LOLytics Process Book

https://github.com/Entranco/dataviscourse-pr-lolytics

Dhanpaul Nandkeshwar

u1089230

dpnandkeshwar@gmail.com

Andrew Fraser

u1171089

u1171089@utah.edu

Background and Motivation

The League of Legends Worlds tournament is one of the largest eSports events to date, garnering millions of viewers every year. One of the most exciting features of the game is that it never stagnates. New champions (which are controlled by each player) are released each year, often significantly influencing the metagame. Updates are usually released every other week or so with a slew of buffs and nerfs (changes that make champions stronger or weaker). Although these changes are often minor, their effects on professional play and the metagame of professional play is monumental. For example, a very small buff to a champion called Maokai right before Worlds 2022 has led to him being one of the most selected champions of the entire tournament. Before this buff, he was seeing little to no competitive play. Professional players have such extensive game knowledge that they will abuse every small change to any champion in the game.

This dynamic metagame has fostered a gold mine of data, especially during Worlds season. As avid League of Legends viewers and players, we wish to analyze how the metagame at Worlds has evolved over time with respect to champion picks and strengths.

Project Objectives

At its core, League of Legends is a very strategic game. Making the best decision and play to make is entirely based upon the information presented to you and assessing risk versus reward. The main components of what decisions you are allowed to make are based on the champion you are playing, how much gold you have, and how much gold the team has. Our objective is to analyze data related to these two fields at the highest level of play: the League of Legends Worlds Tournament. We would like to see how the changes to the champions and changes to the game that happen very consistently affect play at the highest level. Specifically, we aim to analyze what champions are being picked during the Worlds tournament, what champions are being banned (not available to be picked

for that game, a total of ten champions can be banned for any given game), and how these champions affect win rates. We also aim to analyze the reason for these outcomes, such as, how much gold did this champion generate, how many kills, deaths, and assists did they have by the end of the game, and so on. We would also like to analyze some of the metadata of the teams that are competing in Worlds, such as which teams picked what champions and what was their success rate with these champions.

Data

We plan to use the Worlds champion datasets from Worlds 2015 up to Worlds 2022. Datasets do exist from the older Worlds tournaments, but they are not nearly as detailed. The new datasets not only contain statistics about which champions were selected, but also their roles, how much gold they had, and their kill/death/assist ratios. This additional data can help us determine whether a champion actually performed as well as expected in the metagame. This could also point to a shift in strategies if a certain champion did not perform well or performs too well. Counterpicks are often developed throughout the tournament as the metagame is explored further, so it would be interesting to see such effects.

Data Processing

Luckily, the data is already easy to extract into csv format. We don't need to perform any additional processing to extract it. We may perform some small transformations on the data based upon certain statistics we care about. For example, we may choose to calculate the most popular champions in each role or the damage/gold generated by each champion. These transformations should be fairly straightforward.

Visualization Design

We want to have two key visualizations. One visualization is used to compare statistics over time, while another is used to compare statistics over various champions. We considered encoding all of this into one, but it is only semi-possible to do so. When more than a few champions are featured, things might get too confusing. You can follow our line of logic in the sketches at the end of the proposal.

Must-Have Features

- A basic bar chart allowing for comparison of any column in the data. This allows us to see a difference between champions.
- Filtering the data by certain key terms (champion, years of the Worlds tournament, roles)
- Sorting visualizations based on the most pertinent columns of the data (gold, K/D/A, damage)
- A chart comparing various statistics concerning each champion with respect to each year of worlds. This allows us to see a change over time.

Optional Features

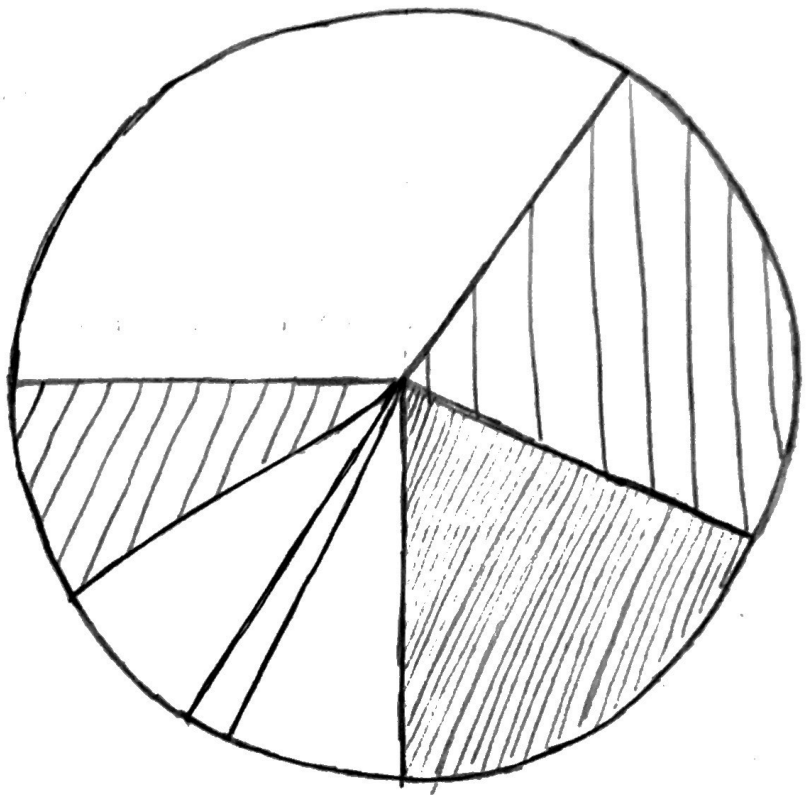
- Sorting visualizations based on every column of the data
- Filtering the data based upon player
- Including an additional exploration of player statistics, potentially connecting them to their most chosen champions
- Box and whisker plots of each column of the data to indicate how common certain values are
- A pie chart illustrating all champions with respect to each statistic
- The ability to analyze matchups between various champions in Worlds (how often champion X defeats champion Y)

Project Schedule

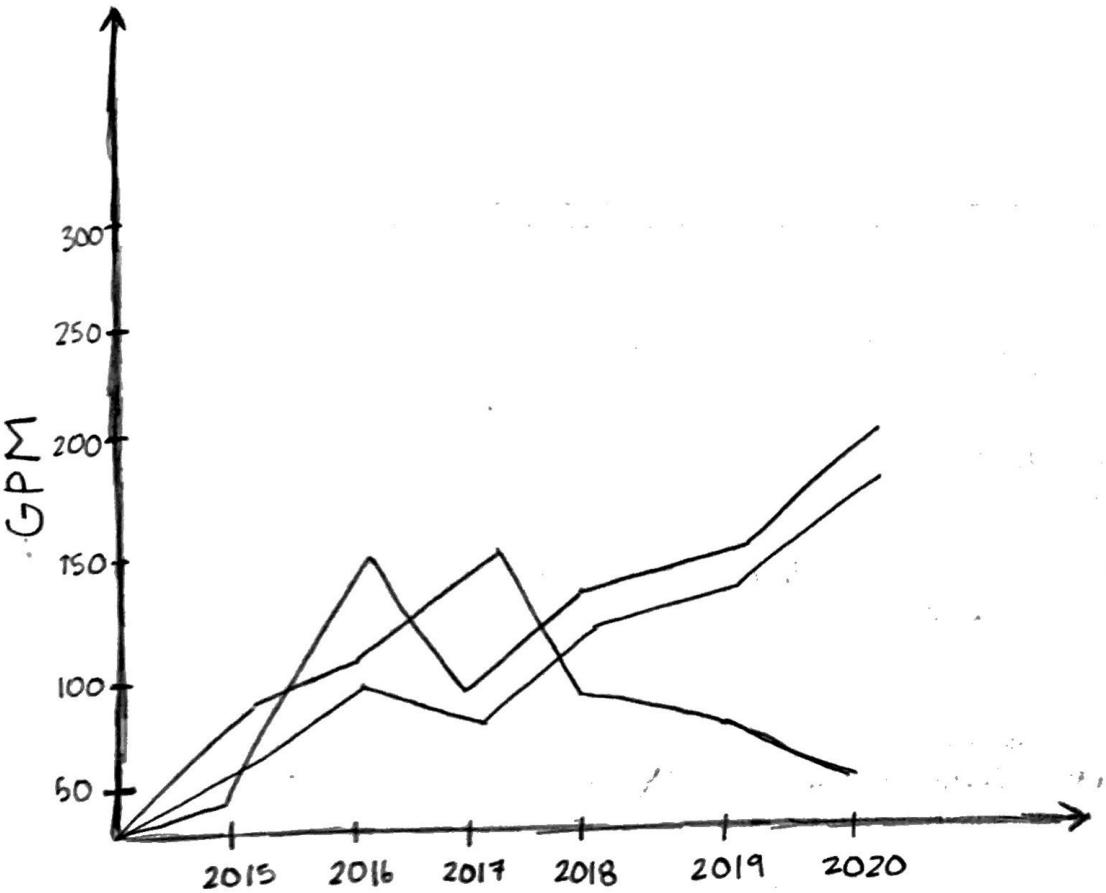
- Week 1: Process the data into javascript, get some kind of visualization on the page
- Week 2: Get the visualization functioning for a single column of the data
- Week 3: Add modularity between different columns, as well as sorting of the data
- Week 4: Work on a visualization over champions to complement the visualization over time
- Week 5: Polish the website, make things looks nicer, add text
- Week 6: Add any optional features that we decide are most important
- Week 7: Prepare for presentation, catch up if necessary

Filter: Year

Filter: Gold/Pick/Win/Loss



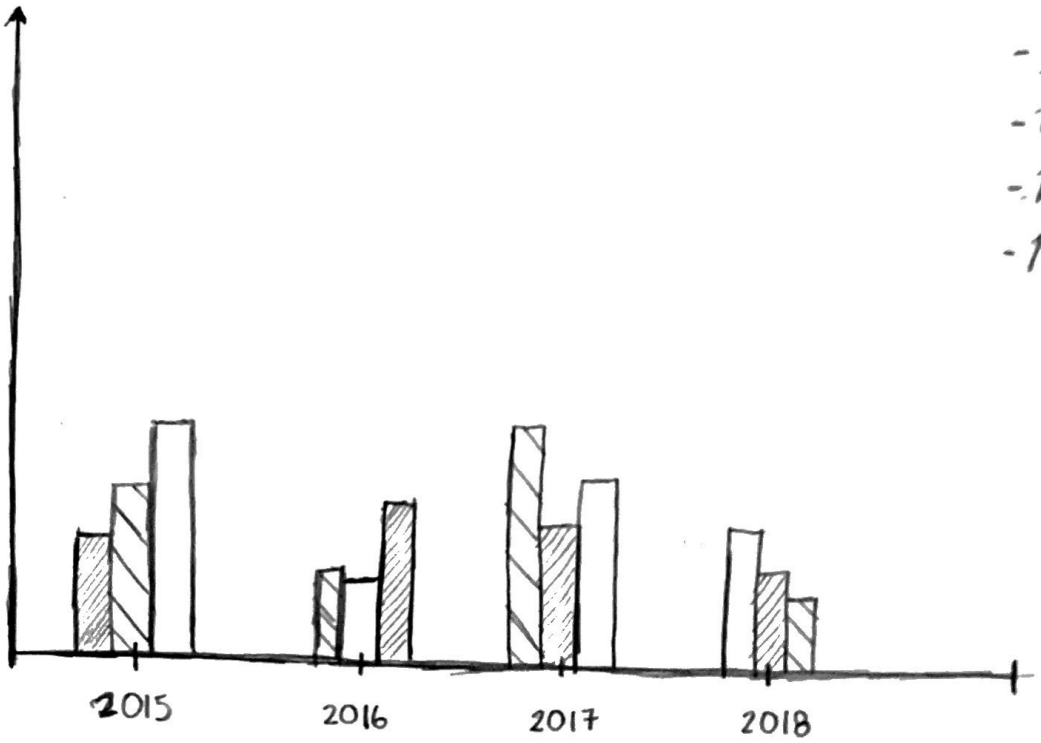
We change what rates we want to change as part of the pie. Each slice of the pie is a champion, coded with color. We can select the year.



* Line per champion, can select specific champions if you want, or show all by default

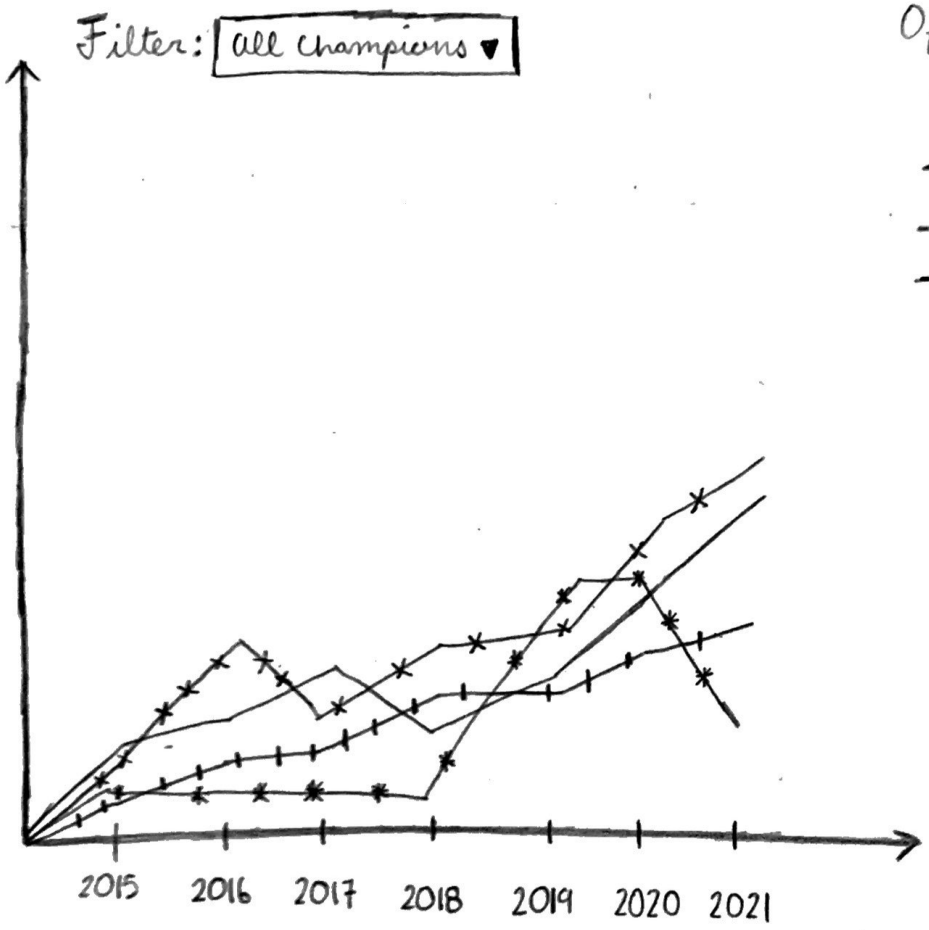
Options for Y-axis:

- Gold per minute
- Win Rate
- Ban Rate
- Pick Rate



* Different bars for different champs, different colors for different roles or other encoded data

We use bars to encode various champions, each with different color. Y-axis is variable. However we can only use a few champions at a time due to cluttering of bars. We could also show more by changing X-axis to show fewer years



* Lines for each champion, can select only ones you want

We combine aspects of all graphs by introducing variable y-axis selection, as well as different line colors and textures to encode additional data. X-axis encodes the year.

Scraping the Data

The first step we took was gathering our data. Luckily, as described above in our project proposal, this is fairly easy to do. Soon after scraping the data, we actually found an even larger gold mine of data. Using the same source of data we planned on initially at LoL Fandom, we found that if we only use the data from the main event, we can actually get the data for every single year of Worlds from 2011 to 2022. This is a huge step for our project, as we wish to plot the data values over time and this gives us as wide of a time span as we could hope for.

Regardless of this discovery and the extra data it required us to scrape, the scraping itself was quite simple. Each year has a website containing the data that we want in a table. By copying and pasting that data directly into a spreadsheet, we were able to download those spreadsheets into .csv files. We did have to add column headers manually, but otherwise this was very smooth.

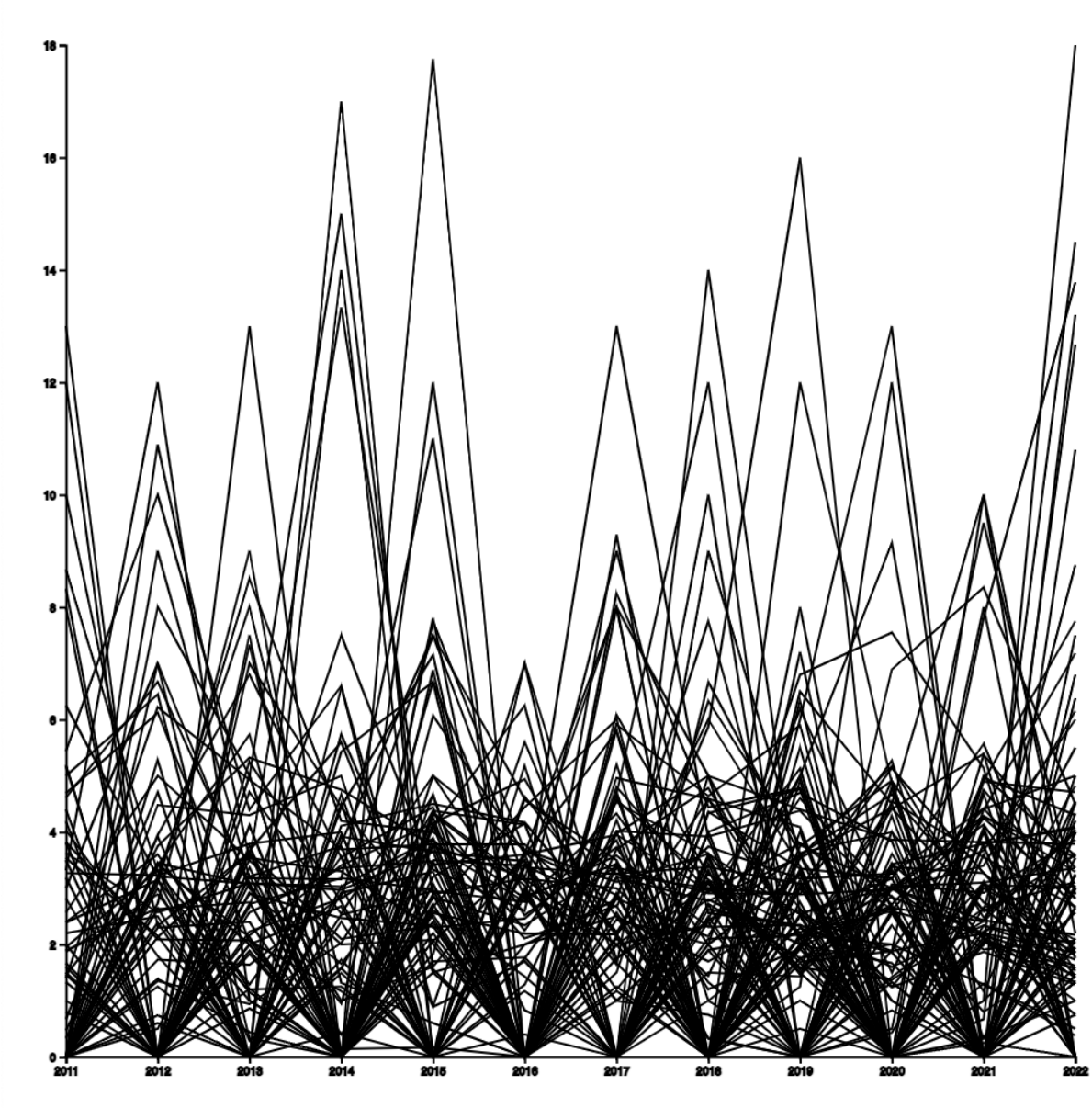
Selection of Data Structures

As LoL fandom displays one table of data per year of worlds, the most natural way to process the data was by year. We created a map called datMap mapping each year to its table of data. However, this turned out to be insufficient for everything we wanted to do. Here is a list of all of the data structures we are using:

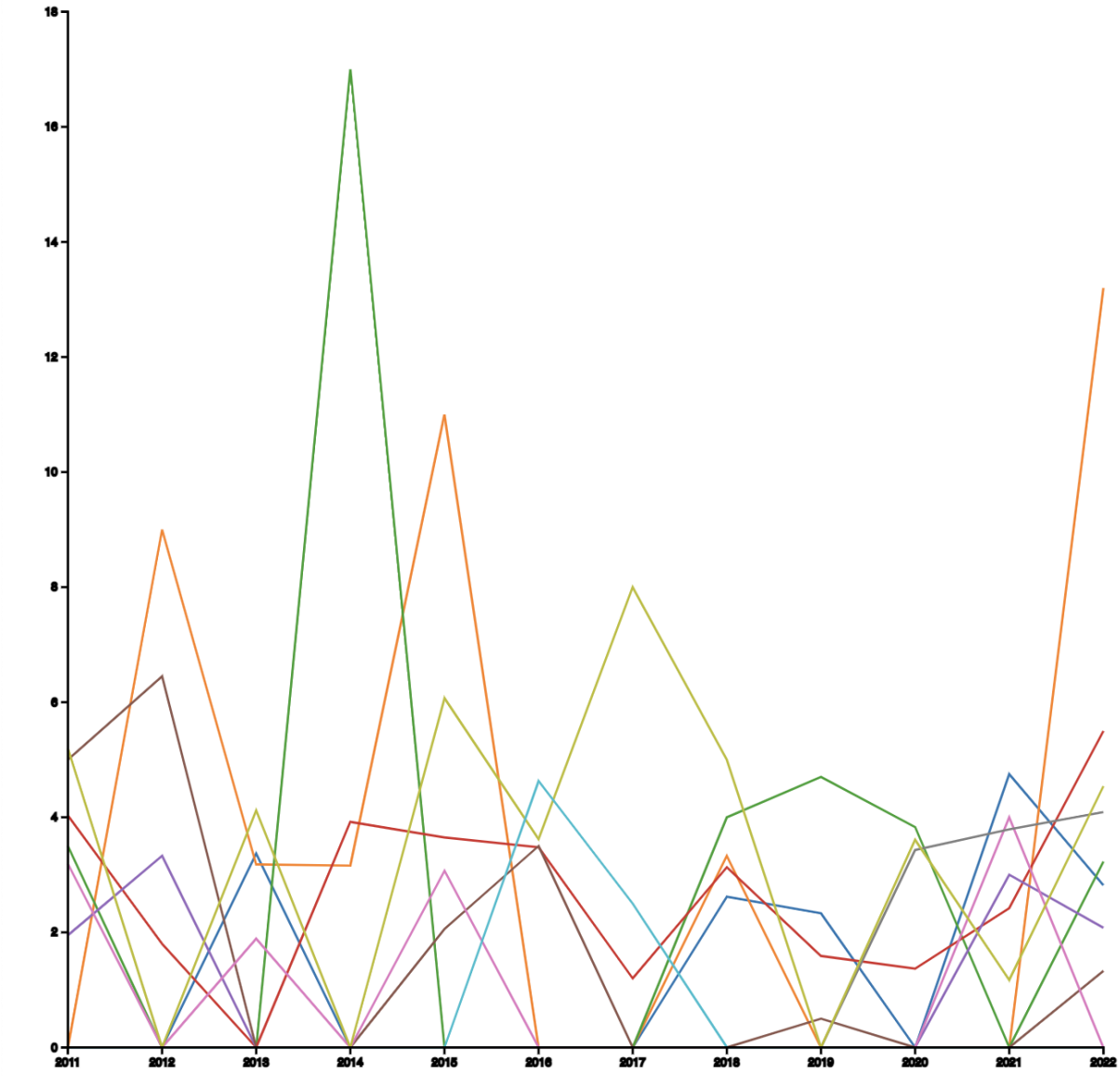
- dateMap: A mapping from year to each table from LoL Fandom. The natural way to process the data.
- champData: A mapping from champion to every row that champion appears in. Very useful when trying to highlight only certain champions in a visualization.
- Years: A list of all years in Worlds. Useful when trying to access items in dateMap.
- Champs: A list of all champions that appeared in at least one Worlds game. Useful when trying to access items in champData.
- Cols: A list with the name of every numerical column. This makes it easy to convert all strings that should be numerical to their numerical values.

Line Chart (Champion Stat / Time)

When creating the line chart, we started by trying to use the dateMap data structure. However, this very quickly got out of hand, as we could not easily access the data that was relevant to one champion and map it to a line. This made it clear that we needed a second data structure, leading to the creation of champData. Using champData, we were able to fairly easily create one line per champion on our line chart. Here is our initial prototype, which graphs the KDA of each champion every year.



As you can see, our prototype clearly has far too many lines, as more than 100 champions have seen play at Worlds. We anticipated this issue during the feedback process and had brainstormed a couple of solutions. To match with our ideas for making the data interactive, we decided to allow the table selections to filter how many champions are shown. To do this, we initially show the first 10 champions alphabetically. Then, as champions are selected from the table, we show that champion's line on the graph. The line and the table row match in color.



Single Year Champion Data

One of our main visualizations, we have implemented a table. To create the table, we create access “dateMap” so we can cleanly get all the data for a single year. Currently the table is only showing about half of the metrics in the data, as using all the metrics appears to clutter the visualization. We have chosen to show all champion data for only a single year at a time in the table, and have incorporated a drop down menu to allow us to change the year that will in turn change the table. For this version of the table, we have highlighted the column names as well as the champion column to help identify the more important sections of the data. We also intend to incorporate a scroll bar for the table, so the entire table doesn’t take up space on the screen. Below is a screenshot of the early implementation of the table.

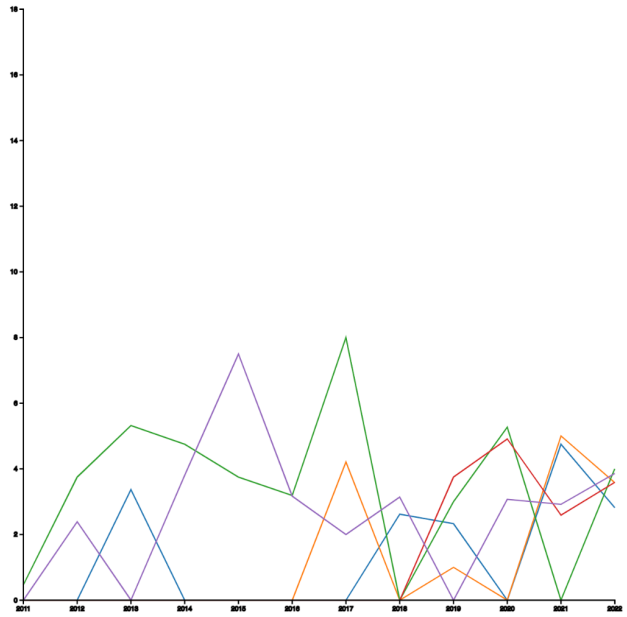
2022

Champion	Pick/Ban	Pick/Ban Rate	Picks	Bans	Players	Wins	Losses	Winrate	Kills	Deaths	Assists	KDA
Aatrox	80	100	58	22	12	12	10	55	2.73	2.59	4.59	2.82
Yuumi	76	95	68	8	6	8	0	100	1.63	1.13	13.88	13.78
Sejuani	72	90	42	30	17	16	14	53.3	2.03	2.6	7.27	3.58
Caitlyn	70	87.5	62	8	5	5	3	62.5	3.25	1.88	4.25	4
Azir	66	82.5	29	37	13	19	18	51	2.95	1.81	4.3	4
Sylas	60	75	28	32	10	21	11	65.6	3.56	2.16	4.19	3.59
Graves	59	73.8	34	25	11	12	13	48	3.16	2.36	5.96	3.86
Maokai	56	70	35	21	14	7	14	33	1.19	3.14	5.62	2.17
Lucian	55	68.8	29	26	9	15	11	57.7	5.12	1.58	4.58	6.15
Viego	48	60	18	30	9	19	11	63.3	3.57	2.13	5.6	4.3
Akali	48	60	24	24	11	12	12	50	3.67	2.17	3.33	3.23
Fiora	42	52.5	23	19	10	9	10	47.4	2.63	2	3.16	2.89
Renata Glasc	41	51.2	20	21	10	13	8	62	1	2.38	9	4.2
Aphelios	39	48.8	8	31	10	17	14	54.8	3.97	2.06	4.48	4.09
Renekton	39	48.8	23	16	9	5	11	31.3	2.44	3.44	3.94	1.85
Viktor	35	44	16	19	10	6	13	31.6	2.26	3.37	4.53	2.02
Kalista	31	38.8	13	18	10	7	11	38.9	3.44	2.72	4.22	2.82
Ornn	31	38.8	15	16	11	9	7	56	1.31	2.19	5.75	3.23
Nami	27	33.8	1	26	9	15	11	58	0.85	2.27	9.85	4.71
Varus	27	33.8	8	19	10	9	10	47.4	3.21	2.16	4.42	3.54
Lee Sin	23	28.7	8	15	9	7	8	47	1.47	2.87	6.13	2.65
Poppy	22	27.5	9	13	9	4	9	30.8	1	2.23	5.77	3.03
LeBlanc	22	27.5	14	8	4	3	5	37.5	3.5	1	4.25	7.75
Camille	21	26.3	10	11	5	8	3	73	3.64	2.64	5.91	3.62
Heimerdinger	21	26.3	13	8	3	4	4	50	0.88	3.38	6.25	2.11
Vi	18	23	6	12	8	6	6	50	1.67	3.08	7.58	3
Nautilus	18	23	11	7	6	1	6	14.3	0.71	4.43	5	1.29
Lulu	16	20	1	15	7	8	7	53.3	0.53	2.2	7.8	3.79
Jax	16	20	5	11	7	5	6	45.5	1.73	2.45	3.36	2.07
Jarvan IV	16	20	8	8	6	2	6	25	1.38	4.13	6.5	1.91
Gragas	14	17.5	3	11	7	8	3	72.7	2	2.09	6.45	4.04
Gnar	14	18	5	9	5	4	5	44.4	1.56	3.44	3.44	1.45
Taliyah	14	18	7	7	4	3	4	43	2.14	3	5.57	2.57
Lissandra	14	17.5	9	5	4	3	2	60	1.6	2.8	7.8	3.36
Kai'Sa	13	16.3	5	8	6	5	3	62.5	5.88	1.38	4	7.18
Thresh	13	16.3	6	7	4	4	3	57	1.14	2.86	9	3.55
Ashe	13	16.3	7	6	3	4	2	66.7	3.17	2.17	6.67	4.54
Draven	13	16.3	7	6	5	3	3	50	4.33	3.33	1.83	1.85

Interaction between Table and Line Graph

Detailed in the screenshot below, we have included interaction between the table and the line graph. By default, the line graph will display the first ten champions by name in alphabetical order. However, when selecting certain champions inside the table, they are highlighted in a certain color. The same highlight color appears

as a line on the line graph next to it. We intend to experiment with better coloring for the table and line graph.



2022	Champion	Pick/Ban	Pick/Ban Rate	Picks	Bans	Players	Wins	Losses	Winrate	Kills	Deaths	Assists	KDA
	Aatrox	80	100	58	22	12	12	10	55	2.73	2.59	4.59	2.82
	Yuumi	76	95	68	8	6	8	0	100	1.63	1.13	13.88	13.78
	Sejuani	72	90	42	30	17	16	14	53.3	2.03	2.6	7.27	3.58
	Caitlyn	70	87.5	62	8	5	5	3	62.5	3.25	1.88	4.25	4
	Azir	66	82.5	29	37	13	19	18	51	2.95	1.81	4.3	4
	Sylas	60	75	28	32	10	21	11	65.6	3.56	2.16	4.19	3.59
	Graves	59	73.8	34	25	11	12	13	48	3.16	2.36	5.96	3.86
	Maokai	56	70	35	21	14	7	14	33	1.19	3.14	5.62	2.17
	Lucian	55	68.8	29	26	9	15	11	57.7	5.12	1.58	4.58	6.15
	Viego	48	60	18	30	9	19	11	63.3	3.57	2.13	5.6	4.3
	Akali	48	60	24	24	11	12	12	50	3.67	2.17	3.33	3.23
	Fiora	42	52.5	23	19	10	9	10	47.4	2.63	2	3.16	2.89
	Renata Glasc	41	51.2	20	21	10	13	8	62	1	2.38	9	4.2
	Aphelios	39	48.8	8	31	10	17	14	54.8	3.97	2.06	4.48	4.09
	Renekton	39	48.8	23	16	9	5	11	31.3	2.44	3.44	3.94	1.85
	Viktor	35	44	16	19	10	6	13	31.6	2.26	3.37	4.53	2.02
	Kalista	31	38.8	13	18	10	7	11	38.9	3.44	2.72	4.22	2.82
	Ornn	31	38.8	15	16	11	9	7	56	1.31	2.19	5.75	3.23
	Nami	27	33.8	1	26	9	15	11	58	0.85	2.27	9.85	4.71
	Varus	27	33.8	8	19	10	9	10	47.4	3.21	2.16	4.42	3.54
	Lee Sin	23	28.7	8	15	9	7	8	47	1.47	2.87	6.13	2.65
	Poppy	22	27.5	9	13	9	4	9	30.8	1	2.23	5.77	3.03
	LeBlanc	22	27.5	14	8	4	3	5	37.5	3.5	1	4.25	7.75
	Camille	21	26.3	10	11	5	8	3	73	3.64	2.64	5.91	3.62
	Heimerdinger	21	26.3	13	8	3	4	4	50	0.88	3.38	6.25	2.11
	Vi	18	23	6	12	8	6	6	50	1.67	3.08	7.58	3
	Nautilus	18	23	11	7	6	1	6	14.3	0.71	4.43	5	1.29
	Lulu	16	20	1	15	7	8	7	53.3	0.53	2.2	7.8	3.79
	Jax	16	20	5	11	7	5	6	45.5	1.73	2.45	3.36	2.07
	Jarvan IV	16	20	8	8	6	2	6	25	1.38	4.13	6.5	1.91
	Gragas	14	17.5	3	11	7	8	3	72.7	2	2.09	6.45	4.04
	Gnar	14	18	5	9	5	4	5	44.4	1.56	3.44	3.44	1.45
	Taliyah	14	18	7	7	4	3	4	43	2.14	3	5.57	2.57
	Lissandra	14	17.5	9	5	4	3	2	60	1.6	2.8	7.8	3.36
	Kai'Sa	13	16.3	5	8	6	5	3	62.5	5.88	1.38	4	7.18
	Thresh	13	16.3	6	7	4	4	3	57	1.14	2.86	9	3.55
	Ashe	13	16.3	7	6	3	4	2	66.7	3.17	2.17	6.67	4.54
	Draven	13	16.3	7	6	5	3	3	50	4.33	3.33	1.83	1.85
	Soraka	12	15	3	9	6	4	5	44	0.33	2.78	7	2.64
	Leona	12	15	3	9	7	2	7	22	0.56	3.78	5.56	1.62
	Trundle	12	15	5	7	4	4	3	57.1	1.57	2	6.57	4.07
	Ryze	12	15	8	4	2	3	1	75	4.75	2.75	4.5	3.36
	Amumu	11	13.8	5	6	4	2	4	33	1	4.17	7.67	2.08
	Bel'Veth	11	13.8	5	6	5	1	5	16.7	1.67	3.33	3.67	1.6
	Sivir	10	12.5	2	8	6	7	1	87.5	5	1.25	8.5	10.8
	Karma	10	12.5	3	7	5	2	5	28.6	1	3.14	5.14	1.95
	Tristana	10	12.5	4	6	6	1	5	17	2	2.83	2.33	1.53
	Yone	10	12.5	5	5	2	5	0	100	5.4	2.2	5.2	4.82
	Miss Fortune	9	11.3	1	8	4	2	6	25	2.38	2.25	3.63	2.67
	Rell	9	11.3	3	6	5	4	2	67	0.33	3	9	3.11