Final Project: An Evaluation of the World Happiness Report

Anna MacFarlane, Edgar Salas, Jerry Fu, Taalin RaoShah

11/4/2020

```
## Rows: 312
## Columns: 11
## $ Overall_rank              <dbl> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12,...
## $ Country                   <chr> "Finland", "Norway", "Denmark", "Icela...
## $ Score                     <dbl> 7.632, 7.594, 7.555, 7.495, 7.487, 7.4...
## $ GDP_per_capita            <dbl> 1.305, 1.456, 1.351, 1.343, 1.420, 1.3...
## $ Social_support            <dbl> 1.592, 1.582, 1.590, 1.644, 1.549, 1.4...
## $ Healthy_life_expectancy   <dbl> 0.874, 0.861, 0.868, 0.914, 0.927, 0.8...
## $ Freedom_to_make_life_choices <dbl> 0.681, 0.686, 0.683, 0.677, 0.660, 0.6...
## $ Generosity                <dbl> 0.202, 0.286, 0.284, 0.353, 0.256, 0.3...
## $ Perceptions_of_corruption <dbl> 0.393, 0.340, 0.408, 0.138, 0.357, 0.2...
## $ year                      <chr> "2018", "2018", "2018", "2018", "2018"...
## $ Region                    <chr> "Western Europe", "Western Europe", "W...
```

Is there a relationship between region of the world and happiness score?

World Map
colored by Happiness Score

https://www.datanovia.com/en/blog/how-to-create-a-map-using-ggplot2/
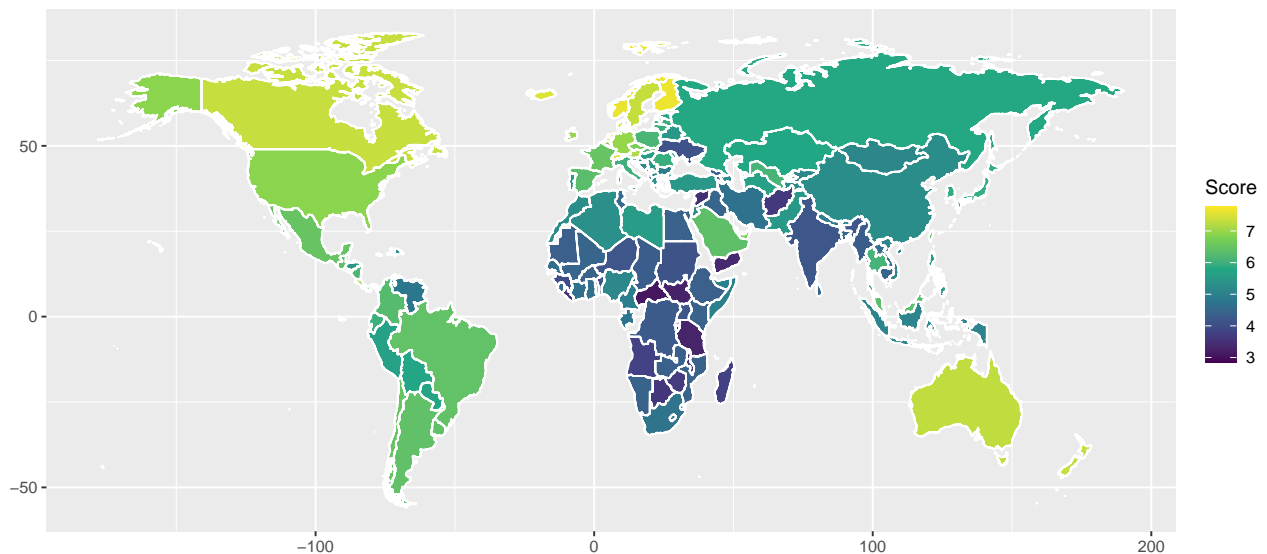
H_0: There is no association between region of the world and happiness score. H_a: There is an association between region of the world and happiness score.

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

```
## data:  worldhappiness$Score and worldhappiness$Region
## X-squared = 2562.2, df = NA, p-value = 0.1024
```

With a simulated p-value of 0.0004998 which is less than our $\alpha = 0.05$, we have sufficient evidence to reject the null hypothesis in favor of the alternative hypothesis and conclude that there is an association between region of the world and happiness score.

Summary Statistics

```
##   WE_mean_happiness median_score        SD
## 1           6.86375        7.003 0.6939077
```

```
##   NA_mean_happiness median_score        SD
## 1          7.189375        7.275 0.1881557
```

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 is less than or equal to the mean score for that same variable in North America, Australia, and New Zealand. Our alternative hypothesis therefore is that the mean score for a particular variable in Western Europe is greater than that of the mean score for that variable in North America, Australia, and New Zealand.

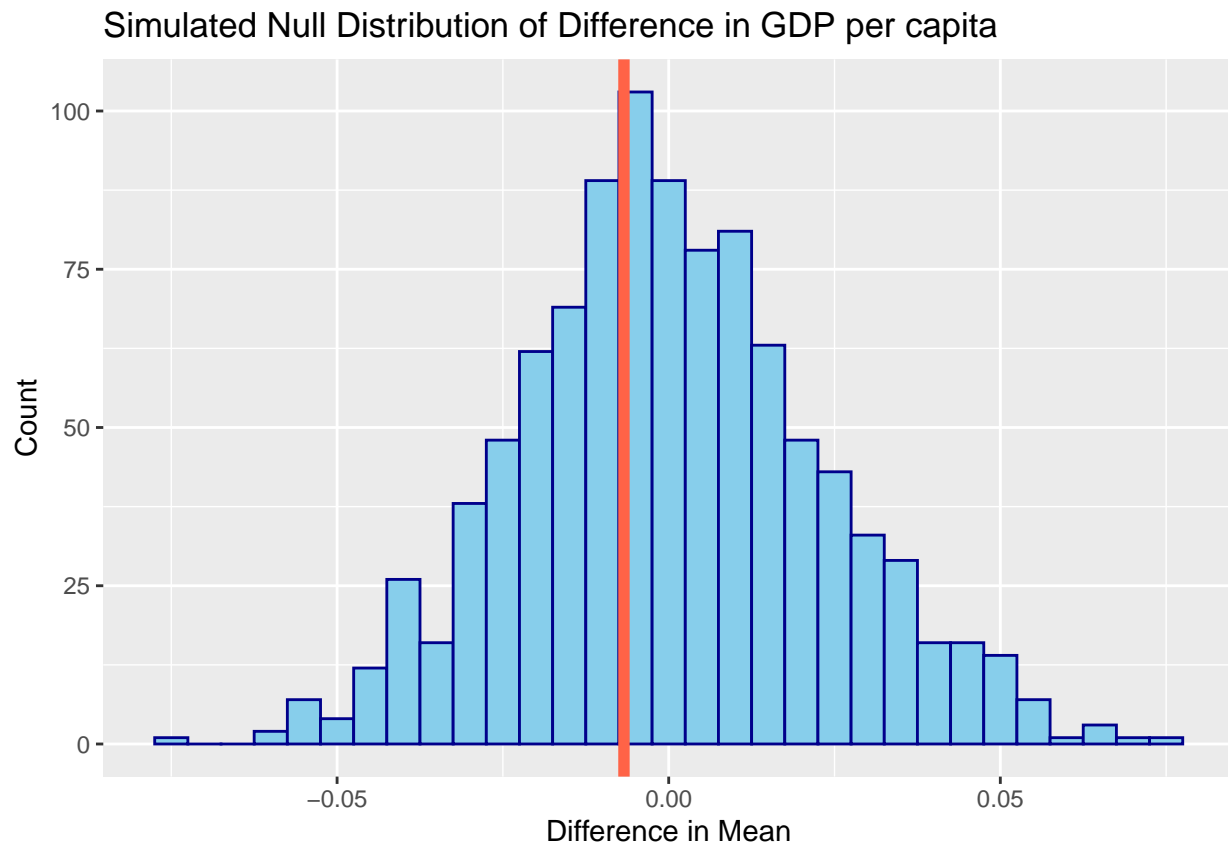H_0: $\mu_{WE} <= \mu_{NA,ANZ}$

H_A: $\mu_{WE} > \mu_{NA,ANZ}$

GDP per capita

```
##    WE_GDP
## 1 1.34385
```

```
##   NA_ANZ_GDP
## 1   1.351125
```

```
##   p_val
## 1 0.366
```

Simulated Null Distribution of Difference in GDP per capita
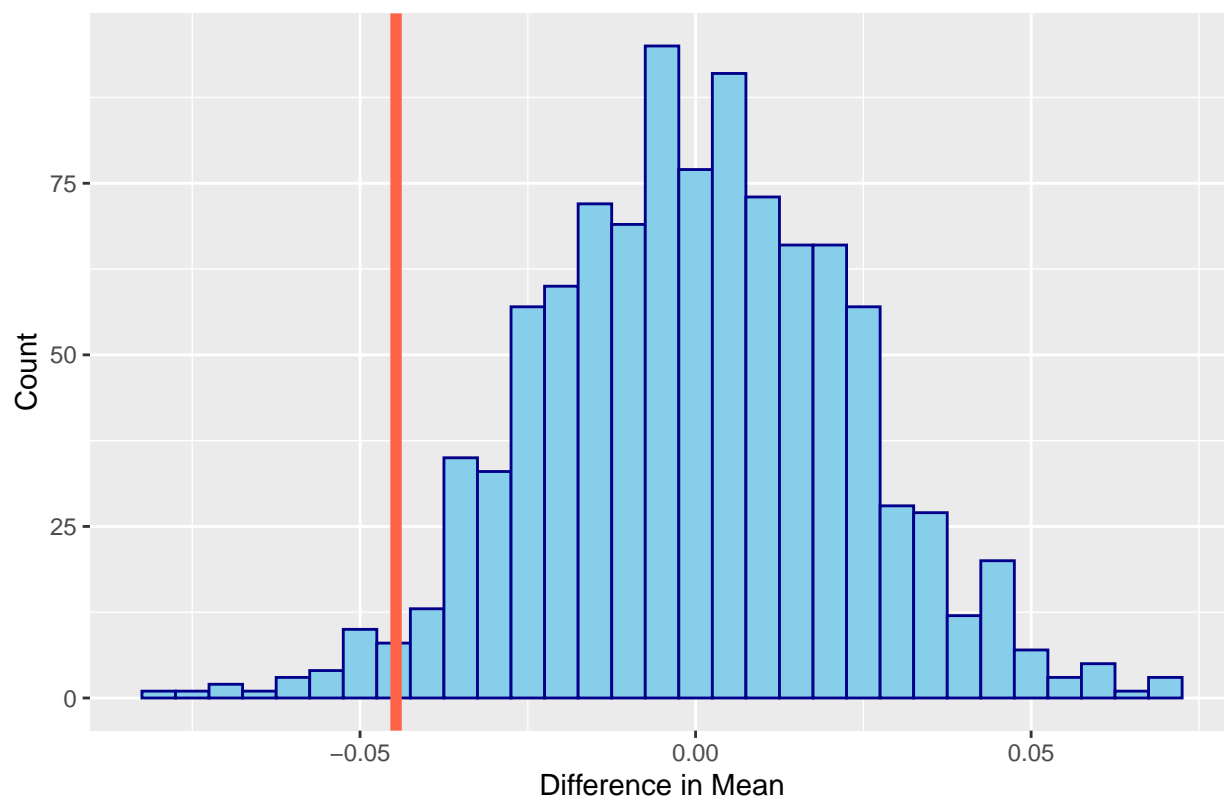
Social Support

```
##   WE_Soc_Support
## 1     1.486825

##   NA_ANZ_Soc_Support
## 1             1.5305

##   p_val
## 1  0.03
```

Simulated Null Distribution of Difference in Social Support
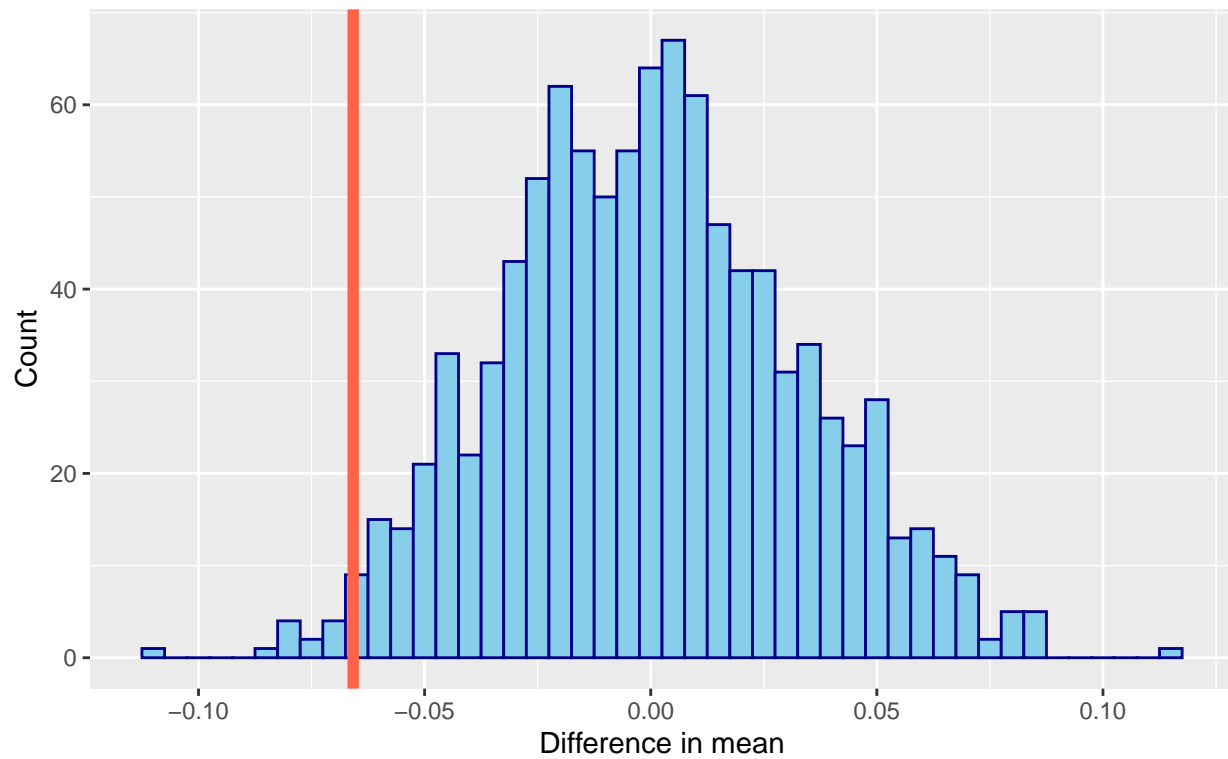

Freedom to make life choices

```
##   WE_Life_Choices
## 1         0.52105

##   NA_ANZ_Life_Choices
## 1               0.587

##   p_val
## 1 0.024
```

Simulated null distribution of difference in mean freedom to make life choices scores
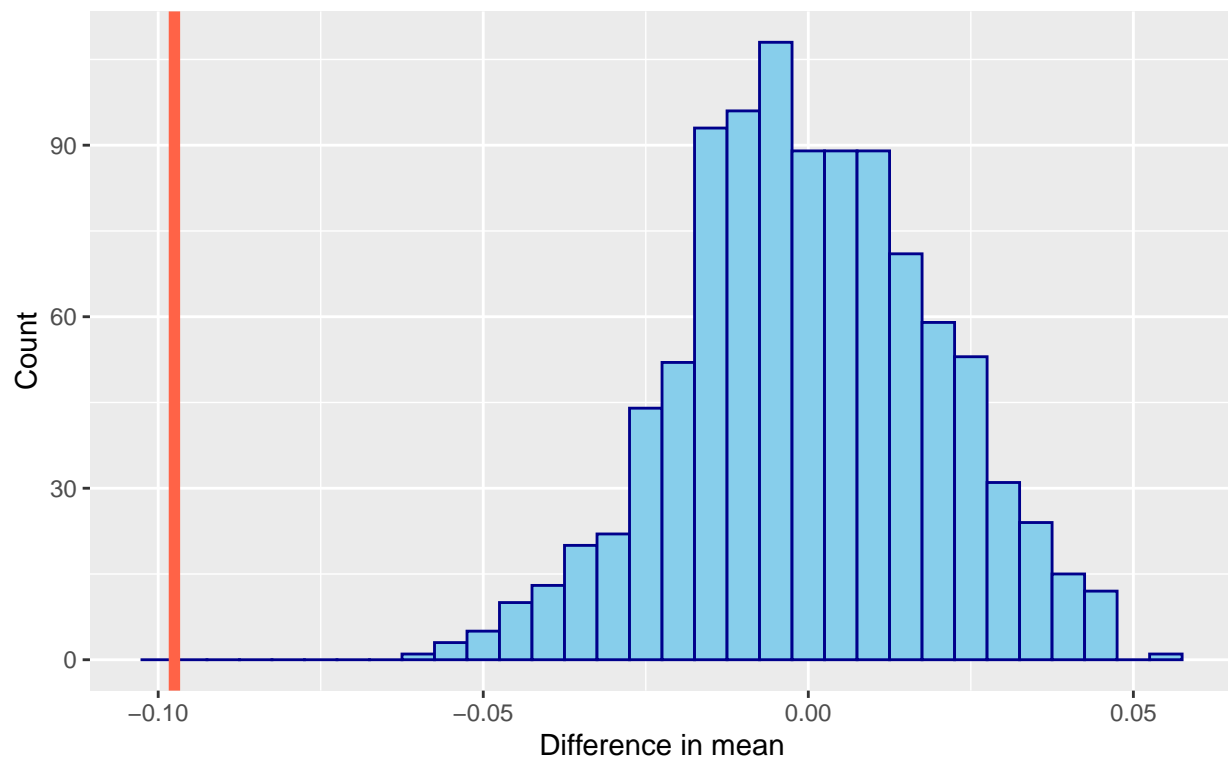
Generosity

```
##      WE_Gen
## 1 0.22245

##   NA_ANZ_Gen
## 1   0.320625

##   p_val
## 1     0
```

Simulated null distribution of difference in mean freedom to make life choices scores

Life Expectancy

```
##   WE_life_expectancy
## 1          0.954825

##   RestofW_life_expectancy
## 1                0.954825

##   p_val
## 1 0.262
```

Simulated Null Distribution of Difference in
 Mean Life Expectancy