What influences the MBTI personality?

1st Author Name
Yafei Wang
Nottingham, UK
alyyw86@nottingham.ac.uk

ABSTRACT
The aim of the project is to investigate the factors that contribute to the different MBTI personalities of each individual, both in terms of innate and acquired factors. In the age of the Internet, a huge amount of data is generated every day. It is vital to visualise this huge amount of data. The project produced information visualisation charts through D3.js and presented them in a web format. The 32 personality profiles based on different countries are presented in bar charts and the personality profiles based on the month of birth are presented in scatter charts.

Author Keywords
Information Visualisation, Web, D3.js, jQuery, JavaScript, HTML5, CSS, CSV

ACM Classification Keywords
HCI

INTRODUCTION
MBTI stands for Myers Briggs Type Indicator. It classifies people into sixteen personalities along four dimensions, namely attention direction, cognitive style, judgment style, and lifestyle. This test is now also widely used in job hunting. It is a tool often used to help individuals understand their communication preferences and how they interact with others. The MBTI assessment model is used in 115 countries, is available in 29 languages and has been used by 88 of the Fortune 100 in the last five years [1]. One of the most intriguing parts of people's ongoing exploration of themselves is their personality. In early 2022, a Korean reality show brought together people with different MBTIs to observe their reactions in everyday life and to explore the differences in the way people with different MBTIs treat each other. People can be divided into 16 personalities, so what are the reasons for the individual differences? Two conjectures are proposed for this question: country of origin and month of birth, which correspond to acquired and innate factors respectively. To further validate the exploration of these two conjectures, the following measures were taken in the project: analysis of MBTI categories, worldwide distribution of personality types, and distribution of personality types by month of birth.

This project visualises information through the use of D3.js. By using D3.js, data visualisation can be achieved very well and using intuitive vectors instead of traditional forms. The visual presentation is achieved through data loading, data binding, parsing of transformed elements and redundant elements.

RELATED WORK
Qi Yang and others collected a large number of social media statements from around the world to analyse their MBTI personality This means that MBTI has a wider and more popular audience [2].

FanBao and Jia Chen use a web visual presentation tool - D3.js - to implement statistical graphing, presenting interactive visual graphs through the book category dimension, the time dimension and other dimensions [3].

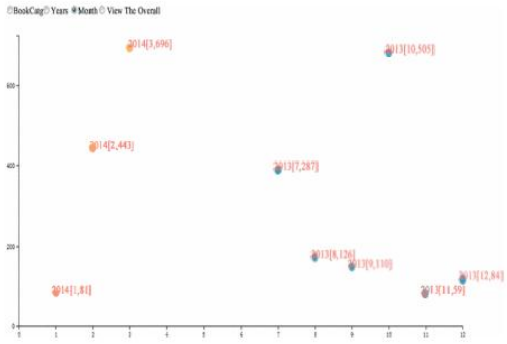

Figure 1. Scatter chart in Fan Bao and Jia Chen work1(https://ieeexplore-ieee-org.nottingham.idm.oclc.org/abstract/document/6845553/authors#authors)

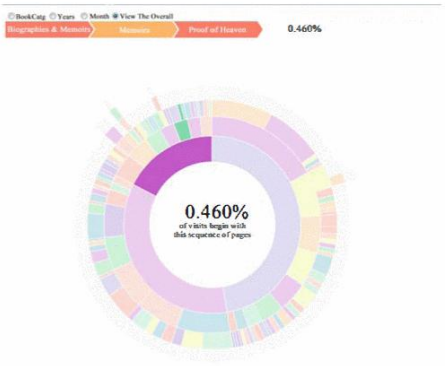

Figure 2. Pie chart in Fan Bao and Jia Chen work1(https://ieeexplore-ieee-org.nottingham.idm.oclc.org/abstract/document/6845553/authors#authors)

DESIGN

The project is to present the results of data visualisation in the form of a web page. Firstly, at the beginning of the project, the basics of web production need to be studied; JavaScript, HTML5, and CSS3 are the essential front-end development languages for web development and are currently the most mainstream development languages. Secondly, for data visualization, there is a library in JavaScript that can manipulate SVG: D3.js. The library can be customised to draw visual charts such as bar charts, line charts, pie charts, scatter charts, force maps, etc. In addition, the relevant data for the problem needs to be found after the technology and tool library have been identified. As D3.js has the ability to read CSV data directly, there is no need for a database or back-end environment. For this project, it is sufficient to process the data via JavaScript. Finally, it is the selection of development tools and the processing of cross-domain data requests. VSCode is a lightweight code editor that supports multiple programming languages by default, including JavaScript, HTML5, and CSS3 required by this project. In this project, programming is made more convenient by downloading extensions. The web development environment needs to be supported by a server to read the data. VSCode has a Live Server plugin that enables the launch of a development local server with life reload functionality for both static and dynamic pages. jQuery was also used in the project to assist in transferring the data read in JS to the components of the HTML page.

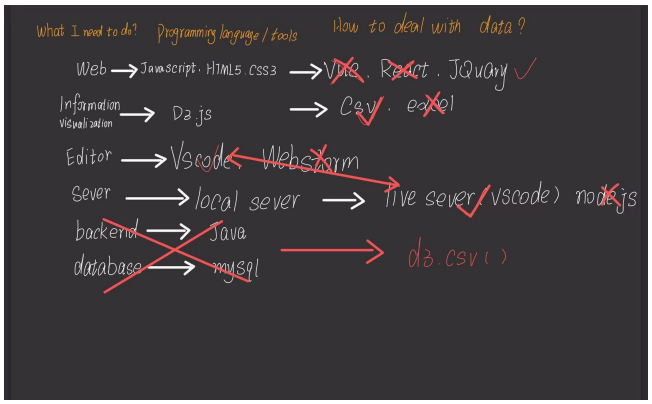

Figure 3. The Choice of tools and programming

There are three main aspects to the design of the visualisation in this project: 1. Introduction to the MBTI personality 2. Distribution of personality in different countries 3. Distribution of personality based on the month of birth. For the introduction of MBTI personalities, the text area tool can be used for the presentation of a large amount of text. It is worth noting that each of the sixteen personalities needs to have a presentation page. For the distribution of personalities in different countries, the bar chart can effectively show the percentage of personalities in each country. As each of the 16 personality types of the MBTI is subdivided into two A and T types, 32 bar charts are drawn up for the chart. To improve the user experience, designing buttons that can be indexed is an effective way to do this. For the distribution of personalities based on the month of birth, dynamic scatter plots add interactivity while visualising the data.

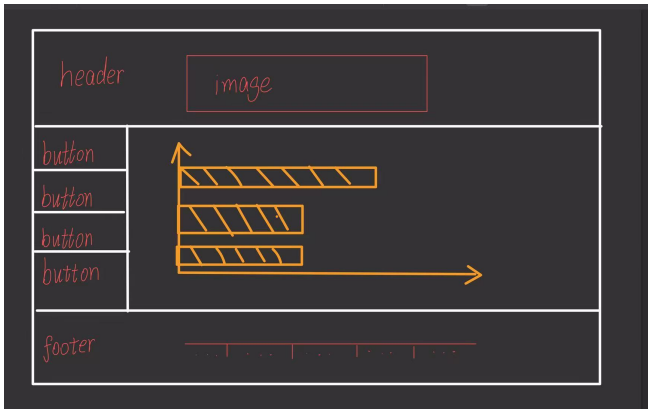

Figure 4. Web Frame Design

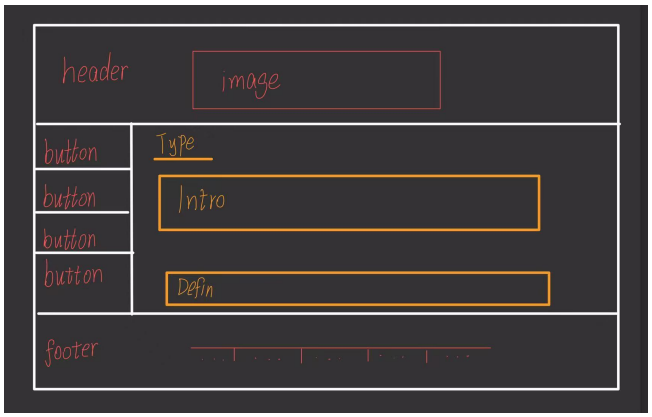

Figure 5. Introduction Part

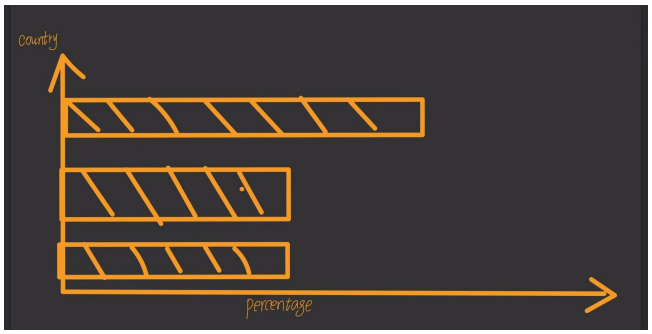

Figure 6. Bar Chart

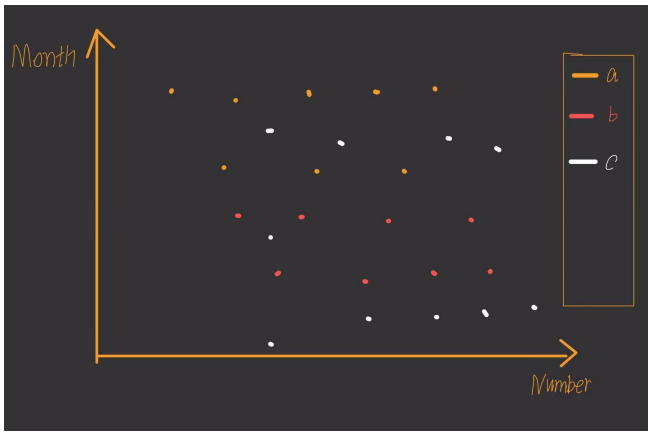

Figure 7 Scatter Chart

IMPLEMENTATION

The implementation of this project is mainly divided into four parts: web framework construction, MBTI personality introduction, country-based MBTI personality distribution, and MBTI personality distribution based on birthday month. The tools used in this project are VSCode Editor, Live Sever, D3.js, and JQuery. The data files are countries.csv, MBTI types and Birthdays (Responses), and types.csv.

Web framework construction

There are four main folders in this project: data (for csv data files), images (for background images), scripts (for js files) style (for csv files).

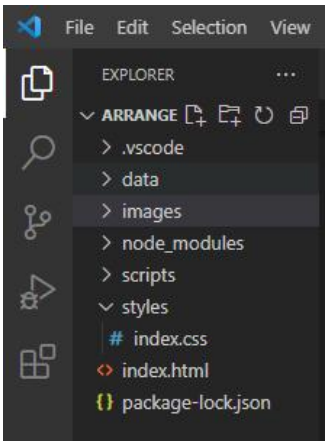

Figure 8. Web framework construction

The layout structure of the page is a three-part structure divided into header, main, and footer. The main part is divided into two columns, the left side is the navigation bar, and the right side shows the main content. The content section contains: a text area, a switchable bar chart, and a dynamic scatter chart.

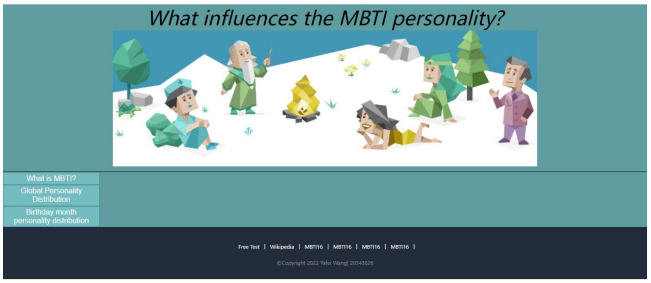

Figure 9.Web Layout

MBTI personality introduction

This section is to read the data of types.csv using d3.csv () and transfer the acquired data to the component corresponding to the id in the html file using jQuery. There are four main sections: "TypeNickname", "Introduction", "Definition", "Celebrities". As there are more text sections, a textarea component is used to display the information. The specific html code is as follows.

```
<div id="detail">
 <h1 id="TypeNickname">refer</h1>
    <p align="center" >
    <textarea     id="Introduction"     rows="50"
      cols="150">refer</textarea>
     </p>
    <h4>Definition</h4>
   <textarea       id="Definition"       rows="3"
      cols="150">refer</textarea>
   <h4>Celebrities</h4>
```

```
<textarea id="Celebrities" rows="3"
    cols="150">refer</textarea>
</div>
```

As there are 16 personalities in total, buttons are set up in the navigation bar to switch data.

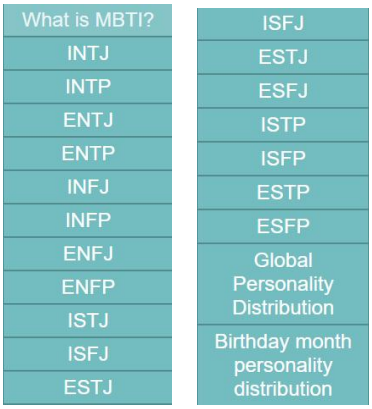

Figure 10. Buttons

Country-based MBTI personality distribution

This section takes a bar chart for visualisation, drawing the SVG canvas in an html file. The bar chart is drawn through the ESTJA.js file, using a linear scale for the x-axis and a strip scale for the y-axis. x-axis represents the proportion of the population of the corresponding personality and y-axis represents the country. Then rectangles are drawn, with the height of the rectangle corresponding to the size of the x-value obtained. The data corresponding to the x- and y-axes is read from counts.csv using d3.csv (). It is worth noting that the data read is a string and needs to be converted from a string type to a float type using parseFloat().The code to draw the bar chart is as follows:

```
Html:
<svg width="800" height="2000" min-height="100px"
id="countryestjasvg"class="svgs"
style="background-color: #ffffff;"></svg>
js:
const xScale = d3.scaleLinear()
  .domain([0, d3.max(data, xValue)])
  .range([0, innerWidth]);
const yScale = d3.scaleBand()
  .domain(data.map(yValue))
  .range([0, innerHeight])
  .padding(0.1);
const g = svg.append('g')
      .attr('transform',`translate(${margin.left},
 ${margin.top})`);
  g.selectAll('rect')
  .data(data)
  .enter()
  .append('rect')
  .attr('y', datum => yScale(yValue(datum)))
                    .attr('width',(datum)=>{return
xScale(xValue(datum))}) // use xSacle to re-scale
data space (domain) and return the rescaled
population;
 .attr('height', yScale.bandwidth())
 .attr('fill', 'steelblue')
```

Because each personality is divided into two types, A and T, a total of 32 bar charts have been created. Thirty-two buttons were drawn on top of the bar charts to switch among the bar charts. To reduce time and space, the 32 bar charts are drawn in the same SVG component. The content of the SVG is destroyed before a new image is drawn when switching. The specific code is as follows.

```
d3.select('#countryestjasvg')
 .selectAll('*')
 .remove();
```

MBTI personality distribution based on birthday month

This section uses dynamic scatter plots to present personality distributions based on the month of birth. The SVG component is used in the HTML file. The dynamic scatter plot is drawn in the JS file, using a line scale for the x-axis and a point scale for the y-axis. The points are plotted through the corresponding positions on the x and y-axis and filled with the corresponding colours according to the y axis attribute. As the scatter plot adds the colour attribute, a legend is drawn to represent the corresponding colour. The legend is drawn similarly to a bar chart, with a rectangle drawn and text added next to it. In addition, a dynamic effect is added by adding a path for the points to move on the x-axis.

The code is as follows:

```
let legend_color = Object.values(color);
            let legend_text = Object.keys(color);
            // draw legend
            var              legend              =
d3.select('#maingroup').selectAll(".legend")
            .data(legend_text).join('g')
            .attr("class", "legend")
```

```
            .attr("transform", function(newdata, i)
{ return "translate(" + (innerWidth1 + 10) + "," +
(i * 25 + 5) + ")"; });


            // draw legend colored rectangles
            legend.append("rect")
            .data(legend_color)
            .attr("x", 0)
            .attr("y", 0)
            .attr("width", 30)
            .attr("height", 20)
            .style("fill", (newdata) => (newdata));
    legend.append("text")
            .attr("x", 40)
            .attr("y", 9)
            .attr("dy", ".5em")
            .attr("text-anchor", "start")
            .text((newdata) => (newdata));
```

With regard to data processing, firstly the data from MBTI types and Birthdays (Responses) was read through d3.csv(). Then, as it was desired to get the number of different personalities for each month, the number of people was calculated using an if-conditional statement and loop statement to get a new JSON array.

RESULTS

The header of a web page consists of text and an image, as shown below.

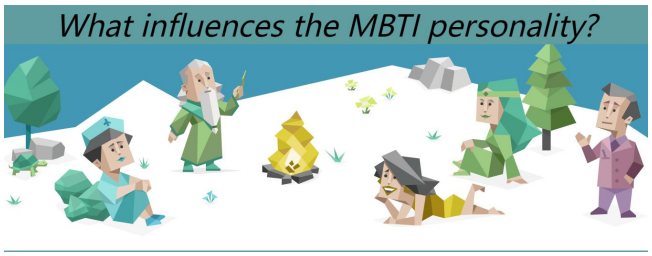

Figure 11. Header of Web

MBTI personality introduction

This section shows the names, nicknames, introductions, definitions and related celebrities for each of the 16 personalities. The navigation bar on the right can be clicked on to switch between personality profiles.

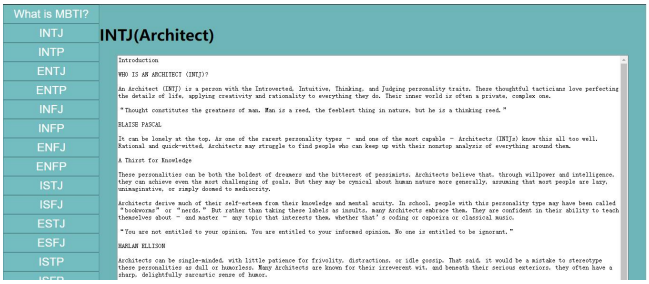

Figure 12. Introduction-1

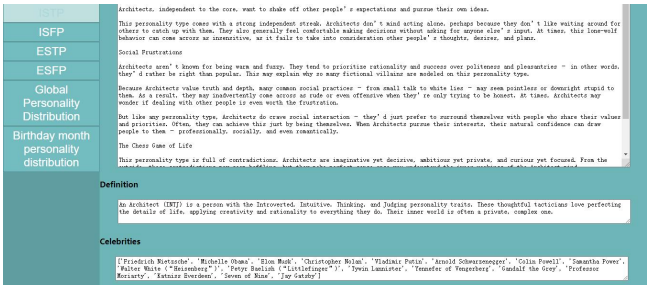

Figure 13. Introduction-2

Country-based MBTI personality distribution

This section presents a bar chart showing the distribution of the 32 personality types across the 158 countries in which each of the 32 personality types exists. There are a total of 32 charts in this section of the project, each of which is arranged in alphabetical order according to the country's initials. According to the personality percentages of 158 countries, geography has a strong influence on the formation of personality, and it is easy to note the disparities between countries. Among all personality types, no more than 15 percent of each country's personality is represented. For example, Afghanistan has the largest proportion of ESTJ-A and ESTJ-T personalities and Albania has the largest proportion of INTP-A and INTP-T personalities.

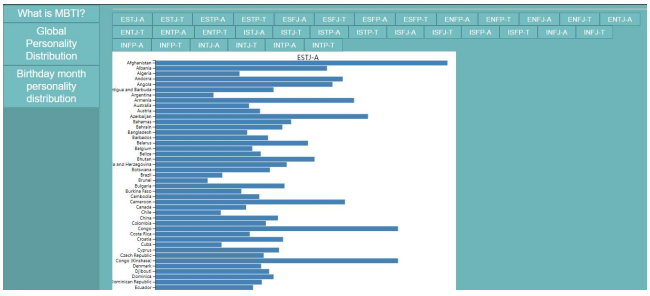

Figure 14. Distribution of ESTJ-A (part)

Figure 15. Distribution of ESTJ-T(part)

Figure 16. Distribution of INTP-A(part)

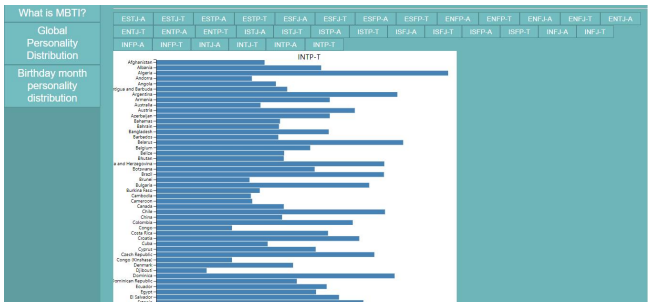

Figure 17. Distribution of INTP-T(part)

Afghanistan has the largest proportion of ESTP-A and ESTP-T personalities

Figure 18. Distribution of ESTP-A(part)

Figure 19. Distribution of ESTP-T(part)

Nicaragua has the largest proportion of ESFJ-A personalities. Rwanda has the largest proportion of ESTP-T personalities. Although the A and T sides of the same personality are present, the distribution varies from country to country and there are large differences.

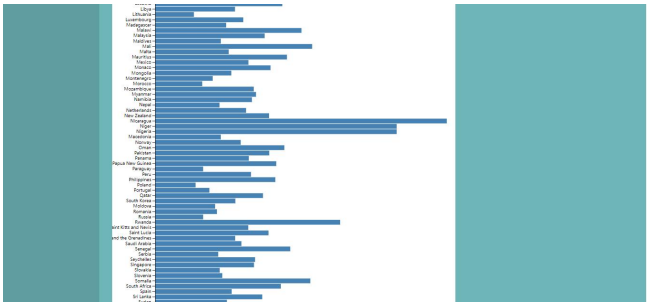

Figure 20. Distribution of ESFJ-A(part)

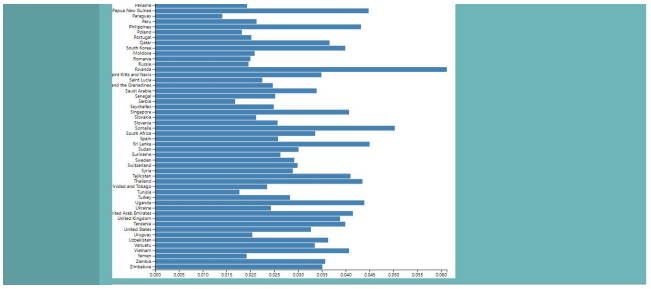

Figure 21. Distribution of ESFJ-T(part)

Due to the length of the images, all of the above images only capture the part containing the highest values. There are 32 distribution charts in the project, and two complete distribution charts are shown below. The chart below shows the distribution of ENFP-A and ENFP-T across 158 countries, with countries sorted by initials. It reflects the data gaps for each country, with Burkina Faso having the most ENFP-A personalities and Lesotho having the least ENFP-T personalities. The largest gap can be 6%, indicating that the country, i.e. the living environment, has a large influence on the formation of personality.

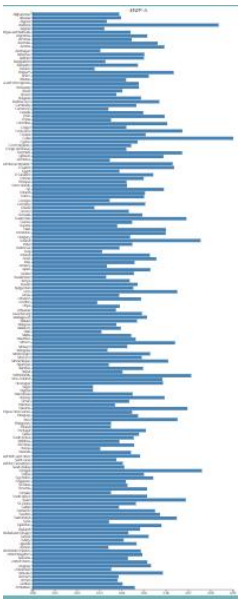

Figure 22.Distribution of ENFP-A(whole)

Figure 23. Distribution of ENFP-T(whole)

Figure 24. Distribution of ISTP-A(whole)

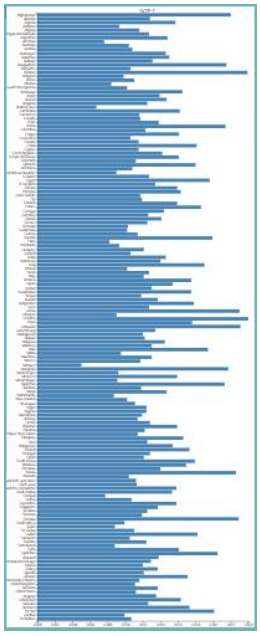

Figure 25. Distribution of ISTP-Twhole)

MBTI personality distribution based on birthday month

Based on the visualisation in this section, it can be seen that the month of birth has a smaller effect on personality, for example, INTJ is very low in all 12 months. Whereas ESFP occupies the most positions in all 16 personalities. And compared to the distribution on a national basis, not every month has 16 personalities, for example January has only seven personalities, while August, which contains the most personalities, contains ten personalities. It is worth noting that each personality has a similar share in different months,

for example, INTJ is concentrated in 2 to 7, ENTJ in 6 to 11 and ESFP in 20 to 24.

The following are process diagrams of the dynamic effect of the scatter plot:

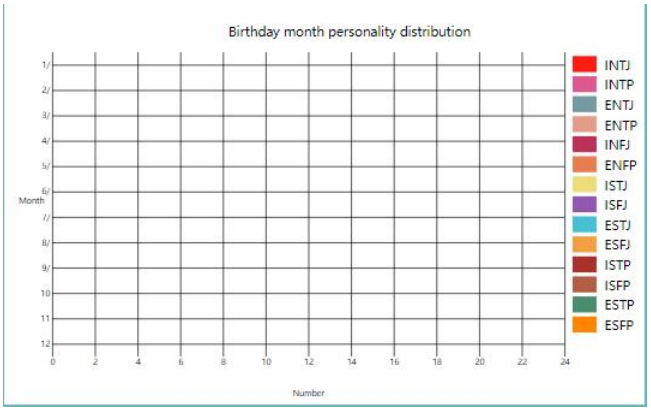

Figure 26. Scatter chart-1

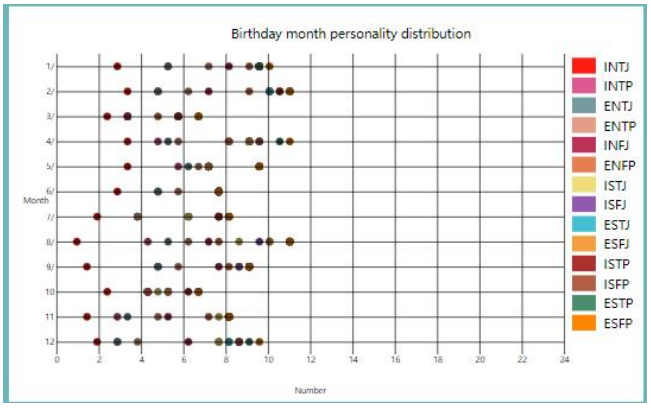

Figure 27. Scatter chart-2

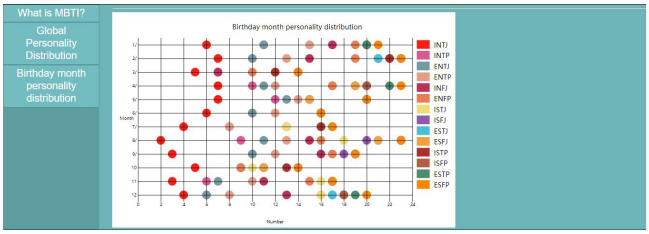

Figure 28. Scatter chart-3

EVALUATION

There are two main methods used to evaluate visualisation projects during the evaluation phase: the user testing method and the non-user testing method. The user testing method is the 'think aloud' method, where participants are asked to complete a series of tasks for the project and discuss with them what it is like to use the visualisation project. The non-user testing method is 'technical performance testing', i.e. I need to do some technical performance tests such as project start-up time, chart switching time, dynamic effect trigger time. These two methods were chosen because they were easy to implement and get feedback on.

The details of the user testing methodology are as follows: two participants were invited to use the visualisation project and no instructional actions were taken while the participants were experiencing the project. As the project was built on a local server, the participants had to use the developer's personal computer to complete the experience. There are some feedbacks from the user tests: 1. The text font in the personality introduction section is rather small and there is excessive text. 2. The bar chart is quite long, which is sufficient for seeing the general distribution, but it is not easy to find the specific countries. 3. The scatter chart is clear, but the dynamic effect is rather short and I hope it will be repeated. 4. The button toggle chart is easy, but the state after clicking is the same as before clicking.

As for the non-user tests, five separate sets of tests were conducted and the corresponding data was recorded for analysis. This test indicates that it has a good response performance, with a start-up time within 1.56 seconds, a switching chart average of 00.49 seconds and dynamic feedback within 1.50 seconds.

Test content	Test1/s	Test2/s	Test3/s	Test4/s	Test5/s
project start-up time	01.53	01.42	01.30	01.56	01.49
chart switching time	00.49	00.61	00.55	00.39	00.42
dynamic effect trigger time	01.35	01.49	01.40	01.43	01.42

Table 1. non-user test

DISCUSSIONS

The project firstly introduces the user to the concept of MBTI and at the end of the page, there is a free test site to support the exploration of personality. Secondly, the user can directly find the distribution of each personality by using a histogram of the personality distribution based on different countries. The distribution is used to further explore the influence of life circumstances on personality formation. Then, the personality distribution by month of birth is used to explore the influence of innate factors on personality formation. A visual analysis reveals that between living environment and innate factors, life environment has a significant effect on personality formation. The charts are interactive and the user is not tedious when looking at the data.

The new idea brought by this project is to explore the reasons for the different MBTI distributions, starting from both innate and acquired factors. And during the exploration, it was found that the differences in the number of people born in different months for each personality were smaller, with the most significant being INTJ.

However, there are some drawbacks to this system: firstly, the number of countries causes the y-axis of the bar chart to be too long and requires some patience if one wants to find specific countries. Besides, the amount of data in the scatter chart is small and further exploration of innate factors requires a larger amount of data to support it.

CONCLUSIONS AND FUTURE WORK

Through this project, I have learned the skills of creating web pages, which is like painting the whole picture before refining the details. In terms of data visualisation, the project has taught me the power of D3.js to manipulate scalable vector graphics (SVG). In terms of data collection, Kaggle is a great platform for data sets, with a wide variety of data available in CSV format.

As a result of this project, I have gained a better understanding of MBTI. Life is a continuous process of self-exploration and the causes of personality are always overlooked. The results of the visualisation show that life circumstances have a much greater impact on personality than congenital (month of birth) influences.

However, there is still a lot of potential for this project. First of all, it is worth thinking about whether there is a better chart to show the distribution of countries than a bar chart. A good attempt would be to draw a global map to show the distribution by the colours marked on the map. Also, finding more data to further explore the effect of the month of birth. Finally, there are many other factors that influence a person's personality and it would be interesting to visualise them all through information to further explore what influences the MBTI. In the age of the internet, with the amount of data being generated every day, it is vital to communicate redundant data in an easy-to-understand manner to users. Information visualisation has emerged as a solution to this problem, with various forms of presenting data, tools, charts, and graphs.

It is my hope that I will be able to do more interactive effects for my subsequent visualisation projects.

Words Count: 3082

Bibliography

1. Kerwin, Patrick L. 2013.True type tales: Real stories about the power of personality type in everyday life. Patrick L. Kerwin.

2. Qi Yang, Aleksandr Farseev, and Andrey Filchenkov. 2021. Two-Faced Humans on Twitter and Facebook: Harvesting Social Multimedia for Human Personality Profiling. In Proceedings of the 2021 Workshop on Intelligent Cross-Data Analysis and Retrieval (ICDAR '21). Association for Computing Machinery, New York, NY, USA, 39–47. https://doi-org.nottingham.idm.oclc.org/10.1145/3463944.3469270

3.Bao, Fan, and Jia Chen. 2014 .Visual framework for big data in d3. js. In 2014 Ieee Workshop on

Electronics, Computer and Applications, pp. 47-50. IEEE.