Final Project: What Makes a Country Happy? Comparing Western Europe with North America

Anna MacFarlane, Edgar Salas, Jerry Fu, Taalin RaoShah

11/20/2020

INTRODUCTION

CodeStone chose to explore the data from the World Happiness Reports (WHR) of 2018 and 2019, published by the Sustainable Development Network. According to the WHR, this data is derived from the Gallup World Poll, systematic telephone surveys and in-person interviews in over 150 countries whose surveys claim to represent 80%+ of the population [1]. The calls are made via random phone number generation and randomly selecting households [2]. The Gallup World Poll changes the phrasing and types of questions they ask to determine a countries' happiness every few years, but the procedure they used remained constant between 2018 and 2019. In order to get a better chance of obtaining accurate results, we chose to include both 2018 and 2019 in the same dataset, so that every country has two observations. The data set in total contains 312 rows, each pair representing the data for a particular country from both 2018 and 2019. The original variables in the data set included the overall rank according to happiness score, country name, happiness score, and then 6 variables that were used to calculate the happiness score in the WHR's analysis. These are Gross Domestic Product (GDP) per Capita , Social Support, Healthy Life Expectancy, Freedom, Generosity, and Perceptions of Corruption, all of which are expanded upon below. Additionally, we added variables for year as well as region to further our analysis.

Generally, we want to explore the following question: what factors contribute to the differences in happiness scores between countries in North America versus Western Europe? Both regions are relatively comparable in terms of economic and political structures, but Western Europe contains the Nordic countries which consistently score highest in the World Happiness reports. Furthermore, there are notable lifestyle differences between North America and Western Europe. Accordingly, we are curious to analyze differences between the specific variables that make up the regions' mean happiness scores.

The countries included in North America for the purposes of this study are the United States, Canada, and Mexico. We altered the original categorizations of regions, in which Mexico was not included but Australia and New Zealand were, to align with common perceptions of the region of North America. Western Europe includes a group of 20 countries, which was not changed from the original categorization.

Considering the concerning state of our world in 2020, including the worsening effects of climate change, threats to democracy, and much more, we found it topical and insightful to evaluate what contributes to happiness within nations and between regions across the globe. The first World Happiness Report, published in 2012, presents the report as a means of grappling with the countless contradictions that exist in modern society such as the balance between pursuing economic success versus protecting the environment or the tradeoffs between personal profit and community trust [5]. Eight years later, these paradoxes persist, and the potential solutions are closely linked to definitions of morality, heightening their controversy. Considering the continued debate over such questions, we believe there are grounds for further investigation into trends of happiness over time and the factors that contribute to it.

Variable Sources and Definitions

Happiness Score: Happiness score is a self-reported measure of overall current life satisfaction. This was measured by asking respondents, "Please imagine a ladder, with steps numbered from 0 at the bottom to 10 at the top. The top of the ladder represents the best possible life for you and the bottom of the ladder represents the worst possible life for you. On which step of the ladder would you say you personally feel you stand at this time?" The average of these values represents respective countries and regions [3][4].

GDP per capita: PPP (purchasing power parity) is a rate of conversion which attempts to equalize the purchasing power across all different currencies. The World Happiness Report sources its GDP per capita in PPP values from the 2017 and 2018 World Development Indicators respectively. This value is logged. [3][4].

Social Support: Social support is a self-reported measure of whether or not the respondent feels they can be helped. Specifically, respondents were asked, "If you were in trouble, do you have relatives or friends you can count on to help you whenever you need them, or not?" and responded with 0 or 1 (no or yes). The average for each respective country or region creates this value [3][4].

Life Expectancy: Life expectancy data was extrapolated from the WHOs health observation data up to 2016. Where missing, life expectancy data for certain countries was found using research and government tools [3][4].

Perceptions of Corruption: Perception of Corruption is a self-reported measure of whether or not respondents feel there is active corruption within government and business. Specifically, respondents were asked, "Is corruption widespread throughout the government or not" and "Is corruption widespread within businesses or not?" and responded to both with 0 or 1 (no or yes). The average for each respective country or region creates this value [3][4]. Note: A higher value for this measure represents a lower perception of corruption.

Generosity: Generosity is a measure of whether or not respondents act generously. Respondents were asked, "Have you donated money to a charity in the past month?" and responded with 0 or 1 (no or yes). Generosity is the residual of regressing the national average of this question with GDP [3][4].

Freedom: Freedom is a self-reported measure of whether or not respondents feel they can do what they want with their lives. Respondents were asked, "Are you satisfied or dissatisfied with your freedom to choose what you do with your life?" and responded with 0 or 1 (no or yes). The average for each respective country or region creates this value [3][4].

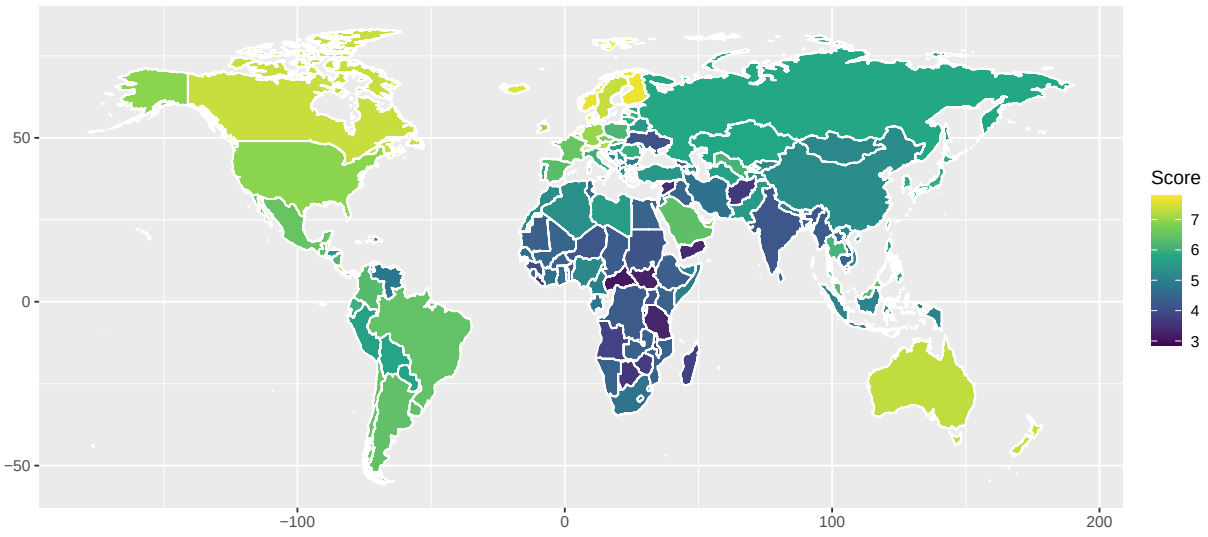
METHODOLOGY

First, we underwent data cleaning duties and loaded in the data sets, merged 2018 and 2019, and added in regions for countries which were missing a region.

World Happiness Scores Visualization

To initially visualize the spread of happiness scores across the regions of interest, we first created a choropleth map of the world, shaded to show differences in average happiness score across each country.

Figure 1: World Map
colored by Happiness Score



Chi-square test

Next, we looked at a general overview of the distribution of Happiness Scores across different countries and regions. In a map of the world colored by Happiness Score, we noticed that countries in different regions of the world tended to have different happiness scores. In particular, Western Europe seemed to have the happiest countries, followed by North America and Central and South America. Eastern Europe, East Asia, South Asia, North Africa, sub-Saharan Africa and Southeast Asia contained some of the least happy countries in the world. A country's happiness score seemed to depend at least partially on its region. In order to see if our conclusions from the map were statistically significant, a chi-squared test was used to test if there was an association between Happiness Score and Region. However, a chi-squared test is applied to two categorical variables. Therefore, we mutated a new variable called Happiness, which gave a country a category of 1, 2, 3, or 4 depending on which quantile the country fell into, 4 being the highest.

A chi-squared test statistic is computed by adding up the squares of the difference between the observed and expected values in a table, divided by the expected values. The expected values are equal to the row total multiplied by the column total, and divided by the overall total. Finally, p-values are derived from degrees of freedom and the chi-square test statistic. Alternatively, a simulated p-value can be created if test conditions are not satisfied. A very small chi-square test statistic means that your observed data fits your expected data well and that a relationship exists, while a large chi-square test statistic means the opposite. In our case, our null hypothesis was that there isn't an association between region and happiness, and our alternative hypothesis was that there was an association between region and happiness. We opted to use a simulated p-value. Keeping the conclusion of the chi-squared test in mind, we moved on to looking at the individual metrics which went into determining happiness score.

We conduct the chi-square test at the $\alpha = 0.05$ level, using the hypotheses below:

H_0: There is no association between region of the world and happiness score.

H_a: There is an association between region of the world and happiness score.

```
##
##  Pearson's Chi-squared test with simulated p-value (based on 2000
##  replicates)
##
```

Table 1: Figure 2: Mean Scores for World Happiness Report Categories

	Happiness Score	GDP	Social Support	Healthy Life Expectancy	Generosity	Corruption	Freedom
Western Europe	6.863750	1.343850	1.486825	0.954825	0.22245	0.2176750	0.52105
North America	6.911167	1.272333	1.423333	0.875000	0.22000	0.1713333	0.52500

```
## data:  chisq$happiness and chisq$Region
## X-squared = 273.46, df = NA, p-value = 0.0004998
```

Summary Statistics

Next, we obtained the following summary statistics for both of the regions of interest, North America and Western Europe. These values will be used in further testing.

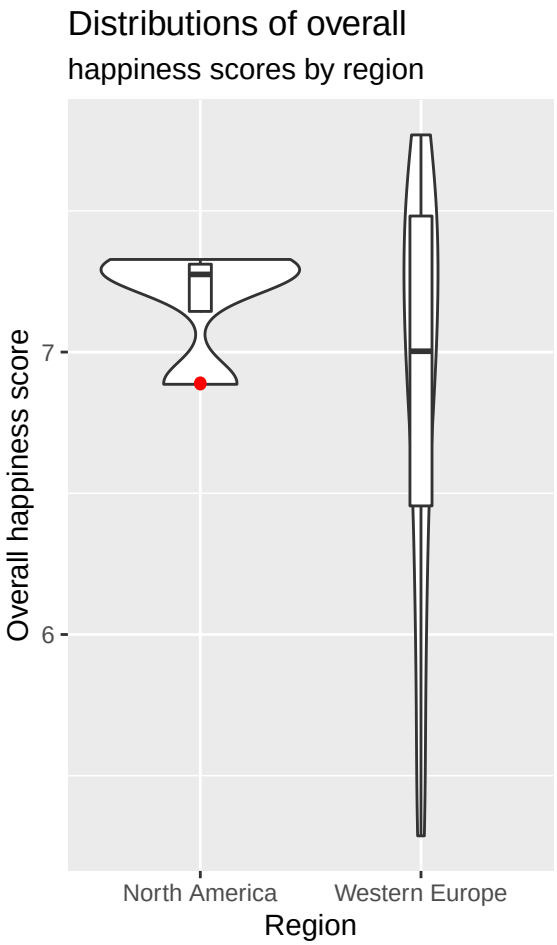
Function

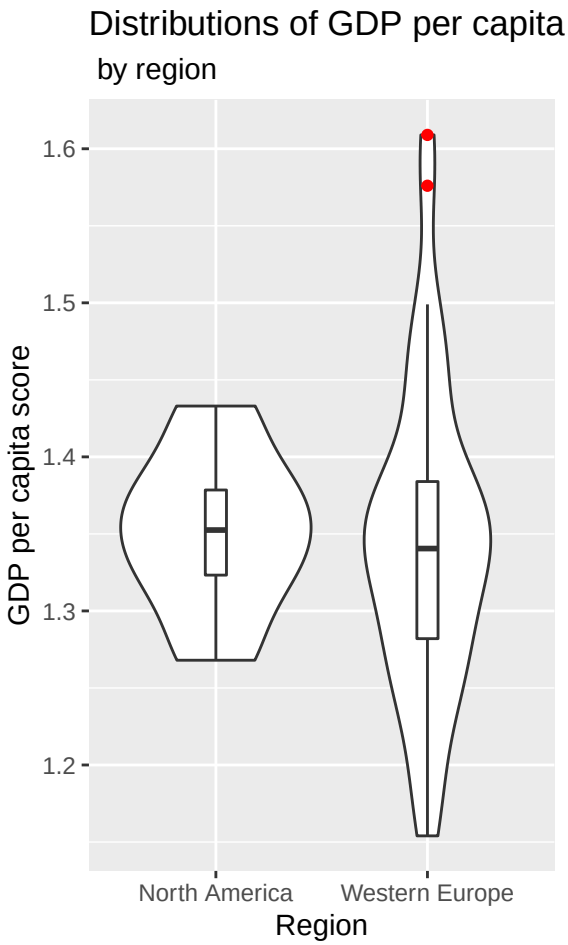
We chose simulation-based testing to calculate the p-value for difference in means of a variable between Western Europe and North America. We wrote a function with the arguments of variable and n_sims, and applied this function to each of the 6 variables which determine Happiness Score. First, we sampled observations from both the Western Europe and North America rows in the dataset, with replacement, and subtracted the means of these two regions, repeating this 1000 times so we would be less likely to commit type I and type II errors. Next, we shifted these differences and converted them to a null distribution centered around 0. After storing the observed difference between the two regions in the original dataset, we found the proportion of these bootstrapped differences which were greater than the observed difference, assuming the null hypothesis was true.

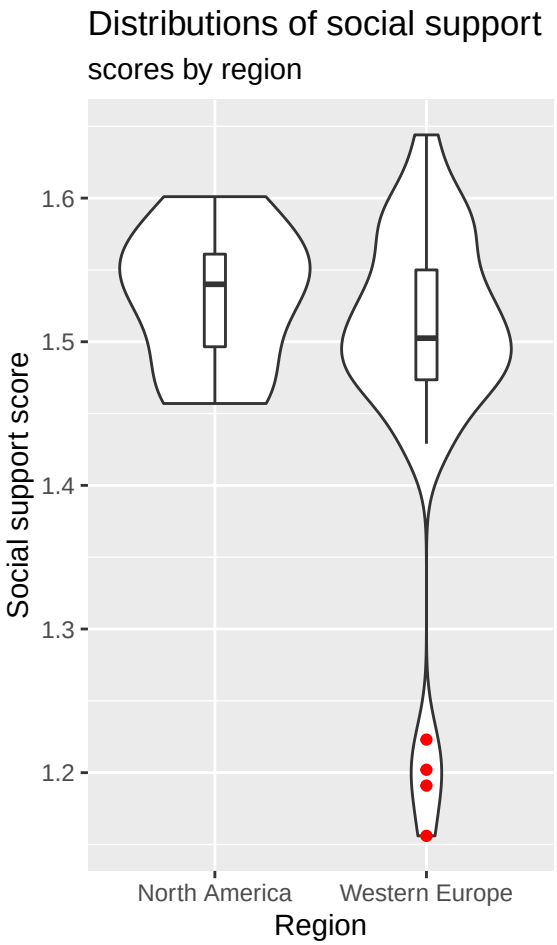
Violin Plot Visualizations

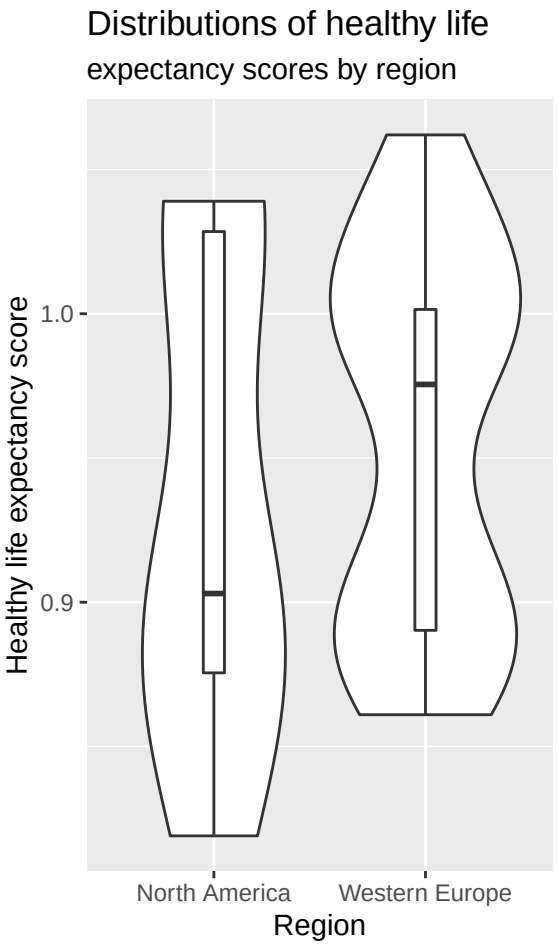
Next, we construct violin plots to further visualize the differences in distribution of respective happiness scores, and sects that might contribute to a country's overall happiness. Here, we plot mean happiness scores, GDP per capita scores, Social Support score, Life Expectancy score, Freedom to make Life Choices score, Generosity score, and Perceptions of Corruption score. Western Europe had a larger spread for all the variables, and slight top heavy or bottom heavy skews for some of the variables such as Freedom to make Life Choices and Social Support.

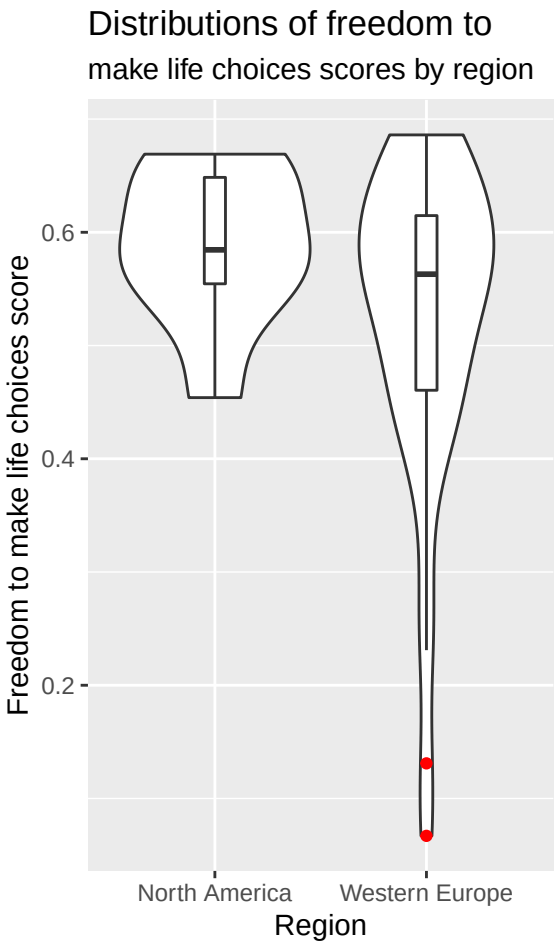
Figure 3: Violin Plots for Happiness Score Distribution

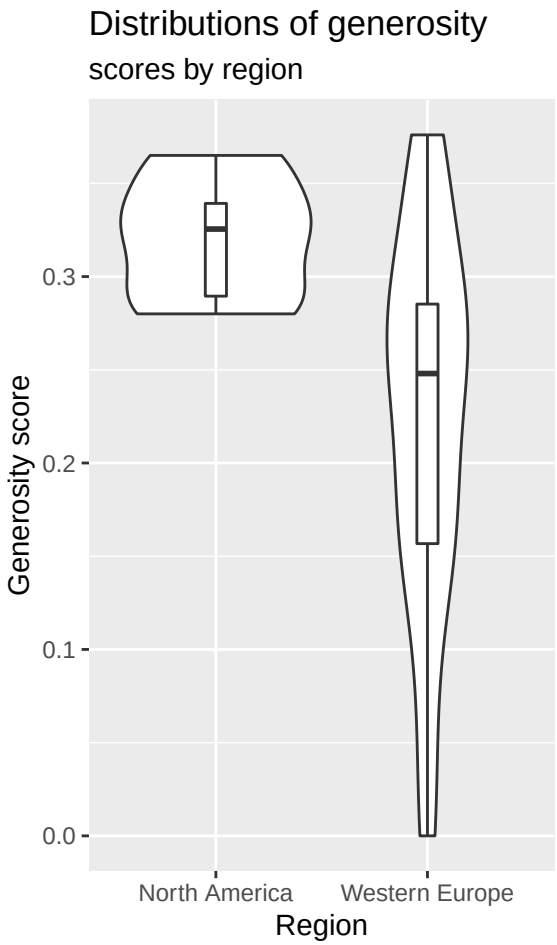

Distributions of overall
happiness scores by region
Overall happiness score
7
6
North America
Western Europe
Region

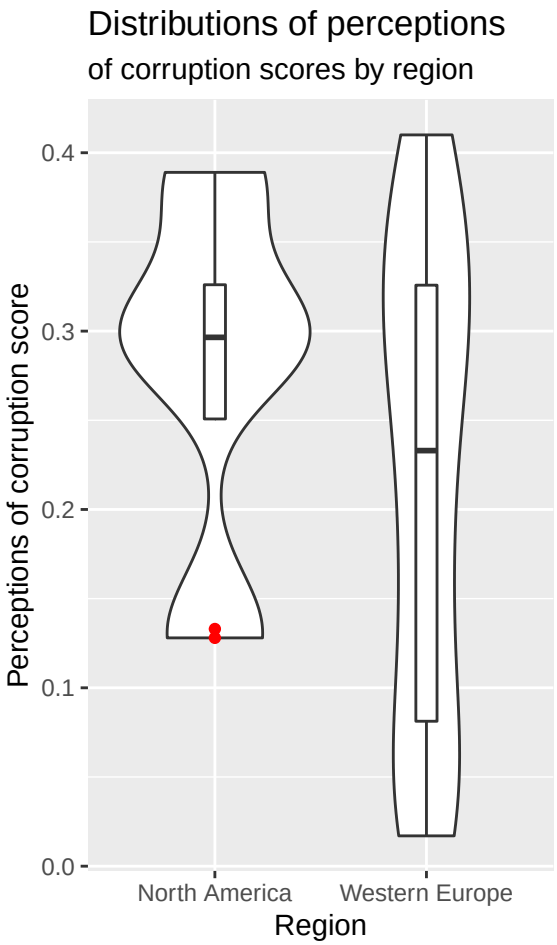

Distributions of GDP per capita
by region
GDP per capita score
1.6
1.5
1.4
1.3
1.2
North America
Western Europe
Region


Distributions of social support
scores by region
Social support score
1.6
1.5
1.4
1.3
1.2
North America
Western Europe
Region


Distributions of healthy life
expectancy scores by region
Healthy life expectancy score
1.0
0.9
North America
Western Europe
Region


Distributions of freedom to
make life choices scores by region
Freedom to make life choices score
0.6
0.4
0.2
North America
Western Europe
Region


Distributions of generosity
scores by region
Generosity score
0.3
0.2
0.1
0.0
North America
Western Europe
Region



Simulation based hypothesis testing

Next, we conducted simulation-based hypothesis tests for each of the variables contributing to overall happiness score to evaluate whether there is a statistically significant difference the regions of Western Europe and North America.

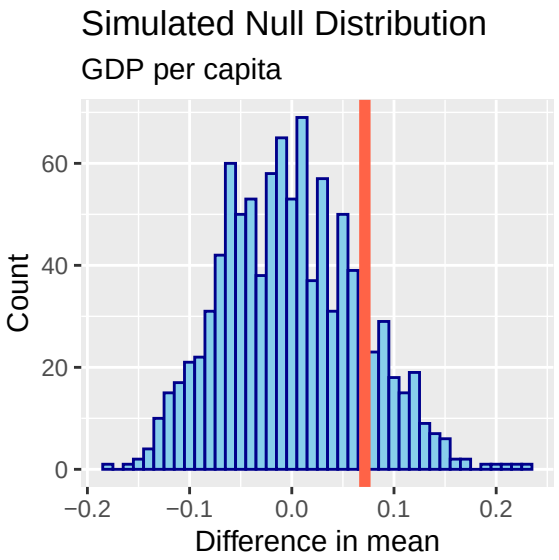
Each of the simulation-based hypothesis tests used to evaluate the variables contributing to overall happiness score will be evaluated at the $\alpha = 0.05$ level.

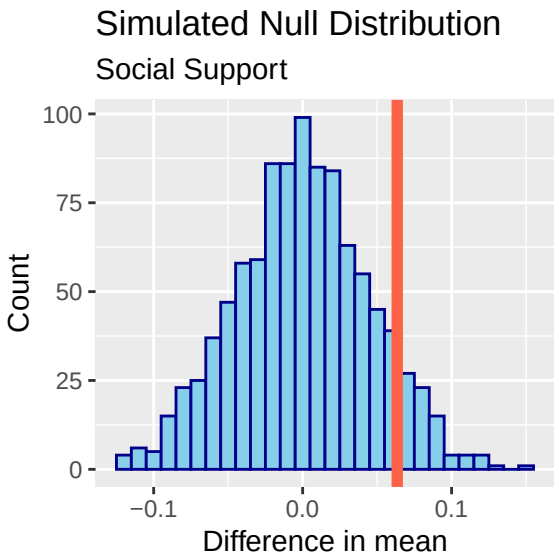
Generally, our null hypothesis states that the mean score for a particular variable in Western Europe equal to the mean score for that same variable in North America. Our alternative hypothesis therefore is that the mean score for a particular variable in Western Europe is not equal to that of the mean score for that variable in North America.

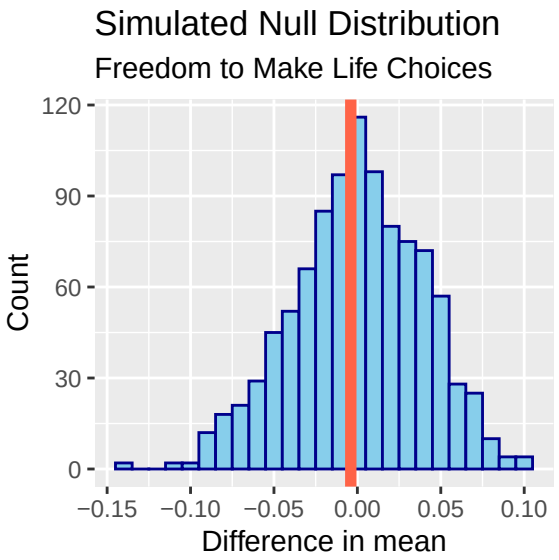
H_0: $\mu_{WE} = \mu_{NA}$

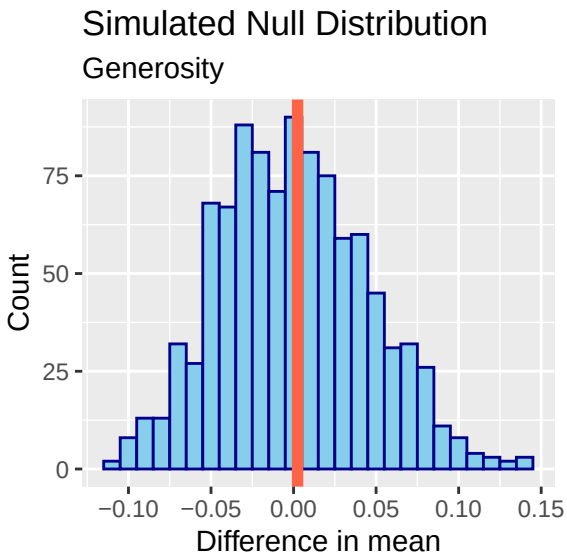
H_A: μ_{WE} != μ_{NA}

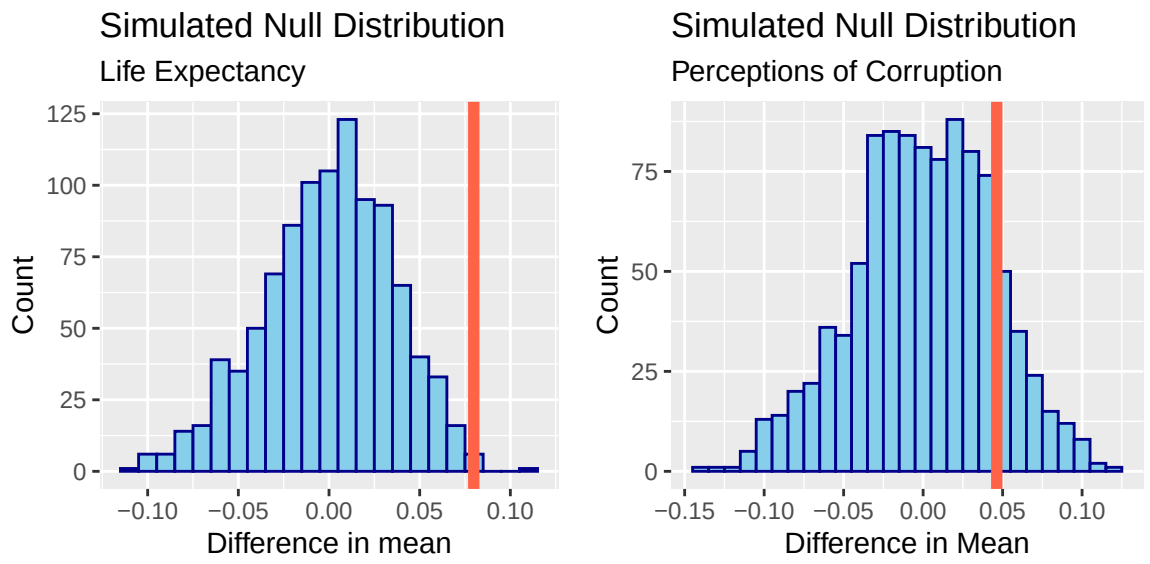
Figure 4: Simulated Null Distributions of Difference in Means


Simulated Null Distribution
GDP per capita
Count
Difference in mean
0
20
40
60
−0.2
−0.1
0.0
0.1
0.2


Simulated Null Distribution
Social Support
Count
Difference in mean
0
25
50
75
100
−0.1
0.0
0.1


Simulated Null Distribution
Freedom to Make Life Choices
Count
Difference in mean
0
30
60
90
120
−0.15
−0.10
−0.05
0.00
0.05
0.10


Simulated Null Distribution
Generosity
Count
Difference in mean
0
25
50
75
−0.10
−0.05
0.00
0.05
0.10
0.15



Based on our bootstrapped distributions, we calculated these p-values.

Table 2: Figure 5: P-values for Hypothesis Tests Difference in Means

GDP	Social Support	Freedom	Generosity	Healthy Life Expectancy	Corruption
0.145	0.083	0.556	0.468	0.007	0.142

RESULTS

Chi-square test

With our simulated p-value of 0.0004998 which is less than our $\alpha = 0.05$, we do have sufficient evidence to reject the null hypothesis in favor of the alternative hypothesis and conclude that there is an association between which region of the world a country is in and happiness score. Our chi-squared test statistic is 273.46 with NA degrees of freedom. The pattern we noticed in the map is statistically significant.

Simulation-based hypothesis tests

GDP per capita:

Since the calculated p-value of 0.145 is not significant at the $\alpha = 0.05$ level, we fail to reject the null hypothesis. Therefore, there is not statistically significant evidence to suggest that the mean GDP per capita score in Western Europe is different than that of North America.

Social support:

The calculated p-value of 0.083 is not significant at the $\alpha = 0.05$ level, so we fail to reject the null hypothesis. Therefore, there is not statistically significant evidence to suggest that the mean social support score in Western Europe is different than that of North America.

Freedom to make life choices

Since the calculated p-value of 0.556 is not significant at the $\alpha = 0.05$ level, we fail to reject the null hypothesis. Therefore, there is not statistically significant evidence to suggest that the mean freedom to make life choices score in Western Europe is different than that of North America.

Generosity

The calculated p-value of 0.468 is not significant at the $\alpha = 0.05$ level, so we fail to reject the null hypothesis. Therefore, there is not statistically significant evidence to suggest that the mean generosity score in Western Europe is different than that of North America.

Life expectancy

The calculated p-value of 0.007 is significant at the $\alpha = 0.05$ level, so we can reject the null hypothesis. Therefore, there is statistically significant evidence to suggest that the mean life expectancy score in Western Europe is different than that of North America.

Perceptions of corruption

Since the calculated p-value of 0.142 is not significant at the $\alpha = 0.05$ level, we fail to reject the null hypothesis. Therefore, there is not statistically significant evidence to suggest that the mean perceptions of corruption score in Western Europe is different than that of North America.

DISCUSSION

Based on the data from the World Happiness Reports of 2018 and 2019, we aimed to explore the differences in the particular variables contributing to happiness score. Of the 6 variables evaluated, statistically significant results were found at the $\alpha = 0.05$ level only for life expectancy, supporting the claim that the mean life expectancy in Western Europe is not equal to that of North America. For all other variables, though, we failed to reject the null hypothesis, indicating that the mean scores for these variables in Western Europe were not statistically different than those of North America.

Initially we believed that there would be a difference in happiness for a number of reasons. First, Western Europe includes the Nordic countries, societies in which capitalist elements are combined with more socialist concepts such as a strong social safety net [5]. While not all of Western Europe follows the Nordic Model, it stands out as a notable ideological difference between the regions of the world being considered. We also believed that the United States' emphasis on individualism over collectivism would lead to differences in the main variables. Yet, for almost every variable, our results did not support the claim that the mean scores in Western Europe were statistically significantly different than those of North America.

Also, in our initial thoughts, we failed to take into account that Canada has very high scores across the board, bringing the North America average up by a lot. Furthermore, a decent amount of countries in Western Europe, like France, have medium levels of mean happiness scores, bringing the Western European average down. Interestingly, the mean scores for most variables ended up being quite similar between the two regions, revealing a level of comparability. Part of the reason for this might be because each region includes a heterogeneous group of countries with varying levels of happiness. For instance, North America includes Canada (high levels), the US (medium-high levels), and Mexico (medium levels). Western Europe similarly has an even distribution of countries in these three categories—the variety of social and political cultures across Western Europe might have offset the effects of the Nordic countries being in that pool. The one variable that did differ significantly, life expectancy, can be explained by the fact that Mexico has a very low score in that category. Also, it should be noted that Western Europe is known for a healthier lifestyle as opposed to the countries in North America which may also contribute this. However, the failure to reject the null hypothesis for the vast majority of the variables points to similarities between the two

regions which aligns with the understanding that both regions include relatively heterogeneous developed countries. Despite differences in political or economic structures both among and between the regions, the findings are not statistically significant, and therefore potentially suggest that these different models both can provide for national happiness.

Looking at the violin plots for each variable across the two regions, we can see that the distribution is consistently wider in Western Europe, likely due to the greater number of countries included in the region. For the variables GDP per capita, social support, and freedom to make life choices, the medians and IQRs are relatively similar between the two regions. However, the medians for life expectancy are visually quite different between the two regions, which ends up aligning with the findings of our hypothesis, even though the IQRs are similar. Then, for the variables of perceptions of corruption and generosity, the IQR is much larger in Western Europe as opposed to North America with the medians being greater in North America for both variables.

Furthermore, we wanted to evaluate whether happiness was independent of region. Essentially, does the average individual's happiness depend on the region of the world that they live in? Based on an initial survey of the data, we found that happiness varies from country to country. On top of that, some regions have high concentrations of high-, mid-, and low-happiness, suggesting that region also plays a role. For example, while North America, Nordic countries, and Oceania have very high levels of overall happiness, regions like Central Africa and South Asia have lower levels of overall happiness. Through our chi-square test, we were able to conclude that happiness was not independent of region, suggesting that happiness score and region are linked.

When understanding our findings, it's important to evaluate the integrity of the dataset. Considering that the data was collected through Gallup, a group well-respected for proper data collection, it can be deemed reliable. However, it should be noted that the translation of particular interviews and surveys into quantitative numbers leaves room for subjectivity to affect the data. While a variable such as GDP per capita is more concrete, something such as Perceptions of Corruption may be dependent on other factors such as access to education or level of freedom of the media. Consequently, the results of data analysis should be interpreted with this in mind.

In order to perform a Pearson's chi-squared test, four conditions have to be satisfied: simple random sample, sample size, expected cell count, and independence. Based on the description of the dataset, the survey was random, enough people were sampled, and the people sampled were independent of each other. Due to the nature of our data, however, expected counts for several cells were smaller than 5. Therefore, care has to be taken when interpreting the results, because the test loses a lot of accuracy. Because our test conditions were not all satisfied, we chose to add the field simulate.p.value = TRUE to the chi-squared test, which uses the Monte Carlo method to obtain a p-value from random samples from the original sample. However, regardless of whichever method we tried, the chi-square test still yielded a p-value less than 0.01.

After trying different analytical strategies, such as the linear model, we learned that the range of methods applicable to analyze this particular dataset is limited by the ways in which the raw data was manipulated in the creation of the report. While we originally started this project with an interest in looking at correlations between variables and the overall happiness score, we soon learned that this is impossible with this dataset considering how the overall happiness score is composed as the sums of the scores from the individual categories. So, we shifted our strategy to instead compare regions. However, a further step for this analysis would be to dive into the raw data from Gallup in order to look at some of the other potential patterns between these variables across the world. Furthermore, our insights into the comparison of regions is also relatively limited considering the small number of countries within each region, such as only having three countries within North America.

Furthermore, going forward, we would hope to study what specific factors lead to happiness among a general population within a country. In order to do this, we would need more variables that affect people's lives: e.g. level of infrastructure, quality of healthcare, quality of education, safety. We might also seek to answer questions like the following: do the factors that predict happiness vary by region? Why are some regions happier than others? What factor is the most important for happiness across the board?

Generally, quantifying happiness is a fascinating concept that takes into account a range of human experi-

ences. However, we do want to note the potential for cultural biases in defining what makes one "happy," so it would be interesting to explore in more detail the ways in which this is evaluated. Regardless, however, the countless potential paths for further research reflect an opportunity to better understand the human condition, and could potentially inform legislation to better people's lives across the globe.

BIBLIOGRAPHY

[1] Gallup World Poll. (2014, October 14). How does the Gallup world poll work? Gallup.com. https://www.gallup.com/178667/gallup-world-poll-work.aspx
[2] Helliwell, J., Layard, R., & Sachs, J. (2018). World Happiness Report 2018, New York: Sustainable Development Solutions Network.
[3] Helliwell, J., Layard, R., & Sachs, J. (2019). World Happiness Report 2019, New York: Sustainable Development Solutions Network.
[4] Helliwell, John F., Richard Layard, and Jeffrey Sachs, eds. 2012. World Happiness Report 2012. New York: UN Sustainable Development Solutions Network.
[5] https://www.investopedia.com/terms/n/nordic-model.asp