

## Statistical Analysis Project - How do you decide “Who’s the Best”?

In order to choose the best shot put thrower to go to IASAS, the three candidates were asked to throw the shot put for a total of 22 times -- this data were then collected. The data found from each of these trials were then gathered and organized into a frequency table for each of the athletes. The frequency tables created were then used to create a frequency histogram that would serve as a visual representation for the incidence of the distance that the athlete was able to throw the shot put.

In order to conduct further analysis to find out who the best thrower was, the original data set was organized to find the measures of central tendency in order to have closer comparison between the candidates.

Also, in order to aid in visualizing the candidates the distribution of the data is displayed through a box-and-whisker plot and was found using the five number summary.

Raw data of Throwers 1, 2, and 3 (in meters)

Thrower 1	8.74	8.94	9.66	10.01	10.01	8.43	10.25	10.14	9.04	9.3	8.69
	8.85	9.25	9.46	10.23	8.95	9.65	8.79	10.62	9.78	9.26	9.39
Thrower 2	10.39	10.86	10.94	9	9.15	9.35	9.35	8.25	8.85	8.95	9.1
	10.2	9.53	8.76	8.03	8.96	9.25	9.98	10.82	10.1	8.96	9.68
Thrower 3	8.79	9.39	9.94	11.47	9.72	8.49	9.63	9.49	9.83	8.82	9.24
	9.13	9.56	9.94	9.75	9.12	8.96	8.83	9.25	9.38	9.62	9.98



Below, one would find 3 tables: a *frequency table* depicting the incidence that **Thrower 2** was able to throw the shot put for each of the intervals given, a table that includes *the measures of central tendencies* of the distance that the Thrower was able to throw the shot put, and a table that depicts *the 5-number summary* of the distance of the shot put.

Below one would also find two visual representations of Thrower 2's 22 trials, by way of a *frequency histogram* and a *box-and-whisker plot*. All of this data is displayed in order to represent the performance of Thrower 2 in order to have a better analysis on who would be best to go to IASAS.

Table 7: Frequency Table depicting the frequency of the distance (in meters) that Thrower 2 throws the shot put for a total of 22 trials.

Distance of the shot put (m)	Frequency
$8 \leq x < 8.5$	2
$8.5 \leq x < 9$	5
$9 \leq x < 9.5$	5
$9.5 \leq x < 10$	4
$10 \leq x < 10.5$	3
$10.5 \leq x < 11$	3
$11 \leq x < 11.5$	0

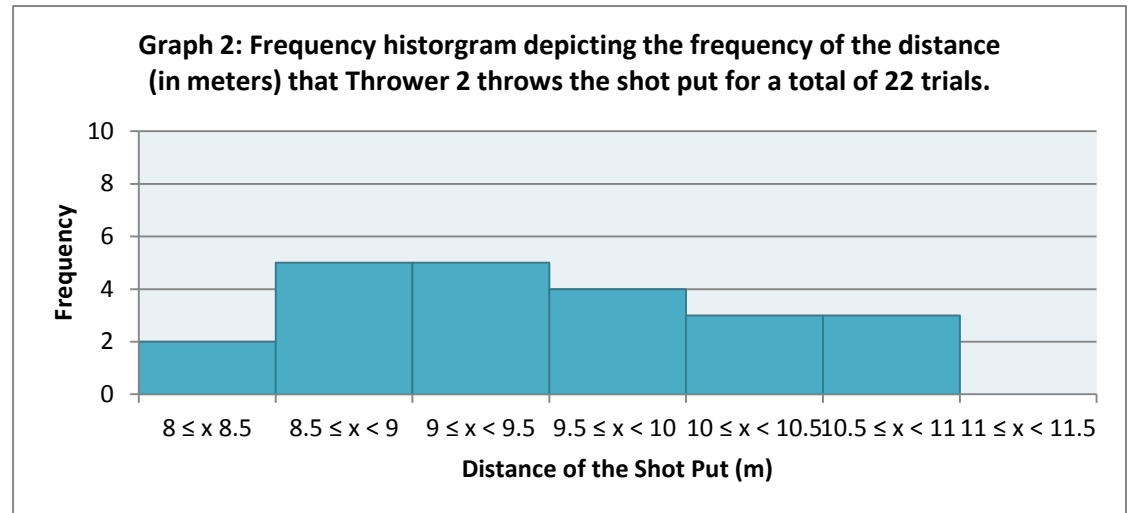


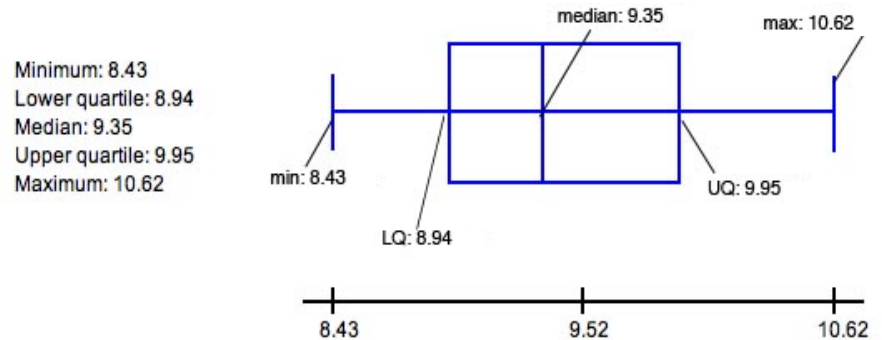
Table 8: Measures of Central Tendencies of Thrower 2

mean	9.475 m
median	9.3 m
mode	9.35 m

Table 9: 5-number summary of Thrower 2

min	8.03 m
Q1	8.96 m
med	9.3 m
Q3	10.1 m
max	10.94 m

Box-and-whisker plot displaying the 5-number summary of Thrower 2



Below, one would find 3 tables: a frequency table depicting the incidence that **Thrower 3** was able to throw the shot put for each of the intervals given, a table that includes the measures of central tendencies of the distance that the Thrower was able to throw the shot put, and a table that depicts the 5-number summary of the distance of the shot put. Below one would also find two visual representations of Thrower 3's 22 trials, by way of a frequency histogram and a box-and-whisker plot. All of this data is displayed in order to represent the performance of thrower 3 in order to have a better analysis on who would be best to go to IASAS.

Table 7: Frequency Table depicting the frequency of the distance (in meters) that Thrower 3 throws the shot put for a total of 22 trials.

Distance of the shot put (m)	Frequency
$8 \leq x < 8.5$	1
$8.5 \leq x < 9$	4
$9 \leq x < 9.5$	7
$9.5 \leq x < 10$	9
$10 \leq x < 10.5$	0
$10.5 \leq x < 11$	0
$11 \leq x < 11.5$	1

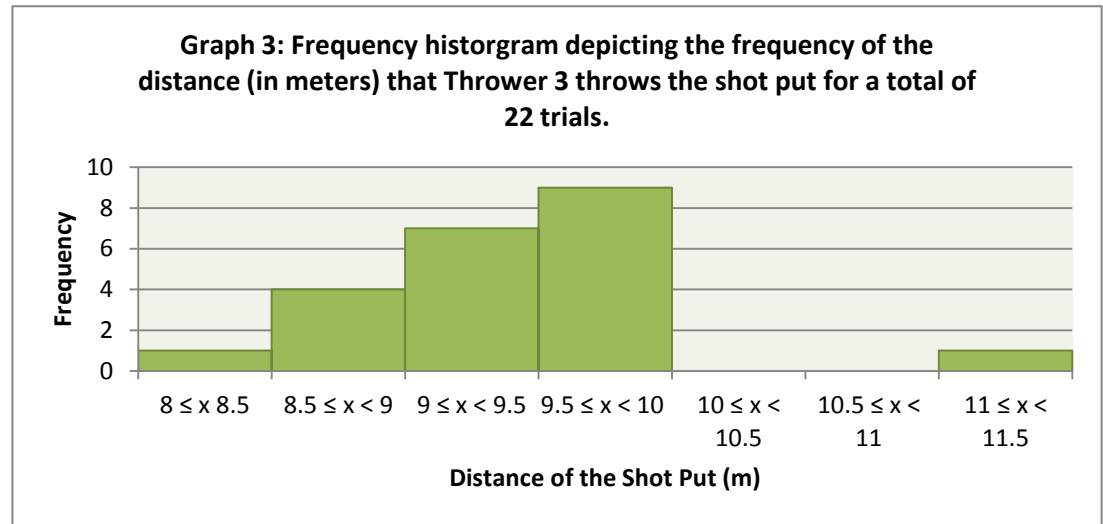


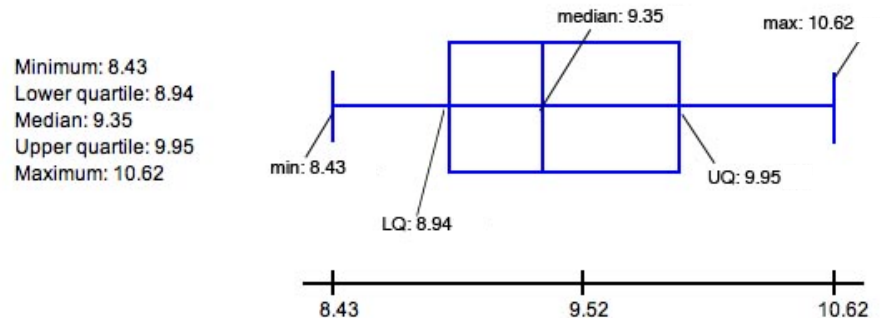
Table 8: Measures of Central Tendencies of Thrower 3

mean	9.47 m
median	9.35 m
mode	9.94 m

Table 9: 5-number summary of Thrower 3

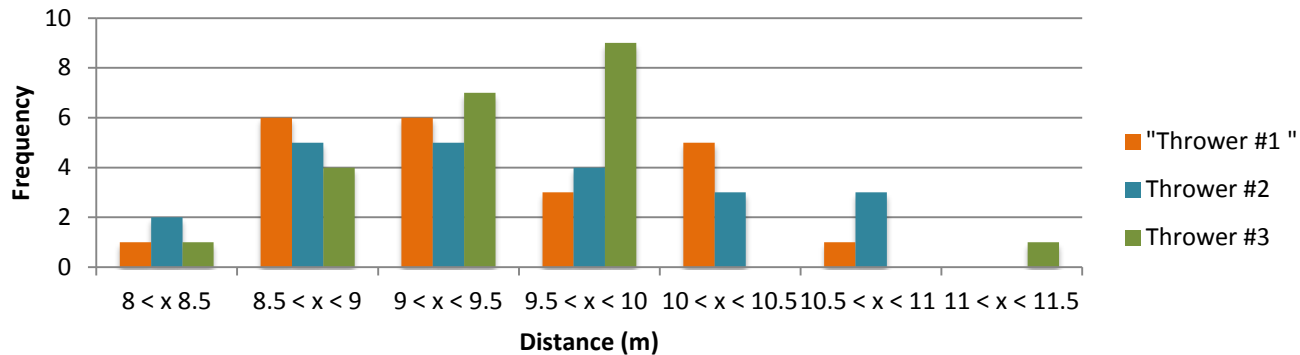
min	8.49 m
Q1	9.12 m
med	9.44 m
Q3	9.74 m
max	11.47 m

Box-and-whisker plot displaying the 5-number summary of Thrower 3



In order to better gage the relationship of the frequencies of each of the throwers, they have all been compiled to be displayed on a bar graph below. From the bar graph below one is able to comprehend that Thrower 3 mostly throws at distances from 9.5 meters to 10 meters, and was only able to reach the 11 meters to 11.5 meters once. Thrower 2 seems to be consistently throwing at distances of 8.5 meters to 9.5 meters, and it appears that he/she is able to hit the further distances less consistently. Out of the three, it is thrower 2 that appears to have the most number of throws at the distance of 10.5 m to 11 m. His/her data appears to be spread out through the intervals but the data displayed on the graph shows that out of the three throwers, in the latter intervals Thrower 2 is the most consistent the data found here, will be much help when choosing the final thrower.

Graph 8: Bar graph that depicts comparison between the frequencies of all three throwers for each of the intervals.



### Decision Making and Justifying

Who is the best? Before we choose the best thrower, we must define what constitutes the best thrower.

In this specific situation, the best thrower would be the one that is able to throw the furthest distance the most consistent amount of times. This criterion would cause one to look closely at the maximum distance of each of the thrower, but also look at the frequency tables and histograms to see at which interval the frequency is highest in order to gage the frequency of the data.

The thrower that I would bring to IASAS would be THROWER 2. This is because Thrower 3 would not be feasible because, although he has the furthest distance from the three throwers at 11.47 m he has no history of throwing the shot put from 10 m to 11 m, which means the thrower's maximum of 11.47 may have just been an outlier or a 'lucky throw,' which means that the statistical basis of the thrower achieving this is slim. Thrower 3's mode was also at 9.94 -- much lower than the other two throwers, and the interval at which Thrower 3 performs consistently in is the  $9.5 < x < 10$  meters, which means he will most likely be getting scores in that range. This leaves Thrower 1 and Thrower 2, when looking at both the box-and-whisker plot and the frequency histogram of the data collected from the trials that both throwers conducted they appear to be almost identical, however the details in these graphs seem almost identical so one must look at the finer details and notice that Thrower 2 was throwing at further distances more consistently than Thrower 1 as in the frequency histogram it shows that Thrower 2 was able to have 3 successful throws at the 10.5 m to 11 m interval while Thrower 1 was only able to hit it once.

Table 10: Distance that Throwers 1, 2, and 3 were able to throw the shot put at for Competition 1, 2, 3, 4, and 5

	Thrower 1	Thrower 2	Thrower 3
Distances for COMPETITION 1 (m)	8.84	10.45	9.12
	9.04	10.76	9.72
	9.76	10.84	9.49
		10.5	
		10.78	
Distances for COMPETITION 2	10.05	9.12	8.48
	10.15	9.05	8.8
	9.43	9.42	9.12
Distances for COMPETITION 3 (m)	9.43	9.61	9.75
	10.45	8.35	9.7
	10.16	8.97	9.92
	10.26		
	10.42		
	10.56		
Distances for COMPETITION 4 (m)	10.04	10.28	9.25
	9.3	9.83	10.662
	9.68	9.76	10.06
			10.25
			10.61
Distances for COMPETITION 5 (m)	9.46	10.85	10.52
	10.67	10.15	10.98
	9.87	9.96	11.26
	10.25	10.68	10.34
	10.26	10.49	11.45
	9.39	10.25	11.02

Because new data has been given, the decision made would be refined through further analysis of the data collected from the 5 different competitions that the throwers participated in. This added information is valuable because in order to have the most thorough understanding of who is the top thrower, one must see how the thrower performs with the added pressure of a competition. These scores were then analyzed through the number of times each thrower made it to the top eight.

The distances of each of the throwers were also compared between the three of them to see who would have won each competition had it only been the three throwers competing. The measures of central tendencies and the 5 number summary were found through collecting all the distances and putting them in one data set.

From the data gathered, a frequency histogram with all the values of the histogram were displayed and box-and-whisker plot were created in order to have a visual representation of the data.

Table 11: Table displaying the places each of the throwers would get by comparing the scores that each of the 3 throwers got from each competition.

	Thrower 1	Thrower 2	Thrower 3
Competition 1	3rd Place	1st Place	2nd Place
Competition 2	1st Place	3rd Place	2nd Place
Competition 3	1st Place	3rd Place	2nd Place
Competition 4	3rd Place	2nd Place	1st Place
Competition 5	3rd Place	2nd Place	1st Place

Table 12: Table depicting the number of times each thrower made it in the top 8 and the total number of successful throws made during the competitions for Throwers 1, 2, and 3

Thrower	Number of Times he/she made it to the top 8	Total Number of Throws made during the competitions
1	2	21
2	1	20
3	1	20

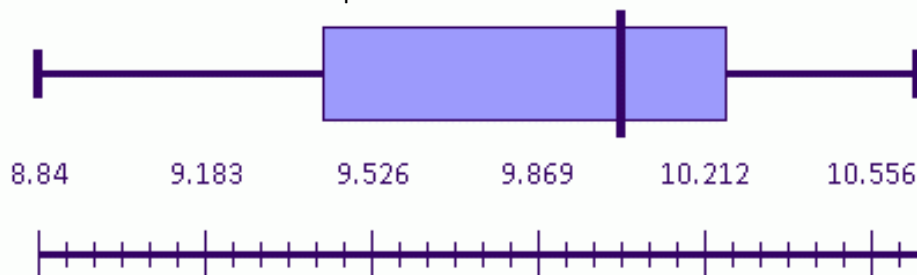
The tables below include the measures of central tendencies (mean median and mode) and the 5- number summary of the data that was collected from the 5 competitions. Also, to better have a better visual of these data values, a box-and-whisker plot was created for each thrower.

Thrower 1	
mean	9.88 m
median	10.04 m
mode	9.43 m, 10.26 m
min	8.84
Q1	9.43
med	10.04
Q3	10.26
max	10.67

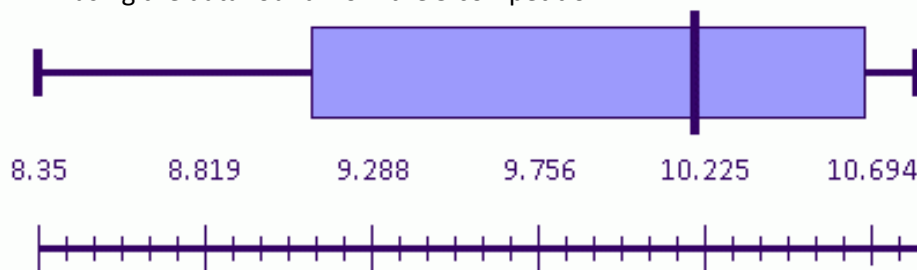
Thrower 2	
mean	10.005 m
median	10.59 m
mode	n/a
min	8.35 m
Q1	9.515 m
med	10.2 m
Q3	10.59 m
max	10.85 m

Thrower 3	
mean	10.0235 m
median	9.99 m
mode	n/a
min	8.48 m
Q1	9.37 m
med	9.99 m
Q3	10.615 m
max	11.45 m

Box-and-whisker plot displaying the 5-number summary of Thrower 1 using the data found from the 5 competition.



Box-and-whisker plot displaying the 5-number summary of Thrower 2 using the data found from the 5 competition.



Box-and-whisker plot displaying the 5-number summary of Thrower 3 using the data found from the 5 competition.

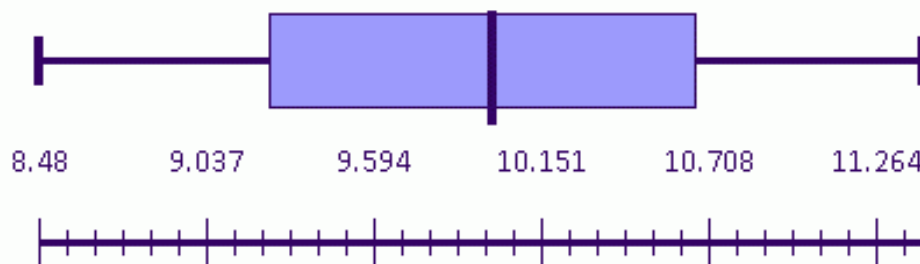
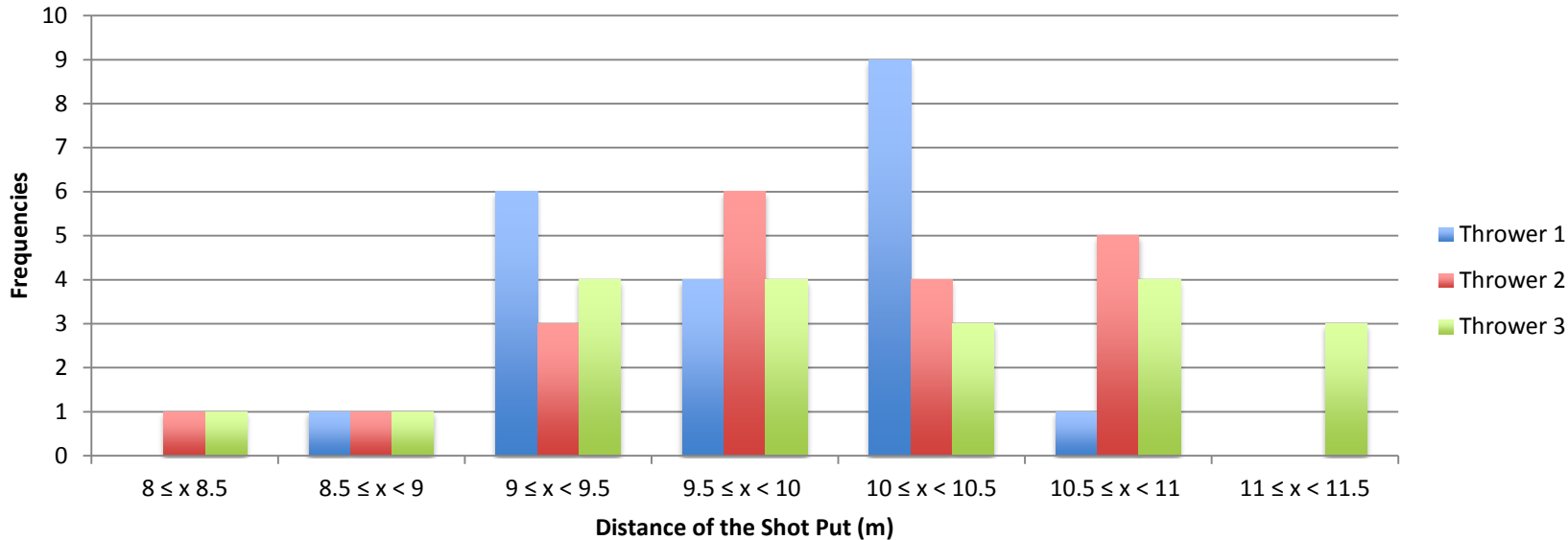


Table 13: Frequency data table depicting the frequency of throws for Throwers 1,2 and 3.

Distance of the shot put (m)	Thrower 1	Thrower 2	Thrower 3
$8 \leq x < 8.5$	0	1	1
$8.5 \leq x < 9$	1	1	1
$9 \leq x < 9.5$	6	3	4
$9.5 \leq x < 10$	4	6	4
$10 \leq x < 10.5$	9	4	3
$10.5 \leq x < 11$	1	5	4
$11 \leq x < 11.5$	0	0	3

Again, in order to find which thrower throws the furthest throws at the most consistent frequency, a Frequency data table and frequency graph were created in order to visually represent the thrower's data.

Graph 10: Bar Graph that depicts the comparison between the frequencies of the distances that each of the 3 throwers were able throw at the 5 competitions.

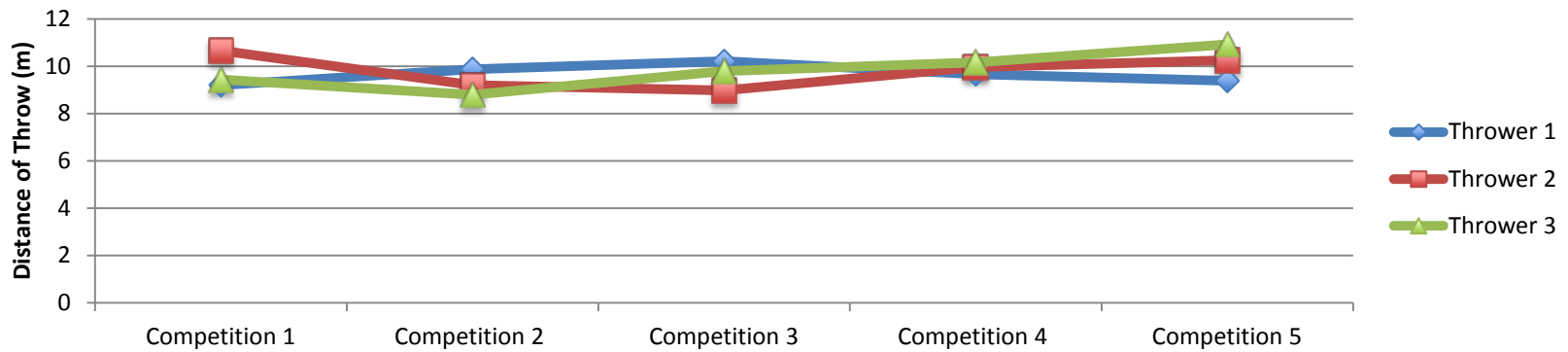




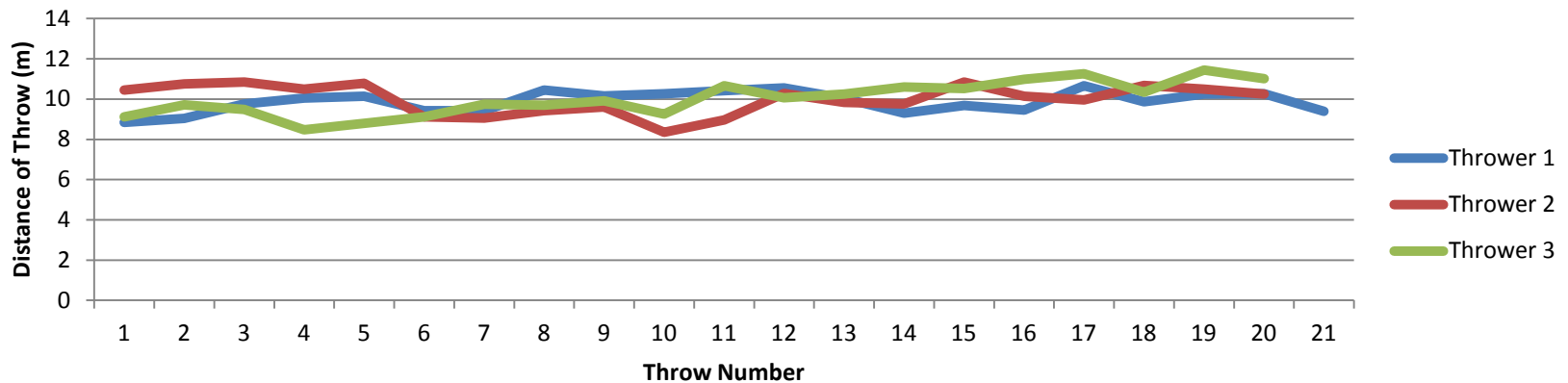
	Thrower 1	Thrower 2	Thrower 3
Competition 1	9.21 m	10.66 m	9.44 m
Competition 2	9.88 m	9.2 m	8.8 m
Competition 3	10.21 m	8.98 m	9.79 m
Competition 4	9.67 m	9.96 m	10.17 m
Competition 5	9.39 m	10.25 m	10.93 m

In order to determine the Thrower's performance over the season and through the various competitions two line graphs were made in order to depict the average distances that the throwers were able to make at these competitions as well as to depict all the raw scores that they achieved through the competition in order to see their progress.

Graph 12: Line graph displaying the average distances that Throwers 1, 2, and 3 made at Competitions 1, 2, 3, 4 and 5



Graph 13: Line graph displaying the distances that Throwers 1, 2, and 3 made in sequential order from Competition 1 to Competition 5.



## Revision of Decision

From the further analysis done, it would be best to revise my decision from Thrower 2 to Thrower 3. From the data collected and through the analysis conducted it appears that Thrower 3, is the best thrower. Again, let us revisit what constitutes the best thrower: The best thrower is the one that is able to throw the furthest distance the most consistent amount of times. And from the various graphs and tables created it appears that Thrower 3 follows this criterion most closely.

When I compared all the data collected together, it appeared that all throwers made it to the top 8 twice, but when their scores were compared against each other it was Thrower 3 and Thrower 1 that got 1<sup>st</sup> place between the three candidates twice.

It also shows in the box and whisker plot that Thrower 3's scores consisted more of further throws as the data. The values labeled toward the end of the box-and-whisker plot for Throwers 1 and 2, 10.556 m and 10.694 m are what would only be found in the third quartile of the Thrower 3's box and whisker plot as Thrower 3 had the furthest throw at 11.45 m. The frequency histogram also shows that Thrower 3's throws were more consistent than the other two throwers in throwing further distances as Thrower 1 and Thrower 2 were not even able to get any throws at the interval of 11m to 11.5 m, whilst Thrower 3 was able to get 3 throws at that distance. The line graph shows that thrower 3's performance kept getting better and better through the various competitions as the line appeared to be consistently moving upwards through the competitions, whilst the other two throwers were fluctuating through the season.

The line graph also must mean that Thrower 3 must have been listening to training methods and is currently at the peak of his/her throwing capability through his/her last 2 performances in Competition 4 and Competition 5, achieving higher scores than Thrower 2. Also, through the data collected from the competitions it appears that Thrower 3 appears to perform better under pressure.