Mathematics Internal Assessment

IB Level

Mathematical Relationship between Food Security Index and Healthcare Index in a Country

Candidate's Name:

Candidate's Number:

Date of Submission:

Introduction

I did not like to follow world health patterns until the coronavirus outbreak in early 2020. Like in many other parts of the world, our country was not spared by the deadly disease. Within a few weeks, the government instituted the lockdown, and I had no choice but to remain at home. During this time, I developed a habit of following world news, especially those revolving around health patterns. I could read different online publications to grasp what was happening. I established that coronavirus spread was accelerated being accelerated by different factors, including old age, pre-existing conditions, and nutrition, among other factors. From one publication, I learned that adequate nutrition is essential for good health and the prevention of diseases. Poor nutrition can lead to various health problems, including malnutrition, anemia, and stunted growth in children. Good nutrition can help to improve immune function, reduce the risk of certain chronic diseases, and promote overall health and well-being. Besides this, there are also linkages between food security and healthcare in terms of the economic and social factors that can impact both. Healthcare systems and policies can impact food security, as they may influence the availability and affordability of food and the ability of people to access it. I could not, however, ascertain this information. Fortunately, in higher mathematics class, I had an opportunity to learn different statistical concepts, including the chi-square, correlation coefficients, and regression analysis, among others. Of more importance, I learned that these concepts could be applied in different real-life applications. Thus, to gain more understanding, I decided to align my mathematics investigation with these concepts, aiming to determine the mathematical relationship between food security and healthcare indexes in a country.

The aim and Statement of the Plan

This mathematical investigation aims to find out the exact mathematical relationship between food security index and the healthcare index in a country. First is statistical data on the food security and healthcare indexes from online databases. This data will then be subjected to an outlier test to identify and select the datasets that do not conform with the rest of the dataset. The remaining dataset will then be subjected to statistical tests of chi-square and correlation to determine whether any interdependence exists between the food security index and healthcare index and the nature of the relationship, respectively.

Data Collection

The food security and healthcare index data were collected from Economist Impact and Numbeo online databases, respectively. A sample of 20 randomly selected countries from different regions was used. Additionally, the base year selected was 2022, a year without many issues of the coronavirus pandemic that may have impacted the health indices. The collected data were recorded in a raw table, presented in table 1 of this investigation below.

Table 1: Data of Food Security and Healthcare Index in 20 Selected Countries in 2022

Country	Food Security Index	Health Care Index
Norway	80.50	76.83
Netherlands	80.10	75.56
United Kingdom	78.80	74.83
Portugal	78.70	71.97
United States	78.00	69.06
Australia	75.40	78.14
United Arab Emirates	75.20	68.73
China	74.20	66.40
Uruguay	71.80	68.36
Saudi Arabia	69.90	60.92
Indonesia	60.20	60.44
Brazil	65.10	57.84
Morocco	63.00	46.69
South Africa	61.70	63.97
Ukraine	57.90	53.34
Pakistan	52.20	59.58
Chile	74.20	63.97
Bulgaria	73.00	56.41
Peru	78.80	52.20
Argentina	64.80	69.32

https://www.numbeo.com/health-care/rankings_by_country.jsp?title=2022

https://impact.economist.com/sustainability/project/food-security-index/

Interquartile Outlier Test

The data in table 1 above will be used in this section to find out the outliers using the general interquartile formula from the mathematical study of Courtney (2018).

$$Upper\ Limit,\ UL = 1.5\ IQR + Upper\ Quartile, UQ$$
$$Lower\ Limit,\ LL = Lower\ Quartile, LQ - 1.5\ IQR$$

Where;

$$IQR = Interquartile\ Range$$

The lower limit, the upper limit, the maximum, and the minimum food security and healthcare indices of the datasets were then calculated using the Microsoft Excel application and recorded in a summary table below.

Table 2: Summary Table for Food Security and Healthcare Indices

Quantity	Food Security Index	Healthcare Index
Lower Quartile, LQ	64.35	59.15
Upper Quartile, UQ	78.18	69.19
Maximum	80.50	78.14
Minimum	52.20	46.19

Considering the dataset of food security;

$$Interquartile\ range,\ IQR = 78.18 - 64.35 = 14.43$$
$$Upper\ Limit,\ UL = 1.5\ (14.43) + 78.18 = 98.38$$
$$Lower\ Limit,\ LL = 64.35 - 1.5\ (14.43) = 42.75$$

These results show that all the food security data falls within the accepted range of 42.75 and 98.38, considering that the minimum and maximum food security indices are 52.20 and 80.50; hence no significant outlier.

Considering the dataset of the healthcare index:

$$Interquartile\ range,\ IQR = 69.98 - 59.15 = 10.83$$

$$Upper\ Limit,\ UL = 1.5\ (10.83) + 69.18 = 85.43$$

$$Lower\ Limit,\ LL = 59.15 - 1.5\ (10.83) = 42.91$$

These results show that all the healthcare index data falls within the accepted range of 42.91 to 85.43, considering that the minimum and maximum healthcare indices are 46.69 and 78.14; hence no significant outlier.

Chi-Square Test

The chi-square statistical tool is primarily used to determine the nature of interdependence between any two categorical and measurable datasets, independent or not independent. The independent choice works with the null hypothesis, while the no independent choice works with the alternative hypothesis. Additionally, the null hypothesis is usually deemed to be supported if a chi-square calculated is more in value than the corresponding chi-square from the distribution table. With respect to this internal assessment, the null hypothesis and the alternative hypothesis would be defined as;

Null Hypothesis, H_0: The food security index and healthcare index in a country are independent of each other.

Alternative Hypothesis, H_1: The food security index and healthcare index in a country are not independent (are dependent) of each other.

The general mathematical relation for computing the chi-square calculated was retrieved from the mathematical study of BMJ (2019).

$$Chi - square,\ X_{calc.} = \frac{\left(F_o - F_E\right)^2}{F_E}$$

$$F_o = Observed\, measurements$$

$$F_E = Expected\, measurements$$

The observed and expected measurements of the healthcare index, based on the food security index, would be calculated through a 3 by 3 table of contingency;

Table 3: Observed Healthcare Index Measurements based on Food Security Index

Food Security Index	Healthcare Index			Sum
	$45.0 \le y \le 58.0$	$58.1 \le y \le 71.1$	$71.2 < y \le 84.2$	
$50.5 \le x \le 60.5$	53.34	120.02	0.00	**173.36**
$60.6 \le x \le 70.6$	46.69	252.05	0.00	**298.74**
$70.7 \le x \le 80.7$	108.61	268.16	377.33	**754.10**
Total	**208.64**	**640.23**	**377.33**	**1226.20**

The results of observed measurements were used to compute the expected healthcare index through a general formula;

$$Expected\ \ Index = \frac{Sum\, of\, Healthcare\, index\, in\, a\, row \times Sum\, of\, Healthcare\, index\, in\, a\, column}{effective\, Sum\, of\, Healthcare\, index}$$

For a sample mathematical calculation, taking the observed healthcare index that had been obtained in the second row and second column;

$$Expected\ Index = \frac{298.74 \times 640.23}{1226.20} = 155.97$$

Table 4: Expected Healthcare Index Measurements based on Food Security Index

Food Security Index	**Healthcare Index**			**Sum**
	$45.0 \leq y \leq 58.0$	$58.1 \leq y \leq 71.1$	$71.2 < y \leq 84.2$	
$50.5 \leq x \leq 60.5$	29.50	90.52	53.35	**173.36**
$60.6 \leq x \leq 70.6$	50.83	155.98	91.93	**298.74**
$70.7 \leq x \leq 80.7$	128.31	393.74	232.05	**754.10**
Total	**208.64**	**640.23**	**377.33**	**1226.20**

From the data presented in tables 3 and 4:

$$X_{calc.} = \frac{(53.34 - 29.50)^2}{29.50} + \frac{(120.02 - 90.52)^2}{90.52} + \ldots + \frac{(377.33 - 232.05)^2}{232.05}$$
$$= 367.70$$

The critical value of the same chi-square would be calculated using a degree of freedom (DOF) as a function of columns, and rows;

$$DOF = \left(Rows'Number - 1\right)\left(Columns'Number - 1\right)$$
$$= (3 - 1) \times (3 - 1)$$
$$= 4$$

Using a DOF value of 4, and an alpha value of 0.05, the critical value of chi-square would be read from the distribution table;

$$X_{critical} = 9.49$$

The two different values of chi-square obtained under this section were then compared;

$$_{Scale.} > X_{critical}$$

Based on this clear difference, there was sufficient evidence to reject the null hypothesis and accept the alternative hypothesis, indicating that the food security index and healthcare index in a country are not independent (are dependent) of each other.

Determination of Correlation Coefficient

Scatterplot Diagram

From the preceding section, it had been established that a country's food security and healthcare index are not independent. Thus, a statistical relationship could be determined between the two datasets. A scatterplot diagram is one of the main statistical tools that can be used to find out this relationship. The healthcare index values were plotted against food security index values using the Microsoft Excel application, as illustrated in the figure below.

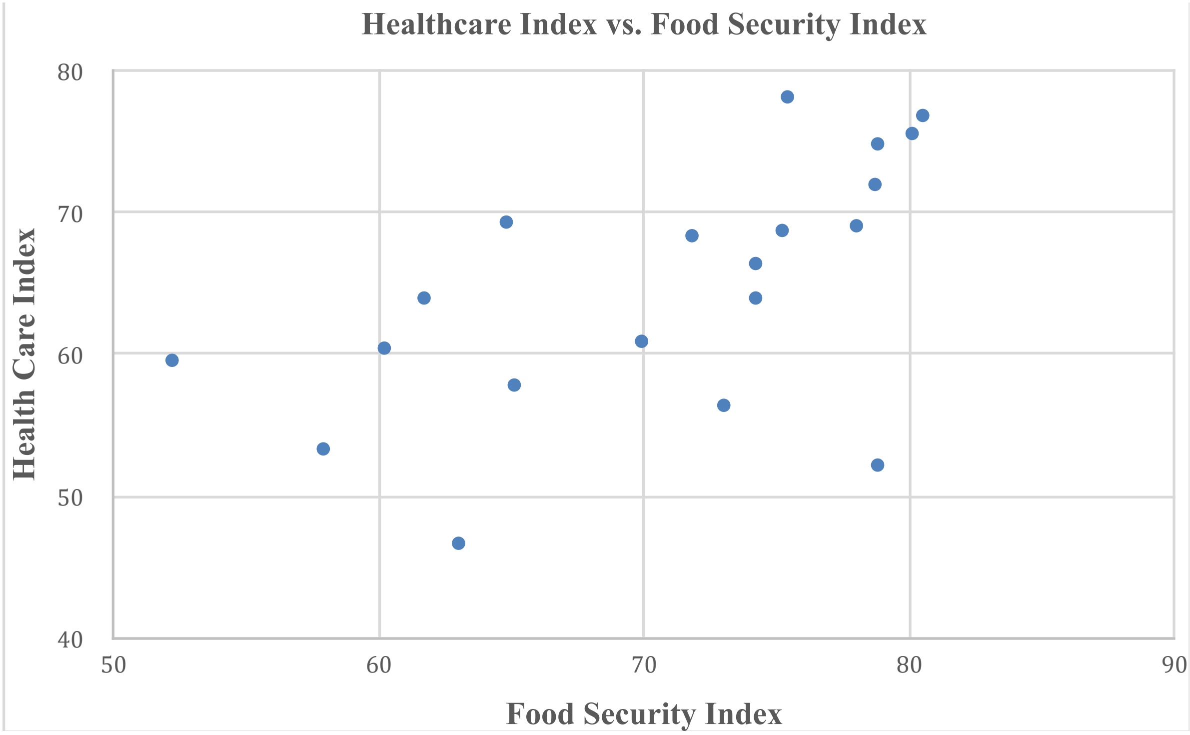

Figure 1: A Scatterplot of the Healthcare Index against the Food Security Index

The scatterplot in figure 1 above demonstrates a correlation between a country's healthcare and food security index since the data points follow a particular trend toward one direction (from the lower left to the upper right). However, it would be important to use a further statistical test to establish the exact value of correlation for a conclusive determination of a relationship between the pair of parameters or variables.

Pearson Product Moment Method

Pearson's product-moment method is used to find the exact correlation or relationship level between two categorical and measurable datasets. With respect to this investigation, the method would be used to find the correlation or relationship between the healthcare and food security indexes in a country. According to the study of Turney (2022), Pearson's product-moment correlation method is such that;

$$Correlation\ Coefficient,\ r = \frac{S_{xy}}{\sqrt{S_{xx}}\sqrt{S_{yy}}}$$

$$S_{xx} = \sum x^2 - \frac{\left(\sum x\right)^2}{N}$$

$$S_{yy} = \sum y^2 - \frac{\left(\sum y\right)^2}{N}$$

$$S_{xy} = \sum xy - \frac{\sum x \sum y}{N}$$

$N = The\ size\ of\ the\ sample;\ in\ this\ mathematics\ investigation,\ N\ is\ 20\ relating\ to\ the\ 20\ countries$

Additionally, the dependent, and the independent variables will be taken as:

$$x = Food\ Security\ Indices$$

$$y = Healthcare\ indices$$

The data in table 1 was then used to summarize the summations of the abovementioned variables.

Table 5: Summary Table for the Pearson's Correlation

x	Y	x^2	y^2	xy
80.50	76.83	6480.25	5902.85	6184.82
80.10	75.56	6416.01	5709.31	6052.36
78.80	74.83	6209.44	5599.53	5896.60
78.70	71.97	6193.69	5179.68	5664.04
78.00	69.06	6084.00	4769.28	5386.68
75.40	78.14	5685.16	6105.86	5891.76
75.20	68.73	5655.04	4723.81	5168.50
74.20	66.40	5505.64	4408.96	4926.88
71.80	68.36	5155.24	4673.09	4908.25
69.90	60.92	4886.01	3711.25	4258.31
60.20	60.44	3624.04	3652.99	3638.49
65.10	57.84	4238.01	3345.47	3765.38
63.00	46.69	3969.00	2179.96	2941.47
61.70	63.97	3806.89	4092.16	3946.95
57.90	53.34	3352.41	2845.16	3088.39
52.20	59.58	2724.84	3549.78	3110.08
74.20	63.97	5505.64	4092.16	4746.57
73.00	56.41	5329.00	3182.09	4117.93
78.80	52.20	6209.44	2724.84	4113.36
64.80	69.32	4199.04	4805.26	4491.94
$\sum x = 1413.50$	$\sum y = 1294.56$	$\sum x^2 = 101228.79$	$\sum y^2 = 85253.48$	$\sum xy = 92298.74$

$$S_{xx} = 101228.79 - \frac{(1413.50)^2}{20} = 1329.68$$

$$S_{yy} = 85253.48 - \frac{(1294.56)^2}{20} = 1459.20$$

$$S_{xy} = 92298.74 - \frac{(1413.50)(1294.56)}{20} = 805.71$$

$$Correlation\ Coefficient,\ r = \frac{805.71}{\sqrt{1329.68}\sqrt{1459.20}}$$

$$= \frac{805.71}{36.46 \times 38.20}$$

$$= \frac{805.71}{1392.78}$$

$$= 0.5785$$

Pearson's correlation coefficient obtained; 0.5785 was greater in value than 0.5, indicating a fairly positive relationship between the food security and healthcare indexes.

Test of Significance of the Correlation Coefficient

The value of the Pearson's correlation coefficient obtained above would be tested for significance through a comparison mathematical test between the t-value calculated, and the t-critical value. If the t-value calculated is greater in size than the t-critical value, then the correlation coefficient would be considered statistically significant; otherwise, the correlation coefficient would be considered not statistically significant (Turney, 2022). The null hypothesis and the alternative hypothesis for this situation were first defined such that;

$$Null\ Hypothesis,\ H0:\ The\ Pearson\ correlation\ is\ not\ statistically\ significant, \rho = 0$$

$$Alternative\ Hypothesis,\ H1:\ The\ Pearson\ correlation\ is\ \ statistically\ significant, \rho \neq 0$$

The general formula for finding the t-value from the mathematics study of Turney (2022) will be used;

$$t_{calc.} = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

$$t_{calc.} = \frac{0.5785}{\sqrt{\frac{1-0.5785^2}{20-2}}}$$

$$= \frac{0.5785}{\sqrt{0.19226}}$$

$$= 1.319$$

For the t-critical value;

$$The\ degree\ of freedom,\ DOF = n - 2$$

$$= 20 - 2$$

$$= 18$$

The t-distribution table with DOF of 18, and taking a significance level of 0.05;

$$t_{crit.} = 1.734$$

Comparing the two values of t-values:

$$t_{crit.} > \ t_{calc.}$$

Since the t-value calculated was smaller in size than the t-critical value, the null hypothesis was deemed to be supported, indicative that the correlation coefficient defining the relationship between the food security and healthcare index was statistically insignificant. Based on this finding, the correlation coefficient could not be used to model the trend between the two datasets.

Conclusion and Evaluation

This investigation has been aimed at finding out the exact mathematical relationship that exists between between food security index and the healthcare index in a country. Different tools of statistics were used. From the chi-square test, it was established that the food security index and healthcare index in a country are not independent (are dependent) of each other. The scatterplot diagram and Pearson's correlation method revealed a fairly positive relationship between the food security index, and healthcare index, even though the relationship was found to be statistically insignificant to be used to model the relationship. The main strength of this investigation was the use of a large sample size of data, 20 countries, a factor that increased the accuracy of the correlation coefficient. However, using only one year's datasets limited the investigation's validity. It would have been better to use at least food security, and healthcare indices from at least three years. Overall, the objective of this investigation was fully met. A suitable extension would be finding the mathematical relationship between the crime rate and food security index through the same approach.

References

BMJ. (2019). *8. The Chi-squared tests*. https://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/8-chi-squared-tests

Courtney, T. (2018). *Detect the Presence of Outliers with the Interquartile Range Rule*. ThoughtCo. https://www.thoughtco.com/what-is-the-interquartile-range-rule-3126244

Turney, S. (2022, May 13). *Pearson Correlation Coefficient (r) | Guide & Examples*. Scribbr. https://www.scribbr.com/statistics/pearson-correlation-coefficient/