Does the infant mortality rate of a country depend upon its GDP per capita?

Plan of Investigation:

I have chosen this topic, as I am interested in global development. I firstly began researching about health inequalities throughout the world and realised that there were significant differences between low-income developing countries and moderate to high-income developed countries. Upon researching I became deeply absorbed and surprised by the huge variations in infant mortality (IMR per 1,000 births) across the world, leading me to consider whether the IMR may depend upon the Gross Domestic Produce (GDP US dollars) per capita of a country. Therefore the type of exploration I will be conducting is an application investigation using statistics to uncover whether or not IMR (per 1,000 births) depends upon GDP per capita (US dollars). The first key mathematical techniques I plan to use in my exploration will be univariate and bivariate mathematics to assist me in generating observable relationships.

C: Personal engagement.

Firstly I will explore the univariate data of my exploration; the IMR (per 1,000 births) and the GDP per capita (US dollars) separately for developing and developed countries using box plots to examine the range and spread of the data. This will also enable me to compare and contrast the IMR (per 1,000 births) for developed and developing countries and the GDP per capita (US dollars) for developed and developing countries using parallel box plots. In addition using the five-figure summaries obtained from univariate mathematics I will calculate the range, interquartile range and test for outliers. Following this, I will use bivariate mathematics to investigate the relationship between the two variables: IMR (per 1,000 births) and GDP per capita (US dollars) within both developed and developing countries. I plan to do this through using scatter diagrams, which will allow me to use the Pearson's Product Correlation to determine if there is a correlation between the data and if an equation can be generated to represent the relationship. If an equation can be formulated I plan to model the expected IMR (per 1,000 births) of a country. Finally to further explore the relationship found between the two variables I will combine all data of developed and developing countries and generate a scatter diagram to see if there is an overall correlation between IMR (per 1,000 births) and GDP per capita (US dollars) regardless of whether the country is developed or developing. If a linear correlation cannot be obtained I plan to use alternative mathematics such as a means to obtain a linear relationship that can be explored using statistics. This will ultimately allow me to answer my research question 'Does the infant mortality rate of a country depend upon its GDP per capita?

A: The student describes the mathematics to be used and the reasons.

B: Key terms defined.

Definitions:

These definitions were obtained from the World Data Bank Organisation (WDBO), County and Lending groups page. However all data from WDBO is obtained and classified ultimately by the World Health Organisation – WHO[1]

Developing Country:

The term *developing* is used to denote low and middle-income countries with a GDP per capita commonly less than $12,000 US dollars

Developed Country:

The term *developed* is used to denote high-income countries with a GDP per capita commonly more than $12,000 US dollars.

Infant Mortality Rate per 1,000 live births (IMR):

The infant mortality rate of a country is the number of deaths of infants that are less than one year of age per 1000 live births. The rate for a given region or country is given by the number of infants dying under one year of age, divided by the number of live births during the year, multiplied by 1,000.[2]

---

[1] The World Data Bank. *County and Lending Groups.* 2015. http://data.worldbank.org/about/country-and-lending-groups (accessed January 15, 2015).

Gross Domestic Produce per capita (GDP per capita):

GDP per capita is the gross domestic product of a country divided by its midyear population. GDP is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources.[3]

Obtaining data:

In order to obtain a random sample of countries, which would be representative, an online random sample generator was used to randomly select 30 developed and developing countries. I did this to enable me to examine the data values of 30 developed and developing countries, and to conclude whether there is indeed a relationship between the IMR (per 1,000 births) and GDP per capita (US dollars) of developed or developing countries. I did this because the data provided by the World Health Organisation (WHO) and the World Data Bank (WDB) had the data of 214 countries worldwide. This sample size would be too large for my exploration therefore I deduced that I would obtain a sample of overall 60 countries, so my exploration was representative of a third of the population countries provided. So, firstly using the 214 countries provided by World Data Bank, I classified each country into developing or developed according to the definitions provided above and in accordance to the current WHO and WDB classifications. I then placed all the developed countries into the online random sample generator. The generator randomly assorted all of the countries in ascending order, in which I took the first 30 countries to be my sample. I repeated this same process for the developing countries list, which allowed me to collect a random sample of both my developing countries and developed countries. The online website used for this process was https://www.random.org/lists/ and served to be extremely effective and easily satisfied.

A: Describes the random sampling process to aid the flow.

[2] The World Data Bank. *Mortality rate, infant (per 1,000 live births).* 2015. http://data.worldbank.org/indicator/SP.DYN.IMRT.IN (accessed January 15, 2015).
[3] The World Data Bank. *GDP per capita (current US$)* . 2015. http://data.worldbank.org/indicator/NY.GDP.PCAP.CD (accessed January 15, 2015).

Univariate Data:

Table 1: The sample populations of thirty developed and thirty developing countries

Developed Countries	Developing Countries
1. Korea Republic	1. Bangladesh
2. Singapore	2. Malaysia
3. Denmark	3. Nepal
4. Chile	4. Botswana
5. Ireland	5. Uganda
6. Australia	6. India
7. Poland	7. Thailand
8. Netherlands	8. Sierra Leone
9. Portugal	9. Angola
10. Iceland	10. Iran
11. United States	11. Costa Rica
12. Canada	12. Nigeria
13. New Zealand	13. Liberia
14. Israel	14. Central African Republic
15. Czech Republic	15. Sri Lanka
16. Greece	16. Brazil
17. Switzerland	17. Bulgaria
18. Spain	18. Belarus
19. Italy	19. Mongolia
20. Sweden	20. Bolivia
21. Cyprus	21. Ethiopia
22. Germany	22. Cambodia
23. Austria	23. Indonesia
24. Finland	24. Moroccan
25. Slovenia	25. Argentina
26. United Kingdom	26. Zambia
27. Japan	27. Afghanistan
28. Belgium	28. China
29. France	29. Nicaragua
30. Luxembourg	30. Algeria

After generating a sample for the exploration, I extracted the GDP per capita in US dollars and the IMR per 1,000 births from World Data Bank using the 2013 statistics for both data values. I used this online statistical organisation, as it is the most reliable source, which collaborates with the World Health Organisation WHO to annually produce statistical health indicators of countries.

Table 2: The GDP per capita (US dollars) and IMR (per 1,000 births) of developed countries

Developed Countries	GDP – per capita (US dollars)	Infant Mortality rate per 1,000 live births
Korea Republic	25977	3.2
Singapore	55183	2.2
Denmark	59832	2.9
Chile	15732	7.1
Ireland	50503	3.2
Australia	67458	3.4
Poland	13648	4.5
Netherlands	50793	3.3
Portugal	21733	3.1
Iceland	47461	1.6
United States	53042	5.9
Canada	51958	4.6
New Zealand	41556	5.2
Israel	36051	3.2
Czech Republic	19845	2.9
Greece	21956	3.7
Switzerland	84815	3.6
Spain	29863	3.6
Italy	35953	3.0
Sweden	60430	2.4
Cyprus	25243	2.8
Germany	46269	3.2
Austria	50547	3.2
Finland	49147	2.1
Slovenia	23297	2.0
United Kingdom	41788	3.9
Japan	38634	2.1
Belgium	46878	3.5
France	42503	3.5
Luxembourg	110697	1.6

A: The data is neatly presented in tables with a key and titles.

Table 3: The GDP per capita (US dollars) and IMR (per 1,000 births) of developing countries

Developing Countries	GDP – per capita (US dollars)	Infant mortality per 1,000 live births
Bangladesh	958	33.2
Malaysia	10538	7.20
Nepal	694	32.2
Botswana	7315	36.3
Uganda	572	43.8
India	1499	41.4
Thailand	5779	11.3
Sierra Leone	679	107.2
Angola	5783	101.6
Iran	4763	14.4
Costa Rica	10185	8.40
Nigeria	3006	74.3
Liberia	454	53.6
Central African Republic	333	96.1
Sri Lanka	3280	8.20
Brazil	11208	12.3
Bulgaria	7499	10.1
Belarus	7576	3.70
Mongolia	4056	26.4
Bolivia	2868	31.2
Ethiopia	505	44.4
Cambodia	1007	32.5
Indonesia	3475	24.5
Moroccan	3093	26.1
Argentina	14715	11.9
Zambia	1845	55.8
Afghanistan	665	70.2
China	6807	10.9
Nicaragua	1851	30.0
Algeria	5361	21.6

Stage 1:

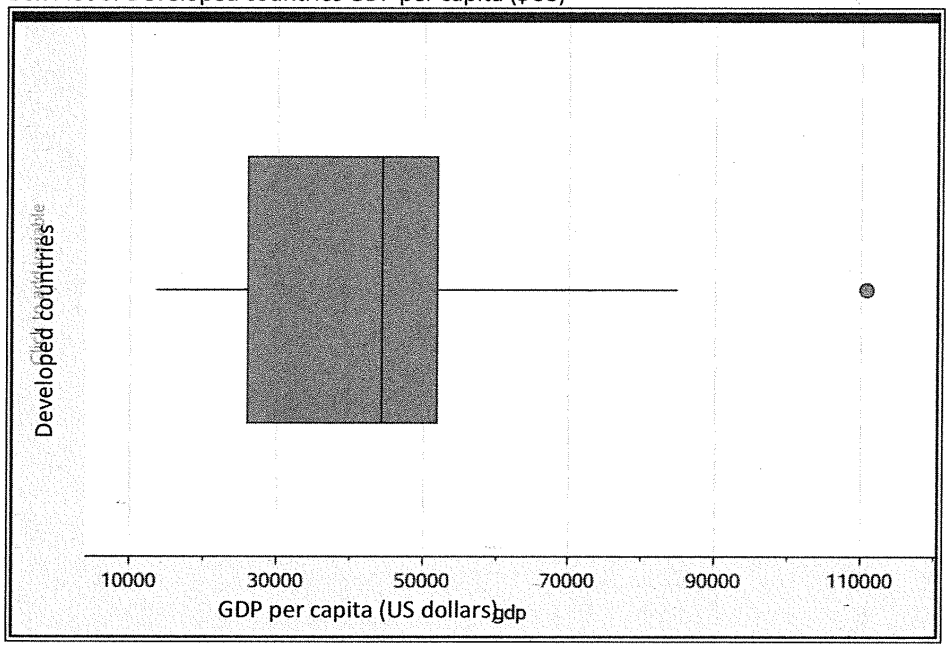
After obtaining a random sample of developed and developing countries, and their respective data values of IMR and GDP per capita, I decided the first statistical stage of my exploration would be to find out the distribution of their GDP per capita in US dollars. In order to find the distributions of developed and developing countries, I decided to use my Ti-Nspire CX calculator to individually graph the GDP per capita data set of my developing and developed countries in the form of a boxplot. Graphing the data in the form of a boxplot would then allow me to visually find the spread of the data and five-figure summaries. This information would allow me to contrast and compare the GDP per capita (US dollars) of developed and developing countries and gain a greater understanding of the distributions of each data set, developed and developing.

E: The calculations are correct.

Table 4: Five figure summaries of GDP per capita (US dollars)

Five Figure Summaries of GDP per capita (US dollars)		
	Developing Countries	Developed Countries
Minimum value	333	13648
Quartile 1 value	958	25997
Median value	3816.5	44386
Quartile 3 value	6807	51958
Maximum value	14715	110697

Graph 1:
Box Plot of Developed countries GDP per capita ($US)



E: Outlier correctly identified on graph.

Range= maximum value – minimum value
= 110697 - 13648
= $97049 US

IQR= $Q_3 - Q_1$
= 51958 – 25977
= $25981 US

Test for Outliers:

B: Terms used are not defined.

The upper boundary= upper quartile Q_3 + 1.5 x IQR
= 51958 + 1.5 x 25981
= $90929.5 US

- There is one outlier, as one value exceeds the upper boundary, this outlier has a GDP per capita value of $110697 US dollars

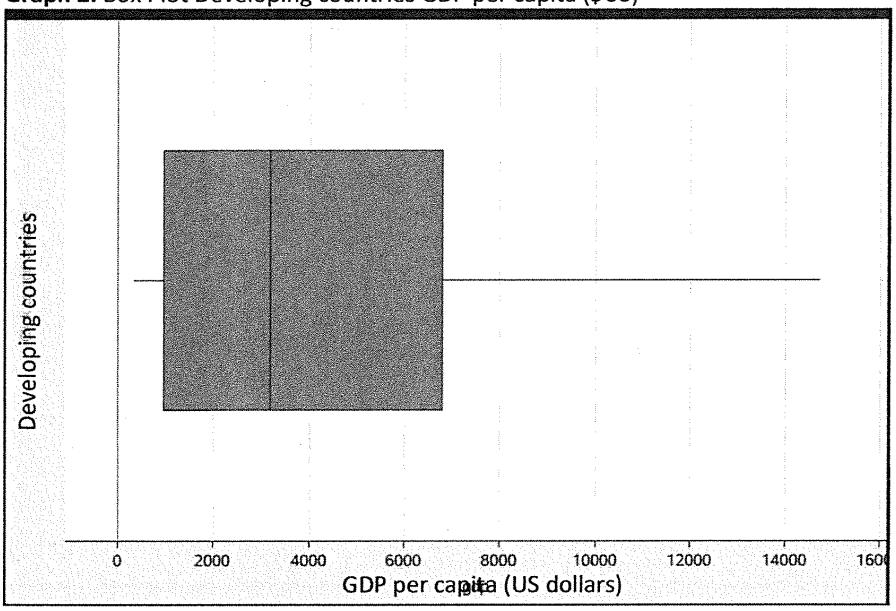
The lower boundary= lower quartile Q_1 – 1.5 IQR
= 25977– 1.5 x 25981
= $-12994.5 US

- There are no outliers as no values less than the lower boundary

The shape of this box plot is negatively skewed in the IQR with a median value of $44386 US dollars. The spread of the data has a range of $97049 US dollars and an IQR of $25981 US dollars.
There is an extreme value, as the GDP per capita value $110697 US dollars is considered an outlier as it exceeds the upper boundary of $85100.5 US dollars. This value belongs to Luxembourg. The reason Luxembourg is rendered an outlier is because upon further research it was learnt that Luxembourg has an extremely productive economy. The main reasons for Luxembourg's economic productivity is because they have a very small population whom are very well educated due to their thriving education system, thus unemployment rates are very low. In addition the government invests heavily, keeping taxes low creating a favourable financial environment within Luxembourg for investors. Thus overall Luxembourg's government debt to GDP ratio is one of the smallest in the world and lowest in Europe, as they do not spend much within the economy or military, which allows them to continue to prosper economically, furthermore explaining why Luxembourg has been rendered an outlier in my exploration.

D: The reason for the outlier has been researched and discussed.

Graph 2: Box Plot Developing countries GDP per capita ($US)



Range = maximum value – minimum value
= 14715 – 333
= $14382 US

IQR= $Q_3 - Q_1$
= 6807 - 958
= $584 US

Test for Outliers:
The upper boundary= upper quartile Q_3 + 1.5 x IQR
= 6807 + 1.5 x 5849
= $15580.5 US

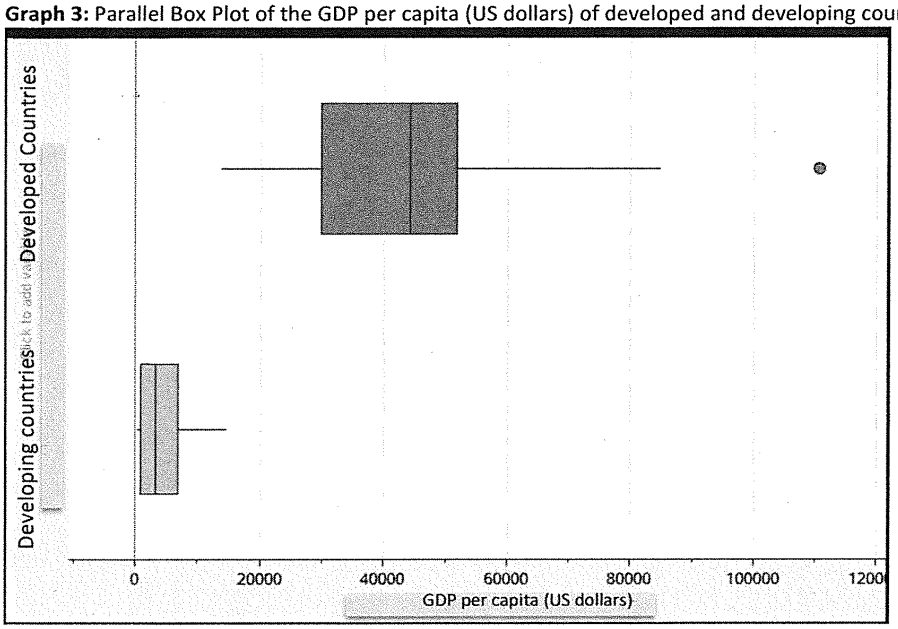
- There are no outliers as no values exceed the upper boundary

The lower boundary= lower quartile Q_1 – 1.5 IQR
= 958 – 1.5 x 5894
= $-4891 US

- There are no outliers as no values less than the lower boundary

The shape of this boxplot is positively skewed with a median value of $3816.5 US dollars GDP per capita. The spread of the data has a range of $14382 US dollars and an IQR of $5489 US dollars. There are no extreme values creating outlier's for this data set.
In order to compare the distributions of developed and developing countries GDP per capita, I created a parallel box plot of the two data sets to enable me to visually compare their respective distributions.

Graph 3: Parallel Box Plot of the GDP per capita (US dollars) of developed and developing countries



This parallel boxplot displays the difference of the GDP per capita distribution of developed countries above and the developing countries below. This boxplot shows my expected projection, that developed countries would indeed have greater GDP per capita's than developing countries, the reason for which is clear, as developed countries are by definition high-income countries with a GDP per capita greater than $12,000 US dollars. This conclusion is confirmed by the median value of $44386 US dollars for developed countries, compared to the lesser median of developing countries $3816.5 US dollars.

D: The student reflects upon the results shown in the box plots

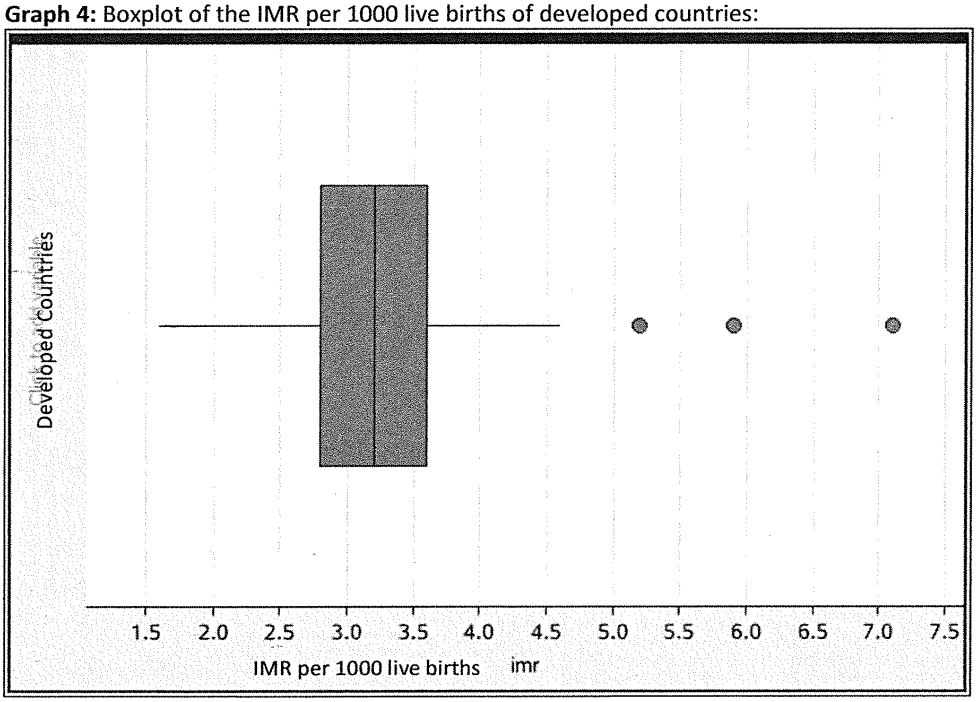
Stage 2:
The second statistical stage of my exploration was to find out the distribution of the infant mortality rate per 1,000 births (IMR) of developing and developed countries. In order to find the distributions I used my Ti-Nspire CX calculator to individually graph the IMR data set of my developing and developed countries in the form of a boxplot. Graphing the data in the form of boxplot would allow me to visually find the spread of the data and gather a five-figure summary, which would allow me to compare and contrast the IMR (per 1,000 births) of developed and developing countries.

Table 5: The five figure summaries for IMR per 1000 live births

Five Figure Summaries for Infant Mortality Rate (IMR) per 1000 live births		
	Developing Countries	Developed Countries
Minimum value	3.7	1.6
Quartile 1 value	11.9	2.8
Median value	30.6	3.2
Quartile 3 value	44.4	3.6
Maximum value	107.2	7.1

E: The calculations are correct.

Graph 4: Boxplot of the IMR per 1000 live births of developed countries:



**The outlier below the lower boundary represents the IMR of both Luxembourg and Iceland as they both have an IMR per 1,000 births value of 1.6

Range= maximum value – minimum value
= 7.1 – 1.6
= 5.5

IQR= $Q_3 - Q_1$
= 3.6 – 2.8
= 0.8

Test for Outliers:

The upper boundary= upper quartile Q_3 + 1.5 x IQR
= 3.6 + 1.5 x 0.8
= 4.80

- There are three outliers, which exceed the upper boundary. The IMR values belong to New Zealand (5.2) The United States (5.9) and Chile (7.1)

The lower boundary= lower quartile Q_1 – 1.5 IQR
= 2.8 – 1.5 x 0.8
= 1.60

- There are no outliers as no values less than the lower boundary

The shape of this boxplot is approximately symmetrically distributed with a median IMR of 3.2 per 1,000 births. The spread of the IMR has a range of 5.5 and an IQR of 0.8. There are a few extreme values, with three outliers positioned outside the upper boundary of 4.80, these values are represent New Zealand with an IMR of 5.2, The United States with an IMR of 5.9 and the most extreme value, Chile with an IMR of 7.1

One of the outliers exceeding the upper boundary, New Zealand can be explained, as Infant mortality rates are generally higher for Pacific and Māori infants. In particular in New Zealand there are significant socioeconomic inequalities within the Māori communities, which is reason for the spike in the IMR.

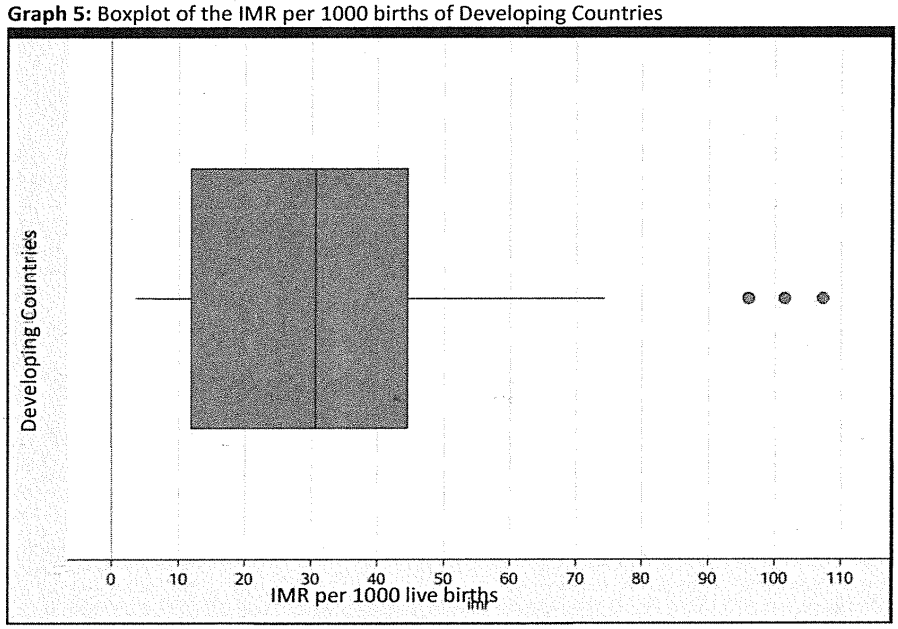
Another value beyond the upper boundary is The United States with an IMR of 5.9. A reason for this is because the US doesn't fund public healthcare. Therefore low-income families, which represent a proportion of the population, do not have access to healthcare, making families are less likely take there infant to the hospital initially, causing more mortalities among infants.

The most extreme outlier of 7.1 belongs to Chile. Chile is ranked first in infant mortality rate of high-income countries, which reflects the level of education, economic development accessibility to healthcare within Chile. The reason for which Chile is an outlier, is because the country has been plagued by the same profound disparities in the distribution of wealth, income and opportunity for years resulting in vast inequalities within the population and its infant mortality rate[4]

D: The student again reflects upon the results.

[4] World Health Organisation. *The Chilean infant mortality decline: improvement for whom? Socioeconomic and geographic inequalities in infant mortality, 1990–2005.* 2007. http://www.who.int/bulletin/volumes/85/10/06-041848.pdf (accessed January 15, 2015).

Graph 5: Boxplot of the IMR per 1000 births of Developing Countries



Range = maximum – minimum
= 107.2 – 3.7
= 103.5

IQR= $Q_3 - Q_1$
= 44.4 – 11.9
= 32.5

Test for Outliers:
The upper boundary= upper quartile Q_3 + 1.5 x IQR
= 44.4+ 1.5 x 32.5
= 93.15

- There are three outliers, which exceed the upper boundary. The IMR values belong to The Republic of Central Africa (96.1), Angola (101.6) and Sierra Leone (107.2)

The lower boundary= lower quartile Q_1 – 1.5 IQR
= 11.9 – 1.5 x 32.5
= -36.85

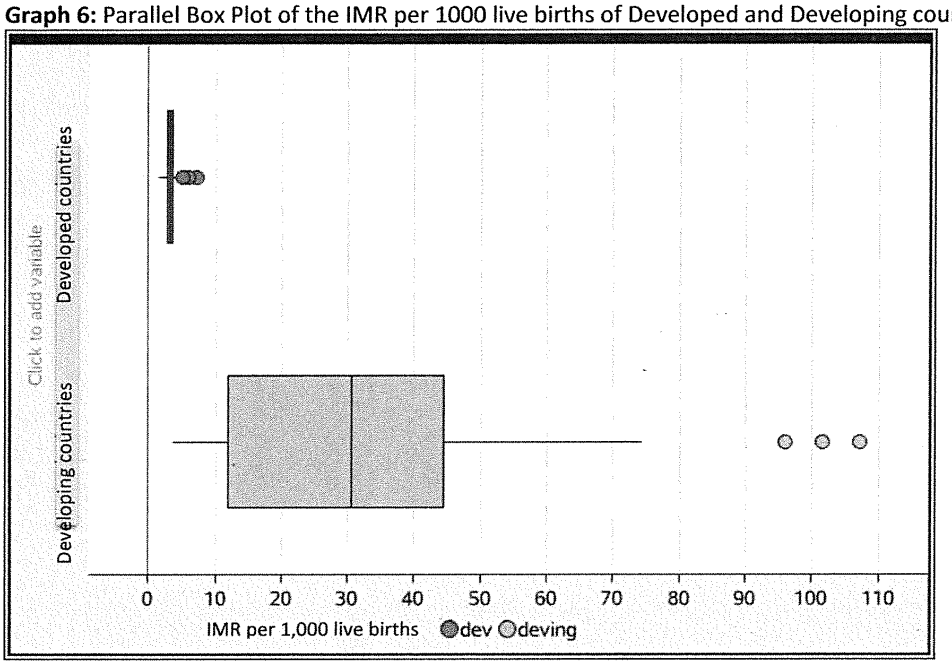
- There are no outliers as no values less than the lower boundary

The shape of this boxplot is positively skewed including its outliers with a median Infant Mortality Rate of 30.6 per 1,000 births. The spread of the IMR has a range of 103.5 and an IQR of 32.5. There are also three extreme values beyond the upper boundary of 93.15. These values are 96.1 infant deaths per 1000 in the Republic of Central Africa, 101.6 infant deaths per 1,000 in Angola and the greatest outlier of 107.2 infant deaths per 1,000 in Sierra Leone.

All of these outliers belong to extremely poor African countries, in which they have large populations and very littler health care resources to sustain the population, for these reasons they are outliers. These extreme values reflect the level of education, economic development and the poor accessibility and deficient quality of health care within these developing countries. Each of these countries also has an incredibly low Human Poverty Indexes (HPI) as there economic and social development is so sparse. And in addition, disease is extremely prevalent in particular HIV/AIDS, which has caused immeasurable amounts of mortalities within adults, children and infants particularly within Africa. Therefore for these reasons, Sierra Leone, Angola and the Republic of Central Africa are rendered outliers.

D: Each result is reflected upon.

Graph 6: Parallel Box Plot of the IMR per 1000 live births of Developed and Developing countries



This parallel boxplot displays the difference of IMR distribution of developed countries above and developing countries below. The boxplot shows my expected projection, that developed countries would indeed have a smaller IMR than developing countries, the reason for which is clear, as developed countries are by definition high-income countries, thus access to healthcare resources is supportable. This conclusion is confirmed by the average infant mortality rate of developed countries, being 3.2 deaths per 1,000, as opposed to the average infant mortality rate of developing countries of 30.6. This indicates that developing countries have an average infant mortality rate approximately 10 times greater than developed countries.

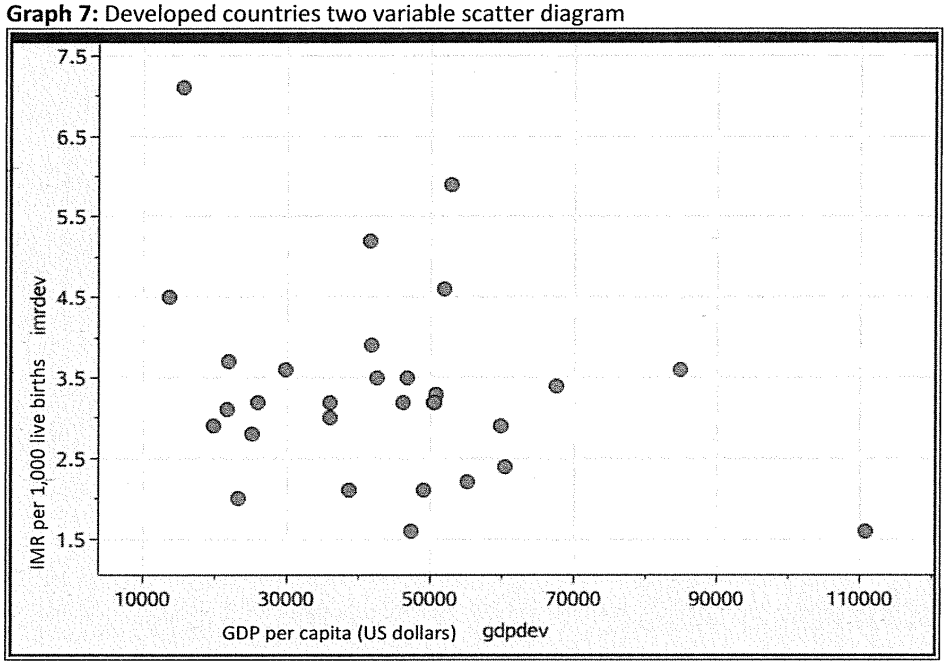
D: Meaningful reflection comparing the box plots.

Bivariate Data:

Stage 3:

The third stage of my exploration was to investigate the relationship between the two variables, infant mortality rate (per 1,000 live births) and GDP per capita (US dollars) within both developed and developing countries. Therefore I generated scatter diagrams of two variables, and examined the extent of the correlation, if any, using Pearson's correlation coefficient and how if there a correlation, I would generate an equation or general summary of the correlations found.

Graph 7: Developed countries two variable scatter diagram



Pearson's correlation coefficient:

$$r = \frac{\sum_{i=1}^{n}\left((x_i - \bar{x})(y_i - \bar{y})\right)}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

"Title"	"Linear Regression (mx+b)"
"RegEqn"	"m·x+b"
"m"	-0.000017
"b"	4.10386
"r²"	0.087285
"r"	-0.29544
"Resid"	"{...}"

E: Calculations are correct

r = - 0.30

The Pearson's correlation coefficient r shows the strength of linear regression, this value for developed countries is -0.30 indicating that there is a weak negative correlation between the GDP per capita (US dollars) and Infant mortality rate per 1000 births of developed countries.

The coefficient of determination r^2

$r^2 = 0.09$

The r^2 value indicates that 9% of the variation in infant mortality rate per 1000 births is explained by the variation in GDP per capita (US dollars). This also signifies that 91% of infant mortalities in developed countries can be explained by other factors.

E: Clear explanation showing good knowledge and understanding.

Least Squares Regression Line:

Using the least squares regression line, which is calculated on the Ti-Nspire CX calculator the equation of the line that best fits the data can be found.

IMR (per 1,000 births) = -0.000017 × GDP per capita (US dollars) + 4.10

Summary:

It can be deduced and summarized after examining the least squares regression line, that for every $1000 (US dollars) increase in GDP per capita the infant mortality rate per 1,000 births decreases by 0.017 deaths per 1,000 live births.

Using this equation I thought it would be interesting to explore and find the infant mortality of a country not used in my investigation and find their expected IMR per 1,000 births.

The country I decided to examine would be a developed nation that has a high-income GDP per capita deeming it a developed country. Therefore the country I selected was Norway.

Norway:

- High income country which renders it comfortably a developed country
- Its GDP per capita in US dollars was 100,898.4 in 2013[5]
- The population of Norway was recorded as 5.08 million in 2013[6]
- Norway is a European country which borders Finland and Russia[7]

IMR (per 1,000 births) = -0.000017 × GDP (per capita US dollars) + 4.10

Substitute Haiti's GDP per capita

Should be Norway.

= (-0.000017 × 100,898.4) + 4.10

= 2.38

Findings:

Using the least squares regression line I formulated earlier, Norway's expected IMR is 2.38 deaths per 1,000 live births.

However according to WHO and World Data Bank Norway's actual IMR for 2013 was exactly 2 deaths per 1,000 live infant births[8]

These results show that the least regression line has been an appropriate means to predict the IMR per 1,000 live births of a country, as there is only a small amount of error in this estimation. The percentage error for Norway's results using the least squares regression line was found to be only 19%.

E: Regression line used appropriately.

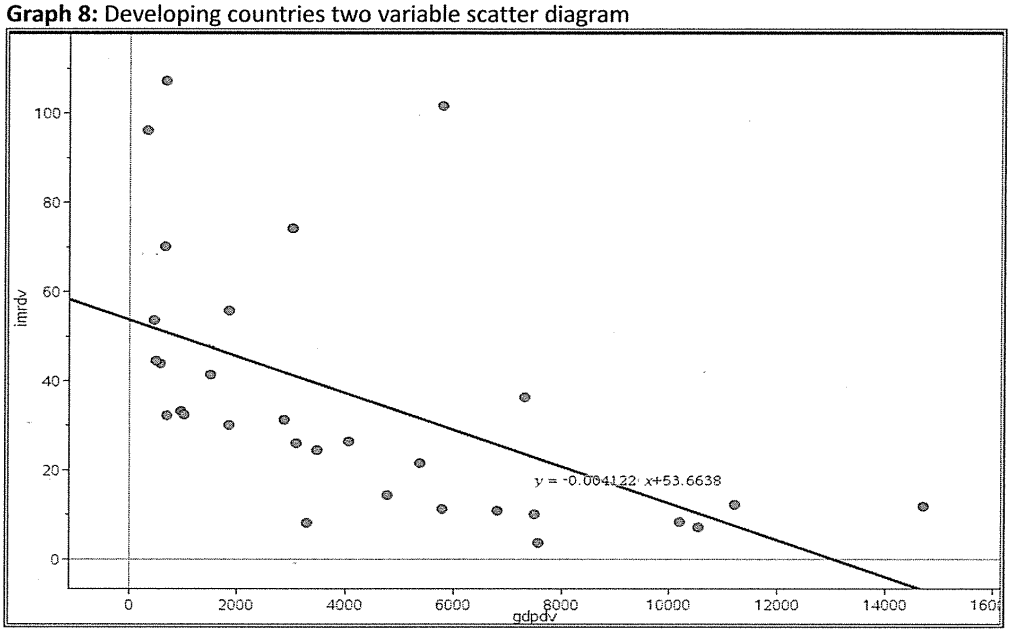
$$percentage\ error = \frac{estimated\ value - real\ value}{real\ value} \times \frac{100}{1}$$

$$percentage\ error = \frac{2.38 - 2}{2} \times \frac{100}{1}$$

$$percentage\ error = 19\%$$

[5] World Data Bank 2015
[6] ibid
[7] Wikipedia 2015
[8] World Data Bank 2015

Graph 8: Developing countries two variable scatter diagram



r value = -0.54 E: Correct calculation.

"Title"	"Linear Regression (mx+b)"
"RegEqn"	"m·x+b"
"m"	-0.004122
"b"	53.6638
"r²"	0.296038
"r"	-0.544094
"Resid"	"{…}"

The Pearson's correlation coefficient r shows the strength of linear regression, this value for developing countries is -0.54 indicating that there is a moderate negative correlation between the GDP per capita (US dollars) and Infant mortality rate per 1000 births of developed countries.

However Pearson's r is only valid for linear relationships, which do not have outliers. Thus this scatterplot appears to be non-linear and may not be an appropriate means to measure the relationship, if any, between GDP per capita and IMR of developing countries. Yet if it were valid the following analysis would be supported.

The coefficient of determination r^2

$r^2 = 0.30$

The r^2 indicates that 30% of the variation in infant mortality rate per 1000 live births can be explained by the variation in GDP per capita (US dollars)

Least Squares Regression Line

IMR *(per 1,000 births)* = m × GDP *(per capita US dollars)* + b

=(-0.00412 × GDP) + 53.7

$$\frac{rise}{run} = \frac{-0.004122}{1} \times \frac{1000}{1000}$$
$$= 4.12$$

Summary:

Overall is can be deduced that for every $1000 (US dollars) increase in GDP per capita, the infant mortality rate per 1,000 births will decrease by 4.12 deaths per 1,000 live births.

Finding the IMR of a country given their GDP per capita

Haiti

- Is a low-income country which renders it a developing country
- It is a Caribbean country
- Haiti has an approximate population of 10.32 million people[9] according to the 2013 World Data Bank statistics
- Its GDP per capita in 2013 $819.9 US dollars[10]

IMR (per 1,000 births) = (-0.00412 × GDP per capita) + 53.7

Substitute Haiti's GDP per capita

= (-0.00412 × 819.9) + 53.7

= 50.3

Findings:

Using the least squares regression line I formulated earlier, Haiti's expected IMR is 50.3 deaths per 1,000 live births.

However according to WHO and World Data Bank Haiti's actual IMR for 2013 was exactly 55 deaths per 1,000 live infant births[11]

These results show that the least regression line is an appropriate means to predict the IMR per 1,000 births for a country however there is some error in this estimation. The percentage error for Haiti's results using the least squares regression line was;

$$percentage\ error = \frac{estimated\ value - real\ value}{real\ value} \times \frac{100}{1}$$

$$percentage\ error = \frac{50.3 - 55}{55} \times \frac{100}{1}$$

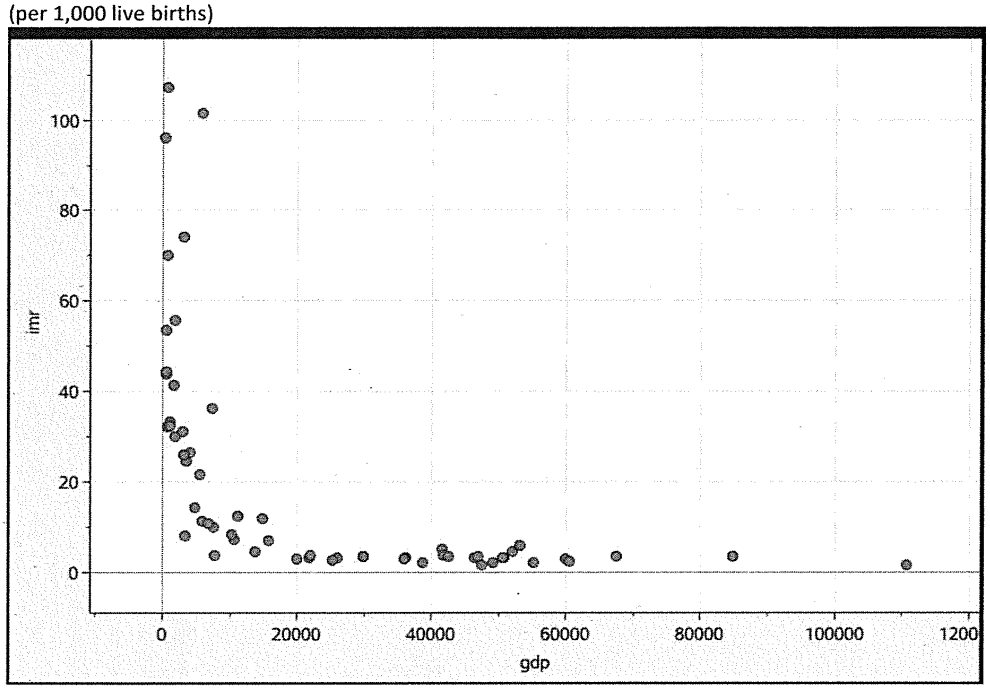
$$percentage\ error = 8.55\%$$

Stage 4:

After concluding that there is some form of a relationship between developed and developing countries GDP per capita and IMR, I wanted to then examine whether there is a general trend or relationship between IMR and GDP per capita. Therefore I merged all data, developed and developing values under the axes of IMR and GDP collectively.

[9] World Data Bank 2015
[10] ibid
[11] ibid

Graph 9: Scatter diagram of all countries (developed & developing) GDP per capita (US dollars) and the IMR (per 1,000 live births)



Stage 5:
However after generating this scatter diagram, it became evident that although there is a relationship between these variables it is a non-linear relationship, which I could not explain further using the data in this form. The relationship between GDP per capita and IMR per 1,000 as displays a logarithm relationship rather than a linear relationship. I considered exploring the data further using a logarithmic mathematics, yet I realised there was another method, known as the Spearman's ranking system, which would allow me to continue investigating my variables in a linear method. I therefore decided to rank the data using Spearman's ranking system, to see if this method would support me in generating a linear graph.

E: Understanding shown.

Spearman's rank order correlation equation:
Spearman's correlation coefficient, (ρ) measures the strength of association between two ranked variables and is derived by the Pearson's product-moment correlation

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

Table 6: Spearman's ranking Table

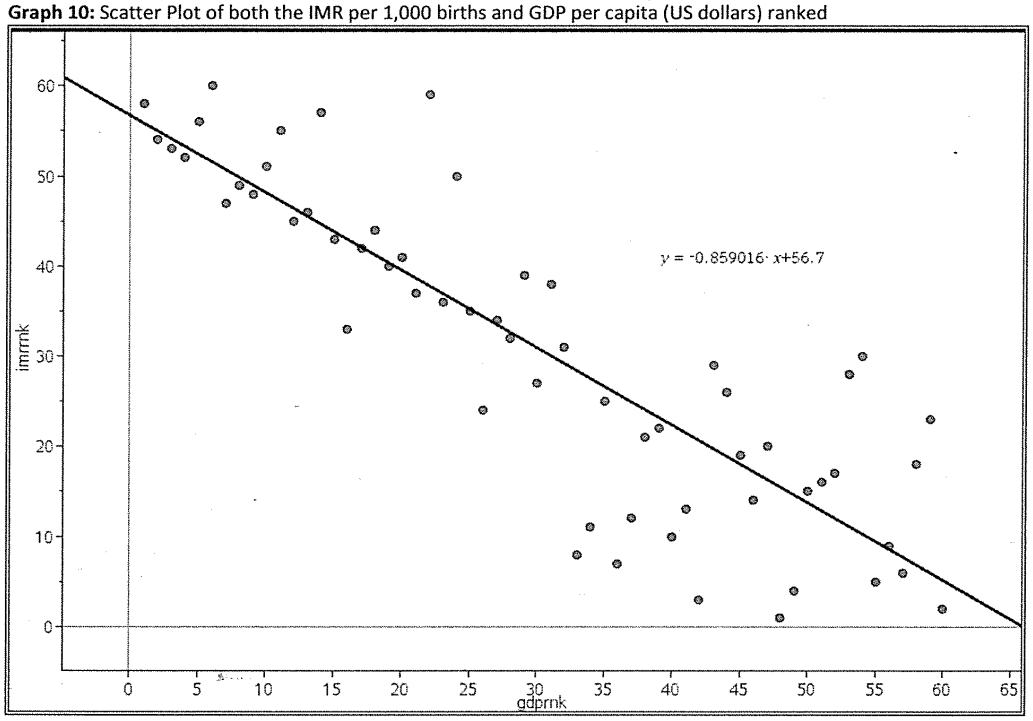
Country:	GDP (US dollars)	IMR (per 1,000 births)	Ranked GDP per capita	Ranked IMR (per 1,000 births)
Luxembourg	110697	1.6	60	2
Switzerland	84815	3.6	59	23
Australia	67458	3.4	58	19
Sweden	60430	2.4	57	7
Denmark	59832	2.9	56	10
Singapore	55183	2.2	55	6
United States	53042	5.9	54	30
Canada	51958	4.6	53	28
Netherlands	50793	3.3	52	18
Austria	50547	3.2	51	17
Ireland	50503	3.2	50	16
Finland	49147	2.1	49	5
Iceland	47461	1.6	48	1
Belgium	46878	3.5	47	21
Germany	46269	3.2	46	15
France	42503	3.5	45	20
United Kingdom	41788	3.9	44	26
New Zealand	41556	5.2	43	29
Japan	38634	2.1	42	4
Israel	36051	3.2	41	14
Italy	35953	3.0	40	11
Spain	29863	3.6	39	22
Korea Republic	25977	3.2	38	13
Cyprus	25243	2.8	37	8
Slovenia	23297	2.0	36	3
Greece	21956	3.7	35	25
Portugal	21733	3.1	34	12
Czech Republic	19845	2.9	33	9
Chile	15732	7.1	32	31
Argentina	14715	11.9	31	38
Poland	13648	4.5	30	27
Brazil	11208	12.3	29	39
Malaysia	10538	7.2	28	32
Costa Rica	10185	8.4	27	34
Belarus	7576	3.7	26	24
Bulgaria	7499	10.1	25	35
Botswana	7315	36.3	24	50
China	6807	10.9	23	36
Angola	5783	101.6	22	59
Thailand	5779	11.3	21	37
Algeria	5361	21.6	20	41
Iran	4763	14.4	19	40
Mongolia	4056	26.4	18	44
Indonesia	3475	24.5	17	42
Sri Lanka	3280	8.2	16	33
Moroccan	3093	26.1	15	43

E: Rankings for equal IMRs should have been averaged.

Nigeria	3006	74.3	14	57
Bolivia	2868	31.2	13	46
Nicaragua	1851	30	12	45
Zambia	1845	55.8	11	55
India	1499	41.4	10	51
Cambodia	1007	32.5	9	48
Bangladesh	958	33.2	8	49
Nepal	694	32.2	7	47
Sierra Leone	679	107.2	6	60
Afghanistan	665	70.2	5	56
Uganda	572	43.8	4	52
Ethiopia	505	44.4	3	53
Liberia	454	53.6	2	54
Central African Republic	333	96.1	1	58

The ranked data was calculated using Microsoft excel, using the rank equation mathematical program.

Graph 10: Scatter Plot of both the IMR per 1,000 births and GDP per capita (US dollars) ranked



Ranking key:

GDP ranking x-axis: The GDP per capita in US dollars is ranked in order from highest to lowest, meaning countries with high GDP per capita values have a high ranking. An example of this is the country Luxembourg; it has the highest GDP per capita ($110,697 US dollars) therefore it is ranked 60 for GDP per capita US dollars.

IMR ranked y-axis: The IMR per 1,000 births is ranked in order from to lowest to highest, meaning countries with low IMR per 1,000 births have a low ranking value. An example of this would be the country of Iceland as it has the lowest IMR per 1,000 births (1.6 deaths per 1,000 births) therefore it is ranked 1 for IMR.

B: Ranking explained.

E: Calculation is correct for ranked data.

Pearson's Product Correlation:

r= -0.86

"Title"	"Linear Regression (mx+b)"
"RegEqn"	"m·x+b"
"m"	-0.859016
"b"	56.7
"r²"	0.737909
"r"	-0.859016
"Resid"	"{...}"

This the Pearson r value indicates that there is a strong negative linear relationship between the ranked GDP per capita (US dollars) and the ranked IMR (per 1,00- births)

The Spearman's rank correlation graph above, demonstrates that although the logarithm graph produced earlier was not linear, there was indeed some form of a relationship between the GDP per capita and the IMR, and the ranked scatter graph shows that there is a monotonic relationship. A monotonic relationship is seen, as when the value of one variable increases, so does the value of the other variable, or as the value of one variable increases the other variable value decreases.

The graph above shows that as GDP per capita increased, so do the infant mortality rate, thus this is a monotonic relationship, in addition to also fitting the requirements of a strong linear negative correlation.

In addition the r^2 value indicates that 74% of the variation in infant mortality rate per 1000 births is explained by the variation in GDP per capita (US dollars), thus signifying that 74% of infant mortalities depends upon the GDP per capita of a country.

The Least Squares Regression line: form (y=mx+b)

Ranked IMR (per 1,000 births) = m (GDP per capita) + b

Ranked IMR (per 1,000 births) = -0.859 (GDP per capita) + 56.7

Using the line of best fit to find the ranked position of other countries:

Norway:

- GDP per capita 2013 $100,898.4 US dollars [12]

$$\begin{aligned}\textit{Ranked IMR (per 1,000 births)} &= -0.859\ \textit{(ranked GDP per capita)} + 56.7\\ &= -0.859 \times 59 + 56.7\\ &= 5.16\end{aligned}$$

Findings:

Norway's ranked IMR per 1,000 births was found to be 5.16. This rank can be found on the Spearman's Rank table, in which it correlates to the rank position of Finland, 5. Therefore using Finland's ranked IMR per 1,000 births, Norway would be expected to have an IMR of approximately 2.1 infant mortalities per 1,000 live births. According to WHO and World Data Bank Norway's actual IMR for 2013 was exactly 2 deaths per 1,000 live infant births[13], therefore the Spearman's Rank equation is a valid means to find the IMR per 1,000 births for Norway.

Moreover, there only 5% error between the estimated IMR per 1000 live births and the literal value of Norway's IMR per 1000 births of 2.0, which reinforces the validity of Spearman's Rank equation in estimating a countries expected IMR using its GDP per capita.

[12] World Data Bank 2015
[13] ibid

$$percentage\ error = \frac{estimated\ value - real\ value}{real\ value} \times \frac{100}{1}$$

$$percentage\ error = \frac{2.1 - 2}{2} \times \frac{100}{1}$$

$$percentage\ error = 5\%$$

Haiti:

- GDP per capita in 2013 $819.9 US dollars[14]

Ranked IMR (per 1,000 births) = -0.859 (ranked GDP per capita) + 56.7
= -0.859 × 7.5 +56.7
= 50.3

Findings:

Haiti's ranked IMR per 1,000 births was found to be 50.3. This rank can be found on the Spearman's Rank table, in which it correlates to the IMR ranked country of Botswana. Therefore using Botswana's ranked IMR per 1,000 births, Haiti would be expected to have an IMR per 1,000 live births of 36.3 using this Spearman's ranked equation. However according to WHO and World Data Bank Haiti's actual IMR for 2013 was exactly 55 deaths per 1,000 live infant births[15].

Moreover, it was found that in this estimation, there was an error of 34% between the estimated IMR per 1000 live births and the literal value of Haiti's IMR per 1000 births. This shows that The Spearman's Rank system has not been an appropriate method to estimate the actual IMR per 1000 births of Haiti, as the error involved shows a lack of reliability and validity.

$$percentage\ error = \frac{estimated\ value - real\ value}{real\ value} \times \frac{100}{1}$$

$$percentage\ error = \frac{36.3 - 55}{55} \times \frac{100}{1}$$

$$percentage\ error = -34\%$$

[14] World Data Bank 2015
[15] The World Data Bank. *Mortality rate, infant (per 1,000 live births).* 2015. http://data.worldbank.org/indicator/SP.DYN.IMRT.IN (accessed January 15, 2015).

Concluding Findings:

In conclusion throughout my exploration it was found that to some extent that IMR per 1,000 births does depend upon the GDP per capita (US dollars) of a country. This was shown in the Graph 8 *'Developing countries two variable scatter diagram'* as the Pearsons Product Correlation coefficient r was found to be -0.55 demonstrating a moderate negative correlation between GDP per capita (US dollars) and Infant mortality rate per 1000 live births. Moreover the coefficient determinate of r^2 for developing countries indicated that 30% of the variation in infant mortality per 1,000 births is explained by the variation in the GDP per capita (US dollars). This is also signifies that 70% of infant mortalities in developed countries can be explained by other factors. In addition the least squares regression line was also used as a means of modeling the expected IMR per 1,000 births using the equation found. The literal IMR per 1,000 births of Haiti was 55 deaths per 1,000. This literal value was extremely close to the expected calculation of 50.3 deaths per 1,000 births using the least squares regression line. This finding highlight that's the least regression equation is a valid means of predicting the IMR of developing countries with small error of 8.55%.

On Graph 7 *'Developed countries two variable scatter diagram'* the Pearsons Product coefficient r was found to be -0.30 indicating that there is a weak negative correlation between the GDP per capita (US dollars) and Infant mortality rate per 1,000 live births of developed countries. Despite the weak correlation reflected in the r value, the least squares regression was an appropriate means of modeling an equation to predict the IMR per 1,000 births for developed countries. This was highlighted when modeling the IMR of Norway, as the predicted value was 2.38 deaths per 1,000 live births in contrast to the literal value of 2 deaths per 1,000 births. This shows that the least regression line whilst was an appropriate means to predict the IMR per 1,000 births for developed countries, as the percentage error for Norway was only 19%, thus it can be concluded that the predictions were largely valid findings. Moreover the coefficient determinate r^2 reflected the weak correlation between IMR per 1,000 births and GDP per capita (US dollars) as only 9% of the variation in IMR per 1,000 births could be explained by the variation in GDP per capita (US dollars), also emphasizing that 91% of infant mortalities can be explained by other factors in developed countries. This may also suggest that IMR only depends upon the GDP per capita to an extent, underlining that at a certain point, increasing GDP per capita of a country does not cause a significant reduction in its IMR per 1,000 live births.

Moreover it was found that there is a significant difference in the GDP per capita of the developed and developing countries, which was found both in the univariate and bivariate mathematical processes used. It was found that for every 1000 (US dollars) increase in GDP per capita the infant mortality rate per 1,000 births decreased by 0.017 mortalities for high-income developed countries. These findings highlight that high wealth will decrease the infant mortalities within developed countries. In contrast for developing countries it was found that for every $1000 (US dollars) increase in GDP per capita, the infant mortality rate per 1,000 births decreases by 4.12 mortalities. These findings show that within both developing and developed countries increases in GDP per capita causes the IMR to decrease. However increase in GDP per capita whilst is less in developing regions, significantly reduces the infant mortalities. This finding could impact Sierra Leone substantially, as the IMR is 107.2 infant mortalities per 1,000 births. The current GDP per capita within Sierra Leone is only $679 (US), therefore if the GDP increased by $2000 (US) the IMR would decrease below 100 infants mortalities per 1,000 births, which would be a significant improvement in infant health within Sierra Leone.

D: Reflection upon the results.

Finally using the Spearman's Rank system the ranked IMR position of both Norway and Haiti could be obtained. This was found using the actual GDP per capita (US dollars) to find the expected GDP per capita rank position on the Spearman's Rank table. Then, using the expected GDP per capita rank, the Spearman's rank equation was used to find the expected IMR rank position, revealing the expected IMR per 1,000 live births of Haiti and Norway. This was a valid method of finding Norway's expected IMR per 1,000 births but not for Haiti. The Spearman's Rank system highlighted that to an extent IMR per 1,000 births does depend upon GDP per capita as outlined by Norway's findings.

If I could conduct this exploration again with unlimited resources and knowledge I would use all of the countries provided by World Data Bank rather than use a sample. In addition I would explore the data and the correlation found between GDP per capita US dollars and IMR per 1,000 live births using a logarithmic model. In this investigation using the r^2 value 9% of the variation in the IMR per 1,000 births in developed countries related to the GDP per capita. However this revealed that 91% of variation was caused by other factors, therefore an improvement would be to investigate the other factors which caused this variation.

Overall throughout this exploration it was found that to an extent Infant Mortality Rate per 1,000 live births does depend upon GDP per capita (US dollars). However it was also found that at a certain point, the wealth of country, represented by its GDP per capita (US dollars) does not cause a significant reduction in the IMR per 1,000 births of a country.

Bibliography

Quora. *What are the main sources of Luxembourg's wealth.* 2013. http://www.quora.com/What-are-the-main-sources-of-Luxembourgs-wealth (accessed January 15, 2015).

Random Org. *List Randomizer* . 2015. https://www.random.org/lists/ (accessed January 15, 2015).

The Perfect World Organisation. *Message of the Perfect World.* 2014. http://www.theperfectworld.org/living-standards-today.html (accessed January 15, 2015).

The World Data Bank. *GDP per capita (current US$)* . 2015. http://data.worldbank.org/indicator/NY.GDP.PCAP.CD (accessed January 15, 2015).

The World Data Bank. *County and Lending Groups.* 2015. http://data.worldbank.org/about/country-and-lending-groups (accessed January 15, 2015).

The World Data Bank. *Mortality rate, infant (per 1,000 live births).* 2015. http://data.worldbank.org/indicator/SP.DYN.IMRT.IN (accessed January 15, 2015).

The World Data Bank. *Population, total.* 2015. http://data.worldbank.org/indicator/SP.POP.TOTL (accessed January 15, 2015).

Wikipedia. *Norway.* 2015. https://en.wikipedia.org/wiki/Norway (accessed January 15, 2015).

World Health Organisation. *The Chilean infant mortality decline: improvement for whom? Socioeconomic and geographic inequalities in infant mortality, 1990–2005.* 2007. http://www.who.int/bulletin/volumes/85/10/06-041848.pdf (accessed January 15, 2015).

Appendix:

The countries World Data Bank provided data on for GDP per capita US dollars and IMR per 1,000 births. This list of countries was categorized into developed and developing countries based on the definitions provided by World Data Bank and then a sample size was obtained.

1. Afghanistan
2. Angola
3. Albania
4. Arab World
5. United Arab Emirates
6. Argentina
7. Armenia
8. Antigua and Barbuda
9. Australia
10. Austria
11. Azerbaijan
12. Burundi
13. Belgium
14. Benin
15. Burkina Faso
16. Bangladesh
17. Bulgaria
18. Bahrain
19. Bahamas, The
20. Bosnia and Herzegovina
21. Belarus
22. Belize
23. Bolivia
24. Brazil
25. Brunei Darussalam
26. Bhutan
27. Botswana
28. Central African Republic
29. Canada
30. Central Europe and the Baltics
31. Switzerland
32. Chile
33. China
34. Cote d'Ivoire
35. Cameroon
36. Congo, Rep.
37. Colombia
38. Comoros
39. Cabo Verde
40. Costa Rica
41. Caribbean small states

42. Cyprus
43. Czech Republic
44. Germany
45. Djibouti
46. Dominica
47. Denmark
48. Dominican Republic
49. Algeria
50. Ecuador
51. Egypt, Arab Rep.
52. Euro area
53. Eritrea
54. Spain
55. Estonia
56. Ethiopia
57. European Union
58. Finland
59. Fiji
60. France
61. Micronesia, Fed. Sts.
62. Gabon
63. United Kingdom
64. Georgia
65. Ghana
66. Guinea
67. Gambia, The
68. Guinea-Bissau
69. Equatorial Guinea
70. Greece
71. Grenada
72. Guatemala
73. Guyana
74. High income
75. Honduras
76. Croatia
77. Haiti
78. Hungary
79. Indonesia
80. India
81. Ireland
82. Iran, Islamic Rep.
83. Iraq
84. Iceland
85. Israel
86. Italy

87. Jamaica
88. Jordan
89. Japan
90. Kazakhstan
91. Kenya
92. Kyrgyz Republic
93. Cambodia
94. Kiribati
95. St. Kitts and Nevis
101. Korea, Rep.
102. Kuwait
103. Latin America & Caribbean
104. Lao PDR
105. Lebanon
106. Liberia
107. Libya
108. St. Lucia
109. Latin America & Caribbean
110. Low income
111. Sri Lanka
112. Lesotho
113. Lithuania
114. Luxembourg
115. Latvia
116. Morocco
117. Moldova
118. Madagascar
119. Maldives
120. Middle East & North Africa
121. Mexico
122. Marshall Islands
123. Middle income
124. Macedonia, FYR
125. Mali
126. Malta
127. Middle East & North Africa
128. Montenegro
129. Mongolia
130. Mozambique
131. Mauritania
132. Mauritius
133. Malawi
134. Malaysia
135. North America
136. Namibia

137. Niger
138. Nigeria
139. Nicaragua
140. Netherlands
141. Norway
142. Nepal
143. New Zealand
144. Oman
145. Other small states
146. Pakistan
147. Panama
148. Peru
149. Philippines
150. Palau
151. Papua New Guinea
152. Poland
153. Portugal
154. Paraguay
155. Pacific island small states
156. Qatar
157. Romania
158. Russian Federation
159. Rwanda
160. South Asia
161. Saudi Arabia
162. Sudan
163. Senegal
164. Singapore
165. Solomon Islands
166. Sierra Leone
167. El Salvador
168. Serbia
169. Sub-Saharan Africa
170. South Sudan
171. Sub-Saharan Africa
172. Small states
173. Sao Tome and Principe
174. Suriname
175. Slovak Republic
176. Slovenia
177. Sweden
178. Swaziland
179. Seychelles
180. Chad
181. Togo

182. Thailand
183. Tajikistan
184. Turkmenistan
185. Tonga
186. Trinidad and Tobago
187. Tunisia
188. Turkey
189. Tuvalu
190. Tanzania
191. Uganda
192. Ukraine
193. Uruguay
194. United States
195. Uzbekistan
196. St. Vincent and the Grenadines
197. Venezuela, RB
198. Vietnam
199. Vanuatu
200. World
201. Samoa
202. Yemen, Rep.
203. South Africa
204. Congo, Dem. Rep.
205. Zambia
206. Zimbabwe
207. Afghanistan