[redacted]

COUNTRIES AND THEIR NATIONAL PLANS ADDRESSING VIOLENCE

INTRODUCTION TO PROBABILITY AND STATISTICS

EAST CENTRAL UNIVERSITY

DR. NICHOLAS JACOB

06/05/2020

COUNTRIES AND THEIR NATIONAL PLANS ADDRESSING VIOLENCE

Table A2: General information and national action plans addressing violence

Country/area	GENERAL INFORMATION				NATIONAL ACTION PLANS								
	Population (2012)[b]	Gross national income per	Income level[c]	Income inequality (Gini index)[d]	Interpersonal violence	Armed violence	Gang violence	Organized crime	Child maltreatment	Youth violence	Intimate partner violence	Sexual violence	Elder abuse
Afghanistan	29824536	690	Low	27.82	No	No	No	Subnational	No	Don't know	No	Subnational	No
Albania	3162083	4520	Middle	34.51	Yes	Don't know	Don't know	Don't know	Yes	Don't know	Yes	Yes	Yes
Algeria	38481705	4970	Middle	35.33	Yes	Yes	Subnational	Yes	Yes	Yes	Yes	Yes	Yes
Armenia	2969081	3770	Middle	31.30	No	No	No	No	Yes	No	Yes	Yes	No
Australia	23050471	59790	High		Yes	No	Yes	Yes	Yes	Subnational	Yes	Yes	Subnationa
Austria	8463948	47960	High	29.15	No	No	No	No	No	Yes	No	No	No
Azerbaijan	9308959	6290	Middle	33.71	Yes	No	Subnational	No	No	No	Subnational	Subnational	No
Bahrain	1317827	19560	High		Yes	Yes	Yes	Yes	Yes	Yes	Subnational	Subnational	Yes
Bangladesh	154695368	830	Low	32.12	No	No	No	No	No	No	No	No	No
Belarus	9405097	6400	Middle	26.48	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Belgium	11060095	44810	High	32.97	No	No	No	No	Yes	No	Yes	Yes	Yes
Belize	324060	4620	Middle	53.13	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Benin	10050702	750	Low	38.62	Subnational	No	Subnational	Subnational	Yes	Yes	Subnational	Yes	No
Bhutan	741822	2420	Middle	38.73	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Bolivia (Plurinational	10496285	2220	Middle	56.29	Yes	Yes	Subnational	Subnational	Yes	Subnational	Yes	Yes	Yes
Botswana	2003910	7650	Middle		No	Subnational	Subnational	Subnational	Yes	Subnational	Yes	Yes	No
Brazil	198656019	11640	Middle	54.69	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Brunei	412238	31590	High		No	No	No	No	No	No	No	No	No
Bulgaria	7277831	6850	Middle	28.19	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes
Burkina Faso	16460141	670	Low	39.79	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Burundi	9849569	240	Low	33.27	No	No	No	No	No	No	Yes	Yes	No
Cambodia	14864646	880	Low	36.03	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Above, is a segment of a data set that shows a list of countries and the existence of national plans against certain crimes. There are about 100 countries listed as well as their population and their capita income. I found the data set on the World Health Organization website under violence and injury prevention. I then downloaded the data set provided on the website and converted it to an excel document. I find this data set interesting because it looks at different countries and whether or not they have in place plans to combat certain crimes. This would show which countries think certain crimes need to have policies against them. There has always been an issue of countries following plans/suggestions by WHO and those who do not. Variables included in the data set are country, population, gross national income per person, income level, income inequality, Interpersonal violence, armed violence, Gang violence, Organized crime, Child maltreatment, Youth violence, Intimate partner violence, Sexual violence, Elder abuse.

This data set has a lot of variables that can be compared to each other, particularly I would like to compare if a country's gross national income per person motivates the government to make plans against sexual violence crimes. The variety of crimes in it provides more than enough variables to work with. The country, their national plans, income level would be categorical variables while population, gross national income per person, income inequality would be quantitative variables. The country, national plan for sexual violence would be nominal variables. Income level would be an ordinal variable. GNI, population, and income inequality would be continuous. The explanatory variable is the gross national income per person. The response variable is the national plan for sexual violence. I want to find out if a nation's income affects their government to create response plans to sexual violence crimes.

REFERENCE

World Health Organization. (2014) Table A2. Retrieved from https://www.who.int/violence_injury_prevention/violence/status_report/2014/data/en/

Link to pdf: https://www.who.int/violence_injury_prevention/violence/status_report/2014/data/Table_A2_General_info.pdf?ua=1

PROJECT 2 06/12/2020 CORRECTED ON 06/21/20

Action **on I. V**	Frequency	Relative F.
Interpersonal v.	68	0.55737705
NO interpersonal V.	54	0.44262295
Total	122	

Income Level	Frequency	Relative F.
High	39	0.295454545
Middle	73	0.553030303
Low	20	0.151515152
Total	132	

Action on Sexual V.	Frequency	Relative F.
Sexual V.	87	0.725
NO Sexual V.	33	0.275
Total	120	

TWO TABLE ANALYSIS

		Interpersonal No	Sexual Violence No	Total No
	122	54	33	87
Interpersonal Yes	68	122	101	
Sexual Violence Yes	87	141	120	
Total yes	155			

The first three tables show the frequencies of select categorical variables from my data set. The first and third table sample the number of countries whose governments have plans to fight interpersonal violence and sexual violence respectively. I made use of the count if function on excel and counted the number of yes's and no's from the data set. I then found the relative frequencies of the values. More countries have more national plans to fight against Interpersonal violence, the value is however a little above half. Sexual violence on the other hand has more countries with national plans. This may mean that more countries pay attention to sexual violence compared to interpersonal violence. I also sampled the frequencies of the income levels of the countries. The result showed that a majority of the countries are in the middle class, followed by high class and lastly, low class. The two way variable table Shows that there are more yes's compared no's. I used the "what if analysis" under stat in excel. Take into account some "I don't know values" and "subnational". This may show that in general there are more countries with national plans against Interpersonal violence and Sexual Violence.

PROJECT 3 06/21/2020

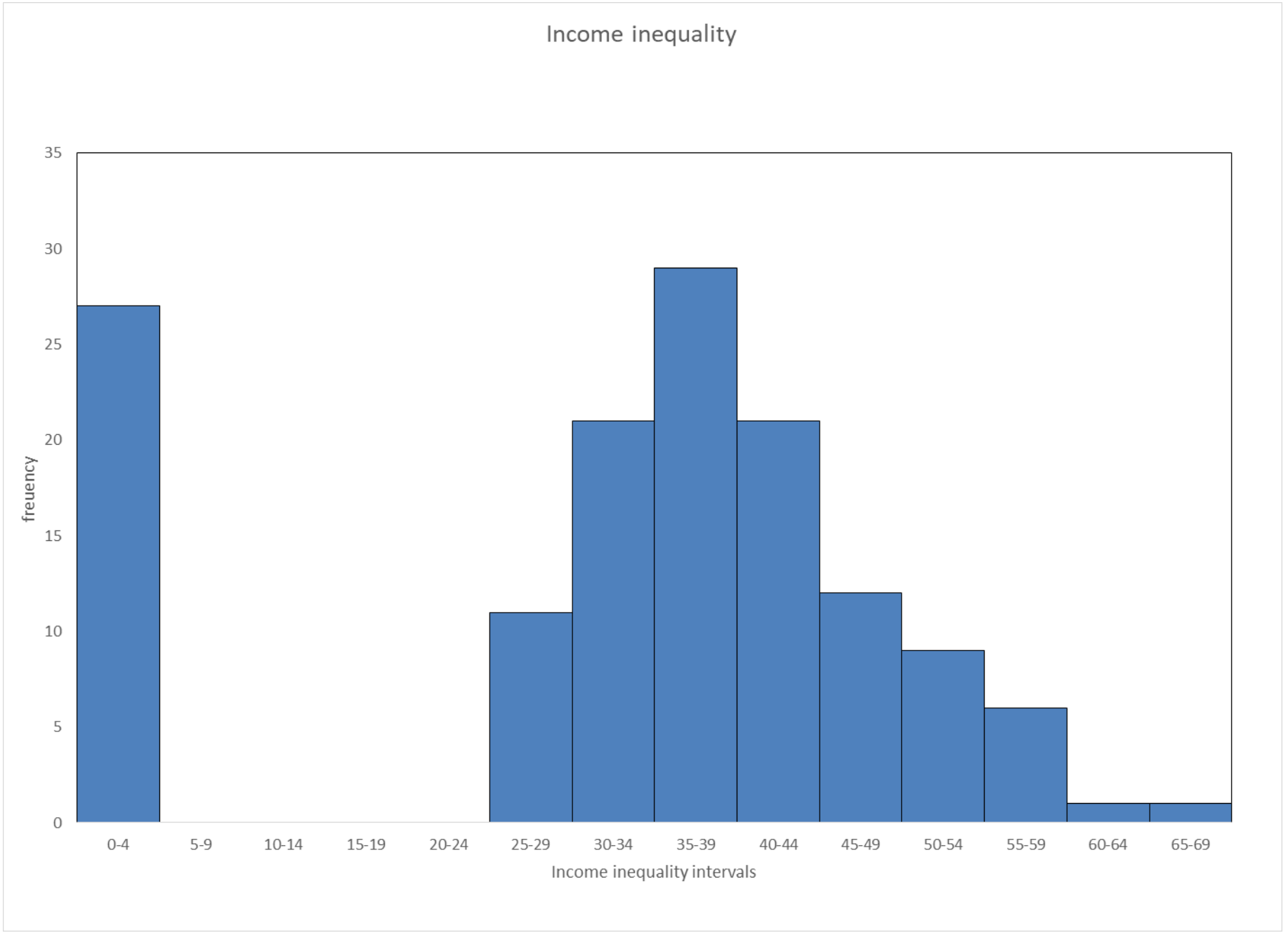
HISTOGRAM

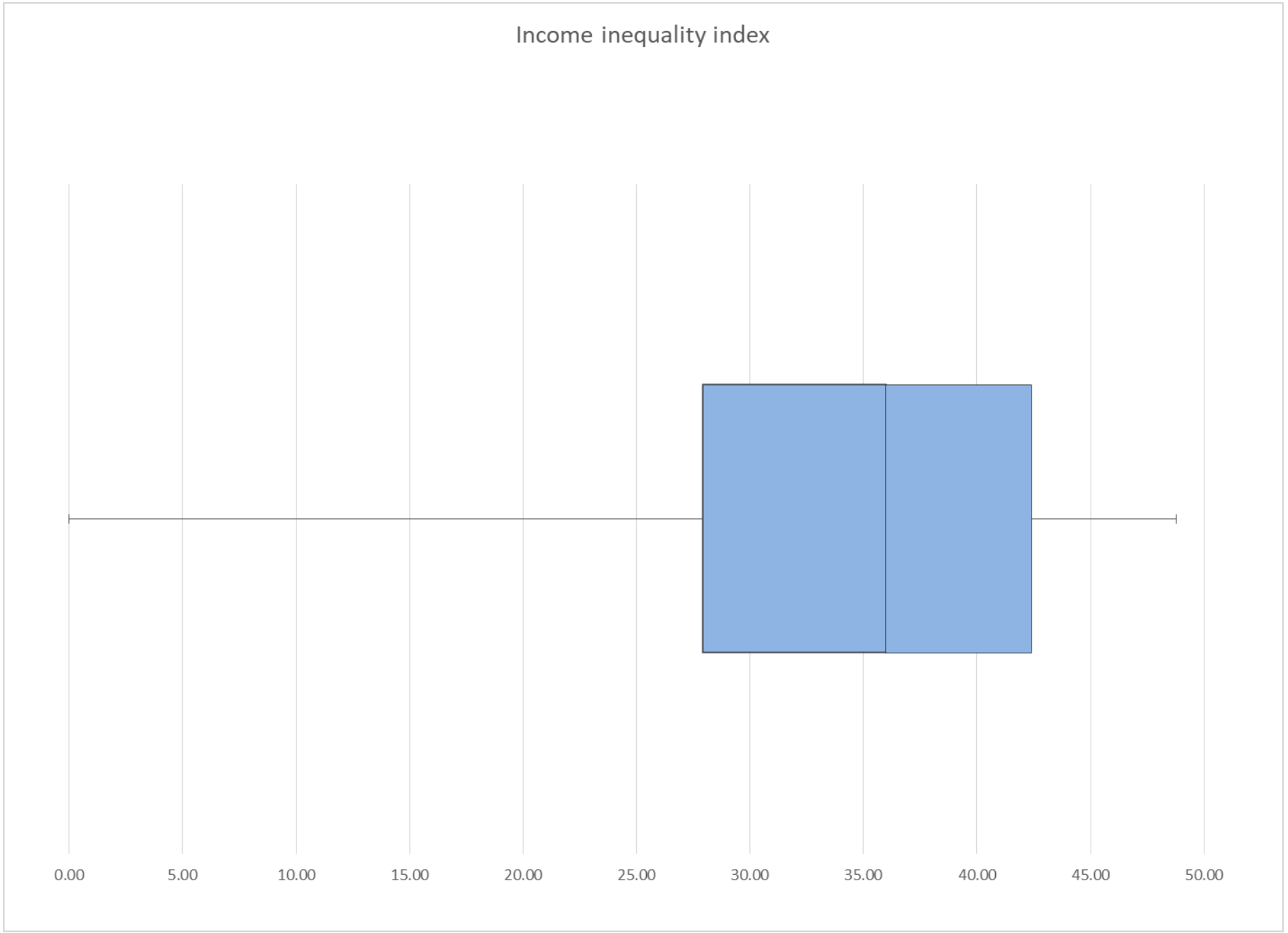
interval	Bin	Frequency
0-4	4	27
5-9	9	0
10-14	14	0
15-19	19	0
20-24	24	0
25-29	29	11
30-34	34	21
35-39	39	29
40-44	44	21
45-49	49	12
50-54	54	9
55-59	59	6
60-64	64	1
65-69	69	1

FIVE NUMBER SUMMARY AND BOX PLOT

Mean	31.73	
Median	35.99	
Mode	0	
Standard Deviation	17.4925644	Difference
Min	0.00	0.00
Q1	27.91	27.91
Median	35.99	8.07
Q3	42.385	6.40
Max	65.77	23.39

For this assignment, I used the income inequality index row. I decided to make an interval range and group my data based on the range and construct a histogram and box plot. This is the first table I included in this document. The first picture below is the histogram I made with the first table. I used the frequency formula and the data array and bin array for form the table on excel. In my opinion if not for the outliers in the data set (values with 0 and really low inequality index) the histogram would have been medium with most countries having medium equality index and fewer countries at the extreme end. However, due to the outliers, the histogram is skewed to the left (the median is greater than the mean). The second table is my five number summary. I calculated them with the excel function. I found the differences between the numbers and used it to create a box plot. I first created a right sided histogram, removed the colors from the boxes and then used error bars to show the whisker lines. I then thickened the outer edges and made finishing touches. The box plot matches my tables, and calculations.


Income inequality
35
30
25
20
15
10
5
0
freuency
0-4
5-9
10-14
15-19
20-24
25-29
30-34
35-39
40-44
45-49
50-54
55-59
60-64
65-69
Income inequality intervals


Income inequality index
0.00
5.00
10.00
15.00
20.00
25.00
30.00
35.00
40.00
45.00
50.00

Project 4 06/27/2020 corrected on 07/12/2020

Quantitative variable

H0: μ GNI= $25000

Ha: μ GNI ≠ $25000

GNI Average	13830

My hypothesis here is that the average amount of gross national income is $25000. To know if my hypothesis is true, I will calculate the mean of GNI of the countries to compare. The mean was $13830. This means that my hypothesis is wrong and the average is lower than $25000.

Categorical variable

H0: p (middle Income) = 0.33

Ha: p (middle income) ≠ 0.33

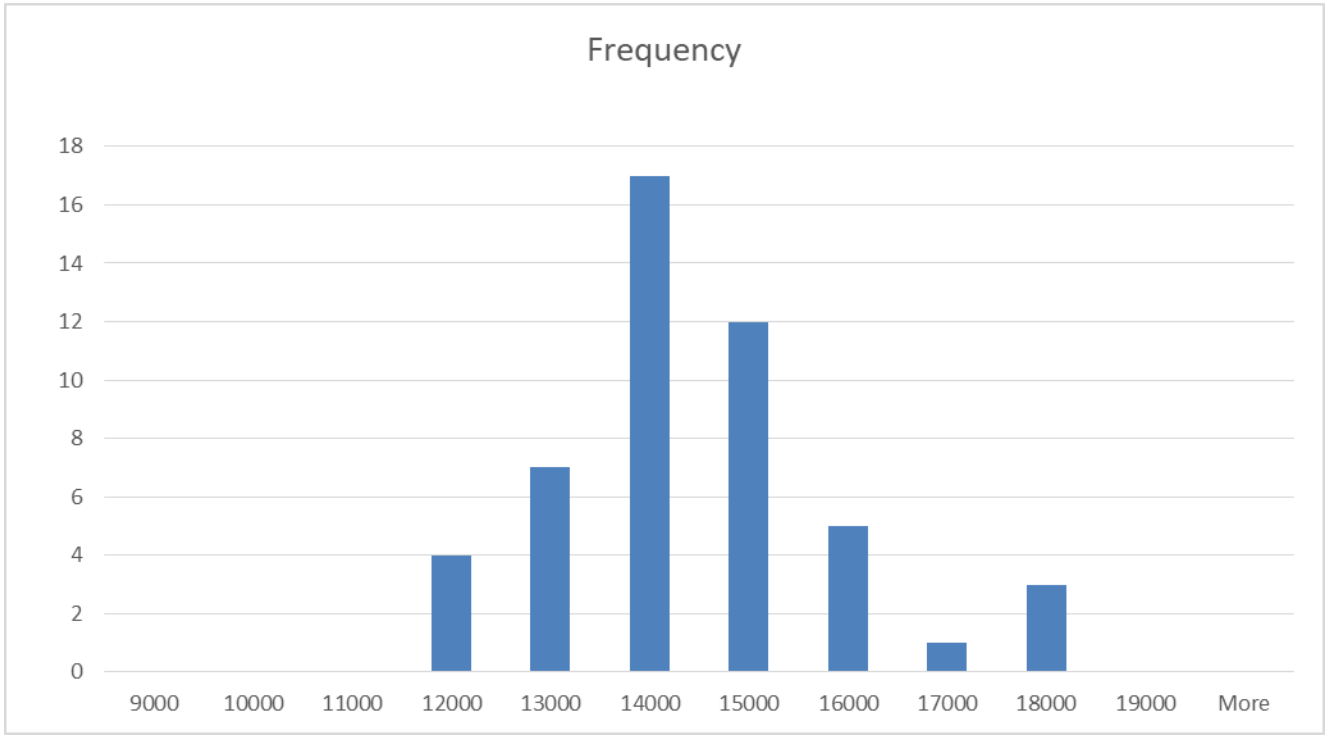
Income Level	Frequency
High	39
Middle	73
Low	20
Total	132

My hypothesis is that a third of countries in the world would have their middle income. I choose this as my hypothesis because I felt that about one-third of countries would have middle income and the remaining one-thirds would be that of high and low income. However my hypothesis is wrong because 73 out of 132 countries have middle income which gives 55% of the countries with middle income. This means that my hypothesis is wrong.

PROJECT 5 07/05/2020 (Histogram corrected 07/19/2020)

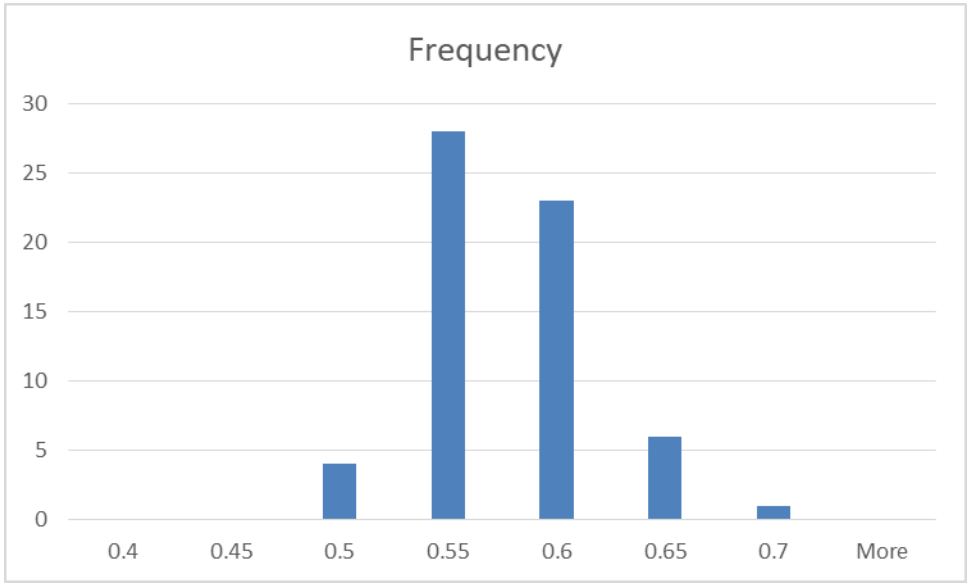
(a) Compute the standard error for the quantitative variable that you set up the hypothesis test for in part 4 using bootstrapping. Report a 95% confidence interval and decide if you are able to reject or fail to reject your null hypothesis. Create a histogram for your bootstrap distribution.

I found the standard error of the gross national income to be 2085.781812. The 95% confidence interval is 9703.649877 to 18046.77712. Since my null hypothesis was that the average GNI for countries was $25000, my null hypothesis would be rejected because it does not fall in the range of my 95% confidence interval. This means that the average GNI for the countries is not $25000.



(b) Compute the standard error for the categorical variable that you set up the hypothesis test for in part 4 using bootstrapping. Report a 95% confidence interval and decide if you are able to reject or fail to reject your null hypothesis. Create a histogram for your bootstrap distribution.

I found that the standard error of the middle income countries to be 0.043511142. The 95% confidence interval is 0.46354769 to 0.637592257. My null hypothesis was that 33% of countries would be middle income countries. My null hypothesis would be rejected because it does not fall in the range of the 95% confidence interval. This means that 33% of the countries are not middle income countries.



PROJECT 6 07/12/2020

(a) Repeat your hypothesis test on the categorical variable utilizing the appropriate formulas for your situation. Compute 95% confidence interval and compare to results from bootstrapping.

I tested my hypothesis by performing a z test by hand calculations. First I subtracted the sample proportion from the hypothesis (0.55-0.33). I then divided the answer by the square root of p * (p-1)/n i.e (0.33*0.67)/132. I then calculated the z score to be 5.38 That means the p(0.55) is 5.38 S.D away from Po (0.33). I found my p value to be 0.00001. I then compared it to 0.05 which was the alpha. I was able to reject Ho because the p value is below threshold. I also calculated the 95% C.I with the formula p +- (z *square root ((p*p-1)/n). Using 1.96 (since 95% CI), I found the CI to be 0.46512 to 0.63487. This value was not far from my bootstrap answers (0.46354769 to 0.637592257). They are almost the same if considering two decimal places. My manual calculations also point towards rejecting my Ho hypothesis because 0.33 does not fall in the 95% CI interval. This means that 33% of the countries are not middle income.

PROJECT 7 07/17/2020

I performed my calculations by calculating a T-test(one tailed). I found my sample size, n= 130. Sample mean x= 13829.84 My standard deviation SD= 19121.5573. The significance level is 0.05.

Using the formula **T test= (x – u)/ [SD/ square root(n)].** I got the value to be **T test= 6.6605.**

I then calculated the p value which gave me 0.00001. This made it significant at p< 0.05. Since p value is less than significance level, I reject my null hypothesis. Also to calculate the 95% C1

Using the **formula 95 % CI = x +- Zc*[SD/(square root n)]. (Zc = 1.96)** I got my **95% CI** to be **10542.78 to 17116.91.**

My null hypothesis (H0: μ GNI= $25000) is wrong since it does not fall in the 95% CI range. In conclusion the average gross national income of 130 countries is NOT 25000. In comparison of my bootstrap answers, 95% CI= 9703.649877 to 18046.77712. My formula calculations not very far. I trust both methods since they both disproved my hypothesis.

PROJECT 8 07/25/2020

2. Use the two-way table from part 2 to complete the following:

(a) Create at least 2 conditional probabilities from your two-way table. Interpret their meanings and explain how they were computed. Include the following formula for conditional probability

If A and B are two events such that given event B has happened and we have to find the probability of event A, then probability of A given B, denoted by P(A/B) is given by $P(A/B) = P(A \cap B)/P(B)$

A conditional probability would look at these two events in relationship with one another.

National Plan	**Yes Y**	**No N**	**Total**
Interpersonal Violence IV	68	54	122
Sexual Violence SV	87	33	120
Total	155	87	242

1) What is the probability that a country has a national plan i.e. a yes?
   [total countries with a plan (yes)] / [total number of countries] = 155/ 242 = 0.64 This means that 64% of countries are likely to have a national plan for interpersonal violence and sexual violence.
2) What is the probability that a country with both IV and SV violence has a sexual violence plan i.e. yes for SV **P(SV/Y)?**
   [SV's with yes]/ [Total yes's] = 87/155 = 0.56 This means that 56% of countries with a national plan for both crimes are likely to have a plan in place for sexual violence.
3) What is the probability that there will be no national plans for interpersonal violence i.e. **P(N/ IV) ?**
   [IV's with no]/ [Total IV's] = 54/122 = 0.44. This means that 56% of countries that consider interpersonal violence are less likely to have a national plan for interpersonal violence

This table, as well as probabilities are a little different because not all countries have a plan in place or are bothered with the area of crime.