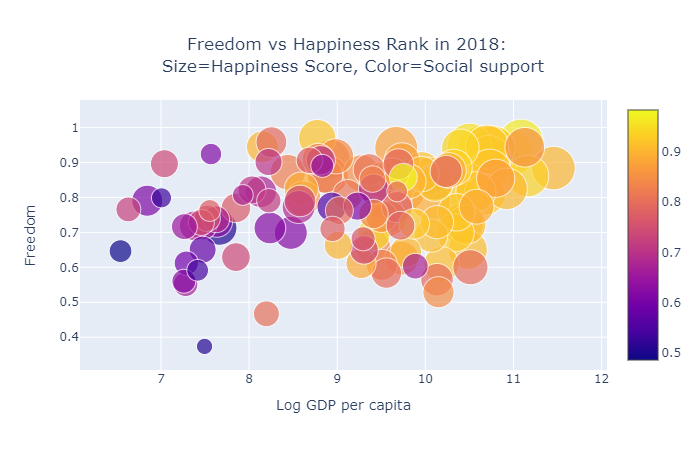
The Happiness Hunters

Vaishall Pradeepkumar, David Nassif, Solicia Xu

**On the Subject of Happiness and its Correlation to Various Factors, by Country and Year**

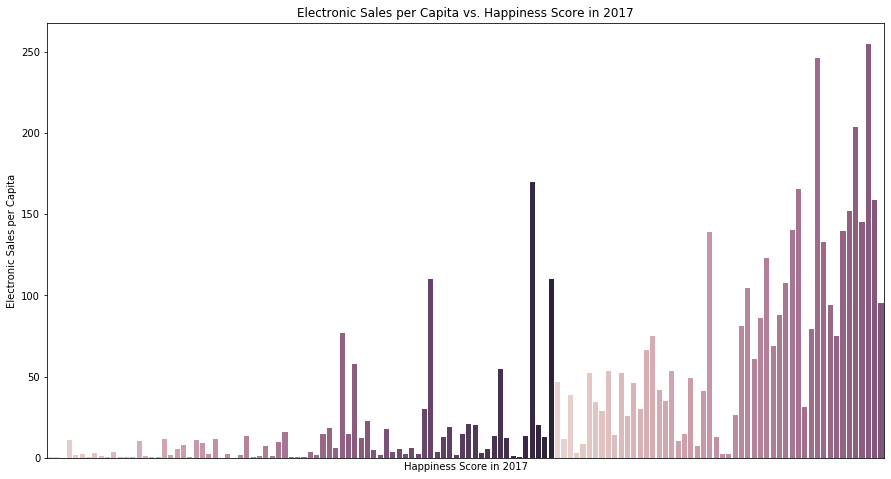
During our team’s investigation, we sought to answer four questions:

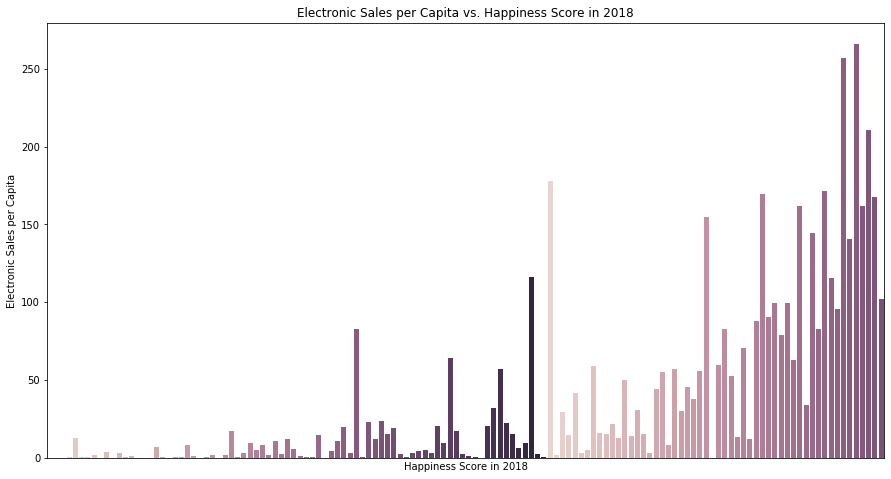
1. **Does freedom or money correlate better with happiness?**

A statistical analysis showed that Log GDP per capita correlates more closely with happiness (p-value could not be viewed as it was too small and displayed as 0, we concluded that the correlation could not be actually 0 after a look at a graph showed that this could not be the case as our data was not perfectly linear) compared to freedom and happiness (p-value of ). Interestingly, the correlation between Log GDP per capita and freedom was also high, although not as high as the two correlations mentioned previously (p-value of ). [](http://vaishall.github.io/)

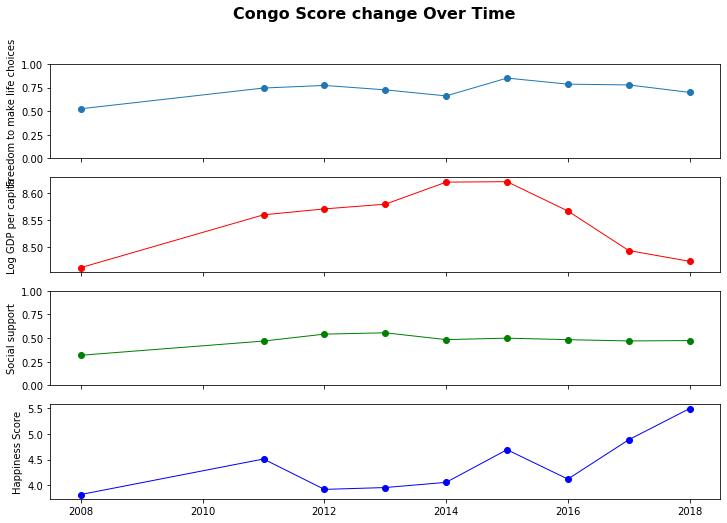
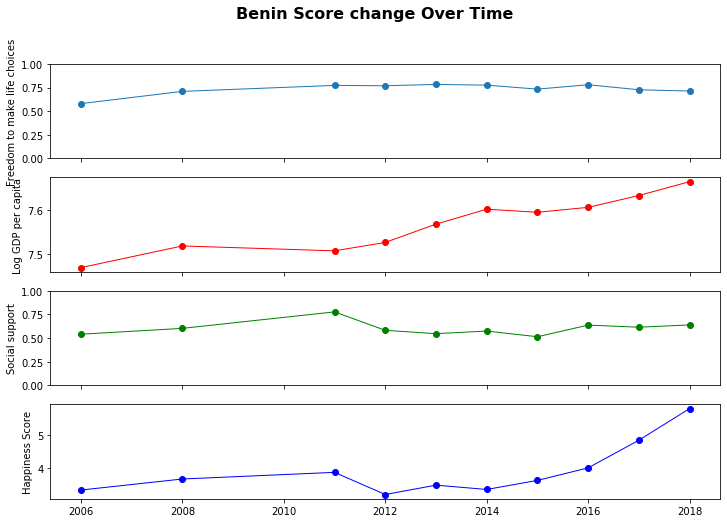
For an interactive version, click the image or go to <http://vaishall.github.io>

1. **Is Consumer Electronics Revenue related to a country’s happiness?**

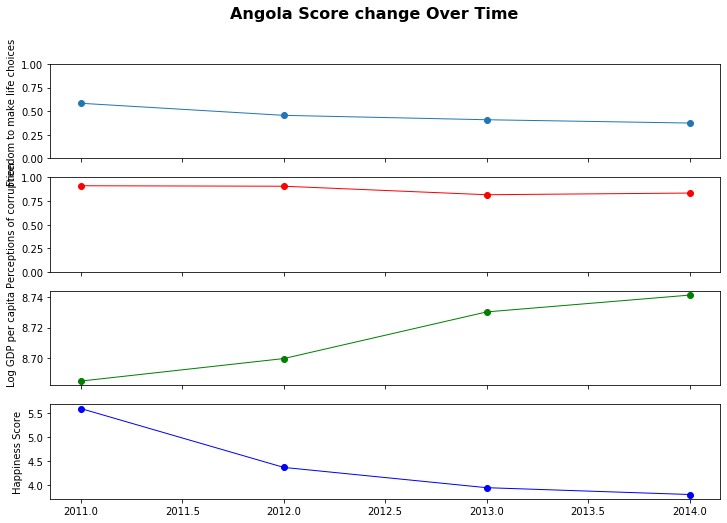
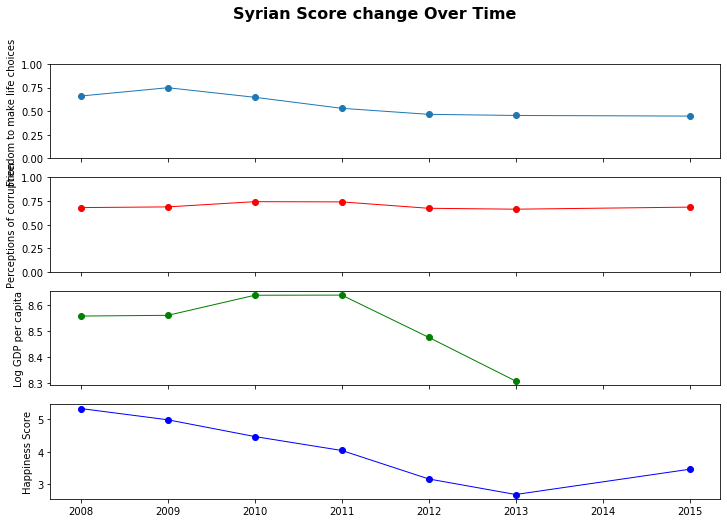
A statistical analysis showed that the correlation between Electronics Revenue per capita and happiness to be strong as well (p-value of ). This was using linear regression, however using non-linear regression would have likely yielded even stronger results (this analyst would suggest an exponential model, with likely a few obvious outliers. Something for further analysis with scipy, perhaps). 



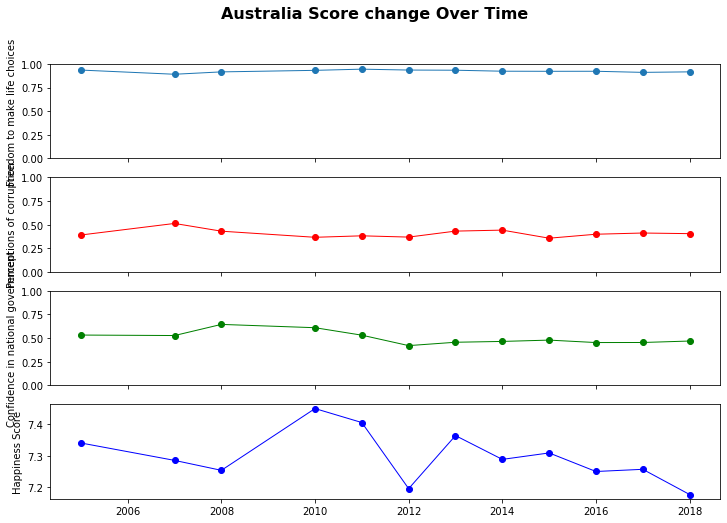
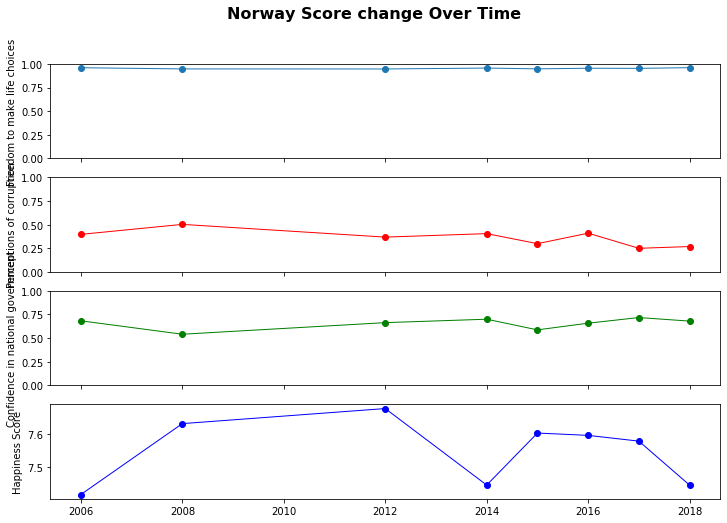
1. **For countries whose happiness data changes the most (or least), what factors influence those changes?**

First, let’s take a look at the countries who increased the most:

Now the countries that decreased the most:



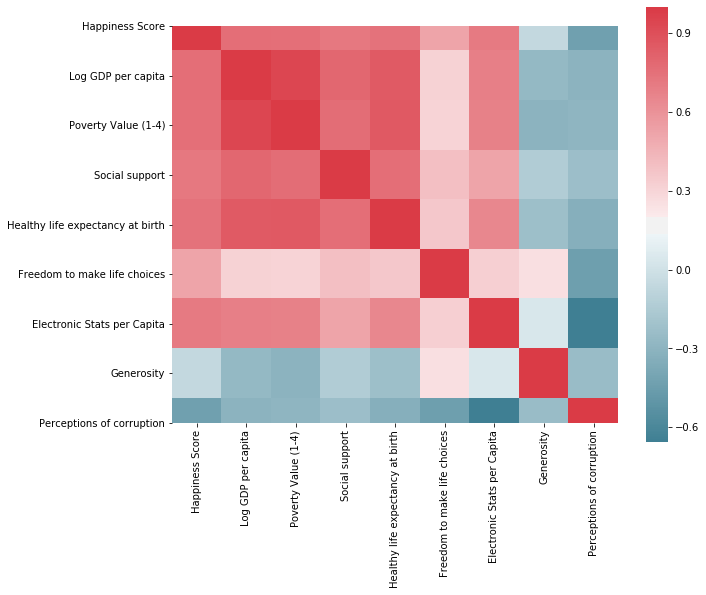
And the countries that had the least change:



It seemed that developed nations with stable economies seemed to have the least percent change in happiness over the years studied. Countries that had the highest percent increase in happiness seemed to be developing countries whose economies started to take off, whereas those that had the highest percent decrease seemed to be those plunged into war, or with increasing levels of perceived corruption in government.

1. **How closely are different variables related to happiness correlated with each other?**

It would have been tedious to calculate the correlation between every variable individually, although it would have been possible to loop over them and save their p-values to a matrix. Thanks to the following graph, we were able to determine that many variables are correlated to each other. Happiness, Log GDP per capita, poverty value, social support, life expectancy at birth, perceived freedom, and Electronic Sales per capita were all directly correlated, and they were inversely correlated with perceived generosity, and perceptions of corruption.



**Notes:**

In addition to the python modules we learned in class such as pandas, numpy, matplotlib, and scipy.stats, we also used a few tools that we had to learn ourselves. These included, but were not limited to:

1. lxml: for web scraping the 2019 Consumer Electronics Revenue data
2. Beautiful Soup: for the 2017 and 2018 data when it couldn’t be done with lxml
3. Seaborn: for Electronic Sales charts and Variable Correlation charts (Questions 2 and 4)
4. Plotly: for our bubble plot of Freedom vs Log GDP per capita vs Happiness vs Social Support (Question 1)
5. html: for a website to host our interactive bubble plot for our presentation

We could have proceeded without using these, but we wanted to go above and beyond and make our project as advanced as possible, even if it meant teaching ourselves all sorts of new things!

**Further Research:**

While our investigation gave us a few answers, it also raised some questions that, given more time, we might have performed research on:

* Why are there some countries that have high Electronic Revenue Sales but relatively low happiness?
* Use machine learning techniques to predict happiness scores for the future.
* Find further supporting data sets for Electronic Sales i.e. availability of electricity, internet, etc. to perform research on.
* Are there alternatives to the World Happiness Report whose findings corroborate those found in the World Happiness Report?
* Why does generosity correlate negatively with happiness?

