Name: Benjamin Eyison
Date: 7th April 2008
Grade: IB2B
Subject: Mathematical studies

Comparing and analyzing the correlation of the number of novels read per week and the modal grade of T.I.S students

Statement of task

This project investigates the number of story books read per week correlated with the modal grades of 30 girls and 30 boys of T.I.S. As mentioned, 30 girls and 30 boys will be surveyed, all from grade 9. Data on the number of books and genres read per week, as well as their modal grades earned last semester. This data is then compared and analyzed to yield necessary results

Statistics

In this section, the mean, mode and median are found. These were done 4 times in

- Number of book genres girls often read
- Girl's modal grade
- Number of book genres boys often read
- Boy's modal grade

Person's Correlation coefficient (PCL)

The PCL was used to find out how and if the amount of books read on the modal grades of students. Both girls and boys data were done separately

Chi-squared test for independence (CST)

The CST was used to test if the modal grade and number of books read are dependent on one another, as well as modal grade and type of books. This shall be done by gender vs modal grade, gender vs. number of books read and type of books read vs. gender.

Table of contents

BOYS DATA

Boys	Number of books read	mark
1	2	81-100
2	1	40-60
3	1	61-80
4	2	61-80
5	2	61-80
6	5	40-60
7	1	40-60
8	2	40-60
9	1	81-100
10	6	61-80
11	6	61-80
12	6	61-80
13	4	61-80
141	2	61-80
15	1	40-60
16	2	40-60
17	4	61-80
18	1	40-60
19	1	61-80
20	1	below40
21	1	61-80
23	1	61-80
24	2	81-100
25	1	40-60
26	6	61-80
27	1	61-80
28	6	below 40
29	1	40-60
30	6	61-80

MARK AVERAGES (BOYS)

Modal mark (%)	Frequency	Fx
Below 40	2	70
40-60	9	450
61-80	16	1120
81-100	3	270
Total	30	1910

Mean= $\frac{\sum fx}{N} = \frac{1910}{30} = 63.66 = 63.7$

A mean of 63.7 shows that boys have a low mark average in English

Mode: 61-80

This shows that more boys exhibit an average performance in English

Median: $\frac{n+1}{2} = \frac{30+1}{2} = \frac{31}{2} = 15.5^{th}$ value= 61-80

This shows that there is a lesser spread of higher marks attained by the boys in English

NUMBER OF BOOKS READ AVERAGES (BOYS)

No of books	Frequency	Fx
1	14	14
2	7	14
3	0	0
4	2	8
5	1	5
6	6	36
Total	30	70

Mean= $\frac{\sum fx}{N} = \frac{77}{30} = 2.33$

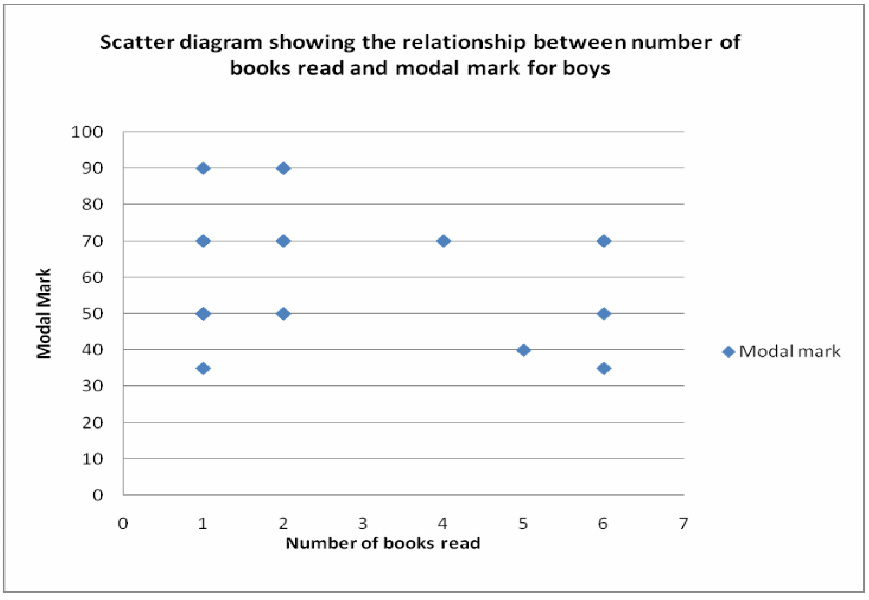
A low mean number of books indicate a very low amount of boys who read

Mode: 1

This means the most frequent number of books read is 1, indicating that boys read less number of books per week

Median= $\frac{n+1}{2} = \frac{30+1}{2} : \frac{31}{2} : 15.5^{th}$ value: 2 as 15-16 represent values present in that frequency

Since the median is 2, this shows that books read by boys were of low numbers.



As shown above, there is no correlation between number of books read and modal mark for boys

GIRLS' DATA

Girls	Number of books read	Mark
1	2	61-80
2	3	61-80
3	2	61-80
4	1	61-80
5	1	61-80
6	3	40-60
7	2	61-80
8	3	61-80
9	1	61-80
10	5	61-80
11	1	40-60
12	2	61-80
13	3	40-60
14	6	40-60
15	2	61-80
16	1	40-60
17	2	61-80
18	4	80-100
19	4	80-100
20	2	40-60
21	3	80-100
22	3	80-100
23	4	61-80
24	3	61-80
25	2	61-80
26	2	40-60
27	6	61-80
28	1	40-60
29	2	40-60
30	3	61-80

MARK AVERAGES (GIRLS)

Modal mark%	Frequency	Fx
Below 40	0	0
40-60	9	450
61-80	17	1190
81-100	4	360
Total	30	2000

Mean=$\frac{\sum fx}{N} = \frac{2000}{30} =$ 66.7, thus average achieved by the female body was from 61-80.

Mode= 61-80, meaning that the largest number of girls gained marks within that range

Median: = 15.5^{th} value: 40-60.

median value shows that midway modal mark % of the girls was lower, showing poorer academic status of the girls

NUMBER OF BOOKS READ (GIRLS)

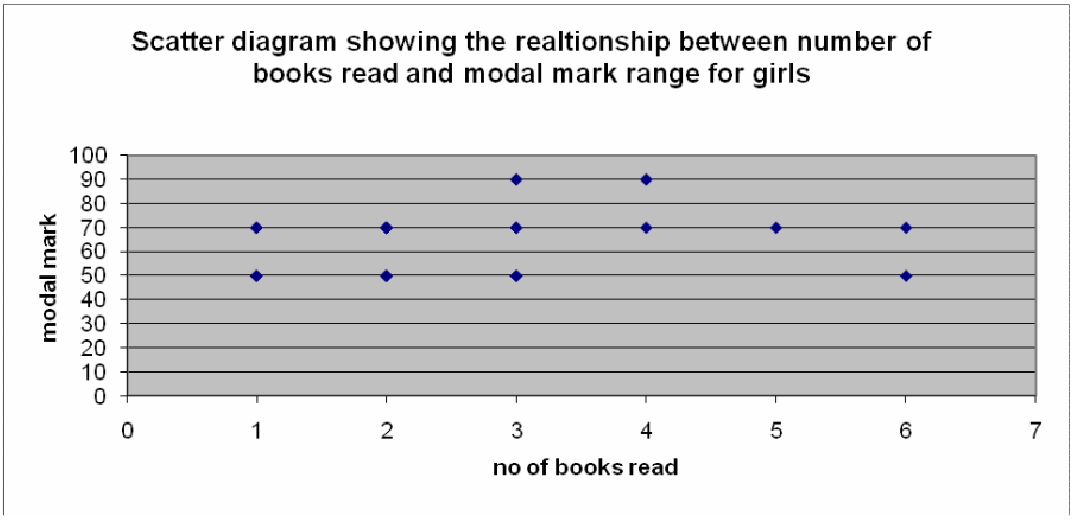
Number of books	Frequency	Fx
1	6	6
2	8	16
3	6	18
4	3	12
5	1	5
6	2	12
Total	26	69

\

Mean==$\frac{\sum fx}{N} = \frac{69}{26}$

Mode: 2, which shows that most girls preferred a lower amount of books to read

Median: 3.5th value: 1. This shows that the middle amount of books read was 1, which is even lesser than the median number of books by boys.



As shown, there is no correlation between number of books read and modal mark for girls

Step 2

The next step was to do the chi squared test. It was necessary to see which were dependent and which were independent. This will show if there is relationship between no of book genres read and modal grade marks. Modal mark groups were made into below and above 61-80. This is because calculations made previously for girls and boys generally lay round 61-80. Those with the exact average were added to the above group.

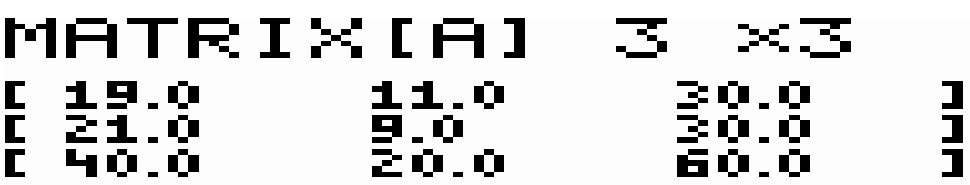
Phase 1: Mark & gender

*H*o = Modal mark and gender are independent

*H*1= Modal mark and gender are dependent

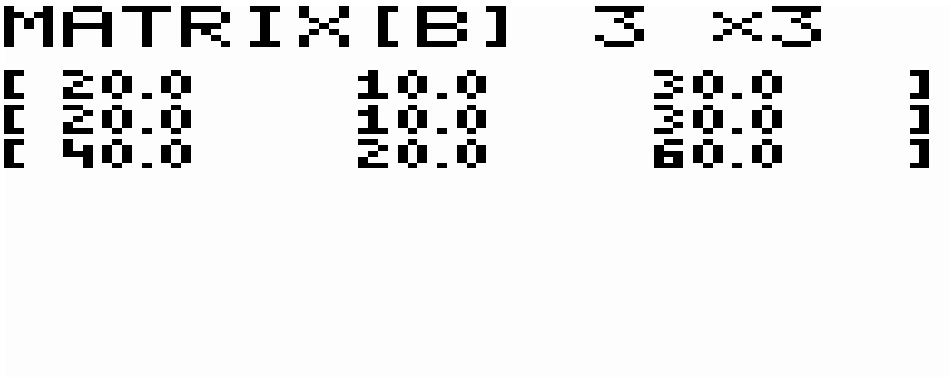
Observed Values

	Modal mark < 61-80	Modal mark > 61-80	Total
Boys	19	11	30
Girls	21	9	30
Total	40	20	60

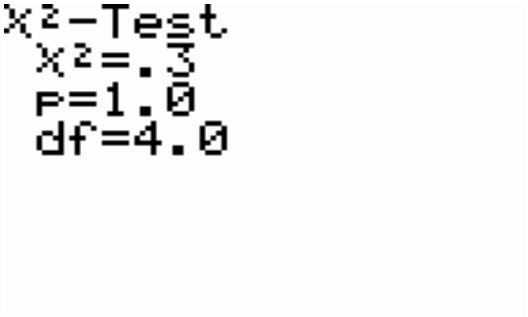


The expected values were calculated using the GDC:

Expected Values



Using the GDC we can find the value of X^2 calc using the GDC. The following results are shown



From above, the following was deduced
X^2 calc: 0.3

P: 1.0

Using the P value at 5%, decision taking was made

Since P >0.05,

Thus accept *H*o

Therefore modal mark and gender are independent

The same test was used for the number of books read and gender. The groups were the number of books<3 and >3

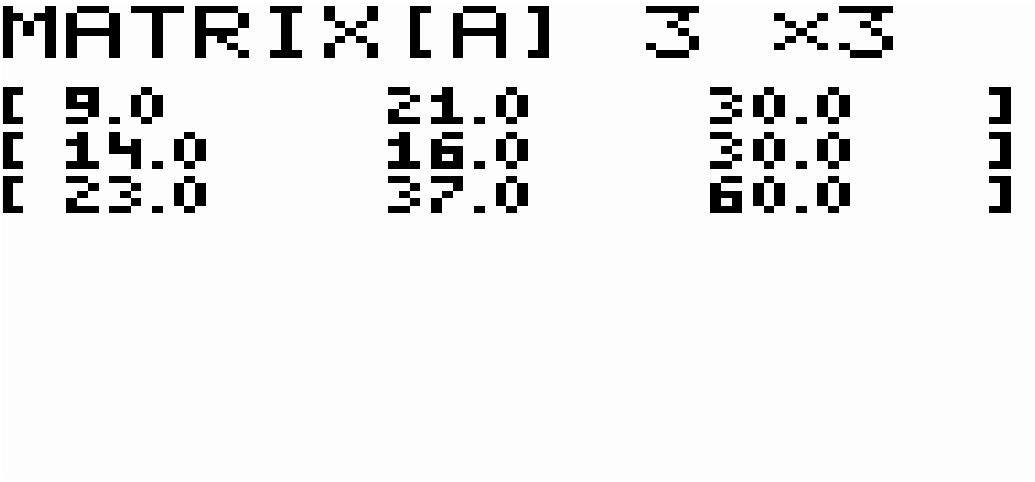
Phase 2: Number of books read & Gender

*H*o: Number of books read and gender are independent

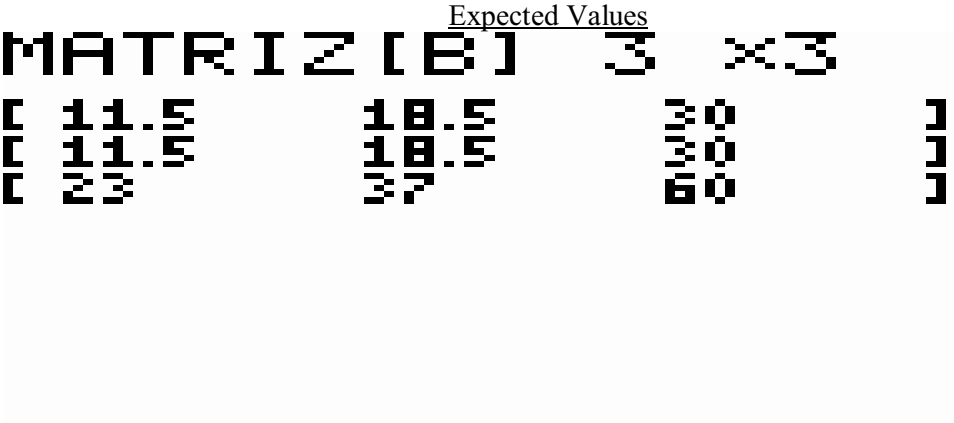
*H*1: Number of books read and gender are dependent

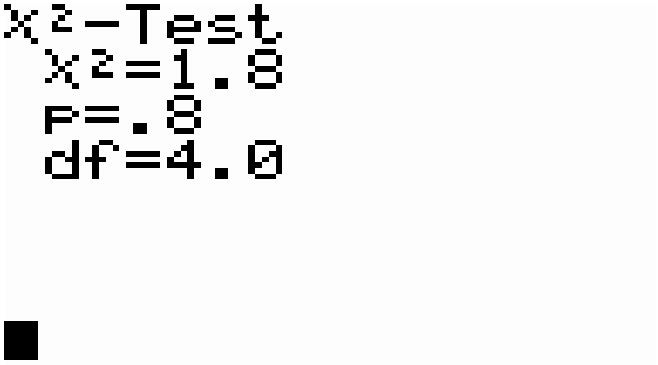
Observed values

	Number of books <3	Number of books >3	Total
Boys	9	21	30
Girls	14	16	30
Total	23	37	60



Expected Values





So $X^2 = 1.8$

P=0.8

Since P>0.05, at 5% expectancy level, we accept *H*o

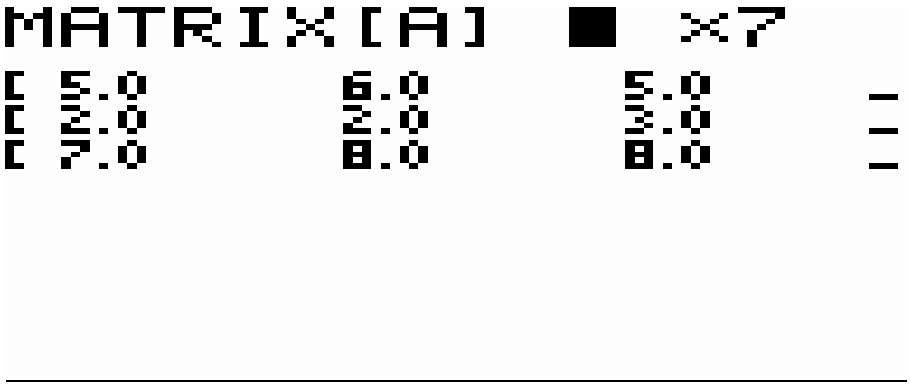
Therefore number of books read and gender are independent

Phase 3: Type of books read & Gender

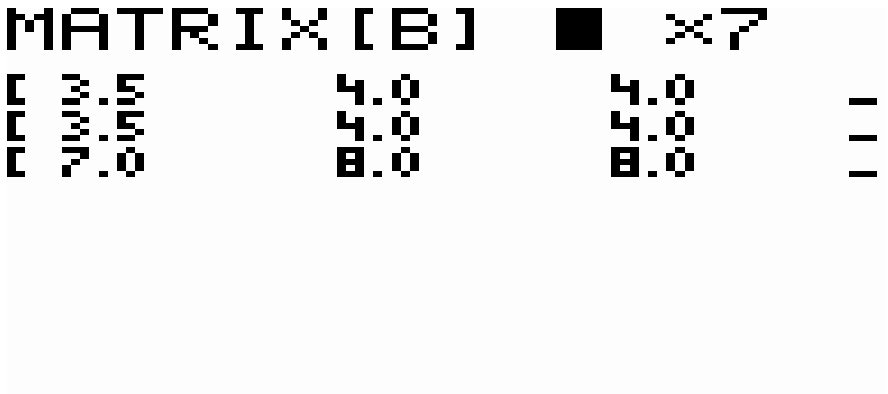
This phase focused on the different type of books read against the gender

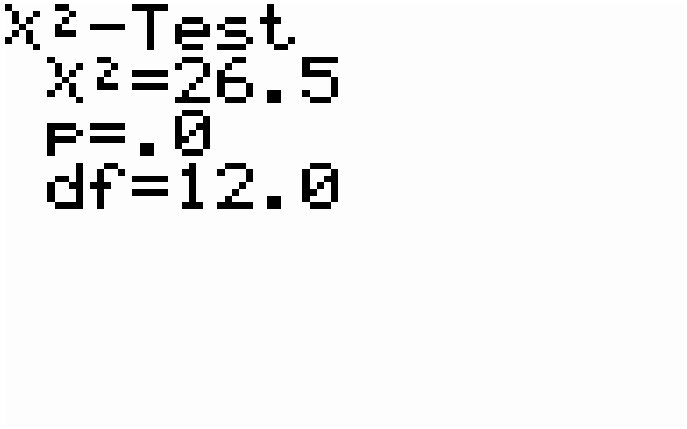
Observed values

	Horror	Science fiction	Adventure	Comedy	Romance	Action	Total
Boys	5	6	5	3	1	10	30
Girls	2	2	3	4	18	1	30
Total	7	8	8	7	19	11	60



Expected values





So $X^2 = 26.5$

P=0.0

Since P<0.05, at 5 % level, we reject H_0

There for type of books read and gender are dependent

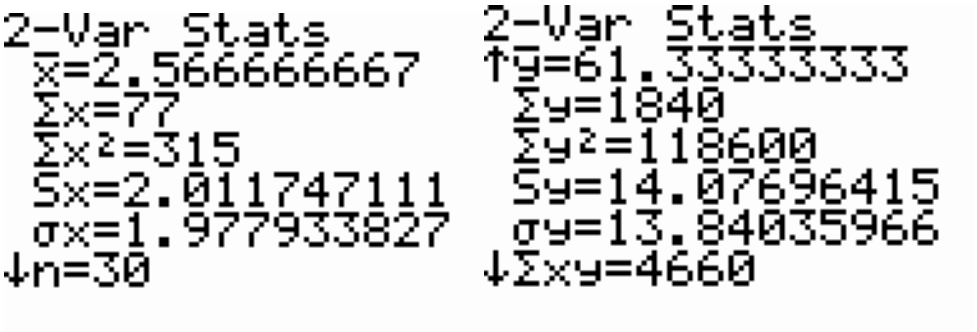
Pearson's correlation coefficient

Next was the Pearson's correlation co-efficient. This was also used to find out correlation between the modal mark and no of books read, but using a different method.
NOTE: for (y), the median of the selected range of modal marks was used to represent (y)

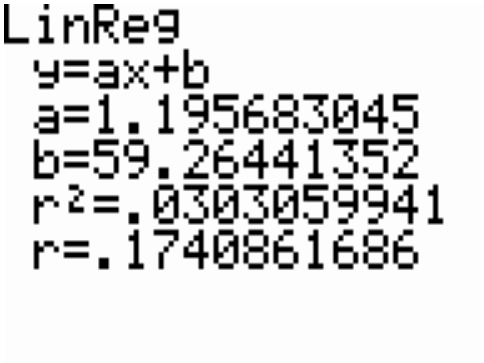
BOYS' DATA

No of books(x)	mark (y)	(xy)	x^2	y^2
2	80	180	4	8100
1	50	50	1	2500
1	70	70	1	4900
2	70	140	4	4900
2	70	140	4	4900
5	50	250	25	2500
1	50	50	1	2500
2	60	100	4	2500
1	90	90	1	1800
6	70	420	36	4900
6	70	420	36	4900
6	70	420	36	4900
4	70	280	16	4900
2	70	140	4	4900
1	70	50	1	2500
2	70	100	4	2500
4	70	280	16	4900
1	50	50	1	2500
1	50	70	1	4900
1	30	35	1	1225
1	70	70	1	4900
1	70	70	1	4900
1	60	70	1	4900
2	80	180	4	8100
1	50	50	1	2500
6	50	420	36	4900
1	50	70	1	4900
6	30	35	36	1225
1	50	50	1	2500
6	50	420	36	4900

The statistic for the table for boys was shown using the 2-Var Stat function on the GDC:



Now the values obtained were used to calculate the correlation coefficient of r using the [LinReg] function of the GDC

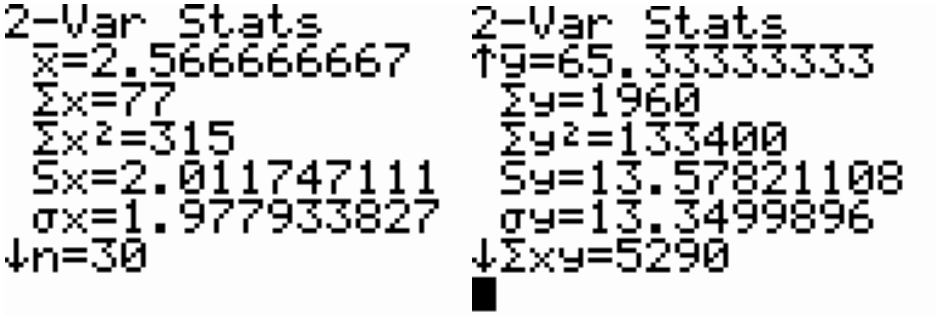


Therefore r= 0.174 and r^2= 0.303

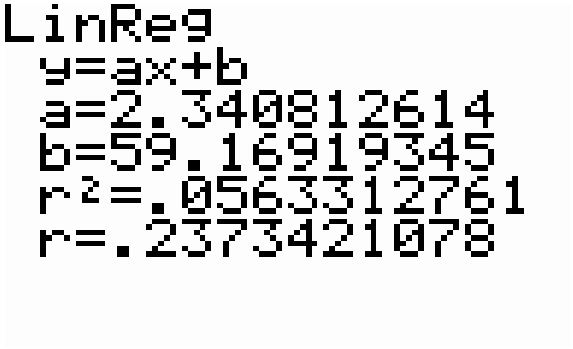
Value of r is a very low positive value. It shows a very weak positive correlation between modal marks and number of books read for the male gender. In other words this means that the amount of books read are independent on the gender for males, implying that, instead of the time spent reading, several factors could have affected the performance of the boys in English Language.
The value of r^2 shows that 30% if the variation in the number of books read can be accounted for by the variation in the modal mark. In other words, approximately 70% of the variation is attributed to other factors.

GIRLS

No of books(x)	Modal mark(y)	(xy)	(x^2)	(y^2)
2	70	140	4	4900
3	70	210	9	4900
2	70	140	4	4900
1	70	70	1	4900
1	70	70	1	4900
3	50	150	9	2500
2	50	100	4	2500
3	50	150	9	2500
1	70	70	1	4900
5	70	350	25	4900
1	50	50	1	2500
2	70	140	4	4900
3	50	150	9	2500
6	50	300	36	2500
2	70	140	4	4900
1	50	50	1	2500
2	70	140	4	4900
4	90	360	16	8100
4	90	360	16	8100
2	50	100	4	2500
3	90	270	9	8100
3	90	270	9	8100
4	70	280	16	4900
3	70	210	9	4900
2	70	140	4	4900
B	50	100	4	2500
6	70	420	36	4900
1	50	50	1	2500
2	50	100	4	2500
3	70	210	9	4900



The values were tabulated on the T.I calculator, then using the [LinReg] function, found the following:



r= 0.24

r^2= 0.56

The value of r is a very low positive value. It is shows a weak positive correlation between modal mark and number of books read for girls. As such, it shows a hint of books affecting the rate at which girls do well in English, but due to the weakness of the correlation, this assumption could be otherwise.
The r^2 value shows that 56% of the variation in the modal mark earned by girls can be accounted by the number of books read. Unlike the boys, this figure shows that approximately 50% of the variation is related to other factors. With this, it can be said that there is half a possibility that girls get better marks by reading more books.

Analysis and conclusion

Step 1

The 30 boys surveyed: Most gain generally from 61-80 mark and most read a little of amount of 1 book per week. Their average mark is 63.7 and their median is 61-80. Their modal value is also 61-80. This mark with correlation to number of books read shows that number of books read has very little to do with modal mark earned.
The 30 girls surveyed: Modal mark is 61-80. Most read a higher amount of books of 2 per week than boys, but is still quite low. Their median value for the marks is 40-60 and their average mark is 66.7. With this mark, in relation to the number of books read, there is little correlation between the two. Moreover

Step 2

Fist chi-squared test done was for gender and modal. There is no connection as the value of $x^2 = 1.76$ and it was less than the critical value. Because of this, both gender and modal grade are independent. The second was gender and no of books read. This also showed no connection as $x^2 = 0.3$, which is less than the critical value. Thus there is little correlation between modal mark and gender, as both are independent of each other. There were also various results for the type of books each gender preferred to read. For instance, romantic books were the most books read by the girls, while horror was preferred most by the boys. In short, the girls seem adapt a certain attitude reading as the books they read are of higher literary value, with complex vocabulary and writing style. However the boys mostly read horror books that are of little vocabulary, and of little distinct literary value

Step 3

The coefficient value for girls was very low, with but that for boys was high. For the boys, $r = -0.96$, showing a strong negative and $r^2 = 0.84$, also showing strong negative correlation between male gender and the number of books read/modal marks. For the girls, although positive, the r value=0.247, showing a very weak positive correlation between the female gender and no of books read/modal mark Also for the boys, the r^2 value means that' there is a 84% chance that the value could be correlated. In other words there is high possibility that the number of books read by the boys did not affect their academic status in English Language. Furthermore, the r^2 value for girls, it shows that ether is a lesser 56% chance that the values cannot be correlated. This means that there is about an even probability that the number of books read by girls affects their earned modal mark in English language.
Thus of all the steps to find the relationship between number of books read and modal mark in English language for both gender has proven that they are independent of one another and there is little or no correlation. The chi squared test showed that all the variables are independent of one another, thus no correlation between them. The type of books read notable showed on genre read often by a gender. From this, one can assume that the type of books read could somehow affect the mount of marks earned in the subject. The mode median and mean showed that even those who read less number of books often got good marks, contrasting the earlier hypothesis. Therefore, these facts show that there is no relationship between n of books read and modal grade earned by

both genders in English Language. Instead, it is assumed that the academic performance of both genders could be affected by several other factors, such as the number of hours slept or the subject preference.

Reliability and Validity

There are a few factors which influence the reliability of this experiment.

Only 30 girls and 30 boys were chosen, as there was the need to use a an equal amount of people for each gender in order to gain fairer results

A specific grade was chosen: grade 9 as it has the largest class population, thus more data gained

There is a chance that those who filled in the questionnaire may have not been honest and wrote false marks or number of books they read per week

Students seemed to have more of higher marks than lower, showing the unreliability of their answers in the questionnaire, as well as lower amount of books read per week. This may be the reason why there is little or no correlation at all for the girls, and a string negative correlation for the boys. In fact, it is plausible to say that from the results gained, the girls who have positive, but low correlation seem more trustworthy than that of the boys'

In the modal grade, the only categories given were that of below 40 to 100. This is because as confirmed with the English Language teacher, no student gained below 40 very often. Thus values of 10-30 were summarized into that of "below 40"