A visualization of gender and cultural biases in artwork collections

Tessa Cramwinckel
University of Amsterdam
Stud. nr.: 10679294

Ewoud Vermeij
University of Amsterdam
Stud. nr.: 11348860

Rico Mossinkoff
University of Amsterdam
Stud. nr.: 12805157

Yke Rusticus
University of Amsterdam
Stud. nr.: 11306386

Hannah Lim
University of Amsterdam
Stud. nr.: 10588973

ABSTRACT

The majority of art collections, musea and art expositions are still overrepresented with male artists and western artworks. However, the precise scale of these overrepresentations are still rather unknown. This paper introduces a visual application to create a clear view on the ratios between male and female artists and western and non-western artworks for any artwork dataset it is given. We introduce a simple design, with clear bar charts and floating donit charts to create a playful, clear visualization. The main contribution of our application lies in bringing awareness to the users of the male and western dominance in artwork collections. Effectiveness of our application is validated through observations of test users. Observations have shown that there is room for improvement in terms of usability and informativeness.

1 INTRODUCTION

Even in this modern century, the artworks presented nowadays are mostly from male artist with a western background [7][16]. For example, in art collections in America, 85% of the artists are white and 87% of the artists are male [3]. This imbalanced representation of artists over the years may be problematic, as female and non-western artist might feel unrepresented and unwanted.

Although it is possible that most paintings in history are made by men or non-western artists, it is still important to be aware of this inequality. Therefore, museums or other art collectors that have a big influence in art education, should take this into account when presenting their artwork and perhaps they might want to consider informing the public with this inequality.

Therefore, in this project a visualization application is created that visualizes gender and cultural biases in artworks. In this project the OmniArt dataset is used, but it can be applicable for any artwork dataset. The application is interactive, such that users are able to play with settings such as time or which bias, gender or culture, should be shown in the current setting. In section 1.4 we will address the details of the application to a greater extend.

In the following of section 1.1, the dataset used for building the tool will be explained, hereafter the inequality regarding gender and cultural background will be discussed and finally the visualization tool will be introduced. In section 2 some relevant related work will shortly be reviewed. Then, in section 3 the relevance of the tool will shortly be supported with a use case. Thereafter, in section 4 the design of the concept will be thoroughly explained, followed by the design of the actual product in section 5. To evaluate the tool, a small empirical research is done, from which the results will be presented in section 6. Finally, the important findings and improvements will be discussed in section 7 followed by a short conclusion in section 8.

1.1 Dataset

For this project the OmniArt dataset is used. This dataset contains over 1.9 million artworks from all over the world. Every artwork is annotated with meta-information, such as creation year and the art school from which the work originates. This information can be used to filter and retrieve artworks with certain characteristics. The dataset contains different types of artworks, including paintings, images and sculptures. Various subgenres from every type are also annotated to help finding specific niches. The dataset is meant to be integrated with popular deep learning frameworks. It can be used to test model performances and experiment with different settings. Because there are quite some features per artwork, there

are many possibilities to experiment with the data. This also holds for visualizations.

1.1.1 Preprocessing data. For the purpose of our visualization, we focused only on the artworks of the general type 'painting'. This subset consisted of around 330 thousand artworks. Although many features were already included in the data, gender and culture were not explicitly annotated. We therefore implied gender and culture labels, based on the other annotations. Note that this is not ideal, and should not be taken as absolute truth, but suffices for the purpose of this work. For gender we implied, or guessed, the gender of each artist based on their first name. For this we made use of a public library[1] in python, that based it's predictions on over 60 thousand common first names. We labeled each artist as "male", "female" or "unknown". For culture, we relied on the already present annotations in the category "school". We manually labeled each school as "western", "non-western" or "unknown". For example, the "Dutch" school was labeled as western, the "Japanese" school as non-western. Many artworks had an unknown school, or were annotated as something unclear such as ", 1866", so these were given the label "unknown". The subset, together with gender and culture annotations was used for the following work.

1.2 Gender bias

With the analysis on the first names of artists, it was possible to estimate the number of female artists in the OmniArt dataset. It showed that 5% of the pieces were created by woman, 36% by men and 59% of the pieces couldn't be identified. This arises a couple of questions:

(1) What is going on in the 'undefined part'? Can we assume that it has the same distribution as the defined part?
(2) How does this distribution change over time?
(3) Did woman in reality make less art than man or does this dataset give us a distorted view of the art history?

The first question is hard to answer. The part that is undefined contains names like 'xynphix'. It requires a lot of extra background research to figure out the gender behind names like this. Nevertheless, since this is a significant part of our dataset, ignoring this could seriously effect the results. For example, it is a plausible scenario that female names have more variations and thus are harder to identify by a simple script. If this is the case, the defined part of the dataset is not a good representation. We are aware of this limitation but due to limited time, we decided to not spend time to do extra research on this undefined part.

The second question, 'How does this distribution change over time?' will be answered with the visualisation where the user can choose the time period and the visualization then shows the distribution of men versus woman. Section 4 will elaborate on the design of the visualisation.

The third question, 'Did woman in reality make less art than man or does this dataset gives us a distorted view of the art history?', can best be answered by consulting sources in the field of art history.

[1]https://pypi.org/project/gender-guesser/

One the oldest identified art pieces are cave paintings. These paintings often have human hand prints. Studies were able to identify 75% of these as woman hand prints[17]. In the time that follows, the ancient historical area, art hardly mention specify individual and it is hard to describe the role of woman during that period. Although some studies show that in India, art was primarily made by woman than by men [19]. From the medieval period, woman seem to play a significant smaller role in art history. This can be explained by a number of factors: for a long period, woman tend to make more 'decorative art' like textile which was not categorized as 'fine arts', which has mostly been occupied by men. Secondly, many woman where not able to follow any education as well as art school. And lastly, society seem to acknowledge art made by men more than woman art. The artist Hans Hoffman wrote to the feminine painter Lee Krasner "This is so good you would not know it was done by a woman." [6][15]. When the feminist movement began around 1960, more woman started art school and contemporary art is now represented by both men and woman [1]. But still, nowdays the art museums are represented by male artists for 87% [3] and at auctions, pieces from male painters are higher valued than female painters [2].

Referring back to the initial question, it seems that our two scenario's (1. woman made less art than men and 2. the dataset is distorted) are intermingled. Sources tell that woman did made less art than man in history, but since woman were less acknowledged over time than man, this may have let to a selffulfilling prophecy.

1.3 Cultural bias

From our dataset, with manually annotated labels concerning culture, we found that overall 65% of the artworks were labeled as "western", 1% as "non-western" and 34% as unknown. The same questions as before arise:

(1) What is going on in the 'undefined part'? Can we assume that it has the same distribution as the defined part?
(2) How does this distribution change over time?
(3) Did non-Western artists in reality make less art than Western artists or does this dataset gives us a distorted view of the art history?

Again, the first question is hard to answer. We will see later on that an unbalanced distribution is expected since this dataset has Western origin. Investigating the undefined part could be very interesting but we decided not to explore the dataset any further.

As with the gender bias, the distribution over time will be visualized with our tool. Section 4 elaborates on this.

The third question, 'did non-Western artists in reality make less art than Western artists or does this dataset give us a distorted view of the art history?', needs some context before answering.

The earliest 'art pieces' that have been found date from around 100,000 BC. These pieces are classified as 'Upper Paleolithic Art' and contain sketches carved in rocks. Later on cave paintings were developed [13]. Almost every textbook about art history starts with this introduction about the development of art as a process that took place in societies all over the world. But as soon as the text

books arise in more modern eras, a lot of writers start to focus on art crafted by Western artists. [11].

This tendency of focusing on Western art is also visible at the prominent art museums like Musée du Louvre in Paris and the Metopolitan Museum in New York. Nine out of ten most popular art museums are located in Western cities and focus mainly on Western art [16]. The most popular exhibition in a non-Western art museum was at the Shanghai Museum in 2018. The title of the exhibition was 'Masterpieces from Tate Britain 1700-1980'.

A third area where the underrepresentation of non-Western art is visible is at auctions. The most expensive painting ever sold is painted by the Italian artist Leonardo da Vinci and sold for over 450 million dollars [8]. A list of 89 highest known prices paid for paintings is largely represented by the Western artists Vincent van Gogh, Pablo Picasso and Andy Warhol [2]. Only three artists from this list have a non-Western background. Also notable, woman don't appear in this list. The highest price paid for a painting painted by a woman is 44 million dollar. This shows that not only are Western artist are more represented in textbooks and museums, they are also valued higher globally.

Is it because non-Western cultures made less art? This is a somewhat complex question. Art is of all times and it is spread all over the world [13]. Every culture has produced art in some way in civil history. The dominant medium and subject of their art is different in every culture. For example, the Western culture likes to make its paintings on canvas while Chinese paintings are mostly made on rice paper. This has an influence on the conservation of art: oil on canvas preserves very good in comparison to the water paintings on rice paper [5]. This could have an effect on today's view on art history. Furthermore, Long and Bradford write about a relationship between welfare and art productivity and state that Western societies produced more art because of the higher welfare [9]. Although Western cultures seem to have had a higher productivity in making art, it is clear that even taking this in account, non-Western art is still underrepresented in textbooks, museums and auctions.

Cultural bias in art history is an issue that is often criticized. Students from different (non-western) backgrounds learn from these textbooks and feel under-represented [12]. This also confirms the unequal contributions of western and non-western art within the collections that our dataset covers. Our visualization is designed to interactively discover these inequalities, among collections and through time.

1.4 The Application

Even without clarifying why there is a gender and cultural bias in most common and well known art, one would know that there is one. The question that arises is how biased art actually is. The goal of our application is very clear. We want to give the user of our application a clear overview of gender and cultural biases through time for different art schools, art collections and artwork types such as paintings, photos and statues. Our application should be able to show the user the ratio of western and non-western art through time, and the same with gender. For this we will use an intuitive combination of a timeline and bar- and donut charts, each showing a ratio. All charts will change fluently when time periods are changed or when a specific bias needs to be shown, if the user pleases so. This will give a nice smooth experience and clear overview of what is happening to all ratios.

To construct our application, we will make use web application techniques such as HTML, CSS and JavaScript. In order to make our application run smooth, we will make use of the D3, a well known visualization library in JavaScript. We did consider the library Bokeh as well, but chose D3 for its freedom in webdesign.

2 RELATED WORK

The OmniArt dataset has been used before for interesting applications, aiming to retrieve some new insight. The following sections will discuss some of those applications in order to get acquainted with the possibilities with the OmniArt dataset. We will also touch the D3 JavaScript library to show the libraries possibilities.

2.1 OmniArt examples

On the website of OmniArt[2], some visualizations of the dataset are displayed.

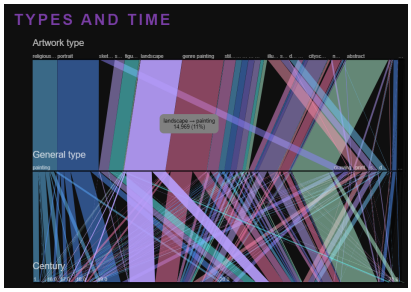

Figure 1: Similarity graph

For example, figure 1 shows a similarity graph that links artwork types to the centuries in which they were built. Hovering over the connections highlights them, to show exactly where the links are. This gives an intuitive idea of the distribution of certain artworks through time. Furthermore, there is a timeline that shows the activity of various artist throughout the years (figure 2).

[2]http://www.vistory-omniart.com/visualization

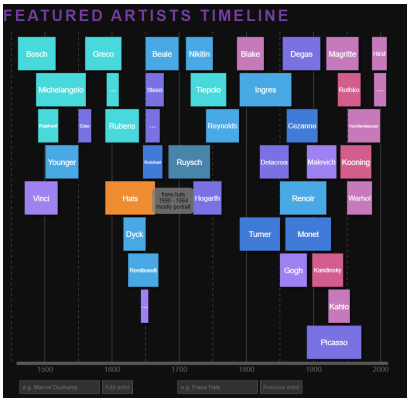

Figure 2: Visualisation that shows activity of painters over time

When hovering over the artist bars, it shows the most common artwork that the artist produced with a time frame in which the artist was active. You can also click on the bars. This will show a timeline with dots for all the works that the artist produced (Figure 3). When clicking on one of the dots, it will show the artwork and

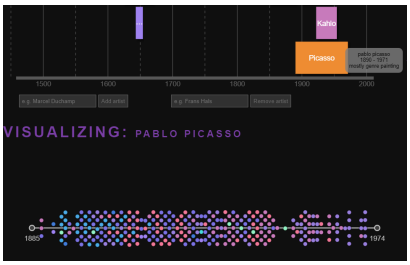

Figure 3: Artist timeline

give some basic information. These examples show some of the possibilities within information visualization. However, there are many more possibilities when using different attributes of the data and different visualization techniques.

2.2 Other visualizations

On the website of D3 [3] there are quite some other examples that show the possibilities within D3. We will outline three of the relevant visualizations for our project.

2.2.1 Barchart race. This example shows how D3 can be used to smoothly transition bars. The code makes use of an svg box with linear transitions. In our design we also make use of bars in which we display information. We will use similar transitions as in this example. However, in this example the bars are separate. We will use stacked bar charts which will be discussed in section 4.

2.2.2 Donut chart. Another interesting example for our project is the donut chart. In this example there is a donut chart which is divided into several categories. Each part of the donut can show some information. We used this in our project as well. [5]

[3] https://js.org/
[5] https://observablehq.com/@d3/donut-chart

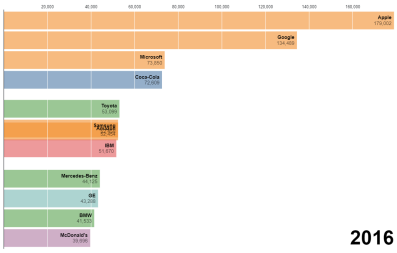

Figure 4: Interactive bars[4]

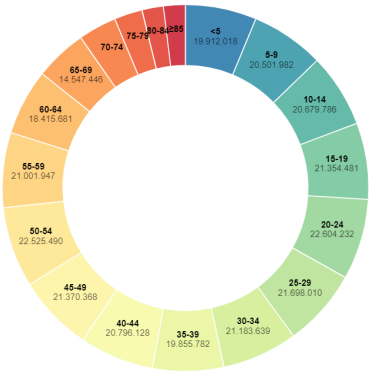

Figure 5: Donut chart

2.3 Icelandic population by age

The following example is also quite interesting for our project. Figure 6 shows a snapshot from an animation that shows information about the Icelandic population. This example has an adjustable timeline and shows the ratio between male and female citizens with their age. Moving the time button shows the changes in this distribution. As we look at the male/female ratio through time in our project, we will use similar techniques to adjust the time and show changes. [6]

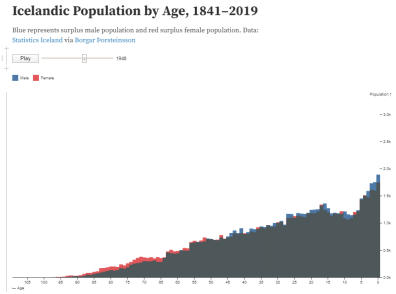

Figure 6: Interaction with time

3 MOTIVATING USE CASE

The art collections of museums in America are for 85% represented by white artists and 87% by male artists [3]. Museums play an

[6] https://observablehq.com/@mbostock/icelandic-population-by-age-1841-2019

important role in art history education in society and therefore they should take extra care about the story they tell by exposing their art. Large museums like MOMA make a serious attempt to change this lack of diversity by setting quota for art purchase and change the ratio of their exposed part of their collection [10] [18]. A tool that visualizes the diversity of their collection could help them to gain insight and to adjust their policy if necessary. Also, if museums attempt to be more transparent about their collection, they could also make this tool public and let visitors play with it.

Not only museums, but also holders of large private collections, online databases and collections owned by companies could benefit from this tool. Again, it gains insight and helps the holders in their management concerning diversity of their collection.

4 CONCEPT DESIGN

In this section we will elaborate on the initial concept of our visualization and all design choices. The goal of the visualization tool is to provide a clear representation of the gender and ethnicity distribution of a painting collection. This is to provide users with more insights into the gender and ethnicity bias in art. Additionally, users will be able to compare the two variables over time to investigate on the (dis)similarities between these variables. To allow users to explore the data and detect relations and trends from the data, the tool is made interactive. In addition, an interactive tool will improve the learning curve of users.

To give users a clear overview of the distribution, the visualization tool will contain two bars on both sides of the screen. These bars are stacked bar charts that represent the gender distribution and the cultural distribution among the paintings in the selected time period. In the rest of the report, selecting these bars will be referred to as selecting a certain group (i.e. gender or culture). Between these bars, the donut charts are visualized that represent the distribution of the group(s). Each donut will represent an art collection (e.g. MOMA New York and WikiArts 17), school (e.g. modern and french) or an artwork type (e.g. portrait and landscape) where the user can choose this selection. Since the data contains temporal information, one important feature of the visualization tool will be the timeline. Users will be able to slide over time to see how the distribution of gender and culture change over time. Furthermore, a feature will be added where users can receive help on how to use the tool. In the following sections the concept will be further explained.

4.1 Timeline

As the data contains historical art items going back as far as 95000 B.C., interaction with time is one of the most important features of our application. Since the data contains only information about the years of the artworks and because long periods are measured best in years, time is considered in years in this application. This can be considered as ordered linear time [4].

Interaction with the timeline should be intuitive. Users should be able to set two dates whereas the tool will use the selected time period to extract the corresponding data and visualize it. Initially multiple options were considered. Option one is a slider, where the user could drag two slider points to set a specific time period. Option two consisted of two separate interaction components where users could scroll in both elements to set a time, or could type a year to select a time period. Both options were considered for the final product. Additionally, setting a certain time interval and being able to move this interval might be a desired feature.

For this interaction it was chosen to implement a slider, since this gives a more natural flow to the interaction. Timeline interaction occurs preferably at a more higher location on the screen.

Adjusting the time period, will cause the donuts and the bars to change in shape, if the data shows change. In the following two subsections these changes will be further explained.

4.2 Flying donuts

In this subsection we will describe the visualization shown in the middle of the screen, the flying donuts. Each flying donut represents a certain category of a variable in our dataset. The data contains the variables: artwork type, school and art collection. The user is able to choose the variable to explore which will be represented by the donuts. Interaction with the bars give the user the possibility to choose between the groups. After having selected a variable and a group, the donuts will be visualized. The donuts are smoothly adjusted in shape, when the time period is changed.

Each donut is a pie chart that represents the ratio between female and male artist or between western and non-western artist. Since we are interested in the distribution within the groups, all the artworks created by artist where the gender or the cultural background are unknown are left out of the donut chart visualization.

The size of each donut represents the quantification of the group (i.e. indicating how many artwork items are present of the group at the selected time period). In addition, visualizing the donuts according to their ratio will give a clearer overview of the distribution within each donut. For example, donuts with a high ratio between male-female will all be presented at the bottom of the screen and donuts with a low ratio between male-female will all be presented at the top of the screen.

4.2.1 Hover and click feature. To gain the information represented in the donut, the user will be able to hover over the donut. Hovering will present a tooltip that contains the amount of artworks created by the groups with its corresponding percentages. In addition, a small subset of the concerning collection (i.e. the paintings by the artists displayed in the donut) clearly seperated by the groups will be shown in the tooltip.

Clicking on the subset of artworks, could lead to two possible options for visualizations. The first option will simply show the full art collection on the screen where the artworks by different groups are clearly divided. To go back to the donut chart, the user simply clicks somewhere on the screen so the full art collection disappears. The second option will change the background to paintings from the art collection of the donut.

Furthermore, clicking on the donut would zoom in and centralize the donut to be the center of the donut chart.

4.3 Bar interaction

In this subsection the bars visualized on the left and right side of the screen are discussed. The left stacked bar chart represents the gender distribution, as the ratio between male, female and the unknown sex of all artists of the corresponding time period. The

right stacked bar chart represents the cultural distribution, as the ratio between western, non-western and the unknown cultural background of all artists of the corresponding time period. The distribution of the bars are dependent on the time period selected with the timeline.

4.3.1 Hover and click feature. To gain the exact amounts of the artists of each group in the distribution, the user will be able to hover over the bars. This will present a tooltip that contains the total amount of artists in each group at the current time period (including the amount of the unknown artists) with their corresponding percentages.

More importantly, the user will be able to click on a bar to interact with the visualization and select the group that will be explored in the donut charts. Furthermore, selecting a bar and filter by a part of the other bar might give additional insights. For example, selecting only female artists and showing the cultural distribution in the 19th century per collection.

4.4 Help

Even with the best, most intuitive user interfaces, it is desirable for users to have the possibility to be helped when questions about the use of an application arises. Therefore, in this application the help section should be easily reached by a clearly visible help button. When clicked, a popup screen should show a help section with all information needed to interact with our application without any problems. It is thrived to make the help section simple and intuitive and as clear as possible.

4.5 Loading screen/icons

The first screen of the application will be the loading screen. This loading screen is necessary because all the paintings need to be downloaded first. Loading screens are convenient as they keep the user entertained while they wait for the tool to open. Having a black screen, will cause the user to exit the tool.

4.6 Background design

Background design should be in balance with all other items shown. A noisy or distracting background can draw the attention of the user in the wrong direction. However, a background can also provide additional information to the user. Since this application provides information about art, it might be interesting to show the artworks on the background. Furthermore, changing the artworks background as mentioned in section 4.2.1 would be a nice and informative feature.

4.7 State diagram concept

In figure 7, the 'State Diagram' for the visualisation tool is shown using the categories of interaction techniques proposed by Yi et al. [14]. When the user opens the tool, the loading screen will appear. After loading the data, the user will enter the exploring mode. In the exploring mode the user has different options. The user is able to filter the data by sliding over the timeline, to explore by selecting a different category (i.e. gender or culture), to gain instructions which can be interpreted as a elaboration process, or to elaborate the data presented with the chosen selection. The user enters the elaboration mode when hovering over the bars or by hovering or clicking on the donuts. When the user is in the elaboration mode, the user is able to choose to go back to the exploring mode or to go further in the deepened elaboration mode by clicking on the artworks. This will set the background to a new selection of artwork or will simply present the artworks of the concerning donut.

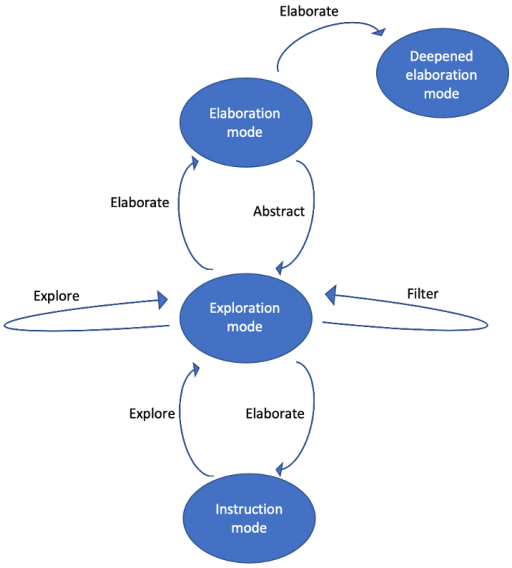

Figure 7: State diagram of the concept of the visualization

5 PRODUCT DESIGN

The actual implemented design differs to some extent from the concept idea. The state diagram in figure 7 is only partly implemented as the deepened elaboration mode is not yet implemented. In figure 8 the view of the visualization tool is presented. The product design of the application will be discussed in this section.

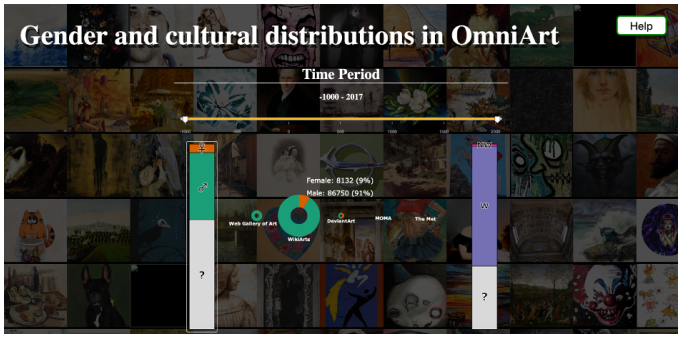

Figure 8: View visualization tool

5.1 Timeline

The final product contains a simple slider timeline consisting of a horizontal line with a width of 60% of the screen. It has two draggable points to set a time period. Above the slider, the exact dates selected with the two draggable points are displayed for clarification. Since the available artworks of our dataset comprise a very broad timeline 1000 B.C. until 2017), it is decided to display the current selected times above the timeline slider. This will give a clear indication on what years are currently selected.

Having such an extensive and broad timeline, it still can be quite a horrendous job to select a very specific timeline. Therefore, a zoom function on the slider is implemented. When the two draggable points on the slider are within a range of 500 years, the slider will "zoom in" on the time around the two draggable points. This will make it easier to select a more specific time. To zoom out, the user can drag the two point to the outer sides of the timeline slider and then the user can select all the times again.

In the implemented timeline, the functionality to select a certain time period and move this time range is not yet implemented. For example, it is not possible to select a time period of 50 years between 1800 and 1850 and slide that time period to a later or earlier time period.

5.2 Flying donuts

The donut charts of the implemented visualization tool, are dependent on the data selected by the user by selecting a time period and a group. Unfortunately, the ability to select schools, artwork type or art collection by the user is not yet possible.

In addition, the size of the donuts represent the quantification of the group. However, visualizing the donuts in a floating structure according to their ratio is not yet achieved. The non-floating part has been solved by putting all donuts next to each other, such that there is a "floating" experience to some extend.

Furthermore, hovering over the donuts will give the exact amount of artists represented in the donuts. However, representation of the artworks is missing. Therefore, the user is also not able to click on the paintings to gain more information about the paintings created by the artists.

5.3 Bar interaction

The bar charts are represented as initialized as they represent the data from the selected time period. Furthermore, the user is able to hover and click on the bars as was intended. Unfortunately, the ability to filter one group with another group is not implemented.

5.4 Help

The help button and section was not the biggest challenge in the development of the visualization application. Therefore, the button and section are developed as planned in the concept.

5.5 Loading screen/icons

The loading screen and icons have been developed as initialized. It has been decided to show the text "OmniArt" in the middle of the screen when the application, including the images for the background, is loaded. Additionally some movement on the screen was needed as a loading indication. Given the fact that we are visualizing certain aspects of an artwork dataset, we decided to draw 4 lines asynchronous around the word "OmniArt" having the colors green, blue, red and yellow. A gesture of honor to the dutch artist Piet Mondriaan. As the loading finishes, it will fade away and the application screen, consisting of the help button, timeline, donuts and side bars, will become visible

5.6 Background design

We had various conceptual ideas about the way we should shape and fill in our background. In the end we were able to implement a background consisting of images randomly picked from the OmniArt dataset. However, the background will stay static after the application is loaded for the first time. We made this decision due to the fact that some images from the dataset were quite larger in terms of memory, making the loading of the images a unpleasant surprise which we solved with a longer lasting loading screen. For our concept, we wanted the background to change on user interaction. For example, when a user clicked on a donut or changed the time period, we wanted the background to show only images of the selection done by the user. The image sizes complicates this. We therefore stick to the first initial background randomly picked from the whole dataset.

6 OBSERVATIONS

We asked 10 test subjects to rate our tool to get more understanding about the operation of the tool. We asked 4 questions about the tool and asked for open feedback. This gave the following results:

6.0.1 *Closed questions.*

(1) In what extent do you find this tool user friendly?
Result: (3.8/5)
(2) Is the goal of the visualization clear?
Result: (4/5)
(3) How interesting do you find the content displayed by the visualization?
Result: (3.9/5)
(4) How well did this tool inform you?
Result: (3.4/5)

6.0.2 *Open feedback.*

- It would be nice if there was a legend for charts and timeline
- The background is a bit distracting for me
- I would prefer to have a more clean and rounded font, e.g. sans serif
- It would be nice if you hover over bars that it only shows numbers of the specific element, e.g. only male for the genderbar, or only unknown
- It was not clear to me that help button and bars can be clicked on

7 DISCUSSION

Most of the ideas in our initial concept design were used in our final product. However, there are some slight differences on which we will elaborate in this section. One of the differences between the original concept and the final product is that the donuts have static locations in the visualization, instead of a desired floating location based on user input. Secondly, the users were only able to browse

the results by collection and not by artwork type and genre. Lastly, we had to refrain from an interactive background, due to limited computational power. Nevertheless, the final product contains all basic elements that were stated in the original design. These include: an intuitive timeline, an interactive selection of groups (gender and culture) with clear ratios, and scalable donut charts that essentially show the same ratios, but spread out over different interesting categories. All these elements help bring the goal of our visualization to the user, that is to visualize biases within art in an understandable and intuitive manner. In addition to this, multiple other elements were added in order to improve general user experience. From a small user study, we found that there were mostly notes about the background being too distracting, or some other aesthetic properties being imperfect. However, the numbers showed that the main goal of the visualization more or less came through, which was to inform the user and to raise interest in the topic. There only were a few elements that were unclear, such as the question marks. These elements need to be adjusted and explained more in order to fully bring the visualization to its right, but this visualization is a step toward a more socially aware art culture. It does not solve the problem of biases in art history, but it can certainly make it appealing to the user to explore the inequalities in art that are present, even today. For future development of this tool, any additional data should be explicitly annotated with culture and gender labels, for more reliable resulting distributions. By doing so, studies could rely on our tool in the future to research the distribution change over time, which is a research question discussed in Section 1.3. Since we used manually annotated data with limited knowledge of art theory, we leave this question to following researchers.

8 CONCLUSION

Currently, there exists an inequality among artists regarding gender and cultural background as the artworks presented nowadays are mainly from male and/or western artists. To make people aware of this inequality, an attempt is made to clearly visualize this. For this project, an interactive visualization tool is implemented for the OmniArt dataset. The visualization is made interactive to make users explore the distributions between the categories (i.e. gender and cultural background) trough time.

Our evaluation found that users were moderately informed by the tool, but that the goal of the visualization was rather clear. Feedback on the visualization tool has revealed that an improvement of the design is needed.

Overall, this project has made a contribution to bring awareness of male and western dominance in artwork collections. However, to utilize this application in a real world environment, an improvement of the design of the application is desirable. Furthermore, adding multiple variables to the application may make the application more interesting and may give the user more insights in the distributions.

REFERENCES

[1] [n.d.]. A Brief History of Women in Art (article) | Khan Academy. https://www.khanacademy.org/partner-content/tate/women-in-art/history-of-women-in-art/a/a-brief-history-of-women-in-art. (Accessed on 03/17/2020).

[2] [n.d.]. List of most expensive paintings - Wikipedia. https://en.wikipedia.org/wiki/List_of_most_expensive_paintings. (Accessed on 03/19/2020).

[3] 2019. Museums in US still failing with artist diversity, study finds | Art and design | The Guardian. https://www.theguardian.com/artanddesign/2019/mar/20/museums-in-us-still-failing-with-artist-diversity-study-finds. (Accessed on 03/18/2020).

[4] Andrienko N. Demsar U. Dransch D. Dykes J. Fabrikant S.I. Jern M. Kraak M.-J. Schumann H. Tominski C. Andrienko, G. 2010. Space, Time, and Visual Analytics. *International Journal of Geographical Information Science* 24, 10 (2010), 1577–1600.

[5] Gustav A Berger and William H Russell. 1984. The new stress tests on canvas paintings and some of their implications on the preservation of paintings. In *Icom 7th triennial meeting. Copenhagen, 10-14 September 1984. Preprints*. 84–2.

[6] Norma Broude. 2018. *Feminism and art history: Questioning the litany*. Routledge.

[7] Goetzmann W.N. Nozari M. Cameron, L. 2019. Art and gender: market bias or selection bias? *J Cult Econ* 43 (2019), 279–307.

[8] Kelly Crow. 2017. Leonardo da Vinci Painting 'Salvator Mundi' Smashes Records With $450.3 Million Sale - WSJ. https://www.wsj.com/articles/leonardo-da-vinci-painting-salvator-mundi-sells-for-450-3-million-1510794281. (Accessed on 03/19/2020).

[9] J Bradford De Long. 1988. Productivity growth, convergence, and welfare: comment. *The American Economic Review* 78, 5 (1988), 1138–1154.

[10] Lisa Dupuy. 2019. 'Drie musea hanteren quota voor werken vrouwelijke kunstenaars' - NRC. https://www.nrc.nl/nieuws/2019/12/27/drie-musea-hanteren-quota-voor-werken-vrouwelijke-kunstenaars-a3985042. (Accessed on 03/18/2020).

[11] Ernst Hans Gombrich and EH Gombrich. 1995. *The story of art*. Vol. 12. Phaidon London.

[12] Catherine Grant and Dorothy Price. 2020. Decolonizing Art History. *Art History* vol 43 (2020), 8–66. Issue 1.

[13] Hugh Honour and John Fleming. 2005. *A world history of art*. Laurence King Publishing.

[14] Y. a. Kang J. S. Yi and J. Stasko. 2017. Toward a Deeper Understanding of the Role of Interaction in Information Visualization. *IEEE Transactions on Visualization and Computer Graphics* 13, 6 (2017), 1224–1231.

[15] Linda Nochlin. 1971. Why have there been no great women artists? *The feminism and visual culture reader* (1971), 229–233.

[16] Emily Sharpe and José da Silva. 2019. Art's Most Popular: here are 2018's most visited shows and museums | The Art Newspaper. https://www.theartnewspaper.com/analysis/fashion-provides-winning-formula. (Accessed on 03/18/2020).

[17] Dean R Snow. 2006. Sexual dimorphism in Upper Palaeolithic hand stencils. *antiquity* 80, 308 (2006), 390–404.

[18] Lucette ter Borg. 2019. Het nieuwe MoMA bewijst dat geschiedenis te herschrijven is - NRC. https://www.nrc.nl/nieuws/2019/11/05/geschiedenis-is-dus-te-herschrijven-a3979201. (Accessed on 03/18/2020).

[19] Yves Vequaud. 1977. *Women Painters of Mithila*. Thames & Hudson.