That is because compared to all of the other components, the adjacency matrix is by far the least complicated, and that is by design. In our initial renderings of the data this way (see fig. 21), one main thing became obvious: we needed sorting.

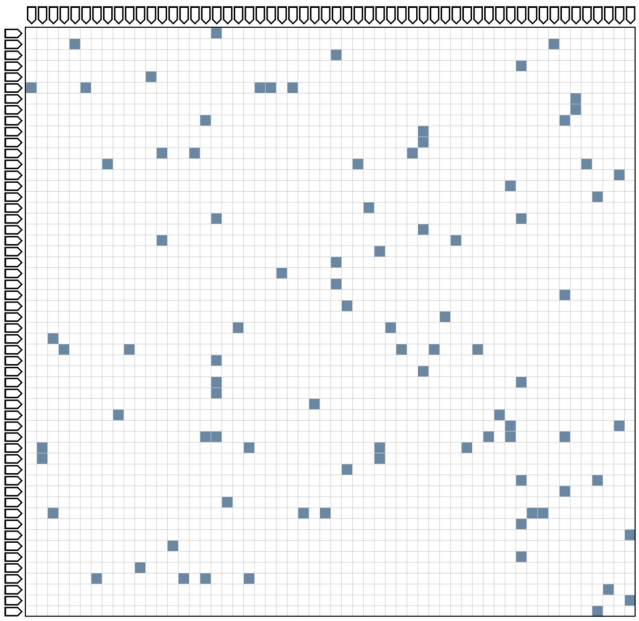


Figure 21. *Our early adjacency matrix with real data*

As we learned in class, sorting is a great addition for most adjacency matrices because it makes otherwise unseen trends become obvious. More specifically, we wanted to implement sorting based on the number of influences. That is what we show in this matrix, and thus sorting on any other data attribute would probably not be very helpful. If we sort by number of influences on the other hand, regions of the matrix become more interesting because that is what we are encoding by color.

Minding the Gap

The last meaningful insight we gained from early visualizations of our data was how apparent the temporal gap in our data set really was. For most of our work on this project, our two earliest artists temporally were separated by a few centuries. This was fine in our eyes until we got an early version of

our area chart (to convey artistic medium data with a chord visualization component). The result with all of our data was somewhat awkward (see fig. 22).

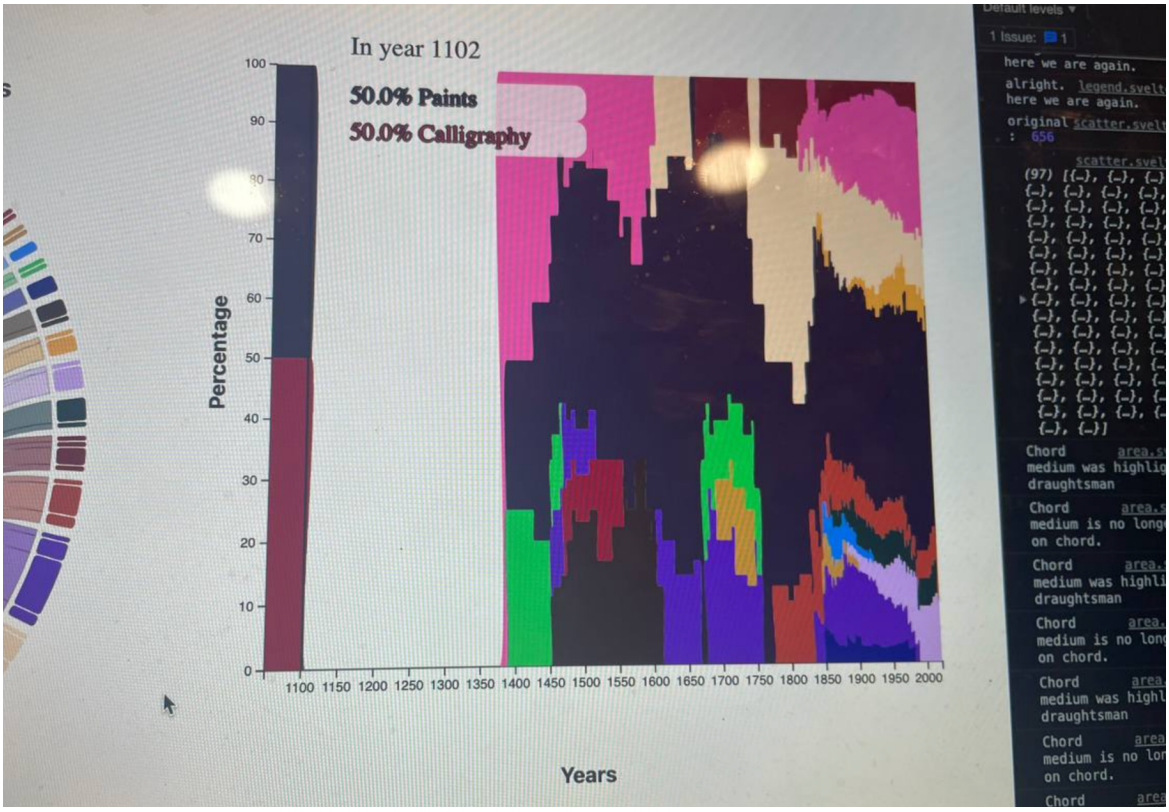


Figure 22. *The massive temporal gap in our data made apparent*

We had two methods of remedying this: we could have the area chart only show data from the 1300's onward, or we could simply get rid of the data from the 1000's. We decided to go with the latter solution because it would still feel awkward to just not show some of our data on a component like this, and this solution only involved removing one artist from our set (who practiced two artistic mediums, hence the appearance of fig. 22).

Design Evolution

Phase 1: Early Stages

Our elements that remained constant include a world map, timeline, arrows showing influences, and artist identifying markers, but the rest of the design has certainly changed. The first major development occurred before we wrote much code at all. We synthesized our first three sketches into a single agreed-upon design. Our first prototype emphasized the timeline a bit too much, and lacked filtering to avoid influence arrow clutter (see fig. 23). The second sketch carved out some good ideas for avoiding that clutter, but it did not utilize geospatial encoding well (see fig. 24). Then our third design emphasized a niche part of our data that we may not even have in the end (that being 'lost' artworks) (see fig. 25)

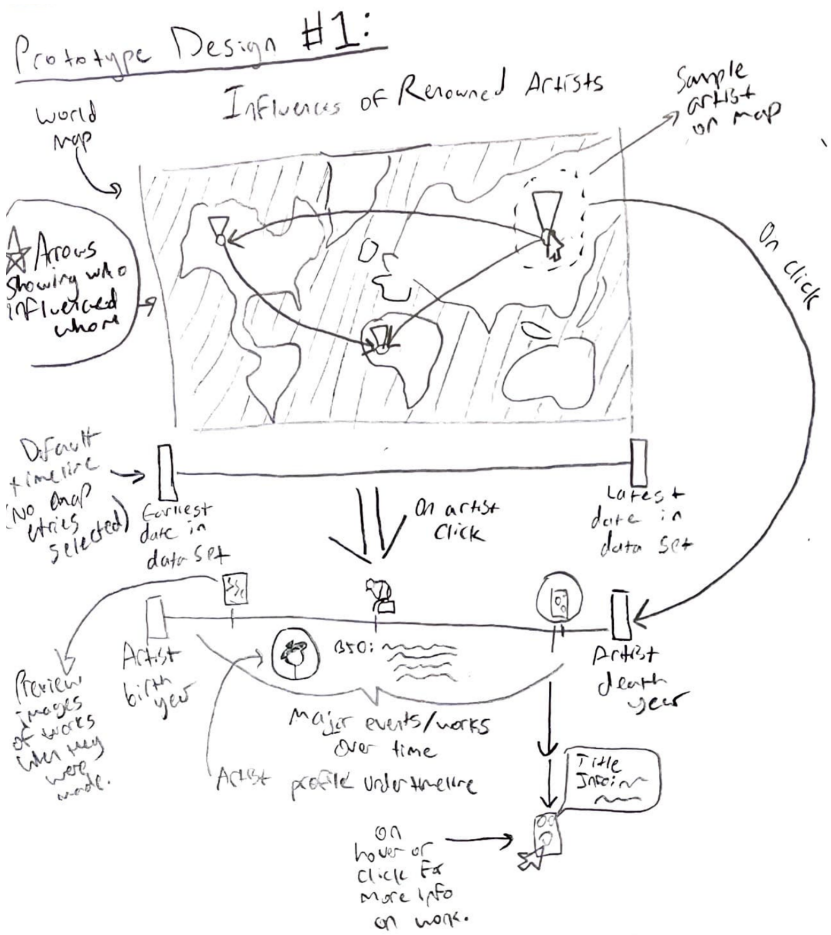


Figure 23. *Mockup design sketch number one.*

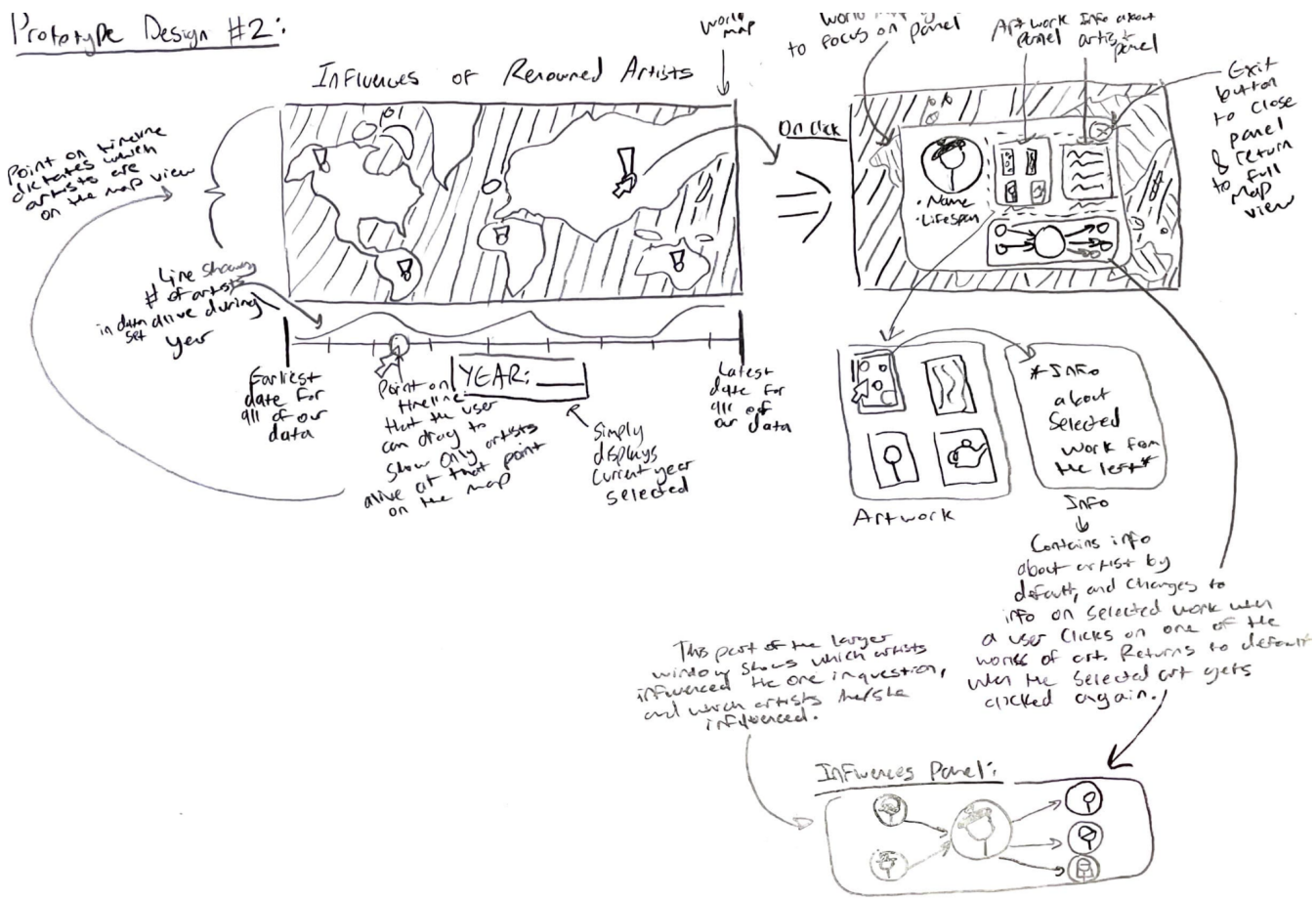


Figure 24. *Mockup design sketch number two.*

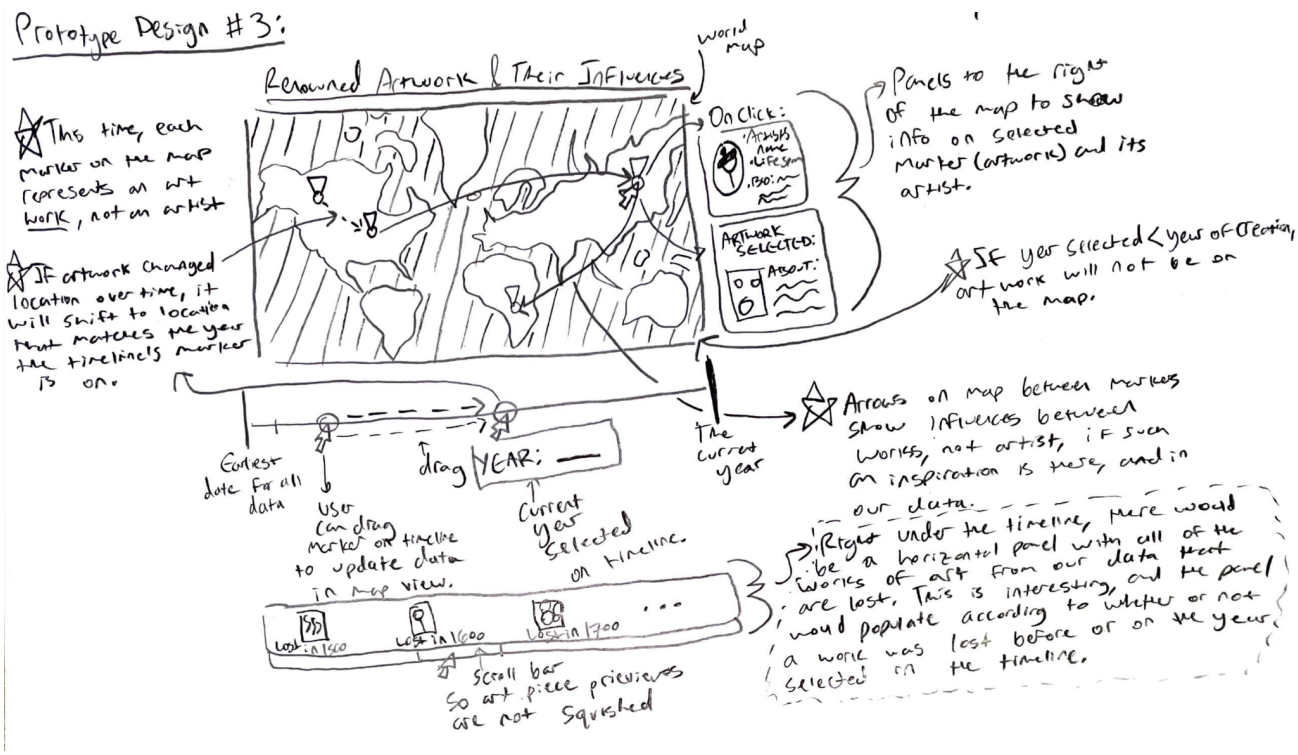


Figure 25. *Mockup design sketch number three*

The last major sketch we drew before starting to code leverages the best encodings from the 3 prior designs. Our world map view uses the fairly strong 2D positioning as a visual encoding for artist location. Plus, the fact that it is a world map effortlessly establishes context that everyone can infer. Additionally, the use of a world map conveys the actual distance that the influence between two artists spans. The orientation, or shape, of our influence arrows encode the kind of influential relationship between artists. The timeline uses 1D length (which is even more easily readable than 2D) to encode time well.

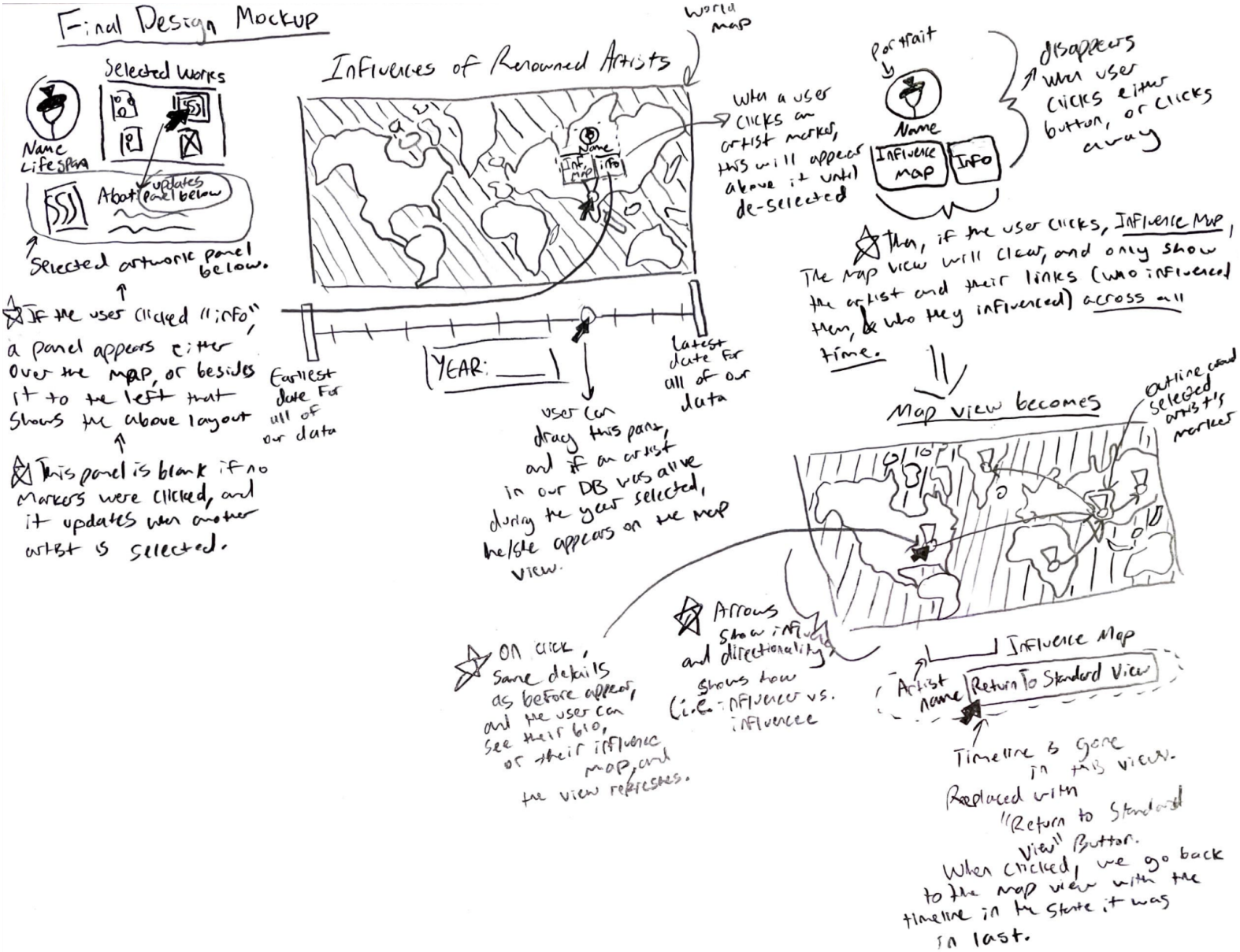


Figure 26. *Our initial "final" design*

Phase 2: Post-Proposal: Reworking What's There

After getting a little bit of code off the ground and scrutinizing our proposal's ideas with Pranav, we massively retooled our design. Starting small, we quickly sketched the idea of a D3 force physics simulation on our map view's nodes to avoid clutter (see fig. 27 and *The Map and the Old Timeline* for more details).

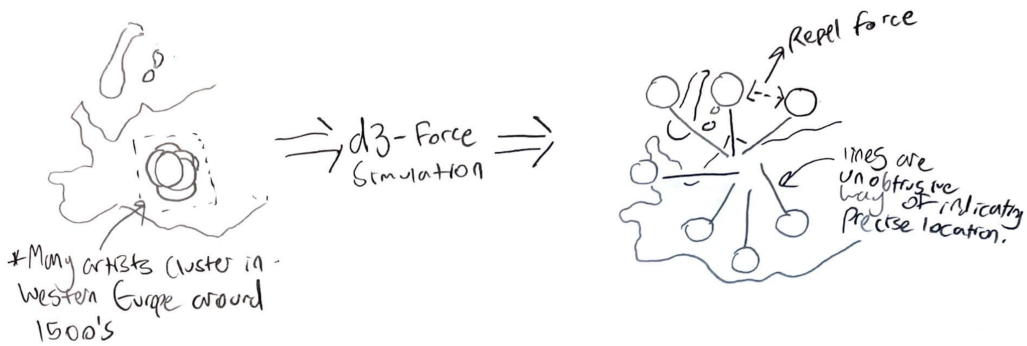


Figure 27. *What we conceptually wanted to do with our nodes and D3 physics simulations*

Another simple alteration that would enhance our design relates to the influence arrows between artists. In addition to using directionality to encode the type of influence relationship, we want to use the width of the arrows to encode the time an influence has traveled (see fig. 28).

Figure 28. *Our planned new encoding of time gap with arrow width*

This encoding would add depth to the influence map and even contribute identifiable trends. For example, if the arrows are wide, that conveys that the artist's popularity saw a resurgence long after death. This also eliminates the need for users to calculate differences in years themselves.

With those minor final design aspects addressed, we can now elaborate on our major design overhaul between our proposal and now. The updated overall design (see fig. 29) now has more views and some minor removals.

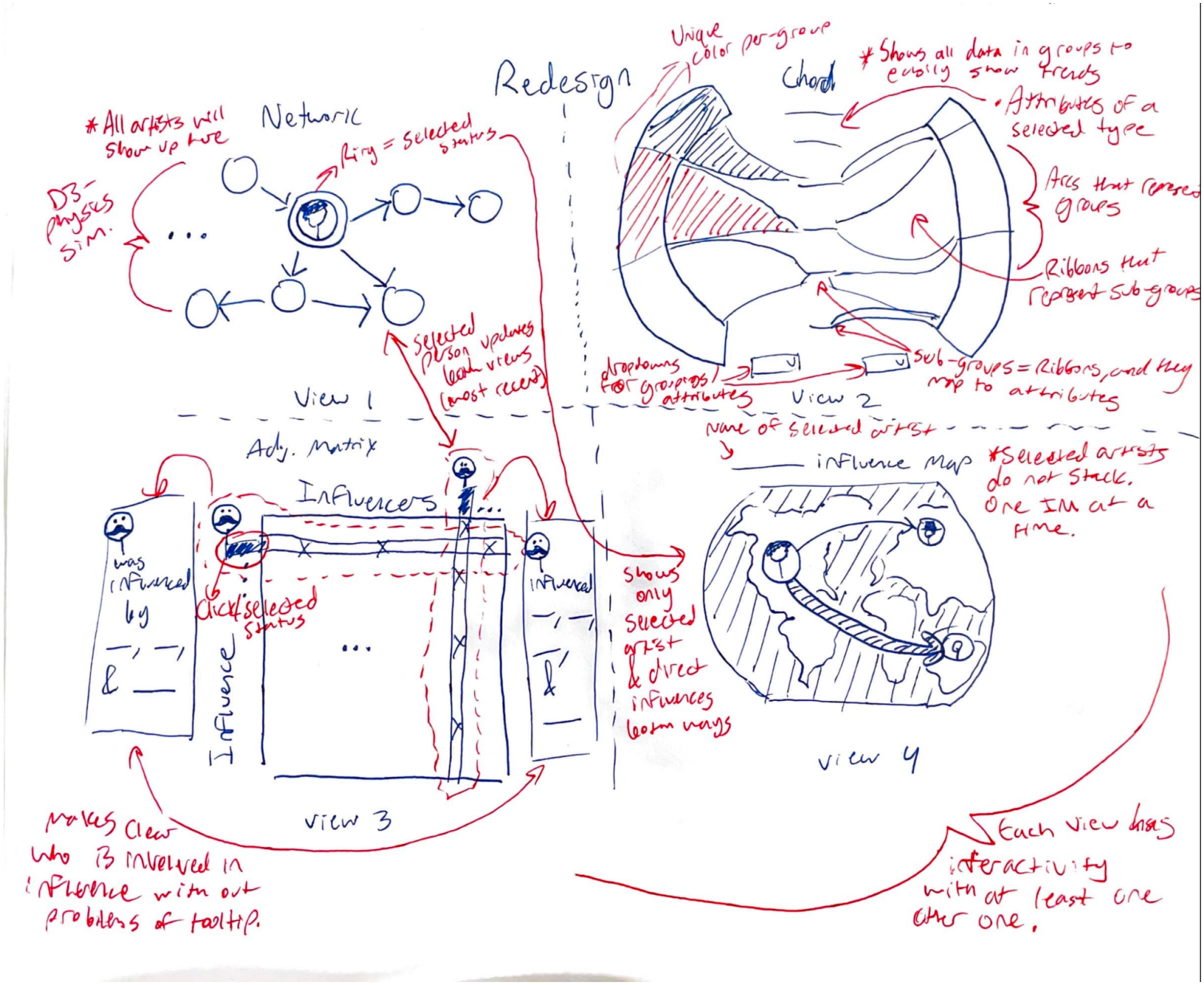


Figure 29. *Our last big design for this phase of development*

This redesign drops the aforementioned timeline component entirely, and opts for a more open-ended approach with greater variety compared to our proposal. Meanwhile, the previous map view is now split in two: the network view and the influence map view specifically. This way we do not have to

worry nearly as much about node spacing (see *The Map and the Old Timeline*), and we have essentially modularized the complex map into a simple high level and a low level view. The former shows all of the data with ideal spacing, and the latter zooms in on the nature of a particular person's influences temporally and geographically.

Additionally, the chord component provides both more color and another kind of high-level overview. Meanwhile, the adjacency matrix provides both a high-level view by showing all of the data, and an individual view through its row and column highlighting. This overall lends itself much better to user exploration, and the level of interactivity we were aiming for here makes all of the components more connected in a way that complements them. However, it is still worth going over the designs of our brand new components more specifically.

Phase 2: Post-Proposal: Brand New Parts

Back when we finished our proposal's design, we were fully prepared to devote a lot of time to conveying information about both the artists' artworks and their biographies. In light of new design ideas, we have redirected the time that we would have spent on that to exciting brand new visualizations. The first of these brand new views is our chord visualization. Our intention with this one is to provide a complete overview of our data in an aesthetically pleasing and unique way. One can see how we intended on accomplishing both of those things by looking at our fleshed out design for it (see fig. 30).

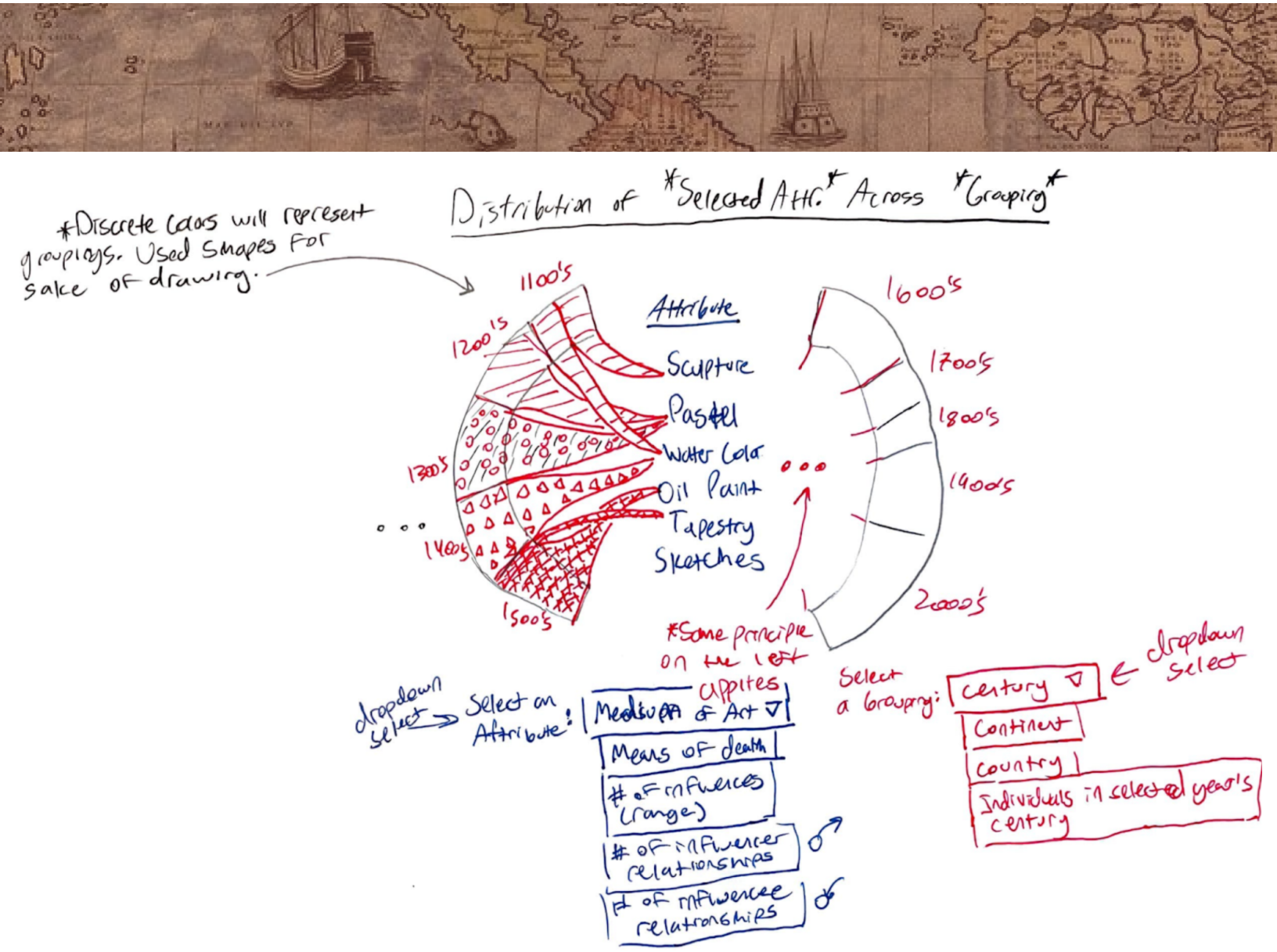


Figure 30. *An early sketch that reflects both the core and some eventually abandoned ideas for the chord component*

A large part of the chord component's design is color. Color serves to encode each discrete grouping, and we want to choose colors that would make something nice to look at. Meanwhile, the ribbons in the visualization, that connect the arcs to their respective attributes, were a specific choice because of how stylish and effective they are. In relation to their effectiveness, they utilize 2D size and positioning, which are fairly strong encodings. Additionally, shape is a key encoding here, and curvature more specifically easily conveys trends among groups in how they map to the center.

Something else unique about the chord visualization is the kind of data that it shows. At this point, we had a lot of ideas for potential data to show here (see mock dropdowns for mixing and matching possibilities in fig. 30), but in the end, we ditched the mix and match idea for something more simple.

That being encoding the century in which the artist was born in the center, and the artists' mediums or means of death by the arcs on the perimeter and their ribbons. At this point in time, we planned on just having one chord visualization that you could change in terms of data being shown, but in the end we decided on having two that show distinct data (see [Phase 3: Final Designs: Big Picture]() for more details).

Our second and last brand new designed component from this phase of design is our adjacency matrix. While it is more simple in some ways compared to the other components, we actually managed to get creative in a few ways during the sketch process (see fig. 31).

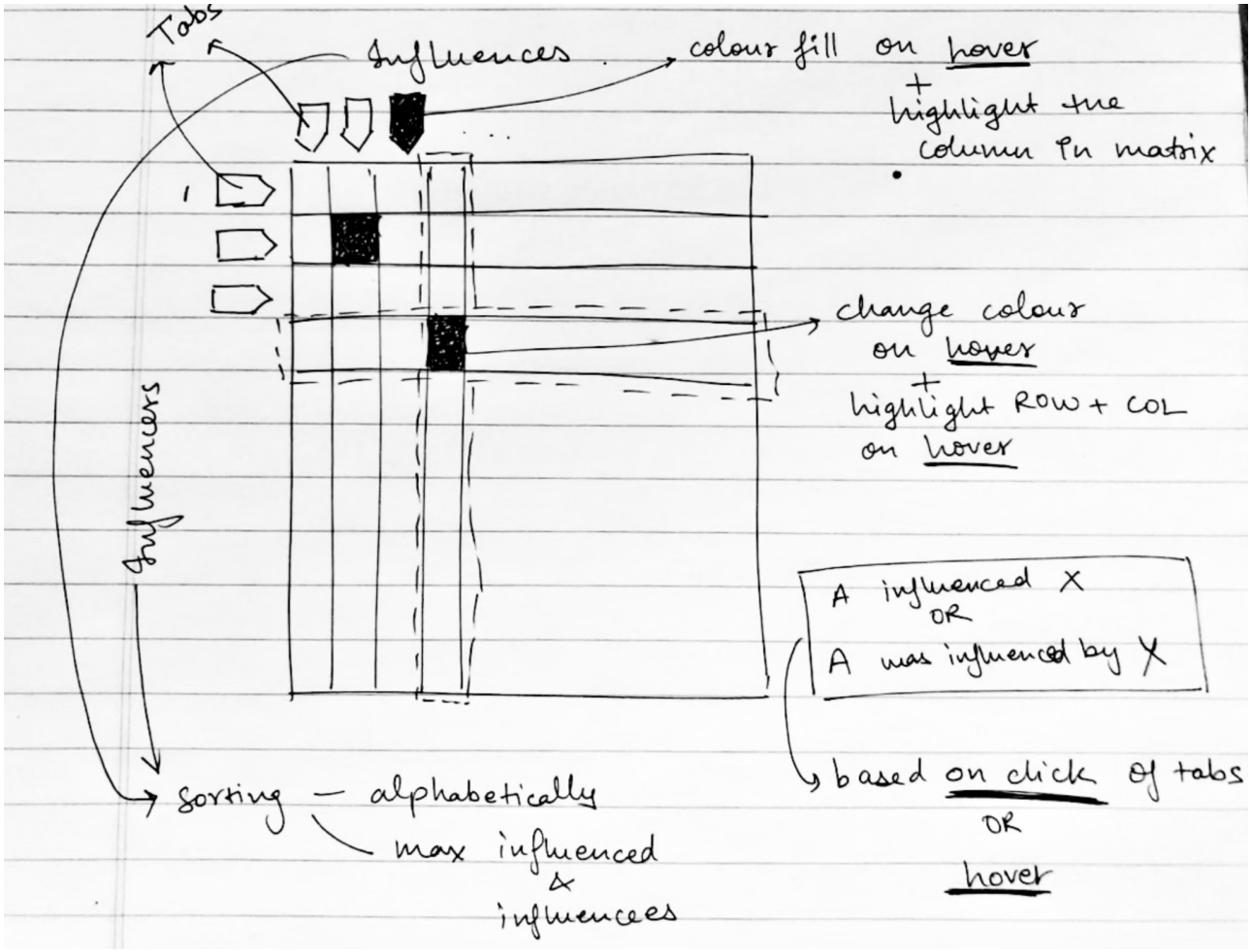


Figure 31. *Our first fleshed out concept sketch for the adjacency matrix*

Our design has many of the standard features of an adjacency matrix, and those fit our influence data well. One can see both the people who influenced somebody and those whom they influenced after looking at the person's row and column respectively. We designed this identification to be easy with color as an encoding for a selected cell or an entirely selected row or column.

However, the place where we get really creative is with the tooltip, or lack thereof. When contemplating how to show the information behind a cell explicitly, our minds immediately went to tooltips, but we thought of a less fleeting way to do the same thing. We designed this view to have one or two boxes off to the side that show the information behind a cell or a row and column combination when clicked. This allows for more calm analysis on the end of the users, and giving them the ability to sort the matrix can allow for easy insights.

Phase 3: Final Designs: Big Picture

After encountering some of our last bits of inspiration, we finally carved out the last design using components that we largely already had implemented. We realized that our last agreed-upon design (see fig. 29) was too busy. Everything was on one page, and we questioned whether or not it was the best approach. Dr. Lex helped put this into perspective when he pointed out how jarring it would be to see influence-related data, and then fairly depressing statistics about artists who committed suicide for example. It was a sort of tonal whiplash, and we remedied it with multiple pages and some rebranding.

This point marked the end of *Influences of Renowned Artists* and the beginning of *Influences, Deaths, and Mediums of Renowned Artists*. While we were already showing our death and medium-related data, it was always second fiddle to the influence data. The visualizations showing that were the star, and everything else was sort of just icing. Based on the aforementioned feedback, we changed our design to incorporate three different pages that emphasize influences, deaths, and mediums of our artists equally. The big picture for the design (see figs. 32, 33, and 34) is something that required extra work to implement, but made more sense than our previous “final” design.

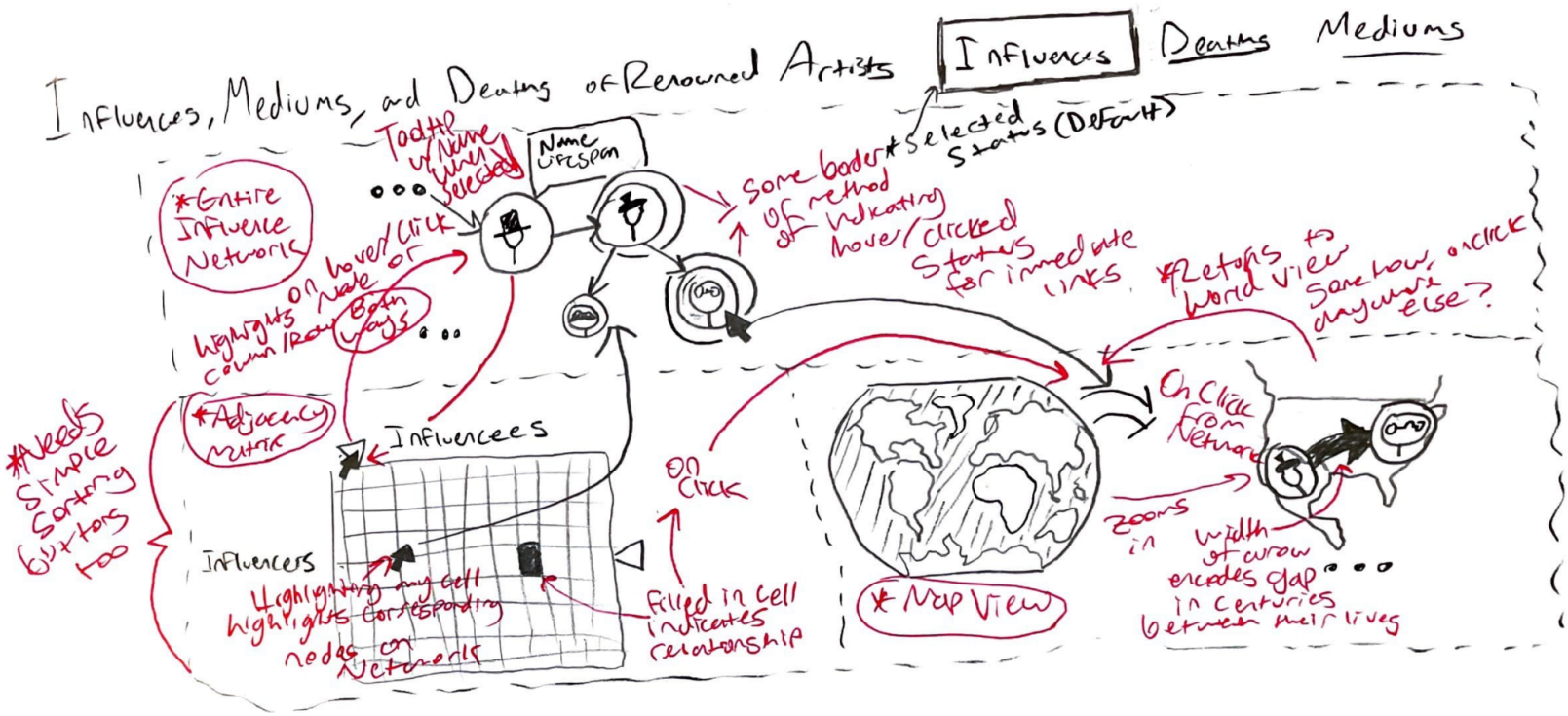


Figure 32. *Our detailed concept of our Influences Tab*

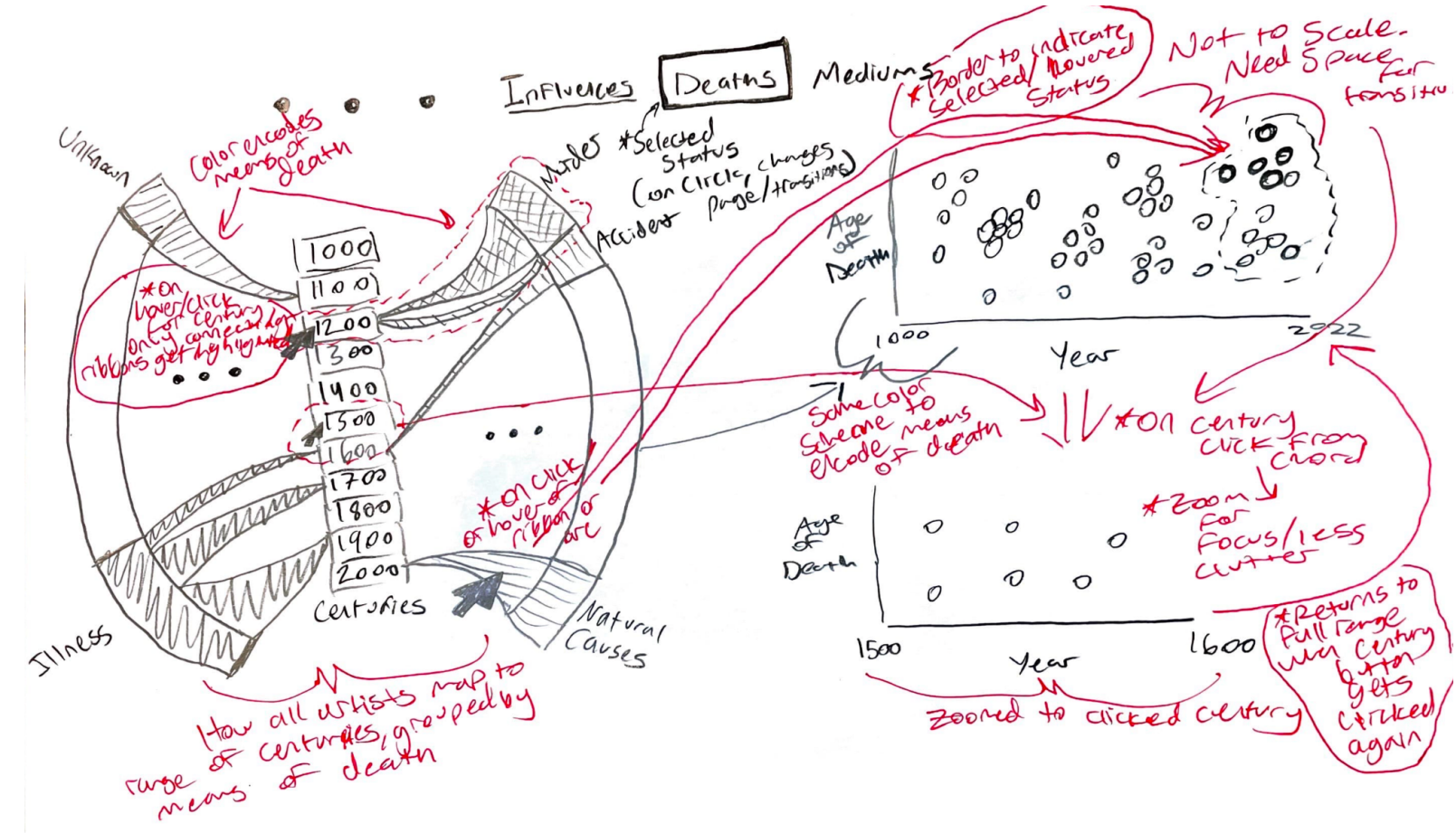


Figure 33. *Our detailed concept for our Deaths Tab*

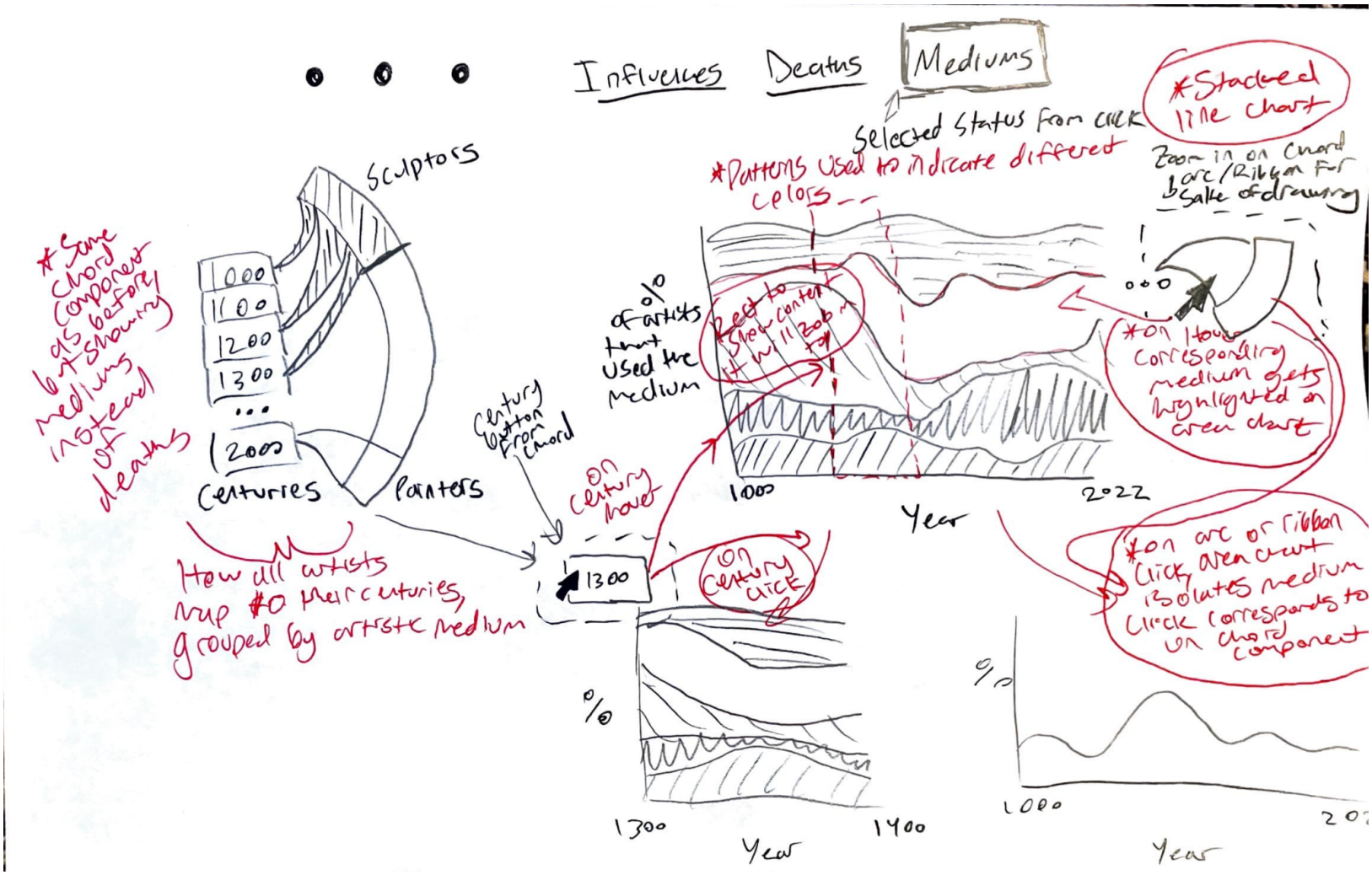


Figure 34. *Our detailed concept for our Mediums Tab*

This modularization of our work allows a user to focus more deeply on one facet of our data. Whether that be influences, means of death, or artistic medium. Some minor exclusions include the lack of dropdowns for the chord component and the lack of text boxes for the adjacency matrix. The former would no longer make sense due to how each page is only supposed to show a certain kind of data, and the latter was a fairly redundant encoding due to the tooltip for the artist network nodes (see top half of fig. 32).

The biggest inclusions for this design iteration are the scatter plot and the stacked area chart. After deciding to split up our three types of data like this, we needed something more to fill up our pages on deaths and mediums while seeming like a natural fit and conveying more information.

Phase 3: Final Designs: Traditional Tools

The scatter plot is a simple chart, but it pairs nicely with our death data chord visualization. It provides two views with different strong suits. The chord visualization is more aesthetically pleasing and organized, but a scatter plot uses 2D space as strong encodings. We figured quantitative data would take advantage of that encoding the best, and death-related data lends itself naturally to that via dates and ages at which an artist died. Then, a great way to make those two views feel more connected is to encode color in the same way for both of them. The fill of the dots and the ribbons both indicate the type of death. Further, these views are intended to feel inseparable through the focusing and axis zooming functionality (more on that in *Implementation*).

As for the stacked area chart on the view for artistic mediums, we are going for a colorful visualization that avoids blank space, unlike the scatter plot. The fact that this area chart's y-axis is percentage based means that it will always be filled with some color. More technically, this uses 2D area as an encoding for the percentage of artists in a given year that were both alive and used a given medium. This is a pretty straightforward encoding for quantitative values like this, and it can even lead to some interesting shapes without compromising information. We intended to also link this component with its view's chord visualization through identical color encoding for the mediums themselves. x-axis zooming is also on the table, and highlighting an area is a fairly clear parallel to highlighting scatter plot dots based on the chord component's ribbons.

Adding these two components this late into the process was feasible for us due to their inherent simplicity compared to our other components that we already made. Plus in hindsight, they both seem like perfect fits and provide some aforementioned traditional visualization strengths to our somewhat non-traditional assortment.

Implementation

First Tab: Influences

It seems logical to discuss our implementation as a whole by focusing on one tab at a time. Upon loading our final product, which should only take a few seconds, the default tab selected is our influence tab, and thus the first view you should see is the one we made for the influence data (see fig. 35). Firstly, we implemented a D3 physics force simulation on all of the nodes in the top network component. There is no animation for it, but this ensures good spacing. The positioning will thus be different each time a user loads the entire visualization web page, but this does not affect anything negatively, and even adds some slight visual variety.

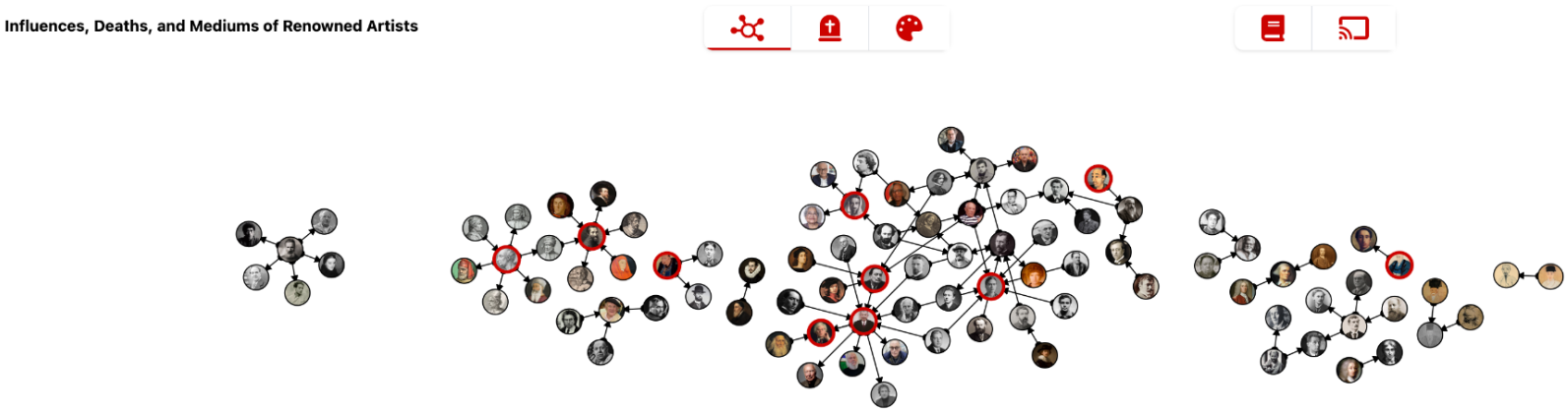


Figure 35. *The top half of our Influences Tab (the first tab of 3)*

The first bit of interactivity for the network component, in and of itself, is the node's hovering capability. When hovering over any artist node, it increases in size and a tooltip appears that provides the artist's identity (see fig. 36). After exiting the node, it simply returns to normal and the text vanishes.

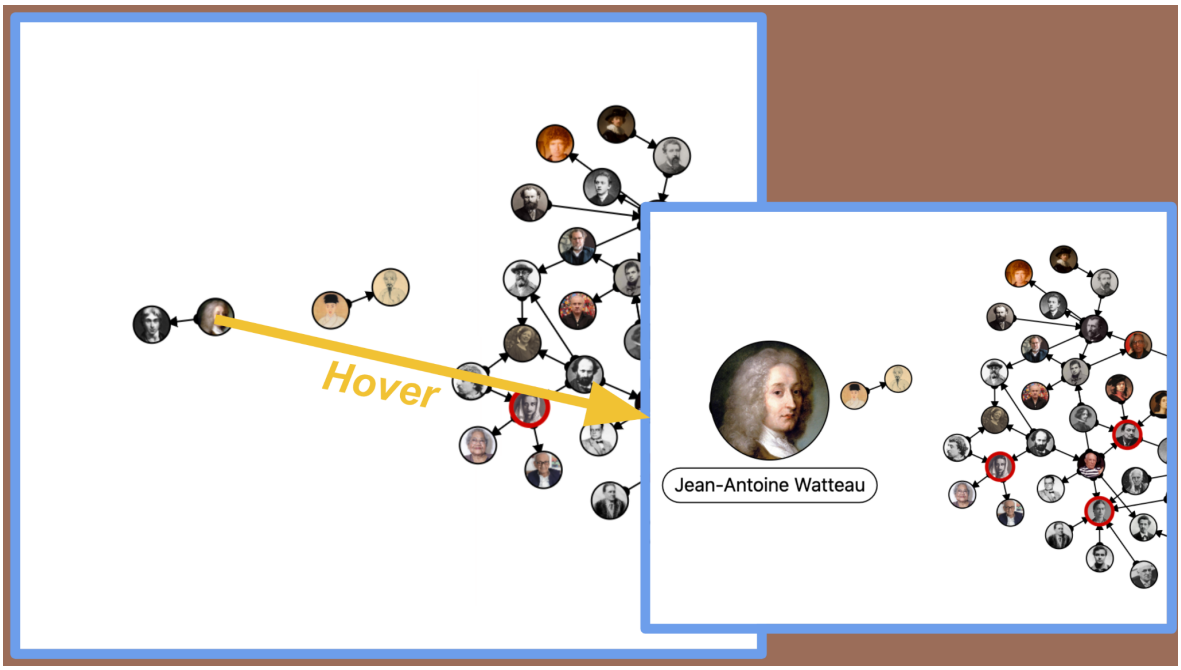


Figure 36. *A simple visual progression between hover and regular states for each network node*

On the topic of mouse hovering, if you notice the nodes that have a red outline, that means that hovering over them will convey a little piece of storytelling relevant to the artist (see fig. 37).

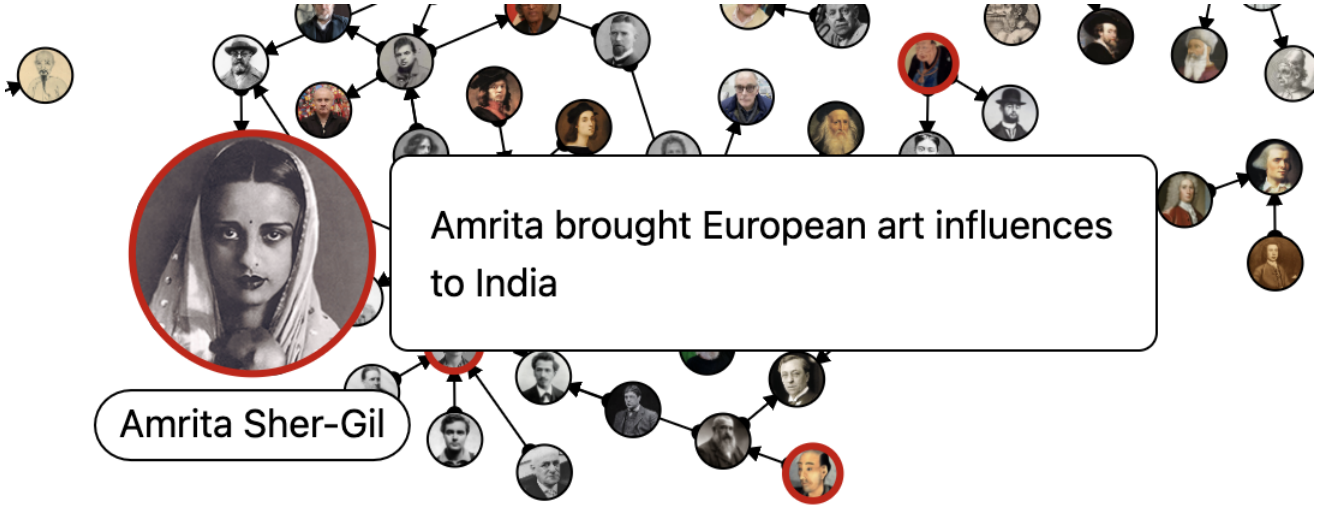


Figure 37. *An example of our bits of storytelling using the red border on our nodes*

Clicking is where our functionality becomes a bit more interesting. For the network component alone, clicking an artist's node will bring them, and their direct influences, into focus through a colored area (see fig. 38).

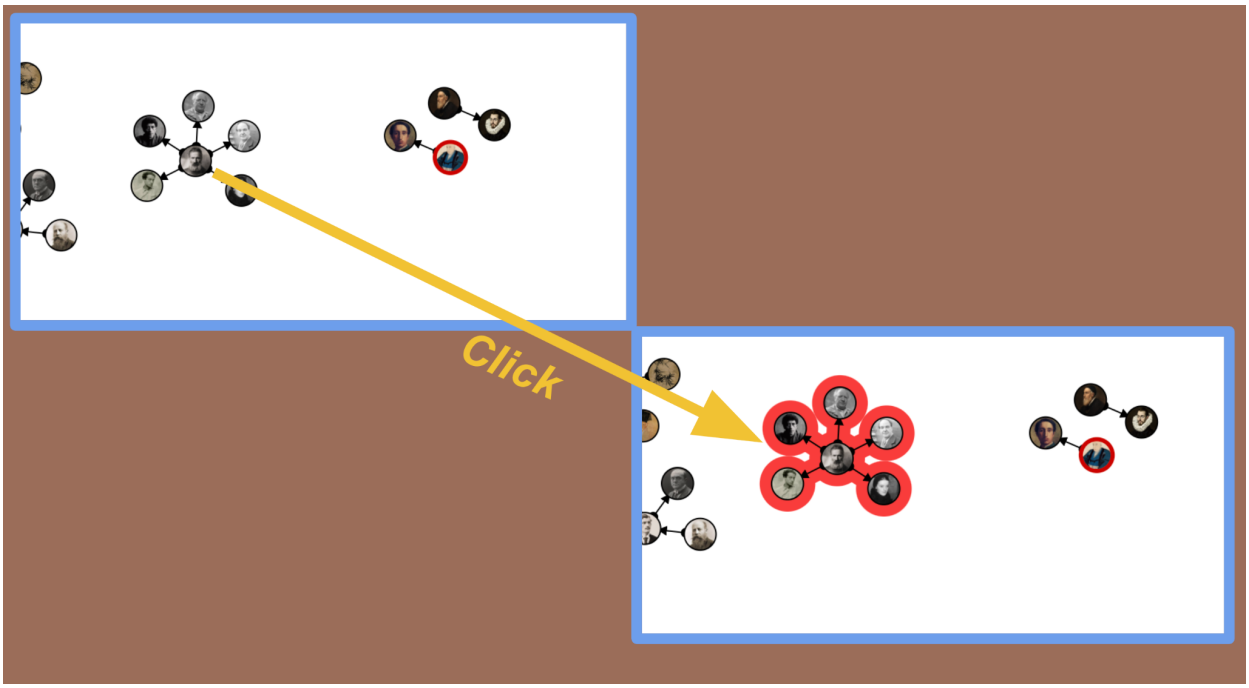


Figure 38. *Similar to figure 36, this just shows the progression between the default state and clicking a node*

Additionally, and more interestingly, the other two components on this tab update accordingly to reflect the selected cluster of direct influences (see fig. 39).

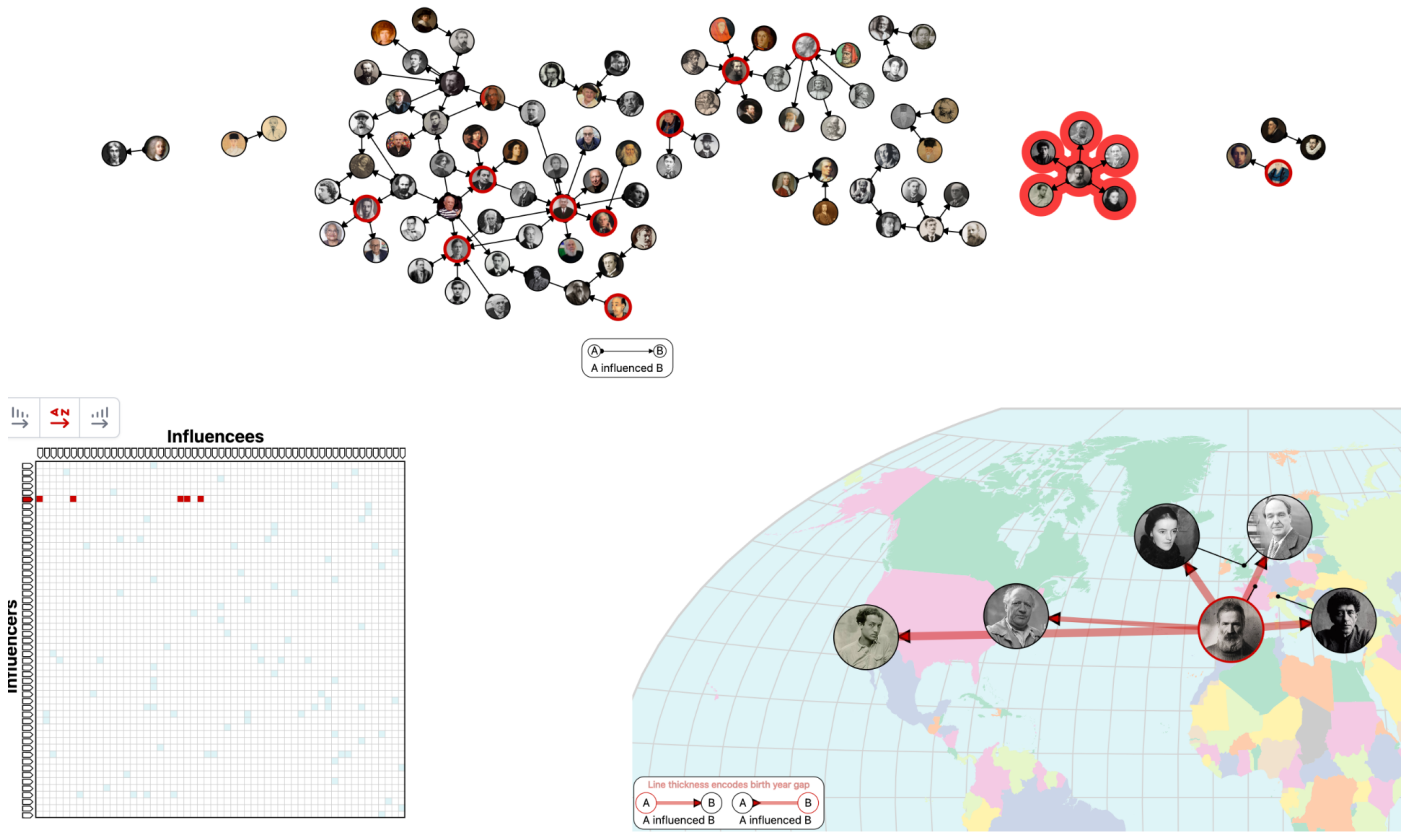


Figure 39. *Every component on the Influences Tab interacting with each other after a click*

The adjacency matrix on the bottom left corner updates to highlight the row that represents the artists that you just clicked on. In this case, we clicked on an influencer among our data. If instead we clicked on an influence (i.e. a sort of leaf node with no artists that they influenced, a column reflecting them would highlight instead (see fig. 40).

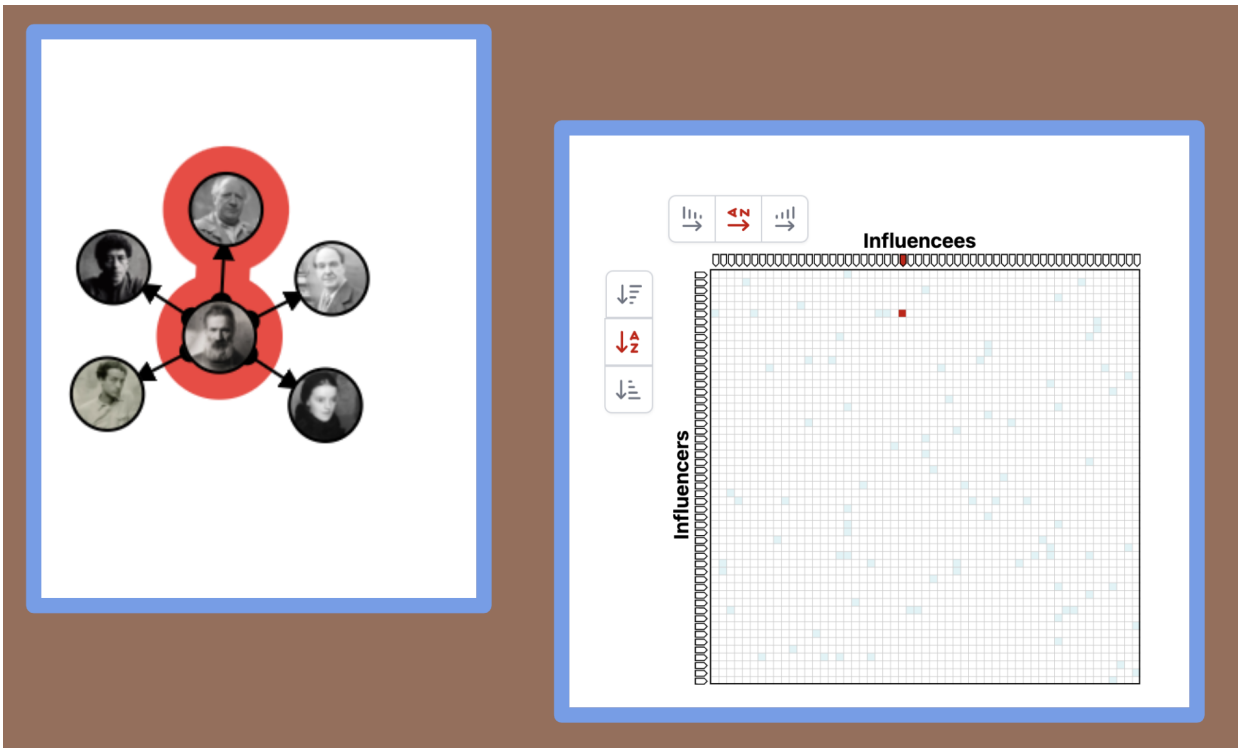


Figure 40. *A highlighting of how when you select a node on the network, its corresponding column highlights on the adjacency matrix (see red dots)*

Simultaneously, the world map view has been updated to zoom in on the parts of the world that the selected artist and their direct influences map to geographically. Clicking anywhere on this map or the network components will return the world map to the default full view. Further, on that map view, one can hover over the nodes there in order to get information about what country each underlying artist is from and their lifespan (see fig. 41).

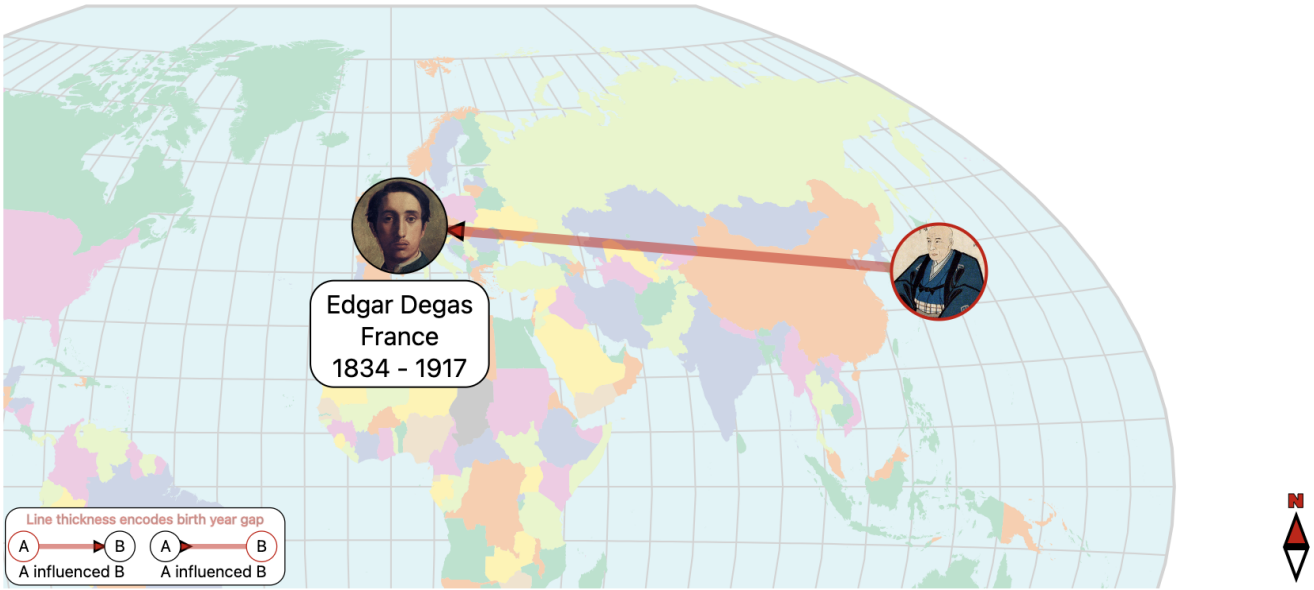


Figure 41. *An example of how you can get specific textual information on the influence world map*

As for the adjacency matrix, it is highly interactive with the network component. We implemented highlighting on a cell level, and so the artists associated with that cell on either axis get highlighted on the network (see fig. 42).

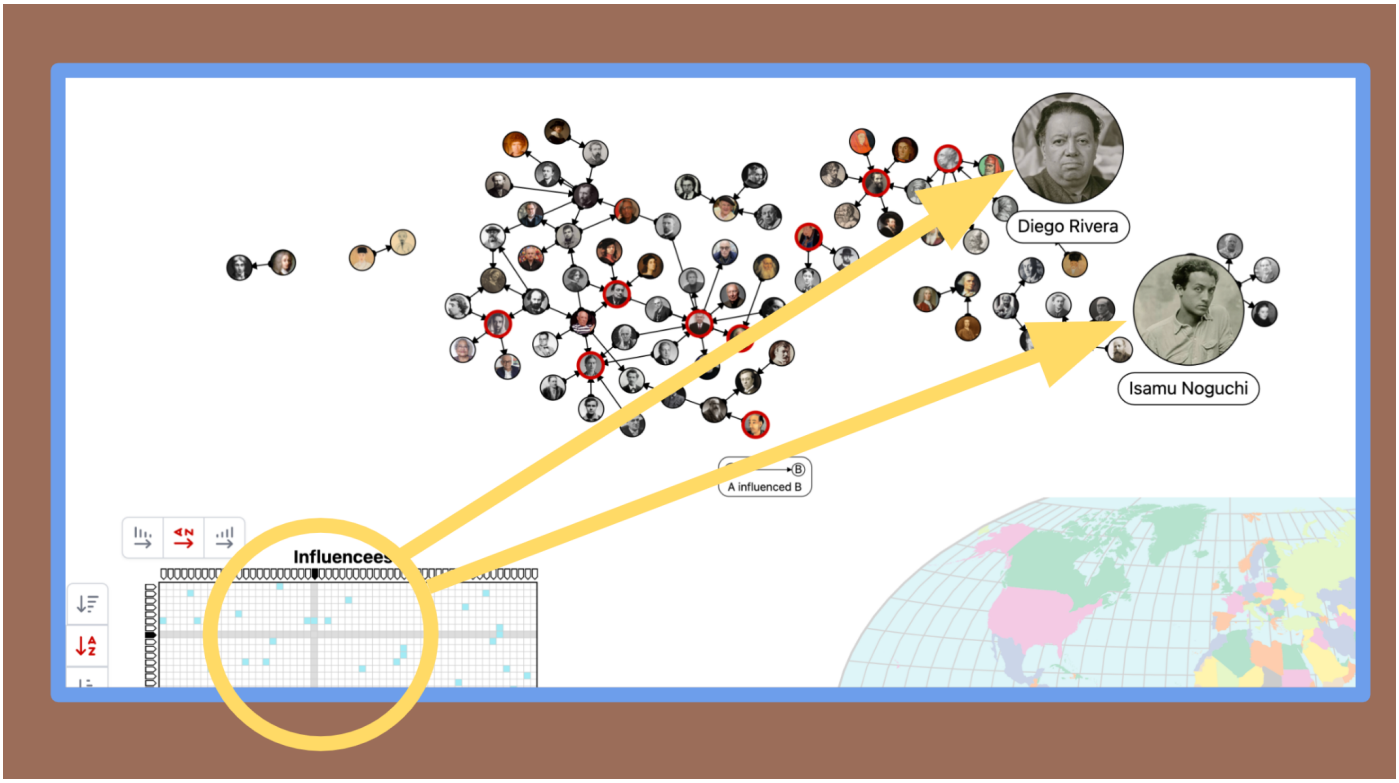


Figure 42. *An annotated screenshot highlighting how hovering on the adjacency matrix focuses the appropriate nodes on the network view*

Something similar can be said about the tabs on the top and left sides as well. Hovering over them highlights only the single artists associated with that tab's row or column (see fig. 43).

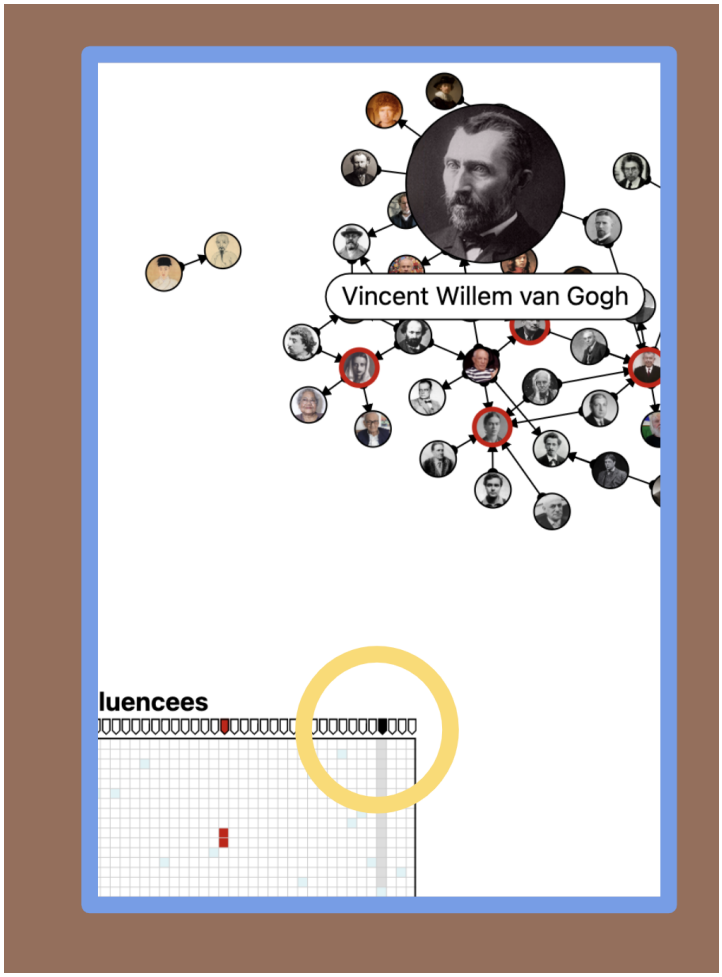


Figure 43. *An example of something similar to figure 42, but with the adjacency matrix's tabs*

Lastly, when a user clicks on a cell that is filled in (i.e. it represents an influence relationship), the map and network components update in the same way that they all do when a node on the network gets clicked, similar to figure 39.

Second Tab: Deaths

After switching to the death tab (indicated on the top by the gravestone icon), one can see three very closely connected visualizations: a chord chart, a scatter plot, and a legend for both of them (see fig. 44).

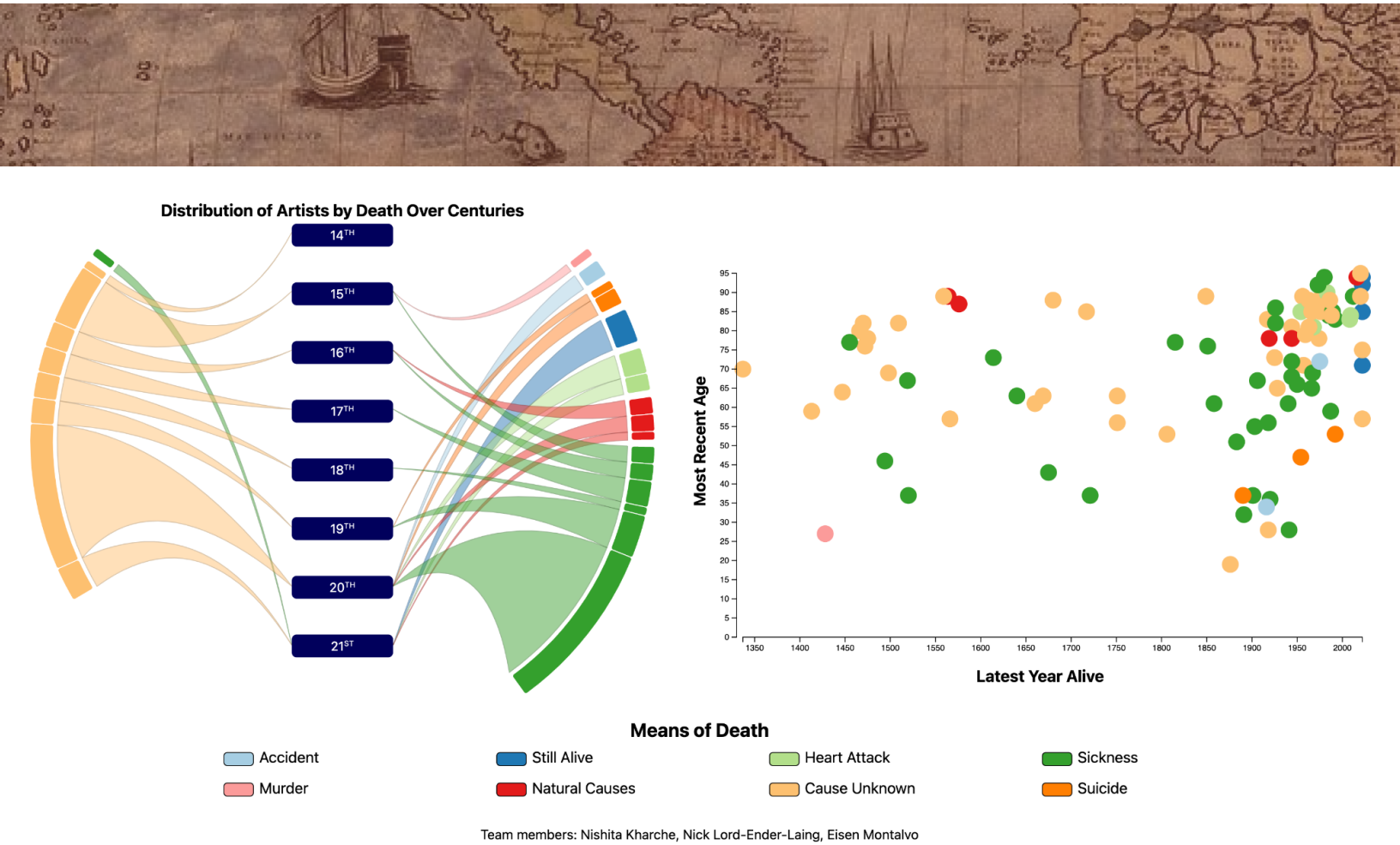


Figure 44. *The default view when switching to the Deaths Tab*

The legend is a new very minor component that only serves a couple of purposes. Firstly, it plainly conveys to a user which colors correspond to which means of death on both the chord and scatter plot components. More interestingly, if you hover over any of the legend's items, both charts filter to just that type of death (see fig. 45)

Figure 45. *An example of filtering by death type*

This type of filtering by death type is also a functionality of the chord component. If one hovers over the arc that represents a death type, the exact same filtering as figure 46 occurs. However, if one goes deeper into this component, they will see how highlighting an arc's individual ribbon highlights only the dots on the scatter plot associated with it (see fig. 47).

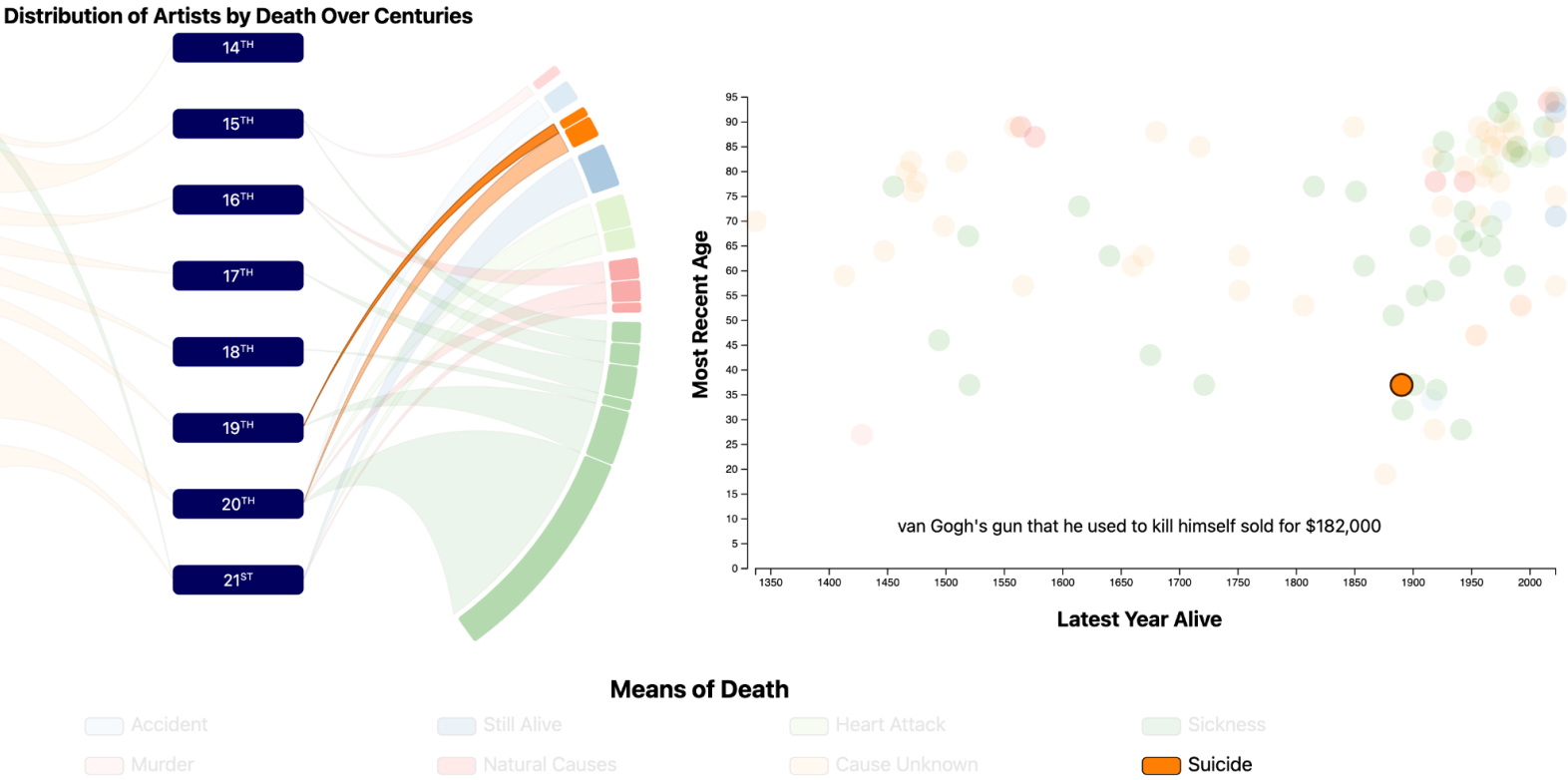


Figure 47. *A narrow example of how hovering over a ribbon affects the legend and the scatter plot by filtering them. Also notice the storytelling bit.*

In figure 47, because that example's ribbon only has one artist, van Gogh, we also coded in an interesting fact about his death. This type of storytelling appears on the scatter plot similarly for every other ribbon that only encapsulates one artist (try and find the rest!).

Perhaps the most interesting interaction between components on this page has to do with the century buttons in the middle of the chord chart. When hovered over, both the scatter plot and the ribbons filter to focus on those connected to that century (see fig. 48).

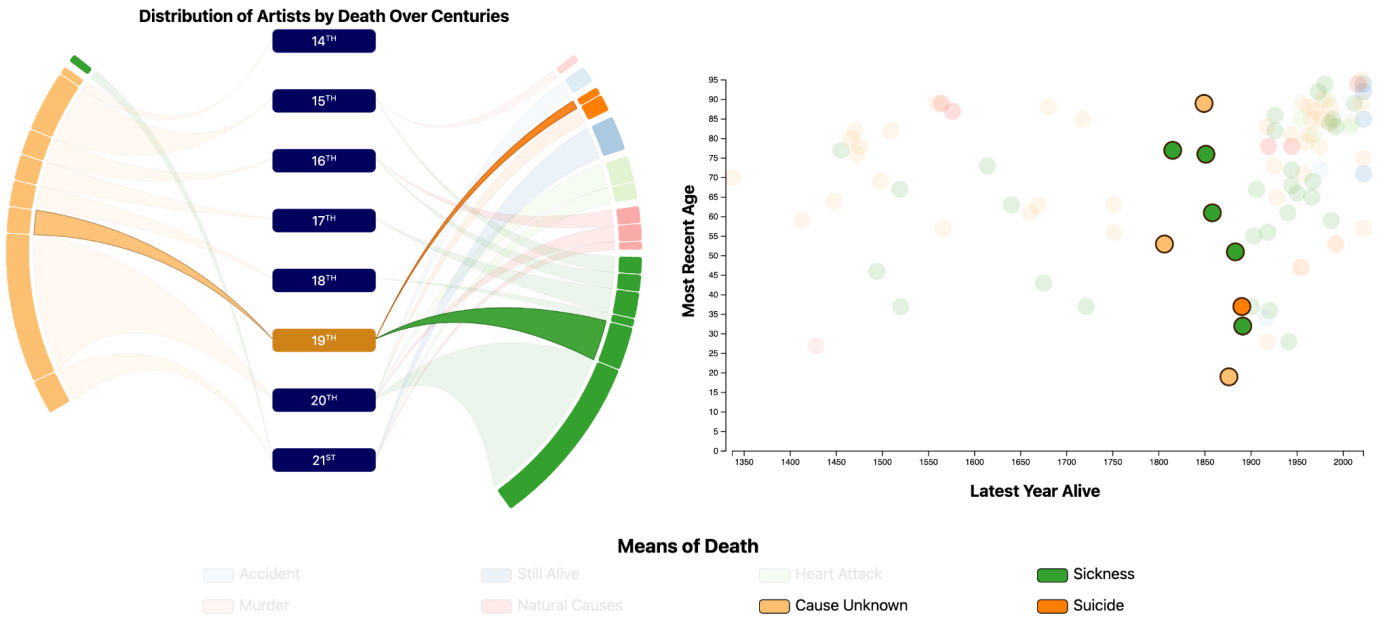


Figure 48. *An example of the century preview highlighting*

When actually clicking one of the century buttons, the scatter plot zooms in to only that selected century (see fig. 49). We even implemented a fluid transition animation that looks really good.

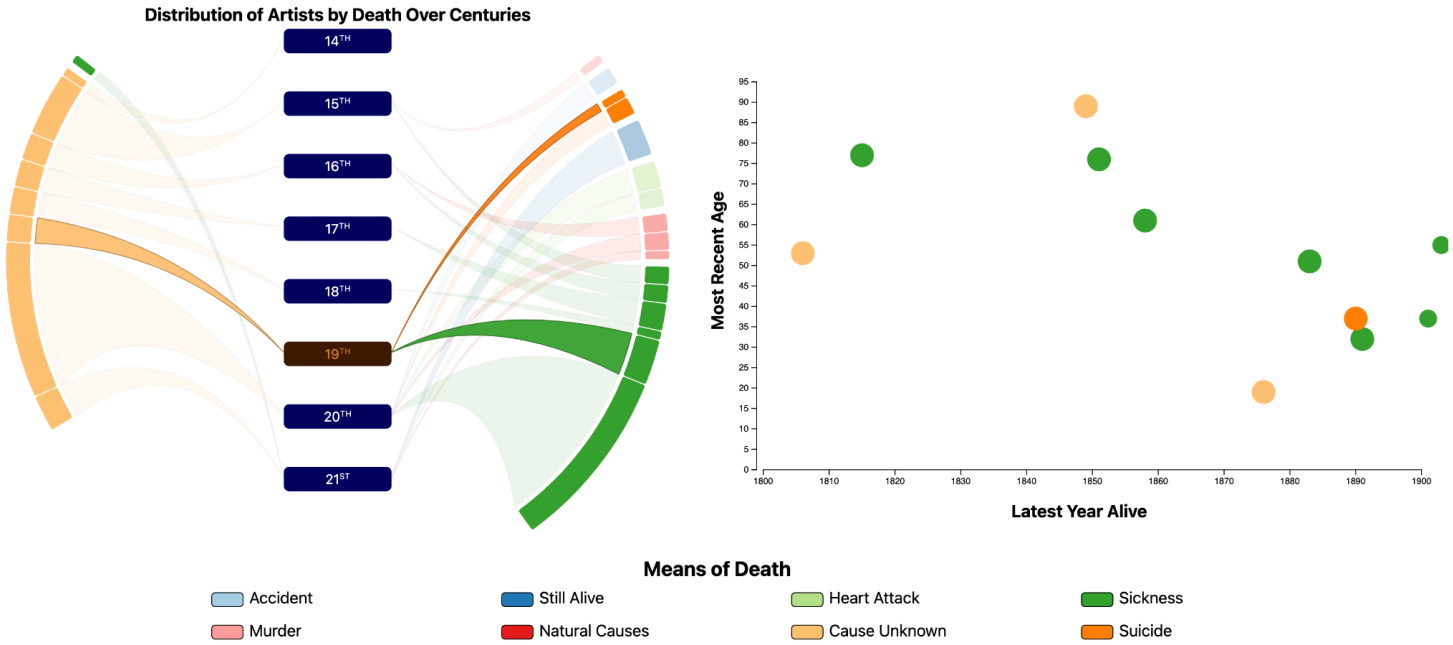


Figure 49. *An example of what happens when a user clicks a century, in this case, the 19th century*

This adds a lot to focusing one's viewing experience, and the state of the page stays this way until the user de-selects that century or clicks on a new one.

Another lasting state change that a user can cause is focusing on dots of a type of death. If a user clicks one of the chord component's arcs, the points that correspond to that death type have a permanent border around them until the arc gets deselected by another click (see fig. 50).

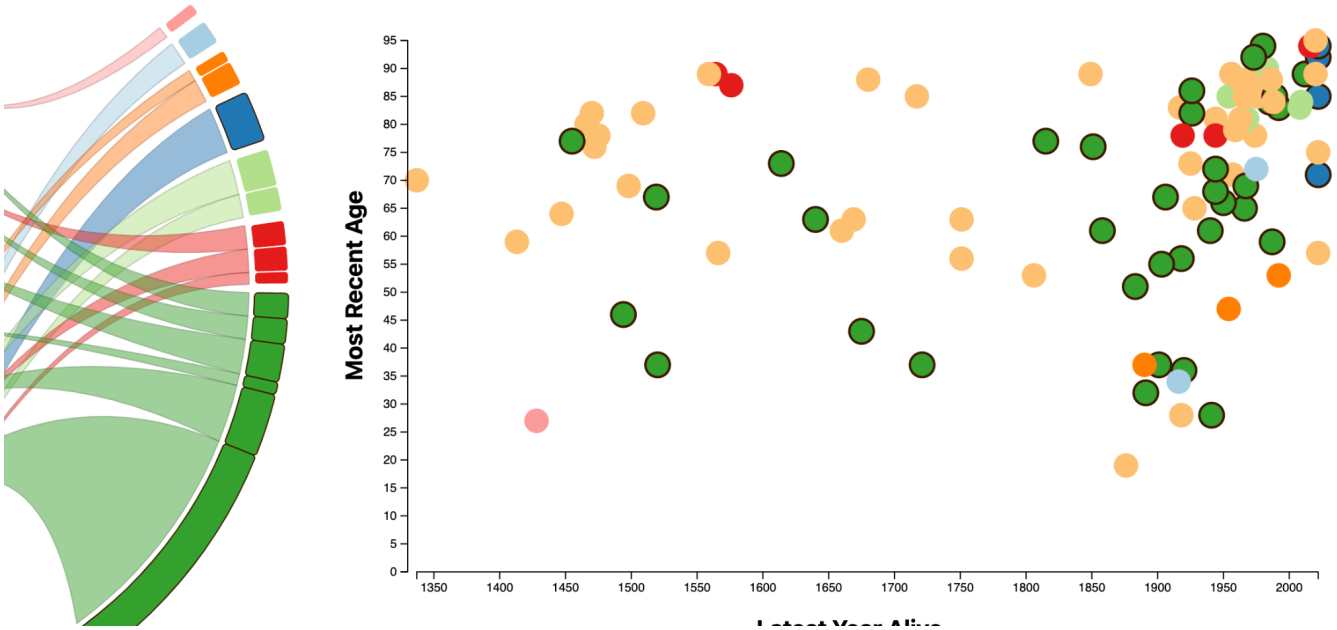


Figure 50. *A zoomed in look at how clicking on arcs keeps their corresponding points bordered*

Also, one final bit of functionality we implemented is a simplistic tooltip that lets a user know which artist they are looking at when hovering over a dot (see fig. 51). Overall, the chord chart is the steering wheel for this view, but this tooltip adds good information that is not there whatsoever.

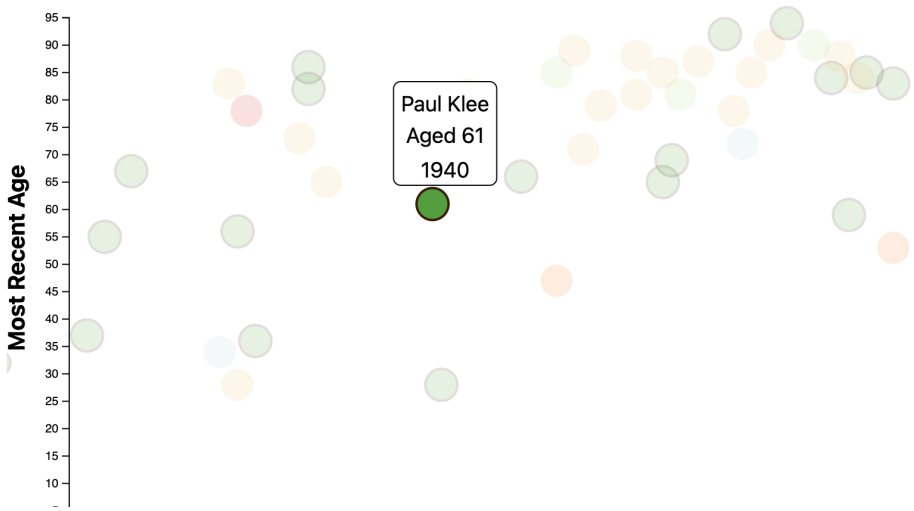


Figure 51. *A simple demonstration of the scatter plot's dot focusing and tooltip*

Lastly, the chord component's technical implementation was a challenge for us. When looking at D3 chord visualizations (see [Inspirations While Revising](#)), we soon realized that we wanted something different than what D3's chord and ribbon code provided. So we went into D3's chord and ribbon code, understood it, and developed something similar. This process was interesting because it relied on D3's support for quadratic Bezier curves, and the arc() function. For a somewhat detailed explanation, see figure 52 (do keep in mind that this only details the general process, not the many minor adjustments that we needed for fine tuning).

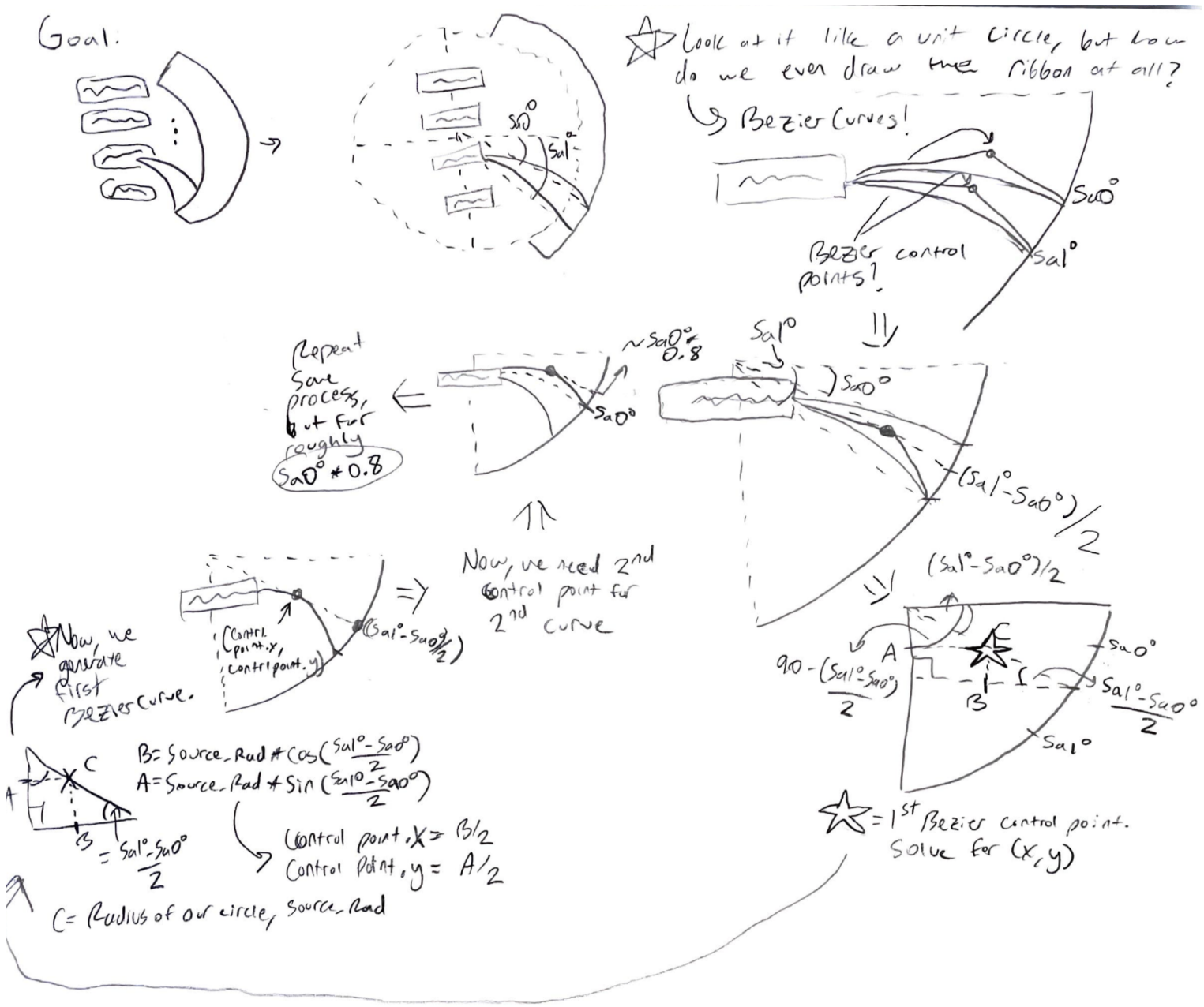


Figure 52. *A high-level overview of how the ribbon's calculations were done.*

Third Tab: Mediums

After clicking the third tab, indicated by the artist's palette icon, one can see our third and last view that conveys data about the chosen artistic mediums of our artists (see fig. 53).

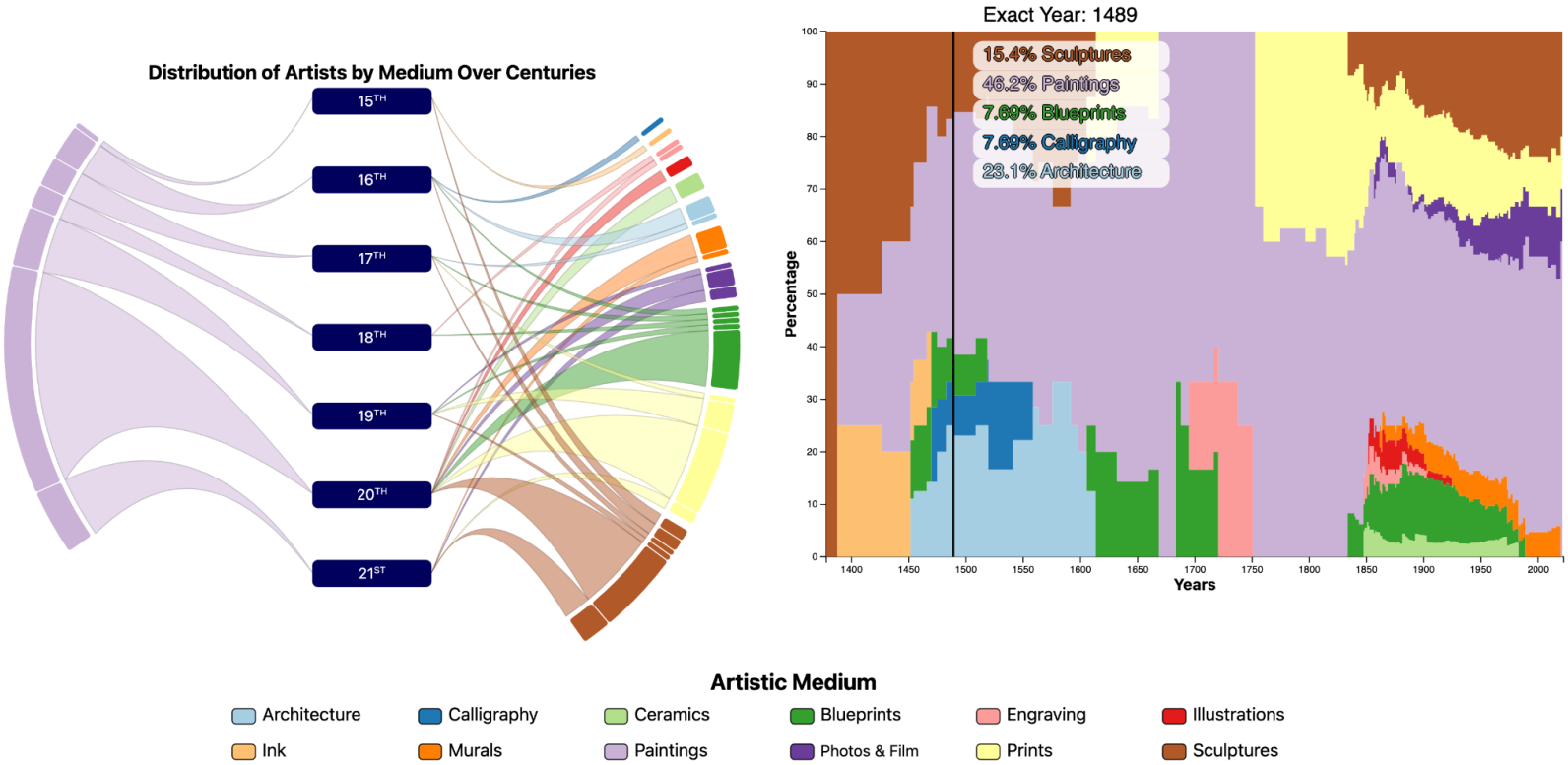


Figure 53. *The default view when selecting the mediums tab (in this case, the user is also highlighting on the area chart)*

One of the first things of note is how we have a line that indicates what the exact percentages are for a precise year. This is helpful in providing precision, and it builds off a similar functionality from a previous homework assignment from this class.

This view has an identical legend component compared to the deaths view. The only difference here is that it now describes our color-based encoding for artistic mediums instead of death types. It also functions similarly in that it filters both the chord component and the stacked area chart to whatever medium a user hovers over (see fig. 54).

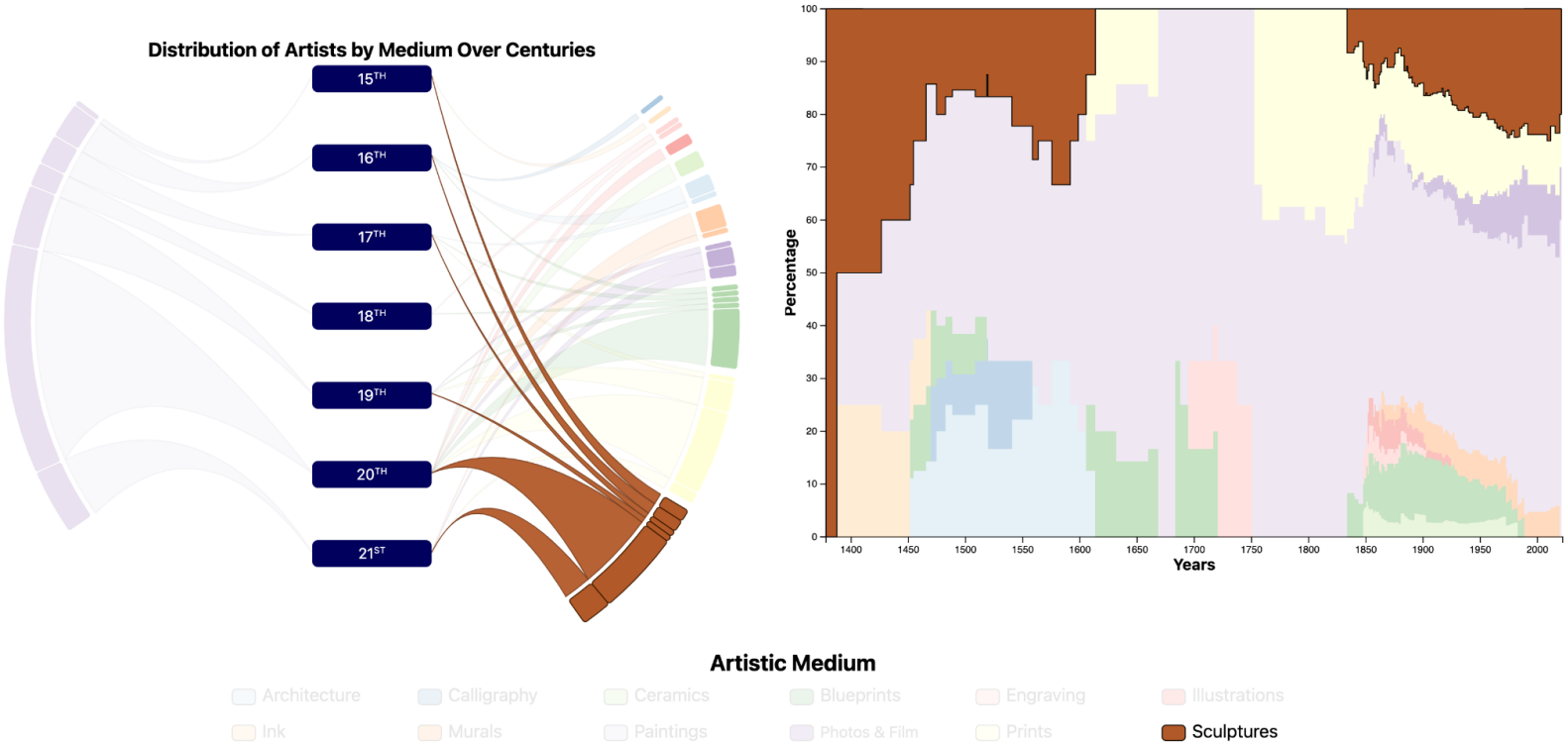


Figure 54. *Medium focusing across all components on the tab*

In another similarity between this view and the death view, hovering over an arc on the chord component accomplishes the same functionality as figure 54. Meanwhile, hovering over a ribbon does the same thing as highlighting an arc here.

However, clicking an arc does something entirely different compared to the death view. Here, it isolates the stacked area chart to only show that medium's area and how its percentage of usage alone fluctuated over time (see fig. 55).

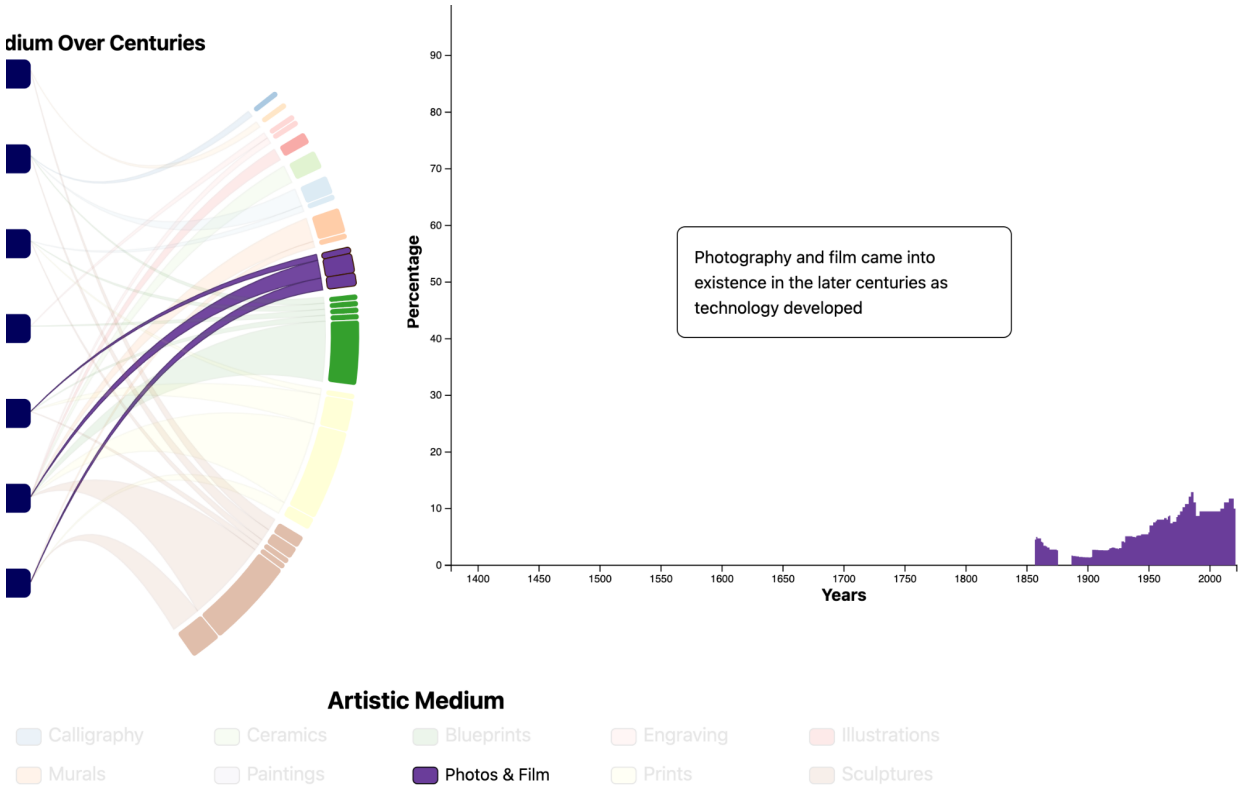


Figure 55. *An example of clicking the arc for a medium and seeing a tidbit about it*

One can also see in figure 55 how this view has minor storytelling components as well in the form of text boxes to explain certain trends, like how film does not show up earlier because it was not invented yet. Clicking another arc changes the area chart to focus on that medium, and then clicking the selected arc again returns that chart to its default stacked area look.

Another new functionality for this view relates to the century buttons. They do the same thing as before in relation to the chord chart, but hovering over them outlines where on the stacked area chart that century corresponds to (see fig. 56).

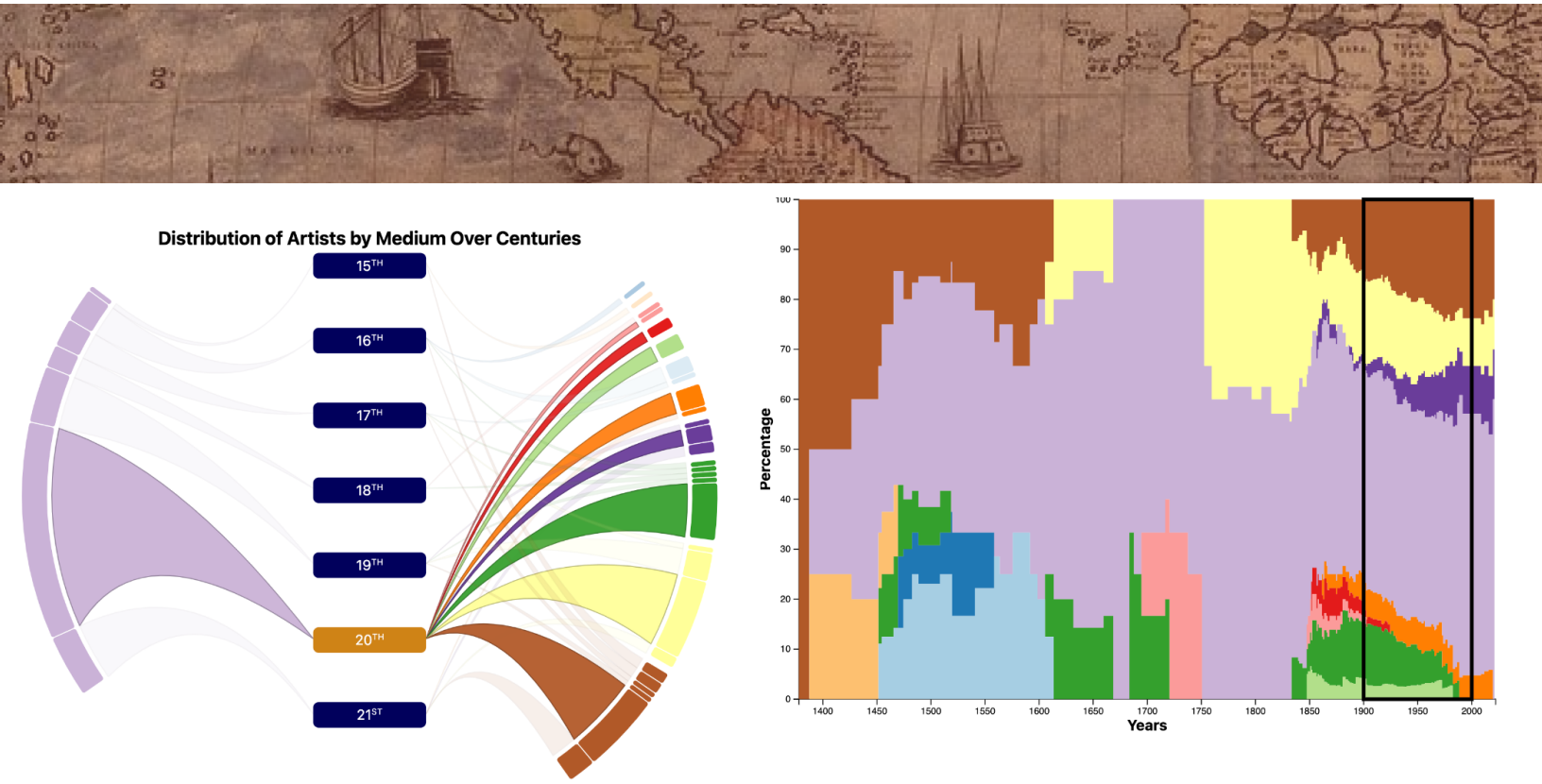


Figure 56. *The area chart's century preview in the form of a rectangle*

This provides a nice little preview of what is to come when a user then clicks that century. When this happens, we have a similar axis zoom to the scatter plot (see fig. 57).

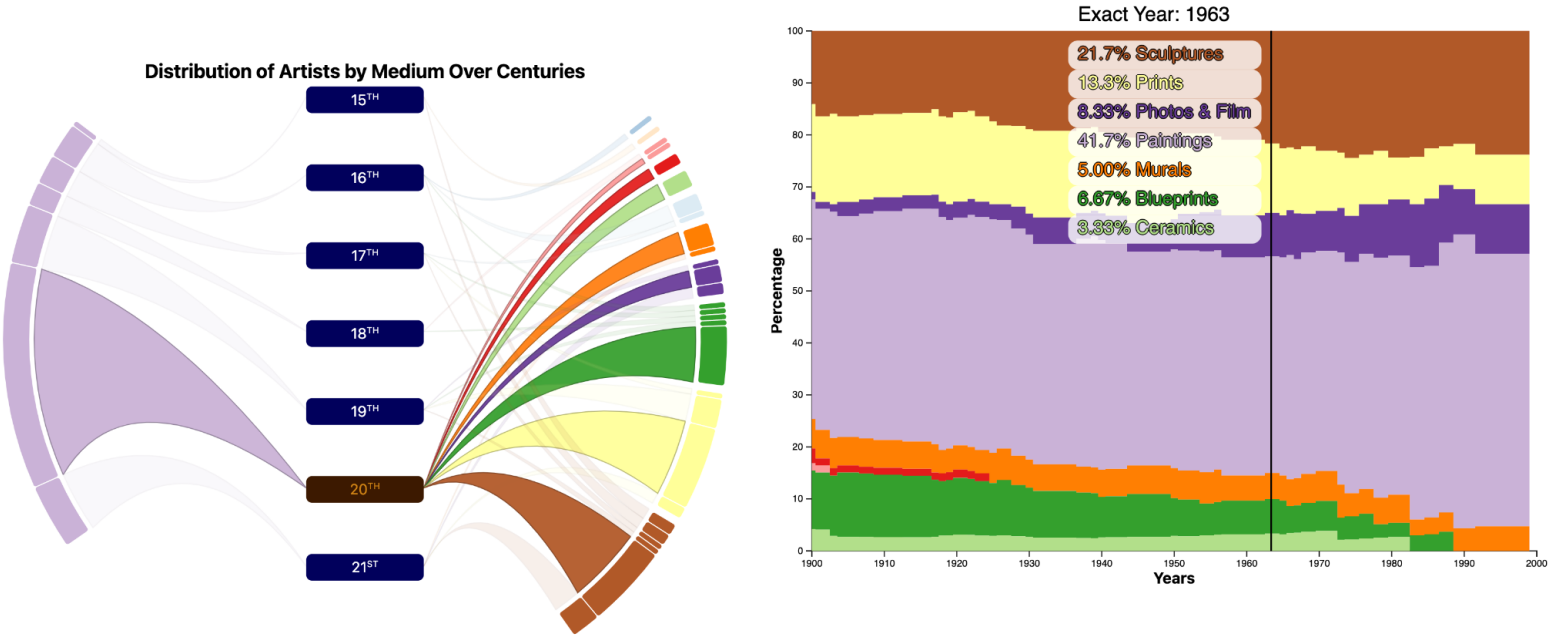


Figure 57. *An example of zooming into a clicked century on the area chart based on the chord chart*

Going back to the default full time range works identically to how it did on the death view, and so there is some coherency there for someone learning how to use this tool. Overall, adding the area chart component was more involved compared to the scatter plot, but they both enhance the chord component nicely.

Evaluation

What Did We Learn About Our Data?

While collecting data we experienced how fascinating it is to see how art influences other artists, whether it be by public murals with social messages, political agenda, crisis situations of nations, or high art as a status symbol. However we concentrated mostly on the influences of the artists in relation to each other. We specifically learned about their personal lives including their nationality, how they died, which century they were born in, etc. The more personal slightly unrelated data ended up playing a role though, as it informed our minor storytelling bits (see fig. 58).

Figure 58. *An example of an interesting obscure fact that we turned into a storytelling opportunity*

Through our visualizations, we can infer many aspects about the lives of the artists, and how they affected art all around the world. Using our network chart's interactivity with our world map, we can see how location was not a limitation for art to spread across the globe. The map chart makes it easy to understand influence over geography, and the width of the arrow showing the influence of one artist on the other shows how they are spread temporally. We also got to see how the geographical distance of the influence increases between artists over time due to technology allowing future generations to learn about past artists.

Looking at the deaths tab, we can understand how tragic some of the artists' lives were. We noticed how a significant number of them died of illness, suicide, or heartattack. One of them was even

allegedly murdered by their possible rivals. The majority of them died of tragic deaths. We also noticed, because of the scatter plot, that a significant amount of them died before they even reached middle age.

Switching to the mediums tab, we can see how the different mediums available to artists changed over time. Using hover over the area chart we can understand what percentage of a particular medium was used in that year. If we want to see more information about a certain medium, then clicking on it shows the data clearly which helps in understanding the trend of that particular medium over time. A single century can also be studied on the area chart (by clicking the century in the chord chart). We can see which medium was used more than the other in a specific century, which can depend on the current technology and resource availability at that time. We can clearly see how calligraphy was prominent in the olden days (clearly shown when a user hovers over calligraphy on the legend or chord chart), and it was more prevalent then compared to today. Another observation involves the introduction of newer media including photography and film that took place relatively recently. In the area chart, we can see how it exploded in popularity after it was invented.

How Did We Answer Our Questions?

For our visualization, we collected data from many sources including books and the internet for our visualizations, after which we decided upon what aspect of it we wanted to visualize in order to get the observations we are looking for. Once we had that we decided on which kind of visualization to use based on what kind of parameter we have, if it's spatial, temporal, categorical etc. This made the visualization more useful in inferring data instead of just being an SVG showing data on a web page.

Additionally, our range of questions is so large because you could apply many of them to each medium or each means of death, or even every artist. Thus, we believe that we have provided justification that we have answered our questions by explaining the functionality that we implemented in relation to those questions and our data (see the previous subsection).

How Well Does Everything Work?

In a way, our final product surpassed our expectations. This is because we met all of our must-haves, and even implemented some minor storytelling, which was a stretch goal of ours.

In terms of functionality, there is little to be worried about. The modularization of our components across multiple tabs actually helped us in relation to this because we did not have to account for edge cases involving certain chart combinations. For example, we never had to worry about how the chord visualization impacts the network chart because the two are now completely separate. We even made that chord component as flexible as possible to the point where one of the only differences between the two versions under the hood is one string parameter specifying what kind of data to visualize.

That being said, there are still a couple of edge case bugs that we have not completely squashed. These are very rare, and even replicating them is difficult. However, there are indeed some rare cases in which the horizontal zooming for the x-axis of the scatter plot stops because of some other interaction. We have a delay function that handles this, but there were a couple of times when it inexplicably did not work. Granted, these are only a couple of times compared to our hundred and hundreds of both test and regular interactions. All of that being said, our implementation works well.

How Can We Improve?

We can improve at the resizing and adjusting of the visualization when the size of the window changes. Right now it correctly resizes when the page is refreshed after the window size is changed, but we can have it resize dynamically as well. Browser compatibility is also a concern as it may be faulty at some places for different browsers.

Another possible improvement could be the loading of the page because right now it takes a bit to load the data. Maybe it can be structured in a different way or parallelized to speed up the loading.
We can also have more data in the database (eg. artists from Africa, South America, etc. that are absent currently) to make it more exhaustive. That would give more insights into the continental travel

and influence over time. We can also have artists from older centuries to show temporal changes outside of our current roughly 700 year range. Including cave paintings somehow would be fun. It will also make the curves on the area chart less step-like looking if we have simply more data.

We can also add data about the artists moving to different places during their lives and see if they come together in groups at a place to work and learn together. Another thing we can add is the location of the artworks created by them as then we can see that if many of the paintings end up at the same place, there is a high chance it is a collectors place or a museum. We also think that it would be fascinating to represent lost artworks somehow.

We can also add a custom map to look similar to the map shown on the title page of this process book, and the styling overall could have been a bit more historic.