Global Happiness: It's all in the Data

Presented by: Elisa Gipe, Patrick Plaisted, Jay Sueno, Ryan Little, Jason Gabunilas

Project Summary:

The pursuit of happiness is a common human interest. Though it is fairly easy to comprehend *what* happiness is, it is often difficult to define *why* we experience it. In this investigation we break down national Happiness Scores from the World Happiness Report from the calendar years 2015 through 2017 and compare it to a variety of metrics, including but not limited to GDP, life expectancy, perception of freedom, and degree of air pollution. As a fun exercise, we also investigate whether the national viewership of the film 50 Shades of Gray is associated with national happiness.

Questions this analysis seeks to explore:

- What are some key contributing metrics to happiness and how does each influence the happiness of a population on a country-level?
- What are some of the strongest metrics that correlate with the happiness of a population overall?
- Which countries are the happiest?
- Does the average air quality (as defined by particular matter concentration) in a given country correlate with happiness? Why or why not?
- Is there a correlation between the number of deaths due to air pollution and happiness? Why or why not?
- Is there a correlation between particulate matter concentration and number of deaths due to air pollution? Why or why not?
- 50 Shades of Grey is considered to be one of the worst movies from the past decade. Do countries that have the lowest (per capita?) sales for movie tickets for 50 Shades of Gray have a higher level of happiness than countries where more people patronized the film in theaters? Does a bad movie affect Global Happiness? (control)

Sources of Data:

- World Happiness Report up to 2020 from Kaggle: https://www.kaggle.com/mathurinache/world-happiness-report?select=2020.csv
- Our World in Data: https://ourworldindata.org/air-pollution
- Pollution data (source: https://ourworldindata.org/air-pollution) https://docs.google.com/spreadsheets/d/1cuZsAMPJmQi6iS-8LSWHCjrs2ruc087xuDYLceTwgR8/edit?usp=sharing
- Movie ticket sales: Box Office Mojo (https://www.boxofficemojo.com/releasegroup/gr4024783365/)

Overview of Data Sets:

Happiness Data Set

The World Happiness Report is produced using the Gallup World Poll (GWP) and contains scores for 153 countries along with the factors used to explain the score. The Gallup World Poll uses the Cantrill Ladder, which asks respondents to think of a step with the most excellent conceivable life for them being a 10 and the most exceedingly bad conceivable life being a 0 and to rate their claim current lives on that scale.Each of the factors are scored relative to other countries in the study. The factors included are: GDP per capita, Healthy Life Expectancy, Social Support, Freedom to make life choices, Generosity, Corruption Perception.

Pollution Data Set

The air pollution datasets provide the levels of air pollutant concentration (in ug/m^3) and number of deaths per 100,000 people due to air pollution annually for over 250 countries beginning in 1990. In some years data for one or both of these metrics was not available. The analysis presented here focused on the years 2015 and 2016.

50 Shades Data Set

The data for 50 Shades of Grey and 50 Shades Darker were taken from Box Office Mojo. The data shows sales data by country, region, release date, opening day sales, and gross sales. 50 Shades of Grey was released first in 2015 so we compare 50 Shades of Grey to the 2015 world happiness data. 50 Shades Darker was released in 2017 so we compare that sales data with the 2017 world happiness report.

What are some key contributing metrics to happiness and how does each influence the happiness of a population on a country-level?

From the Gallup World Poll Data, the explored influencing metrics collected are: GDP per capita, Healthy Life Expectancy, Social Support, Freedom to make life choice, Generosity, and Corruption Perception. With those metrics in mind, we did a series of regression tests to see how each of these metrics compared to the overall ladder score of the country. As described below, the metrics that presented the strongest correlation were the size of family and the overall GDP. The weakest correlations were the generosity and corruption rankings.

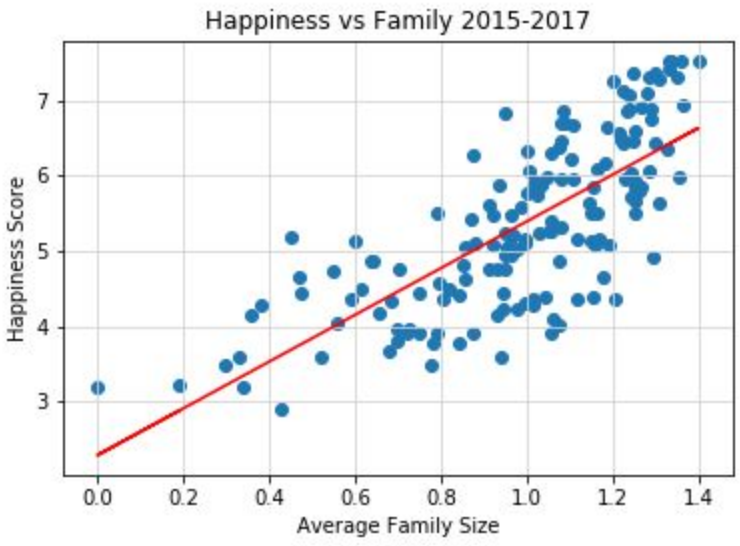


Figure 1: Overall Happiness score compared to average family size.

Which countries are the happiest?

Over the 3 years analyzed, 2015-2017, the following countries, descending from the highest score, had the highest average overall happiness score: Switzerland, Iceland, Denmark, Norway, Canada, Finland. The following countries, ascending from the lowest average overall happiness score, are as follows: Chad, Guinea, Ivory Coast, Burkino Faso, Afghanistan, Rwanda.

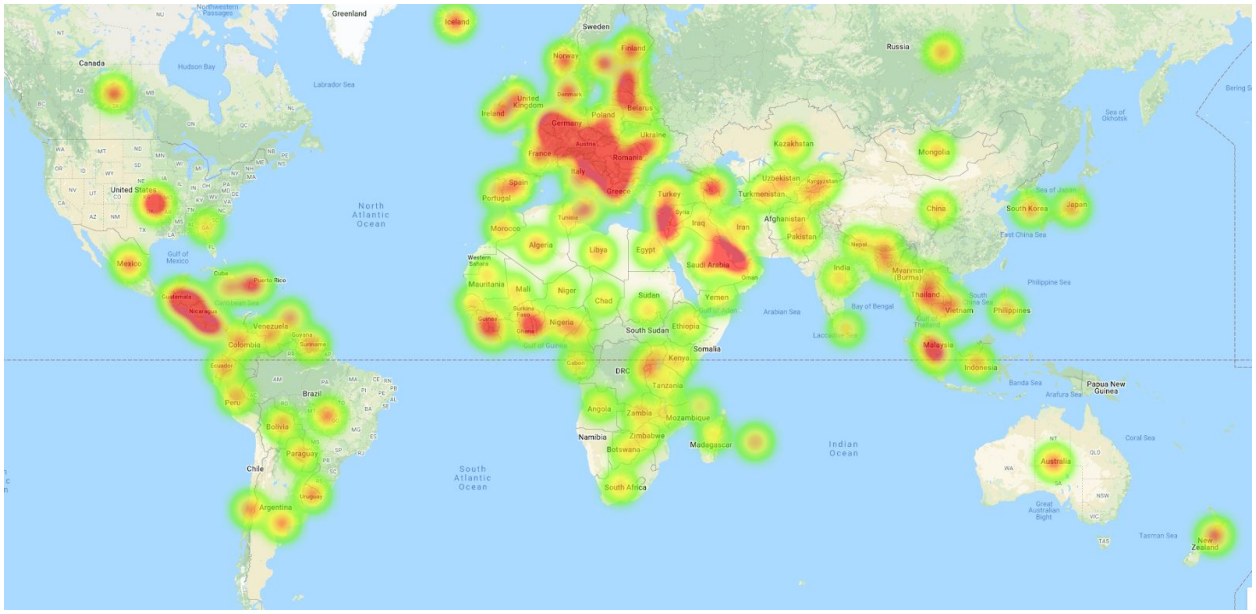
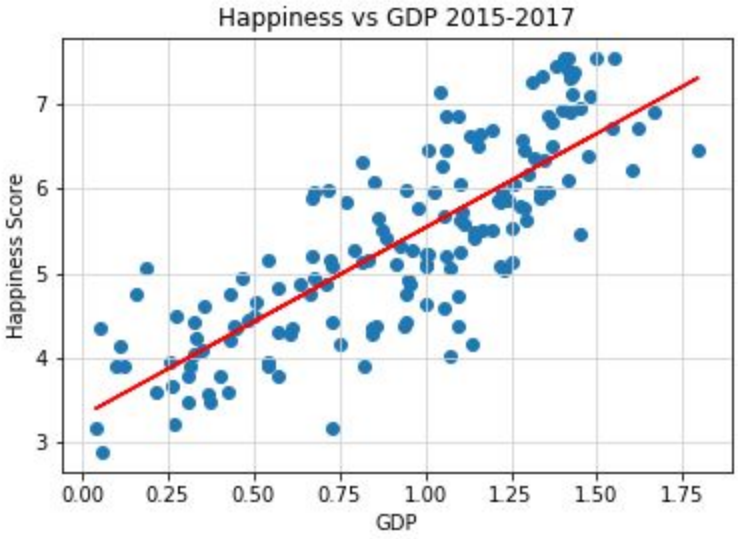
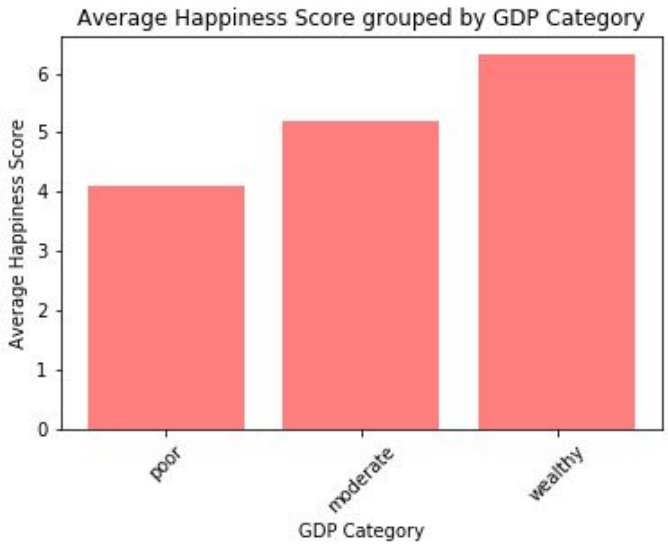


Figure 2: Happiness Data ranked by Country 2015-2017

What are some of the strongest metrics that correlate with the happiness of a population overall?

Overall, the data seems to suggest that the metrics collected that have the highest correlation probability with happiness are GDP and Family Size, whereas those that represented the weakest correlations with happiness are corruption and generosity.





Figures 3 & 4: Happiness metrics compared to country GDP.

What was the general overview of global air pollution in 2016?

Our source data from Our World in Data quantifies a country's air pollution as the average concentration of particulate matter of 2.5 microns in diameter or smaller (PM2.5), provided in units of micrograms per cubic meter (ug/m3). Using the Google Places and Google Maps APIs, we created heatmaps representing the concentrations of air pollutants across the globe.

Global view of air pollution in 2016

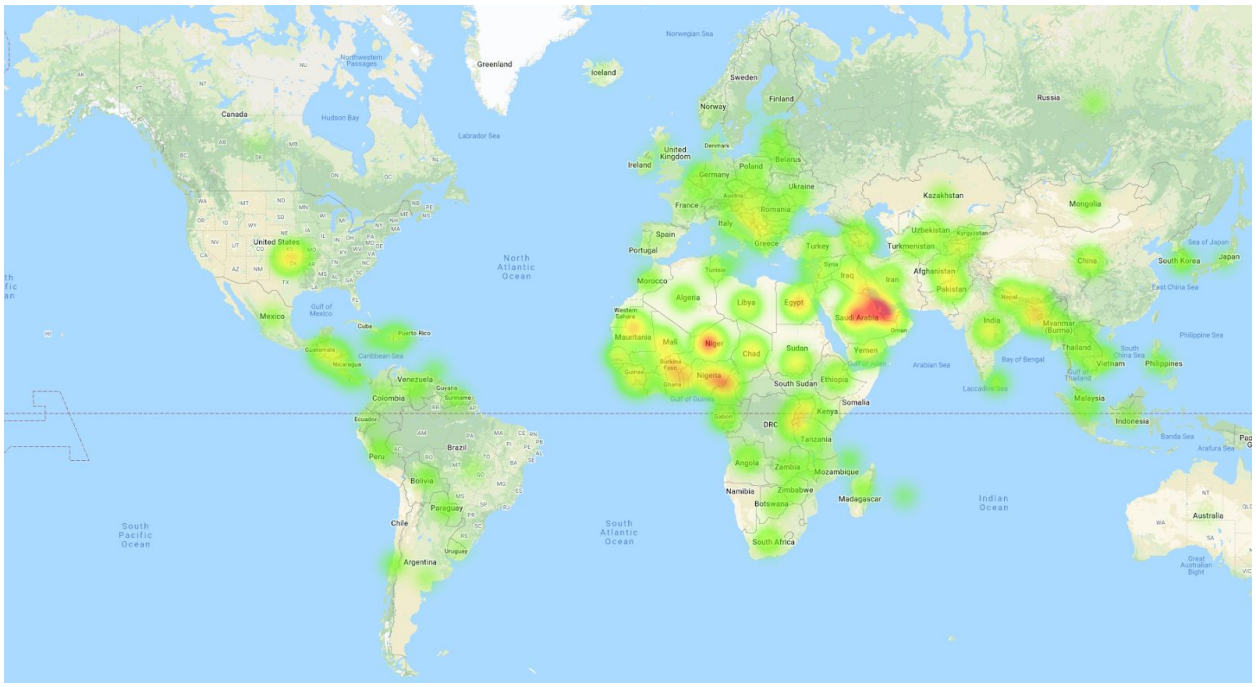


Figure 5: Average particulate matter concentrations by country, 2016

As the visualization shows, concentrations of air pollutants were highest in the Saharan and central regions of Africa as well as in the Middle East, followed by countries in South Asia. Per our source, PM2.5 tends to be higher in drier climates due to the higher presence of dust and sand particulates. Air pollution in the Middle East may be explained in part by crude oil mining and refining. Note that this heatmap displays the average PM2.5 values for each country and treats each country as a point location on the map. It does not show where pollutants are localized within each country.

Is there a relationship between the happiness of a country and the annual number of deaths within that country?

Pollution causes a notable amount of deaths every year. In this study we want to see if air pollution deaths have a correlation to happiness within a country. To do this we plotted the air pollution deaths per 100,000 people in a country versus how that country ranks for happiness.

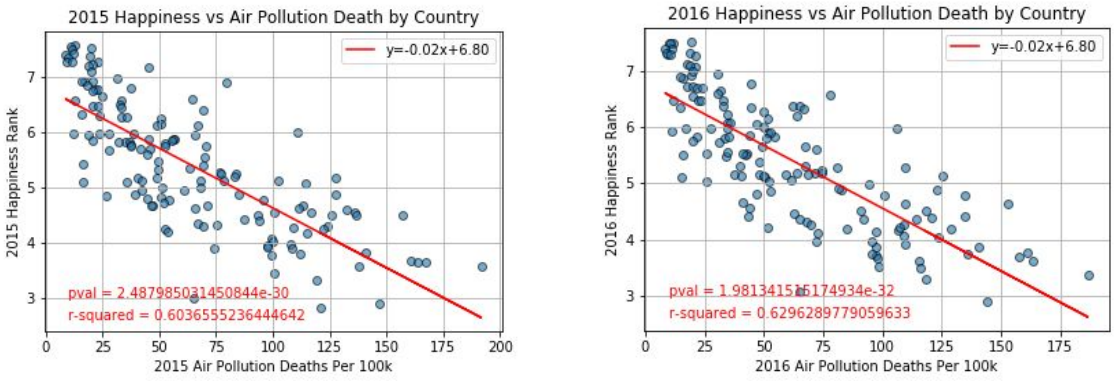


Figure 6: Country and Air Pollution Deaths

As we can see in the figures above, there is a strong correlation between happiness and deaths due to air pollution. The both the P-values are well below the 0.01. This means that we can REJECT the Null Hypothesis that states that there is NO correlation between happiness and deaths due to pollution. In other words, we can be fairly confident that there IS a correlation.

The R-squared value's of 0.60 in 2015 and 0.63 suggests that there is low variability of the datasets to the regression lines. This means that there is a high likelihood that the data models significant correlation between happiness and deaths due to air pollution. In layman's terms, there is probably a likelihood that people are happier in countries where there is less death due to air pollution vs those countries that have higher pollution death rates.

We conclude that if there is a relatively high level of death due to pollution, then there will be a downward pressure on happiness. As we will see there are other elements of pollution that factor into happiness. Other questions that we can study to understand this relationship is the quality of healthcare and level of investment within a country.

Are residents of countries with higher levels of air pollution less happy than in countries with cleaner air?

Breathing is a necessary function, and it is not unreasonable to postulate that one's happiness could be influenced at least in part by the air one is exposed to day after day. To examine this question, we graphed the average air pollutant concentrations against the happiness scores for countries where both of these data points were available. This analysis was performed with data from 2015 and 2016.

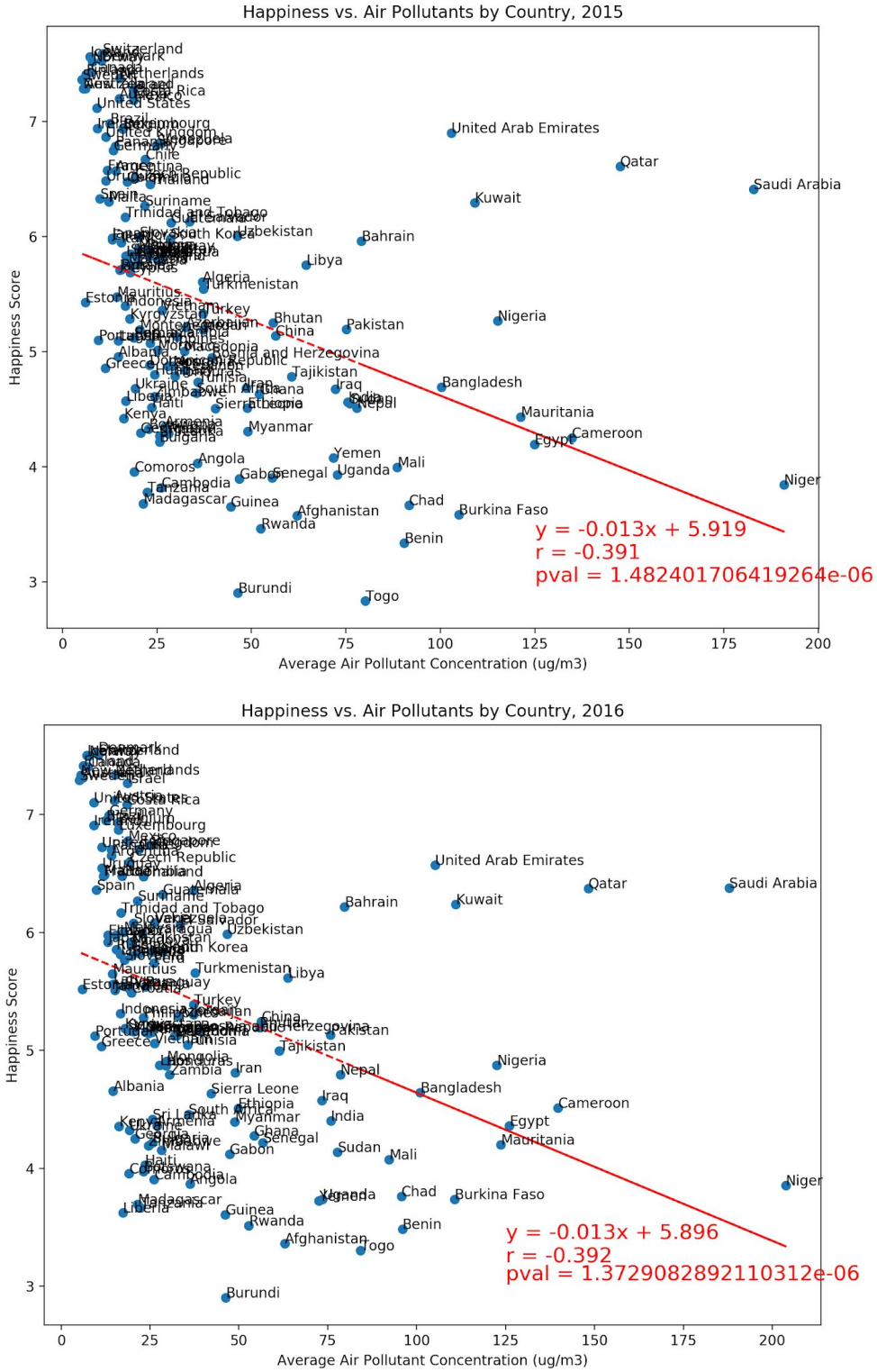


Figure 7: Scatterplot and linear regression analysis of average air pollutant concentration and happiness score in 2015 (top) and 2016 (bottom)

As one may have expected, this analysis suggests that there is a negative correlation between air pollutant concentration and happiness score, with the general trend being that countries having the lowest levels of air pollution also having the highest happiness scores and vice versa. While the correlation is moderate at best, it was apparent and consistent for both the 2015 and 2016 data sets. Interestingly, several countries with relatively high levels of air pollution were also some of the happiest, such as Qatar, Kuwait and Saudi Arabia. These countries benefit from large oil and natural gas reserves that, while possibly contributing to air pollution, also fuel a high-income economy, providing a degree of economic prosperity for its residents that may contribute to increased happiness. Conversely, there are also multiple countries that were quite unhappy in spite of having low levels of air pollution, including Afghanistan, Burundi, and Rwanda. It is likely that these countries suffer from other ailments that impact happiness scores, some possible examples being poverty, war, or famine. On a more positive note, the greatest clustering of countries is found in the top-left corner of the plot, representing low pollutant concentrations and high happiness.

Do countries with greater air pollution also experience higher death rates due to polluted air?

Per our data source, small particulate matter PM2.5 can lead to respiratory issues and other pollution-related diseases. With this in mind, it absolutely would stand to reason that with greater levels air pollution comes higher death rates due to air pollution. We interrogated this idea by plotting air pollutant concentrations against death rates from outdoor air pollution per 100,000 residents for each country with available data in 2015 and 2016.

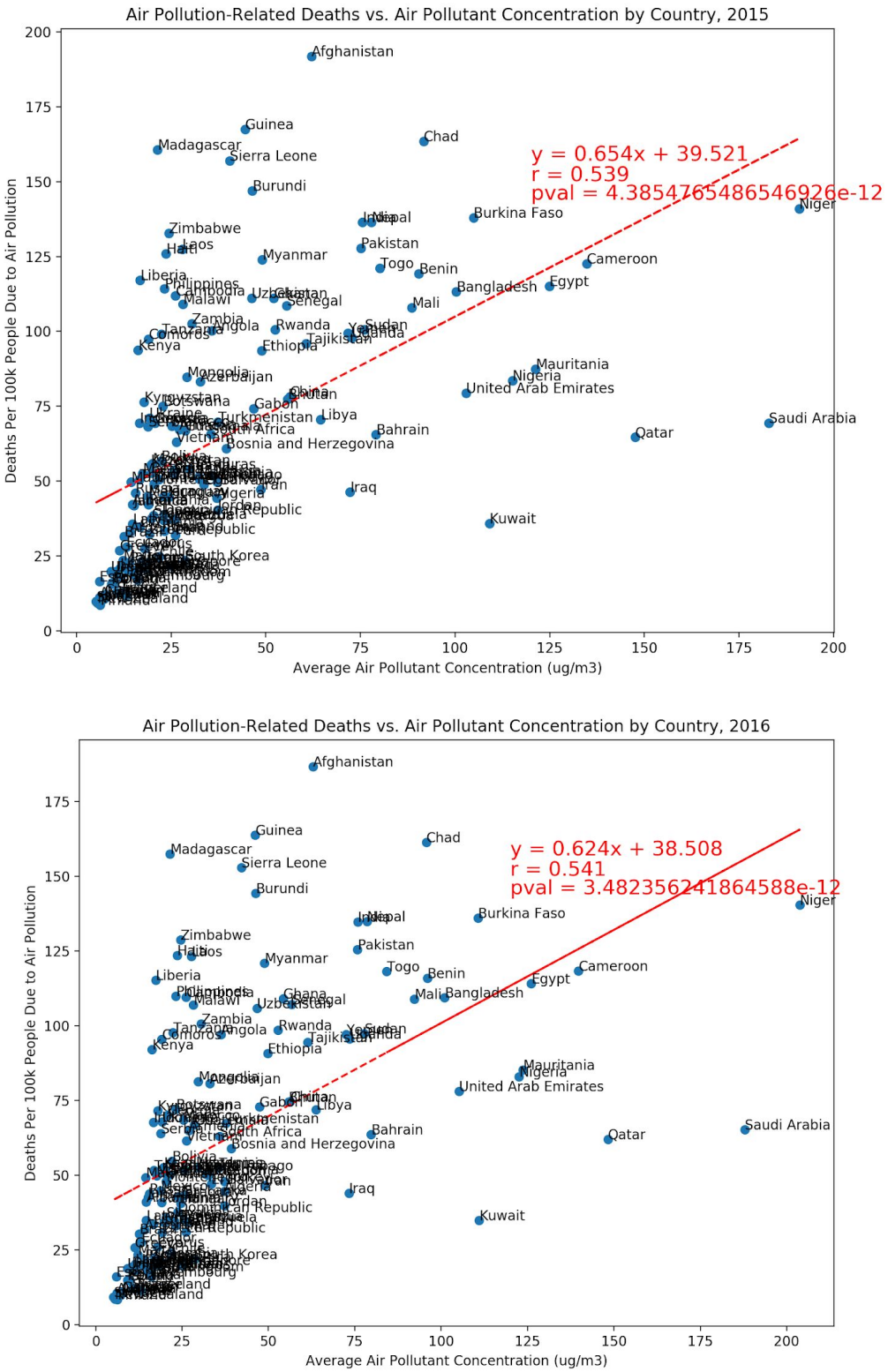


Figure 8: Scatterplot and linear regression analysis of average air pollutant concentration and deaths per 100,000 people in 2015 (top) and 2016 (bottom)

To no surprise, this analysis shows moderate positive correlation between air pollutant concentration and air pollution-related death rates. As expected, death rates due to air pollution are very low in countries that have low levels of air pollution, and fortunately this appears to be the case for many if not most countries in this dataset. Additionally, countries that were happy in spite of high levels of air pollution also seem to also have low relatively low death rates associated with air pollution. Residents of these wealthy countries may spend more time indoors where they can avoid air contaminants, and the healthcare systems of these countries may also be well-developed to effectively treat pollution-related illnesses when they do arise. Perhaps the most surprising revelation from this analysis is that several countries with low levels of air pollution, such as Madagascar, Burundi, Zimbabwe, and Haiti, nevertheless experience high death rates due to air pollution. These deaths may be reflective of a deteriorated state of general health of the population for any number of reasons (malnourishment, other diseases, etc.) and healthcare systems that are unable to provide adequate care.

Does Happiness Correlate with Ticket Sales to the Movie 50 Shades of Grey?

Happiness does not correlate to the ticket sales of the movie 50 Shades of Grey. The United States is the outlier with the most sales - because the movie was produced in the United States, the movie takes place in the United States, and there are more people in the United States, this makes the United States a likely outlier.

Only the United States and Philippines sold more than half of gross sales on opening day. The United States is 15th in world happiness rank, and Philippines is 90th. There is not a correlation between sales as a dollar amount in sales to happiness, or percentage of total sales made on opening day to happiness. Countries that have lower sales for 50 Shades of Grey do not have a higher level of happiness than countries.

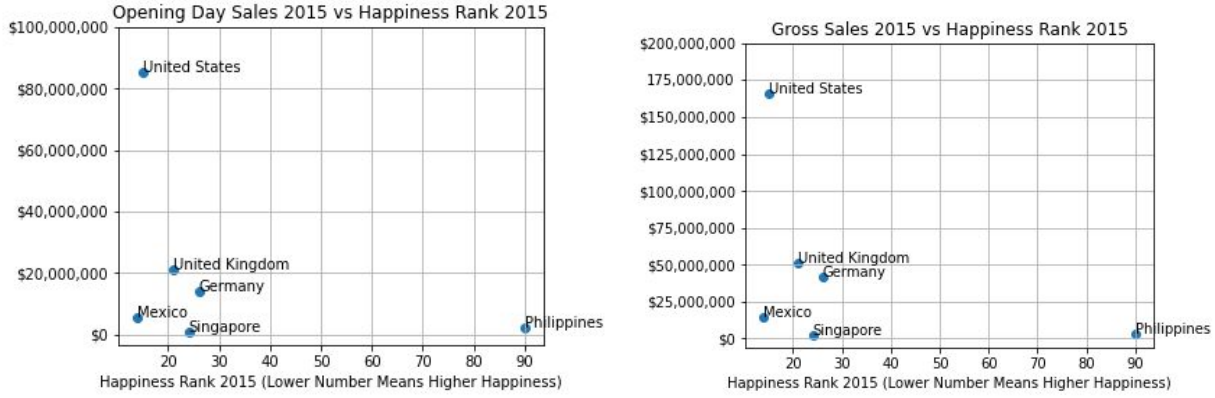


Figure 9: Scatterplot and linear regression analysis of average air pollutant concentration

Does Happiness Correlate with Ticket Sales to 50 Shades Darker, the sequel to 50 Shades of Grey?

Happiness does correlate to ticket sales with the sample of countries pulled from the world happiness report population. An increase in happiness rank was found with countries that decreased their purchase of opening day movie tickets to the sequel of 50 Shades of Grey, 50 Shades Darker.

Germany sold 8% less opening day movie tickets and increased 10 spots in world happiness. The Philippines sold 4% less opening day movie tickets and increased 18 spots in world happiness. The United Kingdom sold 8% less opening day movie tickets and increased 2 spots in world happiness. The United States sold 11% less opening day movie tickets and increased 1 spot in world happiness. Mexico sold 3% more movie tickets on opening day and decreased 11 spots in world happiness. Singapore sold 1% more tickets on opening day and decreased 2 spots in world happiness. What does this mean? This data points to the conclusion that when you see a bad movie on opening day, if you don't see the sequel you will be happier.

Country	Happiness Rank 2015	Gross Sales 2015	Opening Day Sales 2015	% Opening Day Sales to Gross Sales 2015	Happiness Rank 2017	Gross Sales 2017	Opening Sales 2017	% Opening Day Sales to Gross Sales 2017
Germany	26	41430653	14135994	34	16	32709880	8529213	26
Mexico	14	14148888	5495154	39	25	6534100	2692197	42
Philippines	90	3565511	2047933	57	72	3399902	1803492	53
Singapore	24	1837395	641621	35	26	1031387	369593	36
United Kingdom	21	51603774	20876148	40	19	28583660	9241277	32
United States	15	166167230	85171450	51	14	114581250	46607250	40