Contents

List of Graphs

1 Introduction

Super Smash Bros. Ultimate, a game for the Nintendo Switch is an immensely popular game with an amazing competitive community nationwide. This report was created with the aims of giving the Australian competitive community a better insight into character changes and skill gaps between each quarter of 2019 and differences between rising skill levels of states. This insight will be properly visualised through multiple data graphs including that of univariate, bivariate and multivariate representations.

2 The Data Set

Hello and welcome to my report about data visualisation. My name is Natalie and I am a third year Software Engineering student and I'm enjoying this course so far.

The data set I have chosen is that of character usage, skill gain and played matches revolving around that of a local competitive gaming community in Australia. The way I have chosen to present this data set is through that of local JavaScript which calls the Google Chart library. This means that I will be able to deploy a live website showing this data for later viewing and later analysis.

As data is always changing in database I was using, I decided to only pull data from last year (2019) to give a more accurate representation of data and split them into each quarter (every 3 months). Since I am an active player within this community, I believe that releasing this data publically will allow for a more insightful look into the past and how the metagame (accepted norm) of the community has changed.

I hope that you will find this interesting even if you don't play games.

3 Techniques, Methods and Execution

In this section, I will talk about my approach to the data collection; how I handled the data and how I displayed the data in web form.

3.1 Techniques - The Approach to the Data

The data that I was trying to access was that of a public API that I could easily access through Python.

The data was stored in JSON form which meant that I had to iterate through various levels of data to get the data that I wanted.

Since I was on a mac, I had to create a virtual environment to be able to access the API. Once I set up this environment, I created a Python script then called the API using a variety of similar functions to the code block below.

```
headers = {'Content-Type': 'application/json',
           'X-ApiKey': '6PG3OWV9UCVFZNTXQJKR'}

# ID being the character ID
def character_matches(id):
    url = 'https://api.ausmash.com.au/characters/{}/matches'.format(id)
    response = requests.get(url, headers=headers)
```

```
    content = json.loads(response.content)
    # JSON file of the content received and returned. It can now be stored for further
        use.
    return content
```

Using similar functions to the above, I was able to create a large dictionary in Python which contained all my data. This structure was split into quarters, by state and then by characters, with each character containing the unique players, the amount of elo gained and the amount of matches played. How I was able to sort this data into a more readable and interactable form can be seen in the next section.

(Click for API Documentation)

3.2 Methods - How I Handled the Data

Before working towards getting my data, I had decided that I wanted to pull elo gain (skill gain, it will be called elo throughout the report), the unique amount of players who played each character and the amount of matches played for the specific character. Each character's data would then be sorted into states which would then further be stored into quarters (every 3 months e.g. Q1 is from January 1st 2019 to March 31st 2019).

To do this, I iterated through every single character's logged matches and got each unique player, their state and the won elo if the player won the match. I would also get the date and would match the date to the associated quarter. Once this was done, I then added their data to a large dictionary data structure.

Once the data was fully sorted, I wrote the contents of the dictionary into a JSON file which was then to be used for my JavaScript file which would display the data using Google Charts. This JavaScript usage will be explained in the next section.

3.3 Execution - Displaying the Data

Using Google Charts I was able to create various suitable graphs for my collected data using native JavaScript. The chosen graphs that were used for my data visualisation were that of:

- Bar Charts
- Bubble Charts
- Line/Area Charts
- Scatter Plots (using the same library as Bubble Charts)
- Step Graphs

Order of iteration through my chosen data would change depending on which graph I had chosen and and what I wanted to display. This lead to many different variances of the same code which would each have different outcomes. For example, for some code I would need to iterate through quarters rather than the states. When running the script for my data, I also made sure to check the player's state of origin rather than where they competed to account for interstate tournaments and visits.

When I wanted to create the specified chart, I would create an array to hold the data in it called 'overallData'. This array would then hold more arrays which would correspond to different datapoints and axis that I wanted to display depending on the chart that I had chosen.

A demonstration of the JavaScript code used to display the chart can be seen below.

```
function drawChart() {
  var data = google.visualization.arrayToDataTable(
     overallData
  );

  var options = {
     title: chartTitle,
     width: 1800,
     height: 800,
     vAxis: {
        title: 'Y axis name'
     },
     hAxis: {
        title: 'X axis name'
     }
  };

  var chart = new
      google.visualization.ChartType(document.getElementById("chart-div"));
  chart.draw(view, options);
}
```

Using similar code blocks to this, I was able to create the following graphs which will be analysed in the next section.

Note: For a more detailed look into my code check out my

4 Data Analysis

Before beginning my data analysis, I decided to group my data into little subsections. These being:

- Player and Match Data
- Character Data
- Elo Data
- Combined Data

4.1 Player and Match Data Analysis

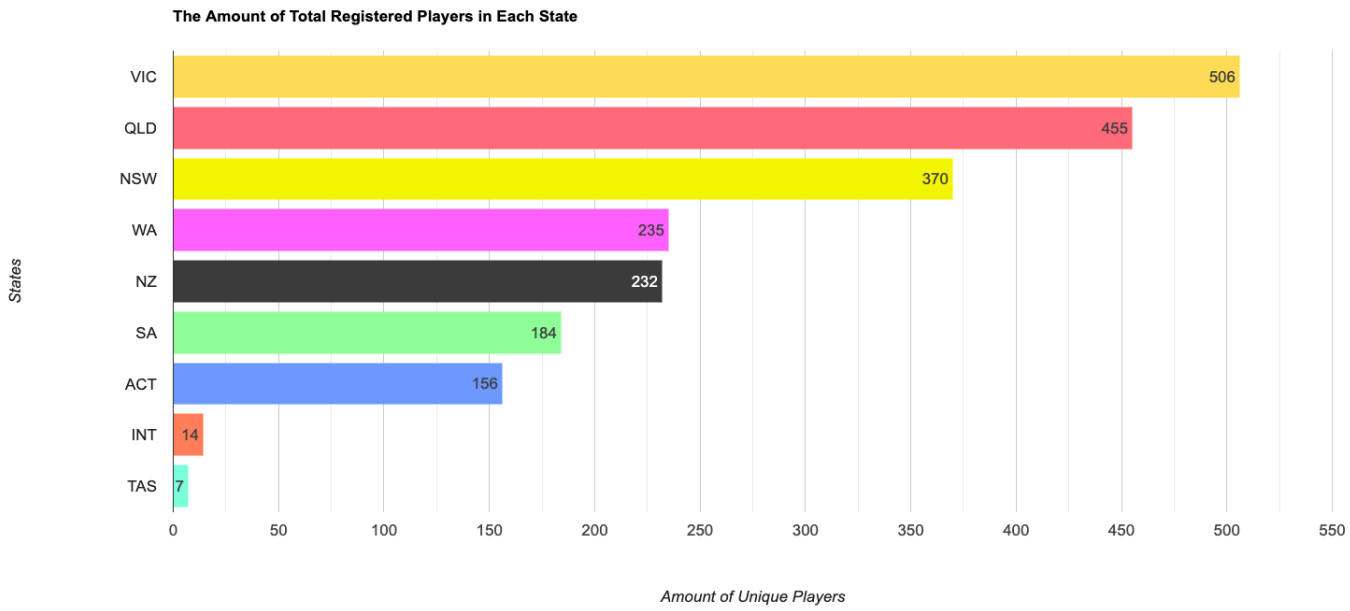

Figure 1: The Total Amount of Registered Players in Each State (2019)

The first graph that I wanted to display was that of unique players that had played in each state. From this graph, it is possible to see that Victoria (VIC) had the most registered players in 2019 while Tasmania (TAS) and International (INT) players have the smallest registered players, this may be due to the populus size of TAS and the lack of international competitors entering Australian tournaments.

It's worth noting that New South Wales (NSW) should have a much larger registered playerbase due to its large population compared to all the other states. This however, does not seem to be the case, making it an outlier in the dataset. This could be due to a lack of tournaments held per week which can lead to a struggle in encouraging new players to join the gaming scene.

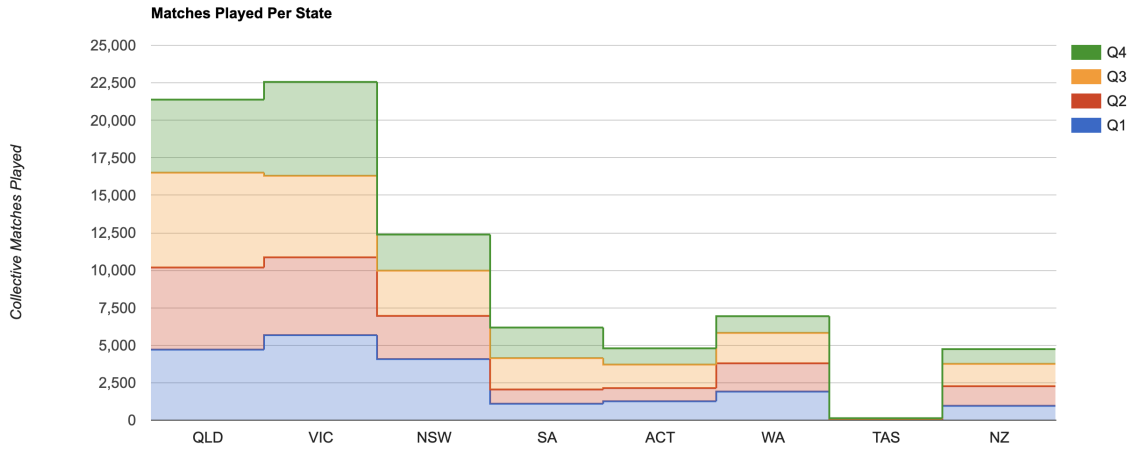

Figure 2: Step Graph Representing The Amount of Matches Played in Each State (2019)

Figure 2 shows the amount of matches that were played each quarter (Q) for each state. From the analysis of Figure 1, it is clear that the amount of players in a state can impact the total amount of matches played each quarter. As per Figure 1, VIC remains the state with the most amount of matches played across the year due to their large registered playerbase.

Something not seen within Figure 1 that is seen in that of Figure 2 is that in the first quarter, NSW and QLD had a similar amount of matches played. This may have been due to the game's recent release as it came out in December which would've have brought many new players to start the year off with but as time went on, NSW players started to drop out of the scene and stop playing which resulted in less matches played in NSW for the rest of the year as the quarters following Q1 resulted in less matches played. One reason for NSW to have such a drop off in player base may be due to its community not being able to engage with new players properly unlike that of VIC and QLD's older communities which have been around since early 2010's, allowing them to be more experienced with new players.

It is also noticeable that TAS has almost no matches played in every quarter, each being around 27 for most of them which was also evident in Figure 1.

Since matches played isn't too much of a good measure of constant activity within a community. I decided to divide the total amount of matches by the players within each quarter to see if any further trends could be found.

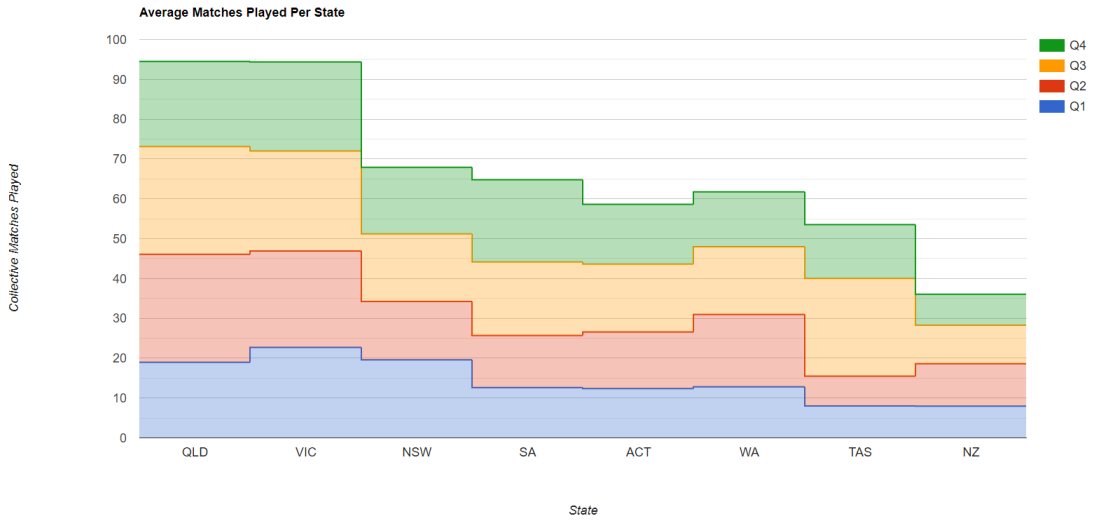

Figure 3: Step Graph Representing The Average Amount of Matches Played in Each State (2019)

This graph differs very differently from the previous one. One noticeable difference being that of Tasmania's matches played. In the previous figure, TAS was barely visible in comparrison to other states, with the average calculation in play. It is possible to see that the TAS' competitive scene is rather active as each individual player is active. Similar to the other graph, New Zealand (NZ) has very steady data, with the amount of matches played and that of active players. From this, it is possible to assume that NZ players all have very similar attendence. This may be due to the fact that NZ does not have weekly tournaments but rather monthly ones.

4.2 Character Data Analysis

When creating the bar graphs for this section of analysis. I decided to create an interactie bar graph which could show the character usage for each quarter per state and for the whole of Australia. Of course, this ended up with 32 different bar graphs to analyse... which I wasn't going to do. Instead, I will be analysing each quarter for the whole of Australia, there is a possibility that each graph will be looked at during the supplied presentation. It is also worth noting that due to the previous figures having a lack of non-average TAS and INT data, I have decided to take it out as they are outliers and are not good for individual analysis.

4.3 Elo Analysis

Elo is that of skill rating within the smash community and acts as a way of seeing who is truly the best. When two players play a set against each other, they are to gain or lose elo based on an algorithm built into ausmash's backend. I however chose to look into purely elo gain as graphs could be more interesting as more elo is usually gained than lost within each state.

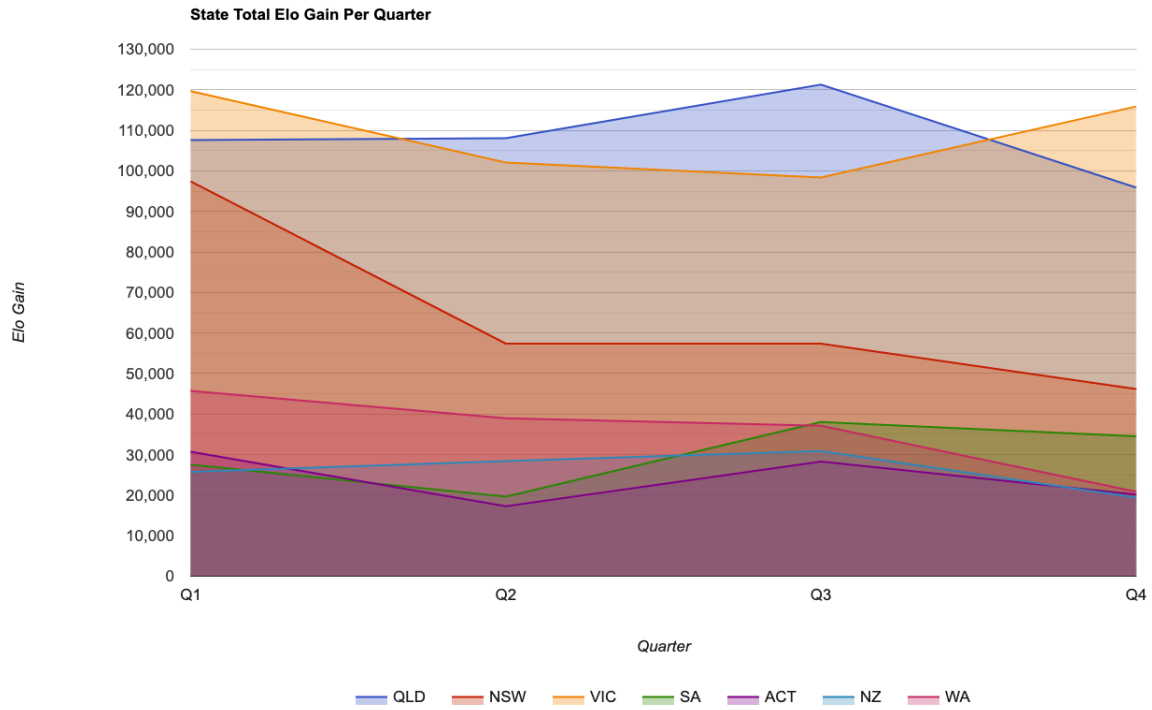

Figure 4: Line Graph Representing Total Elo Gain within each State (2019)

The above line area graph show that of the fluctuating elo gain over each quarter of 2019. As seen in all previous graphs, QLD and VIC are always competing for the most elo/matches played due to their larger playerbases compared to other states. It's worth noting that VIC currently has the most top players residing in it which could result in more elo gain for players due to larger losses. Similar to what was seen in Figure 2, NSW has a large decline from Q1 to Q2 due to the player base greately decreasing which results in less elo being gained. It's also noteable that Q3 (July - September) has the largest difference between VIC and QLD. This was due to a lack of large events (majors) happening during this quarter. Only two majors were present during this quarter with one being in WA and the other being in QLD. This meant there were more opportunities for QLD players to gain elo as the tournament was local.

Since elo gain is similar to that of player size, I decided to divide to total elo gain by the amount of unique players each quarter to get an average elo gain within each state.

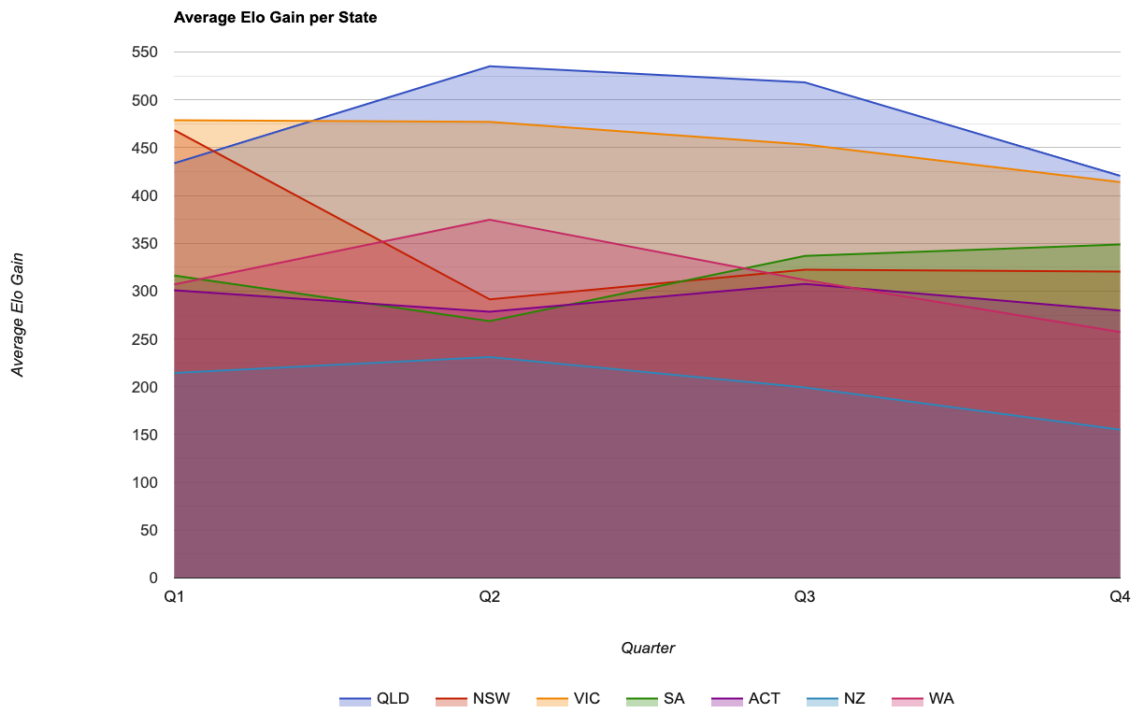

Figure 5: Line Graph Representing Average Elo Gain within each State (2019)

This change in from total to average elo gain can be drastically be seen. A main difference is that of the QLD vs VIC comparison and that of WA's spike in Q2. WA's spike particularly has not been seen in other graphs due to the player base size but due to the average calculation, it is easier to see the spike in Q2 where WA overtakes NSW in average elo gain. This could have been due to new characters being added to the game (Banjo & Kazooie specifically) which could have allowed for players to perform better in tournaments.

From this figure alone, it is safe to assume that QLD had better players all around when compared to all other states in 2019. Although QLD was behind on average elo gain in Q1, as the game evolved, QLD was able to catchup and eventually have a better average elo gain.

As this was hard to analyse due to the line graph's colouring, I decided to use the same data points for a bar graph which can be seen in the figure below

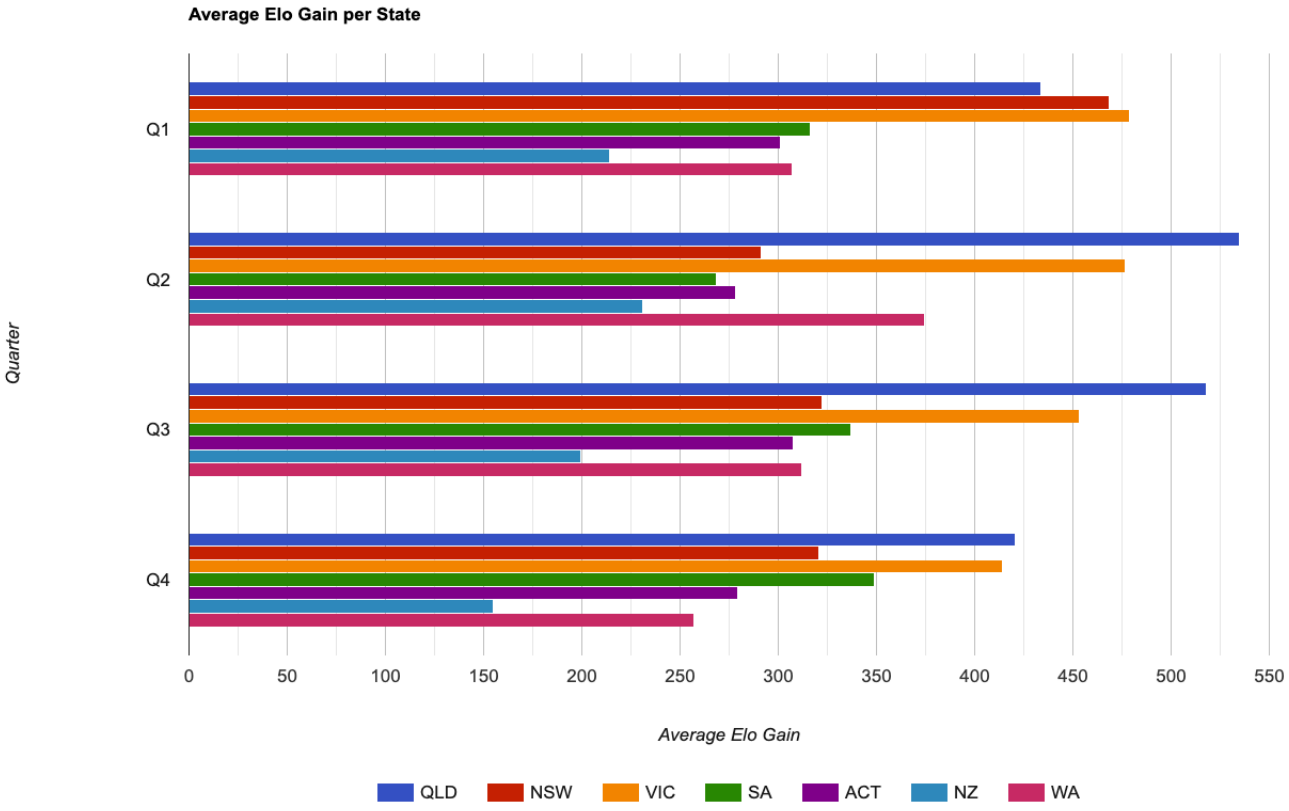

Figure 6: Bar Graph Representing Average Elo Gain within each State (2019)

What can mainly be seen within this bar graph is similar to that of the previous line graph but with states being easier to see. As stated before, WA had a large spike in average elo gain in Q2. This can be seen more clearly in the bottom bar of each quarter as WA slowly increases and spikes in Q2 before slowly going down for Q3 and Q4. It's also easier to see that ACT had a very steady average elo gain with no fluctuations. This may have been due to the relatively small but active player base.

Once again, it is safe to assume that QLD has better players all around due to its average elo gain across 2019 being much higher than other states for the majority of the year.

5 Extras