Forming a Conclusion

Forming a conclusion means answering the investigative question. You need to describe the relationship between the two given variables and state (with reasoning) whether or not that relationship makes sense. Include references to your analysis.

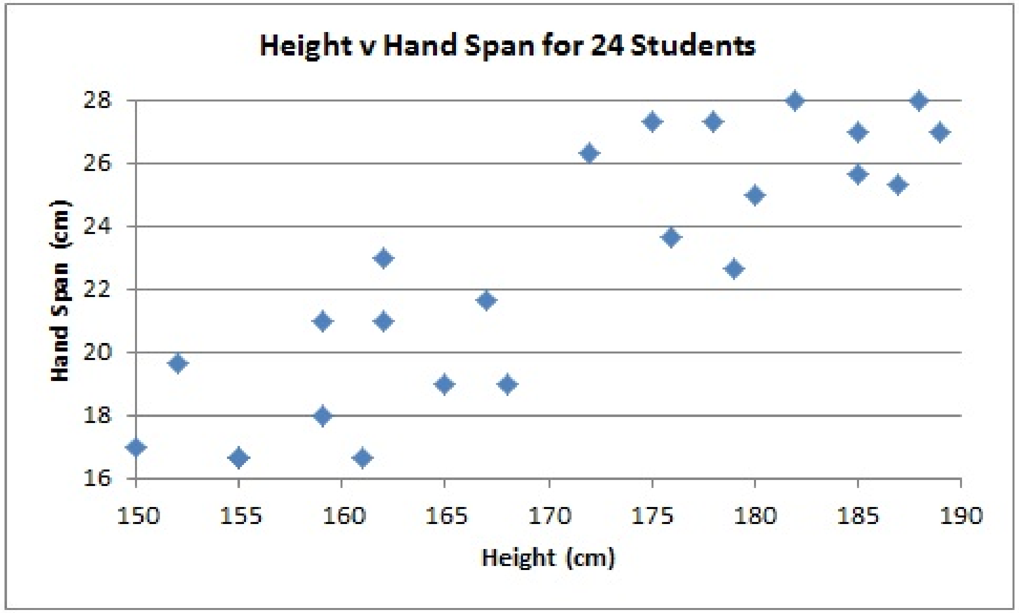
We use the sample statistics and graphs to make statements about the population. We need to be aware that the sample will vary from the population. If the sample shows that there is a relationship between the two variables AND the sample size is reasonable (at least 30) then we can say there probably is a relationship in the population as well.

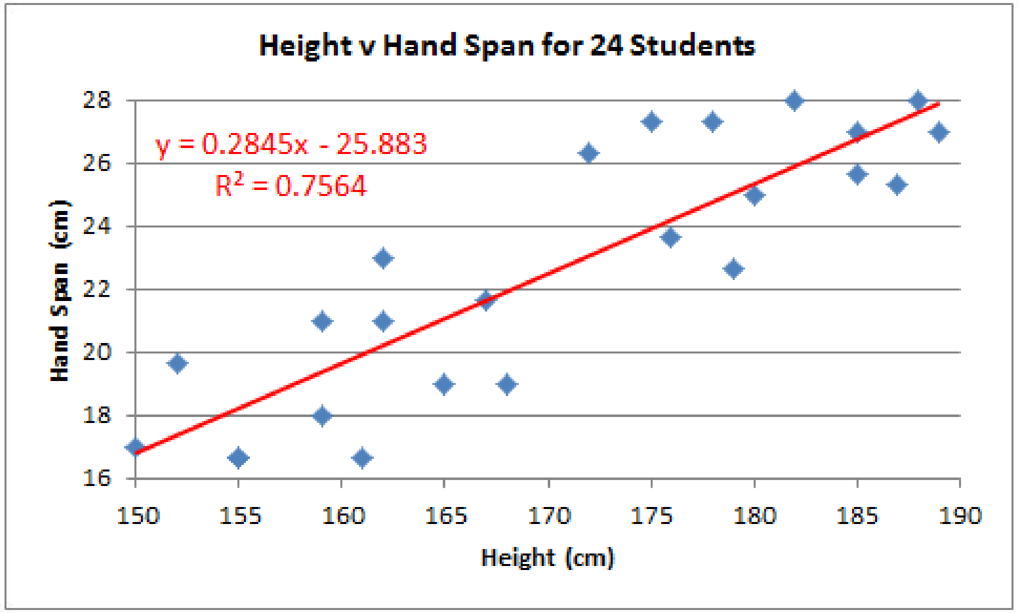
Now you are ready to form a conclusion, which means answering the investigative question. You need to describe the relationship between the two given variables and state (with reasoning) whether or not that relationship makes sense. Include references to your analysis.

For some investigations ¡t might be appropriate to use your trend line to make predictions. This is not essential for the award of any grade for this standard but would demonstrate understanding of a different aspect of analysing bivariate data. If you do this you need to use an x-value to estimate the corresponding y-value. You should discuss how sensible your prediction is.

Forming a Conclusion

Example 1: I wonder if there is a relationship between height and hand span for high school students.





Conclusion

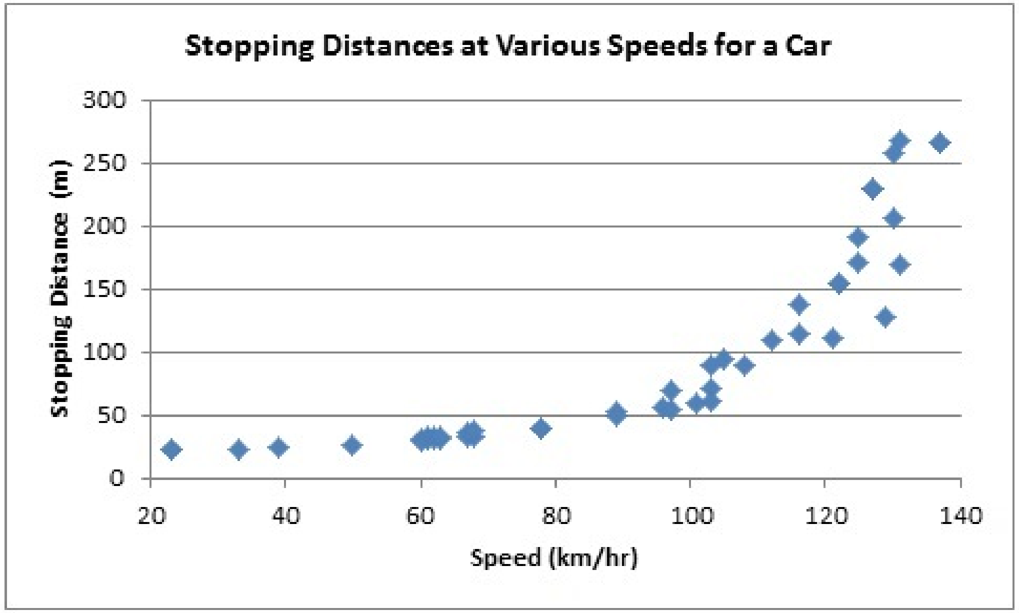
Taller students tend to have longer hand spans. This can be seen by the positive slope of the trend line, as the heights get greater then hand spans get longer. This result is not very surprising. Taller people tend to have greater measurements for other body lengths (arm lengths, leg lengths, etc.) and so you would expect them to also have greater hand spans.

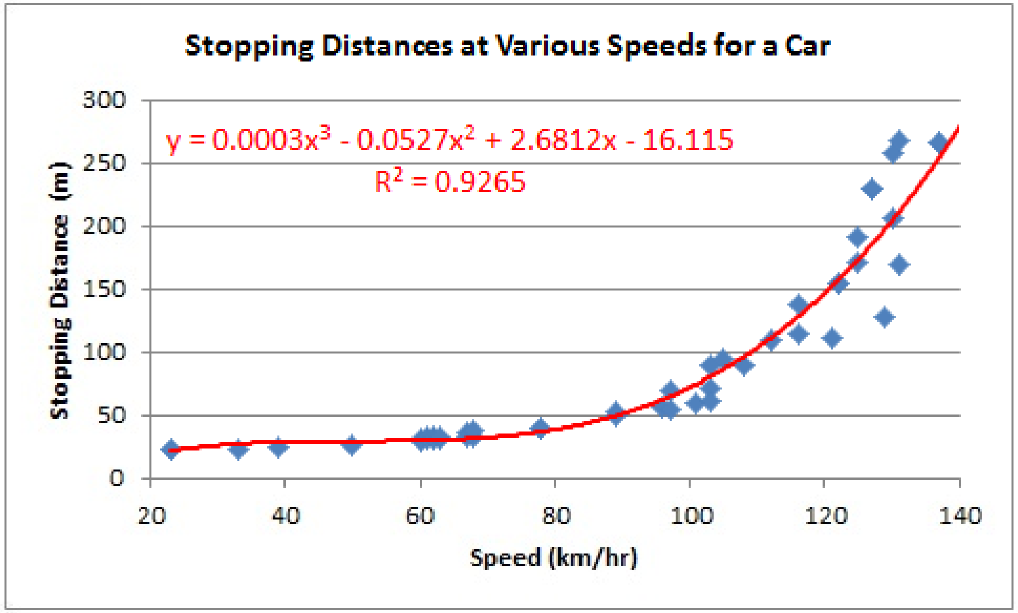
Prediction

The trend line is $y = 0.2845x - 25.883$. $0.2845 \times 157 - 25.883 = 18.7835$. Therefore I predict that a 157cm tall student will have a hand span of about 19cm. Because the relationship is moderate to strong ($r = 0.8697$) I think the estimate will be reasonably accurate for this student. However I notice that the variation is typically about ±2cm for the heights in the sample. Therefore I expect that there will be an error of about ±2cm in my estimate of 19cm. If the two groups are substantially girl and boy, or junior and senior, a more accurate estimate might be obtained by modelling females and males or juniors and seniors separately, as you can't assume the two groups are solely female and male, or junior and senior.

Forming a Conclusion

Example 2: I wonder if there is a relationship between the speed of a car and its stopping distance.





Conclusion

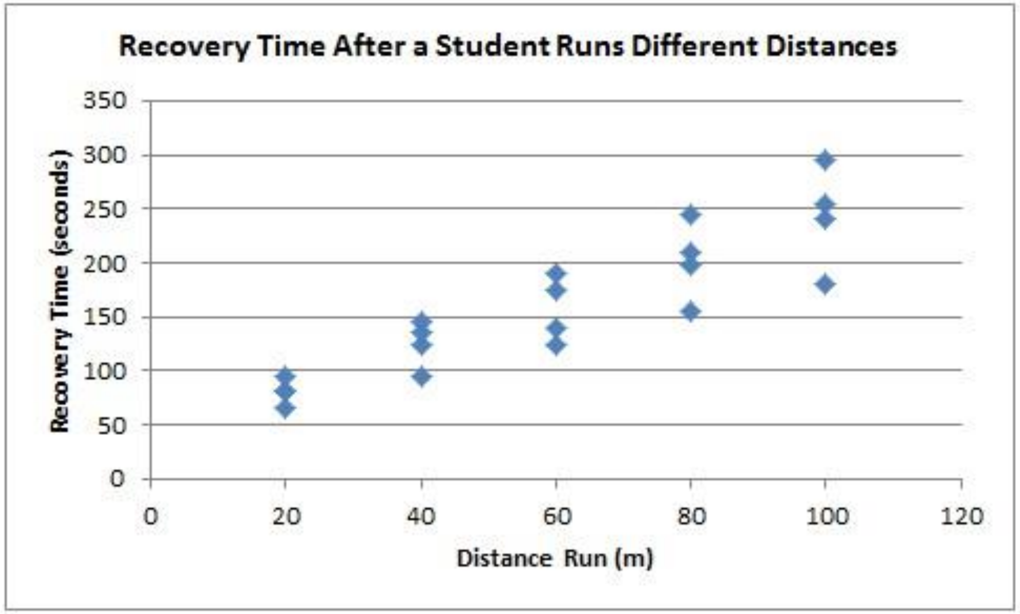
Cars travelling at higher speeds tend to have longer stopping distances. This can be seen by the way stopping distances are greater with higher speeds. However it is not a linear relationship, the rate at which the stopping distance changes increases as speeds increase. You would expect stopping distances to increase with increasing speed. Faster cars take longer to stop.

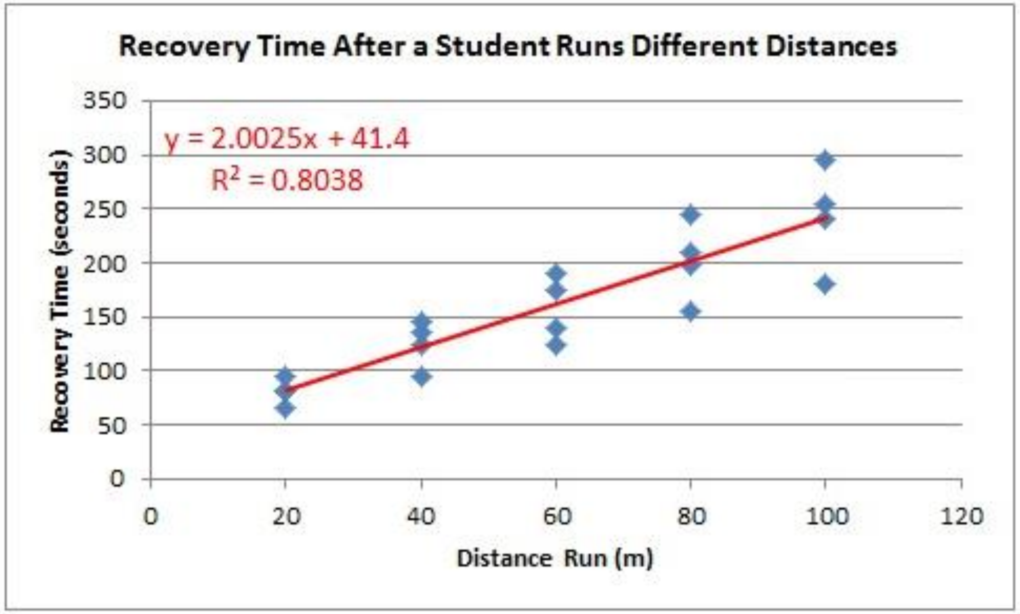
Prediction

The curve is $y = 0.0003x^3 - 0.0527x^2 + 2.6812x - 16.115$. $0.0003 \times 85^3 - 0.0527 \times 85^2 + 2.6812 \times 85 - 16.115 = 15.267$. *However by eye the value from the graph looks like about 40. Therefore I predict that when the car is travelling at 85km/hr the stopping distance will be about 40m. I put the difference between 15 and 40 down to rounding effort in the model as the computer has stated it. At 85 km/hr I think this model will be reasonably accurate because the dots are all close to the curve in that area of the graph. I don't think this model will work at speeds greater and 140km/hr as there is no reason to believe the dots will continue to follow this particular curve. I also think this model will only be reliable for the make, model, tyres and perhaps even driver used in this trial. These are all factors that could change stopping distance significantly.*

1. What would you write in your "Conclusion" section? Answer your investigative question and make a prediction.

I wonder if there is a relationship between the distance a student runs and the amount of time it takes him/her to recover after the run.





Conclusion

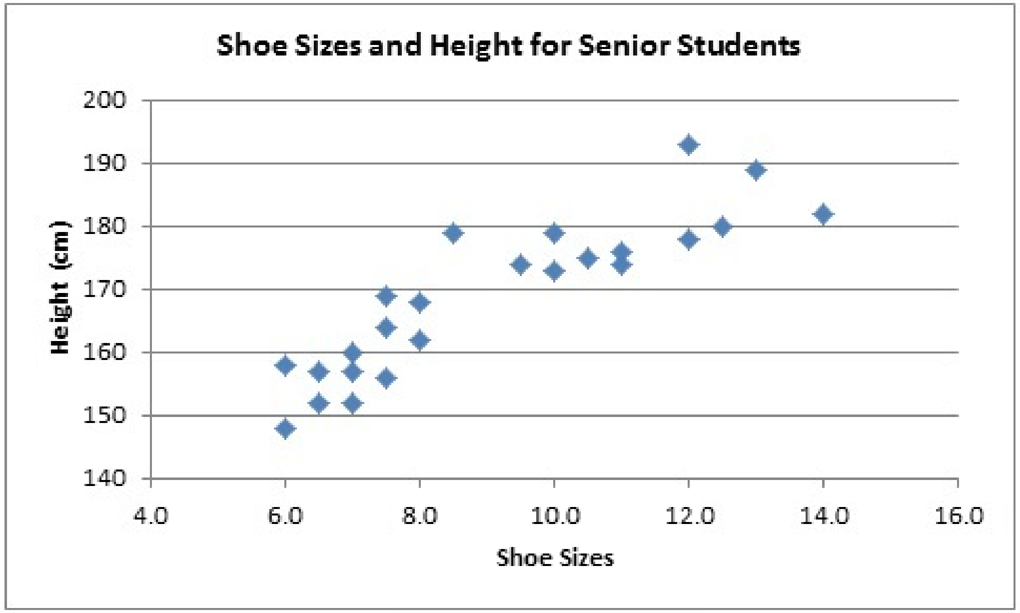
The student tends to take longer times to recover after he runs longer distances, This can be seen by the positive gradient of the trend line, You would expect this because the student's body would work harder if he ran further, resulting in a longer time for his heart rate to return to normal. This conclusion applies only to this student but in general people would take longer times to recover after running longer distances.

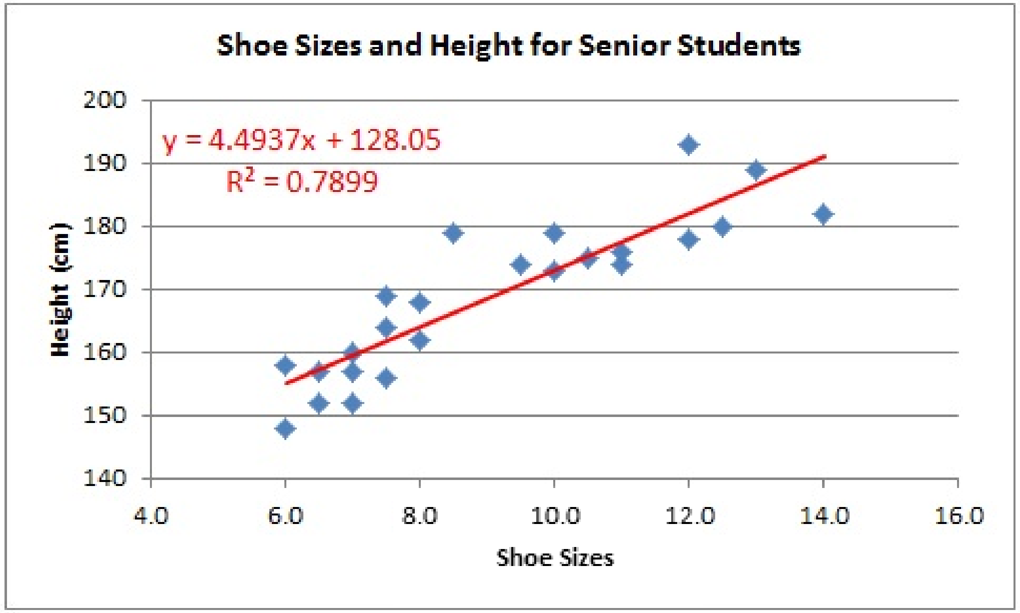
Prediction

The trend line is $y = 2.0025x + 41.4$. $2.0025 \times 50 + 41.4 = 141.525$. *Therefore I predict that if this student runs 50m his recovery time will be about 142 seconds. Because the relationship is strong* ($r = 0.8965$) *I think the estimate will be reasonably accurate for this student. However I notice that for distances 40m and 60m there is variation of about 60 and 80 seconds respectively. Therefore I can expect that there will be an error of about ±35 seconds in my estimate of 142 seconds. I do not expect this prediction to be useful for anyone else as variables like age, fitness and natural athleticism will cause huge variation in results from person to person.*

2. What would you write in your "Conclusion" section? Answer your investigative question and make a prediction.

I wonder if there is a relationship between the shoe sizes and the heights of senior students.





Conclusion

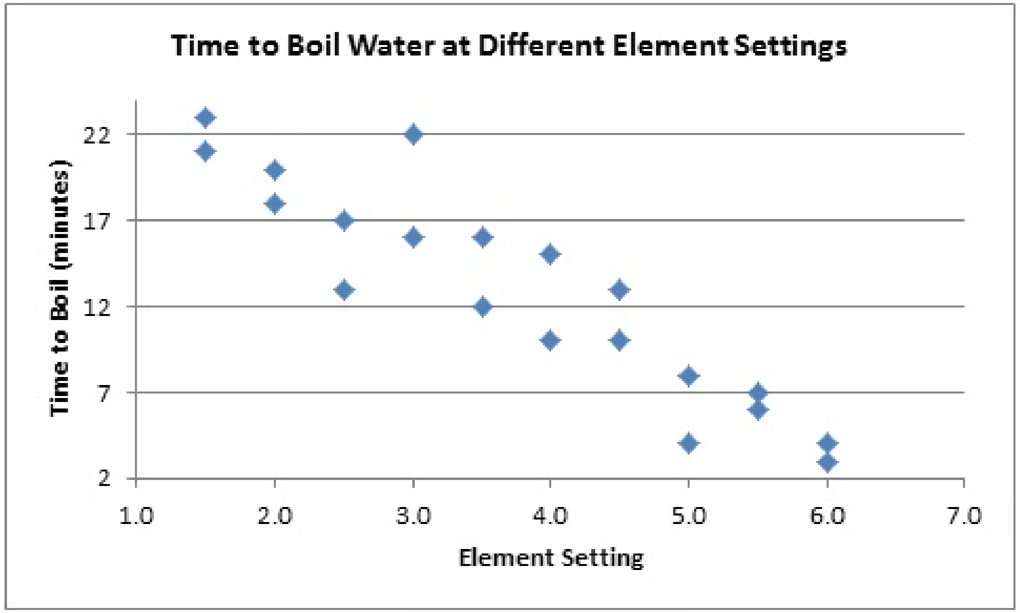
Senior students with larger shoe sizes tend to be taller. This can be seen by the positive gradient of the trend line. This result is not surprising, taller students would usually have greater body measurements including their foot size, which would mean greater shoe sizes.

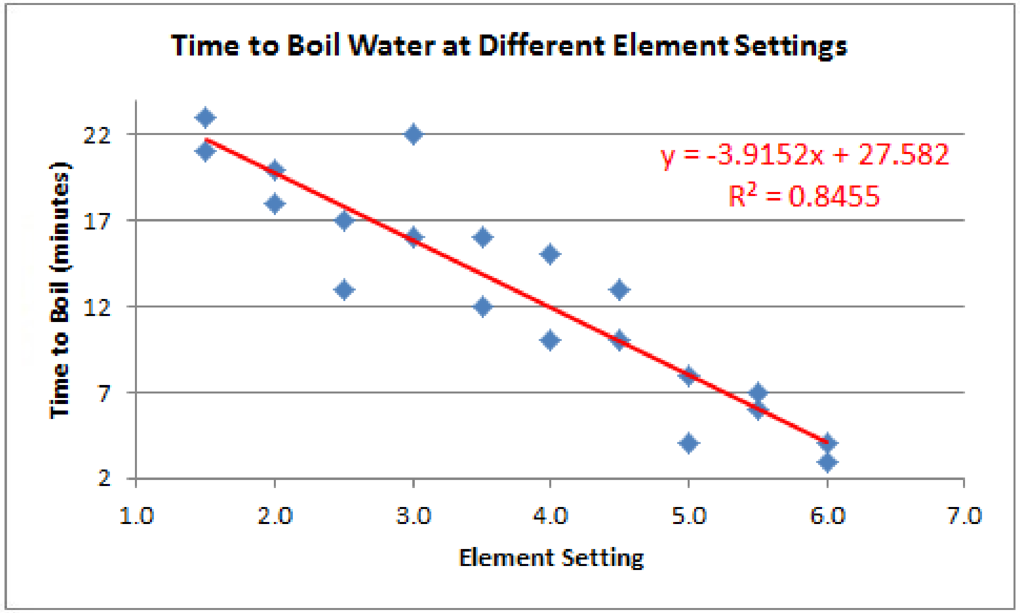
Prediction

The trend line is y = 4.4937x + 128.05. 4.4937 x 11.5 + 128.05 = 179.72755. Therefore I predict that a student with shoe size 11½ will be about 180cm tall. Because the relationship is strong (r = 0.8888) I think the estimate will be reasonably accurate for senior students. However I notice that the variation is typically about ±5cm for the shoe sizes in the sample. Therefore I expect that there will be an error of about ±5cm in my estimate of 180cm. If the two groups are substantially female and male, a more accurate estimate might be obtained by modelling females and males separately as you can't assume the two groups are solely female and male.

3. What would you write in your "Conclusion" section? Answer your investigative question and make a prediction.

I wonder if there is a relationship between the element setting on the stove used to boil water and the time it takes for the water to boil.





Conclusion

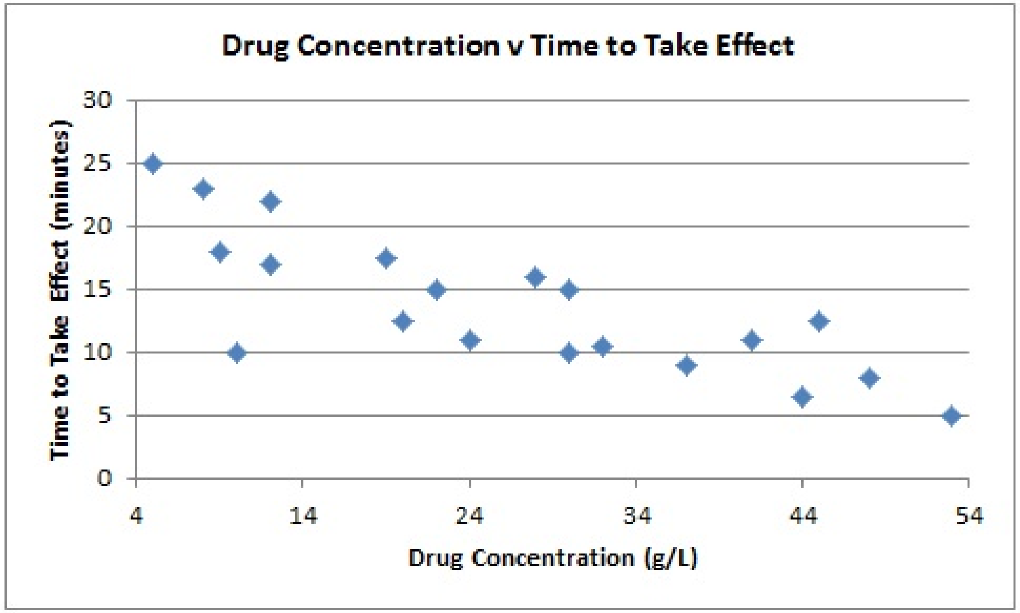
Pots on higher element settings tend to take shorter times to boil. This can be seen by the negative gradient of the trend time. You would expect this result as higher element settings produce more heat and this reduces the time for the pot to boil.

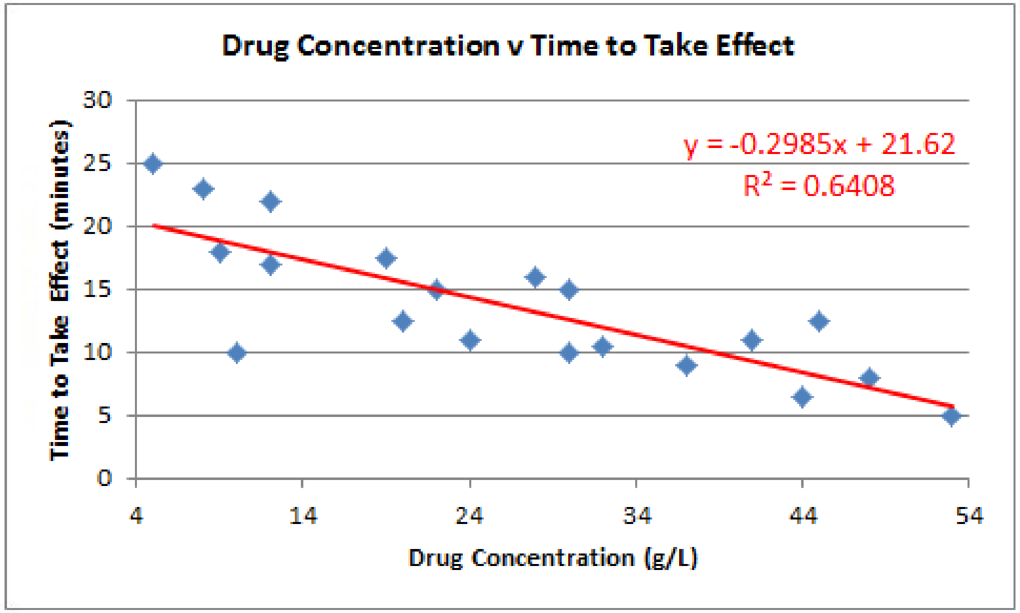
Prediction

The trend line is y = -3.9152x + 27.582. -3.9152 x 7 + 27.582 = 0.1756. Therefore I predict that the water will take 0.1756 minutes or about 11 seconds to boil if the element is set to 7 (presuming it has such a setting). Although the relationship is strong (r = -0.9195) I don't think this estimate will be very reliable. This is because 7 is outside the domain of element setting values in the sample. The model gives a time to boil value of 0 minutes when the element setting is 7.04. This illustrates that the model is not reliable outside the domain of element setting values in the sample. Even within the domain there is typically ±3 minutes of variation on the dataset so it is clear that a prediction of 11 seconds has no degree of reliability.

4. What would you write in your "Conclusion" section? Answer your investigative question and make a prediction.

I wonder if there is a relationship between the concentration of a drug and the time it takes for the drug to take effect.





Conclusion

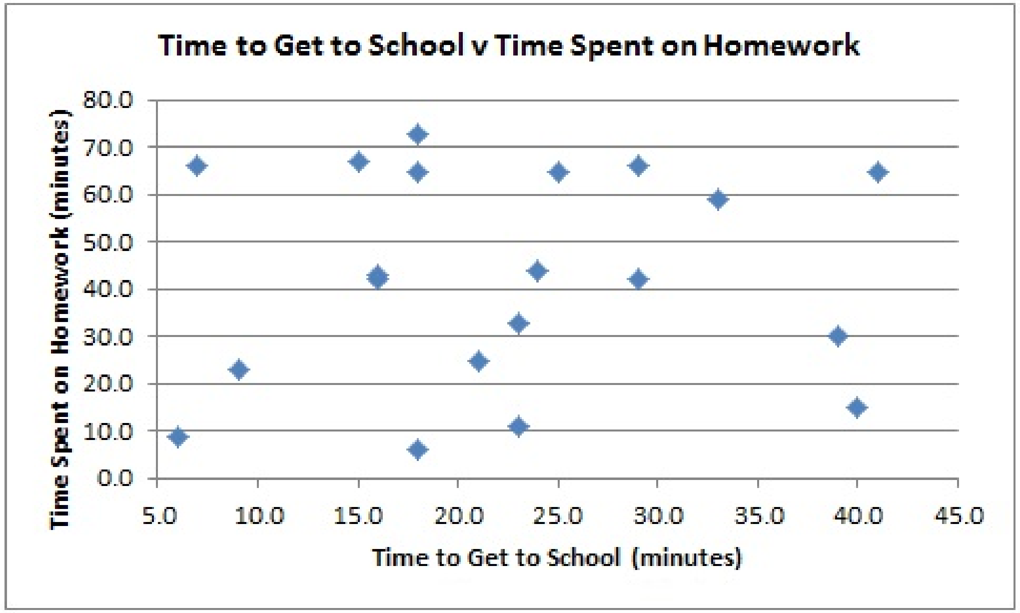
Higher drug concentrations tend to take shorter times to take effect. This can be seen by the negative gradient of the trend line. While you need to know a lot about the drug to be sure, you would expect using a higher concentration would result in a faster reaction. I am not sure if the 24 data points in this trial come from 24 different people but I expect they do and if so I expect that such things as weight, fitness and gender might be variables that would affect the relationship between drug concentrations and time to take effect.

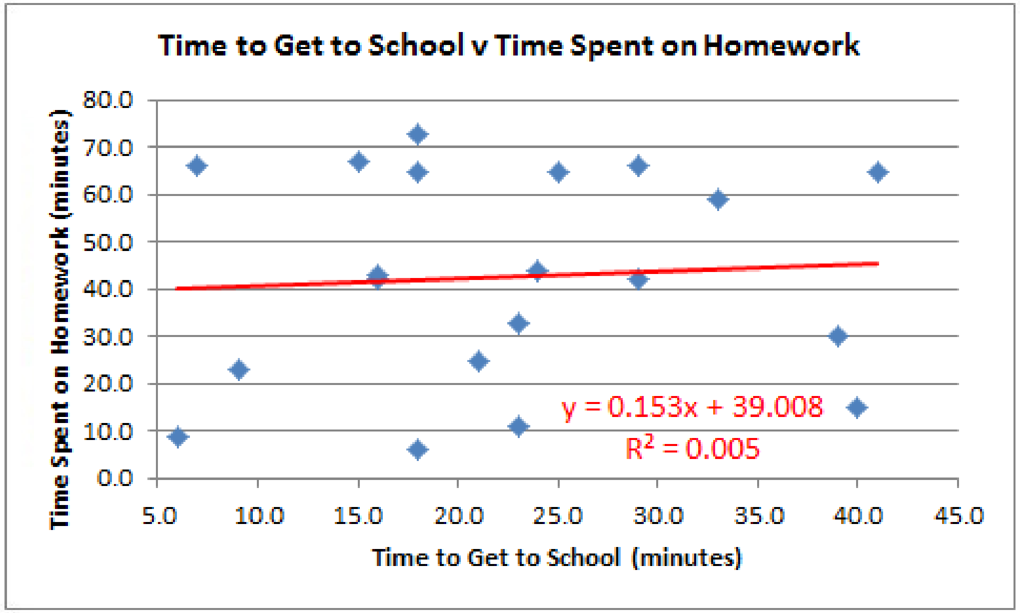
Prediction

The trend line is y = -0.2985x + 21.62. -0.2985 x 15 + 21.622 = 17.1445. Therefore I predict that at a concentration of 15 g/L the drug will take about 17 minutes to take effect. Because the relationship is moderate to strong (r = 0.8005) I think the estimate will be reasonably accurate for the drug used in this study. However I notice that the variation is typically about ±3 minutes for the show sizes in the sample. Therefore I expect that there will be an error of about ±3 minutes in my estimate of 17 minutes. This prediction is obviously only useful for this drug. Also factors like weight, fitness and gender as mentioned above make the model unreliable for predictions for any random person.

5. What would you write in your "Conclusion" section? Answer your investigative question and make a prediction.

I wonder if there is a relationship between the time students take to get to school each day and they amount of time they spend on homework each day.





Conclusion

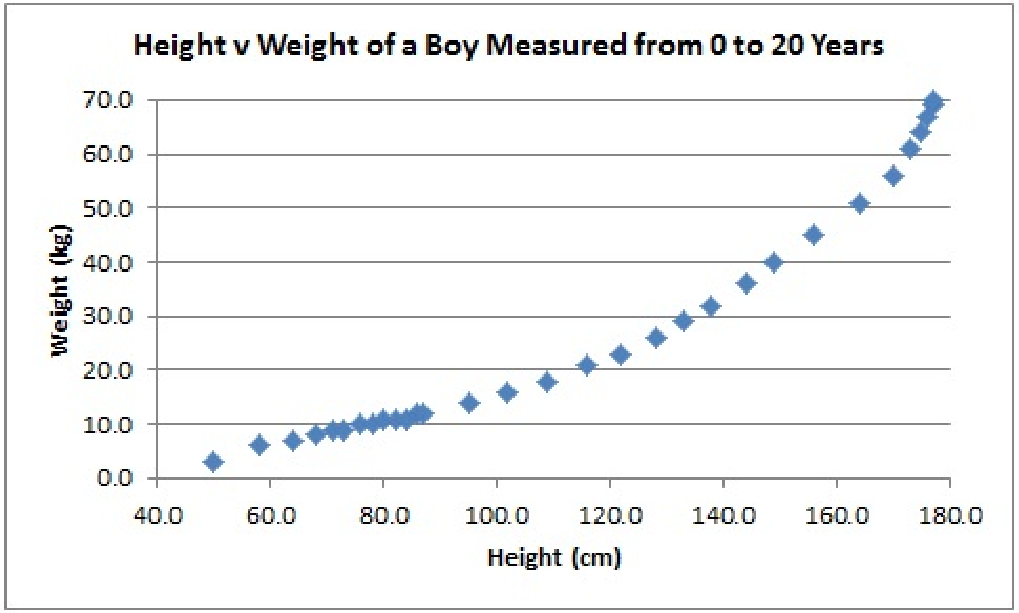
As points on the plot appear to be random there is no relationship between the time taken ot get to school and time spent on homework. One could speculate that taking a very long time to get to and from school means a student has less time to spend on homework and thus tends to spend less time on homework but this study does not support that conclusion.

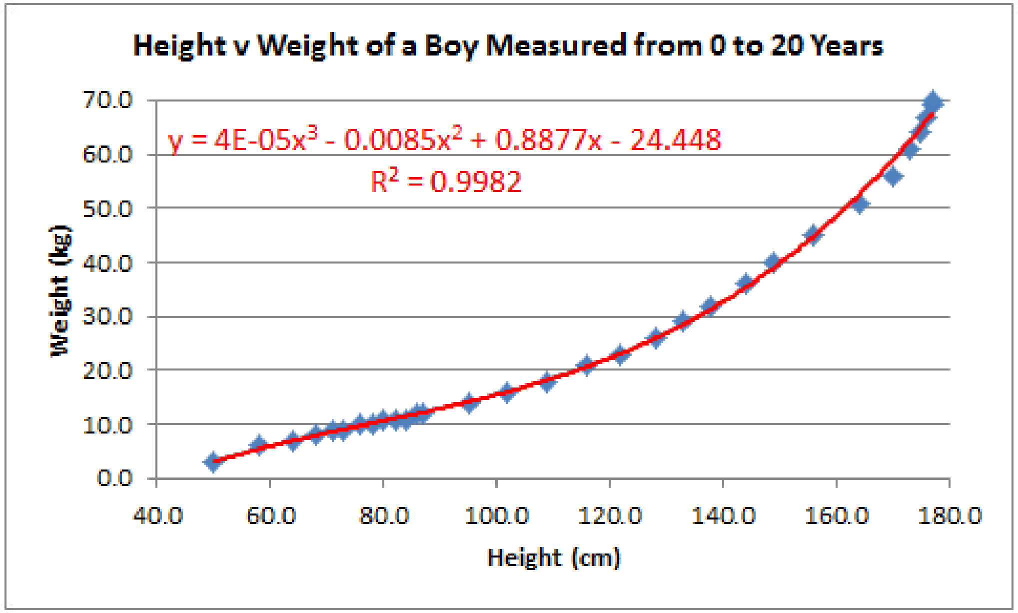
Prediction

The computer gives us a model $y = 0.153x + 39.008$. However there is clearly no relationship between these variables, (the dots are randomly scattered and r is only 0.0707) so it would not be appropriate to make predictions using this model.

6. What would you write in your "Conclusion" section? Answer your investigative question and make a prediction.

I wonder if there is a relationship between a boy's height and his weight measured at significant times from 0 to 20 years of age.





Conclusion

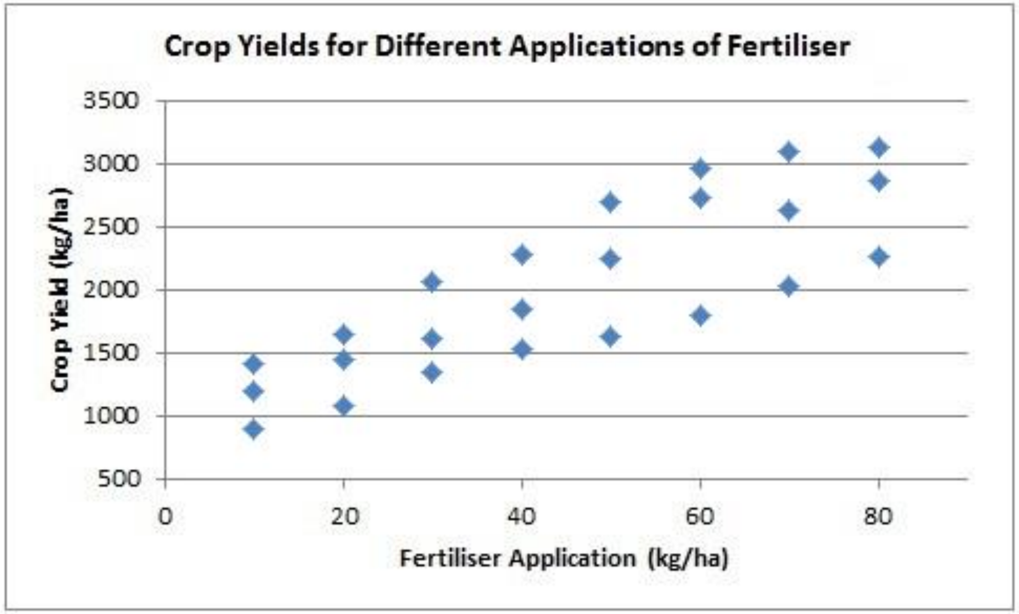
The taller this boy is the heavier he tends to be. The upward curve in the trend line shows this. You would expect this because as height increases, other body proportions (and hence weight) would also increase. The computer has fitted a cubic model to the data points in the sample which makes sense as all 3 dimensions of a body are likely to increase at the same rate so body volume and therefore weight would be roughly cubic.

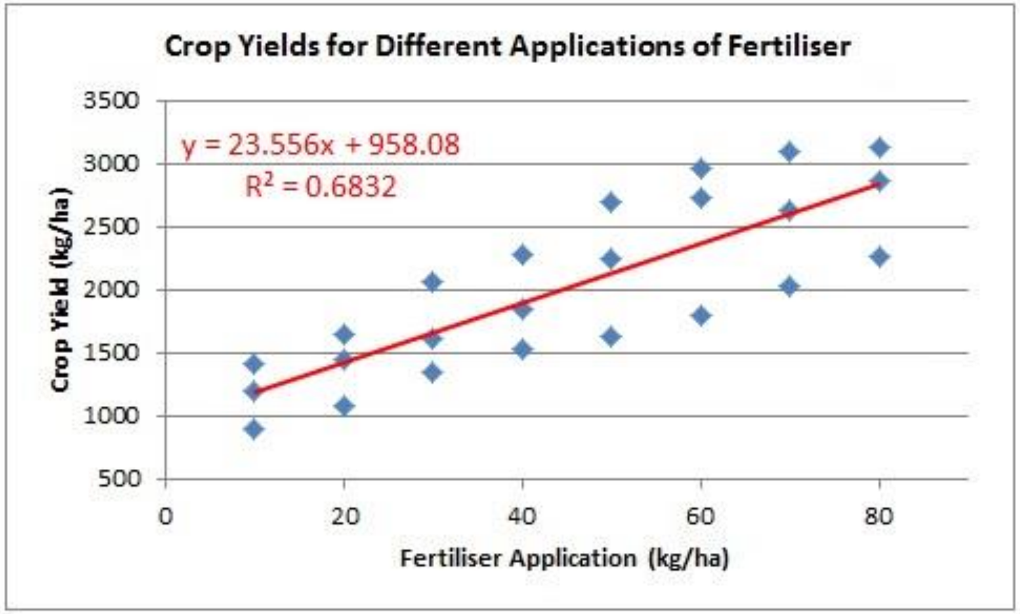
Prediction

The curve is $y = 0.00004x^3 - 0.0085x^2 + 0.8877x - 24.448$*.* $0.00004 \times 92^3 - 0.0085 \times 92^2 + 0.8877 \times 92 - 24.448 = 16.42392$*. Therefore I predict that when this boy's height is 92cm is weight would be about 16kg. The dots in this sample are all very close to this curve so I believe this prediction would be very accurate. I don't think this model would be accurate for other boys or for girls as growth varies a lot between people. Nor would it be accurate for this boy at older ages because there is no way of knowing if the curve will continue to be accurate.*

7. What would you write in your "Conclusion" section? Answer your investigative question and make a prediction.

I wonder if there is a relationship between the amount of fertiliser applied per hectare and the crop yield per hectare.





Conclusion

Higher yields tend to result with higher fertilizer applications. This can be seen by the positive gradient of the trend line. You would expect this but there would be a limit as too much fertilizer could damage the crop. This might be why the points appear to be levelling out.

Prediction

The trend line is y = 23.556x + 958.08. 23.556 x 55 + 958.08 = 2,253.66. Therefore I predict that the crop yield will be about 2,250 kg/ha if fertiliser is applied at the rate of 55 kg/ha. Because the relationship is moderate to strong (r = 0.8266) I think the estimate will be reasonably accurate for this fertiliser and this crop. However I notice that in the data points we have, there is variation of about ±500 kg/ha in yield in this part of the graph. Therefore I can expect that there will be an error of about ±500 kg/ha in my estimate of 2,250 kg/ha. Obviously this model will only be useful for predicting results with this particular fertiliser, this particular crop and close to similar soil conditions. Although the dots show a positive trend I expect that eventually they will start to go downwards as at some point more fertiliser will not help the soil but will in fact damage it.