

Statistics Coursework
Mayfield High School

I am going to develop a statistical investigation and an enquiry which is set on data from a fictitious high school called Mayfield. Furthermore, this data was provided.

Year Group	Number of Boys	Number of Girls	Total
7	151	131	282
8	145	125	270
9	118	143	261
10	106	94	200
11	84	86	170

Firstly, the total number of students at the school is 1183.
A lot of information on the students is available on the computer such as the name of the student, the age, year group, IQ, weight and height.

I will base my investigation on heights and weights.

Furthermore, I will be making 6 hypotheses which will be based on previous knowledge.

Hypothesis

1. I think that the taller the girl the heavier they will be. I have noticed this at my school. It will range between different year groups.
2. I think that the taller the boy, the heavier he will be. I have noticed this at my school.

3. I think that as the younger the student, the less their weight. I have noticed this at my school. I think this because the boys haven't gone into growth spurt yet and grown as much.
4. I think that the older the student the taller they will be. This is because once in puberty they grow very quick and more the weight.
5. I think the boys will be taller than the girls.
6. I think that the boys will be heavier than the girls. I have learnt this from my school.

I have chosen these hypotheses because I have noticed these at my school and these hypotheses seem to be true.

I will try to prevent being bias and make everything fair. To start the enquiry, I will be collecting a random sample of data. A random sample will prevent bias. I am going to get 15% of each set of data of boys and girls in every group so that I will have a sufficient amount of data to investigate.

Each sample will be rounded to the nearest suitable figure.

Firstly I will work out how many students I will collect in each strata which I will need to use the following calculations. I will then round each calculation to the nearest figure.

$$\text{Year 7 boys} = \frac{151}{100} \times 15 = 22.65 - \text{rounded} = 23$$

$$\text{Year 7 girls} = \frac{131}{100} \times 15 = 19.65 - \text{rounded} = 20$$

$$\text{Year 8 boys} = \frac{145}{100} \times 15 = 21.75 - \text{rounded} = 22$$

$$\text{Year 8 girls} = \frac{125}{100} \times 15 = 18.75 - \text{rounded} = 19$$

$$\text{Year 9 boys} = \frac{118}{100} \times 15 = 17.7 - \text{rounded} = 18$$

$$\text{Year 9 girls} = \frac{143}{100} \times 15 = 21.45 - \text{rounded} = 22$$

$$\text{Year 10 boys} = \frac{106}{100} \times 15 = 15.9 - \text{rounded} = 16$$

$$\text{Year 10 girls} = \frac{143}{100} \times 15 = 21.45 - \text{rounded} = 22$$

$$\text{Year 11 boys} = \frac{84}{100} \times 15 = 12.6 - \text{rounded} = 13$$

$$\text{Year 11 girls} = \frac{86}{100} \times 15 = 12.9 - \text{rounded} = 13$$

Year Group	Number of Boys	Number of girls	Total Number
7	23	20	43
8	22	19	41
9	18	21	39
10	16	14	30
11	13	13	26

The strata in year 11 seem to be smaller than the other years. I shall bear this in mind when looking at the year elevens. Their results may not be as reliable.

I will collect the samples using this information:

Year 7:

Year 7 Boys

Pupil No.	Gender	Height	Weight
96	M	1.61	63
108	M	1.51	45
30	M	1.47	50
70	M	1.35	29
262	M	1.52	42
5	M	:	40
176	M	1.5	56
206	M	1.54	42
171	M	1.47	45
53	M	1.59	52
6	M	1.42	26
242	M	1.72	75
155	M	1.54	48
127	M	1.67	63
101	M	1.47	41
78	M	1.43	33
28	M	1.49	38
276	M	1.55	32
153	M	1.5	40
33	M	1.54	43
151	M	1.63	60
49	M	1.55	50
57	M	1.55	38

Year 7 Girls

Pupil No.	Gender	Height	Weight
38	F	1.53	52
237	F	1.47	44
271	F	1.61	47
149	F	1.54	42
277	F	1.75	40
183	F	1.25	33
164	F	1.42	48
110	F	1.63	51
84	F	1.48	37
80	F	1.64	40
116	F	1.48	47
106	F	1.58	50
43	F	1.57	45
150	F	1.48	34
128	F	1.62	40
162	F	1.6	45
159	F	1.42	34
173	F	1.61	52
275	F	1.6	57
186	F	1.59	38
198	F	1.51	50

Year 8:

Year 8 Boys

Pupil No.	Gender	Height	Weight
450	M	1.53	32
468	M	1.72	51
543	M	1.55	38
373	M	1.67	51
540	M	1.54	37
311	M	1.57	50
337	M	1.57	62
380	M	1.2	38
460	M	1.65	59
494	M	1.6	38
291	M	1.52	38
320	M	1.5	39
487	M	1.43	48
455	M	1.57	49
500	M	1.42	48
526	M	1.52	43
548	M	1.68	45
342	M	1.7	69
287	M	1.65	40
348	M	1.65	44

Year 8 Girls

Pupil No.	Gender	Height	Weight
416	F	1.54	52
473	F	1.57	51
387	F	1.7	55
361	F	1.66	72
293	F	1.75	50
334	F	1.6	46
301	F	1.65	52
447	F	1.54	54
545	F	1.45	51
426	F	1.62	46
400	F	1.56	45
371	F	1.63	42
377	F	1.69	61
412	F	1.59	52
429	F	1.66	50
431	F	1.72	43
542	F	1.65	58
497	F	1.61	52
282	F	1.55	42
317	F	1.62	46

Year 9:

Year 9 Boys

Pupil No.	Gender	Height	Weight
663	M	1.7	47
798	M	1.6	48
572	M	1.71	68
726	M	1.52	54
657	M	1.6	55
579	M	1.61	60
619	M	1.74	57
662	M	1.72	62
658	M	1.58	50
679	M	1.7	50
793	M	1.75	52
642	M	1.65	50
776	M	1.56	60
654	M	1.59	68
811	M	1.67	54
623	M	1.65	55
681	M	1.67	42
704	M	1.6	40
797	M	1.64	59

Year 9 Girls

Pupil No.	Gender	Height	Weight
749	F	1.53	57
711	F	1.62	48
569	F	1.51	48
757	F	1.53	48
753	F	1.62	42
714	F	1.63	52
812	F	1.65	48
694	F	1.68	36
557	F	1.5	65
718	F	1.62	54
614	F	1.8	62
656	F	1.2	42
777	F	1.57	52
669	F	1.57	40
743	F	1.58	55
627	F	1.75	56
806	F	1.58	36
766	F	1.58	40
562	F	1.51	48
616	F	1.61	38
773	F	1.5	45
723	F	1.53	58
763	F	1.67	48

Year 10:

Year 10 Boys

Pupil No.	Gender	Height	Weight
50	M	1.57	40
16	M	1.75	56
71	M	1.55	50
157	M	1.67	44
83	M	1.85	70
196	M	1.62	57
65	M	1.65	50
185	M	1.84	62
183	M	1.63	50
82	M	1.79	75
137	M	1.75	57
122	M	1.7	59
11	M	1.66	66
103	M	1.66	63
41	M	1.65	64
48	M	1.55	57
193	M	1.5	65

Year 10 Girls

Pupil No.	Gender	Height	Weight
22	F	1.73	51
99	F	1.67	57
72	F	1.6	50
77	F	1.65	54
189	F	1.7	55
160	F	1.55	60
49	F	1.7	48
95	F	1.68	50
131	F	1.58	36
169	F	1.61	59
27	F	4.65	53
146	F	1.56	63
134	F	1.61	54
136	F	1.63	48
123	F	1.74	47

Year 11:

Year 11 Boys

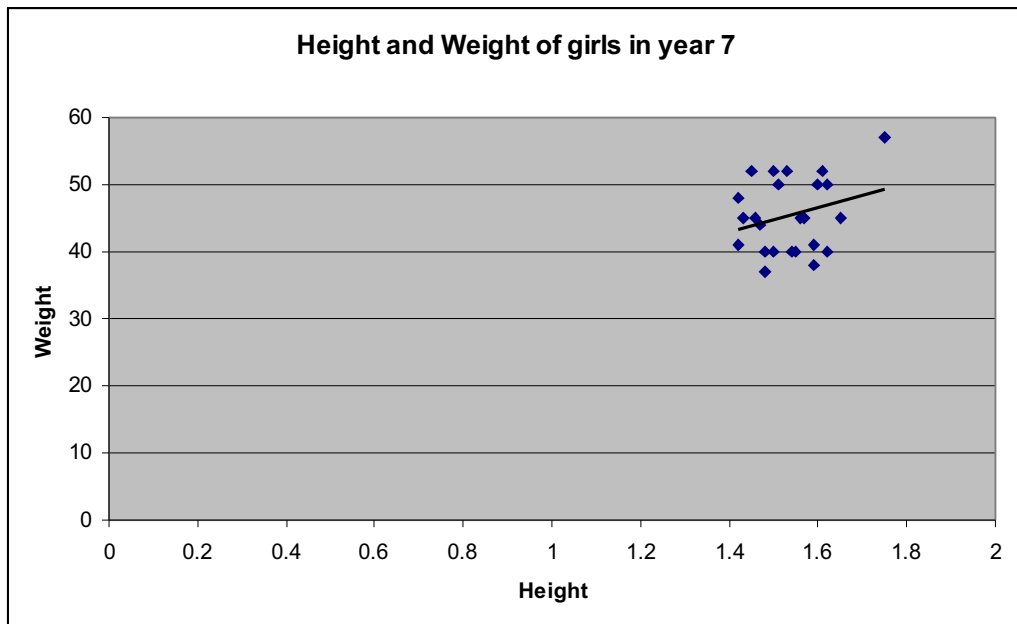
Pupil No.	Gender	Height	Weight
256	M	1.79	72
322	M	1.92	45
350	M	1.73	50
357	M	1.65	45
299	M	1.68	63
347	M	1.65	50
260	M	1.8	72
278	M	1.81	54
280	M	1.65	58
242	M	1.55	54
241	M	1.88	75
321	M	1.64	60
236	M	1.62	52
344	M	1.61	42

Year 11 Girls

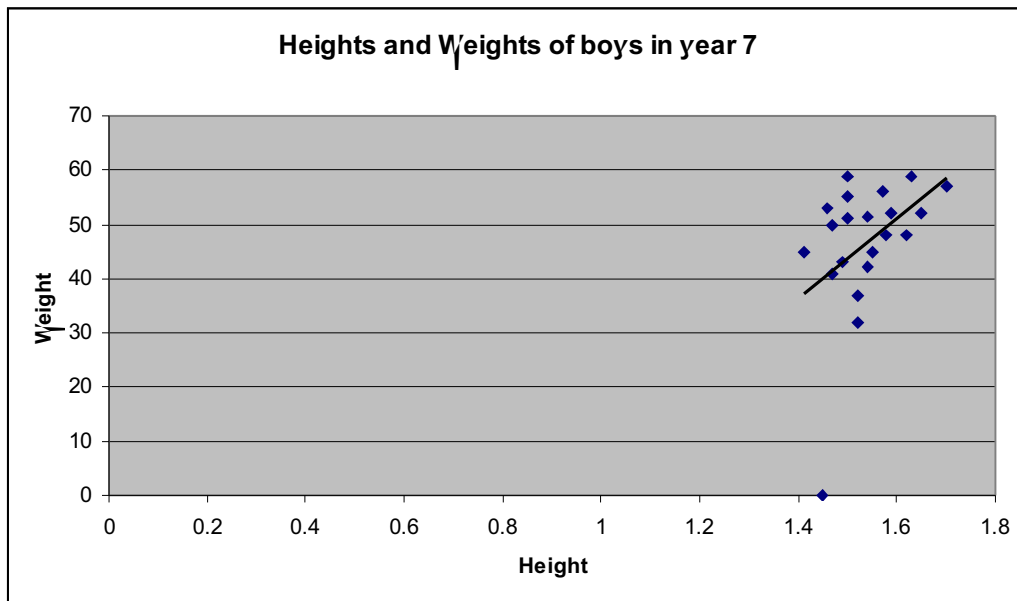
Pupil No.	Gender	Height	Weight
245	F	1.7	50
302	F	1.56	45
261	F	1.62	51
361	F	1.71	42
370	F	1.67	48
203	F	1.6	54
229	F	1.65	66
360	F	1.76	56
257	F	1.72	60
294	F	1.6	45
215	F	1.68	54
359	F	1.59	45
251	F	1.63	45
298	F	1.63	48

This is the number of samples of people I will take from the generator:

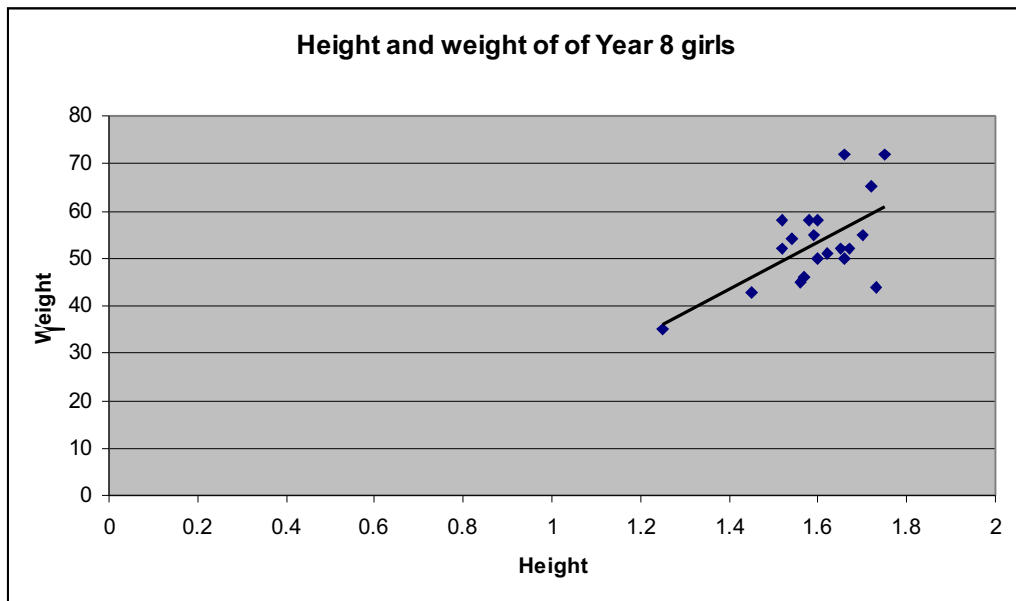
I am now going to put my samples into scatter graphs. I am going to be looking for a link between heights and weights. I will work out the correlation coefficient. The correlation coefficient will tell me how strong the link is. I will look for any outliers. These are results that do not fit the pattern. I must then decide if these outliers are mistakes, or just very big or small students.



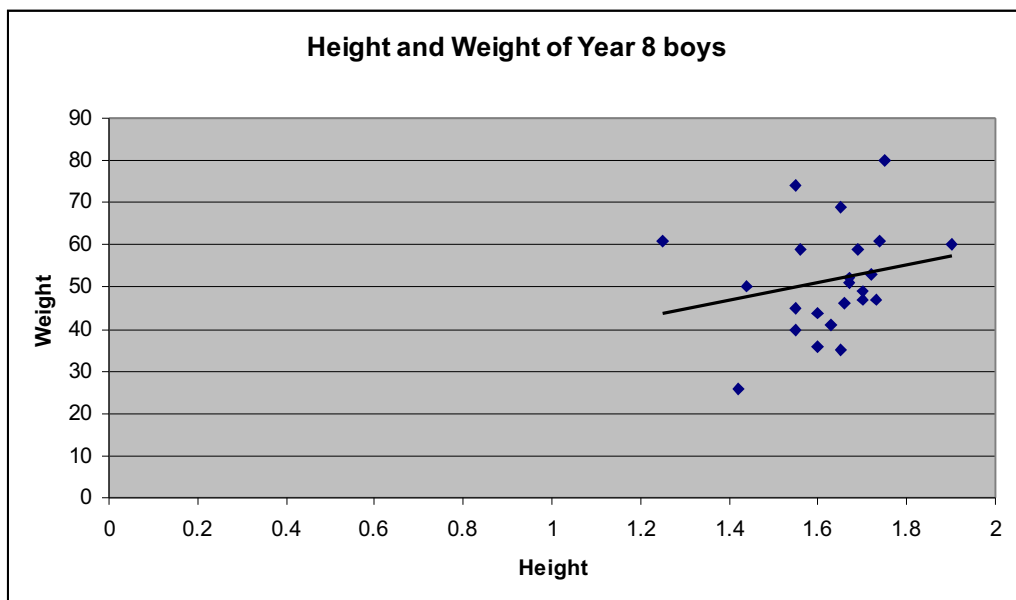
I have drawn a line of best fit. Here we see the more the weight the taller they are. The correlation co-efficient for this graph is 0.36254. The coefficient shows that the correlation is positive and shows a good correlation.



Here I have also drawn on a line of best fit. The line shows that the heavier they are the taller they are. The correlation coefficient for this graph is 0.827073. Here there is strong positive correlation between the heights and weights of boys. The correlation is high

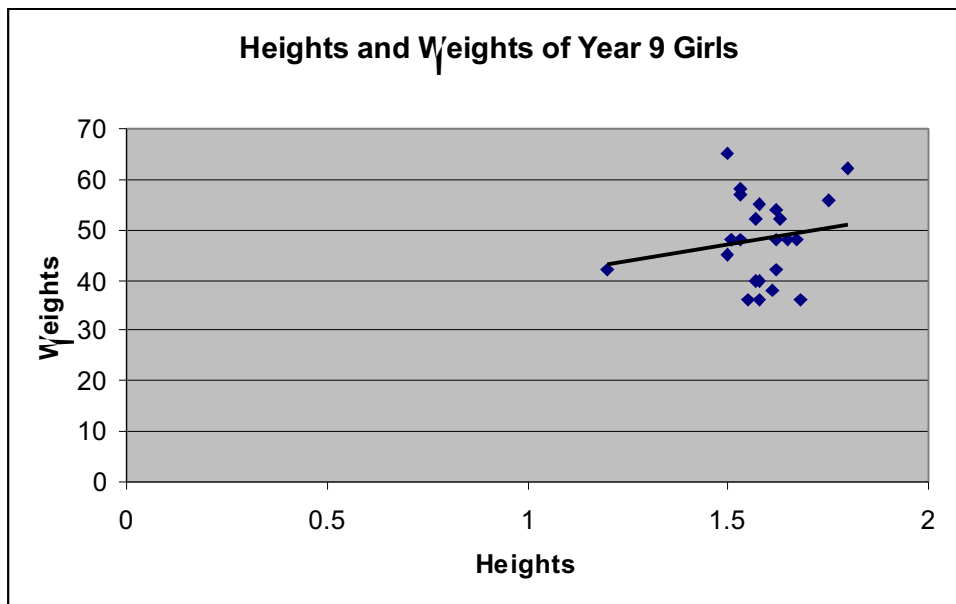


I have put in the line of best fit and the line of best fit shows positive correlation. The more they weigh the taller they are. The correlation co- efficient for this graph is 0.183311. There is a lot of results around my line of best fit. The correlation is quite low for these results and a scattered correlation.

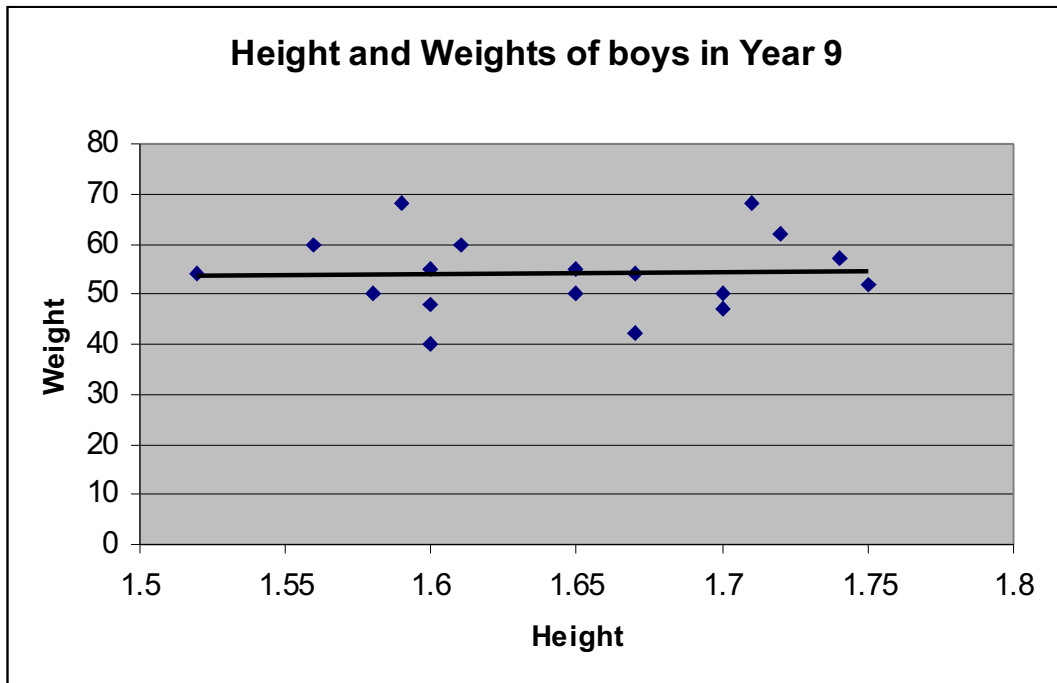


This chart shows a positive correlation. The correlation co- efficient for this graph is 0.407866. This graph contains an odd result. The results look to be scattered. The correlation is good but shows a scattered correlation

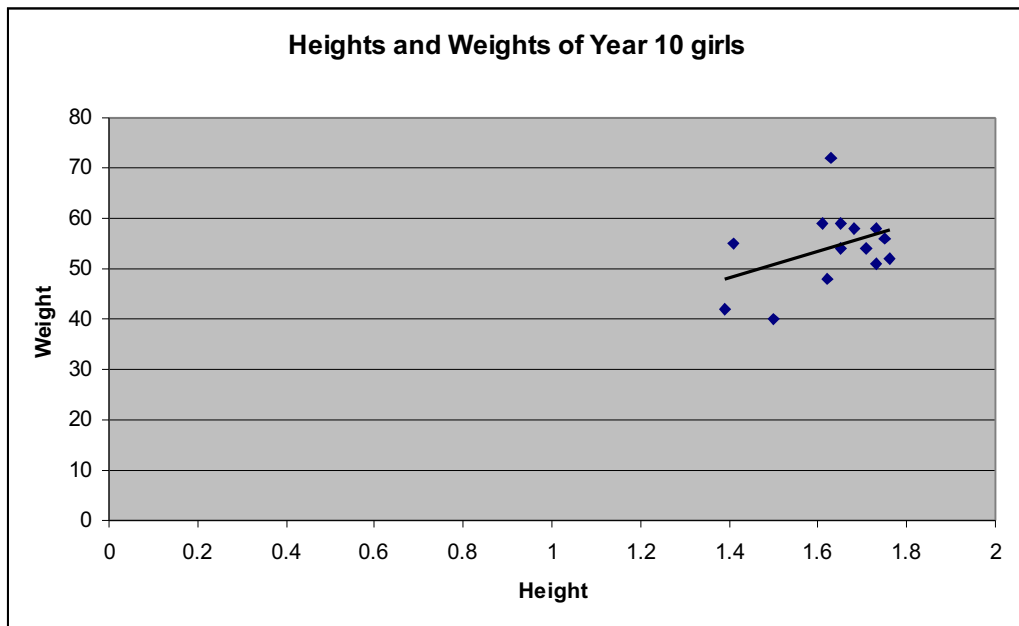
By Ahsan Ahmed



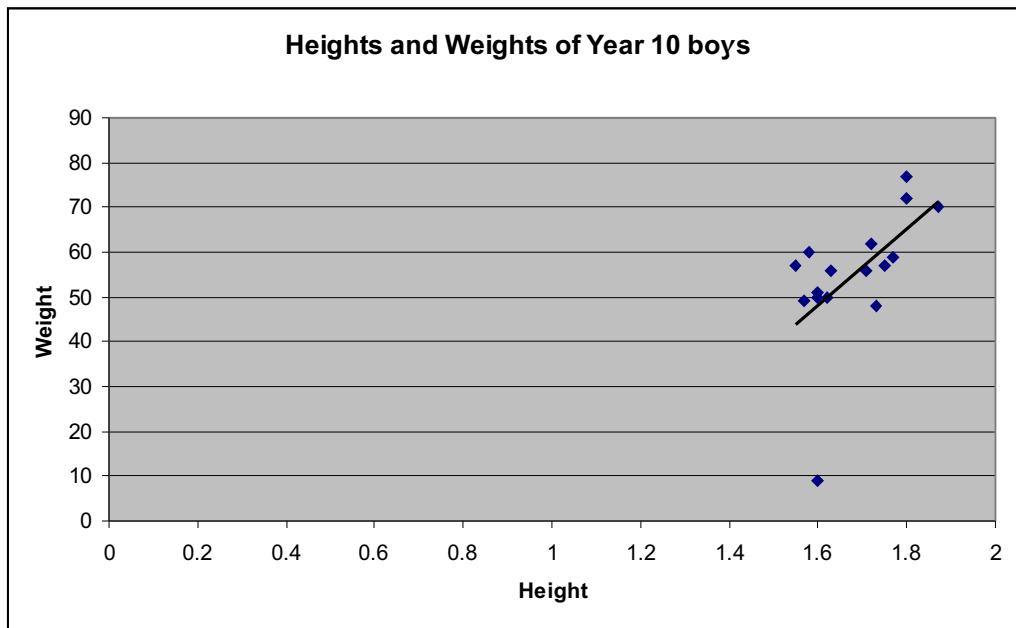
This chart shows a positive correlation. The correlation coefficient for this graph is 0.168134. These results show the smaller the student the more heavily. The results are scattered. Not as I hypothesised even though the line of best fit shows the taller the girl the more they weigh. There are some odd results.



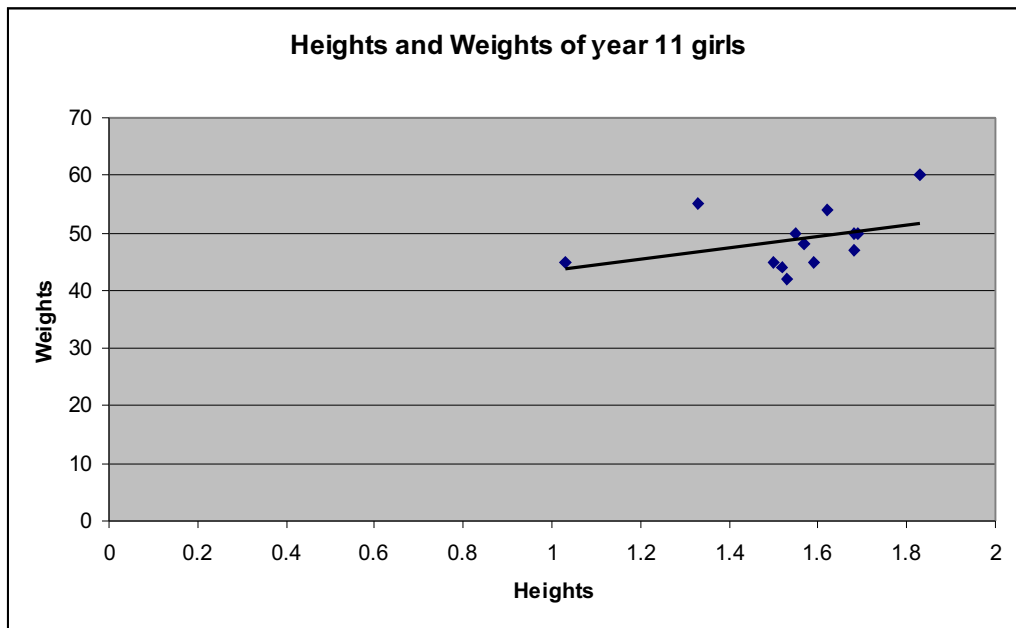
This chart shows a positive correlation. The correlation coefficient for this graph is 0.029926. These results are evenly spread showing that the correlation is poor and there are some boys who have grown quicker than others. This shows no link to my hypothesis as all the results are all scattered as shown by the line of best fit.



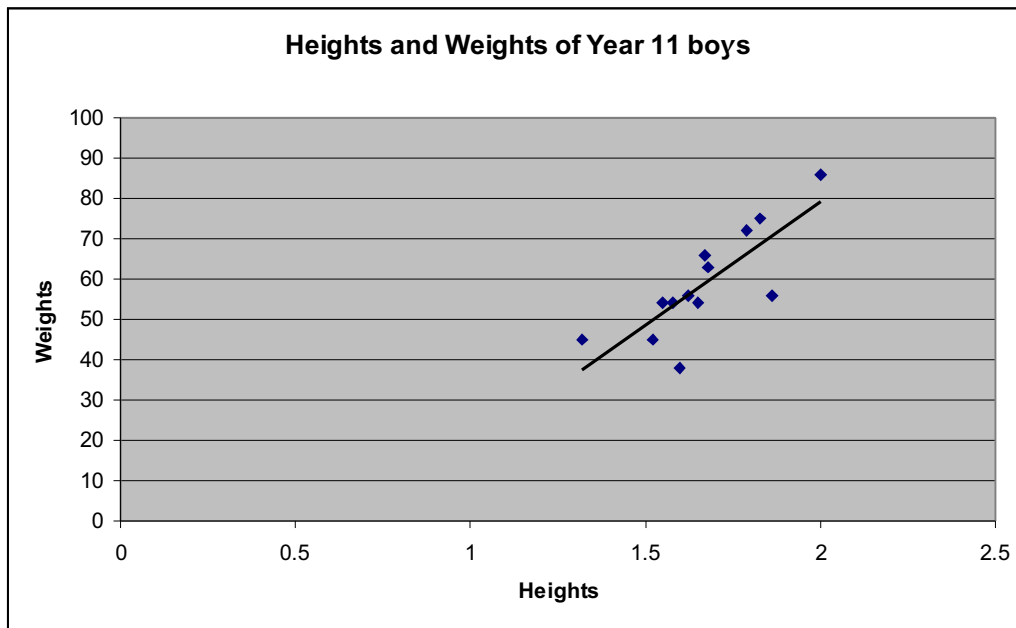
This chart shows a positive correlation. The correlation coefficient for this graph is 0.00899. The coefficient is very low. The results show taller the taller the student the more the weight. This was as I hypothesised the taller the student the more they will weigh.



This chart shows a positive correlation. The correlation coefficient for this graph is 0.449898. This graph shows as the boys get taller they also get heavier. The line of best fit shows exactly as I hypothesised. The taller the student the more they weigh.



This chart shows a positive correlation. The correlation coefficient for this graph is 0.369314. As the height is increasing the weight also increases gradually. This was not as I hypothesised there are a number of students who weigh around 60kg and are the height is around 1.15. I hypothesised that more the students the weigh the taller they are.



This chart shows positive correlation. The correlation coefficient for this graph is 0.379445. As the height increases, they start to weigh more. The correlation is good even though there are a couple of boys outstanding in the line graph, who weigh less and are smaller than others. The chart shows that I hypothesised rightly because the taller the student the more they will weigh although there are a couple of odd results.

KS3 and KS4

I am going to group the Year 7, 8 and 9 pupils and group the next year 10 and 11's which would then show average weights and heights of KS3 and KS4, boys and girls. In doing this, I will be much be a lot easier to compare the differences. I gathered all the boys and girls heights and weights from KS3 and used Excel to work out the average. I did the same to find out the averages the KS4. Here are the charts I created after working the averages out, rounded to 2 decimal points.

KS3		KS4	
Height	Weight	Height	Weight
1.58	48.4	1.72	54.73

These averages show that the taller the boy or girl the heavier they are, as I hypothesised. I also hypothesised that the older the student the taller and heavier they are.

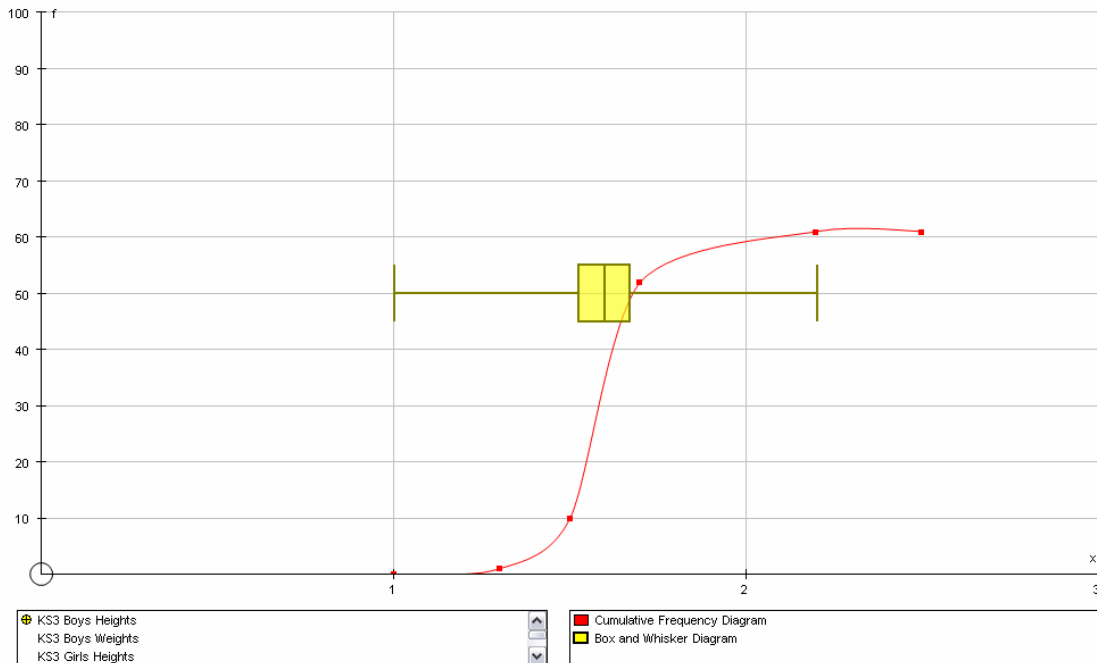
All four averages show just as I expected, the older the students the

I will draw cumulative frequency graphs to allow me to see the lower quartile, the upper quartile and the median; I can also see the interquartile range. To analyse these further, I will produce box and whisker diagrams, which I can use to analyse my hypothesis.

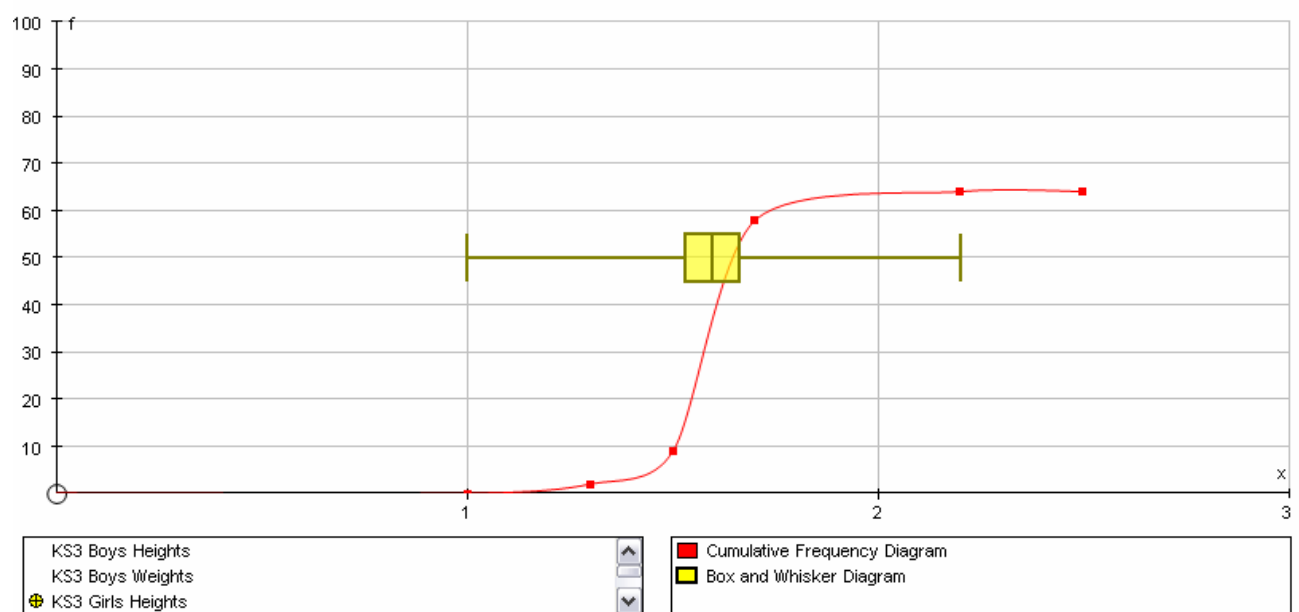
If I produce box plots from each of these I will be able to make comparisons about the spread of data about the median. I will be able to make a comparison between each gender and KS and see if there is a relationship between each of these. Also I will be able to compare each of the range, the inter-quartile range and the median of each group.

Cumulative Frequency Graphs and Box Plots:

KS3 Boys Heights

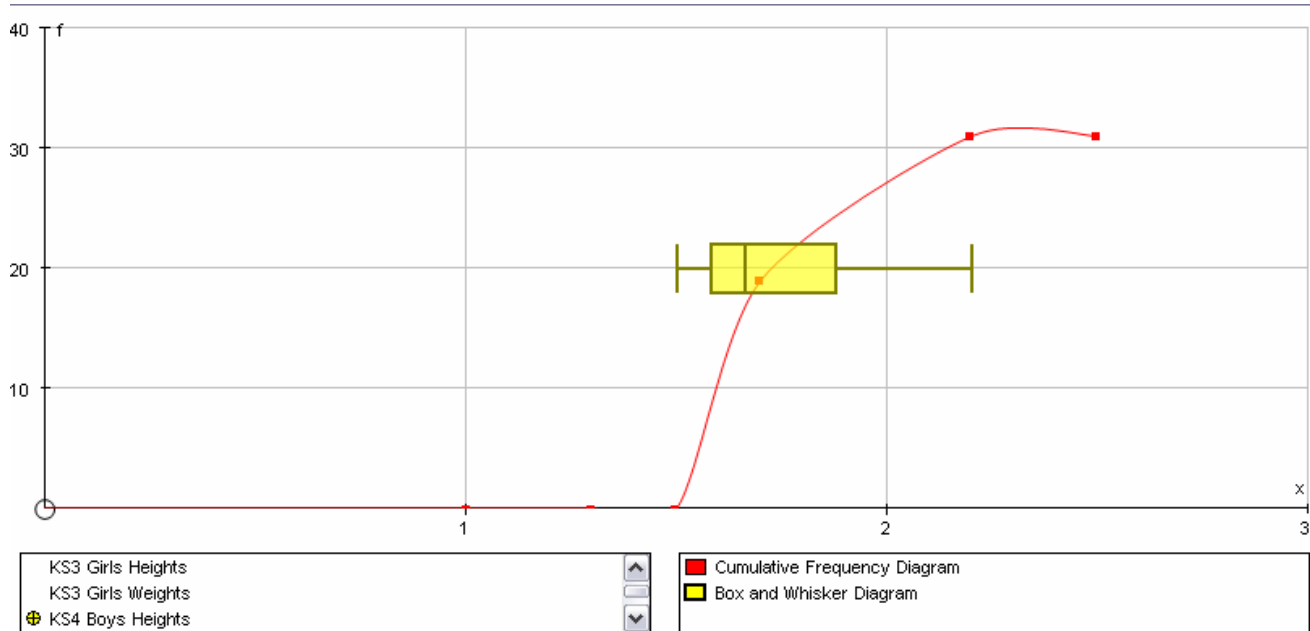


KS3 Girls Heights

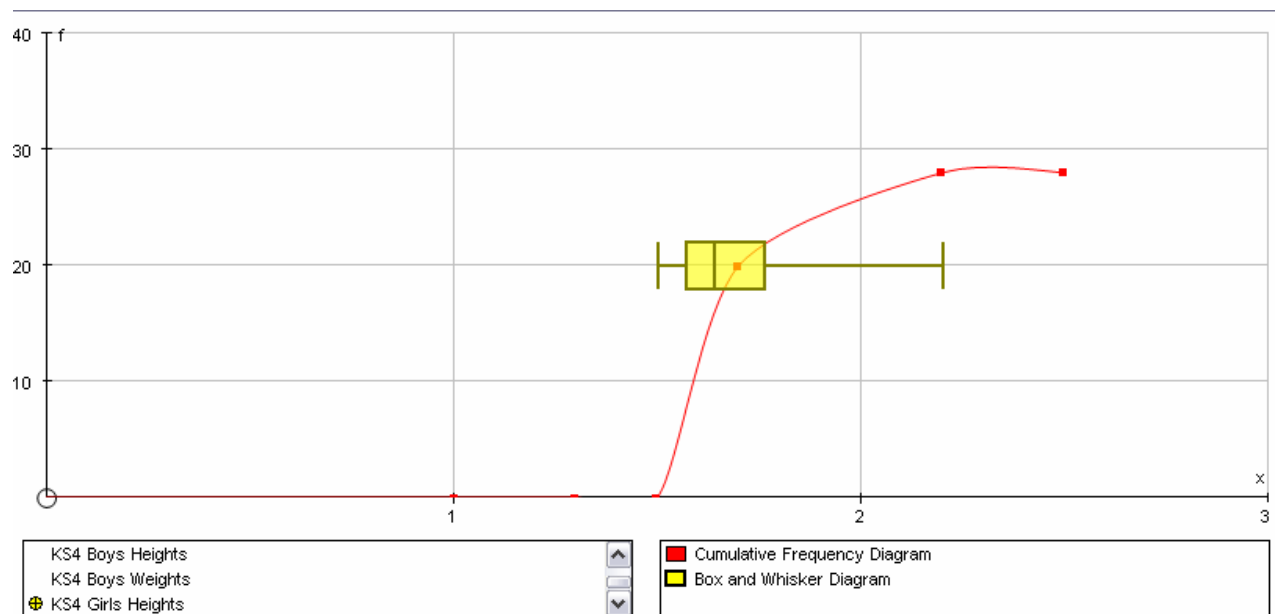


By Ahsan Ahmed

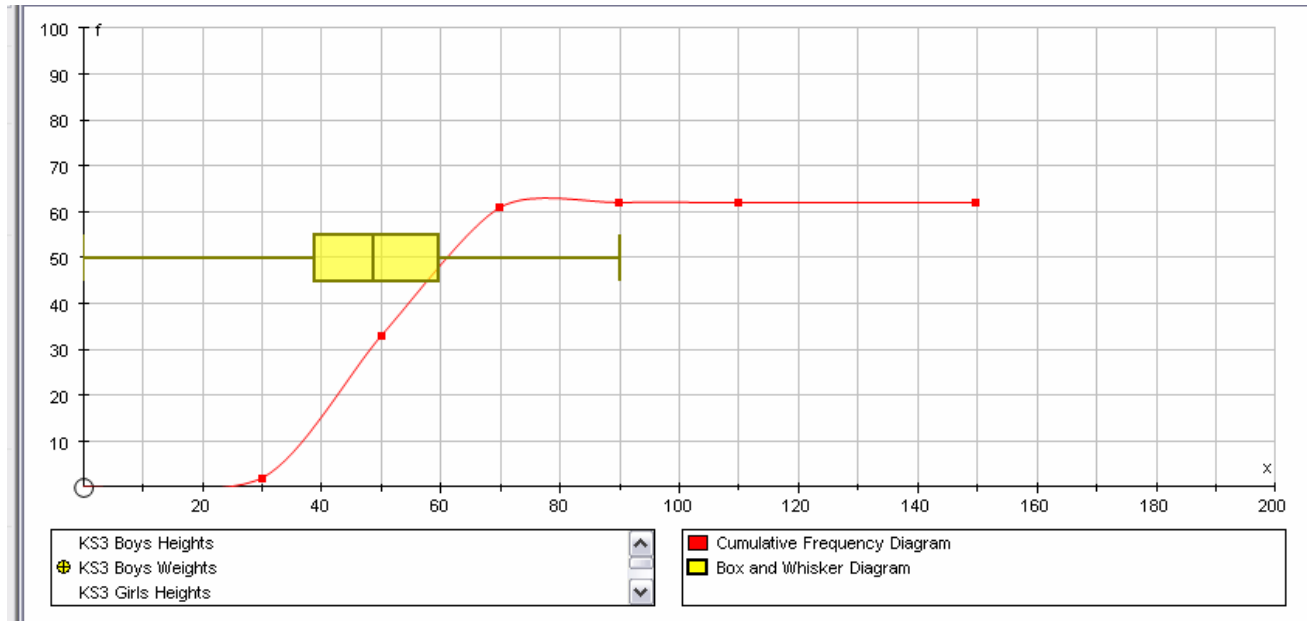
KS4 Boys Heights



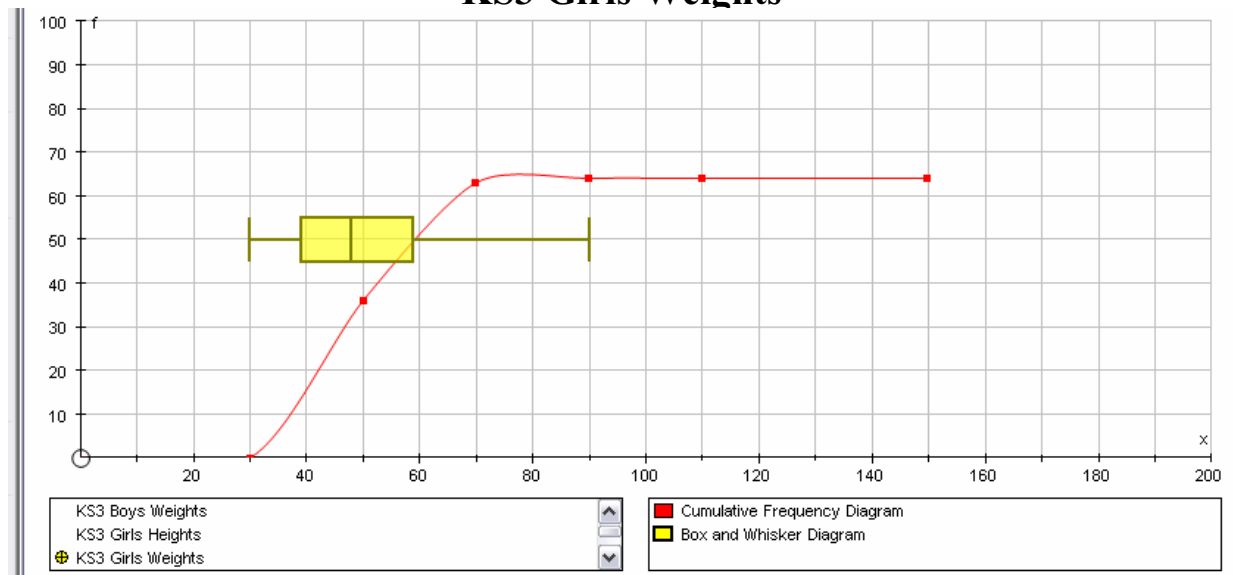
KS4 Girls Heights



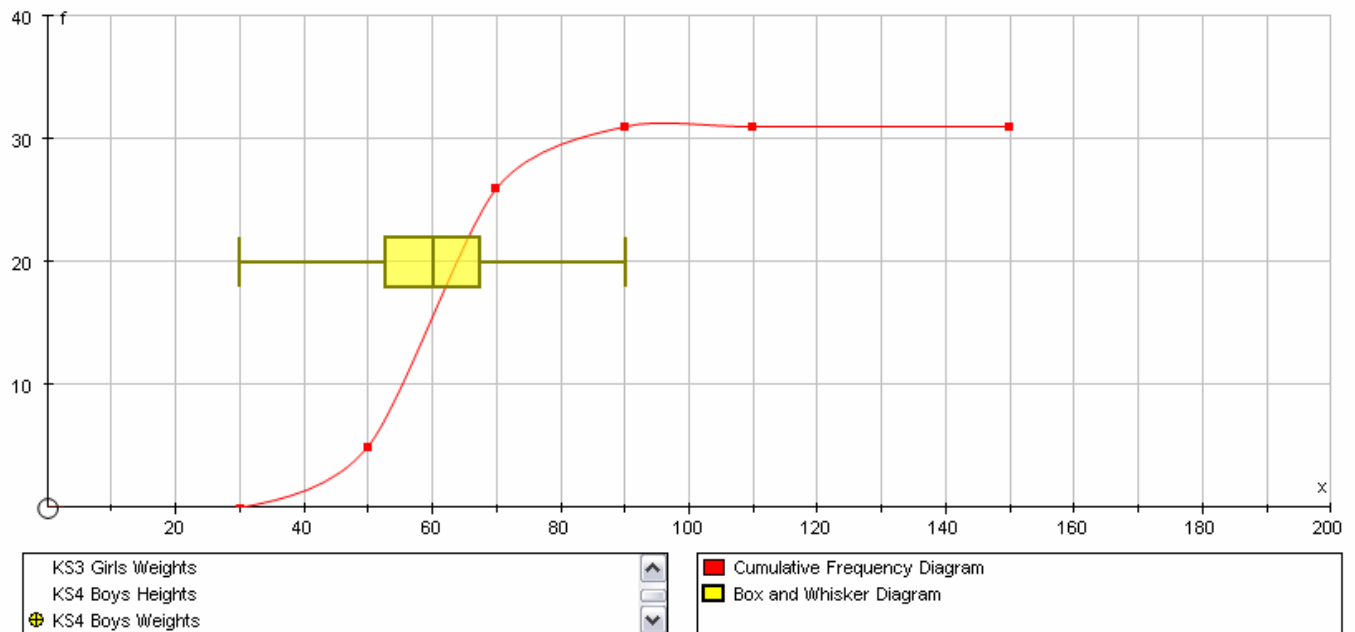
KS3 Boys Weights



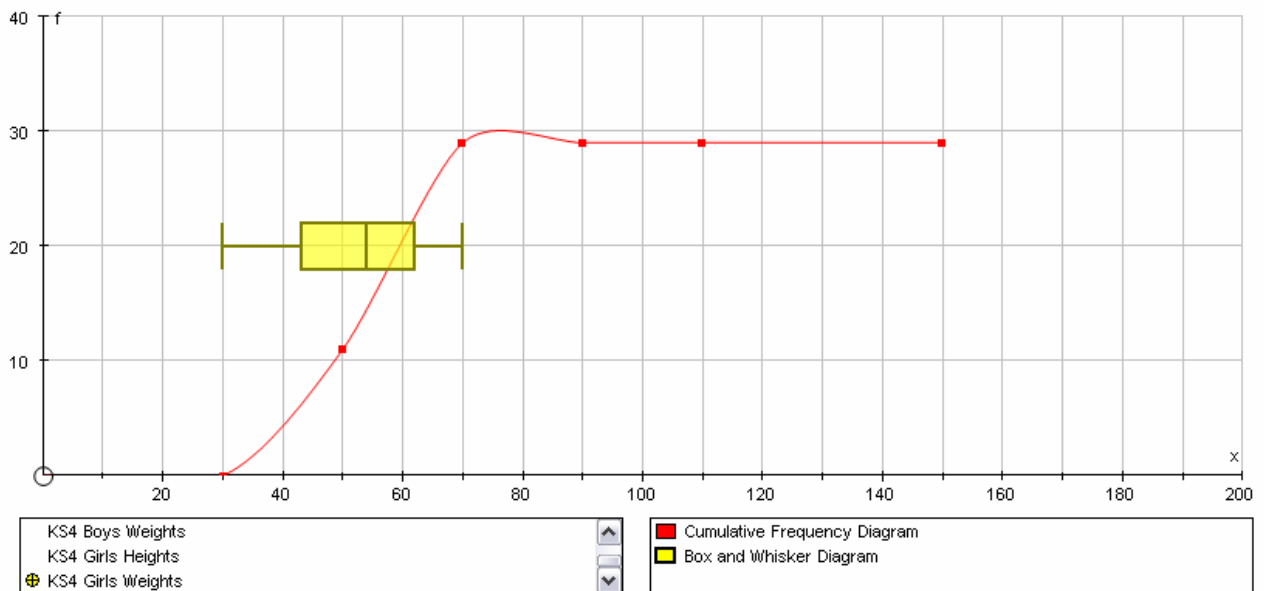
KS3 Girls Weights



KS4 Boys Weights

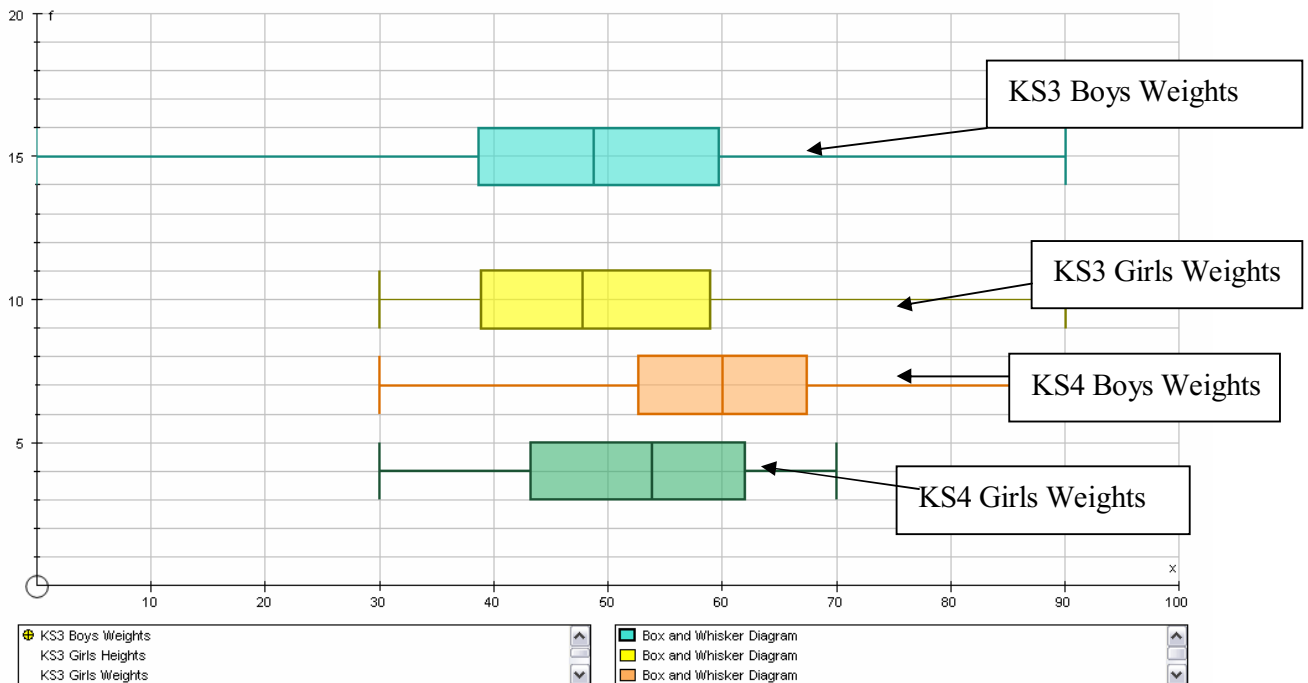


KS4 Girls Weights



By Ahsan Ahmed

Box Plot of Weights (Grouped Data)

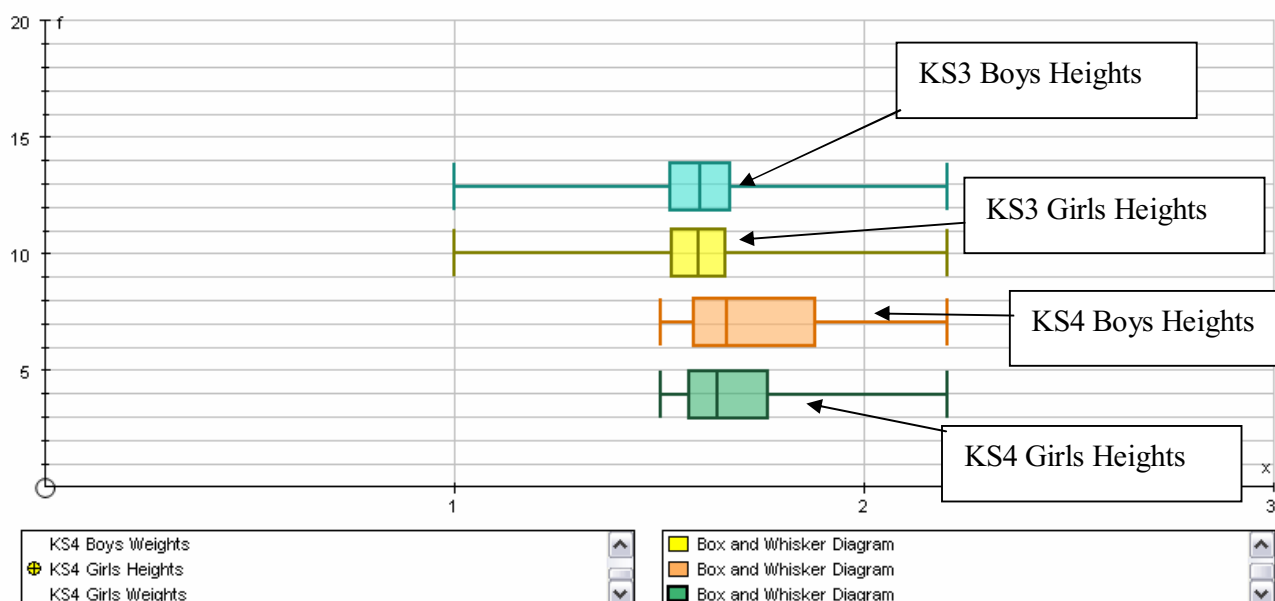


From this diagram it is clear that the KS4 Boys are the heaviest in general and that the Boy's are heavier overall. Judging by the median of each group, the lightest of all groups are the KS3 Girl's. Furthermore the lower bound value for the KS3 Girl's is considerably lower than the minimum KS4 Girl's weight. Again I think the rate of puberty has an effect on this.

Moreover it is clear that the median of each group varies and in general boys are heavier then girls. It is clear that there is a relationship between age and weight however it appears there is only a weak relationship.

Again using these graphs to make a comparison I can interpret that the range of the KS4 Girl's is much closer together than that of the other groups. I think is because they have all gone through puberty now and have fully grown. Furthermore this is emphasized as there is wide range in the KS3 boy's and this could be because they are currently going through puberty.

Box Plots of Heights (Grouped Data)



Using the box and whisker diagrams above, I can now analyse and compare the results. In accordance to the medians, boys are generally taller than girls and Key Stage 4 pupils are taller than Key Stage 3 pupils. There is a big difference in the interquartile range between each group as the KS4 Boys have a very big range. As you can see, the tallest student in most groups fall into the same data set however, Key Stage 3 tallest boy falls into the category of 1.95 to 2.1 and therefore the group appears to be taller.

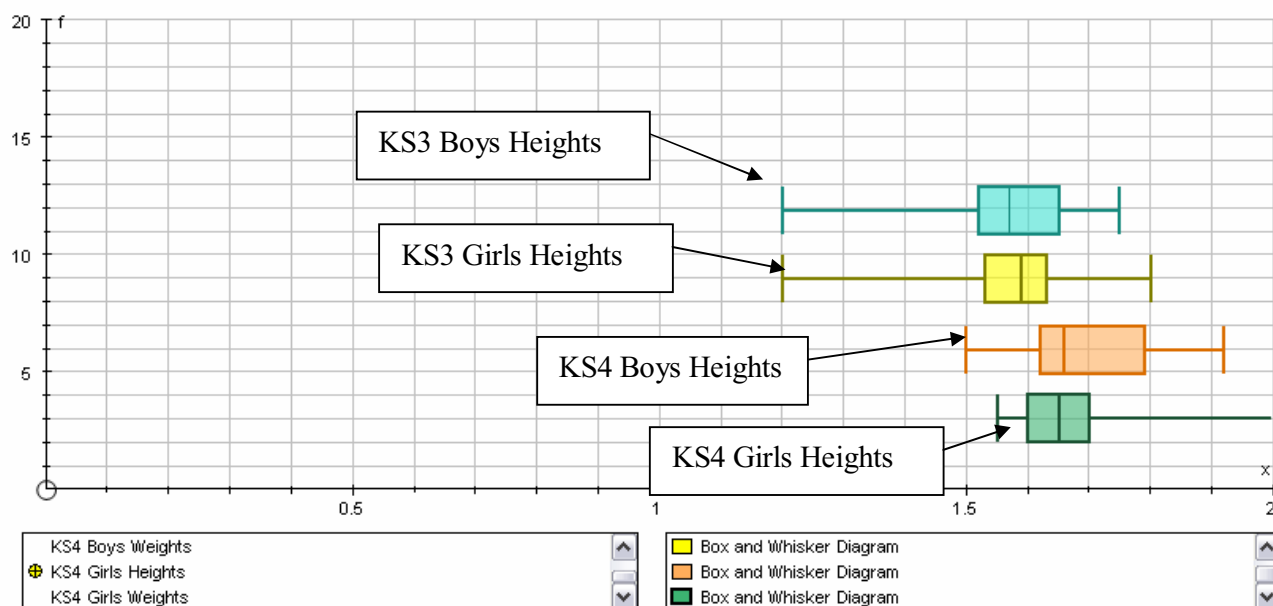
From these box and whisker plots you can see that there is not a wide variation in heights and the range of each group is very similar except for the KS4 boys' heights which are outstanding. Most whiskers start or stop at the same place, this indicates that as I have grouped the data, there is at least one person in this category.

You can see that the range of heights in KS3 Boy's is almost identical to that of the KS4 Girls. However the inter quartile range and the median value is slightly different. Due to the inter quartile range it shows that KS4 Girl's are generally taller than KS3 Boys.

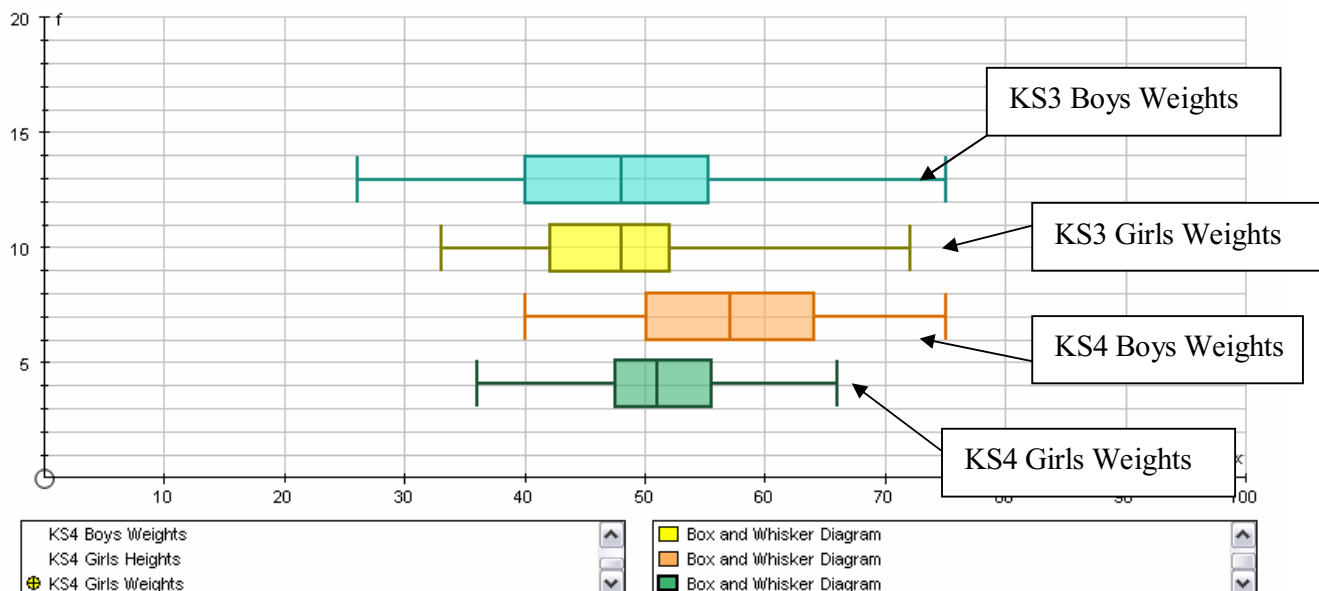
However, making this comparison seems inappropriate; it is obvious that the KS4 Girl's will be taller as they are – in youth years, considerably older.

As a result puberty would have a strong effect Also this graph shows that again KS4 Boy's are the largest and have a higher minimum range than the other groups and the highest median height.

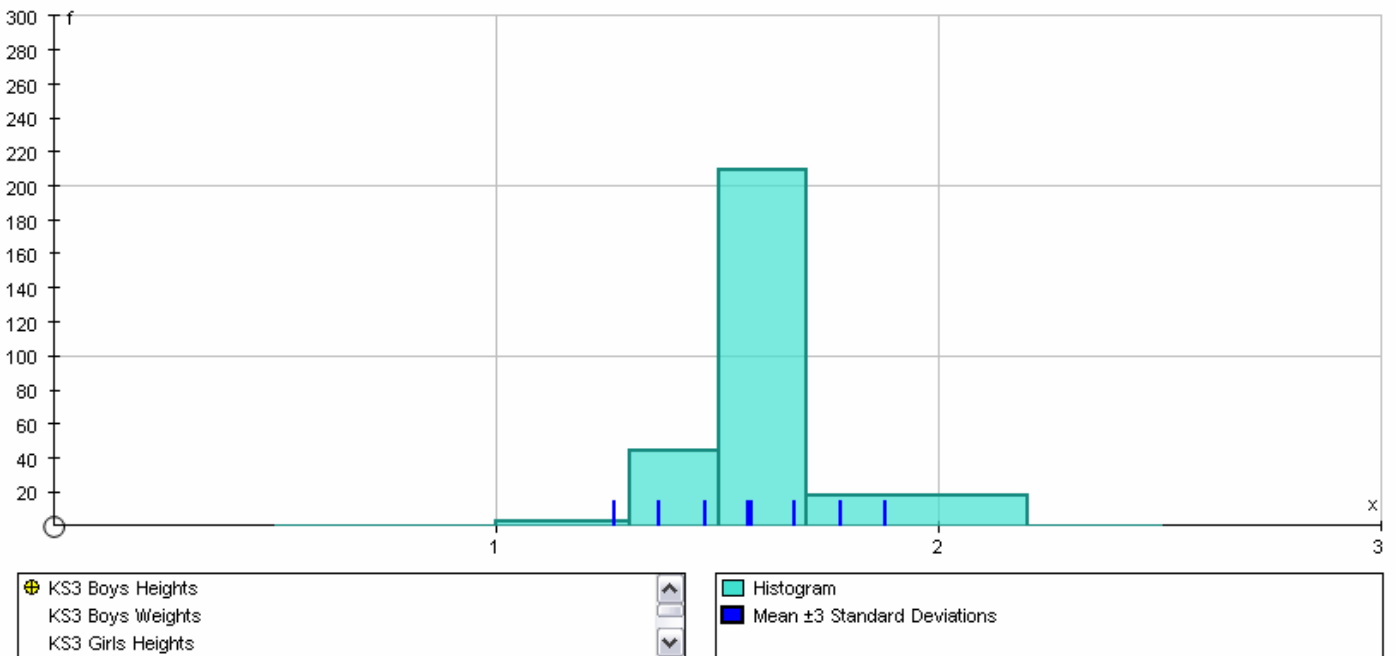
Box Plots of Heights (Raw Data)



Box Plots of Weights (Raw Data)

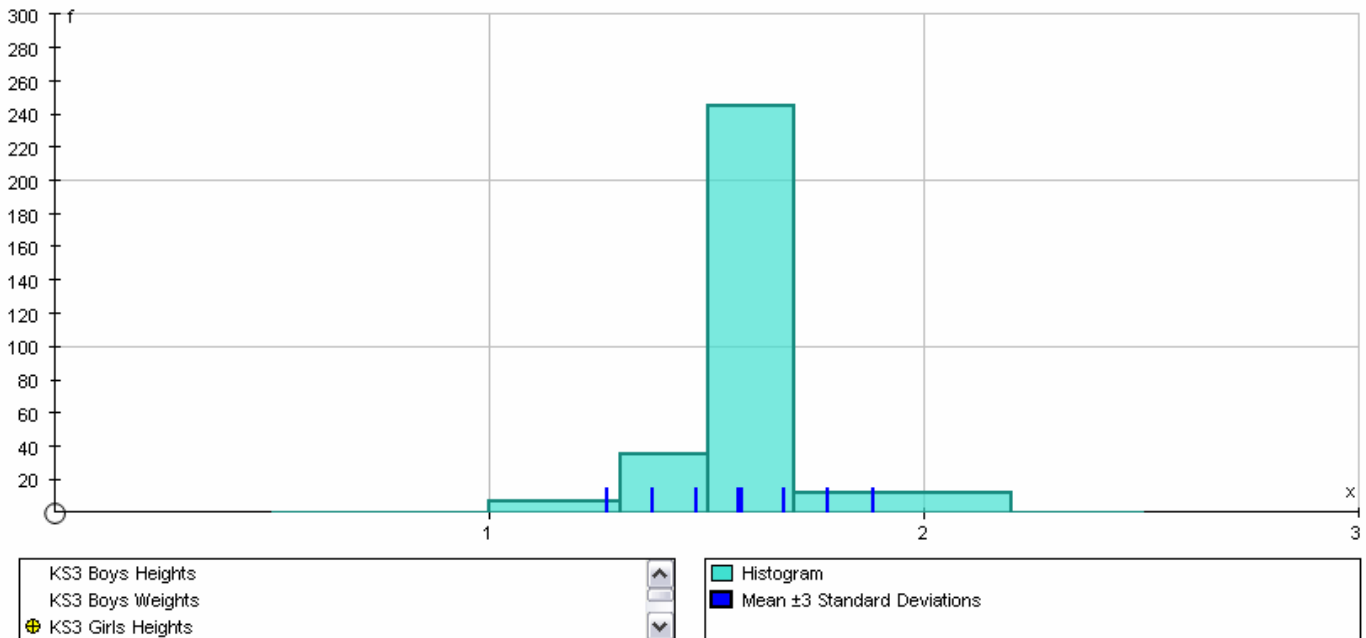


Histograms for heights KS3 Boys



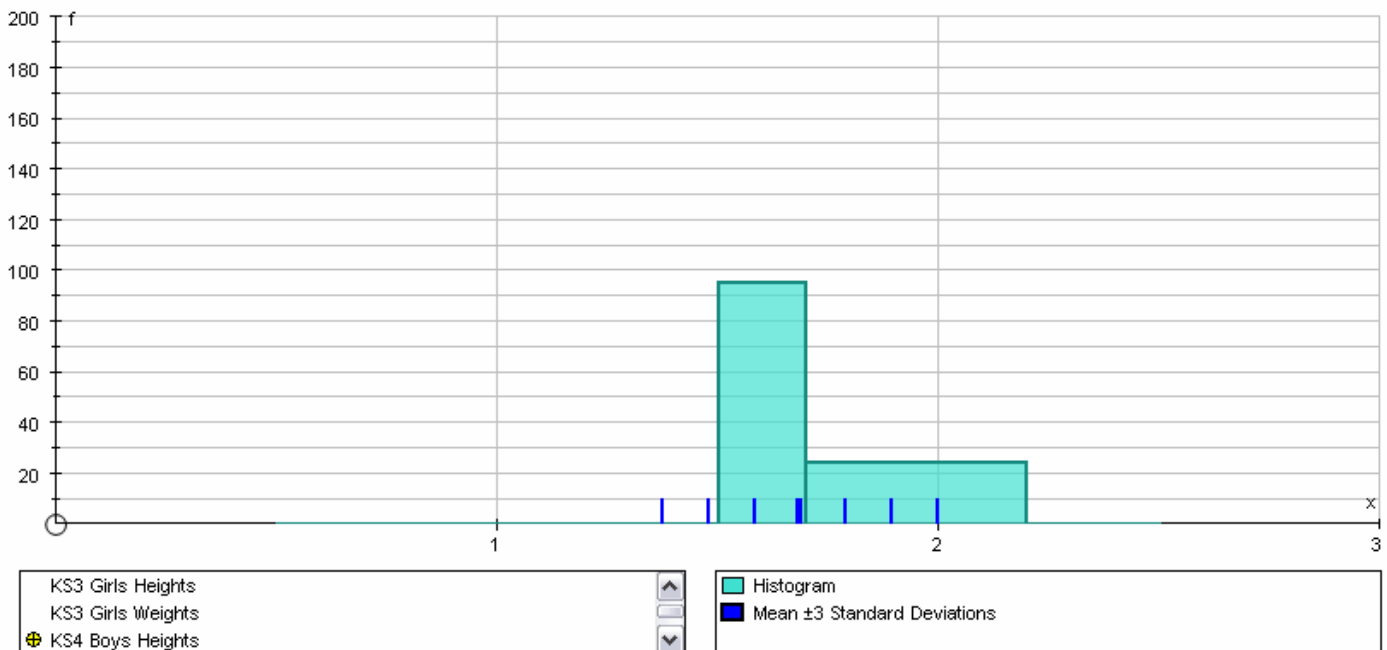
From this histogram you can infer that the data is slightly skewed to the right of the mean – 1.614. In my opinion I feel this is because at this age the boys are constantly growing and are therefore expected to have several people above the average depend on what age they reach puberty.

Histograms for heights KS3 Girls



Again using this histogram to analyze the data I can infer that the data is in general is fairly equal however a minute amount of people are in a lower group, however these could be just under the group barrier however I cannot tell. Nonetheless I feel this data is fairly equal however ever so slightly the majority of boys are in the larger group.

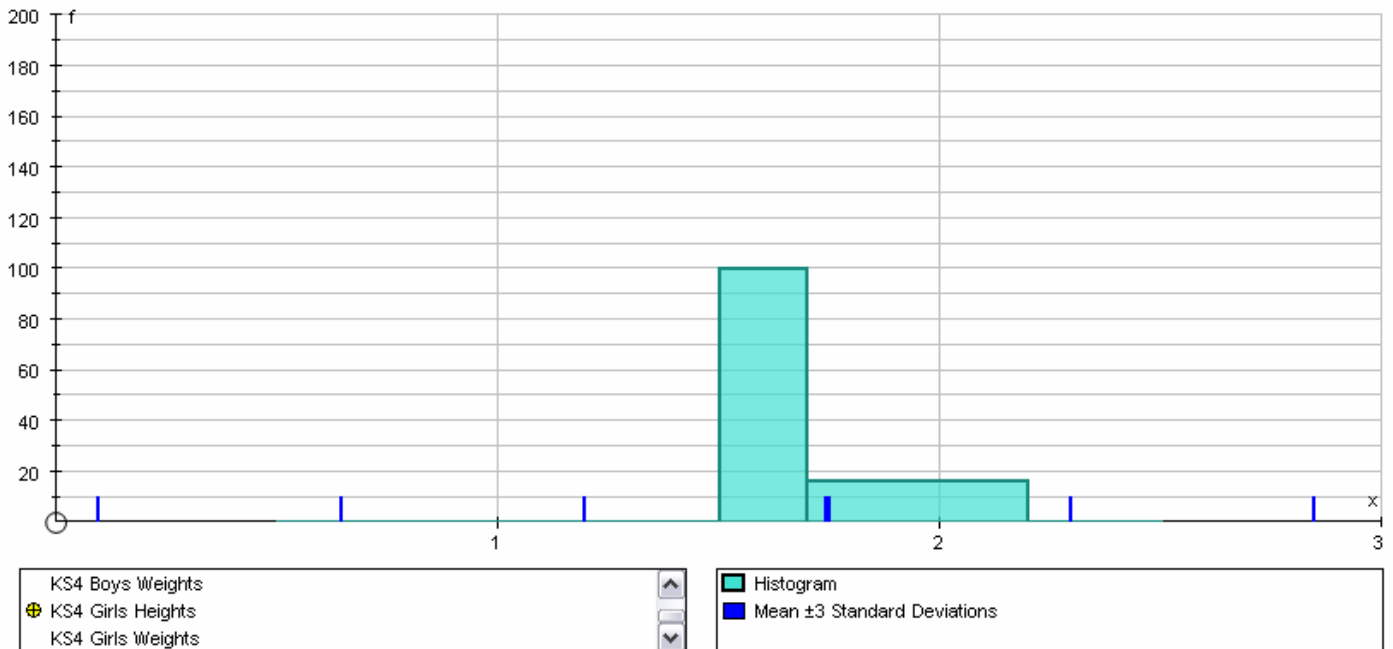
Histogram for heights KS4 Boys



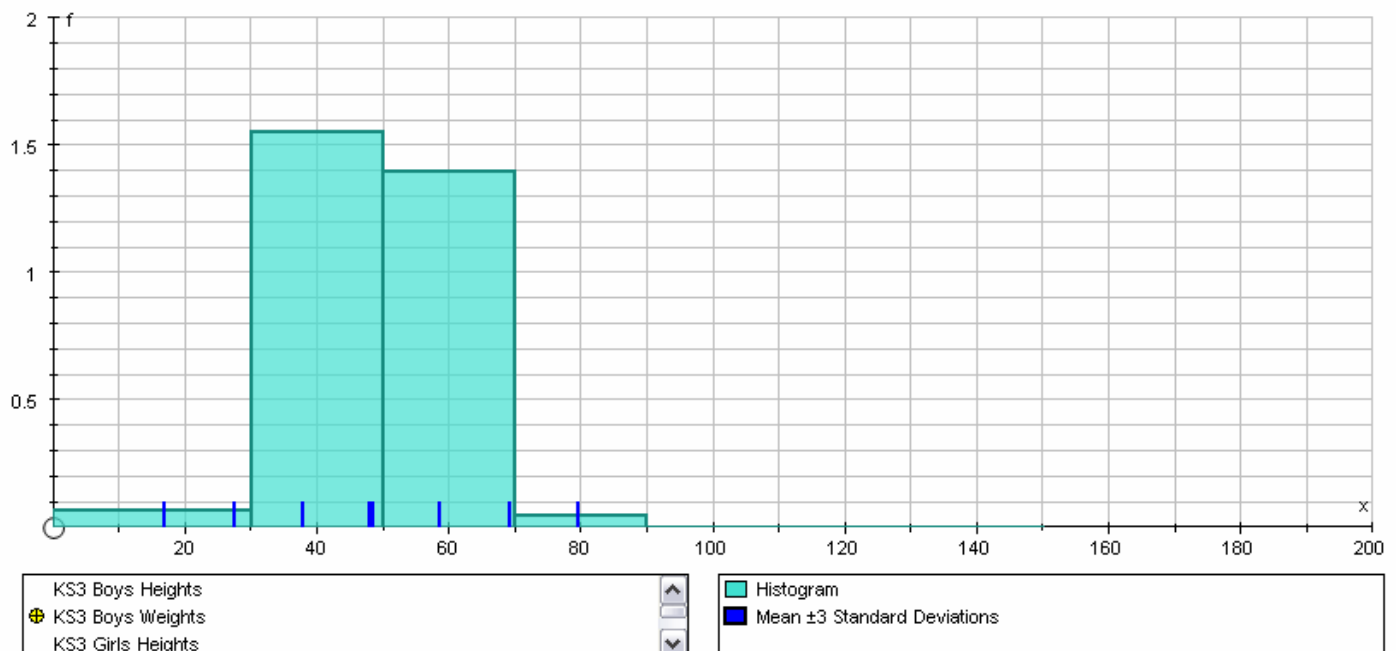
By Ahsan Ahmed

Once more, from analyzing this data I can infer that this group of data is slightly skewed to the right, indicating that the KS4 boys in general are taller and also quite varied. Again this could be because of puberty.

Histogram for heights KS4 Girls

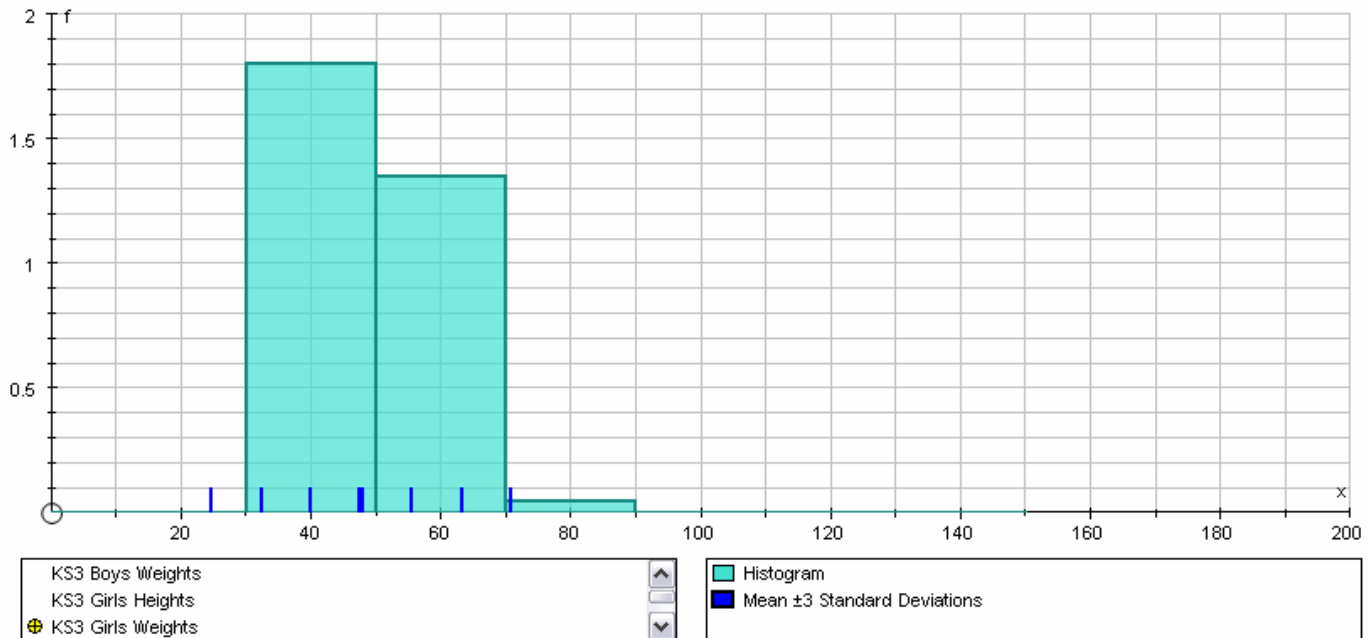


Histogram for weights KS3 Boys



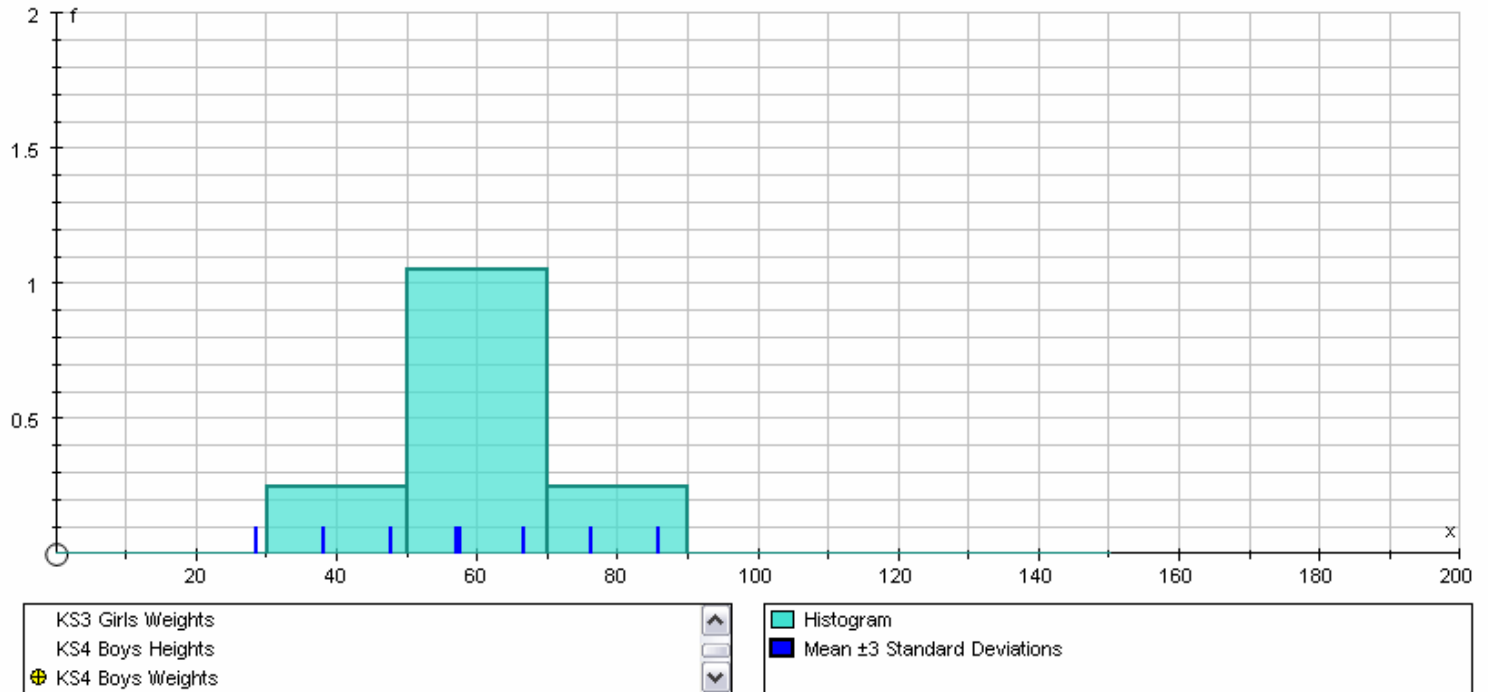
This grouped data is widely varied and doesn't follow a strong correlation, the groups with the most people in are varied and don't show a relationship. Between the weight of 30 and 70 they appear to have a good correlation except for the others which are low consistently.

Histogram for weights KS3 Girls Weights



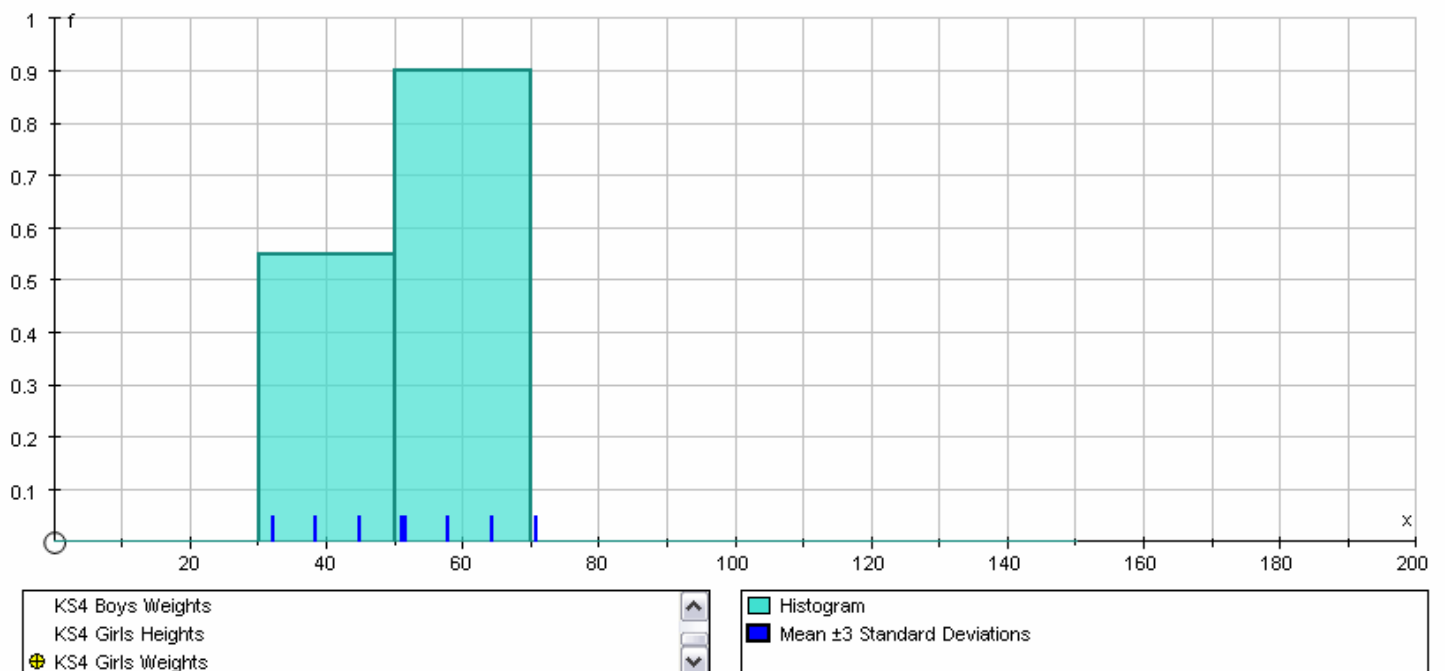
Again the histogram shows that the weights are equally linked in two regions between 30 and 70. When analyzing this data I can infer the data is slightly skewed to the right, indicating that the best part of this sample is heavier than the calculated to the mean of the data.

*Histogram for weights **KS4 Boys***



As you can see in this data there are groups in the table that don't contain any students, however there are groups either side with data in. This indicates the data in this group is widely varied. Furthermore using the calculated mean of 63.75, the data is largely skewed and doesn't show a relation; therefore it is almost impossible to predict a weight for that age category.

*Histogram for weights **KS4 Girls***



This data is fairly equal and shows a strong relationship, it is collated together and shows a strong relationship in the weight of KS4 girls' weights. Also compared to the mean of 54.73 it is clear that the data is equally balanced and shows a correlation. You would be able to make a general prediction of this groups' weights by their age.

Using these Histograms I am able to compare each group using their frequency density. Within the program Autograph, where the data is, I selected to create a histogram. I did this by selected each data group and selecting the appropriate button. I then selected to draw these histogram's calculating the frequency density of each group – autograph calculated this. To calculate the frequency density is:

$$\text{Frequency Density} = \frac{\text{Frequency}}{\text{Class width}}$$

As you can see, in these graphs the data is widely ranged on the Y axis. The numbers are varied because of the frequency density for each group is so different; it has been scaled to an appropriate number for each group. Each graph shows varied results, but nevertheless they all have a similar correlation.

Furthermore I am able to make a comparison through analysing the distribution of the data in the above graphs. From glancing at them you are able to distinguish each class width and the whole range of data. As you can see, in the KS3 Boy's I have added a midpoint and from this I am able to interpret that the data is slightly skewed to the right. In my opinion I believe this is because KS3 boys vary massively in their size and weight due their age and their bodies beginning to change. Each male may hit puberty at a different age and it can affect them differently therefore I believe that as my data is slightly skewed to the right that the majority of KS3 boys have started puberty.

In addition comparing this graph to that of KS4, it again shows the data is slightly skewed to the right. This therefore indicates that for boys overall there is a weak positive correlation, indicating that there is a strong relationship between height and weight.

Conclusion

Now I have a basic idea of the relationship between height and weight of adolescents attending secondary school.

From these results you can infer that the relationship between KS3 Boy's has a positive correlation as it is 0.5, therefore this proves that the relationship in KS3 Boy's is fairly strong.

Also when looking at the results I have analyzed the KS3 Girl's results and this shows that the relationship in their height and weight is a positive correlation however very weak.

The KS4 Boy's relationship shows a fairly strong positive relationship, with it being 0.6. This proves that for this gender and age there is a positive, strong relationship between height and weight.

As you can see the KS4 Girl's shows a weak negative relationship, indicating that there is little correlation for the students in this category.

Overall in this coursework I have established that there is relationship between people's heights and their weights. This relationship however is much stronger in boys overall in both age groups. This is emphasised as KS4 Girl's show a negative relationship between the two factors. The correlation of KS4 girl's is negative therefore this shows that as height increases weight decreases.

My initial prediction was that as height increases weight also increases; I saw this in three of my four data sets; the only one this did not follow for was the KS4 girls. I also predicted that the relationship would be stronger in boys than in girls. My results show me that this was true in both cases. From these results I can make statements that:

“Boys heights and weights are very closely related”

This is proven in the scatter graphs I inserted above when I inserted the line of best fit as it follows a similar proportional ratio – as height increased weight also does.

“Girls heights and weight show very little correlation”

Again referring to the scatter graphs inserted when I collated the two factors together, the KS3 Girls shows a very weak positive correlation and the KS4 Girls show a negative correlation.

“The age can also affect these results especially in girls”

As shown above again in the scatter graph the ks3 Girls show a very weak positive correlation, however the KS4 Girls show a negative relationship. Furthermore comparing these results to the boys this emphasises this point.

Evaluation

In this piece of coursework, I have learnt that predictions are not always right and that to make sure that I follow up every thing that I predict to make sure it is correct. I feel if I was to do the coursework again, I will definitely improve my hypothesis and try to make every thing accurate as possible.

Overall, I enjoyed doing this coursework and learnt a lot of valuable skills in doing so and I have established that there is relationship between people's heights and their weights. My knowledge of Microsoft Excel has increased, and I have learnt how to use the Scatter Graph software.

I completed the scatter graphs, histograms and charts in quick succession and then followed them up with a explanation of the charts and what they showed.

To conclude, I feel I have completed this coursework to the best of my ability, without the help of anyone apart from my maths teacher.