2019

Art is Data is Art

Nicole Orchosky
nmo17@zips.uakron.edu

Follow this and additional works at: https://ideaexchange.uakron.edu/spa

Part of the Archival Science Commons, Digital Humanities Commons, and the Public History Commons

Please take a moment to share how this work helps you through this survey. Your feedback will be important as we plan further development of our repository.

Recommended Citation
Orchosky, Nicole (2019) "Art is Data is Art," Student Projects from the Archives: Vol. 3 , Article 6. Available at: https://ideaexchange.uakron.edu/spa/vol3/iss1/6

This Article is brought to you for free and open access by IdeaExchange@UAkron, the institutional repository of The University of Akron in Akron, Ohio, USA. It has been accepted for inclusion in Student Projects from the Archives by an authorized administrator of IdeaExchange@UAkron. For more information, please contact mjon@uakron.edu, uapress@uakron.edu.
Art is Data is Art

Nicole Orchosky, The University of Akron, Department of Multidisciplinary Studies

Abstract: The Digital Humanities field is rapidly introducing new and innovative ways in which we can analyze and explore large bodies of humanities material in order to make new discoveries and connections. This project serves as an introduction on how to use simple Digital Humanities tools to examine a dataset. In this project, data collected about the body of artwork exhibited in the 1913 Armory Show like medium, subject, or year of creation is analyzed using three different free-to-use tools. The data is then presented in a visual format that brings new questions and connections to light. The limitations and frustrations that can arise when taking on a Digital Humanities project are also addressed. This paper is intended to be read as a supplement to the provided project website. More information about the project in addition to the visual content, sources, and data mentioned in this paper can be found on the project website at https://nicoleorchosky.wixsite.com/artisdataisart

In my major I focused heavily on art history, and I’m particularly interested in the story of the 1913 Armory Show because of how it “scandalized” America yet eventually it greatly influenced what we consider “fine art” today. The 1913 Armory Show was the “first large-scale exhibition of modernist art in the United States” (Mancini 833). It hosted around 1400 artworks, and about a third of the artists were from Europe. This show was essentially the American public’s first introduction to what European avant-garde artists had been creating. Americans were very shocked by how abstract the new modern art was. One story that I think sums up America’s response to the Armory Show is a quote by art critic Julian Street. When Street saw Marcel Duchamp’s cubist painting Nude Descending a Staircase, (No. 2), he described it as “an explosion in a shingle factory” (Street). The very traditional, academic-minded critics of the US expected art to be realistic, beautiful, and pleasing to view. They did not look at a painting such as Duchamp’s and think it was “art.” Today though, numerous modern artists who exhibited in the Armory Show such as Duchamp, Picasso, or Cézanne have become household names and you can’t walk into a major art museum without seeing their work.

When I was working on a project about the Armory Show for another class, I hoped to find some sort of list of all the works in the show, ideally with dates, artist names, and maybe even images of each piece. I really thought this was a simple thing to expect based on how beautifully detailed art museum websites tend
to be and how integral the show was to the history of art. However, the best resource I could find was a website by the New York Historical Society created for the 100th anniversary of the show in 2013. The list of works was transcribed from the show’s original catalogue as well as some information found in Walter Kuhn’s Family Papers collection (an organizer of the show). The data is presented as text only and appears on just one long webpage that is difficult to navigate. On this page, you can’t make any searches, narrow your search, or filter by different fields. Another page on the site is a lot more easily navigable and allows the user to explore the artworks by gallery as well as search by various categories, but the site can be finicky and slow, and I thought that even more could be done using this data. The data was already gathered and transcribed; it just needed to be better organized. I started out wanting to make a searchable database with images of all the artwork, but if I’ve learned anything in this class it would be to accept my limitations. I suspect that I haven’t found an Armory Show list containing images because images may not exist for many of the works, and the images that do exist are on many different websites, all with different copyright rules for image use. Gathering all the images wasn’t going to be possible for me with my limited timeframe. I decided to start by creating a spreadsheet with all the textual information that I did have access to on the Historical Society site and go from there.

I learned pretty quickly that creating a dataset is not easy. I started by copying all the info on the Historical Society site into a spreadsheet, which was simple but time-consuming. The hard part was that I had to make a lot of (sometimes tough, sometimes seemingly arbitrary) decisions and then commit to them in order to maintain consistency. One issue I came across was that I completed my first version of the spreadsheet with the artists’ names separated into first/middle/last names. I thought this would be useful for some reason but it never was, and it just got more complicated the further I got down the list, as some artists had very long names with multiple parts, or suffixes and prefixes to their names, and many women were known by both their own names and their husbands’ names. I got to the end and regretted breaking up the names like I did because it really served no purpose when visualizing the data, so I had to go back through and make a new column simply including the full name of each artist. Another challenge was that many of the artworks were known by more than one name or the list contained a foreign name and a translation of that name. I didn’t want to leave anything out, so I entered all the names available separated by semicolons under the title column. One more issue was that some artworks were
listed as created during one year but were discovered since the show to be created a different year. Again, I didn’t necessarily want to leave anything out and I didn’t know where the new years came from, so I listed all available information separated by semi-colons. I also had to make the decision to include only the most recently discovered years in my point map. Now that I had this massive spreadsheet of nearly 1400 works of art, I had to decide what I was going to do with it.

Inspired by our class time playing with Voyant and learning about other digital humanities data visualization tools, I decided to make a few visualizations with this data set. I also loved the way Kristen and Kacie’s websites were turning out, so I wanted to try presenting my project through a Wix website. I ended up naming my site “Art is Data is Art” because I like the idea of taking a visual work of art, breaking it down into textual and numerical data in a spreadsheet, and then reinterpreting it into visualizations in the form of graphs and charts that are works of art in their own right.

I researched and tried out many data visualization tools, but the three that I ended up using were RAWGraphs, Flourish, and Voyant. I had a few criteria in mind when deciding which programs to use. The program had to be free (or provide a free version), I wanted to work directly in the browser rather than downloading any software, and ideally, I wanted the resulting graphs to look nicely designed and polished. The programs I would end up using also had to have the types of visualizations available that would work well with the type of data I had, so as I browsed programs I thought about my dataset and the categories that I ended up including and I considered how those different categories could relate to one another in interesting ways.

I wanted to start simple, so I started with Voyant, which I already knew how to use from class. I also liked that I could end up with interesting visualizations using only one field of my data, in this case the titles of the artworks. I made a version of my spreadsheet with only one column including the titles of each of the artworks and uploaded that list to Voyant. I wanted to see if certain words were used more often in titles than others. I discovered with Voyant’s word cloud feature that some words are definitely used more than others in titles, especially the word “woman.” It isn’t really surprising that a lot of art features women as a main subject, but it was cool to see it in a visual like this. Of course, one question led to another, and I wanted to see what would happen the more I cleaned my data. There are more details about this process on the website, but I first entered the titles as-is, then entered only the more unique words like nouns,
verbs, and adjectives, then I translated any foreign words into English to more accurately represent the frequency of each subject. I will admit that this word cloud, while cool, doesn’t represent the entire Armory Show. Some artworks were unnamed, some were simply called “drawing 1” or “no. 1,” and some, as I noted earlier, were known by more than one name so I had to decide which name to include. The word cloud does not represent the entire collection, but it does reveal the frequency in subject matter for about half of the artworks in the show and there were still some definite patterns to be uncovered regarding modern art of the period.

Next I used RAWGraphs because it was really easy to use, and it offered a lot of different types of charts and graphs that looked great. RAWGraphs tells you exactly what types of data are needed to create the graph of your choice, whether it be textual, a date, a location, a numerical value, etc., and all I needed to do after uploading my spreadsheet was drag each column I wanted to include to the appropriate category. I created what is called a “cluster dendrogram,” which is used to show hierarchy. I wanted to see where the artists that participated in the Armory Show originated from. For this chart, I only needed to include the columns for City, State or Region, Country, and Artist Full Name. Once organized in the chart, you can get a good sense of where the artists came from and which countries were best represented in the show. Again, this chart has its flaws. For example, I couldn’t find location information for some of the artists, so they aren’t represented in the chart.

I used Flourish for my last two visualizations. Flourish was the hardest of the three programs to get the hang of initially. I had to adjust my spreadsheet a few times to make it work correctly in the program. However, Flourish is very user-friendly because once you choose a type of visualization it shows you an example using its own default dataset. The program shows you both the final visualization and the spreadsheet used to make it, including what type of data needs to be in which column. Studying the example made it easy to adjust my spreadsheet so that it had the correct data in the correct columns, and once that was done my visualizations appeared accurately on the first try.

I first made a hierarchical bar chart because I wanted to know which artistic mediums were present in which galleries at the New York show. Again, this idea emerged just from looking at my spreadsheet and wondering what would happen if I compared one column against another, in this case, “medium” and “gallery.” The hierarchical bar chart gives a good idea of what mediums of art were placed where, and reveals that a handful of the galleries definitely focused on certain
types of art. At a glance you can see that there were a lot of oil paintings overall, and that gallery A had a lot of sculptural work while galleries K and L had more drawings and prints.

The most exciting visualization that I made was the point map in Flourish. For this visualization I wanted to know where each artwork was made and when, and I was able to make an awesome interactive map with a scrubber bar on a timeline. The points each represent works of art and are color-coded by medium. This visualization required the use of location data like the RAWGraphs dendrogram, but Flourish was more work because it required GPS coordinates for each location rather than just location names. I wanted this map to happen, but I didn’t want to type each and every one of hundreds of location names into a map to find their GPS coordinates, so I looked to more computer tools to see if I could convert all of my locations to coordinates at once. I followed a combination of different tutorials and managed to convert my data relatively quickly with the use of a few different programs, and I discuss this process in more detail on the website. As far as the final map goes, I’ll offer another disclaimer: this visualization is by no means perfect. The locations I was able to find were only the birthplace of each artist, not the locations where each artwork was made, so after I did all that work I was disappointed to realize that the final map didn’t actually represent exactly what I hoped it would, the location where each artwork was made. Artists move around, they live in different countries, they travel, there wasn’t enough time for me to research the locations where they made each work, and in most cases that information is probably unknown. I didn’t want all my work on this map to go to waste so, while I presented the map knowing that it’s not necessarily accurate, I tried to emphasize throughout the website that my project is more focused on showing the potential of a dataset like this more than it accurately represents the data itself.

We’ve talked a lot in our class about who the audience is for projects like these. I think that the spreadsheet I’ve compiled can be useful to art historians and art history students, especially as the 1913 Armory Show is so incredibly important in the history of modern art. I also think that the general exploration of data like this, regardless of the subject matter, should be on every humanities scholar’s mind as it can uncover new connections to explore, or at the very least data visualizations can help students understand complicated subjects and concepts at a glance. These tools are beneficial to both students and researchers, and my hope for anyone viewing my website is that it would get the gears turning in their minds about what else they could do with my dataset, or how they could
use any data to visualize all kinds of different subjects across disciplines with the help of free online tools.

There are plenty more things that I wish I had the time or ability to do with my data. For example, I would still love to be able to compile as many images of the artworks as possible into one database. I also think that, because we do have access to the show’s floorplan and gallery locations, a virtual gallery would be incredible, especially if you could explore it using alternate reality. I also imagined at one point how cool it would be to figure out what colors were used most in each artwork and find out if certain colors were used more often based on artist or location, but identifying color is more subjective and I think qualitative data may be more difficult to represent than quantitative. This would also probably require a program in which you submit an image and it extracts the color palette from the image, which I know exists, but would be nearly impossible to use without all the artwork images available in one place. I think that more fields could also be added to the spreadsheet to produce more results. For example, the amount of money for which paintings were bought and sold is known through primary sources. The website where I got my source data also indicates which artists were personally invited to the show and which had to apply themselves. Additional data can only make for more interesting visualizations in the future, the possibilities are endless.
Works Cited


Mancini, Joann M. “‘One Term Is as Fatuous as Another’: Responses to the Armory Show Reconsidered.” *American Quarterly*, vol. 51, no. 4, Dec 1999, pp. 833-870.