Sophie Sawyers

Homework 2 Analysis

In looking at my visualizations mapping the various relationships between nationality, birth decade, and gender of the artists in the dataset, there are many interesting patterns to note. When looking at the first sankey visualization mapping nationality to decade, it is clear that most of the artists in the dataset are American (4,750 artists). France and Germany are roughly tied for second place (France with 725 artists and Germany with 797 artists) with Britain in third with 640 artists, but these European countries fall well behind the United States. Additionally, most artists seem to be born between the 1920s and the 1960s, with the 1940s being the most popular birth decade (1,340 artists) followed by the 1930s (1,140 artists). There are very few artists born in the nineteenth century, especially between the years of 1810 and 1870. While the dataset included American artists born from nearly every decade, some countries, especially those in South America like Argentina and Brazil as well as central Europe like Poland, Denmark, and the Netherlands, had very little artist representation. The prevalence of American artists translated into the sankey visualization mapping nationality to gender as well. America had the most representation (4,860 artists), once again followed by Germany (871 artists), France (782 artists), and Britain (764 artists) though the competition is not close. In this visualization, which had a higher count threshold of 30 artists for visualization purposes, had overall poor representation from countries in South America, Asia (with the exception of Japan), northern and central Europe, and Africa. Perhaps what is more notable in this visualization, though, is the representation of gender in the artist data. Of the data represented in this visualization, about 84.32% of the artists were male (8,930 artists) while 15.68% of the artists were female (1,660 artists). Furthermore, of the female artists that were represented, all but two nodes originated from North American and European countries, meaning there is little to no female representation for female artists in South America, Africa, Australia, and Asia. This trend certainly speaks to gender inequality throughout history as well as possible racial disparities artists from predominantly non-White countries face. The third sankey visualization mapping gender to birth decade continues this analysis of gender over time. As with the second sankey visualization, most of the artists represented in this visualization's data were male; specifically, about 82% of the artists were male (8,750 artists) while 18% of the artists were female (1,920 artists). Every birth decade includes male representation, while the first birth decade with female representation was not until 1870, 70 years after the first recorded male representation in the data. In addition to the observations made about gender and race/nationality in the previous visualization, this visualization confirms this lack of female representation through a chronological lens as marked by the late arrival of female artists in the data. This translates visually as well as the most prominent nodes for male artists seem to map to 1930 (1,210 artists) and 1940 (1,180 artists) while the most prominent nodes for female artists are not very noticeable at all (317 artists in 1940 and 324 artists in 1960). The multi-layered sankey diagram puts all of this together, showing the most representation of male artists from America or Western Europe in the mid to late twentieth century.

Overall, these observations show that there has been a lack of diversity, equity, and inclusion in the art world. The vast majority of artists in the data are white men, with women only starting to be represented 70 years after the first male artists in the data. Even when women are represented, their presence pales in comparison to male artists, which are represented in larger numbers and for more years included in the dataset. This includes "later" years in the dataset, when we would expect more of an emphasis to be placed on female artists (in 1970, there were about half as many female artists as there were male artists). From a nationality perspective, as mentioned, most artists were American followed by countries in Europe with a few from South America and Asia. In other words, most of the artists were from predominantly white countries. With all this in mind, there is a significant bias in the art world for the years represented in this dataset (and for years after) that needs to be addressed to ensure artists from a wide variety of racial, ethnic, and gender backgrounds have an equal chance to display their work for the public.