Interactive Visualization of Gender Inequality in China

Yunyi Zhu
MIT CSAIL
yunyizhu@mit.edu

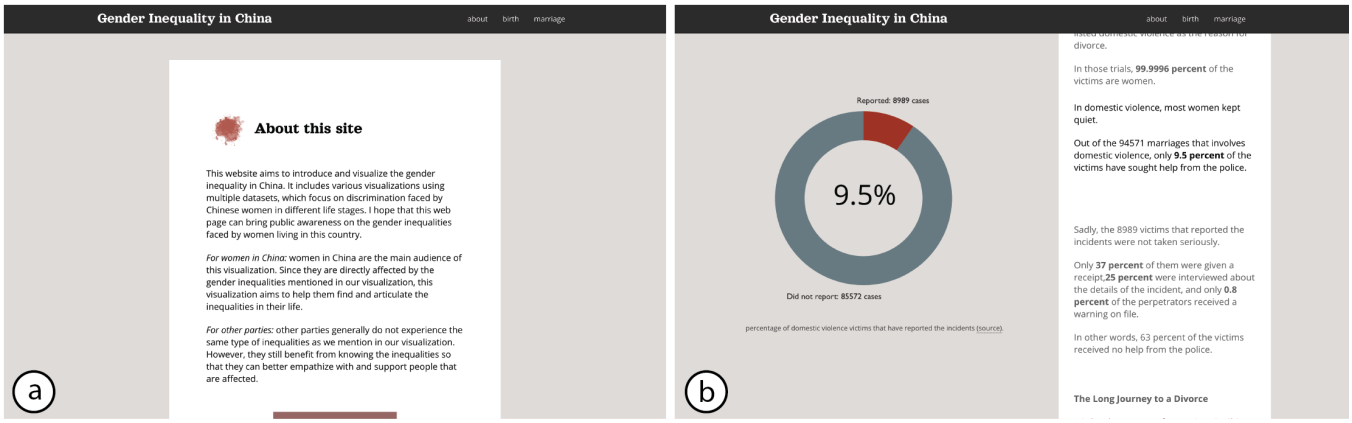

Figure 1: (a) Layout of the website's entry page. (b) Layout of the visualization together with text.

ABSTRACT

In this project, we develop an interactive website that introduces and visualizes the gender inequality in China. It includes visualizations using multiple datasets that focus on discrimination faced by Chinese women in different life stages. For each life stage, it uses scrollytelling to display the data together with paragraphs that provides the contexts of the data. We hope that this website brings awareness to different parties on the gender inequalities faced by women living in China.

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI)**.

KEYWORDS

interactive data visualization; gender equality.

ACM Reference Format:
Yunyi Zhu. 2018. Interactive Visualization of Gender Inequality in China. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/1122445.1122456

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

In recent years, online feminist movements in China have surged due to the increase of female users and female-majority communities. Despite its growth, however, it is still widely misunderstood and stigmatized as "man-hating". Moreover, since feminist movements are generally not supported by the Chinese government, they are in danger of being labeled as anti-government and thus canceled, both online and offline.

In the midst of the debates, many who do not support the online feminist movements argue that gender inequality is nonexistent in China. In particular, they argue that they are not convinced by non-factual data (survey data and indices) and think that the gender equality index from United Nations is inconsistent with their own observation, where Chinese women's life is "better than men's" [9].

To prove that gender inequalities exist, we decide to visualize the inequalities using factual data from either Chinese government database or websites with mostly Chinese users. The audience of this visualization includes both women in China and other parties.

For women in China: women in China are the main audience of this visualization. Since they are directly affected by the gender inequalities mentioned in our visualization, this visualization aims to help them find and articulate the inequalities in their life.

For other parties: other parties generally do not experience the same type of inequalities as we mention in our visualization. However, they still benefit from knowing the inequalities so that they can better empathize with groups that are affected.

For this reason, we have built a scrollytelling-based website (Figure 1) that visualizes discrimination against women in China in different stages of women.

2 RELATED WORKS

Past works about gender equality includes visualizations on gender equality and equity and documentation of feminist movements in China.

2.1 Gender Equity Visualization

Many data visualization have focus on the topic of gender equity and equality. For instance, Unwanted [3] is a scrollytelling website that educates the audience about gender selection in newborns in India (Figure 2). This project uses various data, including statistical data visualized in number, graphs and maps as well as empirical data such as quote and surveys. In our work, we focus on the similar issue with different data sets. We aim to perform more calculation and thus directly visualize the amount the women that are missing.

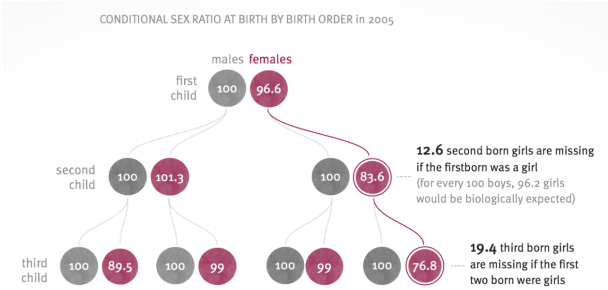

Figure 2: Project "Unwanted" visualizes newborn sex ratios conditioned on birth order.

Many infographics projects also focus on feminism related issue, such as the lack of representation of women in the film industry [2], in politics [5], and in the society in general[8].

In addition, in the creation of the visualization, we try our best to stick to the guidelines introduced in Data Feminism [17] and Ethical Dimensions of Visualization Research [16].

2.2 Documentation of Feminist Movements in China

Many artists, social activist communities and data visualization scientists in China have worked on the issue of gender inequality. For example, online artist Fei-Jia-Luo-Fu-Ren created a set of artworks illustrating the different types of discrimination and humiliation in a woman's life [15] such as intelligence shaming and virginity shaming. Collective efforts by the web users, such as the Questions to JiangShanJiao movement [7], also call into attention on types of gender discrimination that are hard to quantify, such as parent attention and divorce shaming. In addition, data scientists and social scientists also focus on gender inequality. For examples, SiXiangStudio [11] created visualization on a hundred divorce cases, which depicts the lack of divorce freedom and domestic violence protection in China.

3 METHODS

Developing this scrollytelling web page involves multiple steps: (1) acquiring raw data from various sources, (2) calculating the data to fit the need for visualizing, and (3) developing the web page user interface to help with interaction.

3.1 Data Acquisition

In this project, we gathered data sets from various sources. While we tried to stick with numeric and factual data, we also included useful survey data and community collection data. For places where different data sources contradict, we provide a note on the visualizing marking the different sources.

Birth data: data about birth inequality (newborn sex ratio) is the easiest to look for since the topic is well studied. In particular, we used the census data from 2010 for the new born gender ratio data of different child order in 2009 [1]. For the global data, we used data from the world bank [6]. Since the newborn ratio from the World Bank data is different from the census data, we provided a note in the visualization. In addition, we gathered news data from different new pages and display related news pages to accompany the data.

Marriage data: marriage and domestic violence data are harder to find since they are less studied and collected. Therefore, we found existing pre-calculated data that are published by other people. In particular, the domestic violence data are from the report written by Jiezhen Wu [18] and the divorce data are from the White Paper of Domestic Law Service in 2018 [4]. We also collected court papers from the China Court Online [10]. However, since the court papers are mostly natural language data, we put it in future work to analyze and visualize them.

3.2 Data Calculation

We performed various calculation on the data that we gathered (1) directly visualize the magnitude of the inequality and (2) fit the data with the visualization library that we use.

Birth data: the data we collected includes the number of male and female babies born in 2009. In our visualization, we computed the number of missing girls (N) by subtracting the the number of newborn girls (N_F) from the number of girls that should have been born. The number of girls that should have been born were calculated by dividing the world average gender ratio (1.05:1) from the number of newborn boys (N_M). We wrote this in an equation below.

$$N = \frac{N_M}{1.05} - N_F$$

3.3 Web Design and Development

The website is built in the *React* framework [13] and the data visualizations are built with the *nivo* [12] and *Victory* [14], which are data visualization libraries based on *D3.js* that are compatible with React.

The major design challenge of this visualization is to deal with the scarcity and the incoherence of the datasets. In particular, the datasets that I gathered are not very related and are thus hard to function in one interactive system. Furthermore, due to the lack of data on education and work inequality, it is hard to put together one narrative for those datasets, which are all important representatives of the gender inequality in China.

To tackle this, I made two separate narratives, the "birth" and the "marriage" sections. Each narrative includes a scrollyteller on one aspect of the inequality. In the original design as well as the future work, I plan to gather more data and make it into a comprehensive timeline including "birth", "education", "work", "marriage" and "inheritance".

4 RESULTS

In this section, we present some of the visualizations that the website has and the design choices that lead to them.

Explaining the newborn sex ratio: the most important factor about birth inequality is the newborn sex ratio, which represents the ratio between newborn male infants and newborn female infants. To make sure that the audience understand the meaning of newborn sex ratio, and the implication of a high newborn sex ratio, we represented the newborn male and female as dots and visualizes the missing female babies using air bubbles (Figure 3).

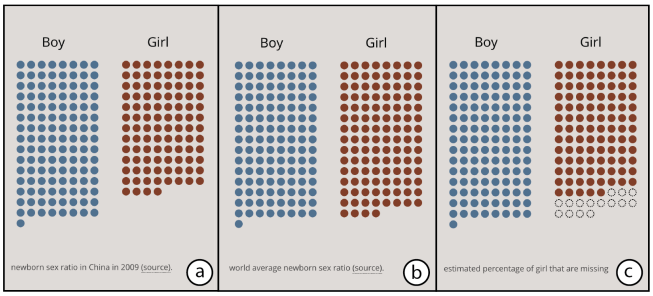

Figure 3: (a) Newborn sex ratio in China; (b) world average newborn sex ratio; (c) visualisation of the proportion of girls that are missing.

Scrolly timeline visualization: to provide contexts for the high newborn sex ratio in China, we visualized data of the worldwide newborn sex ratio. Instead of using a line chart and highlight the line for China, we opt for a more animated interaction, a scrolly timeline. In particular, we allow the user to navigate through the vertical timeline using the basic scroll interaction (Figure 4). The visualization shows a color coded map for that year in each frame. While such interaction obfuscates the exact trends of the newborn sex ratio, it has interesting interaction and highlights the fact that China has the highest ratio in every year.

Visualizing layered ratio: when visualizing the police inaction on the reported domestic violence cases, we needed to visualize layered ratio. In particular, we need to visualize the ratio between reported and receipts, receipts and interviews as well as interviews and

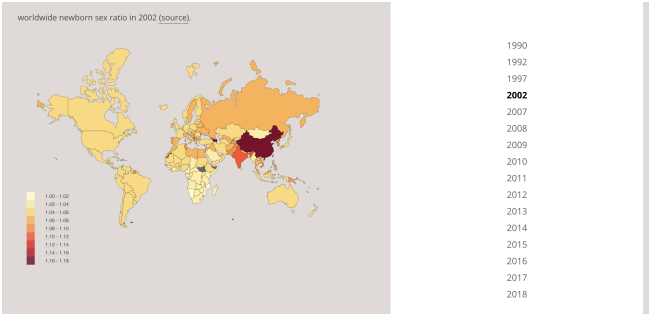

Figure 4: This scrolly timeline combines the interaction of a scrolly teller with traditional dragging timeline.

actual results. We tried using multiple pie charts as well as vertical bar charts, but none were effective. In the end, we found that a horizontal bar chart with sufficient labels effectively illustrate the idea of layered ratios (Figure 5).

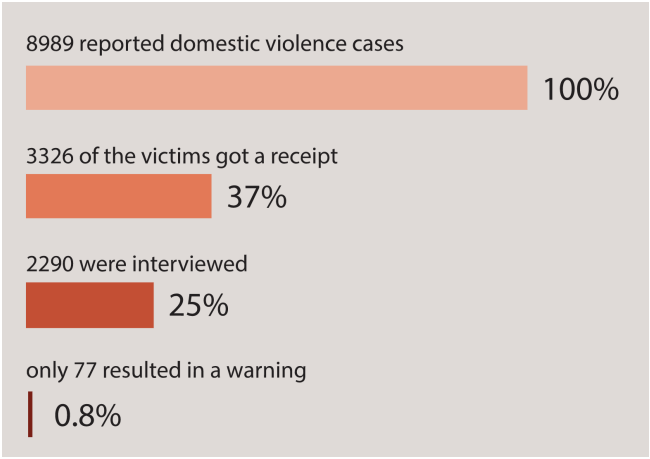

Figure 5: Visualization of police inaction on reported domestic violence cases.

Non-numerical data: apart from numerical data, factual data such as discrimination policies and important news articles also help the audience understand the topic. Therefore, at the end of each section, we display a list of summaries to the related factual data as well as the sources. This serves as a starting point for the audience to learn more about gender inequality in China (Figure 6).

5 DISCUSSION AND FUTURE WORK

While this visualization website is usable, it can benefit largely from additional data and better graphics. In the future, I plan to turn this project into a long-term personal project and perform the following improvements.

Data about education and work inequality: the education and work inequality in China are subtle. From a naive glance at the data, the working participation of women in China is relatively high.

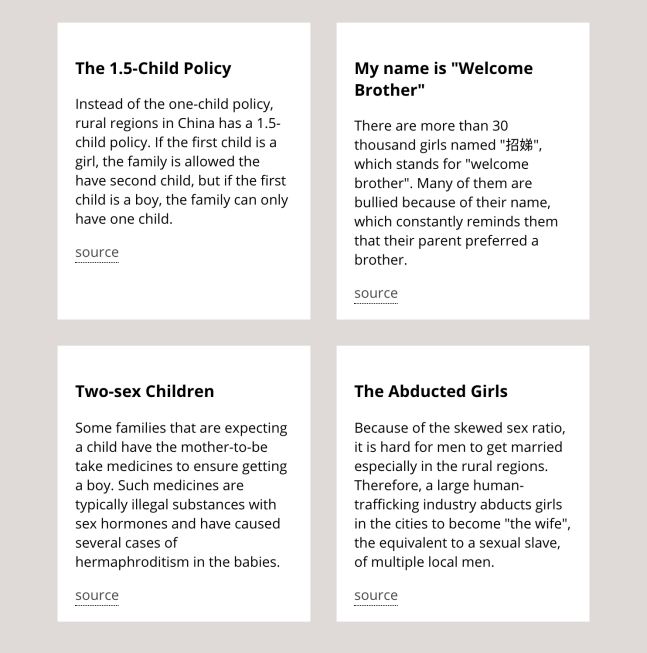

Figure 6: User can browse a list of factual data such as laws and important news articles to learn more about the topic.

However, what typical datasets do not catch is the disadvantage of women when they are looking for a job. For example, some governmental jobs put percentage caps on the number of women they hire. Such data are hard to find because the documents are usually private, but in the long run, I want to find a way to visualize this.

Other ways to visualize: I plan to try out other ways to visualize the existing data that I have. While scrollytelling is a good way to layout the data together with the texts, traditional infographics and articles might also achieve the same effect.

Translate the website to Chinese: since women living in China are the most important audience of this visualization, I plan to translate the website to Chinese. I will still keep the English version and allow the viewer to choose the language they view in.

REFERENCES

[1] 2010. *China Census Data 2010.* http://www.stats.gov.cn/tjsj/pcsj/rkpc/6rp/indexch.htm
[2] 2013. *Gender Inequality in Film.* https://www.nyfa.edu/film-school-blog/gender-inequality-in-film/
[3] 2015. *Unwanted, the ongoing war against daughters in India.* https://unwanted.interactivethings.io/#/introduction
[4] 2018. *2018 White Paper on Domestic Law Service.* http://www.jialilaw.com/file/2018%E5%B9%B4%E5%A9%9A%E5%A7%BB%E5%AE%B6%E4%BA%8B%E6%B3%95%E5%BE%8B%E6%9C%8D%E5%8A%A1%E8%A1%8C%E4%B8%9A%E7%99%BD%E7%9A%AE%E4%B9%A6.pdf
[5] 2018. *Women in the House of Representatives.* https://pudding.cool/2018/07/women-in-congress/
[6] 2018. *World Bank Newborn Gender Ratio Data.* https://data.worldbank.org/indicator/SP.POP.BRTH.MF?locations=CN
[7] 2020. *Questions to Jiang Shanjiao, the Patriotic Virtual Idol.* https://chinadigitaltimes.net/chinese/635845.html
[8] 2020. *Visualizing the data: Women's representation in society.* https://www.unwomen.org/en/digital-library/multimedia/2020/2/infographic-visualizing-the-data-womens-representation
[9] 2021. *Are women more privileged than men in China?* https://www.zhihu.com/question/296532752
[10] 2021. *China Court Online.* https://wenshu.court.gov.cn/
[11] 2021. *How hard is it to get divorced in China.* https://www.sohu.com/a/412403860_157078
[12] 2021. *Nivo.* https://nivo.rocks/
[13] 2021. *React.* https://reactjs.org/
[14] 2021. *Victory: React.js components for modular charting and data visualization.* https://formidable.com/open-source/victory/
[15] 2021. *Women should be whoever they want to be.* https://m.weibo.cn/detail/4637797896685420
[16] Michael Correll. 2019. Ethical Dimensions of Visualization Research. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300418
[17] Catherine Dignazio and Lauren F. Klein. 2020. Data Feminism. (2020). https://doi.org/10.7551/mitpress/11805.001.0001
[18] Jiezhen Wu. 2017. *Analysis on Domesic Violence Lawsuits in 2017.* https://zhuanlan.zhihu.com/p/26749311