# TABLE OF CONTENTS


➔ Introduction p.3

➔ Information/Measurement p.3

➔ Mathematical Processes and Interpretation of Results p. 4 to p.10
- Histograms p. 4 to p. 6
- First Chi-Square test p. 7 to p. 8
- Second Chi-Square test p. 9 to p. 10

➔ Validity p. 11

➔ Appendix p. 12 to p. 15

# INTRODUCTION

**Task:** Based on the 130 sets of data that I have collected from IB students around the world, I am going to determine whether there is a relationship between an IB student's overall grade average and his/her rank in the family as well as the number of siblings that he/she has. I find this project to be of particular interest to me as I am an IB student and because the IB is an academic program which takes place during the two last years of high school, which means that grades are at their utmost importance as they directly affect our future. Therefore, it would be very interesting to know if we are automatically at an academic advantage because of our rank in the family or because of the number of siblings we have.

**Plan:** The data was collected during a summer program at a university in the United States and in math class at my school. All of the students who participated were IB students, and in order to provide me with their overall average over 7, they calculated the mean based on the average grade of each of the subjects they are taking. To be able to find answers to my investigation, I am going to use histograms as they give a clear sense of the distribution of the data and enable me to look at the two factors that I am focusing on (the rank in the family and the number of siblings) at the same time. After I analyze the histograms, I will decide whether to perform a Chi-Square test based on the data that have shown to be the most interesting on the histograms, meaning the data that seem to be dependent and therefore show a relationship with a student's IB grade average.

# INFORMATION/MEASUREMENT

My investigation is based on IB grade averages over a total of 7 maximum marks. Since most of my numerical results represent people, no mathematical unit must be used. However, for the Chi-Square tests' expected values, all of the numbers are to three significant figures, and for the Chi-Square values and P-values, they are given as exact numbers.

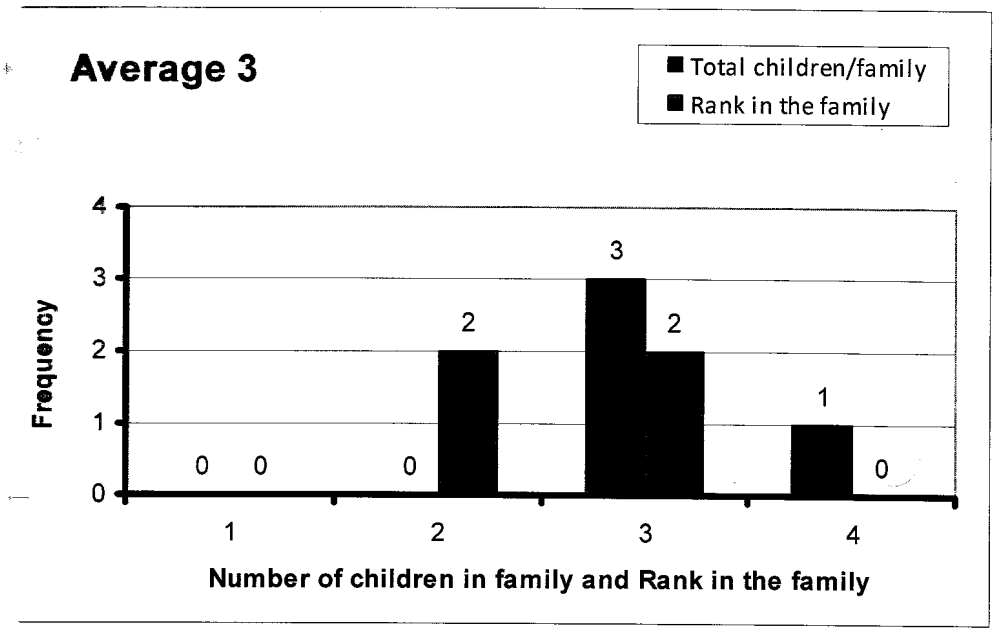
A table containing all of my raw data is given in the appendix.

# MATHEMATICAL PROCESSES AND INTERPRETATION OF RESULTS

Not enough people out of the 130 that had participated in giving me data had an overall IB average grade of 2 or 7 (no one had an average of 1) for a histogram to be worthwhile doing for those grades. Consequently, four histograms have been done: averages 3, 4, 5, and 6.

As I had mentioned in my introduction, I have decided to use histograms because they are a clear and organized way of representing data and because they make any trends apparent.

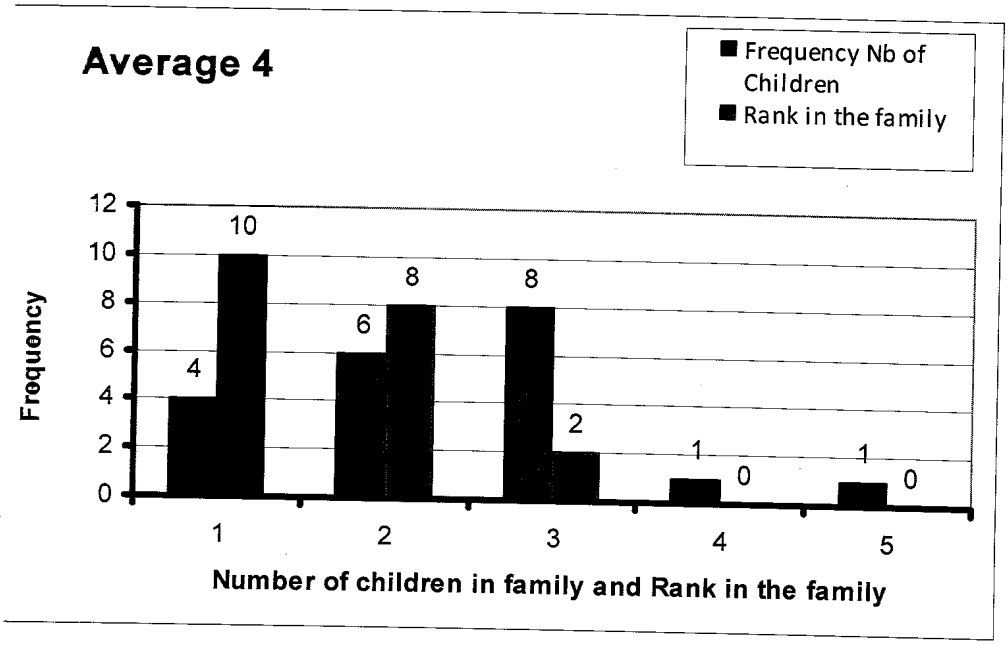
**Histogram representing students who have an overall IB average of 3 over 7:**




This histogram indicates that there are two students who scored 3 out of 7 who are second in the family rank, two students who are third in the family rank, three students who have two siblings, and one student who has three siblings.

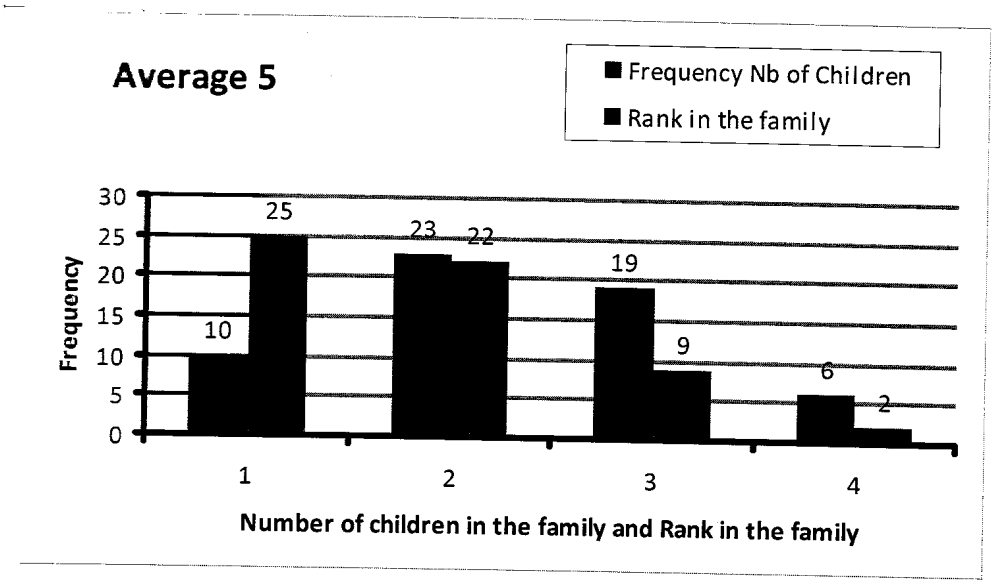
We can infer from the first histogram "Average 3" that a student is more likely to have an overall IB average of 3 over 7 if he/she is either second or third in the family rank and has two siblings. However, this cannot be considered as a real trend as only 4 students out of the 130 that I asked obtained an average grade of 3, which is evidently not enough to show any valid pattern.

**Histogram representing students who have an overall IB average of 4 over 7:**




When looking at the second histogram "Average 4", we can see that one is most likely to obtain such an overall average if one is first in the family rank and has two siblings. Indeed, ten students who are the eldest in their family and eight students who have two siblings scored an average IB grade of 4.

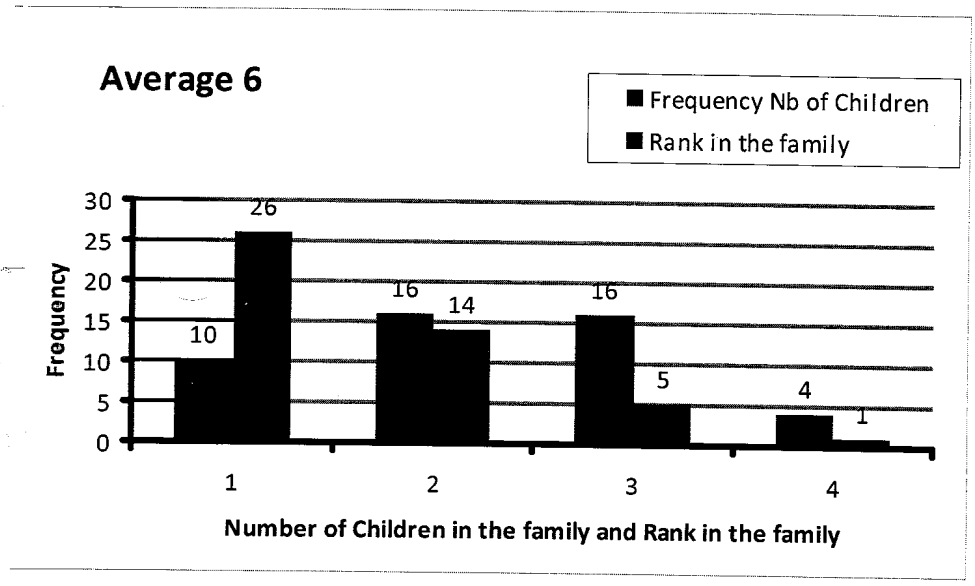
**Histogram representing students who have an overall IB average of 5 over 7:**




The third histogram "Average 5" shows us that twenty-five students who have an average grade of 5 are first in the family rank and only ten come from a family with one child.

We are also shown that twenty-three students who have an average of 5 out of 7 come from a family of two children and nineteen come from a family with three children. Therefore, an IB student is most likely to get an average of 5 over 7 if he/she is first in the family rank and has one other sibling.

**Histogram representing students who have an overall IB average of 6 over 7:**




The fourth and last histogram "Average 6" demonstrates a rather strong correlation between people who are first in their family rank and who obtain an overall average of 6 over 7. Being second in the family and having one other sibling also are quite high in this histogram.

Looking at the four histograms as a whole, we can see that there is a trend in the fact that students who are the first in the family ranking get higher averages. Indeed, no student who is first in the family rank has obtained an average of 3 over 7, and when checking in my raw data to see the ranking of the person who had an average of 2, it has confirmed the trend as that person is second in the family rank. Ten IB students who are the first children born in the family have obtained an average of 4 over 7, twenty-five students who are first in the family ranking obtained an overall average of 5, and twenty-six students who are also first have obtained an average of 6 over 7. Again, I checked with my raw data to see the family ranking of the student who had an overall IB average of 7 over 7, and he is indeed first in his family's ranking.

In order to investigate more deeply on the relationship of a student's rank in the family and his her overall IB grade average, I will perform a Chi-Square test on the averages 5 and 6 with the rankings from first to third, as they were the ones with the most data. Indeed, I have chosen not to incorporate the average of 4 out of 7 as the data is too poor compared to the averages of 5 and 6.

## Chi-Square test

**Null Hypothesis:** A student's average IB grade is independent to his/her rank in the family.

**Alternative Hypothesis:** A student's average IB grade is dependent to his/her rank in the family.

**Degree of significance chosen:** 5%

## TABLE OF OBSERVED VALUES

| | Average of 5 | Average of 6 | TOTAL |
|---|---|---|---|
| 1st child in the family | 25 | 26 | 51 |
| 2nd child in the family | 22 | 14 | 36 |
| 3rd child in the family | 9 | 5 | 14 |
| TOTAL | 56 | 45 | 101 |

**TABLE OF EXPECTED VALUES (to 3 significant figures)**

| | Average of 5 | Average of 6 |
|---|---|---|
| 1st child in the family | 28.3 | 22.8 |
| 2nd child in the family | 20.0 | 16.0 |
| 3rd child in the family | 7.76 | 6.24 |

**Degree of freedom:** (2-1) x (3-1) = 2

**Chi-Square value:** 1.763136551

**Chi-Square critical:** 5.991

**P-value:** 0.4141329278

Since the Chi-Square value is smaller than Chi-Square critical, we accept the null hypothesis. The fact that the p-value is bigger than the degree of significance (0.05) confirms this. Therefore, a student's average IB grade is independent to his/her rank in the family according to the Chi-Square test. Consequently, the trend observed in the histograms may be a coincidence as 130 IB students is not quite representative of the actual IB student population.

I will now perform another Chi-Square test to see whether the number of children in a student's family and his/her IB grade average are dependent or independent, as this first Chi-Square test has proven to be very useful as it contradicted the trend from the histograms. Again, I will only look at the averages 5 and 6 out of 7 as they are the most interesting due to the amount of data. Also, I will only test families who have one, two, or three children as I do not have enough data for students in families with four children. The pattern shown by the histograms concerning this part of the investigation for grades 5 and 6 are that a student is most likely to obtain those averages if they have one or two siblings.

## Second Chi-Square test

**Null Hypothesis:** A student's average IB grade is independent to the number of children in his/her family.

**Alternative Hypothesis:** A student's average IB grade is dependent to the number of children in his/her family.

**Degree of significance chosen:** 5%

## TABLE OF OBSERVED VALUES

| | Average of 5 | Average of 6 | TOTAL |
|---|---|---|---|
| One child in the family | 10 | 10 | 20 |
| Two children in the family | 23 | 16 | 39 |
| Three children in the family | 19 | 16 | 35 |
| TOTAL | 52 | 42 | 94 |

## TABLE OF EXPECTED VALUES (to 3 significant figures)

| | Average of 5 | Average of 6 |
|---|---|---|
| One child in the family | 11.1 | 8.94 |
| Two children in the family | 21.6 | 17.4 |
| Three children in the family | 19.4 | 15.6 |

**Degree of freedom:** (2-1) x (3-1) = 2

**Chi-Square value:** 0.4548712582

**Chi-Square critical:** 5.991

**P-value:** 0.796573696

Since the Chi-Square value is smaller than Chi-Square critical, we accept the null hypothesis. The fact that the p-value is bigger than the degree of significance (0.05) confirms this. Therefore, a student's average IB grade is independent to the number of children in his/her family according to the Chi-Square test.

At the end of this investigation, we can infer that a student's rank in the family and the number of siblings he/she has do not affect overall IB averages, even though some patterns had seemed to appear on the histograms.

# VALIDITY

I think the histograms are a clear and visually pleasing way to efficiently note trends among the data. They are easy to understand and give a straightforward sense of what to expect. However, the results obtained by those histograms cannot be considered to be representative of the whole IB population, as I had mentioned before. This limits the validity of the patterns that emerged on those histograms as the Chi-Square test refuted any relationship between rank in the family and IB average, and number of children in the family and IB average. Nonetheless, having done the two Chi-Square tests to check those patterns has given more accuracy to the conclusions of my investigation.

An important limitation of my investigation is the fact that the grade averages 2 and 7 could not be completely included because of the lack of data associated with those averages. This is why I think that taking the same number of people per average would have helped improve my investigation. Perhaps if I could have obtained 50 IB students per average grade, my results could have been more reliable.

# APPENDIX

## RAW DATA

| | Gender | Rank | No. Of Siblings | Average |
|---|---|---|---|---|
| 1 | M | Oldest | 2 younger sisters | 5 |
| 2 | M | Middle | 1 older sister 1 younger brother | 5 |
| 3 | F | Middle | 1 younger brother 2 older sisters | 4 |
| 4 | M | Youngest | 1 older sister 1 older brother | 6 |
| 5 | F | Younger | 1 older brother | 6 |
| 6 | F | Oldest | 2 younger sisters | 6 |
| 7 | F | Older | 1 younger brother | 5 |
| 8 | M | Middle | 1 older sister 1 younger brother | 6 |
| 9 | M | Only Child | - | 6 |
| 10 | M | Middle | 1 younger brother 1 older brother | 5 |
| 11 | F | Only child | - | 5 |
| 12 | F | Younger | 1 older brother | 5 |
| 13 | M | Older | 1 younger sister | 7 |
| 14 | F | Older | 1 younger sister | 5 |
| 15 | M | Only Child | - | 6 |
| 16 | M | Middle | 1 older brother 2 younger sisters | 2 |
| 17 | F | Youngest | 2 older sisters | 3 |
| 18 | M | Younger | 1 older sister | 6 |
| 19 | F | Younger | 1 older brother | 5 |
| 20 | F | Oldest | 2 younger brothers | 6 |
| 21 | M | Middle | 1 older sister 1 younger sister | 3 |
| 22 | M | Only Child | - | 4 |
| 23 | F | Middle | 1 older sister 1 younger brother | 5 |
| 24 | F | Younger | 1 older sister | 4 |
| 25 | M | Only Child | - | 4 |
| 26 | F | Older | 1 younger brother | 5 |
| 27 | F | Only Child | - | 6 |
| 28 | F | Youngest | 1 older brother 2 older sisters | 6 |
| 29 | F | Oldest | 2 younger sisters 1 younger brother | 6 |
| 30 | M | Middle | 2 older sisters 1 younger brother | 5 |
| 31 | M | Only Child | - | 4 |
| 32 | F | Older | 1 younger brother | 6 |
| 33 | M | Oldest | 2 younger sisters | 6 |
| 34 | F | Youngest | 1 older brother 2 older sisters | 5 |

| | Gender | Rank | No. Of Siblings | Average |
|---|---|---|---|---|
| 35 | M | Middle | 1 older sister 1 younger sister | 5 |
| 36 | M | Younger | 1 older brother | 5 |
| 37 | F | Only Child | - | 5 |
| 38 | M | Only Child | - | 6 |
| 39 | F | Younger | 1 older brother | 5 |
| 40 | M | Oldest | 1 younger sister 2 younger brothers | 6 |
| 41 | F | Middle | 1 older sister 1 younger brother | 6 |
| 42 | F | Middle | 1 older brother 1 younger brother | 5 |
| 43 | F | Younger | 1 older brother | 5 |
| 44 | F | Older | 1 younger brother | 6 |
| 45 | M | Oldest | 2 younger sisters | 4 |
| 46 | M | Only Child | - | 5 |
| 47 | M | Oldest | 2 younger brothers | 4 |
| 48 | F | Middle | 1 older sister 1 younger sister | 4 |
| 49 | F | Younger | 1 older brother | 5 |
| 50 | M | Middle | 1 older brother 1 younger brother | 6 |
| 51 | M | Only Child | - | 6 |
| 52 | F | Youngest | 2 older sisters 1 older brother | 5 |
| 53 | M | Older | 1 younger sister | 4 |
| 54 | F | Oldest | 1 younger sister 1 younger brother | 5 |
| 55 | M | Middle | 1 older brother 1 younger sister | 5 |
| 56 | F | Youngest | 2 older brothers | 5 |
| 57 | F | Only Child | - | 6 |
| 58 | F | Middle | 1 older brother 1 younger brother | 5 |
| 59 | M | Older | 1 younger sister | 4 |
| 60 | F | Youngest | 2 older sisters | 6 |
| 61 | M | Middle | 1 older brother 1 younger sister | 4 |
| 62 | M | Older | 1 younger brother | 5 |
| 63 | F | Only Child | - | 6 |
| 64 | F | Only Child | - | 5 |
| 65 | M | Younger | 1 older sister | 5 |
| 66 | F | Younger | 1 older brother | 5 |
| 67 | M | Middle | 1 older brother 3 younger sisters | 4 |
| 68 | M | Oldest | 2 younger sisters | 6 |
| 69 | F | Oldest | 1 younger brother 2 younger sisters | 5 |
| 70 | M | Older | 1 younger brother | 6 |
| 71 | F | Younger | 1 older brother | 5 |

| | Gender | Rank | No. Of Siblings | Average |
|---|---|---|---|---|
| 72 | F | Middle | 1 older sister 1 younger brother | 6 |
| 73 | F | Younger | 1 older sister | 4 |
| 74 | M | Older | 1 younger brother | 4 |
| 75 | M | Only Child | - | 5 |
| 76 | F | Middle | 1 older sister 1 younger brother | 5 |
| 77 | M | Only Child | - | 5 |
| 78 | F | Youngest | 2 older brothers | 5 |
| 79 | M | Older | 1 younger brother | 6 |
| 80 | M | Oldest | 2 younger sisters 1 younger brother | 5 |
| 81 | M | Youngest | 2 older brothers | 5 |
| 82 | F | Older | 1 younger sister | 5 |
| 83 | F | Younger | 1 older sister | 6 |
| 84 | F | Only Child | - | 5 |
| 85 | F | Middle | 1 older sister 1 younger sister | 4 |
| 86 | F | Only Child | - | 5 |
| 87 | F | Middle | 1 older brother 1 younger brother | 5 |
| 88 | M | Younger | 1 older brother | 6 |
| 89 | M | Older | 1 younger sister | 6 |
| 90 | M | Older | 1 younger sister | 5 |
| 91 | M | Youngest | 2 older sisters | 6 |
| 92 | F | Only Child | - | 6 |
| 93 | M | Younger | 1 older brother | 6 |
| 94 | F | Younger | 1 older brother | 5 |
| 95 | M | Older | 1 younger sister | 6 |
| 96 | F | Middle | 1 older brother 1 younger sister | 4 |
| 97 | M | Oldest | 1 younger brother 2 younger sisters | 5 |
| 98 | F | Youngest | 1 older brother 1 older sister | 6 |
| 99 | M | Middle | 1 older sister 1 younger sister | 3 |
| 100 | F | Only Child | - | 6 |
| 101 | M | Younger | 1 older brother | 5 |
| 102 | F | Only Child | - | 4 |
| 103 | M | Middle | 1 older brother 1 younger sister | 5 |
| 104 | M | Older | 1 younger sister | 6 |
| 105 | M | Youngest | 2 older brothers | 5 |
| 106 | F | Middle | 1 older sister 1 younger brother 1 younger sister | 6 |
| 107 | M | Only Child | - | 6 |
| 108 | F | Middle | 1 older sister 1 younger brother | 6 |