Team Members
Carl Gao
Jesse Chen
Shannon Zhu

LOL VIZWIZ

May 1, 2014

CS171 Final Project

League of Legends VizWiz

League of Legends VizWiz is a sleek, user-friendly site that allows users to explore the characters of the immensely popular Riot game, League of Legends.

Navigate the site with the search-able, sort-able bar across the top of the screen, and peruse interesting statistics on your favorite champions over time. Not sure who to choose for your next lineup? Or just curious to see the history of formerly over-powered champions... or the current standings of who you think are "good" and "bad"? LoL VizWiz is the place to go!

Overview and Motivation

We are fans of the multiplayer online battle arena game League of Legends, and appreciate the game for its strategic depth, variety of play styles, and focus on teamwork. At high levels of competition, players and coaches analyze statistics on the best tactics and team compositions in an attempt to achieve optimal win rates. Many casual players, including ourselves, are interested in improving our level of success as well.

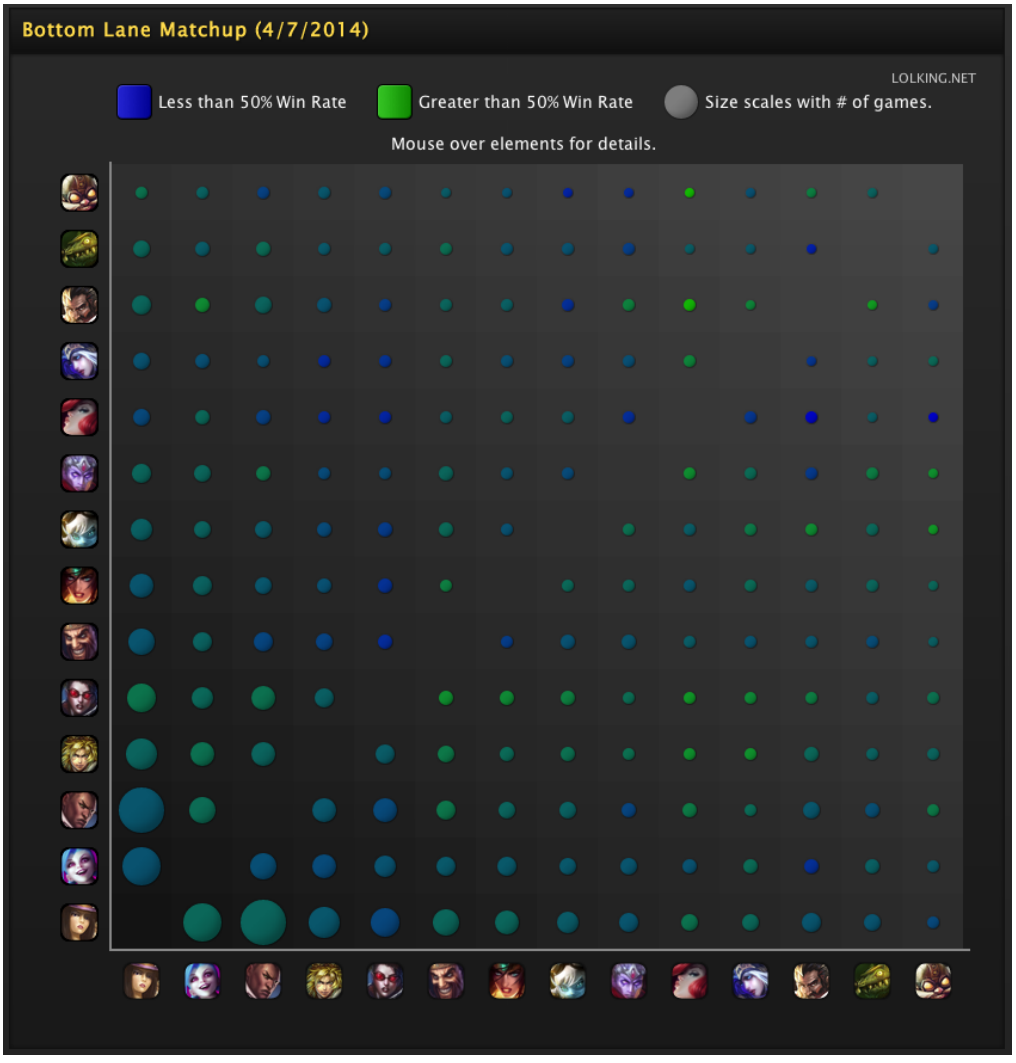
Thus, we would like to discover, through the visualizations that we will produce, the distributions of win percentages for certain combinations of champions on the same team as well as on opposing teams. We will provide visualizations for the pick rates and ban rates for champions over time, which would tell us which champions are the most popular in the current metagame as well as their historical popularity rates.

This would benefit League of Legends players, as they can strategically learn certain champions based on those who show the highest win rate. It would also be beneficial in drafting team compositions, since champion combinations with higher synergies can be shown through these visualizations. Many of the existing related visualizations proved to be very poorly designed and difficult to gather visual information from. We hope to consolidate useful information in a format that is clean, aesthetically pleasing, and most of all, easy to access and read.

Related Work

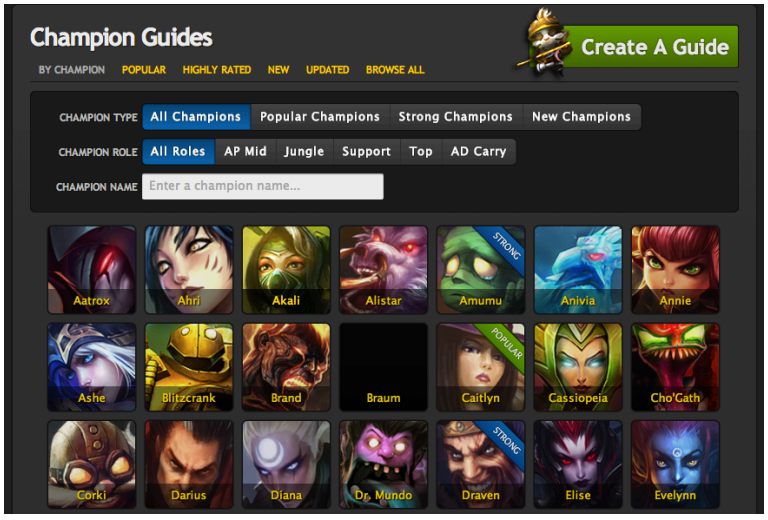
A variety of existing visualizations about League of Legends inspired and motivated us.

1. Visualizing win rates for lane matchups between champions. This visualization has a shockingly bad color scheme. The subtle shades between blue and green are very difficult to read, and do not effectively communicate the difference between winning on average and losing on average, since the 50% mark is an undifferentiated mix of blue and green. This motivated us to be very clear in our own visualizations and use position/length instead of color to represent important continuous variables. [Link]



2. This visualization shows the best teammates a particular champion can have. It's readable, but the bars are confusing because their lengths aren't proportional to anything in particular. Looking at this, we decided that distance from 50% would be a good baseline to extend bars proportionally from, such that bar length is abs(win rate) - 50%. [Link]

3. This visualization allows for the selection of champions based on a variable number of criteria, allowing the user to narrow down on a champion or subset of champions quickly. It doesn't presume a specific or narrow usage – if people are looking to complete a team, they are likely interested in searching for only champions of a particular name. If they are curious about a specific champion, they likely would want to search it directly. If they feel like browsing, that is an option as well. This visualization didn't directly inspire the searchbar idea, but we appreciated the sorting techniques and usability that we wanted to include in our own project. However, this visualization is tucked away in a tab and isn't integrated to the site – ours fixed this by being constantly accessible in a dynamic search bar across the top. [Link]



Questions

We initially wanted to answer the questions: how popular has a champion been over time, and which champions is the champion effective or ineffective with and against? Our first question evolved into our pick/ban rate graphs, and our second question became the synergy/matchup charts. As we explored our data and the possibilities, we realized that another crucial user need would be the overall win rate of a particular champion. So one of our new questions was: "How effective has a champion been over time in winning games?" This new element was incorporated on each champion page, to be linked and brushed along with the pick and ban rates – as they are similar graphs covering percentage-of-games statistics over time.

As we continued to work, we began to feel that it would be useful from a user's perspective, especially one new to the game or simply exploring the site, to be given a place to begin at the homepage. This inspired us to make a Top-5 and a Bottom-5 visualization, alongside a relationships chord diagram, to populate the home page and catch the reader's eye or give them a few champions to look up.

Most of our questions stayed consistent over the course of our work on the project, as the fundamental characteristics of the game remain constant and therefore the information that is "important" is hard to miss and hard to change. We stuck with our original ideas, adding the questions mentioned above, as we moved through! The only significant modification we made was to consider only bottom-lane synergies when looking at the chord chart, as this would create a chart that isn't too overwhelming but also has the most significant information. For background knowledge, the bottom lane is the only lane in which two champions play instead of one – so synergies there matter much more than across the rest of the game map.

Data

Our data sources for visualization data as well as images were lolking.net, loldb.gameguyz.com, and leagueoflegends.com. We used Python's requests and BeautifulSoup 3 library to scrape our data and images. For our data processing and cleanup, we structured our scraping programs in a very modular way to achieve flexibility and speed, having different code and data files for each type of data. The specific resource URLs and methods can be found in our code.

In more detail: to scrape the data for the pic and ban visualizations, we explored sites (after asking permission) that contained statistics on past games. When we navigated those sites to find the data we were looking for, we would Ctrl+F in the DOM inspector to find the HTML element, or find the data array in the site's Javascript files. Programs to parse HTML (as mentioned above) were then used to systematically index into text, extract string IDs, champion names, champion titles, win rates over the

past month, etc. We then populated our own data structures – a JSON dictionary – with the data that was scraped.

Similarly, the synergies and matchups data was scraped from similar sites. A python script was written to programmatically find the HTML divs related to the data we were searching for, and then to selectively grab and filter that data before exporting it as a JSON array. One problem we ran into was that the "top5" and "bottom5" matchups and synergies were not differentiated in the site we scraped from, so we had to re-scrape this data in two batches to make sure our selection criteria were more rigorous for each round of scraping (especially since dictionary data structures are unsorted).

The champion roles were scraped from a page on LoLKing that contained a list of each champions contained within their own divs. These divs contained a variable number of meta tags that contained the lane(s) that champion usually plays in the game. Similarly to the processes above, the HTML was scraped from this URL with BeautifulSoup3, and the resulting text was parsed for certain strings that would locate the tag of the champion name, and a variable number of meta tags. These were added to an array, the names were modified to line up with the names we had in our JSON structure containing all champs, and the two were merged.

The images were also grabbed in this manner from the League of Legends site– two for each champion: a thumbnail and a portrait. With each click, a JavaScript function was written that located the champion that was clicked on in this dictionary and modified the background of the website to the "portrait" value held in that dictionary.

All data not mentioned in its own paragraph was scraped, cleaned, and processed in a similar fashion – however, most substantial sources of data and their scraping processes are detailed above.

Exploratory Data Analysis

We used existing online visualizations to grok our data and look for insights on how to build our own visualizations. As mentioned in Related Work above, we found various flaws that informed our own designs. In addition, we got a feel for the range of values for various data, so we knew what scales to draw them to.

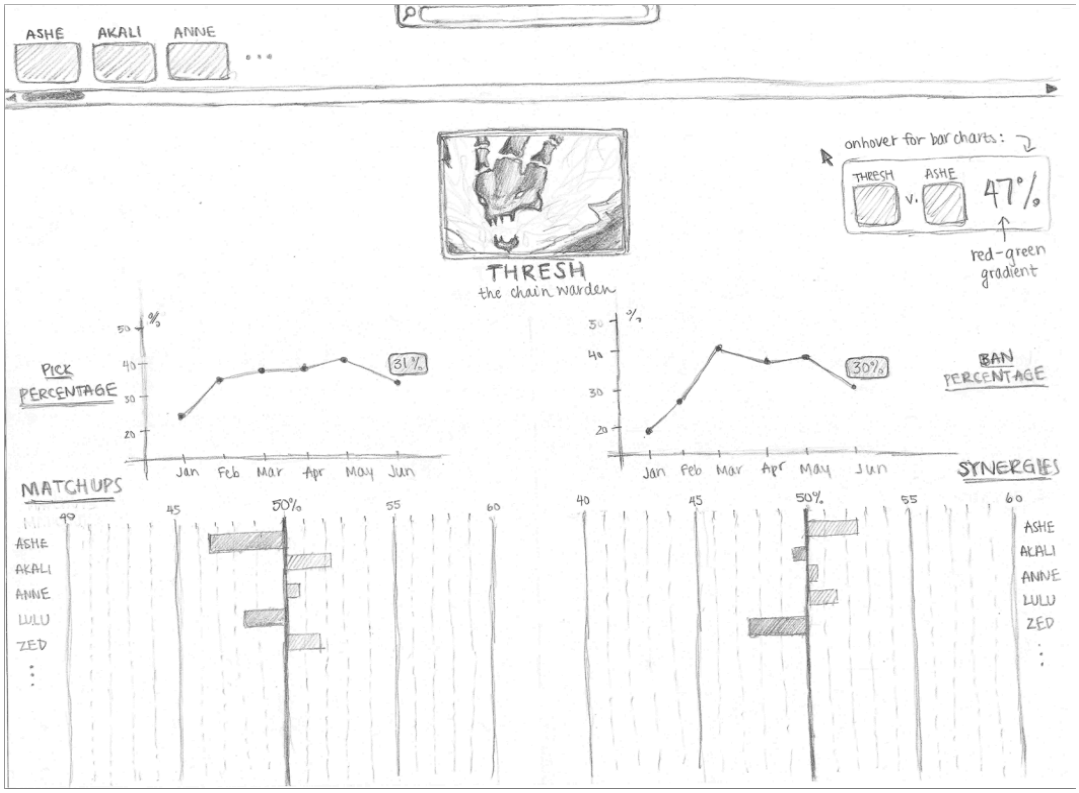
The visualizations that we built to explore our data were the same ones we built to display it in the end. For instance, we built the line charts to display the win-rates, pick-rates, and ban-rates of each champion. We found that, as we expected, most of the data was clustered around 50% - therefore, we were able to restrict the range of the win-rate visualization to the 40% to 65% range in order to focus more on the variability in the data rather than present a less informative, relatively straight line hovering around 50%. This information was gathered when we initially built our crude, un-styled line charts. Upon

confirming that our data looked like this, we continued on with our design (with some y-axis tweaks), and built the line charts with linking and brushing as currently displayed on the website.

Similarly, when we were looking at the "related work" on matchups and synergies, we noticed that most champions' matchup and synergy win-rates also hovered around 50%. However, this 50% mark is essential in determining whether or not a champion is overall "good" or "bad" with someone, and we were inspired by the poor visualizations we came across to make this point very clear. We also explored this data with a crude version of what is now presented in the project, and again found that a bar graph diverging at 50% represented the data cleanly and with the appropriate range, and we were able to move forward comfortably with this insight in mind.

Design Evolution

The original design included a large headshot of the champion in the center of the page, above any visualizations. However, as we began constructing the page with user goals in mind, we realized that keeping one portrait and name at the top was not sufficient to emphasize which champion the page was currently on (this is extremely important, because all the graphs only list names of other champions that the current one is being compared to). As the user scrolls down, we did not want the current champion to become less apparent.



We also felt the user experience would be more compelling with a larger, more colorful page, so we decided to modify our design into a fading background and translucent data boxes floating above it.

We originally planned the search bar to remove all other champions that weren't matching the portion of input the user had already typed. We decided to include them, faded out, both for aesthetic reasons and to make the other champions more accessible even if they were not explicitly searched for in the particular query. We faded them out very significantly, however, to ensure they are not distracting to the eye.

In terms of other features of the search bar design, we wanted to include a home button in a place that would leave the page feeling balanced. We didn't want to add a logo or place the home button on the page itself, as that felt like it was compromising real estate from the visualizations and violated the clean, window-like design we wanted to present to users. Having the home button be visible but unobtrusive was important, but we changed the original design to allow for both the home button and the champion icons to light up on mouseover (and for the cursor to change to a hand) to emphasize the clickability of these buttons.

When designing the graphs for each champion, we didn't want to cover up arbitrary parts of the background, as that would be aesthetically frustrating and also shows the user less information. To remedy this, we decided to overlay black boxes on the background that were dark enough to make the text and visualizations very clear, but also so the background image remained coherent and visually appealing.

We decided to place the win-rate line graph as the first graph that is visible because it presents the most general information possible about the champion. To this extent, we decided to keep the balance of graphs on the page slightly uneven, so the champion's face (which is always positioned to the top right) is shown. To remedy the balance to some extent, and also because this space was very open, we placed a shorter brushing tool that covered all three line graphs at once. This design decision was made with the rationale that a person using the brushing tool would be interested in a specific time in the past, and the other graphs are rendered useless if they still depict the old visualization frame. Win, pick, and ban rates are highly correlated and mutually support each other, and under very few circumstances would anyone decide they wanted to see them displaying separate time frames.

Another change in design we made along the way included the setup for the matchups and synergies. At one point we wanted to display all 118 champions vertically in each chart, but realized that this information is not quite so useful and would also require immense amounts of scrolling. The most important information for a user would be the champions this particular character is the best or the worst with – or the best or the worst against. To keep the page clean and readable and to keep the graphs from being overwhelming, we decided to convey this most essential information in the top 5

best and worst synergies and matchups. The data for the in-between champions was also difficult to find reliably or to come by, so instead of compromising both the design and the reliability of the data we displayed, we ended up deciding that this indeed was the best decision given the situation.

The design for the search bar toggle was also discussed. There were ideas thrown around of making it a drop-down (as we ran into the same problem with the home button – not wanting to make it obtrusive on the clean, scrollable visualization area below the searchbar) but also not wanting to create too much clutter along the top. Eventually, we decided that accessibility was also very important, as people using these buttons are more likely to be users needing a quick decision on who to play – and therefore need to filter their options by lane. We kept the buttons unobtrusive by keeping them small and plaintext, and making it very apparent which is currently selected by filling the entire button in. They were originally placed on the left top corner, but realizing that this seemed unclean and much more obtrusive as it is the "origin" of the page and the first corner that is seen, we decided to move it to the right side.

In terms of general design decisions, we wanted to keep the colors neutral – and when they were not neutral, in the case of the red-and-green synergies and matchups bar charts, that they were colored for a very clear and logical reason (no rainbow scales were used, yay!). We decided on nothing more complicated than a line chart or a bar chart because we felt that these basic but powerful visuals conveyed the data most cleanly and most naturally, not requiring the user to examine keys or color schemes or other unusual contraptions in order to glean the data from the visual. While this was discussed in class as a tradeoff between simplicity/cleanliness and memorability, we felt that other aspects of the site could capture the memorability aspect without negatively affecting the most important part – the data visualization. To this end, we kept the line graphs very simple, and added brushing and linking as a way to keep the user interested and also to make the data more accessible for each user's needs. Likewise, we allowed the bar charts to be sortable by champion name or by value – so if a person is looking for a particular champion to see whether their favorite synergy is supposed to be good, they can do so easily. Or, if someone else is looking to see what champion is the best to play with a particular other one, that is also easy to look up as well – just sort the data to display by value.

The only other major design decision we ended up debating was on the y-axis scales of the line graphs. If we kept all of them consistent, that may present a more comparable picture across multiple graphs. However, this would also make them less readable and create a lot of white space either above or underneath the essential line. We decided to make the y-scale variable to the data it is attempting to display, because cross-graph comparisons are likely to be only pattern based. Ban rates will always be on a very different scale than pick rates do – but that's okay, being able to tell how much one goes up by as opposed to the other is the point of the visualization. The absolute numbers are also highlighted by the tooltips when one point of any of the charts is moused over, making the absolute values also very apparent to the viewer.

Targeted Users

One of our biggest considerations when designing this project was the applicability of our visualizations to a wide user base. We wanted the site to be generally aesthetically pleasing and easy to navigate and generally understand regardless of experience with LoL. However, we were mostly concerned with varying types of users among those who do know LoL – as some users would want to browse the champions out of interest for the game, and others may want to use it in a more "intense" manner by being able to quickly navigate to champion information and make an informed decision

Implementation

The search bar began by scraping the web and data wrangling the result into a json. This created an array of objects, each containing a character id, name, thumbnail image, and portrait image. These images were also scraped from the web into two folders, and their local links were put into the JavaScript object.

Then a textarea was constructed and styled, then programmed to take an input that is not case-sensitive, accommodates acceptable characters such as spaces and apostrophes, and matches the to the array of champion objects by id. This is filtered to create a pair of set and complement arrays, which triggers another function that creates DOM elements that sets up the search bar system. Upon click, these divs are given properties and styles, and trigger side-effects in other divs in the DOM that set the background image, for example. They also change the dataset being passed to D3 code, which the visualizations use to display.

The search bar was also tweaked to accommodate the hitting of "enter" prematurely, which will lead to a champion's data as long as there is no ambiguity over the requested champion. The search bar and all of its features were entirely designed and coded from scratch! A lot of time was spent getting the search feature to catch all corner cases and provide a smooth user experience, as well as thinking about and implementing the structure of the pages themselves.

When brainstorming ideas for our home page, we knew we wanted a large and informative visualization about an aspect of the game encompassing many different champions. So we decided to implement a chord diagram detailing the popularity of various Attack Damage Carry-Support partnerships. We began by scraping data from LolKing, then wrangling the data into a matrix that the chord diagram could take as an input. We then customized the labels on the diagram to correspond to each champion's name, and color-coded the segments to show which role each champion plays.

Another important visualization for our home page was the top 5 and bottom 5 win rate charts. We wanted to give the average user an easy takeaway message about the overall viability of certain champions. We wanted these charts to quickly convey the best and worst champions in the League in the current metagame, so we decided to utilize the unique look of color-coded side-by-side bars where bar length corresponded to extremity of either winning or losing a lot.

Moving onto our individual champion data pages: in order to give the user a way to very finely examine a champion's effectiveness and popularity over time, we created a set of 3 graphs to show win rate, pick rate, and ban rate. These graphs are all linked together, in that hovering over a point in one highlights the points for the same date in the other graphs. We are also proud of our time selector functionality, a time axis that the user can click and drag on to focus in on the time frame he or she is most interested in seeing for each of the 3 graphs. Through the time selector, we implemented linking and brushing on each of the 3 graphs it controls.

We designed our matchups and synergies graphs with empirical practicality in mind; a user interested in a particular champion must want to know which teammates to partner with, and which enemies to avoid. As such, we had each graph sorted by default by win rate, so the user can at a glance note whom the most extreme partners or enemies are. We also decided to offer the user an option to sort alphabetically in case he or she is looking for a particular champion.

As we polished our project, we came to realize another common and important use case was if a user is looking to gain expertise with a new champion role, and would like to see details about champions who play a specific role. Thus, we added the role filters for the search bar at the top-right corner.

Evaluation

Some very fascinating takeaways that we got from having worked with this data include a general appreciation for the producer-change to consumer-behavior delay. For example, as most players of League of Legends know, some champions such as Kassadin were released with abilities that created an unbalanced game. In this case, Kassadin was over-powered, and people responded by banning him at in approximately 90% of the total games played. When Riot, the game producer, realized this imbalance, the eventually "nerfed" him, or decreased his playing power in mid-March until he was very comparable with the other champions. The player reaction to this is shown vividly in the data and is very interesting – ban rats began to drop heavily in a nicely shaped parabola, while pick rates jumped the first week he was "nerfed" as people presumably thought "wow, here is a chance to play Kassadin!" and did so. However, they quickly realized that he was actually significantly worse, than before, and pick rates quickly dropped off and stayed low from then on.

Other patterns also emerge, especially since League of Legends rotates free champions every week. Unless you have accumulated enough in-game credits to buy a champion permanently and play it whenever you like, your selection is limited to a dozen or so champions that are available that week. The patterns in pick rates for these weeks is very interesting – the data shows that at the beginning of a week when a champion is newly free, that champion experiences a distinctive jump in pick rates. However, as the week goes on, pick rates markedly decline as people become less interested in the novelty of the champion and begin to resort to their usual, owned and perhaps non-free, champions to play.

These insights on player behavior, among others that are not explicitly listed here, and extremely interesting to explore and actually answer more than just the questions we posed at the beginning of our project. Those questions were certainly answered too – the pick and ban patterns over time match the history of League of Legends and each individual champion rather closely. These new discoveries also lead to interesting conclusions, such as that games played near the beginning of the week are much more likely to have people playing the free champions and perhaps champions they are less familiar with – and games later in the week are more likely to contain people playing their old favorites.

Besides being purely interesting, patterns such as these that can be found in the data we have visualized can affect game decisions or can influence what people expect when they enter a game. The visualizations themselves work very nicely, and there is no specific visualization goal that we fell short on. All of the linking and brushing and tooltips work very well for the pick-ban graphs, and the bar charts showing synergies sort by alphabetical order and by value. The search bar is very functional and accommodates for special characters in names, etc. If we were to improve the visualizations further, we would like to build in an auto-updating feature that can grab statistics weekly and update the graphs on its own. This would ensure the website remains up-to-date and useful to players at any time.