

Adam Vajdak
Patrick Furlong

Final Report

We are combining global datasets from kaggle to try to understand if internet usage has an impact on happiness around the world. The three datasets that we will be combining are the World Happiness Report 2022, Global Internet Usage and lastly Country Continents. The World Happiness Report ranks each country on a “Happiness Score” which is calculated using variables such as GDP, Social Support and life expectancy for example. The Global Internet dataset contains country names along with income per person and internet usage. The last dataset we pulled in for this project simply shows each country and the region of the world that country is located in. This helps us pull in the continents for a more macro analysis by region.

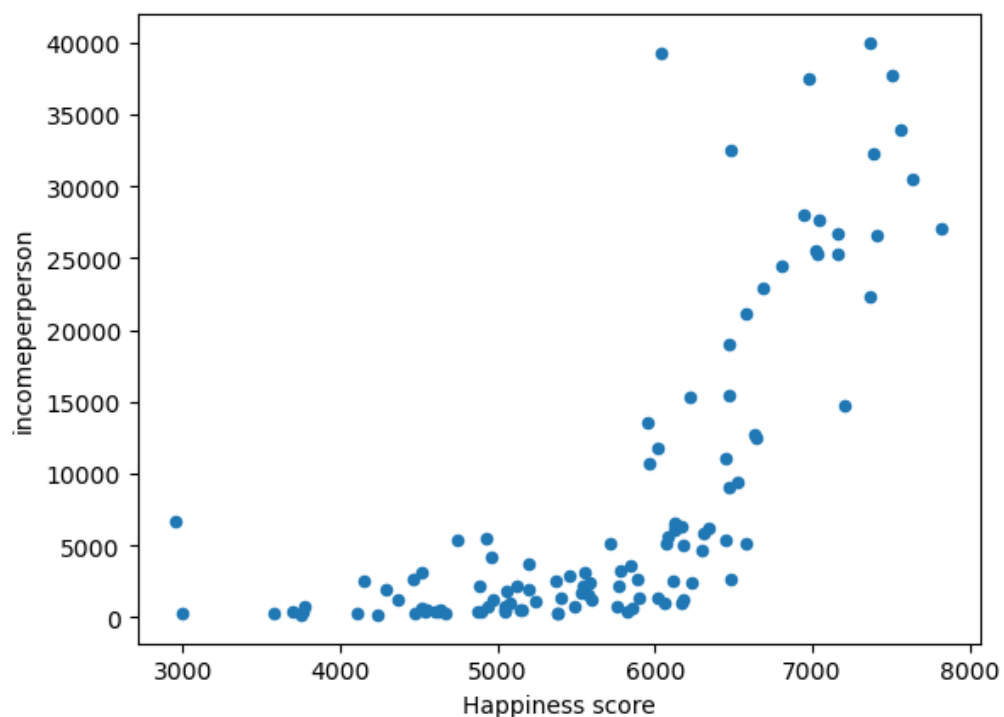
Once we had combined each dataset into a dataframe using the country name as a reference, we set our sights at answering following questions:

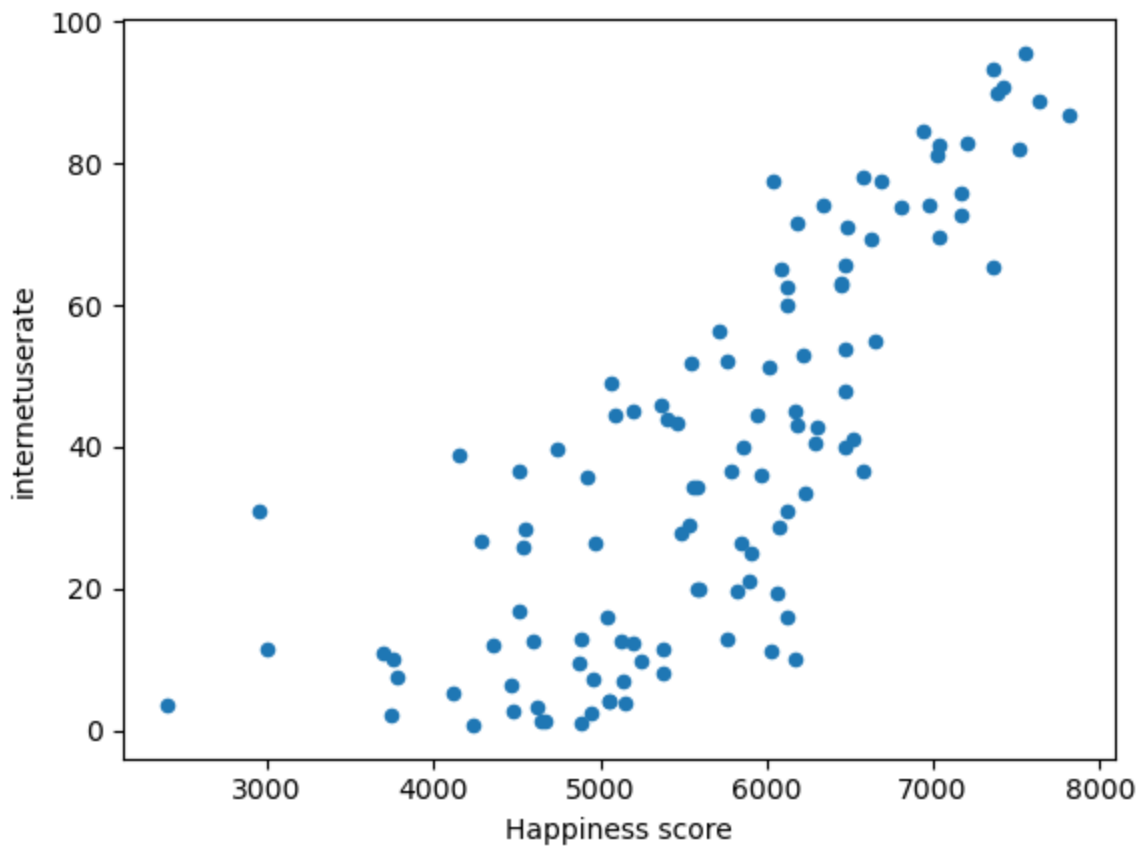
- **What countries are the happiest?- Patrick**
- **What regions are the happiest?-Patrick**
- **What factors contribute to happiness?- Adam**
- **Does Internet usage play a role in a country's happiness score?-Adam**
- **Does personal income play a role in happiness score?-Adam**

First we had to merge the data sets together on country names and remove all the countries that could not be joined. We can see the top 5 and bottom 5 countries below.

| | RANK | Country | Happiness score | Whisker-high | Whisker-low | Dystopia (1.83) + residual | Explained by: 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 | incomeperperson |
|-----|------|-------------|-----------------|--------------|-------------|----------------------------|------------------------------|------------------------------|---------------------------------------|--|--------------------------|---|-----------------|
| 0 | 1 | Finland | 7,821 | 7,886 | 7,756 | 2,518 | 1,892 | 1,258 | 0,775 | 0,736 | 0,109 | 0,534 | 27110.731590 |
| 1 | 2 | Denmark | 7,636 | 7,710 | 7,563 | 2,226 | 1,953 | 1,243 | 0,777 | 0,719 | 0,188 | 0,532 | 30532.277040 |
| 2 | 3 | Iceland | 7,557 | 7,651 | 7,464 | 2,320 | 1,936 | 1,320 | 0,803 | 0,718 | 0,270 | 0,191 | 33945.314420 |
| 3 | 4 | Switzerland | 7,512 | 7,586 | 7,437 | 2,153 | 2,026 | 1,226 | 0,822 | 0,677 | 0,147 | 0,461 | 37662.751250 |
| 4 | 5 | Netherlands | 7,415 | 7,471 | 7,359 | 2,137 | 1,945 | 1,206 | 0,787 | 0,651 | 0,271 | 0,419 | 26551.844240 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 100 | 124 | Namibia | 4,459 | 4,593 | 4,326 | 1,414 | 1,292 | 0,877 | 0,354 | 0,384 | 0,067 | 0,071 | 2667.246710 |
| 101 | 126 | Myanmar | 4,394 | 4,511 | 4,277 | 0,877 | 1,038 | 0,829 | 0,491 | 0,513 | 0,452 | 0,194 | NaN |
| 102 | 127 | Sri Lanka | 4,362 | 4,487 | 4,237 | 0,595 | 1,415 | 0,934 | 0,660 | 0,529 | 0,150 | 0,079 | 1295.742686 |
| 103 | 129 | Egypt | 4,288 | 4,377 | 4,198 | 0,856 | 1,388 | 0,732 | 0,548 | 0,469 | 0,041 | 0,254 | 1975.551906 |
| 104 | 131 | Ethiopia | 4,241 | 4,365 | 4,118 | 1,375 | 0,788 | 0,809 | 0,457 | 0,472 | 0,205 | 0,136 | 220.891248 |

After that some quick scatter plots on the main variables we are looking at which is the correlation between income and internet usage on happiness.





We can see that there is some positive correlation with both but internet usage has a more clear correlation while income only affects happiness after it increases to a certain point. After that we ran some quick correlations on the data to understand what affects happiness and how much. For the two correlation tests we have pearson and kendall correlations respectively.

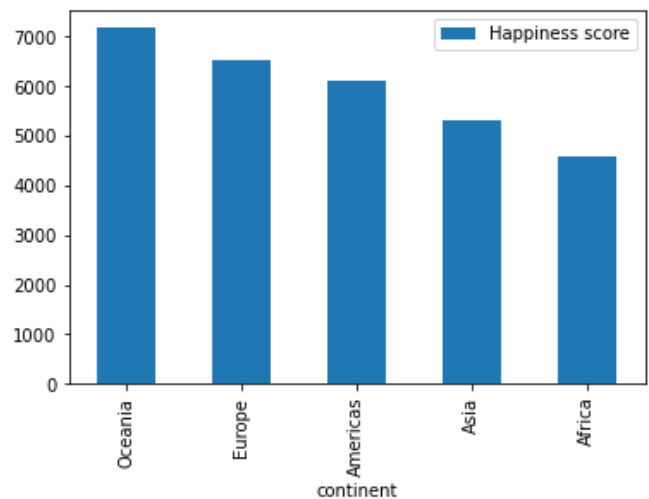
| | Happiness score | | Happiness score |
|--|-----------------|--|-----------------|
| Happiness score | 1.000000 | Happiness score | 1.000000 |
| Dystopia (1.83) + residual | 0.542152 | Dystopia (1.83) + residual | 0.336158 |
| Explained by: GDP per capita | 0.782118 | Explained by: GDP per capita | 0.669770 |
| Explained by: Social support | 0.822994 | Explained by: Social support | 0.654501 |
| Explained by: Healthy life expectancy | 0.754245 | Explained by: Healthy life expectancy | 0.574592 |
| Explained by: Freedom to make life choices | 0.641987 | Explained by: Freedom to make life choices | 0.451657 |
| Explained by: Generosity | 0.035857 | Explained by: Generosity | 0.012689 |
| Explained by: Perceptions of corruption | 0.521924 | Explained by: Perceptions of corruption | 0.293336 |
| incomeperperson | 0.723670 | incomeperperson | 0.620722 |
| internetuserate | 0.786039 | internetuserate | 0.610922 |
| urbanrate | 0.620843 | urbanrate | 0.462114 |

We can see that the factor with the highest correlation was one we did not even look at, Social Support. Which is followed by GDP ,income, and internet user rate. After that we took a look at the summary statistics for each field we did correlation on.

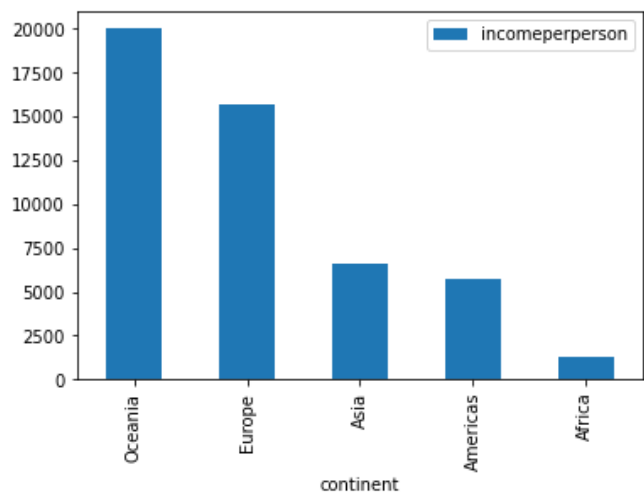
| | Happiness score | Dystopia (1.83) + residual | Explained by: GDP per capita | Explained by: Social support | Explained by: Healthy life expectancy | Explained by: Freedom to make life choices | Explained by: Generosity | Explained by: Perceptions of corruption | incomeperperson | internetuserate | urbanrate |
|-------|-----------------|----------------------------|------------------------------|------------------------------|---------------------------------------|--|--------------------------|---|-----------------|-----------------|------------|
| count | 115.000000 | 115.000000 | 115.000000 | 115.000000 | 115.000000 | 115.000000 | 115.000000 | 115.000000 | 113.000000 | 113.000000 | 115.000000 |
| mean | 5642.895652 | 1841.408696 | 1438.921739 | 920.634783 | 605.817391 | 531.513043 | 150.208696 | 154.513043 | 8167.124111 | 38.665397 | 58.841043 |
| std | 1092.941391 | 485.078415 | 405.610157 | 280.096360 | 159.511732 | 142.081665 | 82.516824 | 129.843182 | 10830.789194 | 27.424695 | 21.942348 |
| min | 2404.000000 | 216.000000 | 0.000000 | 0.000000 | 191.000000 | 0.000000 | 0.000000 | 0.000000 | 184.141797 | 0.749996 | 12.980000 |
| 25% | 4933.000000 | 1609.500000 | 1139.500000 | 732.000000 | 509.500000 | 460.000000 | 89.000000 | 65.000000 | 772.933345 | 12.500073 | 41.880000 |
| 50% | 5768.000000 | 1895.000000 | 1452.000000 | 1005.000000 | 633.000000 | 555.000000 | 138.000000 | 115.000000 | 2667.246710 | 36.499115 | 61.320000 |
| 75% | 6446.500000 | 2145.000000 | 1798.500000 | 1136.500000 | 731.500000 | 649.000000 | 202.500000 | 202.000000 | 11066.784140 | 62.471230 | 74.070000 |
| max | 7821.000000 | 2790.000000 | 2149.000000 | 1320.000000 | 866.000000 | 740.000000 | 468.000000 | 587.000000 | 39972.352770 | 95.638113 | 100.000000 |

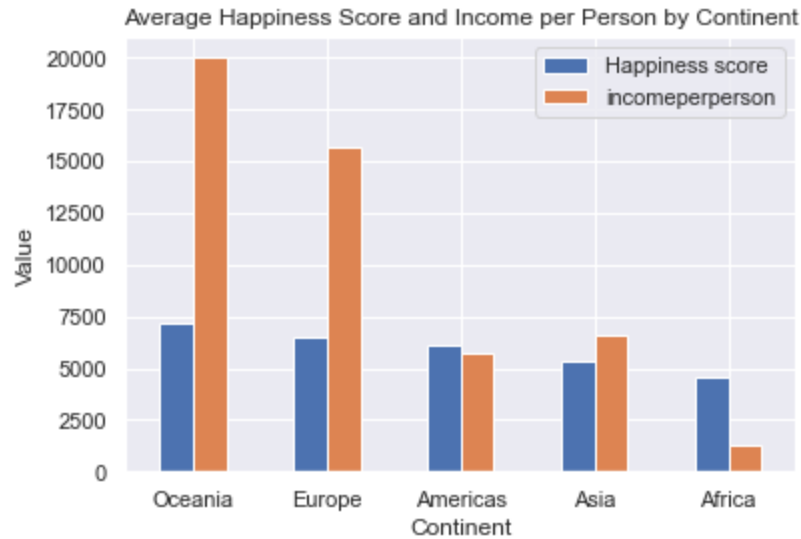
Next, we decided to group each county by continent. The average Happiness Score and average income per person was calculated per each continent and displayed as the two bar charts below:

| | continent | Happiness score |
|---|-----------|-----------------|
| 4 | Oceania | 7181.000000 |
| 3 | Europe | 6531.939394 |
| 1 | Americas | 6107.058824 |
| 2 | Asia | 5352.529412 |
| 0 | Africa | 4519.361111 |

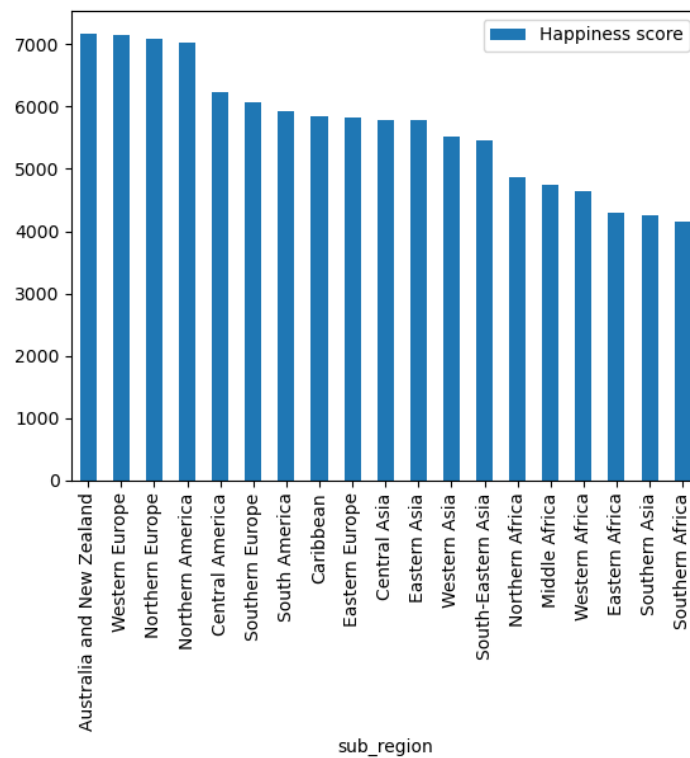


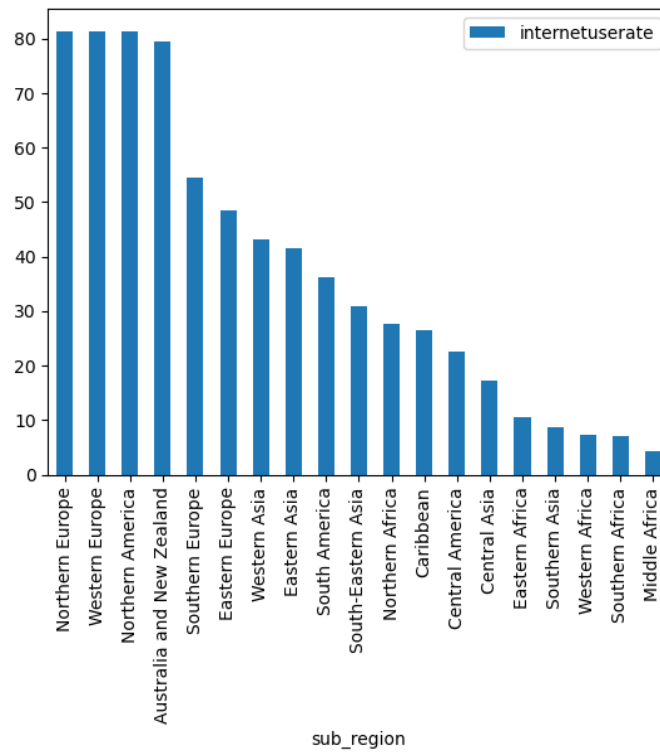
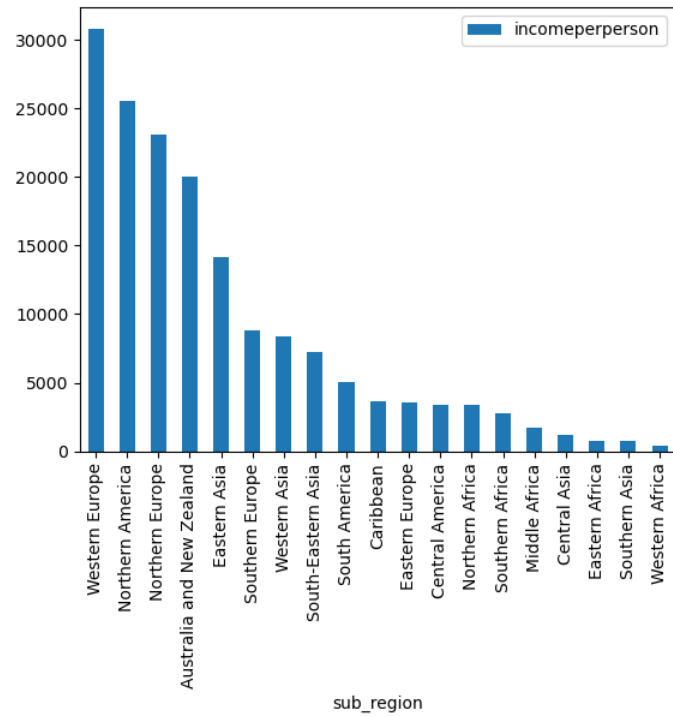
| | continent | incomeperperson |
|---|-----------|-----------------|
| 4 | Oceania | 20014.074995 |
| 3 | Europe | 16439.623446 |
| 2 | Asia | 6343.723863 |
| 1 | Americas | 5520.398253 |
| 0 | Africa | 1292.120641 |





These graphs show us that there is some correlation but not a strong one which we also found with our correlation testing. Since continents are very large, there can be high disparities between the populations on each continent. Because of this we then decided to take a closer look at each sub region to get a more acute view of regional Happiness Score. Below, we have the average Happiness Score by sub region, average income per person by sub region and average internet usage per person by sub region:





In sub regions northern america, australia/new zealand, north and western Europe are the top 4 in the 3 statistics are keeping track of.

Summary:

Happiness is not just determined by one factor, although some are more important than others. The happiest countries are Finland, Denmark, Iceland, Switzerland, and the Netherlands. The factor that supports happiness the most in countries is Social Support. Which is followed by GDP, income, and internet user rate. The Internet and income do have a correlation with Happiness although it is not as strong as we thought it does seem to affect it to a degree.

Sources:

Global Internet Usage: <https://www.kaggle.com/datasets/sansuthi/gapminder-internet>

World Happiness Report: <https://www.kaggle.com/datasets/mathurinache/world-happiness-report>

Countries to continent: <https://www.kaggle.com/datasets/statchaitya/country-to-continent>