

HAPPINESS EXPLAINED: AN EXPLORATION OF HAPPINESS SCORES

Members:

Arlette Varela

Nathan Bolt

Nirupama Shankar

Teshanee Williams



What is the WHR?

–The WHR is a global survey based on the Gallup Poll of 150+ countries that ranks how happy the citizens perceive themselves to be. It digs into citizens' subjective well-being to see how urban, social, and natural environments “combine to affect our happiness.”

–The scale is from 0-10 and is no longer grounded on a definition of “Happiness” but instead a “Happiness Ladder” based on the Cantril Ladder/Scale. Respondents are to imagine they are standing on a ladder with 10 steps, then answer which step they feel on at the current time.

- The six SWBs do not have an impact on the total score, they are sub-categories that aid in answering why some countries have the happiness score they have. They are typically based on 'yes/no' survey questions.

<https://worldhappiness.report/>

[Cantril Ladder](#)

Core Message & Hypotheses

- Core Message
 - According to the **World Happiness Report** there are several **subjective well-being measures (SWB)** that can be used to measure **happiness**.
 - Our main objective was to explore, the measures of subjective well-being to determine which ones had the most influence happiness.
 - We also set out to examine the different factors to determine the correlation of each of these with overall happiness.
 -
- Hypotheses
 - We hypothesized that:
 - The higher absence of corruption would lead to a higher overall happiness score.
 - The higher the GDP per capita would lead to a higher happiness score
- What are the six subjective well-being measures?
Social Support, Healthy Life Expectancy, Generosity, Freedom to Make Life Choices, GDP per capita, and Perception of Corruption

Questions and Motivation

- Questions:

1. Which **countries** had the lowest happiness score? highest happiness score?
2. Which of the SWB (i.e., social support, healthy life expectancy, generosity, freedom to make life choices, GDP per capita, and perception of corruption) are most highly correlated with a higher happiness score?
3. Which of these measures have the most influence on the happiness score?

- What motivated us to answer these questions?

- We wanted some positivity
- Put some perspective into how good some countries have it relative to the lowest
- We discovered Perception of Corruption as an SWB and found it interesting.

Brief Summary

- Were we able to answer the questions to our satisfaction?
 - *Our analysis shows that we were able to answer our questions. Arlette, Niru, and Teshanee will detail how we got there.*
- Brief summary of our findings
 - *We found that GDP, Social Support, and Healthy Life Expectancy had strong correlations*
 - *Moreover, all SWBs were found to be statistically significant except Generosity*
 - (p-value < .05 = statistically significant)

Data Cleanup & Exploration

Describe

We discovered a decades worth of data from [Kaggle](#) and imported the three most recent years of data as CSVs into a Jupyter Notebook.

Clean-up

We compared our three datasets and dropped unnecessary columns. A few countries were not present in all datasets - those were dropped.

Merge

Once all three datasets were uniform, we merged by country into one large dataframe.

Reorder

Lastly, we reordered and titled the columns chronologically and sorted by happiness score descending.

```
10
11 #Import and read csvs - please note that the csv files from kaggle were titled using the year the report was published
12 #The date on the df title reflects the year the actual data applies to (i.e. the happiness2020.csv actually contains data from 2019)
13 file2018 = "happiness2018.csv"
14 file2019 = "happiness2019.csv"
15 file2020 = "happiness2020.csv"
16
17 hap2017_df = pd.read_csv(file2018)
18 hap2018_df = pd.read_csv(file2019)
19 hap2019_df = pd.read_csv(file2020)
```

```
1 #renaming "Country or region" in original 2018 file to "Country" for ease of merging
2 hap2017_df = hap2017_df.rename(columns={"Country or region": "Country"})
3
4 #drop unnecessary columns:
5 clean_2017_df = hap2017_df.drop(columns = ["Overall rank"])
6 clean_2017_df
```

| | Country | Score | GDP per capita | Social support | Healthy life expectancy | Freedom to make life choices | Generosity | Perceptions of corruption |
|-----|--------------------------|-------|----------------|----------------|-------------------------|------------------------------|------------|---------------------------|
| 0 | Finland | 7.632 | 1.305 | 1.592 | 0.874 | 0.681 | 0.202 | 0.393 |
| 1 | Norway | 7.594 | 1.456 | 1.582 | 0.861 | 0.686 | 0.286 | 0.340 |
| 2 | Denmark | 7.555 | 1.351 | 1.590 | 0.868 | 0.683 | 0.284 | 0.408 |
| 3 | Iceland | 7.495 | 1.343 | 1.644 | 0.914 | 0.677 | 0.353 | 0.138 |
| 4 | Switzerland | 7.487 | 1.420 | 1.549 | 0.927 | 0.660 | 0.256 | 0.357 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 151 | Yemen | 3.355 | 0.442 | 1.073 | 0.343 | 0.244 | 0.083 | 0.064 |
| 152 | Tanzania | 3.303 | 0.455 | 0.991 | 0.381 | 0.481 | 0.270 | 0.097 |
| 153 | South Sudan | 3.254 | 0.337 | 0.608 | 0.177 | 0.112 | 0.224 | 0.106 |
| 154 | Central African Republic | 3.083 | 0.024 | 0.000 | 0.010 | 0.305 | 0.218 | 0.038 |
| 155 | Burundi | 2.905 | 0.091 | 0.627 | 0.145 | 0.065 | 0.149 | 0.076 |

156 rows x 8 columns

```
1 #drop columns we dont need in 2020 data and rename "Country name" column to "Country" for ease of merging
2 hap2019_df = hap2019_df.rename(columns={"Country name": "Country"})
3
4 #drop unnecessary columns:
5 clean_2019_df = hap2019_df.drop(columns = ["standard error of ladder score",
6                                           "logged GDP per capita",
7                                           "upperwhisker",
8                                           "lowerwhisker",
9                                           "Ladder score in Dystopia",
10                                          "Dystopia + residual",
11                                          "Social support",
12                                          "Healthy life expectancy",
13                                          "Freedom to make life choices",
14                                          "Generosity",
15                                          "Perceptions of corruption"])
16
17
18 #rename columns:
19 clean_2019_df = clean_2019_df.rename(columns={
20     "Explained by: Log GDP per capita": "GDP per capita 2019",
21     "Explained by: Social support": "Social support 2019",
22     "Explained by: Healthy life expectancy": "Healthy life expectancy 2019",
23     "Explained by: Freedom to make life choices": "Freedom to make life choices 2019",
24     "Explained by: Generosity": "Generosity 2019",
25     "Explained by: Perceptions of corruption": "Perceptions of corruption 2019",
26     "Ladder score": "Score 2019"
27 })
28
29
30 clean_2019_df
31 ..
```

```
1 #no further cleaning of our 2018,2019 and 2020 data is needed. Now we can merge the files.
2 merge1_df = pd.merge(clean_2017_df, clean_2018_df, on = "Country", how = "inner", suffixes=("_ 2017", "_ 2018"))
3 merge1_df.head(10)
```

| | Country | Score 2017 | GDP per capita 2017 | Social support 2017 | Healthy life expectancy 2017 | Freedom to make life choices 2017 | Generosity 2017 | Perceptions of corruption 2017 | Score 2018 | GDP per capita 2018 | Social support 2018 | Healthy life expectancy 2018 | Freedom to make life choices 2018 | Generosity 2018 | Perceptions of corruption 2018 |
|---|---------|------------|---------------------|---------------------|------------------------------|-----------------------------------|-----------------|--------------------------------|------------|---------------------|---------------------|------------------------------|-----------------------------------|-----------------|--------------------------------|
| 0 | Finland | 7.632 | 1.305 | 1.592 | 0.874 | 0.681 | 0.202 | 0.393 | 7.769 | 1.340 | 1.587 | 0.986 | 0.596 | 0.153 | 0.393 |
| 1 | Norway | 7.594 | 1.456 | 1.582 | 0.861 | 0.686 | 0.286 | 0.340 | 7.554 | 1.488 | 1.582 | 1.028 | 0.603 | 0.271 | 0.341 |
| 2 | Denmark | 7.555 | 1.351 | 1.590 | 0.868 | 0.683 | 0.284 | 0.408 | 7.600 | 1.383 | 1.573 | 0.996 | 0.592 | 0.252 | 0.410 |
| 3 | Iceland | 7.495 | 1.343 | 1.644 | 0.914 | | | | | | | | | | |

```
1 #merge 2019 data into 2017,2018 df:
2 merge2_df = pd.merge(merge1_df, clean_2019_df, on = "Country", how = "inner")
3 merge2_df
```

| | Country | Score 2017 | GDP per capita 2017 | Social support 2017 | Healthy life expectancy 2017 | Freedom to make life choices 2017 | Generosity 2017 | Perceptions of corruption 2017 | Score 2018 | GDP per capita 2018 | ... | Generosity 2018 | Perceptions of corruption 2018 | Regional indicator | Score 2019 | GDP per capita 2019 |
|---|---------|------------|---------------------|---------------------|------------------------------|-----------------------------------|-----------------|--------------------------------|------------|---------------------|-----|-----------------|--------------------------------|--------------------|------------|---------------------|
| 0 | Finland | 7.632 | 1.305 | 1.592 | 0.874 | 0.681 | 0.202 | 0.393 | 7.769 | 1.340 | ... | 0.153 | 0.393 | Western Europe | 7.8087 | 1.28519C |
| 1 | Norway | 7.594 | 1.456 | 1.582 | 0.861 | 0.686 | 0.286 | 0.340 | 7.554 | 1.488 | ... | 0.271 | 0.341 | Western Europe | 7.4880 | 1.424207 |

Data Description

- CSV files for each year with a column for each SWB

| Overall rank | Country or region | Score | GDP per capita | Social support | Healthy life expectancy | Freedom to make life choices | Generosity | Perceptions of corruption |
|--------------|-------------------|-------|----------------|----------------|-------------------------|------------------------------|------------|---------------------------|
| 1 | Finland | 7.632 | 1.305 | 1.592 | 0.874 | 0.681 | 0.202 | 0.393 |
| 2 | Norway | 7.594 | 1.456 | 1.582 | 0.861 | 0.686 | 0.286 | 0.34 |
| 3 | Denmark | 7.555 | 1.351 | 1.59 | 0.868 | 0.683 | 0.284 | 0.408 |
| 4 | Iceland | 7.495 | 1.343 | 1.644 | 0.914 | 0.677 | 0.353 | 0.138 |
| 5 | Switzerland | 7.487 | 1.42 | 1.549 | 0.927 | 0.66 | 0.256 | 0.357 |
| 6 | Netherlands | 7.441 | 1.361 | 1.488 | 0.878 | 0.638 | 0.333 | 0.295 |
| 7 | Canada | 7.328 | 1.33 | 1.532 | 0.896 | 0.653 | 0.321 | 0.291 |
| 8 | New Zealand | 7.324 | 1.268 | 1.601 | 0.876 | 0.669 | 0.365 | 0.389 |
| 9 | Sweden | 7.314 | 1.355 | 1.501 | 0.913 | 0.659 | 0.285 | 0.383 |
| 10 | Australia | 7.272 | 1.34 | 1.573 | 0.91 | 0.647 | 0.361 | 0.302 |
| 11 | United Kingdom | 7.19 | 1.244 | 1.433 | 0.888 | 0.464 | 0.262 | 0.082 |
| 12 | Austria | 7.139 | 1.341 | 1.504 | 0.891 | 0.617 | 0.242 | 0.224 |
| 13 | Costa Rica | 7.072 | 1.01 | 1.459 | 0.817 | 0.632 | 0.143 | 0.101 |

| Overall rank | Country or region | Score | GDP per capita | Social support | Healthy life expectancy | Freedom to make life choices | Generosity | Perceptions of corruption |
|--------------|-------------------|-------|----------------|----------------|-------------------------|------------------------------|------------|---------------------------|
| 1 | Finland | 7.769 | 1.34 | 1.587 | 0.986 | 0.596 | 0.153 | 0.393 |
| 2 | Denmark | 7.6 | 1.383 | 1.573 | 0.996 | 0.592 | 0.252 | 0.41 |
| 3 | Norway | 7.554 | 1.488 | 1.582 | 1.028 | 0.603 | 0.271 | 0.341 |
| 4 | Iceland | 7.494 | 1.38 | 1.624 | 1.026 | 0.591 | 0.354 | 0.118 |
| 5 | Netherlands | 7.488 | 1.396 | 1.522 | 0.999 | 0.557 | 0.322 | 0.298 |
| 6 | Switzerland | 7.48 | 1.452 | 1.526 | 1.052 | 0.572 | 0.263 | 0.343 |
| 7 | Sweden | 7.343 | 1.387 | 1.487 | 1.009 | 0.574 | 0.267 | 0.373 |
| 8 | New Zealand | 7.307 | 1.303 | 1.557 | 1.026 | 0.585 | 0.33 | 0.38 |
| 9 | Canada | 7.278 | 1.365 | 1.505 | 1.039 | 0.584 | 0.285 | 0.308 |
| 10 | Austria | 7.246 | 1.376 | 1.475 | 1.016 | 0.532 | 0.244 | 0.226 |
| 11 | Australia | 7.228 | 1.372 | 1.548 | 1.036 | 0.557 | 0.332 | 0.29 |
| 12 | Costa Rica | 7.167 | 1.034 | 1.441 | 0.963 | 0.558 | 0.144 | 0.093 |
| 13 | Israel | 7.139 | 1.276 | 1.455 | 1.029 | 0.371 | 0.261 | 0.082 |

| Country name | Regional indicator | Ladder score | Standard error of ladder score | upperwhisker | lowerwhisker | Logged GDP per capita | Social support | Healthy life expectancy | Freedom to make life choices | Generosity | Perceptions of corruption | Ladder score in Dystopia | Explained by: Log GDP per capita | Explained by: Social support | Explained by: Healthy life expectancy | Explained by: Freedom to make life choices | Explained by: Generosity | Explained by: Perceptions of corruption | Dystopia + residual |
|--------------|---------------------------|--------------|--------------------------------|--------------|--------------|-----------------------|----------------|-------------------------|------------------------------|------------|---------------------------|--------------------------|----------------------------------|------------------------------|---------------------------------------|--|--------------------------|---|---------------------|
| Finland | Western Europe | 7.8087 | 0.031156 | 7.86977 | 7.747634 | 10.63927 | 0.95433 | 71.90083 | 0.949172 | -0.05948 | 0.195445 | 1.972317 | 1.28519 | 1.499526 | 0.961271 | 0.662317 | 0.15967 | 0.477857 | 2.762835 |
| Denmark | Western Europe | 7.6456 | 0.033492 | 7.71124 | 7.579955 | 10.774 | 0.955991 | 72.4025 | 0.951444 | 0.066202 | 0.168489 | 1.972317 | 1.326949 | 1.503449 | 0.979333 | 0.66504 | 0.242793 | 0.49526 | 2.432741 |
| Switzerland | Western Europe | 7.5599 | 0.035014 | 7.62853 | 7.491272 | 10.97993 | 0.942847 | 74.10245 | 0.921337 | 0.105911 | 0.303728 | 1.972317 | 1.390774 | 1.472403 | 1.040533 | 0.628954 | 0.269056 | 0.407946 | 2.350267 |
| Iceland | Western Europe | 7.5045 | 0.059616 | 7.62135 | 7.387653 | 10.77256 | 0.97467 | 73 | 0.948892 | 0.246944 | 0.711171 | 1.972317 | 1.326502 | 1.547567 | 1.000843 | 0.661981 | 0.36233 | 0.144541 | 2.460688 |
| Norway | Western Europe | 7.488 | 0.034837 | 7.55628 | 7.419719 | 11.0878 | 0.952487 | 73.20078 | 0.95575 | 0.134533 | 0.263218 | 1.972317 | 1.424207 | 1.495173 | 1.008072 | 0.670201 | 0.287985 | 0.434101 | 2.168266 |
| Netherlands | Western Europe | 7.4489 | 0.027792 | 7.50337 | 7.394428 | 10.81271 | 0.939139 | 72.30092 | 0.908548 | 0.207612 | 0.364717 | 1.972317 | 1.338946 | 1.463646 | 0.975675 | 0.613626 | 0.336318 | 0.36857 | 2.352117 |
| Sweden | Western Europe | 7.3535 | 0.036234 | 7.42452 | 7.282481 | 10.75879 | 0.926311 | 72.60077 | 0.939144 | 0.111615 | 0.25088 | 1.972317 | 1.322235 | 1.433348 | 0.98647 | 0.650298 | 0.272828 | 0.442066 | 2.246299 |
| New Zealand | North America and Oceania | 7.2996 | 0.039465 | 7.37695 | 7.222248 | 10.50094 | 0.949119 | 73.20263 | 0.936217 | 0.191598 | 0.221139 | 1.972317 | 1.242318 | 1.487218 | 1.008138 | 0.64679 | 0.325726 | 0.461268 | 2.128108 |
| Austria | Western Europe | 7.2942 | 0.033365 | 7.3596 | 7.228804 | 10.74282 | 0.928046 | 73.0025 | 0.899989 | 0.085429 | 0.499955 | 1.972317 | 1.317286 | 1.437445 | 1.000934 | 0.603369 | 0.25551 | 0.281256 | 2.398446 |
| Luxembourg | Western Europe | 7.2375 | 0.030852 | 7.29797 | 7.177031 | 11.45068 | 0.906912 | 72.6 | 0.905636 | -0.00462 | 0.367084 | 1.972317 | 1.536676 | 1.387528 | 0.986443 | 0.610137 | 0.195954 | 0.367041 | 2.1537 |
| Canada | North America and Oceania | 7.2321 | 0.040405 | 7.31129 | 7.152905 | 10.69237 | 0.927177 | 73.6016 | 0.933913 | 0.124771 | 0.390843 | 1.972317 | 1.301648 | 1.435392 | 1.022502 | 0.644028 | 0.281529 | 0.351702 | 2.195269 |
| Australia | North America and Oceania | 7.2228 | 0.041841 | 7.30481 | 7.140791 | 10.7206 | 0.944855 | 73.60454 | 0.915432 | 0.19046 | 0.415169 | 1.972317 | 1.310396 | 1.477146 | 1.022608 | 0.621877 | 0.324974 | 0.335996 | 2.129804 |

Cleanup, Merge & Reorder

- Columns differed by name. Some countries were not represented in all three years - those countries were dropped.
- We compared our three datasets and performed the following tasks:
 - dropped unnecessary columns,
 - verified data types were consistent,
 - renamed columns,
 - reorganized columns,
 - created a new merged data frame with all three years of data.

[illegible]

Final Dataframes

- Our primary analysis is based on the average of three years of data. Our main dataframe, which was used for the majority of our analysis, was created by taking the average of all three years of data and sorting by descending overall happiness score.

[illegible]

Final Dataframes Cont.

We then created two new dataframes from our averaged dataframe to find the 10 countries with the highest and lowest happiness scores. These two data frames were also used to create our stacked bar graphs for the 10 countries with highest and lowest happiness scores.

```
top10_df = averagesorted_df.head(10)
last10_df = averagesorted_df.tail(10)
top10_df
```

| | Country | Region | Score Average | GDP per Capita Average | Social Support Average | Healthy Life Expectancy Average | Generosity Average | Freedom to Make Life Choices Average | Perceptions of Corruption Average |
|---|-------------|-----------------------|---------------|------------------------|------------------------|---------------------------------|--------------------|--------------------------------------|-----------------------------------|
| 0 | Finland | Western Europe | 7.736567 | 1.316667 | 1.559509 | 0.940424 | 0.171557 | 0.646439 | 0.421286 |
| 2 | Denmark | Western Europe | 7.600200 | 1.361667 | 1.555483 | 0.947778 | 0.259598 | 0.646680 | 0.437753 |
| 1 | Norway | Western Europe | 7.545333 | 1.466667 | 1.553058 | 0.965691 | 0.281662 | 0.653067 | 0.371700 |
| 4 | Switzerland | Western Europe | 7.508967 | 1.430667 | 1.515801 | 1.006511 | 0.262685 | 0.620318 | 0.369315 |
| 3 | Iceland | Western Europe | 7.497833 | 1.355333 | 1.605189 | 0.980281 | 0.356443 | 0.643327 | 0.133514 |
| 5 | Netherlands | Western Europe | 7.459300 | 1.372667 | 1.491215 | 0.950892 | 0.330439 | 0.602875 | 0.320523 |
| 8 | Sweden | Western Europe | 7.336833 | 1.365667 | 1.473783 | 0.969490 | 0.274943 | 0.627766 | 0.399355 |
| 7 | New Zealand | North America and ANZ | 7.310200 | 1.279667 | 1.548406 | 0.970046 | 0.340242 | 0.633597 | 0.410089 |
| 6 | Canada | North America and ANZ | 7.279367 | 1.341667 | 1.490797 | 0.985834 | 0.295843 | 0.627009 | 0.316901 |
| 9 | Australia | North America and ANZ | 7.240933 | 1.350667 | 1.532715 | 0.989536 | 0.339325 | 0.608626 | 0.309332 |

Data Analysis

Analysis steps:

01

Identify the top ten and bottom ten countries

- Sorting the “average” data frame by Score
- Creating two distinct data frames for each
- Using stacked bar graphs to represent SWBs

02

Understand data distribution

- Used `.describe()` to look at the data
- Calculating quartiles
- Calculating outliers
- Plotting Box and Whisker for each SWB

03

Identify correlation amongst SWBs, and each SWB with the Happiness Score

- Calculating pearson's correlation coefficient
- Using Seaborn to generate correlation heatmap

04

Study the extent of influence of each SWB on the Happiness Score

- Examining the linear relationship between the variables
- Multiple linear regression analysis to understand influence

Questions & Data

- Q1: Which **countries** had the lowest happiness score? highest happiness score?
 - *The average over the last three years shows Finland as having the happiest overall score. Rounding out the top five is Denmark, Norway, Switzerland, and Iceland.*
 - *On the other hand, the lowest happiness scores fall on South Sudan, Afghanistan, and Central African Republic.*

Top 10

| | Country | Region | Score Average | GDP per Capita Average | Social Support Average | Healthy Life Expectancy Average | Generosity Average | Freedom to Make Life Choices Average | Perceptions of Corruption Average |
|---|-------------|-----------------------|---------------|------------------------|------------------------|---------------------------------|--------------------|--------------------------------------|-----------------------------------|
| 0 | Finland | Western Europe | 7.736567 | 1.316667 | 1.559509 | 0.940424 | 0.171557 | 0.646439 | 0.421286 |
| 2 | Denmark | Western Europe | 7.600200 | 1.361667 | 1.555483 | 0.947778 | 0.259598 | 0.646680 | 0.437753 |
| 1 | Norway | Western Europe | 7.545333 | 1.466667 | 1.553058 | 0.965691 | 0.281662 | 0.653067 | 0.371700 |
| 4 | Switzerland | Western Europe | 7.508967 | 1.430667 | 1.515801 | 1.006511 | 0.262685 | 0.620318 | 0.369315 |
| 3 | Iceland | Western Europe | 7.497833 | 1.355333 | 1.605189 | 0.980281 | 0.356443 | 0.643327 | 0.133514 |
| 5 | Netherlands | Western Europe | 7.459300 | 1.372667 | 1.491215 | 0.950892 | 0.330439 | 0.602875 | 0.320523 |
| 8 | Sweden | Western Europe | 7.336833 | 1.365667 | 1.473783 | 0.969490 | 0.274943 | 0.627766 | 0.399355 |
| 7 | New Zealand | North America and ANZ | 7.310200 | 1.279667 | 1.548406 | 0.970046 | 0.340242 | 0.633597 | 0.410089 |
| 6 | Canada | North America and ANZ | 7.279367 | 1.341667 | 1.490797 | 0.985834 | 0.295843 | 0.627009 | 0.316901 |
| 9 | Australia | North America and ANZ | 7.240933 | 1.350667 | 1.532715 | 0.989536 | 0.339325 | 0.608626 | 0.309332 |

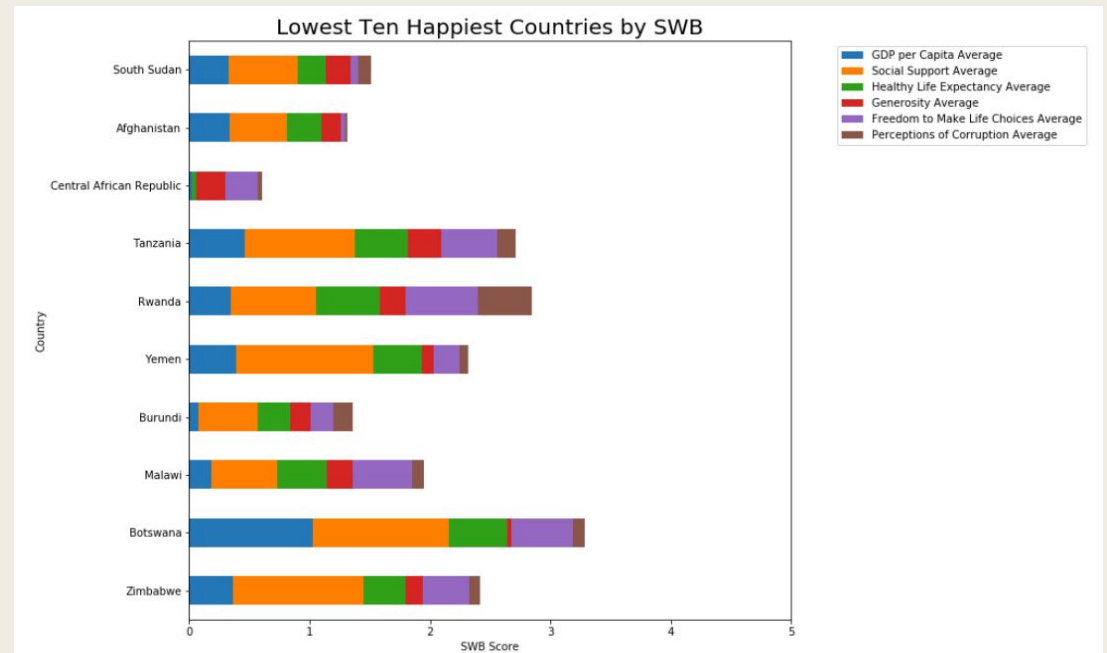
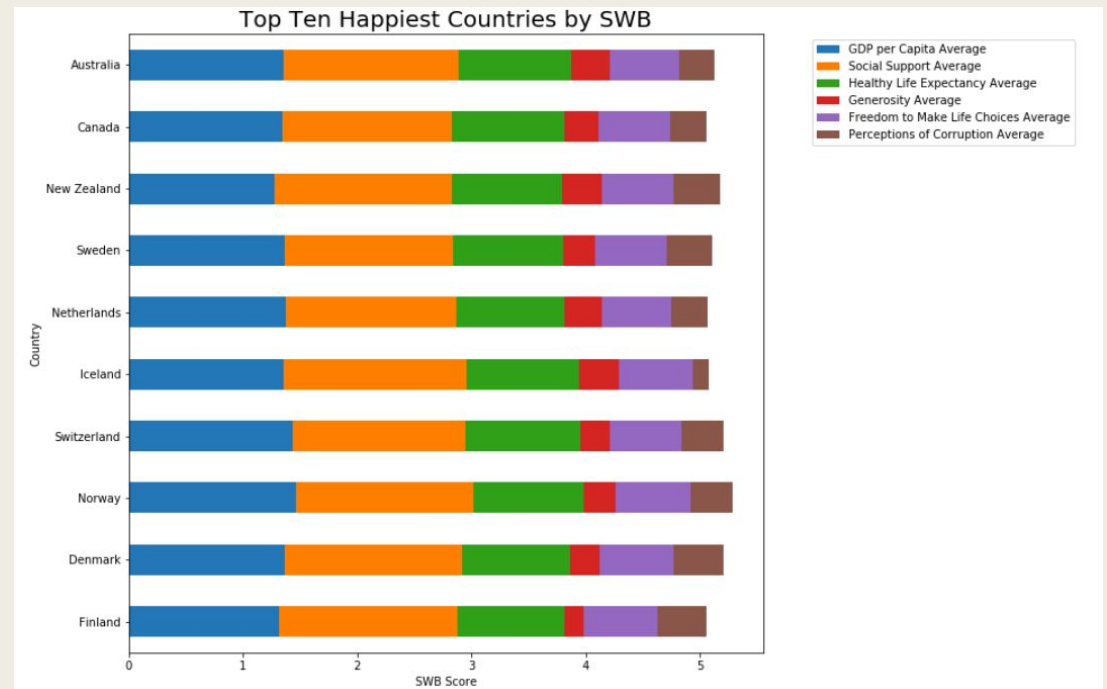
Bottom 10

| | Country | Region | Score Average | GDP per Capita Average | Social Support Average | Healthy Life Expectancy Average | Generosity Average | Freedom to Make Life Choices Average | Perceptions of Corruption Average |
|-----|--------------------------|------------------------------|---------------|------------------------|------------------------|---------------------------------|--------------------|--------------------------------------|-----------------------------------|
| 132 | Zimbabwe | Sub-Saharan Africa | 3.551400 | 0.360000 | 1.085278 | 0.352013 | 0.144783 | 0.381468 | 0.089643 |
| 134 | Botswana | Sub-Saharan Africa | 3.518967 | 1.025000 | 1.134898 | 0.483034 | 0.033469 | 0.507030 | 0.097929 |
| 135 | Malawi | Sub-Saharan Africa | 3.511667 | 0.187667 | 0.543789 | 0.415721 | 0.213728 | 0.487130 | 0.100211 |
| 143 | Burundi | Sub-Saharan Africa | 3.485100 | 0.076000 | 0.492525 | 0.273404 | 0.170801 | 0.186800 | 0.156062 |
| 139 | Yemen | Middle East and North Africa | 3.420800 | 0.390333 | 1.137826 | 0.407000 | 0.095230 | 0.210240 | 0.076117 |
| 138 | Rwanda | Sub-Saharan Africa | 3.351433 | 0.341000 | 0.709959 | 0.528794 | 0.217568 | 0.598363 | 0.446847 |
| 140 | Tanzania | Sub-Saharan Africa | 3.336733 | 0.462000 | 0.916225 | 0.440893 | 0.272514 | 0.469114 | 0.149294 |
| 142 | Central African Republic | Sub-Saharan Africa | 3.213967 | 0.024667 | 0.000000 | 0.038333 | 0.235504 | 0.274271 | 0.033755 |
| 133 | Afghanistan | South Asia | 3.133967 | 0.338000 | 0.470145 | 0.294017 | 0.161412 | 0.028333 | 0.020742 |
| 141 | South Sudan | Sub-Saharan Africa | 2.974533 | 0.326667 | 0.578760 | 0.226936 | 0.211978 | 0.062536 | 0.102719 |

Questions & Data

Q1 cont.

- Q1: Which **countries** had the lowest happiness score? highest happiness score?
 - *Generosity does not appear to occupy a large proportion in comparison to the other SWBs for the top ten countries.*
 - *Social support appears to have a larger presence in the top ten and bottom 10 happiest nations*



Questions & Data: Cont.

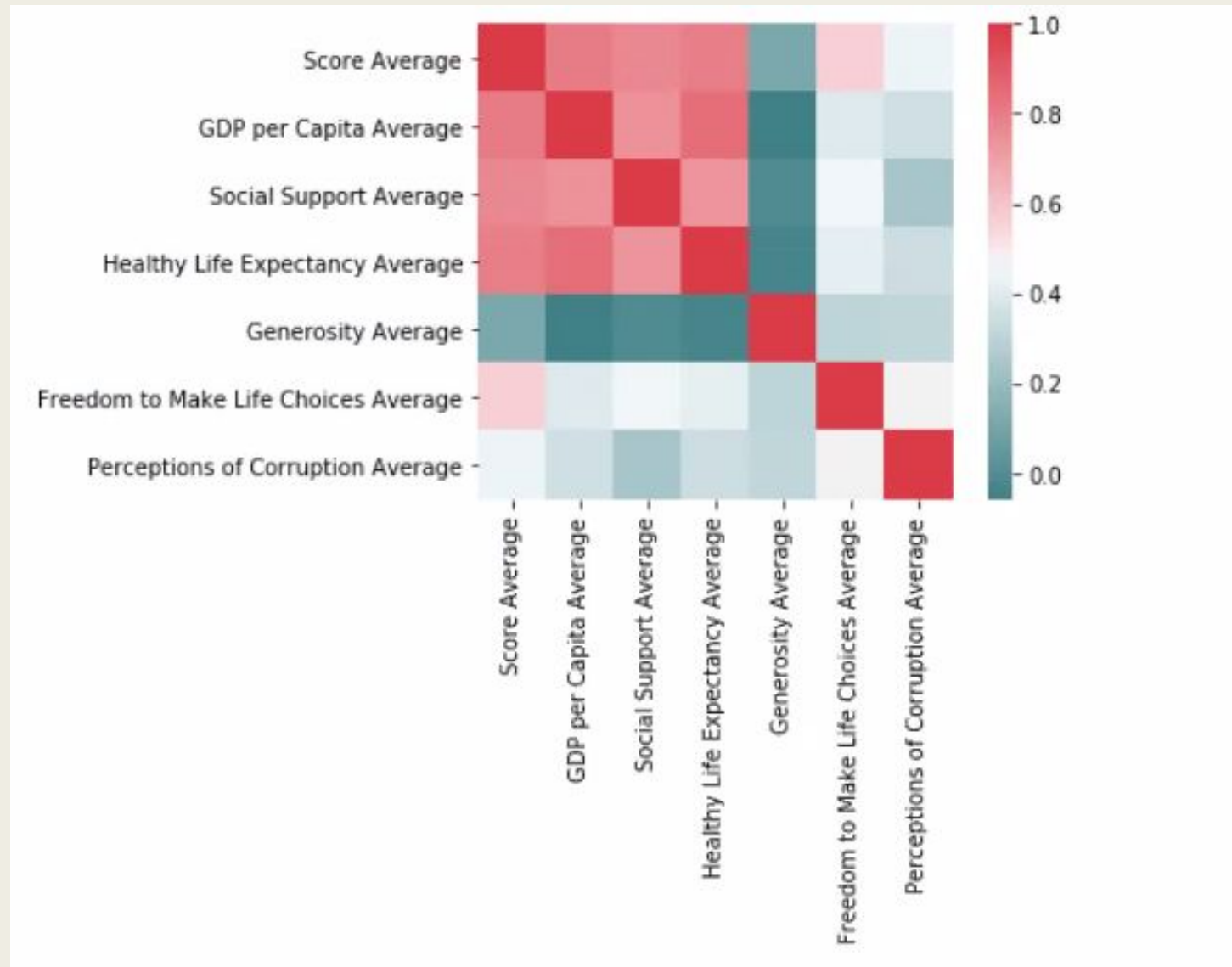
- Q2: Which of the measures of SWB (i.e., social support, healthy life expectancy, generosity, freedom to make life choices, GDP per capita, and perception of corruption) are most highly correlated with a higher happiness score?
 - *We found that GDP had a highest correlation (at .807), Healthy Life Expectancy a close second (at 0.798), and Social Support was third (at 0.774)*
 - *All other SWBs were moderate to weak correlations.*

| | Score Average | GDP per Capita Average | Social Support Average | Healthy Life Expectancy Average | Generosity Average | Freedom to Make Life Choices Average | Perceptions of Corruption Average |
|--------------------------------------|---------------|------------------------|------------------------|---------------------------------|--------------------|--------------------------------------|-----------------------------------|
| Score Average | 1.000000 | 0.807388 | 0.774377 | 0.798818 | 0.112926 | 0.569723 | 0.439359 |
| GDP per Capita Average | 0.807388 | 1.000000 | 0.744840 | 0.848141 | -0.058766 | 0.393532 | 0.352392 |
| Social Support Average | 0.774377 | 0.744840 | 1.000000 | 0.735922 | -0.001536 | 0.449281 | 0.239814 |
| Healthy Life Expectancy Average | 0.798818 | 0.848141 | 0.735922 | 1.000000 | -0.031951 | 0.412639 | 0.348450 |
| Generosity Average | 0.112926 | -0.058766 | -0.001536 | -0.031951 | 1.000000 | 0.301899 | 0.316261 |
| Freedom to Make Life Choices Average | 0.569723 | 0.393532 | 0.449281 | 0.412639 | 0.301899 | 1.000000 | 0.461814 |
| Perceptions of Corruption Average | 0.439359 | 0.352392 | 0.239814 | 0.348450 | 0.316261 | 0.461814 | 1.000000 |

Key:

| | Coefficient <i>r</i> | |
|-----------------------|----------------------|--------------|
| | Positive | Negative |
| Strong | 1 to 0.8 | -0.8 to -1 |
| Moderate | 0.8 to 0.5 | -0.5 to -0.8 |
| Weak | 0.5 to 0.3 | -0.3 to -0.5 |
| No Correlation | 0.3 to 0 | 0 to -0.3 |

Heatmap - correlation amongst SWBs



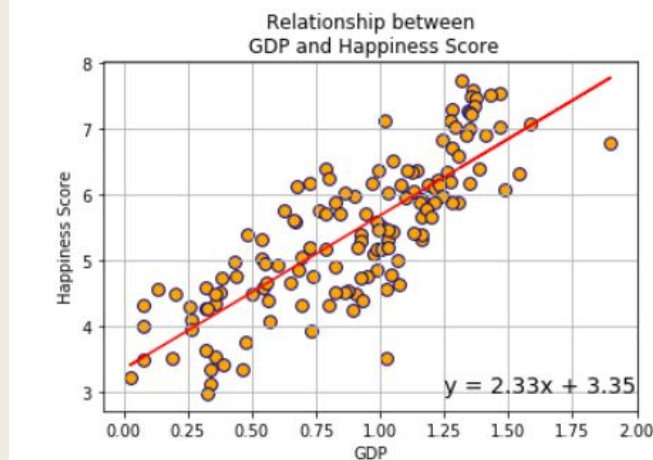
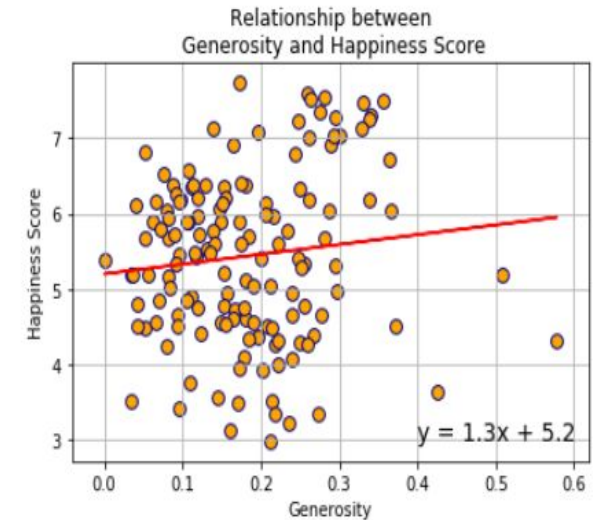
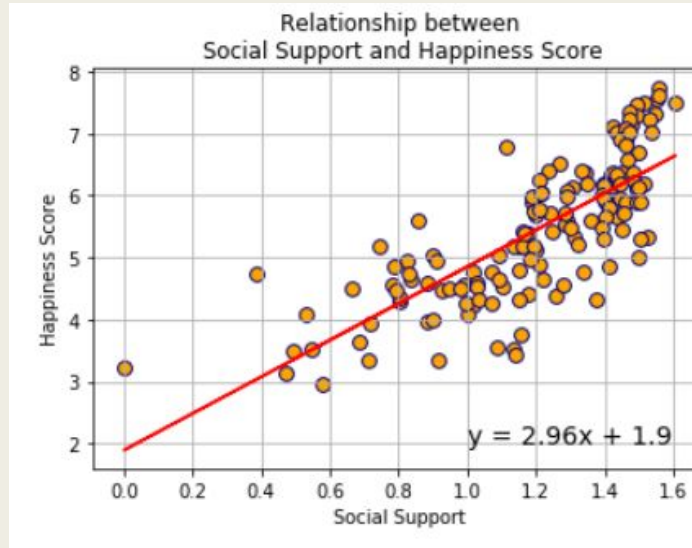
Questions & Data: Cont.

- *In order to study the relationship between the SWBs and the Happiness Score, we ran a series of single linear regressions of each SWB with the Happiness Score*

```
def line_regression_plot(x_values, y_values, x_desc, y_desc,
                        Determinant, file_name, x_notation, y_notation):
    (slope, intercept, rvalue, pvalue, stderr) = linregress(x_values, y_values)
    regress_values = x_values * slope + intercept
    line_eq = "y = " + str(round(slope,2)) + "x + " + str(round(intercept,2))
    plt.scatter(x_values,y_values,s= 60, color="orange", edgecolor = "navy")
    plt.plot(x_values,regress_values,"r-")
    plt.annotate(line_eq,(x_notation, y_notation),fontsize=14,color="black")
    plt.title(f"Relationship between\n {x_desc} and {y_desc}")
    plt.xlabel(x_desc)
    plt.ylabel(y_desc)
    plt.grid(True)
    print(f'The r_squared value is: {(rvalue**2).round(2)}')
    print(f'The slope is: {(slope).round(2)} and intercept is: {(intercept).round(2)}')
    print('The p-value is {:.2f}'.format(pvalue))
    print(line_eq)
    plt.savefig(file_name)
    plt.show
```

Questions & Data: Cont.

- To test for the assumption of linearity, we ran a linear regression to examine the relationship **between Happiness and each SWB.**
- The analysis revealed a **significant relationship** between Happiness and each of the SWB variables **except for Generosity.**



| Statistical Values | r-squared | p-value |
|------------------------------|-----------|---------|
| GDP per Capita Average | 0.65 | 0.0000 |
| Social Support | 0.60 | 0.0000 |
| Healthy Life Expectancy | 0.64 | 0.0000 |
| Generosity | 0.01 | 0.1778 |
| Freedom to make Life Choices | 0.32 | 0.0000 |
| Perceptions of Corruption | 0.12 | 0.0000 |

Questions & Data: Cont.

Q3: Which of these measures have the most influence on the happiness score?

- We utilized a Multiple linear regression model to examine **influence**
 - As expected, our predictive model shows that all of the variables are significant except generosity.

```
=====
                        OLS Regression Results
=====
Dep. Variable:          Score Average      R-squared:                0.792
Model:                  OLS               Adj. R-squared:           0.783
Method:                 Least Squares      F-statistic:              86.82
Date:                   Tue, 21 Jul 2020    Prob (F-statistic):       3.23e-44
Time:                   21:41:13           Log-Likelihood:          -107.11
No. Observations:       144               AIC:                     228.2
Df Residuals:           137               BIC:                     249.0
Df Model:                6
Covariance Type:        nonrobust
=====
```

| | coef | std err | t | P> t | [0.025 | 0.975] |
|--------------------------------------|--------|---------|-------|-------|---------|--------|
| const | 1.8296 | 0.211 | 8.661 | 0.000 | 1.412 | 2.247 |
| GDP per Capita Average | 0.8945 | 0.230 | 3.893 | 0.000 | 0.440 | 1.349 |
| Social Support Average | 1.0368 | 0.243 | 4.272 | 0.000 | 0.557 | 1.517 |
| Healthy Life Expectancy Average | 1.1017 | 0.358 | 3.076 | 0.003 | 0.393 | 1.810 |
| Generosity Average | 0.7055 | 0.499 | 1.415 | 0.159 | -0.281 | 1.692 |
| Freedom to Make Life Choices Average | 1.3344 | 0.387 | 3.446 | 0.001 | 0.569 | 2.100 |
| Perceptions of Corruption Average | 0.9328 | 0.525 | 1.776 | 0.078 | -0.106 | 1.971 |

```
=====
Omnibus:                10.019      Durbin-Watson:           1.787
Prob(Omnibus):           0.007      Jarque-Bera (JB):        10.131
Skew:                    -0.575     Prob(JB):                 0.00631
Kurtosis:                 3.604     Cond. No.                  27.2
=====
```

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Questions & Data: Cont.

Q3: Which of these measures have the most influence on the happiness score?

- To generate a stronger model, we dropped generosity from the model and ran it again.

OLS Regression Results

| | | | |
|-------------------|------------------|---------------------|----------|
| Dep. Variable: | Score Average | R-squared: | 0.789 |
| Model: | OLS | Adj. R-squared: | 0.781 |
| Method: | Least Squares | F-statistic: | 103.0 |
| Date: | Wed, 22 Jul 2020 | Prob (F-statistic): | 8.09e-45 |
| Time: | 16:23:18 | Log-Likelihood: | -108.16 |
| No. Observations: | 144 | AIC: | 228.3 |
| Df Residuals: | 138 | BIC: | 246.1 |
| Df Model: | 5 | | |
| Covariance Type: | nonrobust | | |

| | coef | std err | t | P> t | [0.025 | 0.975] |
|--------------------------------------|--------|---------|-------|-------|--------|--------|
| const | 1.9187 | 0.202 | 9.481 | 0.000 | 1.519 | 2.319 |
| GDP per Capita Average | 0.8543 | 0.229 | 3.734 | 0.000 | 0.402 | 1.307 |
| Social Support Average | 1.0440 | 0.244 | 4.286 | 0.000 | 0.562 | 1.526 |
| Healthy Life Expectancy Average | 1.0802 | 0.359 | 3.008 | 0.003 | 0.370 | 1.790 |
| Freedom to Make Life Choices Average | 1.4671 | 0.377 | 3.891 | 0.000 | 0.721 | 2.213 |
| Perceptions of Corruption Average | 1.1289 | 0.508 | 2.220 | 0.028 | 0.124 | 2.134 |

| | | | |
|----------------|--------|-------------------|---------|
| Omnibus: | 11.596 | Durbin-Watson: | 1.757 |
| Prob(Omnibus): | 0.003 | Jarque-Bera (JB): | 12.159 |
| Skew: | -0.619 | Prob(JB): | 0.00229 |
| Kurtosis: | 3.703 | Cond. No. | 25.4 |

Discussion:

- We set out to examine the influence of the **measures of social well-being** on **Happiness**.
- We found evidence to support our hypothesis. The higher the perceived absence of corruption, the higher the happiness score.
- The four most influential SWBs were:
 - *Freedom to Make Life Choices*
 - *Perceived Absence of Corruption*
 - *Healthy Life Expectancy*
 - *Social Support*
- GDP per Capita is significant but not as effective at predicting a higher happiness score, so less developed countries have higher Happiness scores.
- Previous studies suggest that economic expansion does not lead to increased happiness at the country level (Easterlin, 1974, Shin, 1980, Rose, 2020).

Post Mortem:

- We started out asking a research question about influence but realized that we were not answering it using a simple linear regression. This required us to revisit the question and expand our skills to run more advanced statistical analyses.
- Limitations:
 - *Additional tests of assumption*
 - *Research individual countries to explore and explain variance across years and measures of SWB*
 - *Explore the variance in the Happiness Scores looking at absence of corruption over time.*



OPEN FLOOR Q & A

Please be nice, our study found we are not that happy...nor generous.



Sources and other docs:

- [Kaggle](#)
- [World Happiness Report](#)
- [Github](#)
- [Write-Up](#)
- [Presentation Timeline](#)

Questions & Data: Cont.

- What are the outliers for the measures of SWB: social support, healthy life expectancy, generosity, freedom to make life choices, GDP per capita, and perception of corruption? How do they compare?
 - *Perception of corruption saw the most outliers at the high end. A high rating is interpreted as less corruption. The mean fell around .12*

