Hannah Jauris
IMGD2905
21 April 2018

Project 4: *League of Legends* and Puckhunt Analyses

Section 1: Game Length and Frequency

For Sections 1 and 2, data regarding League of Legends was obtained from the Riot database using the Riot Developer API through RiotWatcher and Python scripts. This data was placed into CSV files and analyzed using Excel. The player followed in the two sections has the Summoner name Aphromoo, and was chosen randomly from a list of competitive League of Legends players found via a Google search. His first 100 games were used in the following analyses. Figure 1 examines Aphromoo's games played to determine if there were any trends in how often these games were played. The data was obtained by extracting from each game the Timestamp attribute, which marked how much time had passed since the first game was played. It was ordered in increasing order, with cumulative percentages of each point calculated in Excel and then graphed.

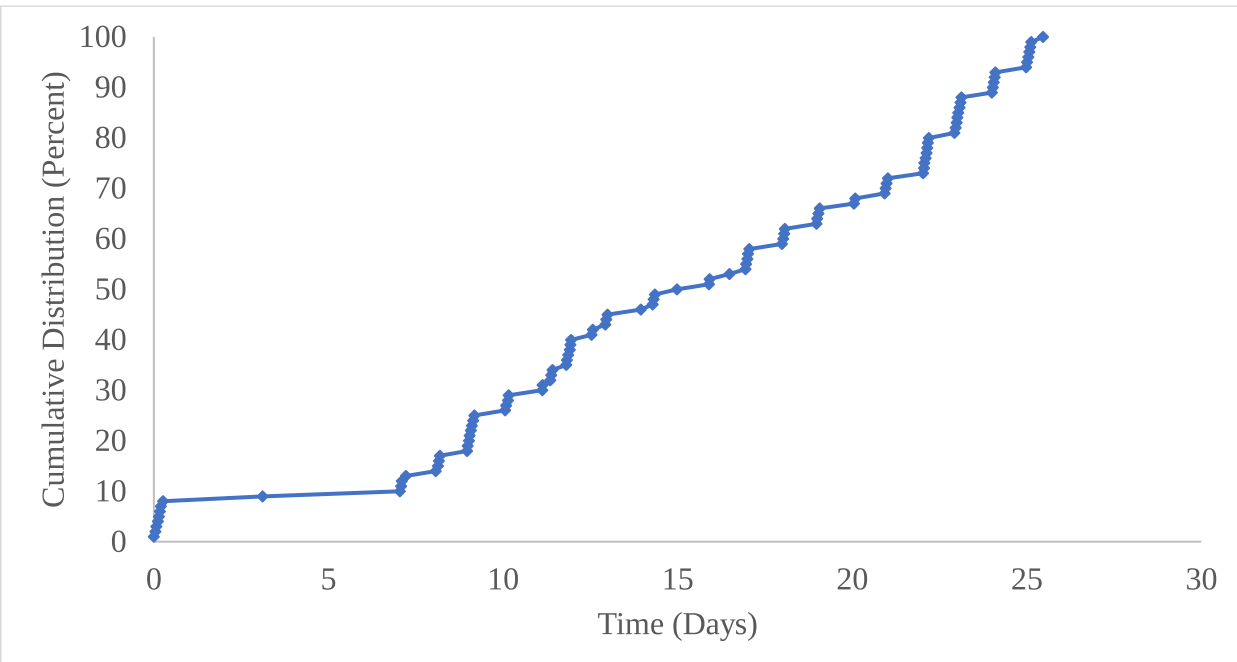

Figure 1: Percentage of games played over time

Figure 1 displays the cumulative percentage of how many games, of the 100 analyzed, were played over time. Sharp increases in the graph indicate that many games were played at around the same time, as the percentage of the games played markedly increased over a small amount of time. Near-horizontal sections indicate times when Aphromoo was not playing, as the percentage of games played is not increasing much during that interval. Figure 1 shows that the first eight games were all played within a day of each other, although there is a significant gap in games played until seven days had passed from the initial game being played, with only one game being played in that interval. After approximately seven days, the time between games decreases significantly, indicating that Aphromoo began to play games more often. There are several other, although smaller, near-horizontal sections of the graph after day seven, and most are a similar length of approximately 24 hours, meaning that Aphromoo seemed to play League of Legends multiple times on a near-daily basis at approximately the same time each day.

Figure 2 investigates the spread of the duration of games played, to determine how long a typical game lasts and if there are visible trends in the distribution. The data was obtained by following the methodology above, although extracting the Game Duration attribute of each match, then converting it from seconds to minutes to be more easily understood, then using Excel to plot the graph.

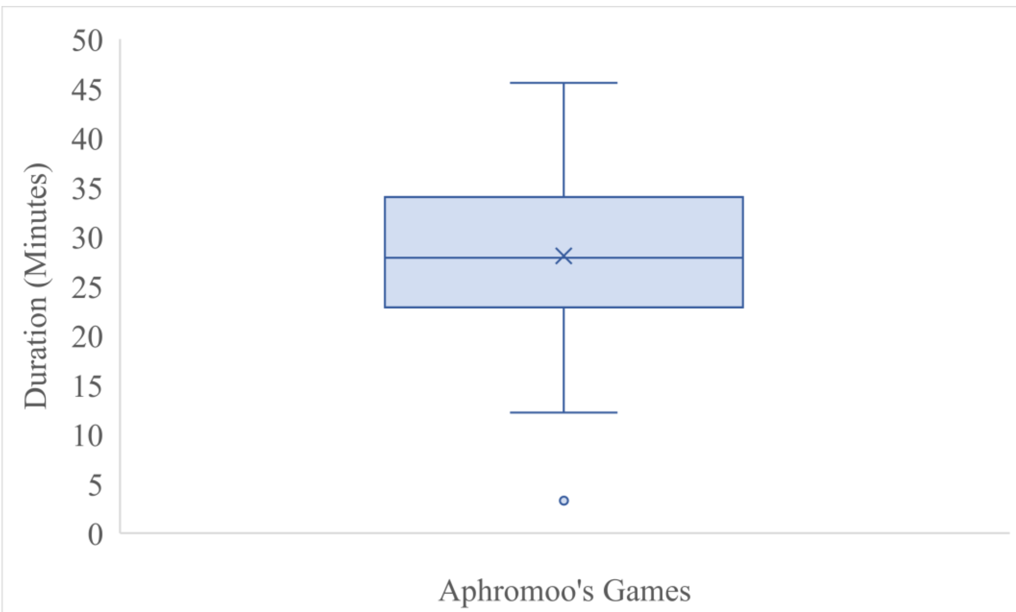

Figure 2: Durations of Aphromoo's first 100 games

The center line of the boxplot indicates the median of the data, with the X marking the mean. The bottom and top edges of the box indicate the first and third quartiles, respectively, and the edges of the whiskers indicate the maximum and minimum values, with the exception of any outliers, shown as individual points. The data has a median of 27.86 minutes, with a semi-interquartile range of 5.58 minutes. The data has a normal distribution, with the median and the mean almost the same values, indicating almost no skew in the data. There are two outliers, both at approximately 3.3 minutes. The distribution of times suggests that, for the most part, the lengths of games are relatively predictable, with many ranging between 22.84 and 33.99 minutes long, and most other games falling within 12.18 and 45.55 minutes.

Figures 3 and 4 explore how many times Aphromoo used various Champions, as well as which Champions he used most, to determine if different Champions were played evenly or if any were favored over others. In each of the 100 games analyzed, the player ID belonging to Aphromoo was found and used to identify which Champion he used in each game, by comparing the Champion ID he used to a static list of Champions to determine the name of the one selected. This was then sorted and frequencies of each Champion were calculated. Over the 100 games, Aphromoo used 25 different Champions. A histogram was created to indicate the frequencies of how many times Aphromoo used a Champion for Figure 3. Since there was not a wide variety of times a Champion was played, bin sizes of one were used for Figure 3's histogram. Figure 4 shows a bar graph of Aphromoo's most-used Champions and orders them by most-used. The Champions used less than five times were not included on the graph to avoid clutter.

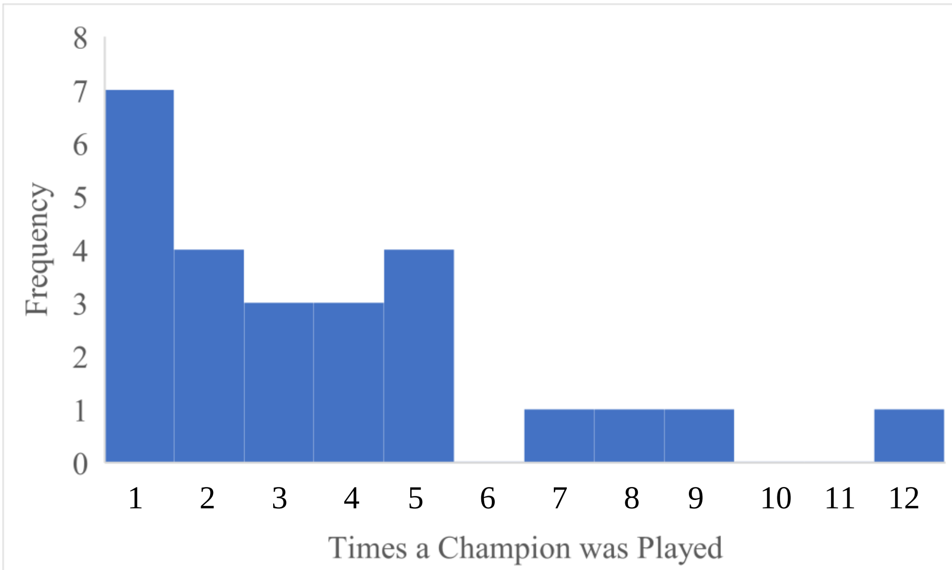

Figure 3: Frequencies of how often Champions were played a certain number of times in Aphromoo's first 100 battles.

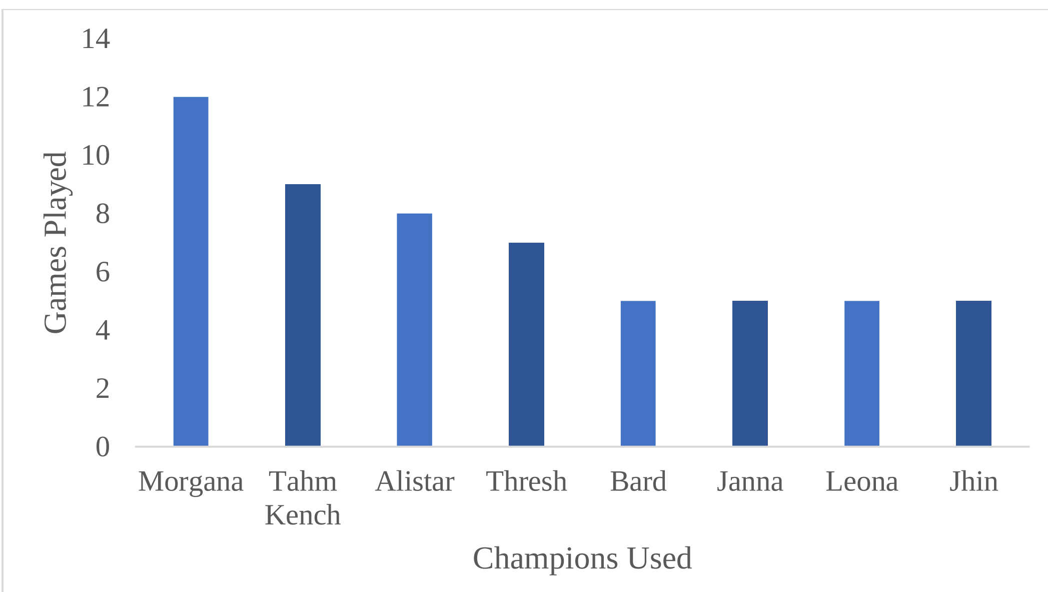

Figure 4: Most frequently played Champions in Aphromoo's first 100 battles.

Figure 3 shows the frequencies of how often Champions were used, with higher bars indicating a higher number of Champions that were used the number of times indicated by that column's bin value. Figure 4 displays the top eight most-often used Champions, with higher bars indicating that that Champion was used more times. As can be seen in Figure 3, Aphromoo played many different Champions for only a small amount of battles, and very few Champions often. The distribution is skewed to the right, with 13 of the 25 Champions played only being used 3 or fewer times, less than the average number of times if Aphromoo had played all of the Champions evenly, and only four heroes being played more than five times. Figure 4 shows Morgana is the most often used Champion, with 12 uses, followed by Tahm Kench, Alistar, and Thresh at 9, 8, and 7 uses respectively. During these matches, Aphromoo seems to have been experimental and willing to try many different Champions, only to discard them and try out other characters, although it seems that certain Champions, such as Morgana and Tahm Kench, were attractive to Alphamoo as he returned to them many times over the sample games analyzed.

Section 2: Gold

Section 2 continues to use the data from the games in Section 1, with the same method of collecting and analyzing data using RiotWatcher and Python scripts as well as Excel. Figure 5 explores how much total gold, obtained by all players, is gathered per game, to determine typical amounts of total gold gathered per game and to discover any trends in the distribution. The data was obtained by extracting how much gold each individual player obtained each match, then adding all players of a single game to get the total for that game.

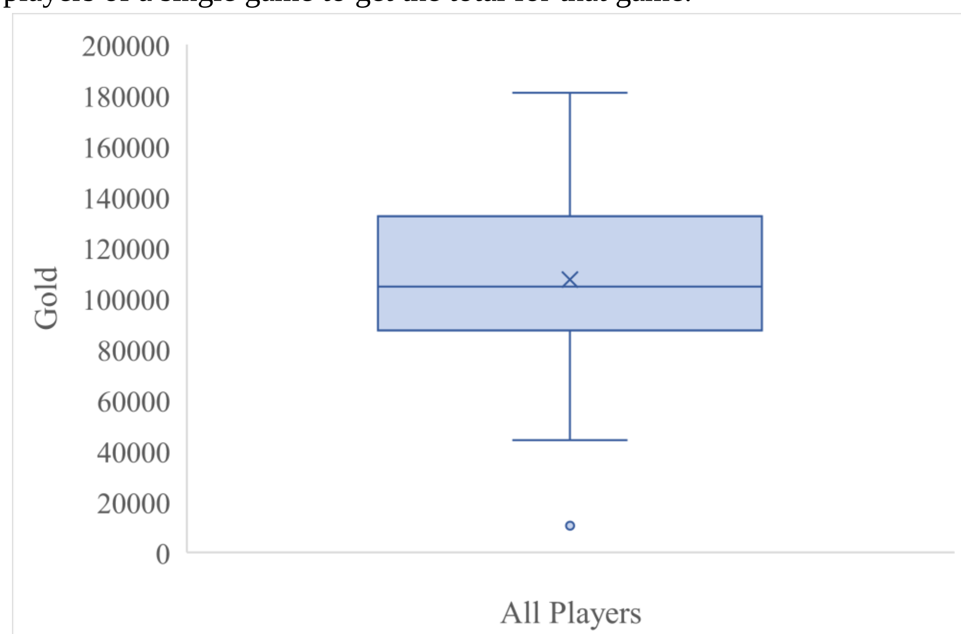

Figure 5: Total gold received by all players combined per game

Figure 5 is a boxplot like Figure 2, and although it deals with gold obtained each game as opposed to duration of games, it can be read in the same way. The total gold obtained follows a nearly normal distribution, with the median gold obtained across all players per match is 104,869 gold, and the semi-interquartile range is 22,494 gold. The mean, 107,611.39 gold per match, is close to the median, indicating that there is little skew in the data. The small amount of skew towards higher numbers does indicate that players are slightly more likely to accrue relatively lower amounts of total gold rather than higher amounts. There is one low outlier at only 10,502 gold and for all players, which may possibly relate to one of the outliers from game duration in Figure 2.

Figure 6 compares the distributions of total gold gathered per game by the winning and losing teams of each game, to determine if there are any differences in how much gold is gathered by a team that wins or lose. This data was obtained using the same data about gold obtained per player per match for Figure 5, except instead of adding the gold obtained by all players per match, only the gold of the winners or losers for each match were added together.

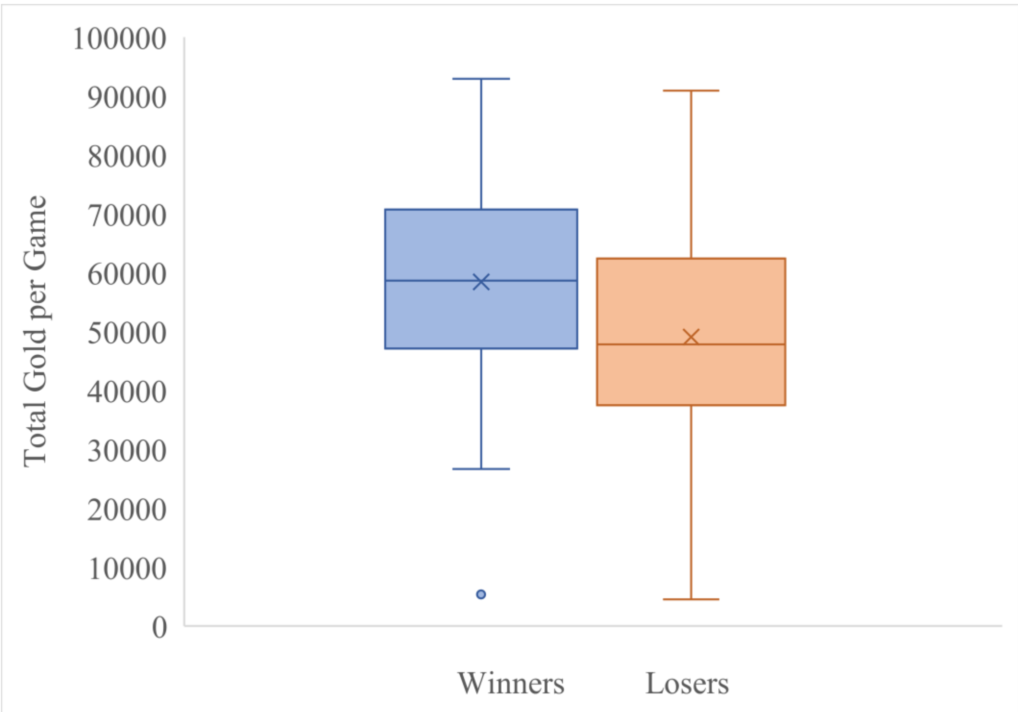

Figure 6: Boxplots comparing total gold won per game for winners and losers

As with Figure 5, Figure 6 also contains boxplots and can be read in the same way, with the two boxplots adjacent to each other for simpler comparison. The median of the losers is lower, at 47,794 gold per game, while the median of the winners is over 10,000 gold higher, at 58,653 gold per game. Further, the winners' gold is barely skewed to the left, with a mean just below the median, and is close to a normal distribution. The losers' gold, however, is clearly skewed to the right, with more players getting lower values of gold. While the semi-interquartile range is about within 1,000 gold for both the winners and losers (11,823 and 12,451 gold per game, respectively), the losers' gold has a wider variance outside of this middle 50%, with the whiskers extending farther in either direction than the winners' gold's distribution. While some losers can gain almost as much gold as the winners, losers will more often obtain less gold throughout the game.

Figure 7 compares the total gold gained per game across all players and the duration of the game, to determine if there is a relationship between a game's length and the amount of gold gathered. It was obtained by combining the data from Figure 2 regarding the durations of games and Figure 5, showing the total gold obtained per game. These values were matched by the game each data point came from to create the scatter plot shown in Figure 7, comparing total gold obtained per match to the duration of the game.

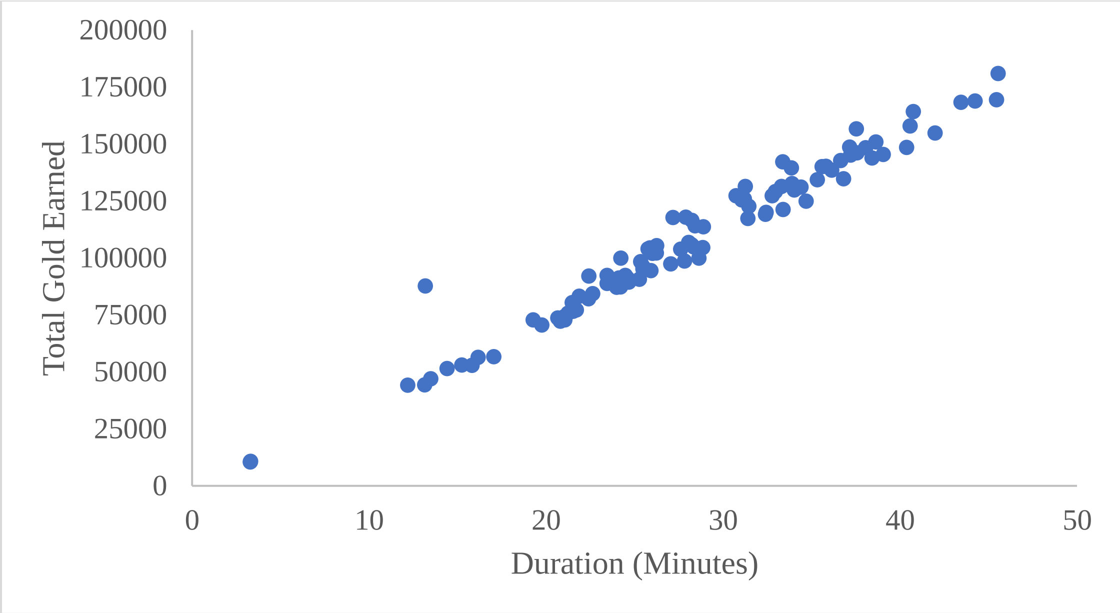

Figure 7: Total gold obtained versus duration of game.

Each point in Figure 7 represents a single game. The x-axis represents the duration, in minutes, a game lasts, and the y-axis shows the total amount of gold obtained by all players combined per game. The further a point is along the x-axis, the longer that game took to complete, and the high a point is on the y-axis, the more total gold was obtained in that game. As shown in Figure 7, there seems to be a relationship between the duration of the game and how much gold is obtained within the game. Most points follow a linearly increasing trend, with more gold being obtained the longer a game goes on. The exception to the trend is the point (13.18, 87,706), in which the players collectively gathered approximately twice as much gold as other games of similar lengths. For the most part, the longer a game is, the more gold players will be likely to obtain throughout it.

Section 3: Puckhunt

This section analyzes the data obtained from members of the IMGD2905 class who participated in an IQP studying affects of latency, in which players had to move the mouse from the center of the screen to click on a red "puck", with varying amounts of lag applied to the cursor in each round of the game. On certain rounds, the players were asked to rate the lag of the mouse, on a scale of 1 to 5, with 1 being low lag and 5 being high lag. The data was taken from the IMGD2905 website zip folder containing all of the class data as of April 17, 2018. There were 11 participants from the class, with each providing 10 ratings for each of the four different lag times, resulting in 110 data points for each lag amount and 440 total points. Figure 8 compares how the players rated the severity of the lag compared to the actual lag applied. The data from all students was taken from a CSV file and imported into Excel to graph and analyze.

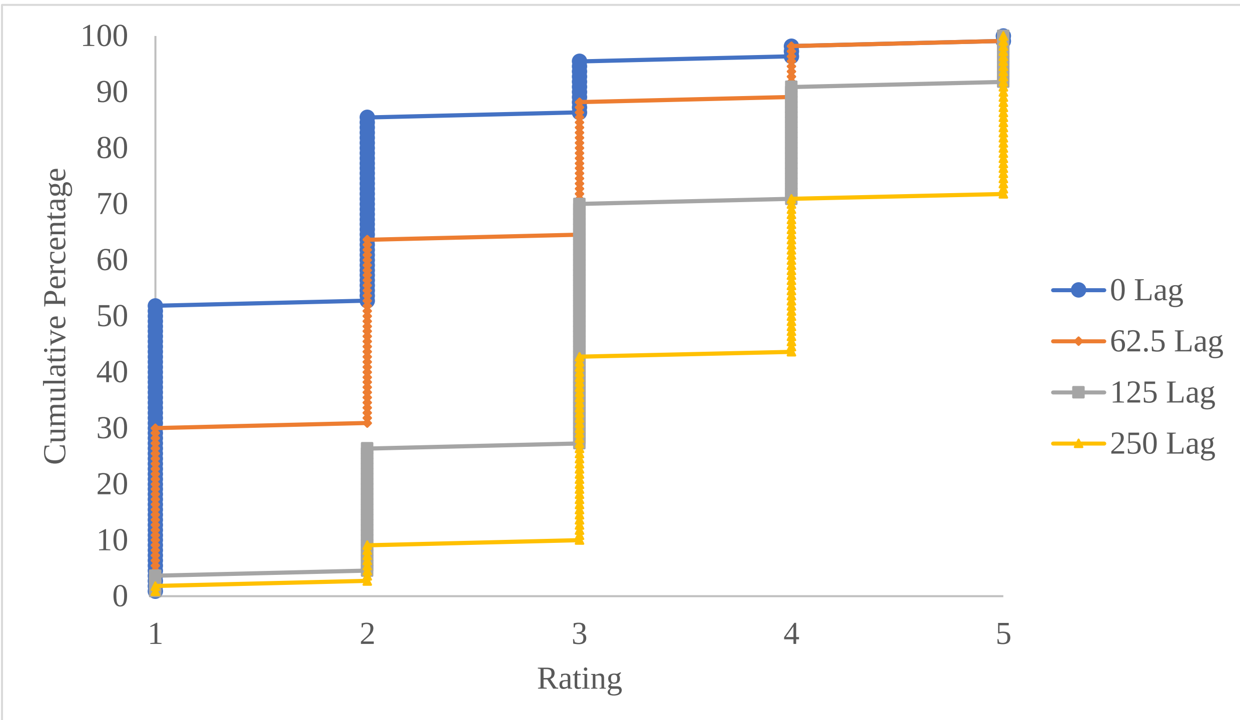

Figure 8: Cumulative Distributions of all player ratings of the four different lag times within the study.

Figure 8 shows the cumulative percentages for the ratings players gave to each of the four amounts of lag, 0, 62.5, 125, and 250 milliseconds. The lengths of the vertical portions of the lines over each rating indicate what percentage of rounds where that lag was given that rating. Figure 8 starts at one because that is the lowest rating a player can give, with lower ratings meaning less lag and greater ratings meaning more lag. The median ratings (at the 50% mark for each line) for each of the four lags mentioned above were 1, 2, 3, and 4, respectively. Over 50% of people thought that rounds with no lag had minimal lag, and it was only rated as high lag (5) two times out of the 110 times the lag was rated. The 62.5 millisecond lag had about a third of players give a rating of one, and another third gave it a lag of two. The remaining third was divided among the higher ratings, although over 80% of people gave the 62.5 millisecond lag a rating of three or lower. The middle lag, the 125 milliseconds lag, had a majority of people rate it at 3, which means that most people accurately labeled it as the middle amount of lag. The 250 millisecond lag had the most ratings at rating 3, although it also received the most ratings of 4 or 5 lag compared to other lags, with approximately 57% of rounds being rated as either 4 or 5.

Figure 9 shows the ratings for lags given by myself, hmjauris, to compare to Figure 8 to determine if there were differences in how I rated the different lags compared to all participants. Since the hmjauris data is from only one person, there are a total of 40 data points, with 10 data points for each lag amount. The data was obtained from the same zip folder as Figure 9's data, although a file that contained only the hmjauris data, as opposed to all data anonymously combined, was used instead to identify my own values.

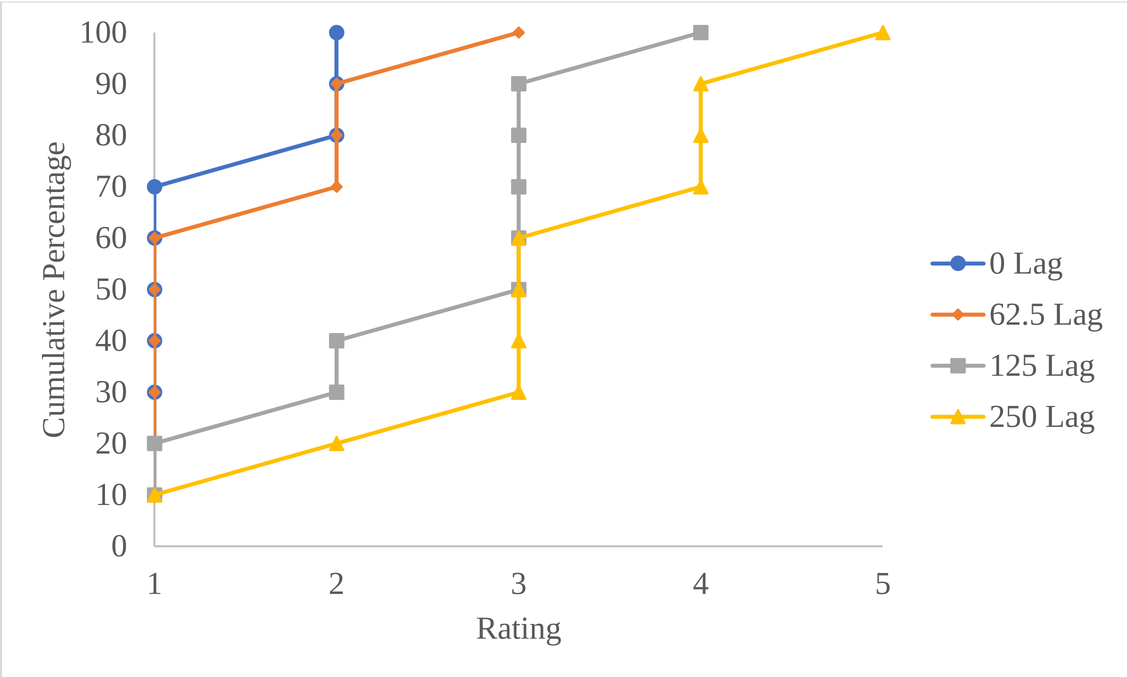

Figure 9: Cumulative Distributions of hmjauris user ratings of the four different lag times within the study.

Figure 9 is the same type of graph as Figure 8, and can be read in the same way. The median values for my ratings for the 0, 62.5, 125, and 250 millisecond lags were 1, 1, 3, and 3, respectively. While the 0 and 125 ratings had the same median rating as the overall scores in Figure 8, the 62.5 and 250 millisecond lags both had a rating one point lower than their overall medians. Compared to the overall ratings of all the players who participated in the study from the class, I rated the zero lag as the lowest level of lag 20% more than other players, and never rated it higher than two, whereas other players rated it as high as five. Similarly, for the 62.5 millisecond lag, only 40% of the ratings were higher than one, and only one rating reached 3. This is double the number of level one ratings for this lag than the overall ratings. The 125 millisecond lag received most of its ratings as 3 from both me and the overall sample, although I only rated it as high as a 4 a single time, whereas approximately 30% of the overall ratings rated the 125 millisecond lag as a 4 or higher. For the 250 millisecond lag, both my ratings and the overall ratings had approximately 30% of ratings at level 3, although where I had only 10% of my ratings for the highest lag as high as 5, 30% of the overall sample population gave the highest lag the highest rating of 5. With every speed of lag, I had a tendency to, overall, rate the lag as lower than the full sample did, regardless of if that lag was large or small.