

Visualising the Global Community: Women in STEM, Education, and the Workplace

Ella Tait

Abstract—Women often face unique challenges, which can go unnoticed by their peers, especially in education, workplace leadership, and the field of Science, Technology, Engineering, and Mathematics (STEM). This can lead to women facing obstacles in their careers due to gender stereotyping and discrimination. This project’s goal is to raise awareness of these issues faced globally by women. To achieve this, a web application has been developed that displays interactive data visualisations using the D3.js JavaScript library. A user study has also been undertaken to assess its effectiveness and usability.

I. INTRODUCTION

A Lack of awareness about gender disparity in the workplace can lead to women being ‘overworked and under-recognized’ [1]. As well as this, those who are aware of this disparity may not know what they can do about it.

A. The Problem

The problem that this project aims to address is the lack of awareness about gender disparity in education, workplace leadership, and STEM by providing evidence and a call to action to explore local initiatives.

It is important to solve this problem as this lack of awareness can lead to fewer opportunities and negative experiences for women, not based on their abilities, but on stereotyping and discrimination [2]. This continued systemic discrimination can cause women to have to work harder than their male peers for less recognition and can lead to burnout [1].

B. The Solution

The solution developed to address the problem of lack of awareness about gender disparity in education, workplace leadership, and STEM is a web application that displays interactive visualisations of data about women over the past decade. It will provide a starting point for people wanting to increase their awareness of these issues. The web application will be made up of three different visualisations, each displaying information that the user can interact with.

The user base of this solution will not be a large group, but rather people who are actively interested in improving either their own or others’ awareness about these issues and need a place from which to begin that will allow them to discover further initiatives, which can also be submitted to the system by users.

This project was supervised by Dr. Stuart Marshall (primary).

C. Final Product

The final implementation of the web application includes three visualisations: the STEM Graduates visualisation is a diverging bar chart (See Fig. 1), the Workforce visualisation is a beeswar graph (See Fig. 2), and the Education visualisation is a connected dot plot (See Fig. 3). All three visualisations allow for filtering by continent and country, and can be hovered over to display further details in the Information and Resources boxes. The Workforce visualisation also allows for filtering by year, while the Education visualisation allows for filtering by gender. New resources can be added by users, and these are sent to the moderation page before being added to the main site.

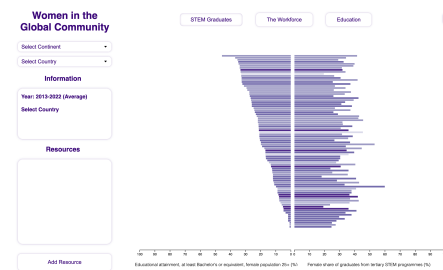


Fig. 1. STEM Graduates Final Visualisation.

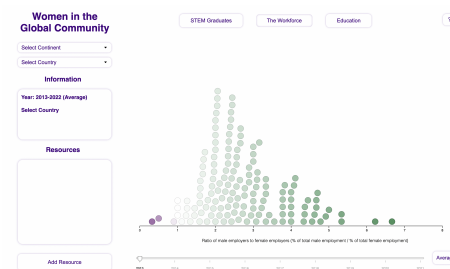


Fig. 2. The Workforce Final Visualisation.

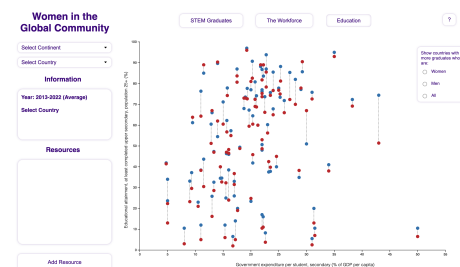


Fig. 3. Education Final Visualisation.

D. Evaluation Key Findings

A user study was undertaken to evaluate the usability of the web application, as well as its ability to convey information to increase awareness about gender disparity. The results showed that while the web application interface was usable and engaging, the information displayed by the visualisations was sometimes difficult to understand for some users. We sought to address this with the addition of a 'help' button, but it is also important for the system to be intuitive to engage casual users.

E. Environment and Sustainability

This project attempts to begin to solve three of the UN Sustainable Development Goals [3]. Goal four, "Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all", is addressed in the first and third visualisations using data that involves women's education. The aim of the visualisations is to bring awareness to inequalities for women in education. Similarly, all three visualisations address goal five, "Achieve gender equality and empower all women and girls", and goal ten, "Reduce inequality within and among countries", as they are trying to solve the problem of lack of awareness about gender disparity. As well as this, the visualisations provide a call to action for users to explore organisations and initiatives that are taking active steps to achieve these goals and submit their own initiatives.

II. REQUIREMENTS

A. User Personas

1) *Primary - Nick*: Nick is a software developer who has recently joined a new team. The women in his team have been discussing the struggles they have faced in the industry, and he is not sure how widespread these problems are. He wants to do his own research about the issues women may face globally in the workplace, leadership, or STEM so that he can understand and support the women in his team better. His goal is to be able to understand the women in his team and their experiences better. He also wants to learn more about issues women have faced globally in the workplace, leadership, and STEM. A barrier for Nick is that he finds existing tools overwhelming and does not know where to start. He also wants to see objective data from a reputable source as he finds just talking to individuals about their experiences very subjective.

2) *Primary - Bea*: Bea is a university student in New Zealand who has started her own initiative to encourage women to pursue careers in STEM. She wants to find information about other initiatives around the world but is not sure where to start. Bea's goal is to be able to see which countries and continents have higher rates of women in STEM so that she can narrow down where she should start looking for examples of successful initiatives. She also wants to be able to find resources about Women in STEM initiatives in other countries and share her own initiative with people around the world. However, she is having trouble getting the word out about her new initiative to anyone other than the people in her class.

3) *Secondary - Taylor*: Taylor is a digital technology teacher at a co-ed secondary school. They want to convince the school to start providing extra support for girls who want to get a degree in a STEM field when they leave school. Taylor's goal is to provide the senior leadership team data from a reputable source and have an engaging and visually interesting presentation. They need to know that the team will find it easy to use themselves and will be able to discover more information if needed. However, none of the senior leadership team have experience in STEM fields. Taylor will also be presenting in-person on a projector screen which can be hard to see when the contents of the screen are low-contrast.

B. Requirements Analysis

In the previous section, I identified the primary and secondary user personas. The main requirements of these users are that the web application:

- 1) Is easy to use
- 2) Is high-contrast
- 3) Is engaging
- 4) Shows global information
- 5) Allows filtering based on region
- 6) Provides access to resources about initiatives that are local to each country
- 7) Allows users to submit their own initiatives to the system

C. Data

Columns:

- Series Name
- Series Code
- Country Name
- Country Code
- 10 columns, one for each year 2022 [YR2022] - 2013 [YR2013] inclusive
- Continent

STEM Graduates Series:

- Female share of graduates from STEM programmes, tertiary (%)
- Educational attainment, at least Bachelor's or equivalent, population 25+, female (%)
- Population, total

The Workforce Series

- Employers, female (% of female employment)
- Employers, male (% of male employment)
- Employers, total (% of total employment)
- Wage and salaried workers, female (% of female employment) (modeled ILO estimate)
- Wage and salaried workers, male (% of male employment) (modeled ILO estimate)

Education Series:

- Educational attainment, at least completed upper secondary, population 25+, female (%) (cumulative)
- Educational attainment, at least completed upper secondary, population 25+, male (%) (cumulative)
- Government expenditure per student, secondary (% of GDP per capita)

III. RELATED WORK

The interactive data visualisations in the project follow the information visualisation taxonomy laid out in Schneiderman's 1996 paper, "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations" [4]. In it, he defines seven high-level tasks that users may wish to perform: "overview, zoom, filter, details-on-demand, relate, history, and extract". This project focuses primarily on the three of the tasks, 'overview', 'filter', and 'details-on-demand'. Each visualisation starts with a high-level overview of the data, and the user is able to filter the visualisation and get additional information about the data. In order to review related work from the perspective of this project, it is necessary to use the requirements generated by the user personas as criteria against which the related work should be evaluated.

A. Literature Review

This literature review seeks to explore and compare existing interactive visualisations of gender-based data across academic institutions, non-government organisations (NGOs), and commercial organisations to locate potential areas of improvement and opportunities for further development within the proposed ENGR489 project.

1) *Understanding Gender Equality Through Visualisations*: The first set of interactive data visualisations is "Understanding Gender Equality Through Visualisations" [5] (See Fig. 4), an ENGR489 project completed in 2017 by Divya Patel. It uses three datasets: global male and female education and labour force participation rates, global male and female gross national income, and the global gender equality index. This data is used to produce four types of visualisations: a choropleth map, dot plot, parallel coordinate graph, and a bubble graph. Three of these are visualisations for the individual datasets, with the fourth, the bubble graph, allowing users to compare their selection of two of the datasets.

The visualisations also have a help button with information about each visualisation which is useful for users who need further clarification. The visualisations which display countries as the data points use colour to group the countries by the continent they belong to. This helps the user get an overall understanding of the data in relation to continents while being able to interact with the graph to get more information, i.e., the country name.

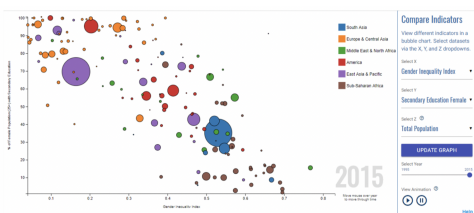


Fig. 4. Understanding Gender Equality Through Visualisations.

2) *Global Gender Gap Report*: The second interactive data visualisation is the Global Gender Gap Report [6] (See Fig. 5), produced by the World Economic Forum, an NGO. It uses histograms and scatter plots to provide an in-depth

look at the data from the Global Gender Parity Index. The visualisation transforms as the user scrolls down the page, using smooth transitions to display the data in new ways. The page has a description on the left side, which changes with the visualisations as the user scrolls and provides further information about the data being displayed. The data points represent countries, and similarly to the first visualisation, are coloured by continent. Hovering over an individual point reveals more information, adding points for that country over the past few years with a trend line.

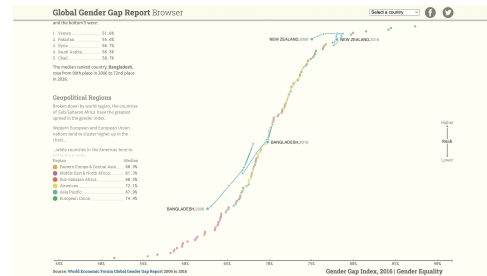


Fig. 5. Global Gender Gap Report.

3) *What's Your Pay Gap?*: The third visualisation is "What's Your Pay Gap?" [7] (See Fig. 3)6), by the Wall Street Journal, a commercial organisation. It focuses on the gender pay gap in the United States, using the median earnings for men and women across different professions to create a scatter plot. There are many points on the scatter plot, so the graph moves to expand where the mouse hovers. However, this makes it difficult to select a specific point to see which profession it represents, as the point moves as the mouse moves towards it. The overall picture of the pay gap is easy to see at first glance, but the method for revealing further information is not user-friendly.

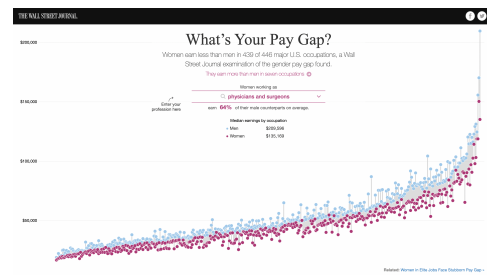


Fig. 6. What's Your Pay Gap?

4) *Gapminder*: Gapminder [8] (See Fig. 7) is a non-profit organisation which has created a website containing many different interactive data visualisations, all relating to its central mission of "Fighting devastating misconceptions with a fact-based worldview everyone can understand". The website's interface allows users to compare a wide range of datasets using different visualisations to display the data. Colour is used to add information from additional datasets and a time slider allows users to see the data from other years both statically and dynamically through an animation sequence. The visualisations provide a large amount of information in an easily understandable interface, and the idea of being able

to filter by certain values will be useful in my own project. However, while the wide range of information available is useful for the organisation’s objectives, it means that the site requires the users to select relevant information relevant to them which could be overwhelming to new users who may not know where to start, like some of the target users identified in Section 2.1.

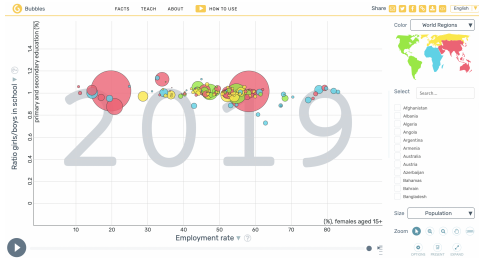


Fig. 7. Gapminder.

5) *OECD Gender Data Portal*: The Organization for Economic Cooperation and Development (OECD) Gender Data Portal [9] was developed by the OECD, an international organisation. The data is kept up-to-date, with data from 2022-23 provided where possible. The website’s interface is a dashboard-style layout, with multiple different interactive graphs available on panels. Users are able to find further graphs and information through links provided at the bottom of each panel. They are also able to view the same graphs with more detail and interactivity by clicking their titles. Within these more detailed pages, users are able to filter the graphs by country, time, and series. There are different types of graph available, including bar charts, dot plots, and maps. Having an initial overview is useful for seeing which data options are available, but the graphs are simplified and therefore do not tell the whole story of the data. The more detailed graphs that can be discovered on demand are much more useful for understanding the data as they are more interactive, which shows the importance of filtering and interactivity.



Fig. 8. OECD Gender Data Portal.

6) *World Bank Gender Statistics Portal*: The final set of visualisations is the World Bank Gender Statistics Portal [10] (See Fig. 9), created by the World Bank, a commercial organisation. This covers a large range of datasets, which the World Bank refers to as “Data on key gender topics. Themes included are demographics, education, health, labor force, and political participation.” The visualisations that can be created

using these datasets are a line graph and a map. The data can also be viewed in a table.

As the datasets can include up to 265 different countries, they can be so large that it is difficult to draw any useful conclusions from the visualisations as the lines of the line graph tend to be densely packed together. There is also no way to compare one dataset to another, which would help to make more in-depth observations. The colours used for the line graph are very similar to each other (shades of blue and green) which makes distinguishing the individual lines difficult even when they aren’t tightly packed. The map does not display any datasets currently, and the table is not a visual representation, so it can be difficult to draw any conclusions from some of the datasets when the line graph is hard to read. The interface provides many different options to customise the data displayed, but very few options to customise the visualisations themselves.

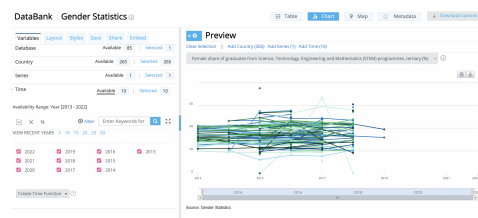


Fig. 9. World Bank Gender Statistics Portal.

7) *Discussion*: The contents of the datasets used to produce interactive visualisations are of great importance. In this case, we are looking for data that focuses on women in STEM, leadership, and education. While they include many different options for gendered data comparison, of the six visualisations presented above, only the OECD Gender Data and World Bank Gender Statistics Portals focus on all three points. This indicates that there could be a need for these three topics to be visualised especially as there are improvements that can be made to how this data is displayed.

As identified previously, the user personas for this project require that the visualisations allow them to see global information, filter based on region, access resources about initiatives that are local to each country, and submit their own initiatives to the system. All except “What’s Your Pay Gap” allow users to see global information and separate the data by region. However, none of these existing resources allow users to access further resources about initiatives, let alone submit their own for others to access.

As well as this, the users require that the visualisations are easy to use, high-contrast, and engaging. The World Bank Gender Data Portal is somewhat difficult to use for the first time, whereas the Global Gender Gap Report has a simple and intuitive interface, with information on how to use the site where needed without overcrowding the page. This is partially thanks to the fact that it is focusing on one dataset only, which allows it to set out the data with plenty of space and limited user choice. What’s Your Pay Gap has clearly been designed to be easy to use, but the implementation does not achieve this in practice. Gapminder and the OECD Gender Data Portal are the most balanced of the group, with a large amount of

TABLE I
COMPARISON OF TOOLS

	Global information	Region filtering	Call to action (further resources)	Users can submit own initiatives	Easy to use	High contrast	Engaging	Women in STEM	Workplace leadership	Education
Understanding Gender Equality	x	x			x	x	x			
Global Gender Gap Report	x	x			x	x	x			
What's Your Pay Gap						x	x	x		
Gapminder	x	x			x	x	x			x
OECD Portal	x	x			x	x	x	x	x	x
World Bank Portal	x	x				x		x	x	x
My ENGR489 Project	x	x	x	x	x	x	x	x	x	x

information being displayed in a relatively simple and intuitive interface.

These differences can be seen in Table I, along with a comparison to this project.

B. Tools and Methodology

The tools that have been used to complete the implementation of the project were chosen for their ability to aid Schneiderman's information visualisation taxonomy or the Agile development process. The web application is being developed using JavaScript, with the visualisations being developed using the D3.js JavaScript library [12]. This library supports filtering and information discovery, which is important to the chosen visualisation methodology. A GitLab repository was being used for Agile development, allowing for issue tracking.

IV. DESIGN

A. Wireframe

The wireframe for the design of the web application, as seen in Fig. 10, is based on the needs of the user personas. It uses a dashboard style layout, which allows for the design to stay simple and intuitive for the user. The user can switch between the three graphs using the buttons and can filter the graph using the options on the left. They can also select a country on the graph and see the relevant information and resources for that country in the bottom left, with the option to add their own resources also available via an 'Add Resource' button. This button will trigger a pop-up form that will allow users to enter all the necessary information about their organisation, including a name, country of origin, description, and link to its website. As with any web form, this input will need to be sanitised to prevent Cross-Site Scripting (XSS). As the input will also contain links, it is important that the suggested resources are able to be moderated before appearing on the main page so that malicious, unwanted, or irrelevant links or information does not get published on the main page. This will be done through a moderation page that will allow admin users to approve or deny new resources added by users.

The three types of visualisation are a diverging bar chart, a beeswarm graph, and a connected dot plot. The first is used because it is important to show the differences in the two datasets. The beeswarm graph is used to help show clusters of values along a single axis, while the connected dot plot allows three datasets to be displayed and compared against each other.

Purple has been chosen as the base colour of the web application, with green as a secondary colour when needed. This is due to purple's historical significance to women's liberation movements [?] and purple, green, and white being the colours of International Women's Day [13]. Red and blue were used for the third visualisation as they allowed for a more clear distinction between the smaller female and male points on the graph.

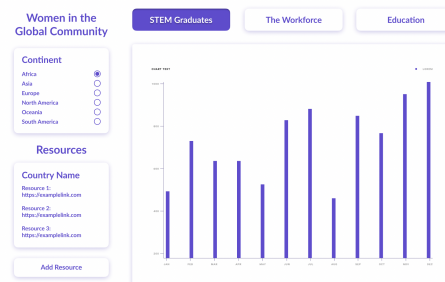


Fig. 10. Wireframe.

B. System Architecture

The system architecture for this web application is a client-server model. The application will pull from the application server storing the data (CSV file), as well as the D3.js library from D3.org, meaning there are two servers and one client.

The application will be run on a local PHP development server and uses the JavaScript library D3.js to create the interactive visualisations. D3.js is a library which works with scalable vector graphics (SVGs) and CSS to create visuali-

sations within the HTML DOM in a standard web browser across multiple different platforms.

The application is made up of multiple components. The server stores the application itself, as well as the data (CSV) because it is not a large amount of data to store (536 kB). Within the application, there are the visualisation, resources, new resource, and moderation components. The visualisation component is made with JavaScript using the D3.js library, while the resources component uses a JSON file for the data and JavaScript for its display. The new resource and moderation components both have a PHP backend, which processes the data, and use HTML to create the form to submit. CSS is used to style all elements of each component.

C. Sustainability Considerations

An important environmental consideration is the consumption of energy caused by computation. Large computations can consume a large amount of energy, but this is not currently a concern as the user base, dataset (536 kB), and D3.js library (873 kB [14]) are small enough that the energy consumption is not high. As the data is unlikely to change quickly as it is updated yearly, caching could be an option in the future to reduce network traffic if multiple sessions were in use.

V. IMPLEMENTATION

The implemented web application is made up of three visualisations, along with a sidebar for filtering and details-on-demand. They are detailed in the following sections.

A. STEM Graduates

The first visualisation implemented is the STEM Graduates graph. The datasets used for this visualisation are 'Female share of graduates from Science, Technology, Engineering and Mathematics (STEM) programmes, tertiary (%)' and 'Educational attainment, at least Bachelor's or equivalent, population 25+, female (%) (cumulative)'. It compares the values of each dataset for each country that has values for both. While the final implementation was a diverging bar chart, this graph went through many iterations.

1) *First Implementation:* The first implementation of the STEM Graduates graph was a stacked bar chart, simply showing the proportion of male to female STEM graduates. However, the stacked bar chart design was not visually appealing, and the limited information shown by the graph was not engaging.

2) *Second Implementation:* The second implementation of this visualisation was when the final graph type, a diverging bar chart, was first implemented. While the graph type allowed for more information to be shown, increasing engagement, and the graph could be filtered by continent, it was limited by the axis showing each country name and the hover function being limited to changing the graph colour.

3) *Third Implementation:* The third implementation tested out using a small multiples layout. Each graph represented countries from one continent. While it was interesting, the overview of all countries was more useful to have as it allowed for a global view, which could then be filtered by continent.

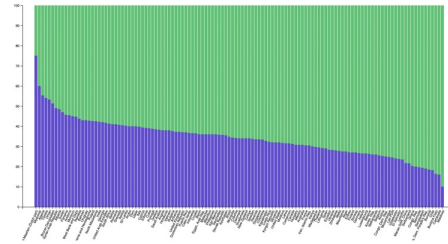


Fig. 11. STEM Graduates Visualisation - First Implementation.

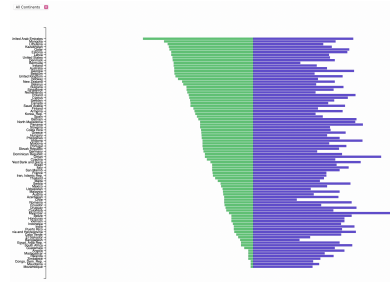


Fig. 12. STEM Graduates Visualisation - Second Implementation.

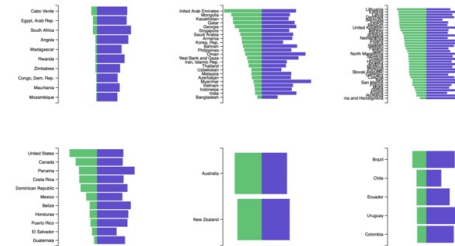


Fig. 13. STEM Graduates Visualisation - Third Implementation.

4) *Fourth Implementation:* The fourth implementation switched the dataset that was put into descending order, added an x-axis, and changed the colour of the bars to reflect the population size of each country. Ultimately, the original layout of the left dataset being in descending order was kept, as it told more of a story about the data - there is no clear increase in the proportion of female STEM graduates when there is a higher percentage of the female population with bachelor's degrees. It also had hover functionality, but this simply turned the active bar green.

5) *Final Implementation:* The final implementation of the STEM graduates visualisation is similar to the fourth, but with key improvements. The descending side of the graph has been switched back, the axis has been split, and the hover functionality fades out the other bars while making the active one darker. Additionally, the bars can be clicked to hold the information shown to the side for longer. Axis labels have been added, and the scale of the bars to the axis has been fixed so it is accurate to the data.

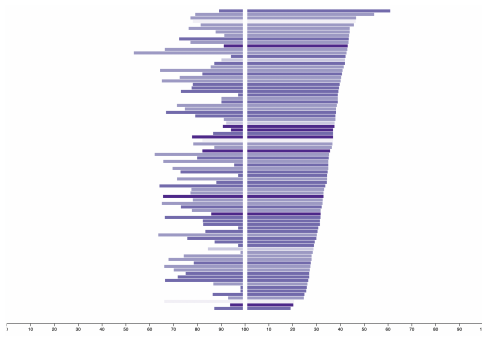


Fig. 14. STEM Graduates Visualisation - Fourth Implementation.

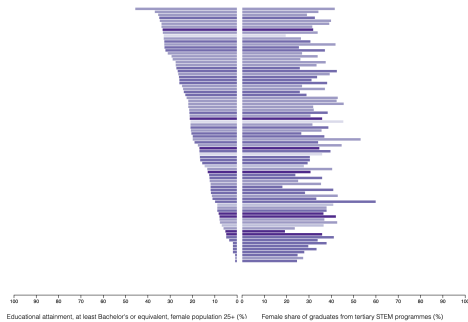


Fig. 15. STEM Graduates Visualisation - Final Implementation.

B. The Workforce

The second visualisation is the Workforce graph. Its final implementation is a beeswarm graph. The datasets used for this visualisation are 'Employers, female (% of female employment)' and 'Employers, male (% of male employment)'. It shows the ratio between male and female employers for each country. This chart was also tested as a bar chart and line graph before the final beeswarm was implemented.

1) *First Implementation:* The first implementation of this visualisation was a bar chart. It used two extra datasets to compare not only male and female employers, but male and female employees. It was very difficult to read as there was so much data that one country could not be distinguished from another. Not only this, but it meant that two out of the three visualisations would be bar charts, which would not be engaging.

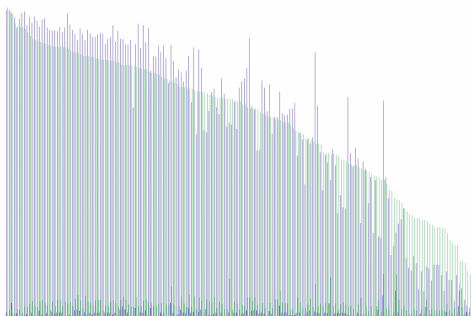


Fig. 16. The Workforce Visualisation - First Implementation.

2) *Second Implementation:* While trying to find a new way to display the data, one of the different types of graph tried out for this visualisation was a line graph. This aimed to show the difference between female employees and employers, but because there were so many countries to represent, the line graph became overcrowded like the previous implementation. This meant that it was nearly impossible to tell country lines apart, so this graph was not fit for purpose.

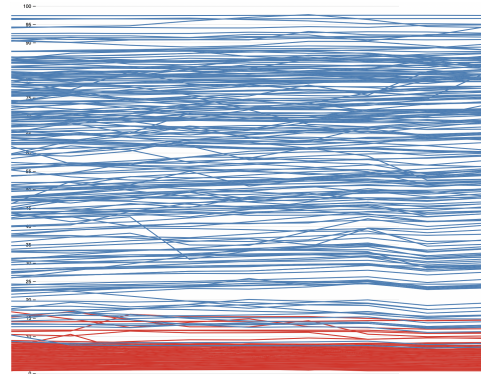


Fig. 17. The Workforce Visualisation - Second Implementation.

3) *Third Implementation:* The third implementation introduced the beeswarm graph as an alternative to the previous styles. The solution to the problem of having too much data was to remove the employee datasets, as they were not required to get across the idea of a workplace power imbalance. Additionally, the remaining two datasets were combined to make one value to display by taking the ratio between the percentage of male and female employers. The beeswarm graph helped to visualise this by showing clusters of data clearly. The colours of the graph represent the ratio for each country, with purple representing a higher proportion of female than male employers, white representing an equal proportion, and green a higher proportion of male to female employers. As the ratio gets further away from the midpoint of 1, the colours get darker.

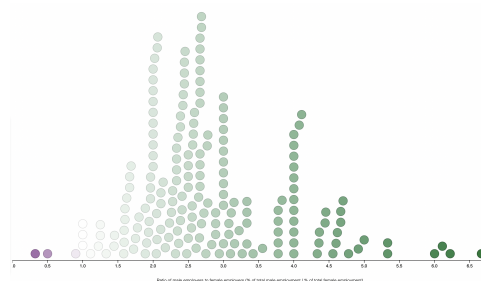


Fig. 18. The Workforce Visualisation - Third Implementation.

4) *Final Implementation:* The final implementation of the Workforce graph kept the beeswarm style, but added a time slider filter so that users can see how the ratios and overall trend change over time, as well as the average over time shown initially.

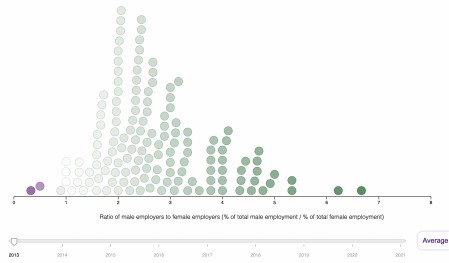


Fig. 19. The Workforce Visualisation - Final Implementation.

C. Education

The third visualisation is the Education graph, which in its final implementation takes the form of a connected dot plot. The datasets used for this visualisation are 'Educational attainment, at least completed upper secondary, population 25+, female (%) (cumulative)', 'Educational attainment, at least completed upper secondary, population 25+, male (%) (cumulative)', and 'Government expenditure per student, secondary (% of GDP per capita)'. It shows the difference between education outcomes for female and male secondary school students compared to the amount of GDP per capita spent by the government on those students.

1) *First Implementation:* The first implementation of this visualisation was a simple dot plot, with colours representing the continent to which the country belonged. Not only was this difficult to distinguish from the background, it did not convey any useful narrative to users.

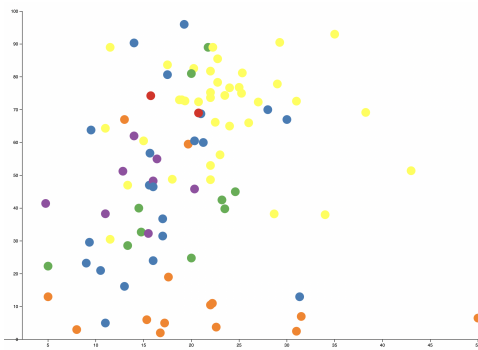


Fig. 20. Education Visualisation - First Implementation.

2) *Second Implementation:* The second implementation of the Education visualisation used the size of the dots to convey population size, and used higher-contrast colours to represent the continents. A world average dot was also added to show which countries were above or below the world average for each dataset. However, this still did not provide any context about female outcomes compared to male outcomes. As well as this, using dot size to convey population size was deceiving, as the difference between the largest and smallest countries is so great that if an accurate size scale was followed, the largest country would be the size of the entire page and the smallest countries would be unable to be seen on the graph at all. As

well as these issues, this implementation used a time slider to show changes in the data over the last ten years. This ran into problems as there were not enough complete datasets for all the countries to be displayed for each year. This was a common problem across all three visualisations and ended up requiring extra processing of the data to get an average value for each country across the past ten years. Where there was not enough data to do this, that country was discarded. Had the datasets been more complete, it would have been ideal to have time sliders for all three visualisations to show the changes in the data over time.

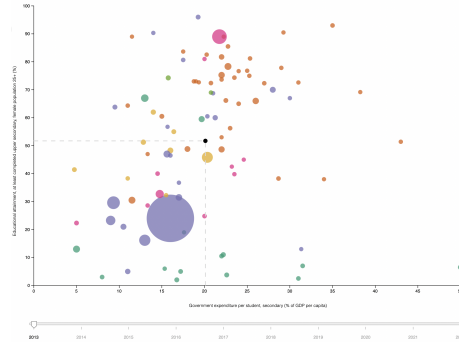


Fig. 21. Education Visualisation - Third Implementation.

3) *Final Implementation:* The final implementation of this visualisation is a connected dot plot. This allows the graph to show both male and female data and compare it to the government expenditure. It also allows users to easily see gender disparity in individual countries by connecting the male and female points for each country with a line. A filter option has been added that allows users to filter the data to only show countries where more women graduate secondary school than men and vice versa.

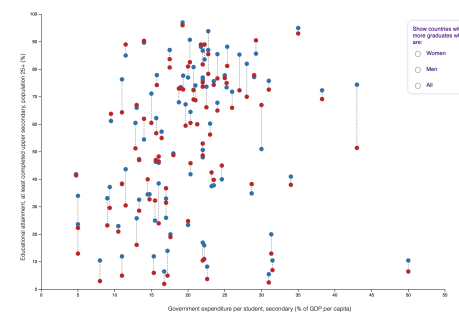


Fig. 22. Education Visualisation - Final Implementation.

D. Filtering

The filtering functionality was also subject to multiple implementations. It is separate from the graph section of the web application and parts of it affect each graph differently. The first implementation was a simple drop-down menu (See Fig. 23) for filtering each graph by continent. However, this did not follow the design, so it was changed to a radio button interface (See Fig. 24) as laid out in the wireframe. As more

functionality was added, such as filtering by country, the radio buttons took up too much space on the page, so the design was upgraded to two styled drop-down menus (See Fig. 25). While they did not match the wireframe exactly, great care was taken to ensure that they fit in with the style shown in the wireframe.

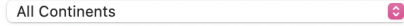


Fig. 23. Filtering - First Implementation.

Women in the Global Community

- all
- Asia
- Europe
- Africa
- North America
- South America
- Oceania
-

Fig. 24. Filtering - Second Implementation.

Women in the Global Community

Select Continent ▼

Select Country ▼

Fig. 25. Filtering - Final Implementation.

E. Details-on-Demand

Details-on-demand requires that users are able to access more information about a graph where they need it. Based on the initial wireframe, this started out as a single box (See Fig. 26) on the left hand side of the graph, underneath the filter. However, as the information was split into two types, information about the graph, and information about additional resources, it made sense to split them into two separate boxes (See Fig. 26) to help users understand what details were available. The final implementation maintains this split, but adds specific values for the country chosen by the user hovering over the graph, as well as example resources, and a fully functional 'Add Resource' Button.

The 'Add Resource' button allows users to submit their own organisations or initiatives to be considered for display in the Resources section. They must fill out all fields in the Add Resource modal that pops up (See Fig. 28), and once they submit the new resource, the form is processed by a

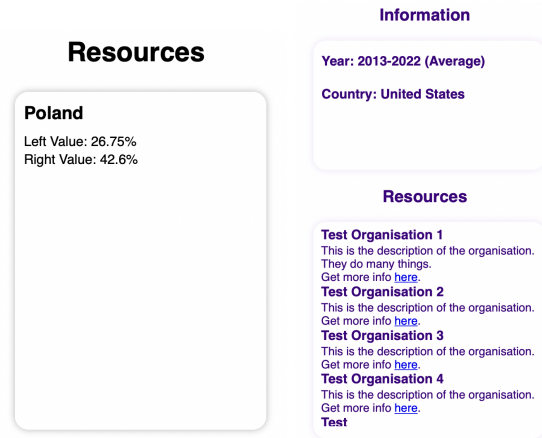


Fig. 26. Details-on-demand - First and Second Implementations.

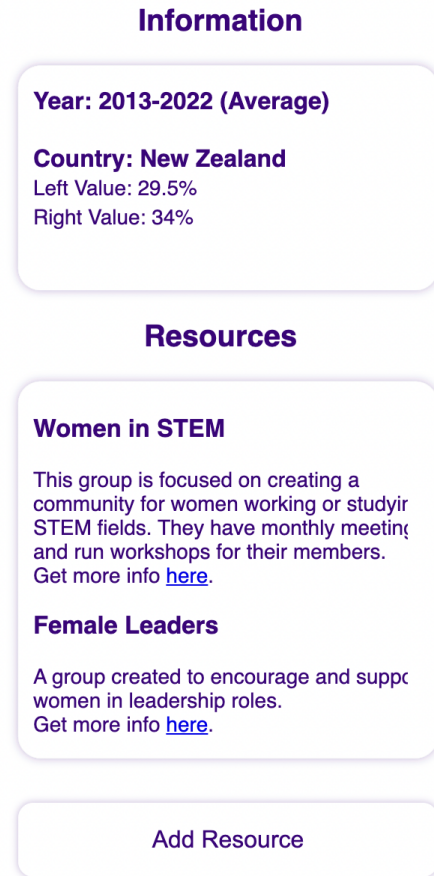


Fig. 27. Details-on-demand - Final Implementation.

PHP backend. This backend encodes the input text as HTML entities to prevent XSS by making sure any input is only stored as text. The new resource is then saved to an intermediary file, which is used to display suggested resources for moderation by an admin user on the moderation page (See Fig. 29). The user inspects the suggested resources and can approve or deny them for display on the main page.

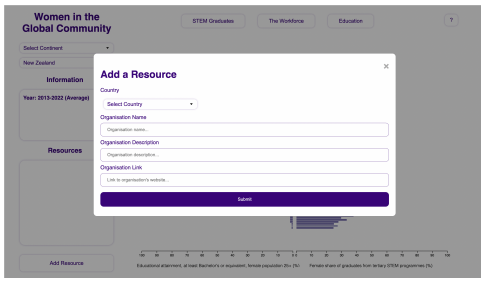


Fig. 28. Details-on-demand - Adding a new resource input form.

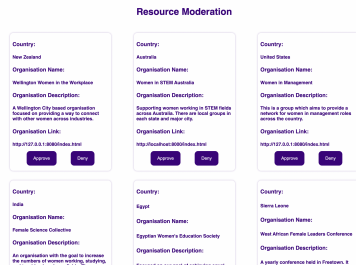


Fig. 29. Details-on-demand - Moderating suggested resources.

F. Meeting the Requirements

All requirements were met in the implementation, as seen in Table II.

TABLE II
REQUIREMENTS AND HOW THEY WERE MET

Requirement	How it was met
Easy to use	Responsive graphs and buttons
High-contrast	Dark text and graphs, light background, drop shadows on boxes and buttons, used colour-blind friendly colour schemes
Engaging	Interactivity (hovering over to learn more, filtering), multiple different graph types
Shows global information	Used datasets with information about countries around the world
Allows filtering based on region	'Select Continent' filter
Provides access to resources about initiatives that are local to each country	'Resources' box
Allows users to submit their own initiatives to the system	'Add Resource' button, input form, moderation page

VI. EVALUATION

A. Performance Metrics

The main performance metrics for this web application were usability and ability to convey information to increase users' awareness of gender disparity. To measure this, a user study was undertaken. The goal of the user study was to identify major usability issues in the system, as well as how well users were able to understand the information being conveyed by the visualisations. It was approved by the Victoria University of Wellington Human Ethics Committee with approval #0000031207. The study involved participants using the web application and answering a two-part de-identified questionnaire.

1) *Participants*: The participants in the study were a mix of men and women, with eight women and two men participating. Their ages ranged from 21 to 23, and were found through personal connections. The ethical issues around this and how they were dealt with are detailed in the Human Ethics Application. The participants were provided with koha (\$20 supermarket vouchers) for participating.

2) *Process*: The first part of the questionnaire aimed to get an understanding of how well the visualisations conveyed information. Participants were given fifteen minutes to find the answers to six questions while interacting with the visualisations for the first time. They were given a brief description of what the tool is and is able to do, but no instructions about how to use it. There were two questions to answer per visualisation. The questions were as follows:

- 1) STEM Graduates
 - a) Do Australia or New Zealand have a higher percentage of female graduates in STEM fields?
 - b) Which country has the greatest difference between the percentage of female graduates in STEM fields and the percentage of women who have a bachelor's degree?
- 2) The Workforce
 - a) Which country in South America has the greatest difference between male and female employers?
 - b) Do any countries have a higher percentage of female employers than male employers? If so, which countries?
- 3) Education
 - a) Between which two countries is there the biggest difference in government expenditure per secondary student?
 - b) What is a conclusion you could draw from this graph?

The second part aimed to understand how usable participants found the web application having used it for fifteen minutes in the first part. They were given ten minutes to rate how much they agreed with ten different statements about the web application. The scale used was a Likert scale:

- 1) Strongly disagree
- 2) Disagree
- 3) Neither agree nor disagree
- 4) Agree
- 5) Strongly agree

The statements rated by participants on the Likert scale were as follows:

- 1) The user interface was intuitive
- 2) The application was easy to use
- 3) I felt confident using the application
- 4) I needed to learn a lot of things before I was able to use the application properly
- 5) The visualisations increased my awareness about gender disparity
- 6) I found the visualisations difficult to understand
- 7) I found it difficult to navigate the application

- 8) I found that the filters helped me to understand the information
- 9) I found that the links to additional resources were relevant
- 10) The visualisations were engaging

B. Part One Results

The first two questions, 'Do Australia or New Zealand have a higher percentage of female graduates in STEM fields?' and 'Which country has the greatest difference between the percentage of female graduates in STEM fields and the percentage of women who have a bachelor's degree?' had mostly correct answers (See Fig. 30, 31). There was one incorrect answer for the second question, and no partially correct or unsure answers. This shows that the STEM Graduates visualisation conveys information clearly, as participants were able to easily find information by interacting with the graph. However, these questions do not test participants' use of filters to discover information and are designed to be simple initial questions to get users familiar with the interface.

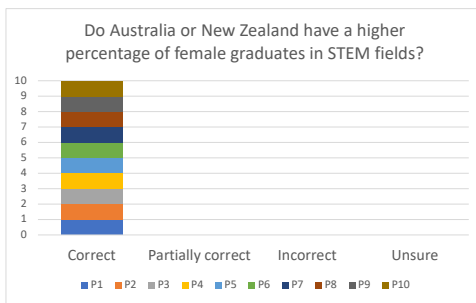


Fig. 30. STEM Graduates question a) results.

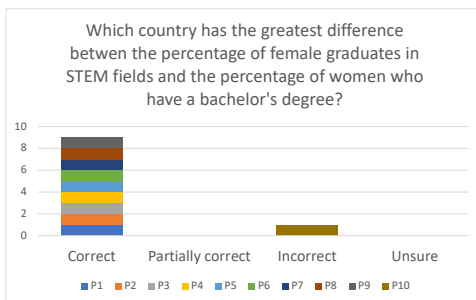


Fig. 31. STEM Graduates question b) results.

The second set of questions, 'Which country in South America has the greatest difference between male and female employers?' and 'Do any countries have a higher percentage of female employers than male employers? If so, which countries?' had a more varied range of responses. The first question received correct responses from all ten participants (See Fig. 32), but the second question only received one fully correct answer (See Fig. 33). The majority of answers for the second question were incorrect, however some were unsure and did not attempt an answer, while some attempted an answer which was only partially correct. The first question simply

required participants to filter the visualisation by continent and find the right-most point, but the second question required an understanding of ratios to answer correctly. The partially correct answers added countries with a ratio of 1, an even split of male and female employers, to their answers. While this could be indicative of a misunderstanding of ratios, it shows that more context may be needed for this graph regardless.

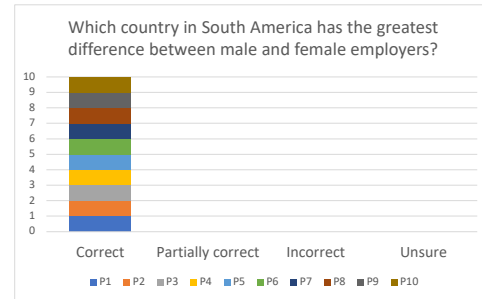


Fig. 32. The Workforce question a) results.

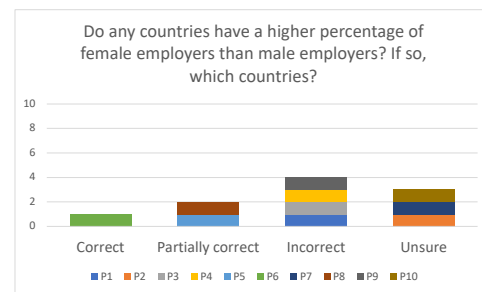


Fig. 33. The Workforce question b) results.

The third set of questions were 'Between which two countries is there the biggest difference in government expenditure per secondary student?' and 'What is a conclusion you could draw from this graph?'. The first question, like the others, had a correct answer, and the majority of participants found it (See Fig. 34), with only two partially correct answers and one incorrect answer recorded. However, the final question did not have a correct answer. It was more open-ended and was designed to see what participants would conclude overall about the visualisation. The answer was expected to be something similar to 'women are less likely to have better education outcomes than men'. As seen in Table III, the actual results recorded in the study varied. Some focused on specific countries or continents, while others focused on the government expenditure. Interestingly, only three participants noted a difference between male and female data, and of these three, only one (P6) correctly identified that the graph shows that women are at an educational disadvantage. While many of the other answers were correct based on the data shown, they missed the overall idea that the graph was showing gender disparity in education. As this was the final question, they may have been able to answer it in a more informed way had they been given more time to understand the graph.

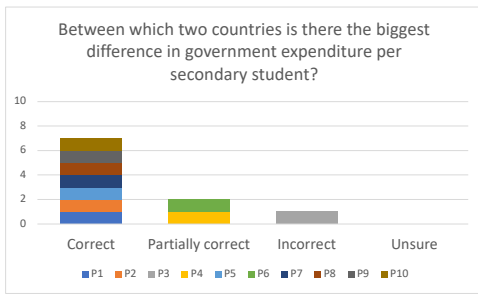


Fig. 34. Education question a) results.

TABLE III
EDUCATION QUESTION B) RESULTS

Participant	Answer
P1	Higher expenditure per student does not necessarily lead to higher educational attainment
P2	Usually the higher the govt. expenditure the higher the educational attainment
P3	Government expenditure is higher for women than men in countries where more graduates are women and vice versa (with the exception of Israel)
P4	Not sure what conclusion to make. Don't know what the graph was showing.
P5	The continent of Africa has a very wide spread of government expenditure per student across its countries, ranging from 5% to 50%
P6	There are fewer countries with more graduates who are women & these countries sit in the average range of government expenditure and educational attainment.
P7	Don't know sorry
P8	The relationship between government expenditure and proportion of male-female graduates. It looks like in the countries where govt expenditure is higher there are more male than female grads.
P9	Most countries spend between 15 and 30% of their GDP on students
P10	Kazakhstan has the highest level of educational attainment of the countries sampled in this research.

C. Part Two Results

The second part of the user study was the usability ratings. From these results we can see that most participants found the interface intuitive and easy to use (See Fig. 35), as well as engaging (See Fig. 36). They also found the filters helpful and were able to navigate the web application easily (See Fig. 36). However, they felt that they were neither confident nor diffident users of the application, and most of them felt as though they needed to learn some things before they were able to use the application properly (See Fig. 35). There were mixed responses about whether their awareness of gender disparity was increased (See Fig 35), as well as whether the links to additional resources were relevant, with one participant noting that they did not find them. Some users also found the visualisations difficult to understand. What these results show overall is that while the web application interface was usable and engaging, the information itself was difficult to understand for some users.

D. Changes Made

I was able to use the results of the user study to add a 'help' button in the top right-hand corner of the interface,

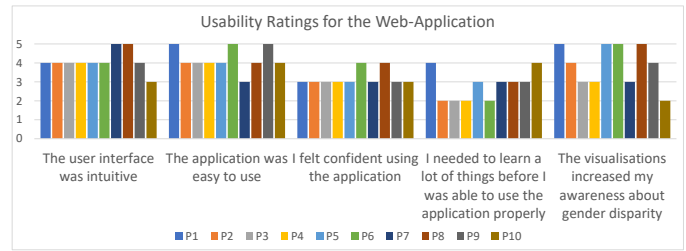


Fig. 35. First five usability rating results.

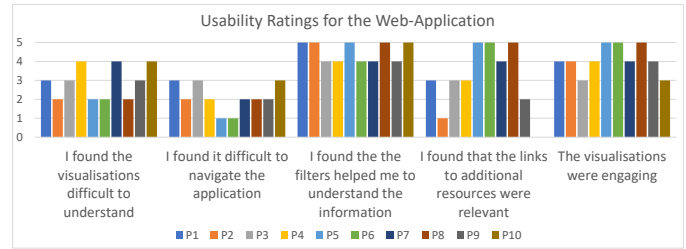


Fig. 36. Second five usability rating results.

indicated with a '?' character. By clicking the button, users trigger a modal (See Fig. 37) that contains instructions for how to use the web application, as well as brief explanations of the graphs and datasets for those unfamiliar with these types of visualisations.

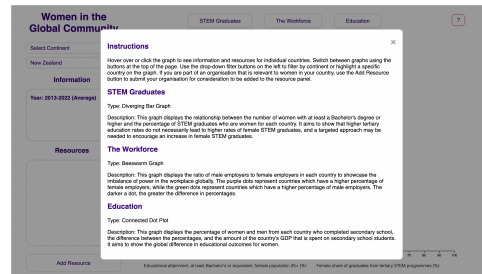


Fig. 37. 'Help' button pop-up.

E. Limitations

The first limitation of the final implementation of the web application is that it is not hosted on a live server. Currently, it has to be run on a local development server and accessed through localhost. This means that potential users would have to clone the GitLab repository, install PHP, and run it on their own device. If the web application were to be developed further, this would be a key feature to implement as it would allow the visualisations to be accessed by anyone and would greatly increase the number of people able to improve their awareness of gender disparity by using the tool.

Another limitation of the current implementation of the web application is the lack of access control for the moderation page. While efforts have been made to not make it accessible directly from the main page, it is just a separate HTML file which can be accessed by directly typing the address into the search bar. If the site was hosted on a live server, there would need to be an authentication and authorisation process

put in place to prevent unauthorised access to the page. This would involve implementing two types of users - admin and regular - which would also help in tracking who is uploading resources for admin consideration. Currently, there is no way to track who uploads a resource and if a malicious or unwanted resource is uploaded, it would be useful to be able to identify the user who uploaded it.

It is also important to acknowledge the limitations of the datasets. The data obtained from the World Bank does not contain data about people who do not identify as male or female and therefore the visualisations are limited by this. If the World Bank updates its Gender Data Portal to include other genders in the future, this should be a key feature to implement in future development. The data is also missing many values, so while an average of any available values for each country is taken to create the visualisations, it must be acknowledged that they (besides the Workforce visualisation) only show average values over the past ten years. A more complete picture of the state of gender disparity might be achieved if the gaps in the data were filled and the timeframe of the data went back further than just ten years.

Additionally, a limitation of the user study undertaken to evaluate the web application was the demographic of the participants. While both male and female participants were recruited, they were all between the ages of 21 and 23, and had either just completed, or were completing very soon, a Bachelor's degree at Te Herenga Waka - Victoria University of Wellington. To get a more comprehensive understanding of how well the web application conveys information and its usability, a future study with a more diverse set of participants could be useful.

VII. CONCLUSION

This project aimed to provide a starting point for people wanting to increase their awareness of gender disparity in education, workplace leadership, and STEM. The main usability requirements identified were that it is easy to use and engaging, while the functionality requirements were that it shows global information, allows filtering based on region, provides access to resources about initiatives that are local to each country, and allows users to submit their own initiatives to the system. The final implementation meets the functionality requirements and adds additional functionality beyond what was required, and the user study evaluated the usability requirements and found that the web application was both usable and engaging. Some small adjustments to users' ability to understand the data were identified in the study and have been made. Overall, this project produced the design, implementation, and evaluation of a web application that displays interactive visualisations of data about women over the past decade.

ACKNOWLEDGMENTS

I am very grateful to Dr Stuart Marshall for supervising this project and for the support and guidance he provided throughout it. I would also like to thank the user study participants for giving up their time to participate in the study.

REFERENCES

- [1] A. Krivkovich, W. W. Liu, H. Nguyen, I. Rambachan, N. Robinson, M. Williams and L. Yee, "Women in the Workplace 2022," McKinsey and Company, 18 October 2022. [Online]. Available: <https://www.mckinsey.com/featured-insights/diversity-and-inclusion/women-in-the-workplace#/>. [Accessed 30/03/2023].
- [2] M. E. Heilman, "Gender stereotypes and workplace bias," *Research in Organisational Behaviour*, vol. 32, pp. 113-135, 2012.
- [3] "THE 17 GOALS," United Nations, 2023. <https://sdgs.un.org/goals>. [Accessed 05/05/23].
- [4] B. Shneiderman, "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations," in 1996 IEEE Symposium on Visual Languages, Boulder, CO, 1996.
- [5] D. Patel, "Understanding Gender Equality Through Visualisations," Victoria University of Wellington School of Engineering and Computer Science, Wellington, 2017.
- [6] World Economic Forum and TWO-N, "Global Gender Gap Report Browser," World Economic Forum, TWO-N, 2016. [Online]. Available: <https://projects.two-n.com/world-gender/>. [Accessed 21/04/23].
- [7] P. Overberg, J. Adamy, L. Thuy Vo, J. Ma, A. Van Dam and S. A. Thompson., "What's Your Pay Gap?," Dow Jones and Company, 17 May 2016. [Online]. Available: <http://graphics.wsj.com/gender-pay-gap/>. [Accessed 21/04/23].
- [8] Gapminder, "Gapminder," Gapminder, [Online]. Available: <https://www.gapminder.org/>. [Accessed 29/05/23].
- [9] OECD, "OECD Gender Data Portal", [Online]. Available: <https://www.oecd.org/gender/data/>. [Accessed 05/11/23].
- [10] The World Bank Group, "DataBank — Gender Statistics," 2023. [Online]. Available: <https://databank.worldbank.org/source/gender-statistics#>. [Accessed 21/04/23].
- [11] M. Bostock, "Data-Driven Documents," 2021. [Online]. Available: <https://d3js.org/>. [Accessed 17/03/23].
- [12] National Park Service, "Symbols of the Women's Suffrage Movement," U.S. Department of the Interior, 6 April 2020. [Online]. Available: <https://www.nps.gov/articles/symbols-of-the-women-s-suffrage-movement.htm>. [Accessed 05/05/23].
- [13] International Women's Day, "About International Women's Day," 2023. [Online]. Available: <https://www.internationalwomensday.com/about>. [Accessed 05/05/23].
- [14] npm, "D3: Data-Driven Documents", 2023. [Online]. Available: <https://www.npmjs.com/package/d3>. [Accessed 06/11/23].