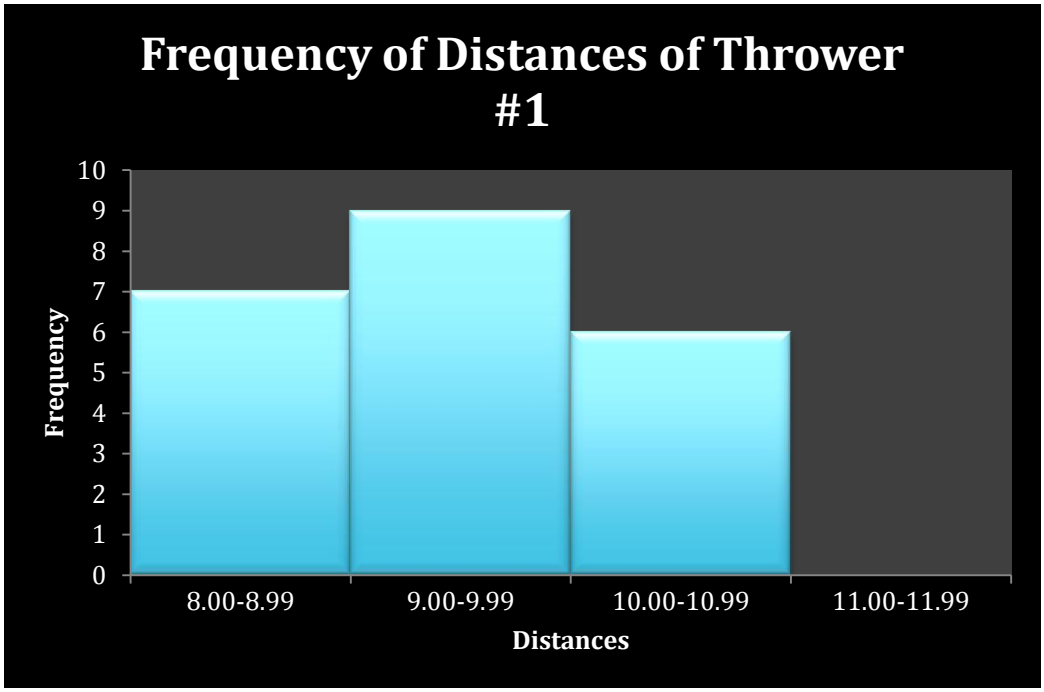
Statistics Project: "Who's the best?'

Track and Field is one of the most athletic sports that everyone loves, especially in ISM. Thus, when it is time for the couches to pick IASAS the top 15 participants, it is often and intricate and rigorous process between the coaches the athletes. Therefore, the most appropriate way to find out who shall and go and who shall stay, a coach would look out for the data (time, or distance) from the whole season and decide from the data found. Therefore, it does not mean that the longest or the fastest runner will be able to go, but more of the most consistent athlete of the whole season. When I say consistent, I mean that in which athlete's data is the closest together. This goes for the for the three Track and Field throwers' data below, in which we have to determine which of the athletes will make to IASAS.

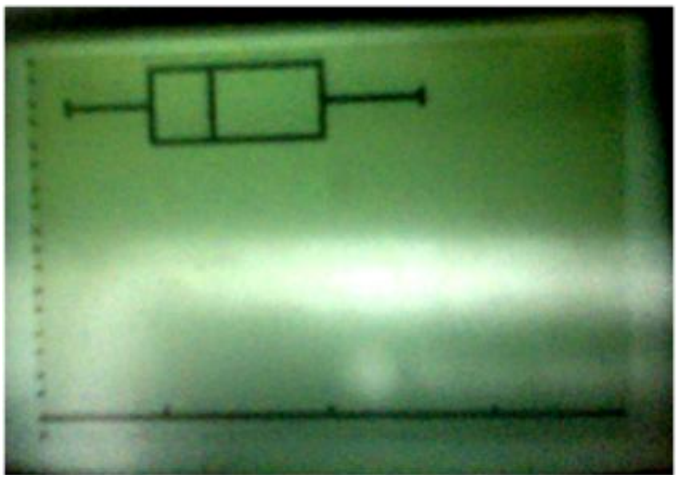
Thrower #1 Data:

Distance of Thrower #1	Frequency
8.00-8.99	7
9.00-9.99	9
10.00-10.99	6
11.00-11.99	0



From the histogram above, we can see that most of Thrower #1"s data is concentrated between the distances of 9.00 to 9.99. However, this is not enough to determine if he is the best fit to go to IASAS. Thus, we need to find different statistical points.

Statistics	Thrower #1
Mean	9.43
Median	9.34
Mode	10.01
Q1	8.94
Q3	10.01
Max Val.	10.62
Min Val.	8.43

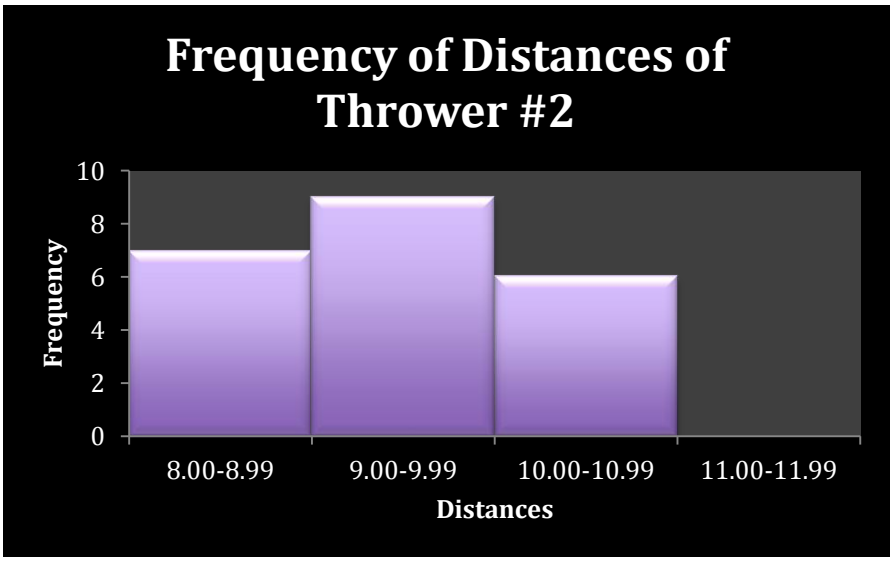


By calculating the Median, Quartile 1, Quartile 3, and the maximum and minimum value, we can find out the spread of data, and further understand the data presented. In addition, we can box and whiskers graph! The box and whiskers provides us a spread of data, showing the Quartile 1 and 3, showing the spread of data between the maximum and minimum.

For thrower #1, we can see that most of the data is above the median. We can also say that 25% of the data is above 10.01. Alternatively, we can also say that from the lower quartile, 8.74, 75% of the data is greater. Moreover, from the box and whiskers, we can see the median, which is 9.34, is leaning more left, or the minimum. This means that most of the data is a lower number.

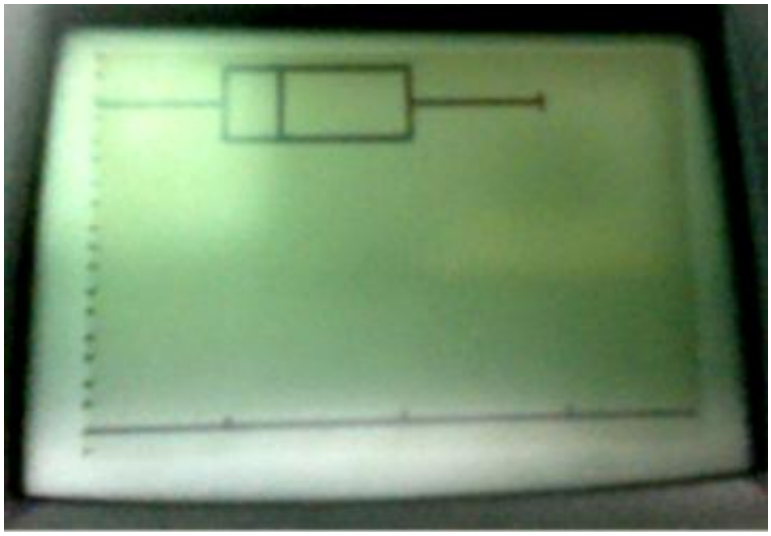
Thrower #2 Data:

Distance of Thrower #2	Frequency
8.00-8.99	7
9.00-9.99	9
10.00-10.99	6
11.00-11.99	0



Once again, we can see from this histogram that Thrower #2 is the most frequent distance of 9.00-9.99. However, this histogram gives us lacking information still to analyze if he is the best, as it is as similar to that of Thrower #1's.

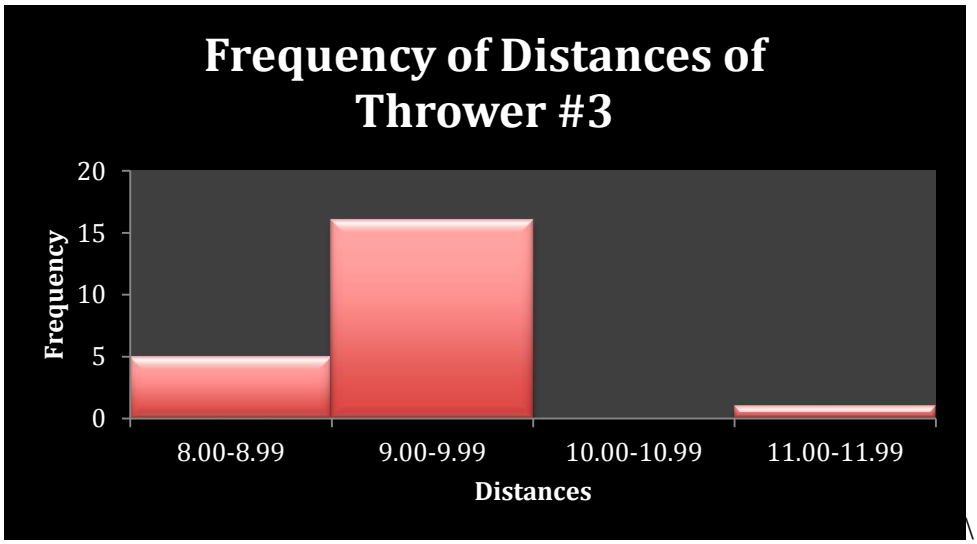
Statistics	Thrower #2
mean	9.78
median	9.3
mode	9.35
Q1	8.96
Q3	10.1
Max Val.	10.94
Min Val.	8.03



As said before, the box and whiskers is an appropriate graph, as it shows more information regarding the data. Thus, we can see that the box, or the "middle half" is more centered between the maximum and the minimum, than that of Thrower #1's box and whiskers diagram.

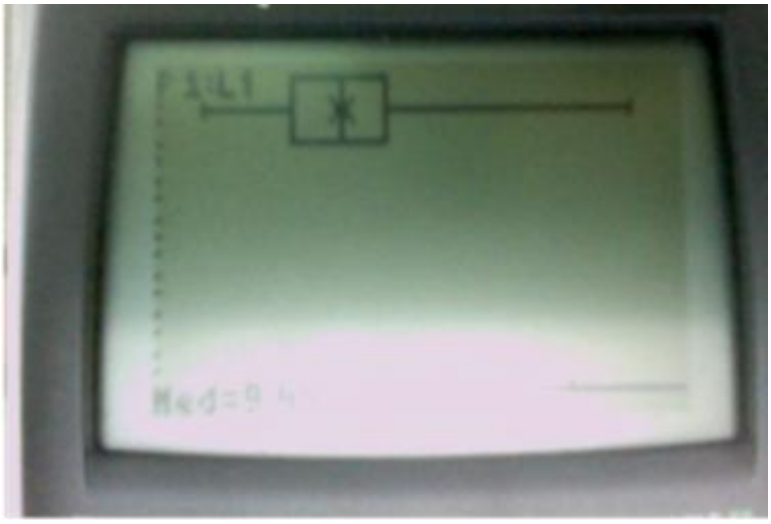
Distance for Thrower #3

Distance of Thrower #3	Frequency
8.00-8.99	5
9.00-9.99	16
10.00-10.99	0
11.00-11.99	1



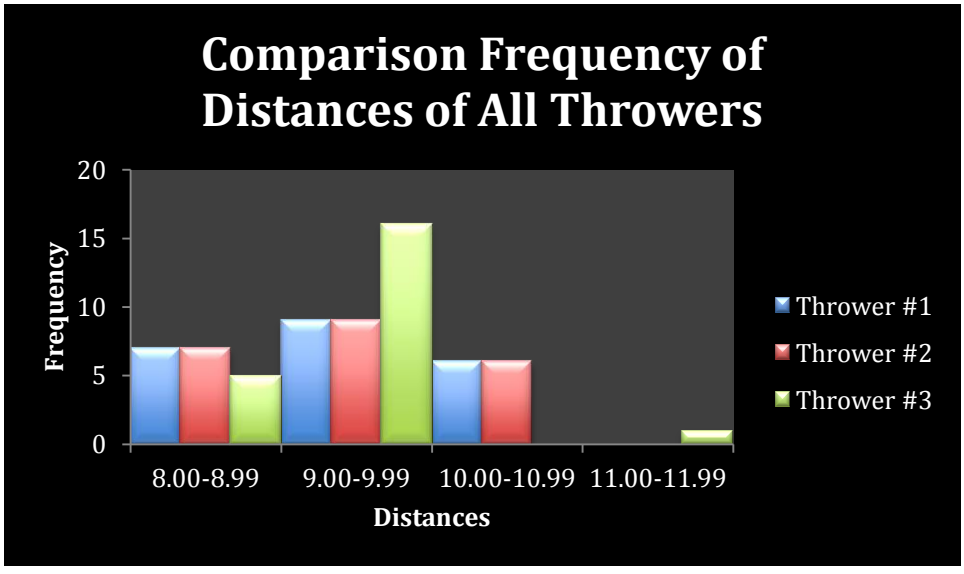
This is the histogram showing the data of Thrower #3. We can see that from the graph, this player is inconsistent. Firstly, from the histogram, we can see that most of this through frequent in the 9.00-9.99 ranges also. However, unlike the first two throwers, thrower 3 does not have any distances in the 10.00-10.99 categories. This, this is also one of the reasons why the data for the 11.00-11.99 can considered an outlier. It is the only data point that is far from the other frequencies as well.

Statistics	Thrower #3
mean	9.47
median	9.44
mode	9.94
Q1	9.12
Q3	9.75
Max Val.	11.47
Min Val.	8.49

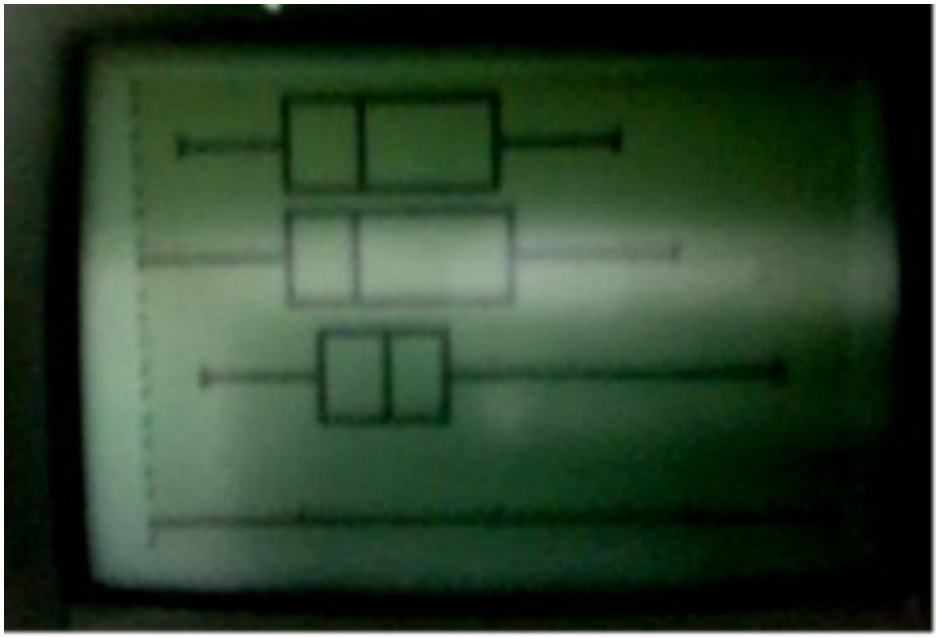


This is the box and whiskers from the median, quartile 1 and 3 and the maximum and minimum value. From this diagram, we can see how unevenly spread the data is. This shows how because of the outlier, the maximum has a loner line to it. Moreover, from the median presented in the graph, we can see that it is more of the lower quarterlies.

Comparing the Three Thrower's distances:



This histogram shows all the frequencies of each thrower, and gives a comparison. As seen, throwers #1 and #2, have an identical data, however after getting the mean, median, mode and other statistical data, we know that thrower #2 has a more consistent and greater value then thrower #1. Moreover, this graph highlights the outlier of thrower #3, as well as the lack of data for the group of 10.00 to 10.99.



From the box and whiskers, we can contrast and compare the three thrower's data. Firstly, we can see that thrower 1 and 2 have similar graphs, but ultimately, thrower 2 has a smaller maximum and larger maximum. Moreover, they also have nearly the same median point. Secondly, we can see how larger a difference thrower 3's box is compared to the other two throwers. This shows how his data among all is

The Best Thrower:

Now we have found the mean, the median, and other statistical terms, we can determine which of the three throwers is fit to go to IASAS. As before, to determine this, we have to find the most consistent player, or the one with the most concentrated data, or the intervals between each point, or with the most evenly spread histogram and box and whiskers. However, we also have to consider how consistently far all three throwers' data are. For the first and second thrower, we know from the histogram that they have similar, if not alike, graphs. This is because of the generalization, or grouping of numbers. Moreover, with Thrower 3, one could easily be deceived and say that Thrower #3 is the best. Because of it's high maximum, which is 11.47. However, we now know, looking at both the histogram and the data, that this is an outlier. If this outlier is taken out, and since he does not have and data for the group of 10.00-10.99 distances, he will have a considerably lower outcome of statistics. For example, his mean, which was originally 9.47, will be, if calculated, to be 9.37, which is now the lowest mean of the three. This shows how the outlier is probably a lucky shot. Therefore, the best person is probably between thrower #1 and #2. However, looking at the data, we can see that thrower 2 has a higher value than thrower #1. Even though their histograms look completely the same, from the mean the median and the mode, we can see that thrower #2 has the highest mean, and the more consistent values, as illustrated in the box and whiskers graph. For example, the mean of Thrower #1 is 9.43 while the mean of Thrower #2 is 9.78.

BUT WAIT!

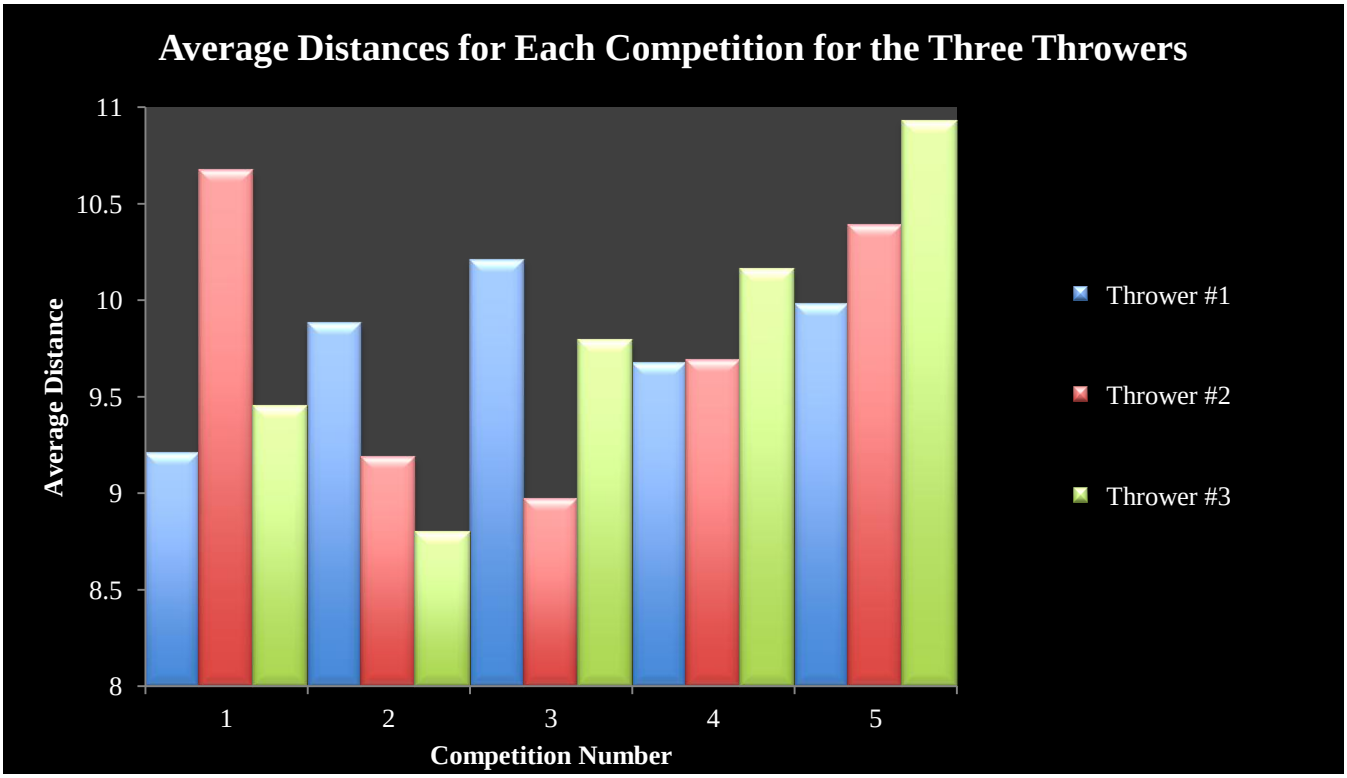
Part 4: New Data and Revisions:

New data is added to all players. As many say, most players usually do better during competition. Moreover, studies have shown that it is through competitions that we can see if throwers have improved over the course of training. Thus, including distances of the throwers during competitions will add to the justification or the revision of who the best thrower is, and thus be part of the IASAS team. The average distance for each of 5 competitions of all three throwers:

Competition Number	**Average Distances**

	Thrower #1	Thrower #2	Thrower #3
1	9.21	10.67	9.45
2	9.88	9.19	8.8
3	10.21	8.97	9.79
4	9.67	9.69	10.16
5	9.98	10.39	10.93

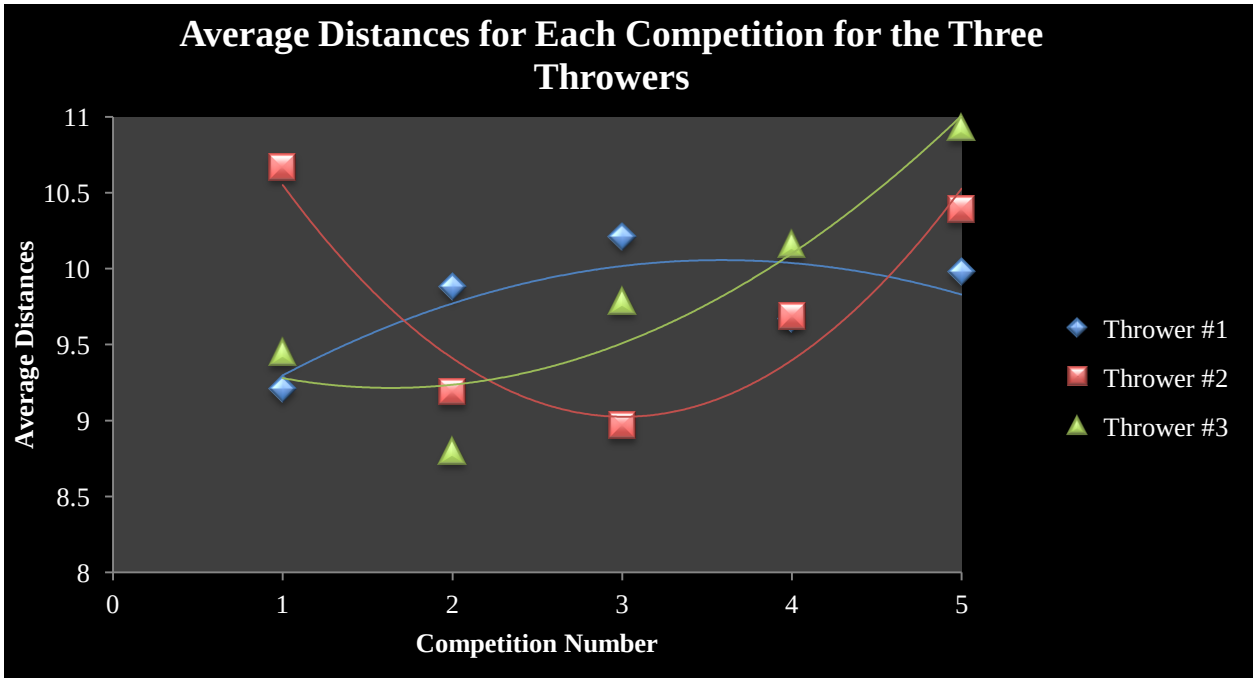
Now, to further understand this data, the most appropriate graph to show, is a comparison histogram of the three thrower's distances. Histograms are the most useful to visualize a broad understanding of the data, as it shows the frequency of the data.



As said before this, we can see the progress of each player for each competition.

For thrower 1, he increases from competition 1 to 3, however it goes back down at competition 4 and increases slightly for competition. For thrower #2, his best is at the first competition, however drops down in competition #2 and drops again for competition 3, and increase at a considerable amount for the 4th and 5th competition. For thrower #3, we can see that see that he starts of low, and drops again for competition #2, however for the 3rd, 4th and 5th competition he has a incredible increase, showcasing different results from the first data set given.

To further understand the data, and see the trends and increase and decrease of results, we can use a scatter plot with trend lines:



This scatter plot graph is also appropriate since it highlights the progress and improvement of each thrower, with the use of trend lines. With the trend lines, we can see ultimately if each of the throwers improve, or get worse, or even stay consistent.

With this scatter plot, we can see that thrower #1, has stayed the most constant in all 5 competitions. This shows that he has improved a little, but ultimately stayed constant, meaning that almost all his points do not have a great distance between them. For thrower #2, we can see a curve as well, this curve is shaped like a bowl, and showing how in the first competition, to the second competition is a steep fall. In addition, if you look at the scatter plot, you will notice how his data is not that a line, or farther apart from each other. Therefore, knowing this, Thrower #2 is not very consistent in competitions. Lastly is thrower #3. Thrower #3 shows data in which he out of all players improve the most, as his trend line shows a steeper, and positive slope.

The Best Thrower (Revised)

At first, I cam to the conclusion after analyzing the first set of results that thrower #2 is the best among the three talented throwers. However, when newer and more accurate data was provided, I now can say that, from the data of the 5 competitions, that thrower #2 has very large range of numbers, most of the data being very far from each other. Furthermore, through the histogram and the scatter plot, I was able come to the analysis that thrower 2 has the steepest slope going down, and the most random heights for the histogram. Therefore, we can say that, in times of competitions, thrower #2 is not the most consistent player, and unless he is able to repeat is performance of the first given data, will not be chosen for IASAS.

Thrower #3, however, is different all together. From the histogram, the calculated averages and the scatter plot, we can see that thrower #3 has drastically improved throughout the season. Moreover, for last competition, he was able to beat the two other throwers, with his maximum of 10.93. Thus, we can say that thrower #3 is the farthest, and most improved thrower for the season. Still, the best thrower is not only determined how far you can throw, but how consistent and constant are the distance of your throws. In addition, from the data, we can see that though he does make a drastic improvement with his throws, the range of his data is too high, or the difference between each data point/distance is too high. For example, in competition #3, he was able to make an average of 9.79 while in the next competition; he made a huge improvement with an average distance of 10.16. Moreover, as we know from the first data set given, he is not a constant thrower, with his data having a outlier in tow. In conclusion, although he may be the farthest through, he could also be as unstable as thrower #2.

Finally, I think that thrower #1 is the most consistent thrower that deserves to be sent to IASAS. Firstly, from the first data series, we know that thrower #2, with the mean of 9.78, had out showed him, even if they had very similar results. However, with the new data present, we can see that in competitions, he has made improvements and his data had the smallest range, meaning that he had the most constant data of all. From the scatter plot as well, the alignment of the data was very clear. Therefore, being the most constant player, with favorable distances, I think thrower #1 is IASAS-worthy.