MG-220 Final Project

FROM: Larry Agyei
TO: Dr. Thomas Craig
DATE: April 25, 2022
SUBJECT: Analysis of the Gallup World Poll Happiness Report

Section 1: Introduction/Overview

There have been so many crazy storylines in the world recently. There always seems to be something new happening every day that we wake up, and stories arriving by the minute. Most of these things can be good, other times not so much. With all that is currently happening in the world right now, it is a wonder to think about how everyone around the globe currently feels at this moment in time. This trail of thought led me to question how many people on earth are happy? Can you measure happiness? Which group of people/which country is currently the happiest?

Section 2: Statement of Hypothesis

There are many factors that one can put into consideration when determining how happy they are. These include factors such as GDP per capita, life expectancy, perceptions of corruption, and many more. Factors that I wanted to dive into out of interest were perceptions of corruption and GDP per capita, as I feel I could make solid predictions about the relationship between these factors and people's happiness.

Hypotheses:

I began my analysis with a prediction of the average world happiness score. I chose the value 4.8.

Null Hypothesis (H_0): The average ladder score for all countries in this dataset is 4.8.
Alternative Hypothesis (H_1): The average ladder score for all countries in this dataset is not 4.8.

My second hypothesis is a prediction that there is no difference in the average GDP per capita for each region from the data we collected. As I will explain later, there are 148 countries

in the dataset for world happiness, and they are split by regions. I wanted to see if there would be a difference or not in the average GDP per capita in each region. My initial prediction is no.

Null Hypothesis (H_0): There is no difference in the average GDP per capita for each region in the dataset.
Alternative Hypothesis (H_1): There is a difference in the average GDP per capita for each region in the dataset.

My third and final hypothesis relates to corruption and its relation to people's happiness scores. I believe that there will be a negative correlation between perceptions of happiness and happiness scores. A greater perception of corruption will give us a lower happiness score, while a lower perception of corruption will give us a higher one.

Null Hypothesis (H_0): There is a negative correlation between perceptions of corruption in a country and their ladder score of happiness.
Alternative Hypothesis (H_1): There is not a negative correlation between perceptions of corruption in a country and their ladder score of happiness.

Section 3: Data and Analysis Methodology

The data that will be used for this analysis is gathered from the 2021 Gallup World Poll World Happiness Index. There is a total of 148 records in this dataset, representing 148 countries. There is also 20 variables in the dataset that relate to various factors that were used to determine happiness levels. Happiness is measured based on a ladder score that is determined by 6 factors. The factors that contribute to a ladder score are GDP per capita, social support, healthy life expectancy, freedom to make life choices, generosity, and perceptions of corruption. Also included are the geographic regions of each country and the standard error of the ladder score for each country.

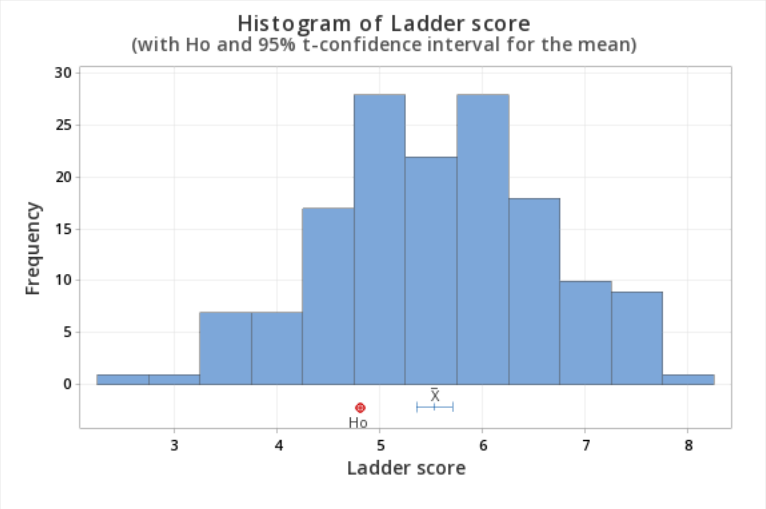
To test the hypotheses above, I will be using a 1-sample t-test, a One-Way ANOVA test, and scatterplot (with a regression fit line) to determine if my hypotheses are true. The 1-sample t-test will be used to determine whether we can reject my claim that the average happiness ladder score is 4.8. The One-Way ANOVA test will be able to determine if there is a difference in

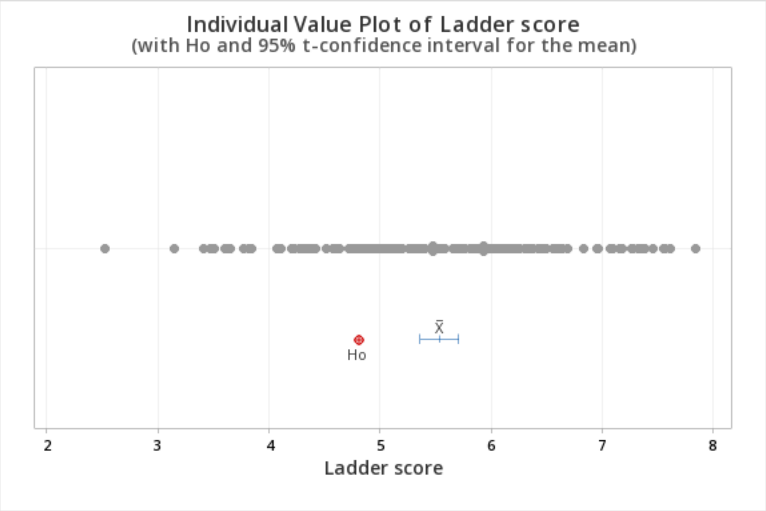
average GDP per capita for each region, and if there are any clusters of data between regions. Lastly, the scatterplot and regression line will be able to provide us information on the relationship between perceptions of corruption and the happiness ladder scores, and the percentage chance that these two variables influence each other. All my work has been done through Minitab (to perform the tests) and Excel (data was formatted onto worksheet).

Section 4: Results

1-Sample t-test: Average Ladder Score for Happiness

The first hypothesis that I tested was the average ladder score for happiness to be equal to 4.8. From the results of the 1-sample t-test, we can confidently predict that there is no chance that there is an average ladder score of 4.8 from the World Happiness Index. There is not a significant amount of statistical data that proves that the average ladder score is the hypothesized one, therefore we can reject that hypothesis claim.





With a p-value of 0.000, which is less than 0.05, there is a 0% chance that the hypothesis is true, as we need to have a p-value higher than 0.05 to confidently suggest that hypothesis is incorrect.

Test

Null hypothesis	$H_0: \mu = 4.8$
Alternative hypothesis	$H_1: \mu \neq 4.8$

T-Value	P-Value
8.33	0.000

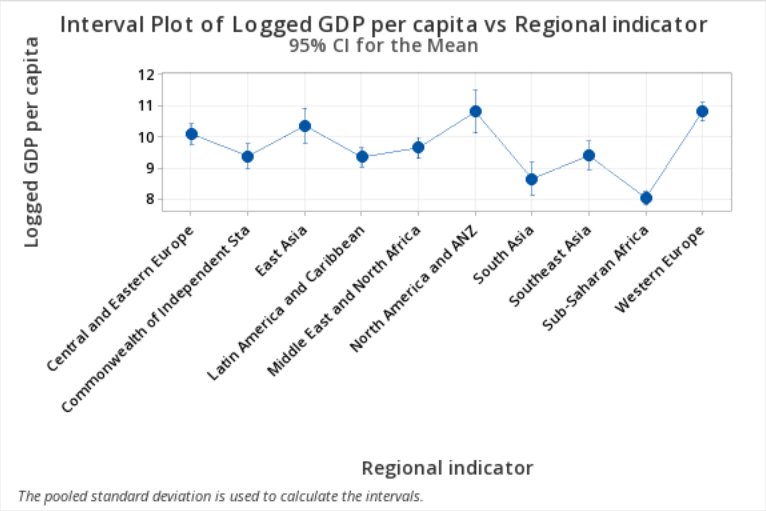
One-Way ANOVA: GDP per Capita Comparison

When testing for my second hypothesis: no difference in the average GDP per capita for each region in the dataset, it was revealed that we have the data to suggest that this claim is not true. There is not a significant amount of statistical data that proves that there isn't a difference in average GDP per capita of each region in the dataset. Because of this we can confidently reject the hypothesis claim.

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regional indicator	9	132.48	14.7197	30.91	0.000
Error	139	66.19	0.4762		
Total	148	198.67			

A p-value of 0.000 that is less than 0.05 gives us 95% confidence to suggest that the hypothesis was incorrect.



Grouping Information Using the Tukey Method and 95% Confidence

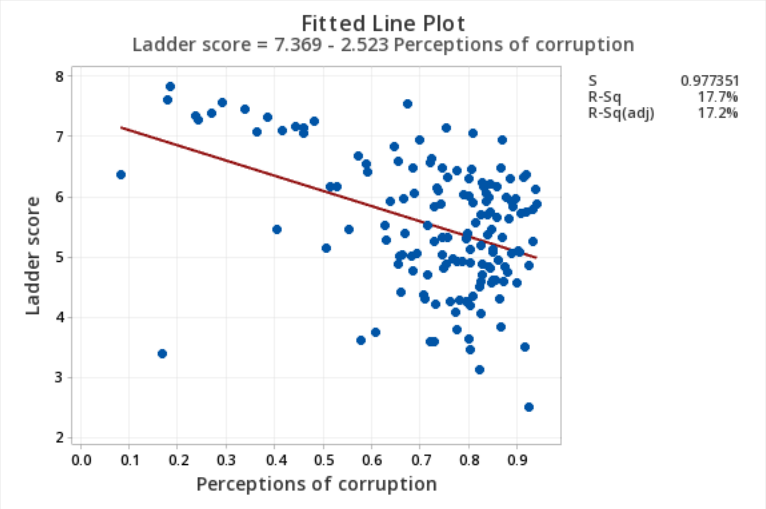
Regional indicator	N	Mean	Grouping				
Western Europe	21	10.8227	A				
North America and ANZ	4	10.8095	A	B			
East Asia	6	10.368	A	B	C	D	
Central and Eastern Europe	17	10.1091	A	B	C		
Middle East and North Africa	17	9.666		B	C	D	E
Southeast Asia	9	9.421			C	D	E
Commonwealth of Independent States	12	9.402			C	D	E
Latin America and Caribbean	20	9.370				D	E
South Asia	7	8.683					E F
Sub-Saharan Africa	36	8.075					F

Means that do not share a letter are significantly different.

The interval plot and grouping information from the One-Way Tukey ANOVA test gives us interesting information regarding the average GDP per capita in the different regions in the world. The 10 regions in the dataset were grouped into 6 different clusters that are statistically similar regarding GDP per capita. This visual representation can also help us in realizing that the average GDP per capita is not the same in all the regions of the dataset.

Fitted Line Plot: Perception of Corruption vs Happiness Ladder Score

Lastly, I analyzed the effect that perception of corruption in a country has on its ladder score for happiness. The hypothesis is that there is a negative correlation between perception of corruption in a country and their ladder score of happiness. The following results show that this hypothesis is correct. The regression line for this fitted line plot is a downward sloping one, meaning that there is a negative correlation between the x-value (perceptions of corruption) and the y-value (ladder score). Further evidence of this is seen in the equation for the regression line (Ladder score = 7.369 – 2.523 x Perceptions of corruption). The subtraction sign in the equation shows that there is a negative relationship between the two variables, and that there will be a downward slope. Our data has shown us that the higher the perception of corruption is in a country, the lower its ladder score will be.



Section 5: Discussion and Conclusion

The World Happiness Index has tons of information that we can churn out and find conclusions for. An attempt to find out how happy the world is ended up becoming an exploration of many factors that contribute to happiness. I became very invested in the topics of GDP per capita and perceptions of corruption in the many countries found in the data. To make the conclusions that I found in my data possible, I used a 1-sample t-test, a One-Way ANOVA test, as well as a fitted line plot. Out of the three hypotheses, only one of them ended up being correct. I was still able to learn a lot from the data even after not getting all my hypotheses correct. The 1-sample t-test revealed that my prediction of an average ladder score being 4.8 was

incorrect. In fact, it was higher than that. I can go to sleep knowing that the world on average is happier than I thought it would be. The One-Way ANOVA test showed me that there is a difference between average GDP per capita in the different regions of the world the dataset provided. I learned that Western/Central and Eastern Europe, North America, and East Asia all have similar, as well as the highest average GDP per capita while South Asia and Sub-Saharan Africa share similar average GDPs. The latter two regions have the smallest average GDP per capita. These regions share the same Tukey clusters, which is why they have similar GDPs. Lastly, the fitted line plot gave us proof of our only correct hypothesis; that there is a negative correlation between perceptions of corruption and a country's happiness ladder score. A negative slope in the regression line equation and a downward slope shows that the negative correlation hypothesis was correct. It was also interesting to see the r-squared value for these two variables. The r-squared value is 17.7%, which means there is a 17.7 percent variation in the ladder score that is caused by variation in the perception of corruption variable. Overall, my analysis of the World Happiness Index was fulfilling in that I was able to draw many conclusions from both the ladder scores and the multiple factors that contribute into making them.

Works Cited

Singh, A. (2021, March 22). *World happiness report 2021*. Kaggle. Retrieved April 28, 2022, from https://www.kaggle.com/datasets/ajaypalsinghlo/world-happiness-report-2021?resource=download&select=world-happiness-report-2021.csv

World happiness report 2021. World Happiness Report 2021. (n.d.). Retrieved April 28, 2022, from https://worldhappiness.report/ed/2021/