Analysis of League of Legends Professional Games

Peter Zhong and Tony Song

Research Questions:

1. What is the most important neutral objective that contributes the most to win a game in league of legends? (Herald, Dragon, Elder Dragon, Baron)

    Elder Dragon > Herald > Baron > Dragon .

2. Who is the most consistent professional player?

    Most Consistent Player: PraY

    Most Consistent Veteran Player: Bang

    Honorable Mention: Uzi

3. Which region has the best all-star team?

    CN (China) > KR (Korea) > EUW (EU West) > NA > TW (Taiwan) > VN (Vietnam)

Motivation and Background:

League of Legends has long been one of the most prominent video game titles in the professional esports scene. It is loved both by the players and the fans for its long professional play history, deep and diverse strategies, and high-octane and action-filled live gameplay. As die-hard LOL esports fans ourselves, we aim to use this opportunity provided by the CSE 163 final project to get a statistical grasp beyond our current figurative understanding of the META (most effective tactics available) of League of Legends.

In the professional esports scene, having mechanically talented players can be advantageous, but every successful team in League of Legends history has a strong team of coaches and analysts. This group of people mostly work in the background, and their main responsibilities include actively going through all the game footages to provide players with information about their opponents' performances as well as their own, making predictions about possible strategies and drafts that will be used by opponents in future games, and performing

statistical analysis to quantitatively come up with effective strategies that often has deep and profound impact on the progression of a game.

On one hand, the first two of our research questions are very similar to what coaches and analysts often need to answer for their professional teams. The first question is very useful because in a professional level LOL game where micro individual mechanical skills are no longer the most significant factor in creating the gap between the winning and losing team, macro maneuvers and neutral objectives are often what gives the winning team the edge to victory. By analyzing this topic, we can decide what objectives contribute the most to winning a high-level LOL game. Consistency is the second question we decide to analyze because it is a very important factor in determining the success of a player, especially when the players are competing at a constantly high level. By analyzing this question, we can see the average impact of the player on the game and also the disparity of the player's contributions between different games.

On the other hand, the third question is closely related to a lot of esports fans' interests as well as our own. Every region has great players for every different position, but they are not always on the same team. It is every esports fan's dream to be able to build an all-star team from the best players in each region. By analyzing this, we will know which region has the strongest all-stars team. Therefore, we believe that these questions are meaningful to our project, as they have real world applications and fit our interest and passion at the same time.

Dataset:

Example web pages that we get our data from:
https://gol.gg/tournament/tournament-stats/World%20Championship%202019/ (Single Tournament Page Example, all data we need are from "Last games" section at very bottom of that page, through those links we can get single game summary)
https://gol.gg/game/stats/20493/page-game/ (Single Game Summary Example, part of Research Question 1's data is in second time graph from "GOLD GRAPH & TIMELINE" section of that page)

https://gol.gg/game/stats/20493/page-fullstats/ (Single Game Players' Stats Example, part of Research Question 2's data are in the table of that page)
https://gol.gg/players/list/season-ALL/split-ALL/tournament-World%20Championship%202019/position-ALL/week-ALL/ (Single Tournament All Players' Stats Example, part of Research Question 3's data are in the table of that page)
https://gol.gg/teams/team-stats/642/split-ALL/tournament-World%20Championship%202019/ (Single Tournament All Teams' Stats Example, part of Research Question 3's data about each team's region are in the table of that page)

Methodology:

1st research question:

Go to the Tournament section of the web page and scrape all the data about the neutral objective each side got and which side won the game from every match from 2018 to 2019 MSI and World Championship games through using the Beautiful Soup. Put all data we get into a single table and compute how much each objective contributes to winning a game.

For every match, use the winning side's numbers of each neutral objective obtained minus the losing side's numbers of each neutral objective obtained. Then sum up all the differences and use the sum to divide the total number of matches to get the relative value each neutral object contributes to win the game. Finally, compare the relative value for each natural object to determine which one is the most important.

2nd research question:

Go to the Tournament section of the web page and scrape all the data about players' stats from every match from 2016 to 2019 MSI and World Championship games through using the Beautiful Soup.

As we further investigated the data, we made some changes to our original plan of filtering out inconsistent players. Firstly, we decided that variance should only be used to eliminate outliers. It should only eliminate highly inconsistent players rather than only leaving

highly consistent players, since the number of matches played significantly reflects more about a player's consistency than the variance of the player's damage per gold. Secondly, we decided that the number of matches played should be our stronger 'gatekeeper' for the reasons mentioned above. After careful consideration of our data, we decided to set the number of game played thresholds at 30 games and variance at 0.2. Additionally, in order to honor the veterans of League of Legends esports, we also decided to conduct another contest where the number of game played thresholds are set at 60 games and no threshold on variance is set.

3rd research question:

Go to the Players section of the web page and scrape all the data from 2016 to 2019 MSI and World Championship games through using the Beautiful Soup. Put all data we get into a single table, remove all the players with 15 or less games, and compute the best player for each game position for each region according to their KDA and win rate using the PT-scoring system named after the creator of this project Peter Zhong and Tony Song. The PT-score of the player is the sum of its KDA ratio and its win rate (0-1). Since win rates of players are normally around 0.5-0.6 and KDA is normally around 5-6, we believe that the PT-scoring system accurately captures the weight of KDA ratio and win rate, since individual performance is normally more important that win rate when choosing an all-star team. Lastly, we compute which region has the best team according to the cumulative PT-score of that region's All-Star team to determine which region has the best team.

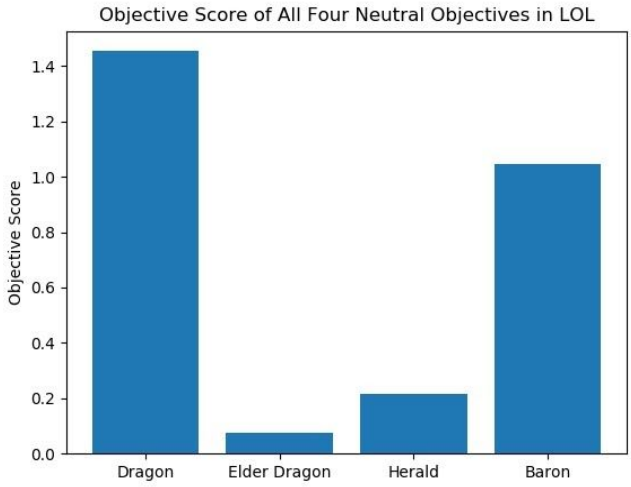
Visualization:

All the visualizations shown below in this report are produced using matplotlib.pyplot. All the barcharts are produced with ax.bar, while the performance curve charts are produced with ax.plot.

Results:

Question 1

Dragon received an objective score of 1.454, while baron received 1.046, herald received 0.217 and elder dragon received 0.075.



Objective score calculates the number of objectives needed to win the game. Therefore, the lower the score, the more impact obtaining one of the objectives have on the progression of the game. Every single objective received a positive score. Therefore, it's no surprise that having additional neutral objectives win games.

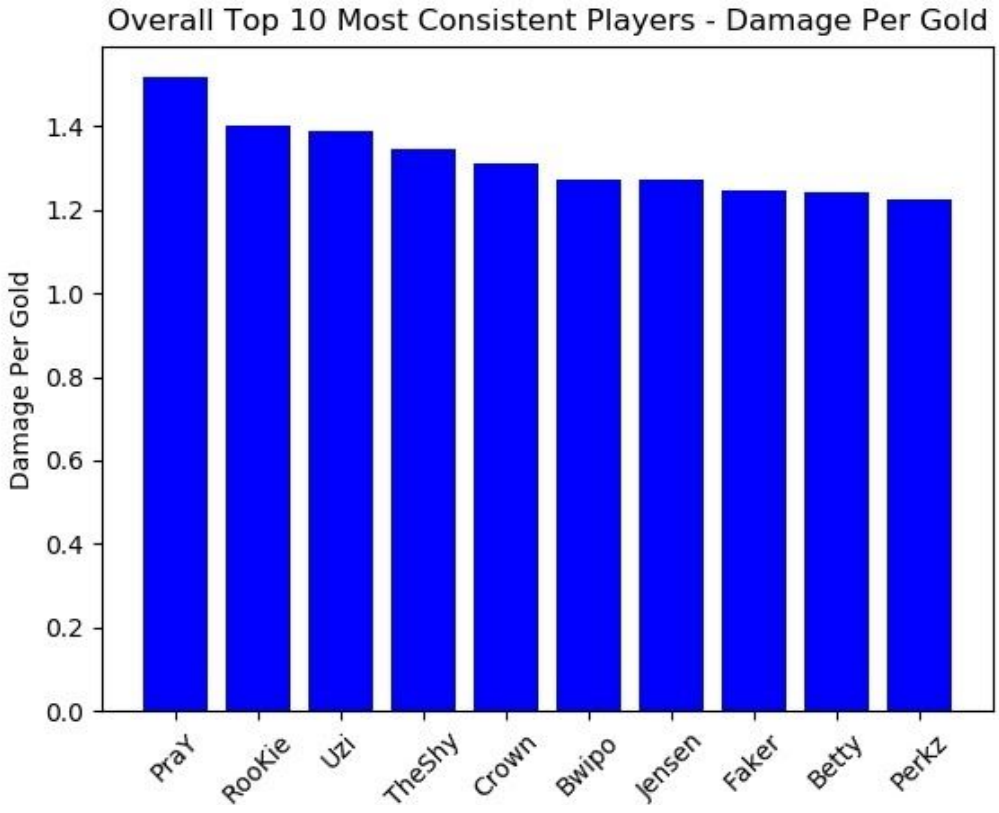
Elder dragons are the most valuable neutral objective, since its execution effect gives an enormous advantage to the team with the buff in late game team fights. The ability to immediately execute enemy champions that are below a certain health threshold is often what gives a team the edge to winning a teamfight and eventually the game.

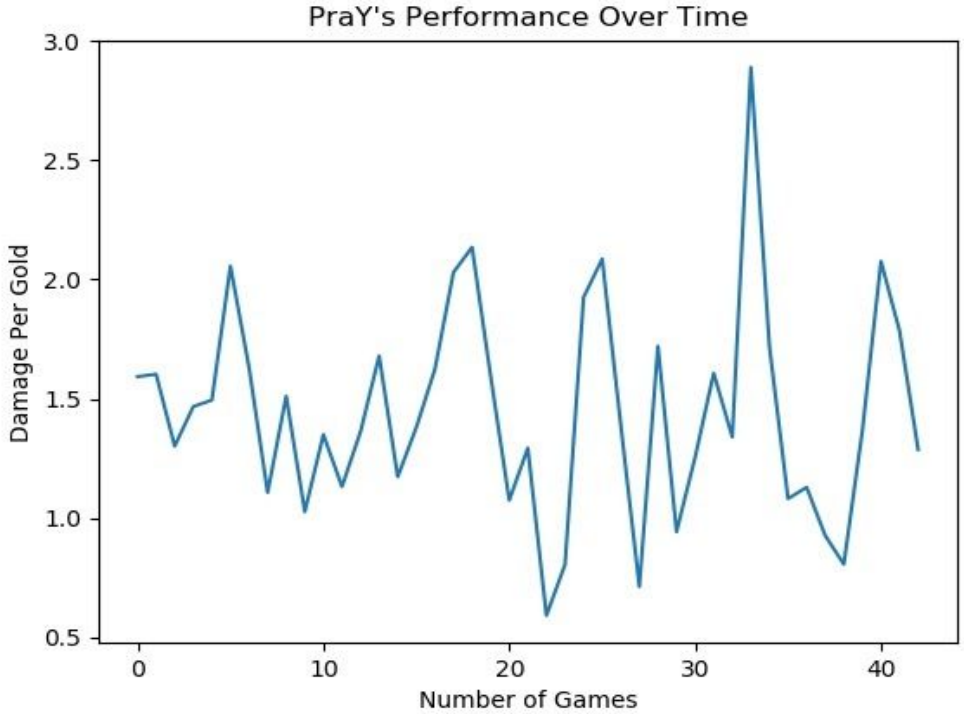
It's a bit surprising, however, to see herald getting such a low score. We conclude that it is because there are only 2 heralds spawning each game, and ganks and dragons are often prioritized in early games over herald. The low number of heralds spawned and captured in LOL reduced its presence overall.

It's also no surprise that Baron turns out to be more valuable than dragons, since Baron buff makes macro maneuvers super effective. With proper utilization (which is the standard for pro plays), the baron buff can crash waves of strong minions into the enemy base and eventually overwhelm the defense.

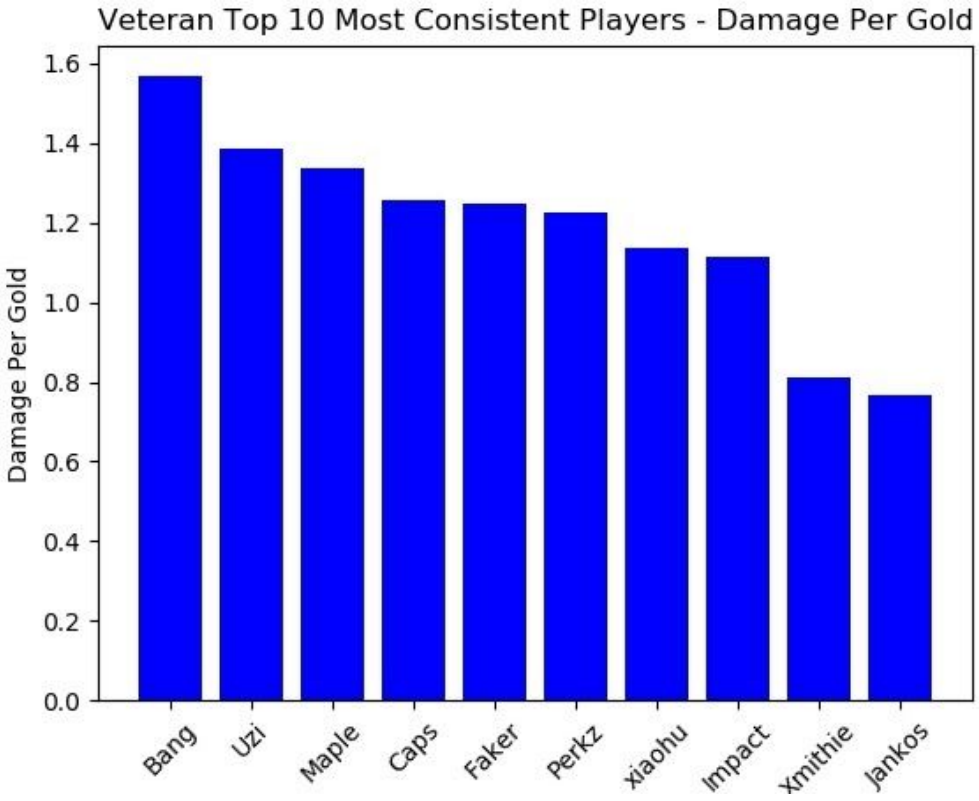
Question 2

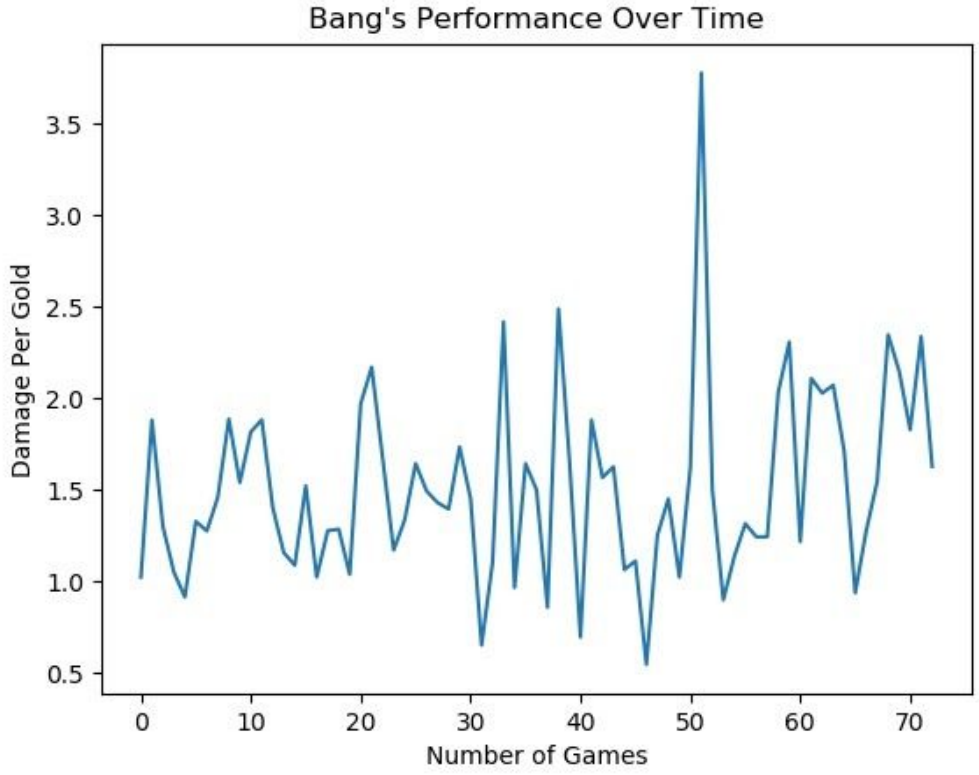
Most Consistent Player: PraY



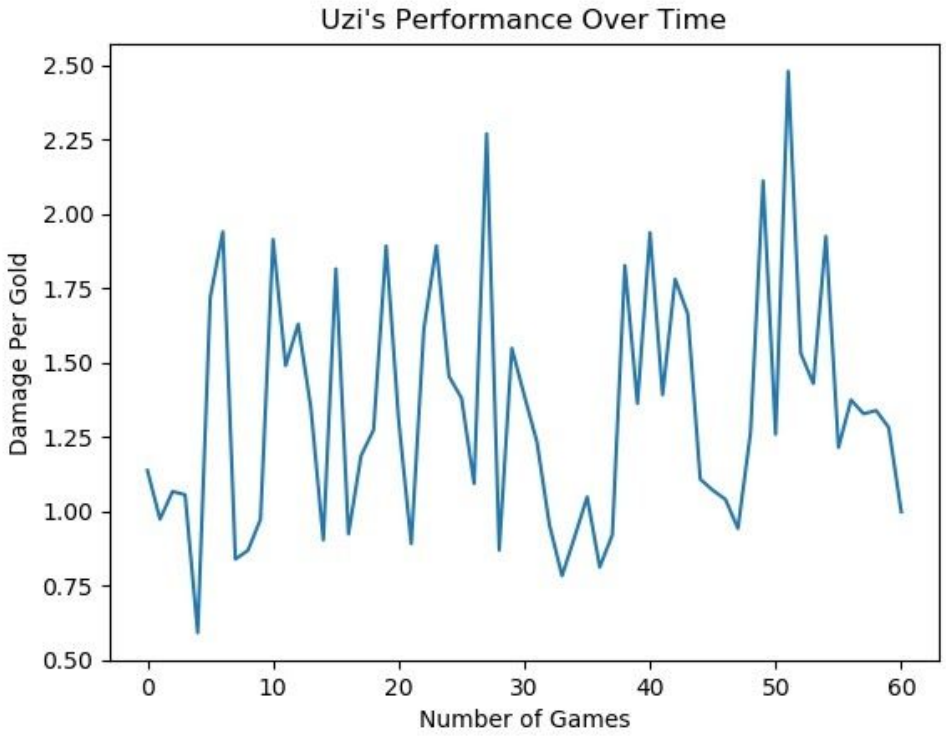


Most Consistent Veteran: Bang





Honorable Mention: Uzi



PraY was the former ADC for the KT roster, who had a terrible regular season in the year of 2019, causing the entire roster and coaching line-up to depart from the organization. PraY retired from competitive play after coming out of his previously announced retirement to play for KT. Despite the recent downfall, his earlier consistent performance in international tournaments with ROX, Longzhu and DragonX made him stand out amongst all other players.

Bang is currently the ADC for Evil Geniuses. He earned his 'Most Consistent Veteran Player' title mainly with his outstanding performance with SKT when they were the unbeatable world champion.

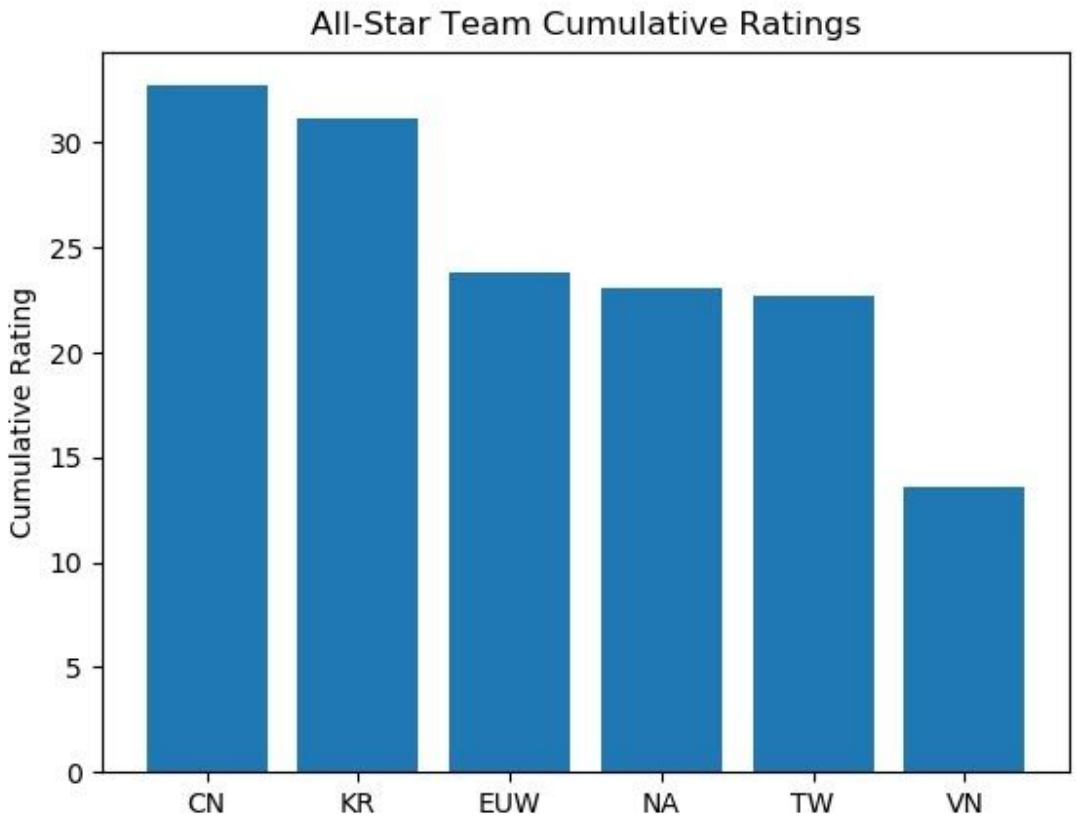
Since Uzi appeared in 3rd place for overall consistency and 2nd place for veteran consistency, we felt like he deserved an honorable mention. He was the ADC for the prestigious Chinese team Royal Never Gives Up, and his short departure from the team followed by the terrible performance of the team without him in the last LPL spring split showed how he has truly been the backbone of this dominating LOL professional squad. Uzi officially announced his retirement from competitive play on the evening of June 2, 2020 due to persisting personal health issues.

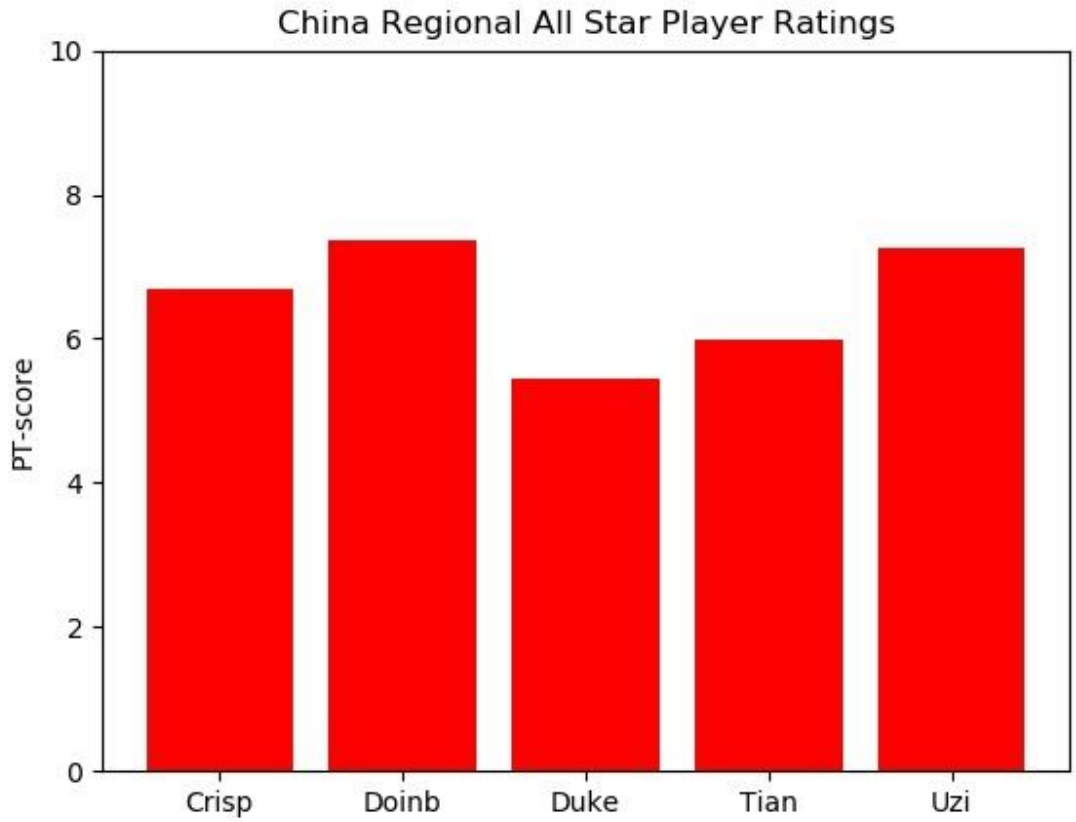
Question 3

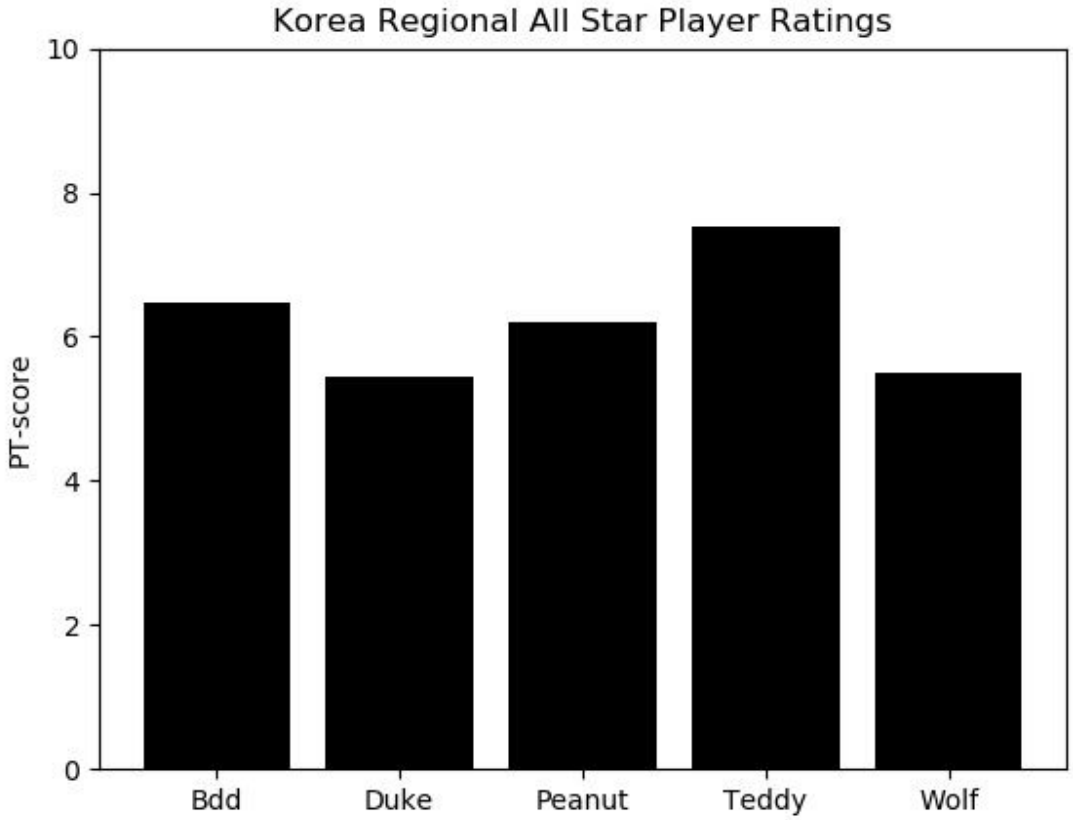
China received a cumulative PT-score of 32.715, while Korea received 31.116, EU West received 23.768, North America received 23.029, Taiwan received 22.722, and Vietnam received 13.616.

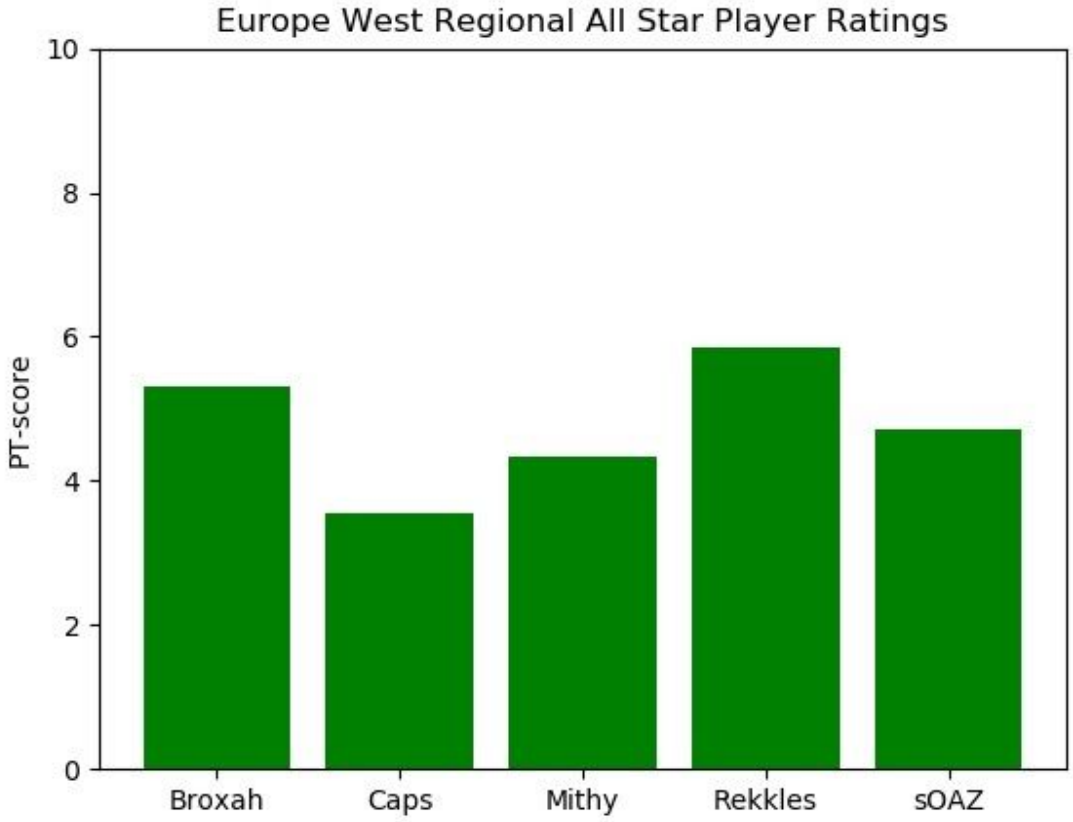
This is not surprising, since China has been crowned world champion for both the year 2018 and the year 2019. Korea has been very strong historically and had a long history of domination. EU was not particularly strong until recently when G2 esports made it to the 2nd place in the 2019 world championship tournament. NA is known to be a mediocre region, since the performance of NA teams at international tournaments is often subpar. TW and VN are really small regions and are normally unheard of.

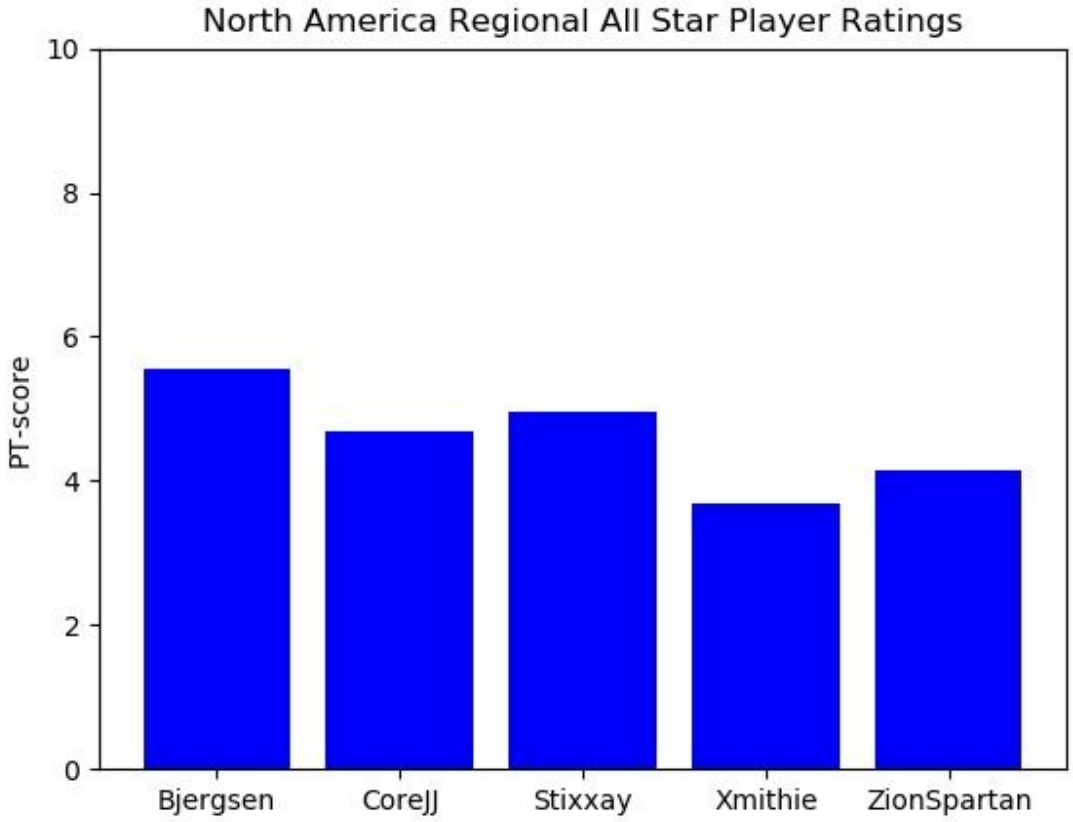
Charts below shows the cumulative PT-score for all-star teams in each region, as well as the PT-score for each individual player from each region.

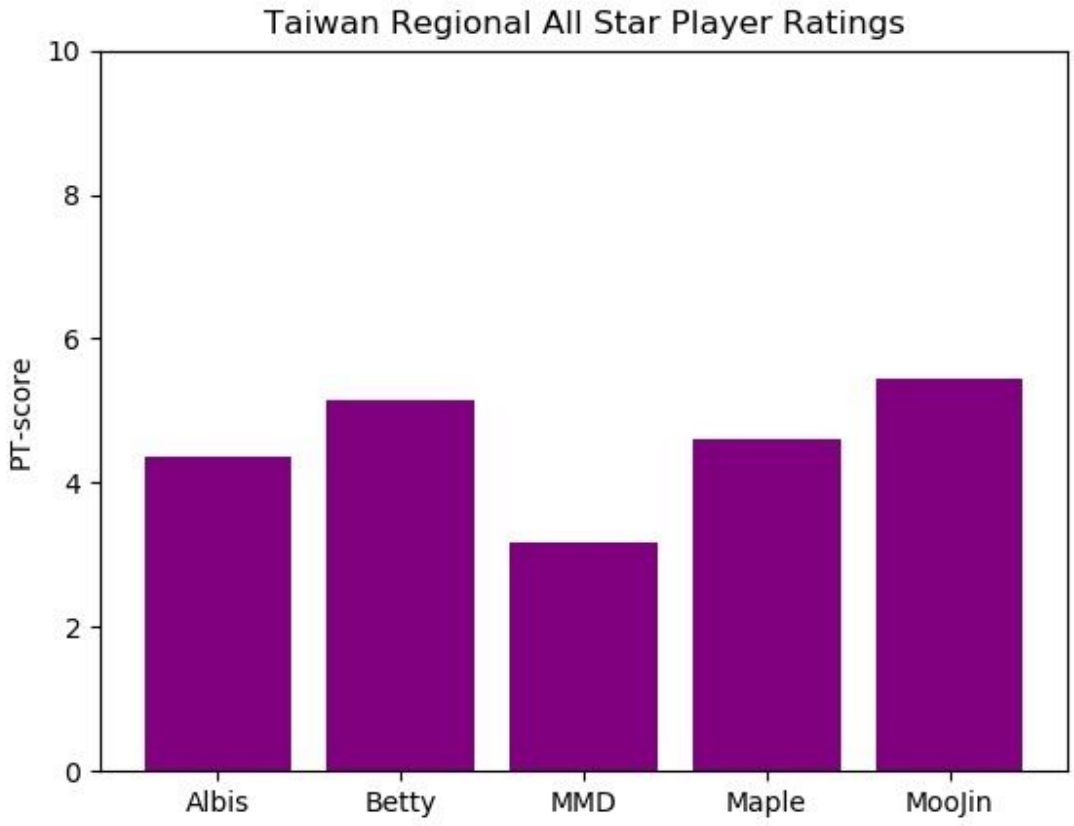


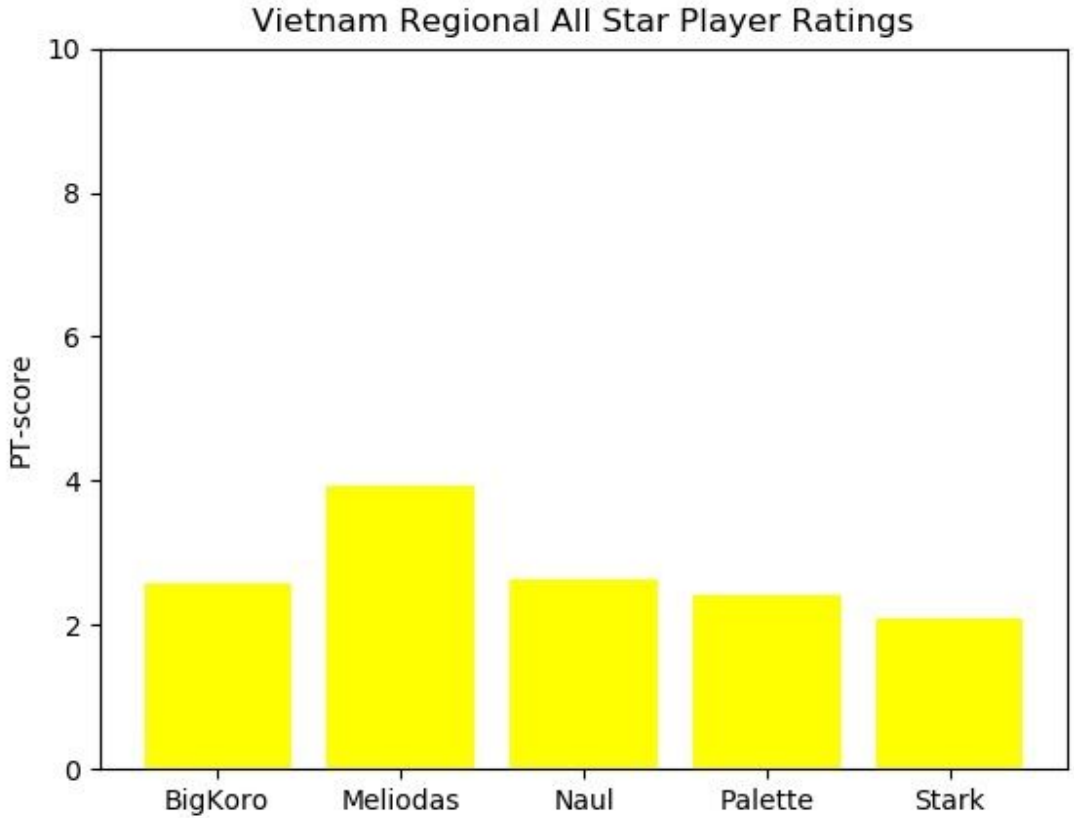

China Regional All Star Player Ratings
PT-score
10
8
6
4
2
0
Crisp
Doinb
Duke
Tian
Uzi


Korea Regional All Star Player Ratings
PT-score
10
8
6
4
2
0
Bdd
Duke
Peanut
Teddy
Wolf


Europe West Regional All Star Player Ratings
PT-score
10
8
6
4
2
0
Broxah
Caps
Mithy
Rekkles
sOAZ


North America Regional All Star Player Ratings
PT-score
10
8
6
4
2
0
Bjergsen
CoreJJ
Stixxay
Xmithie
ZionSpartan


Taiwan Regional All Star Player Ratings
PT-score
10
8
6
4
2
0
Albis
Betty
MMD
Maple
MooJin


Vietnam Regional All Star Player Ratings
PT-score
10
8
6
4
2
0
BigKoro
Meliodas
Naul
Palette
Stark

Challenge Goals:

Messy Data: Since there are very few usable and clean datasets about League of Legends professional games and they only provide very limited data, we decide to scrape the data directly from several esports websites, which requires significant preprocessing before the data becomes usable for the investigation.

We think that we accomplished this goal since after the data processing, we successfully ended up with four very clean csv files that are used by the algorithm file to compile and compute all the results shown above.

New Library: The nature of the goal of our investigation requires us to inspect web page scripts and collect data from there. Therefore, we decided to learn the library named Beautiful Soup to help us accomplish that goal.

We think that we accomplished this goal since we managed to precisely acquire the portion of the data that we want to study from websites with Beautiful Soup. And we also used some other support libraries like request and csv during the process of getting data from url and saving data into csv files.

Work Plan Evaluation:

We ended up dividing the entire projects into two portions: the data collection and the algorithms application, both of which turned out to be around 300 lines of python code in length. Tony mainly worked on data collection, while Peter worked more extensively on algorithms application and report-writing. Through Peter actively communicating with Tony about what data is needed for the calculations and Tony providing timely feedback, the collaboration went very smoothly and the overall workload was evenly divided among the two members of the team. And at last, we went through each other's code to check whether the comments and code itself are easy enough to understand.

The time consumption of each part is underestimated, since we significantly underestimated the complexity of directly scraping data from the web along with very

complicated processing required to complete the project. Both the data collection and the algorithm application took about 8-9 hours to complete (about twice of what we estimated them to be), while the report writing is easier, taking up about 2 hours.

Testing:

We made some very minor changes to our main file named algorithm.py in order to make it more easily testable. These changes include implementing returns to make testing easier and removing all the plotting functionalities that are unnecessary during the testing. No changes are done to the main structure of the application of algorithms.

For the first two functions (objective_score and _consistent_player), we used the first 10 rows of the original data to run through the program while we calculated the correct results by ourselves on a calculator and used our result to compare with the result obtained by the function. For the last function(all_star), it's quite hard to get a correct portion of our original data for it to function properly. Therefore, we handcrafted two new data sets with six regions named A, B, C, D, E and F, each containing 10 players (2 from each of the 5 positions). We exaggerated the data so that the return is obvious, and used that result to test our function.

All the testings are implemented in the file named testing.py.

Collaboration:

We received no additional assistance from people besides the course staff and our group mates. We used several online resources. For using Beautifulsoup library, we went through the documentation online but found it was still very hard to understand since that documentation needs some prior knowledge on web programming. Therefore, we searched for some examples about how to use Beautifulsoup to process the data from a url link and put data into the csv.

Unfortunately, we didn't use the information from the 6/1 lecture about web scraping since we have already finished writing all the code at the time of the lecture. Moreover, when we used pandas to process our data, we found that we needed some functions that we didn't learn so far. Therefore, we went through the panda documentation to get some useful functions like .var() (to calculate the variance of player performance) and .map() (to remove redundant information in the original csv file).