

EXPLORATORY VISUAL ANALYSIS :

HAPPINESS SCORES ACROSS THE WORLD

Data Visualization Project

INTRODUCTION

The purpose of this project is to visualize happiness scores around the world and identify additional factors that contribute to the changing trends of happiness. Here we try to weigh several factors like economic growth, family structures, access to healthcare, etc. that affect the happiness scores of various countries. This project encompasses a wide range of potential users from official institutions to the general public. To understand if factors such as economy, government trust, pandemic, etc. have a significant impact on the happiness score of a country, we will be performing an in-depth visual analysis using Tableau. We will explore our hypothesis further by creating calculated fields and visualizing insights using several different types of visualizations such as boxplots, scatterplots, correlation plots, heatmaps, etc.

$$Happiness\ Score = \beta_1 Economy\ (GDP\ Per\ Capita) + \beta_2 Family + \beta_3 Health\ (Life\ expectancy) + \beta_4 Freedom + \beta_5 Generosity + \beta_6 Trust\ (Government\ Corruption) + \beta_7 Dystopia_Residual + u_i$$

HAPPINESS AND FREEDOM (COVID-19)

As more and more countries are canceling travel restrictions and face mask requirements, the world is gradually recovering from Covid and trying to resume normal life as before. When the outbreak of Covid first struck the eastern countries in 2019, quarantine policy was soon carried out worldwide. Our team was intrigued to know whether quarantine policy was closely associated with happiness scores and to do that, we used one of the variables -- freedom, to assess quarantine policy's impacts.

H₀: Covid's quarantine policy has no significant impact on happiness

[There is no statistically significant relationship between happiness and freedom – Covid's quarantine policy]

H₁: Covid's quarantine policy has a significant impact on happiness

[There is a statistically significant relationship between happiness and freedom – Covid's quarantine policy]

We first plotted each continent's freedom and happiness scores from 2018 to 2020 to see if there were any apparent differences (Appendix 1.1), and we noticed that all continents' freedom scores dropped; especially Europe, Asia, Africa, and South America had significant drops. We then used a world map to further examine the changes in each country's freedom (Appendix 1.2 & 1.3), and it's easy to see that the orange area (less freedom) expanded and the color of blue diluted.

To better observe the pattern, we plotted the averages of freedom and happiness across all years (Appendix 1.4). We observed that starting in 2018, freedom and happiness followed the almost same pattern. Our interpretation was that people might barely think of freedom as an issue before the pandemic, but after Covid hit, freedom became one of the important parameters to evaluate happiness.

We also looked into freedom's correlations with happiness, and the trend line indicates a significant positive relationship between freedom and happiness (Appendix 1.5). Outliers were also investigated by using the sum of freedom each year. Freedom in 2019 and 2020 are considered outliers when using the average with a 95% confidence interval as the reference (Appendix 1.6). It was expected to see freedom in 2019 being an outlier, but it was surprising to see freedom in 2020 was also an outlier for being rated even higher than before.

To get more details on this unexpected finding, we looked into each continent's freedom score from 2018 to 2020 (Appendix 1.7). The averages of freedom in 2018 and 2020 based on continents were about similar. While Australia and Africa's freedom in 2020 was still lower than

in 2018, North America, South America, Europe, and Asia's freedom in 2020 all increased compared to 2018, which could be contributing to a higher sum of freedom and make 2020 become the outlier. To investigate further, we selected North America to see if we could get more answers. We found that Canada and the US's freedom in 2020 was still lower than in 2018, but Mexico's freedom in 2020 was higher than ever (Appendix 1.9). However, we did some research but didn't find any related policies that could explain such a surge in Mexico's freedom in 2020.

While we tried to conclude, we refined our visualizations for better understanding. And we found that all continents' freedom and happiness scores followed the same pattern except for Europe, which always had a discrepancy (Appendix 1.8). Freedom seems not able to explain or reflect much of Europeans' happiness. Overall, we believe that there is a positive relationship between freedom and happiness, but the degree of the relationship can vary based on different continents and pre or post-Covid.

As Covid-policy-related inferences from the Happiness Score and freedom can lack enough data reliability (we don't have access to infer with certainty that Freedom in 2018-2019-2020 changed due to Covid's policies). Therefore, we decide to refine the hypothesis and look at other variables and alternative hypotheses that back up our primary hypothesis.

From the main hypothesis, we discover that there's a downgrade in happiness score from 2018 to 2019 (the year that Covid hit), and then slightly improves in 2020.

As an extension of the study, we analyze if the number of deaths per 1 million due of the population due to Covid is related to the happiness scores of countries:

H₀: The number of deaths per Covid doesn't affect a country's happiness rank

H₁: The number of deaths per Covid affects a country's happiness rank

For this study, we first plot a histogram for all countries' average rank of Happiness per year (see appendix XYY), where there's an observable decrease in the Happiness Score rank from 2018 (5.4647) to 2019 (5.4307) with a relative difference of -0.62% (-0.034 points).

Afterward, with the Data from Stata, we group the top 5 countries with more deaths per 1 million population (Bosnia and Herzegovina, Bulgaria, Georgia, Hungary, and Peru) and we group the top 5 countries with the least deaths per 1 million population (Burundi, Chad, China, New Zealand, and Nigeria). We compare the previous plot with the same plots but grouped by the corresponding countries (therefore, we compare 3 plots: the histogram with all countries (to see the average decrease from 2018 to 2019) vs the histogram with the average happiness rank score for the 5 countries most affected by Covid deaths vs the histogram with the average happiness rank score for the 5 countries most affected by Covid deaths (See appendices XYZ and XYZ)).

To choose the number of countries to be grouped, the study started with wanting to group the top 10 countries but had to be corrected to 5 countries as the data didn't show some of the top 10 countries.

As expected, the histogram that showed the most happiness score difference from 2018 to 2019 was the histogram for the five most affected countries by Covid (regarding deaths per capita), decreasing from 5.137 to 4.90 (-0.237, -4.72% of relative difference).

Nevertheless, contrary to what was expected, the second biggest difference in happiness scores between 2018 and 2019 came with the countries least affected by Covid deaths, decreasing from 4.9862 to 4.8244 (-0.1618, -3.3% of relative difference).

For a clearer understanding of patterns, we divided the grouped plots by country to analyze each country individually (see appendices XYZ and XYZ). For instance, from 2018 to 2019, Bulgaria and Hungary had a steep decrease in happiness score rank, whereas, against all odds, Georgia increased its happiness rank.

Therefore, after analyzing all plots, we conclude that the number of deaths doesn't have an impact on the countries' average happiness ranking score.

The result can be due to:

- The small sampling size (perhaps if instead of choosing the top 5 countries, we had chosen a larger number, the results would've been different).
- Other variables happening between 2018-2020 influence a country's happiness (political stability, wars, crisis, external factors, etc.).
- Countries' Covid concern level wasn't influenced as much by the number of deaths per capita but rather by other factors (media coverage, amount of people infected rather than people dead, economic crisis due to Covid rather than deaths per capita, etc.).
- Lack of reliability in a country's deaths reports (for instance, North Korea reported 0 deaths, but no data shows with exactitude the depth of the impact of Covid in North Korea, affecting the reliability of our study)

ADDITIONAL HYPOTHESIS EXPLORED

I. GOVERNMENT TRUST AND HAPPINESS SCORE

H_0 : Government Trust & happiness scores are independent of each other

H_1 : Government trust & Happiness scores are dependent on each other

To analyze if the happiness score of the 132 countries has any dependence on the government trust, we consider the metrics - average happiness score and the average government trust of the different countries from the year 2015 to 2020.

The values of the Average happiness score range from 3.079 (Burundi, Africa) to 7.546 (Denmark, Europe). The country with a median happiness score of 5.415 is Turkey & it has a government trust value of 0.119. When the Average Happiness score was taken on the Y axis and plotted against the average government trust values taken on the X axis, we can notice a clear pattern in the data. Most of the countries are clustered at the lower end of the X-axis. We can also see a clear pattern as the government trust increases and the happiness score of the countries also increases.

Based on the visual patterns some outliers to the above observations were the countries Singapore and Rwanda.

1. Rwanda - despite having a high government trust value of 0.47755, has only a happiness score of 3.418. So we can understand that probably some other factors like income level/xyz (other than the government trust) affect the happiness score of the country.
2. Singapore - Even though Singapore is much better than Rwanda, it has a much lower happiness score than the other developed countries despite also having a high government trust value.

Thus based on the pattern observed in the graph we can say that the government trust & happiness scores of the countries do have a certain degree of correlation & thus we can reject the null Hypothesis & conclude that the Happiness score of a country does depend on the trust that the people have on the government. (APPENDIX 2)

II. GDP PER CAPITA AND HAPPINESS SCORE

H_0 : GDP does not make a significant impact on the Happiness score

H_1 : GDP does make a significant impact on the Happiness score

Gross Domestic Product is one of the popular measurements of a country. To explore this hypothesis further, we have conducted regression analysis and visualized using correlation plots, regression lines, scatterplots, heatmaps, and more (APPENDIX 3). Some unique insights we observed were - Qatar has the highest GDP score but is not the happiest country as 80% of residents are not citizens. Syria has a relatively lower happiness score compared to other countries with the same level of GDP. This interactive scatterplot shows that there is a strong positive correlation between GDP and Happiness. While per capita output is indeed correlated with these other variables (wealthier countries can provide more social services, for example), the differences in factors other than GDP seem to explain much of the variance in happiness for countries at the top of the list. Just increasing GDP without changing how GDP is spent provides rapidly diminishing returns to human contentment.

If we consider Costa Rica which is the 13th happiest country in the world. Yet GDP per capita only explains 14.1% of the nation's overall happiness score, whereas social support explains substantially more, about 20% of the score. The United States, on the other hand, explains 19% of its happiness score with per capita income and is ranked 5 spots below Costa Rica. Statistically speaking, Costa Ricans "use" substantially less GDP to generate a level of happiness greater than what Americans generate with far more GDP.

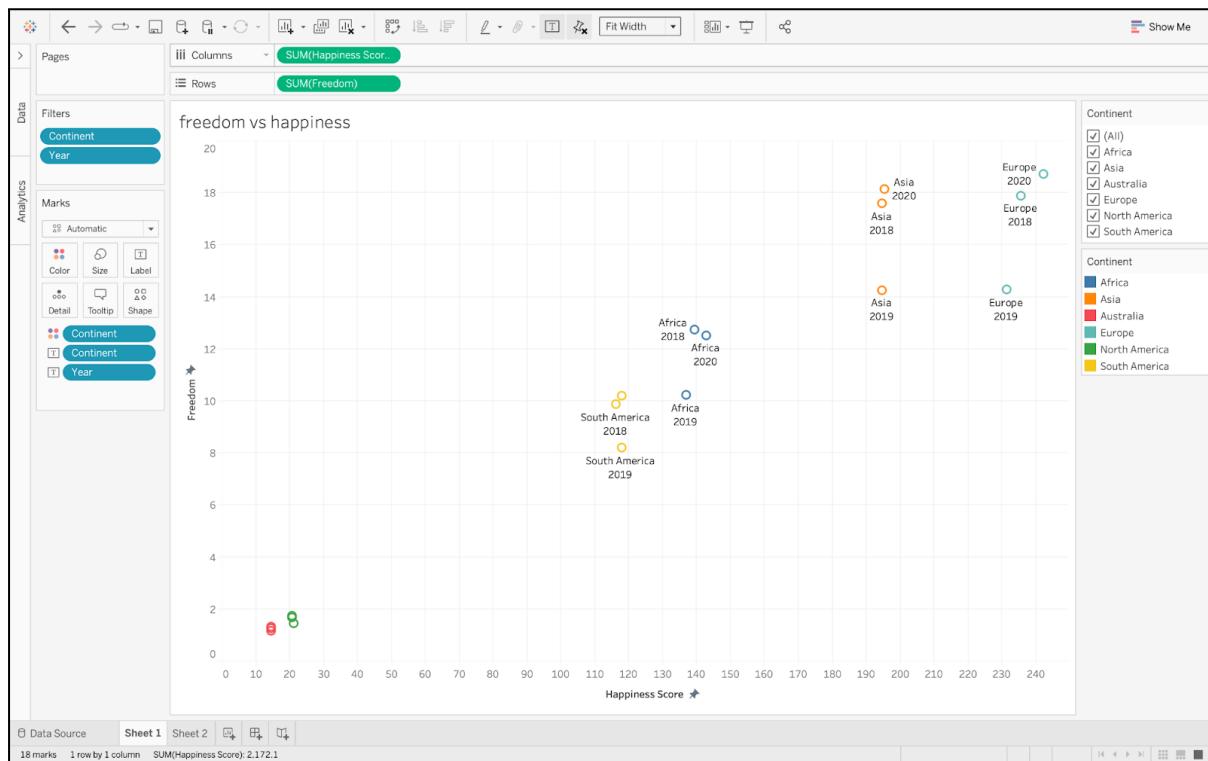
REFERENCES

Best, R. de. (2022, July 27). *Covid-19 deaths per capita by country*. Statista. Retrieved November 16, 2022, from <https://www.statista.com/statistics/1104709/coronavirus-deaths-worldwide-per-million-inhabitants/>

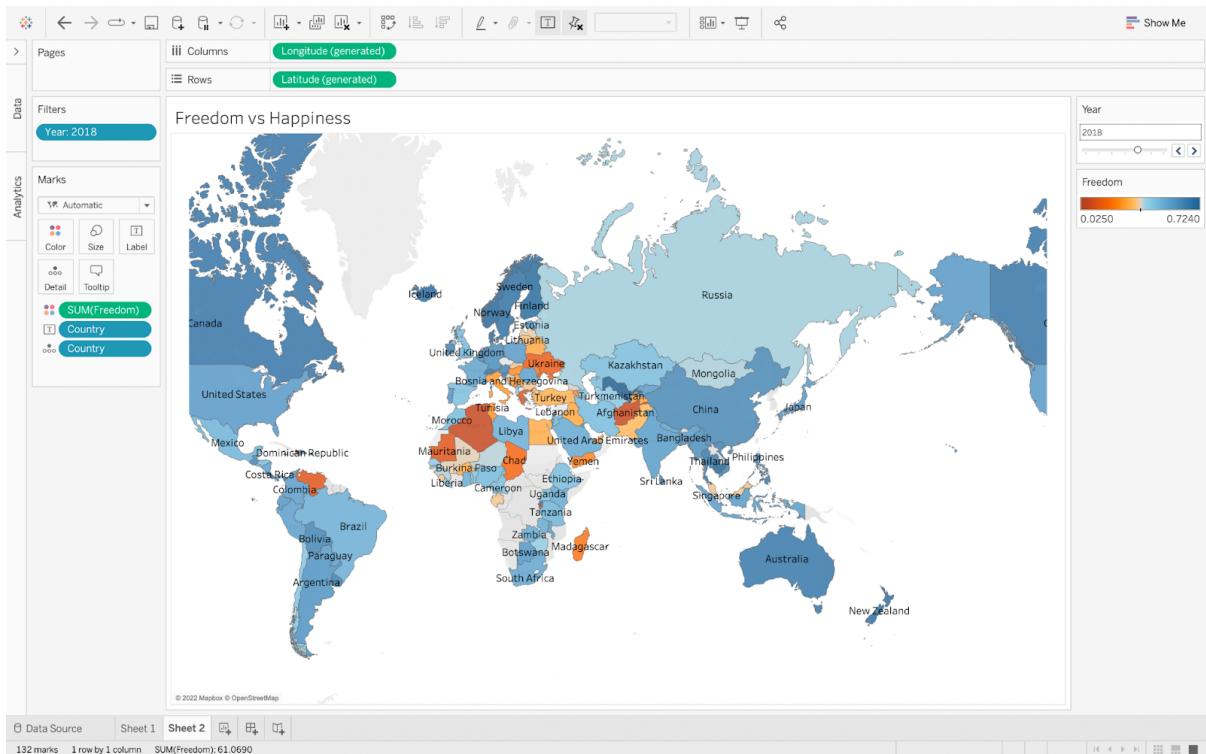
APPENDIX

APPENDIX 1: FREEDOM AND HAPPINESS SCORE

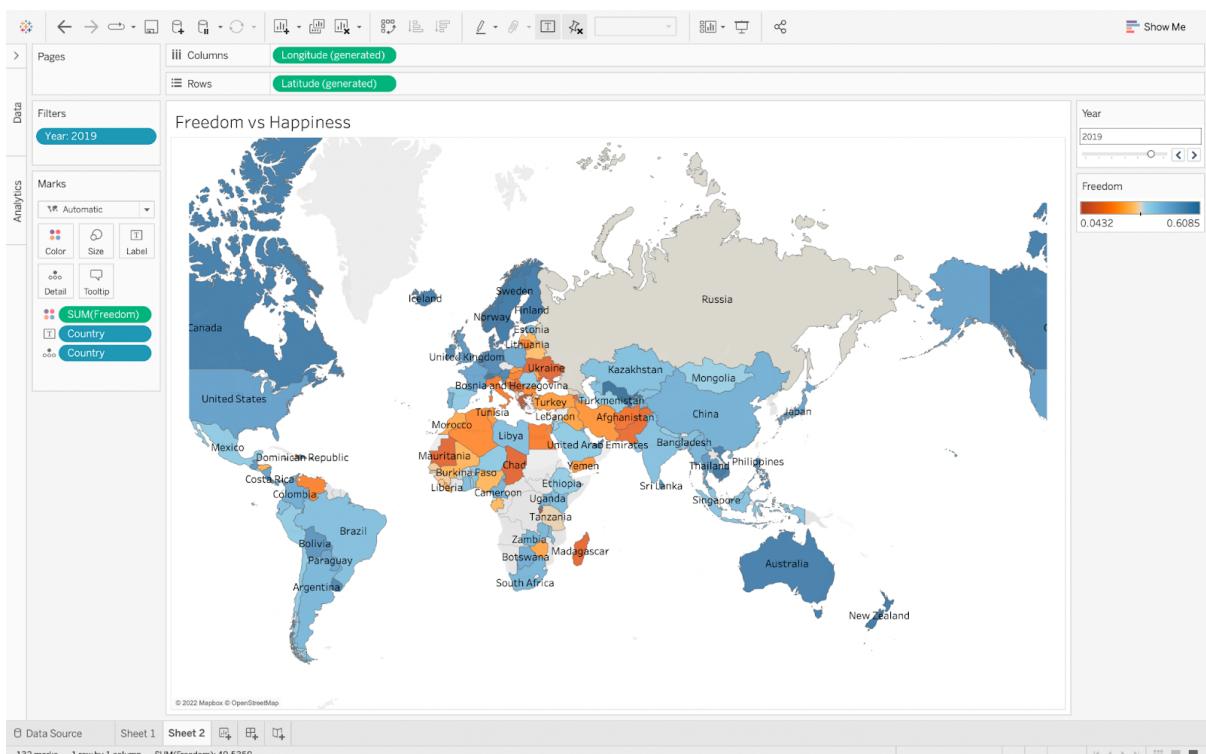
APPENDIX 1.1 - Each continent's Freedom score vs Happiness score (years 2018-2020)



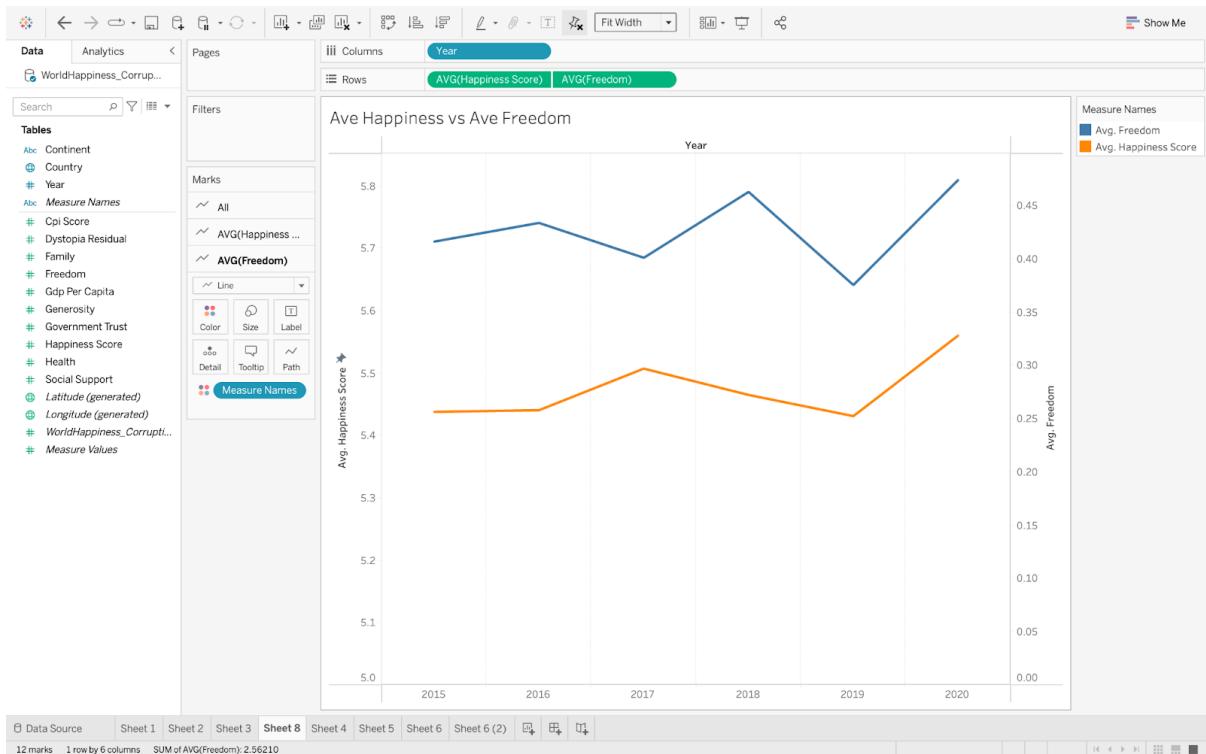
APPENDIX 1.2 - Freedom Map in 2018



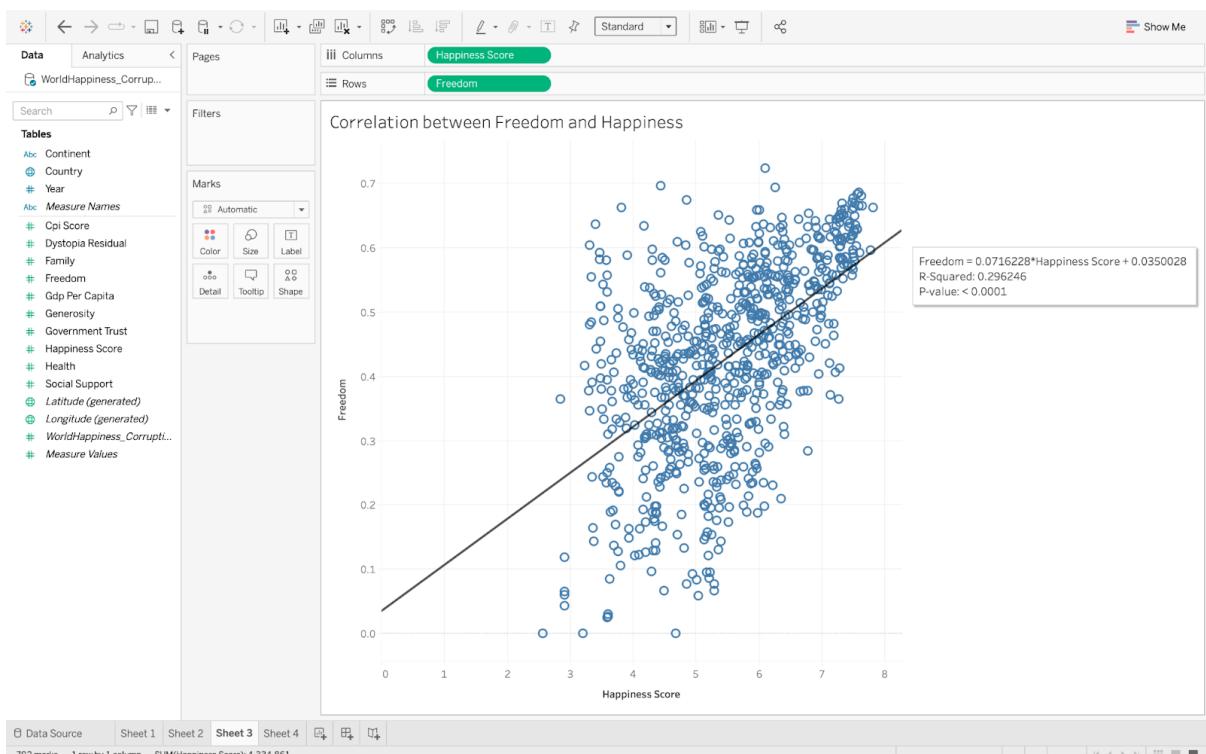
APPENDIX 1.3 - Freedom Map in 2019



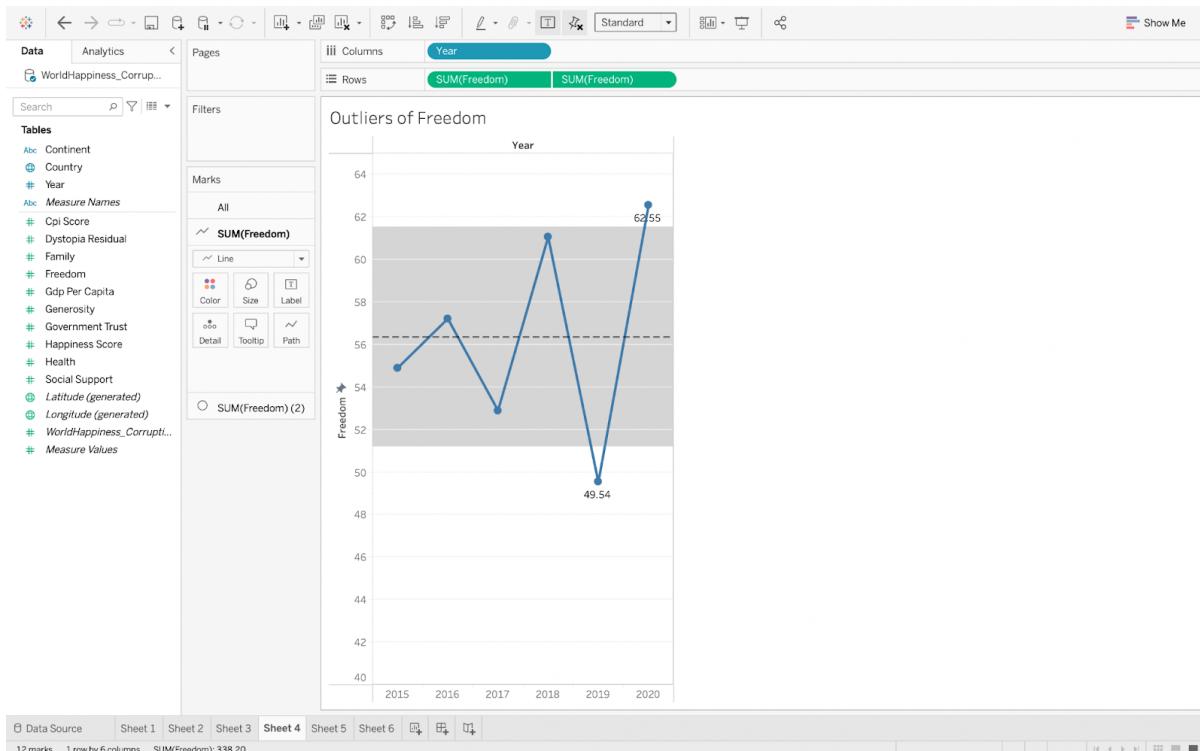
APPENDIX 1.4 - Freedom score vs Happiness score across all years



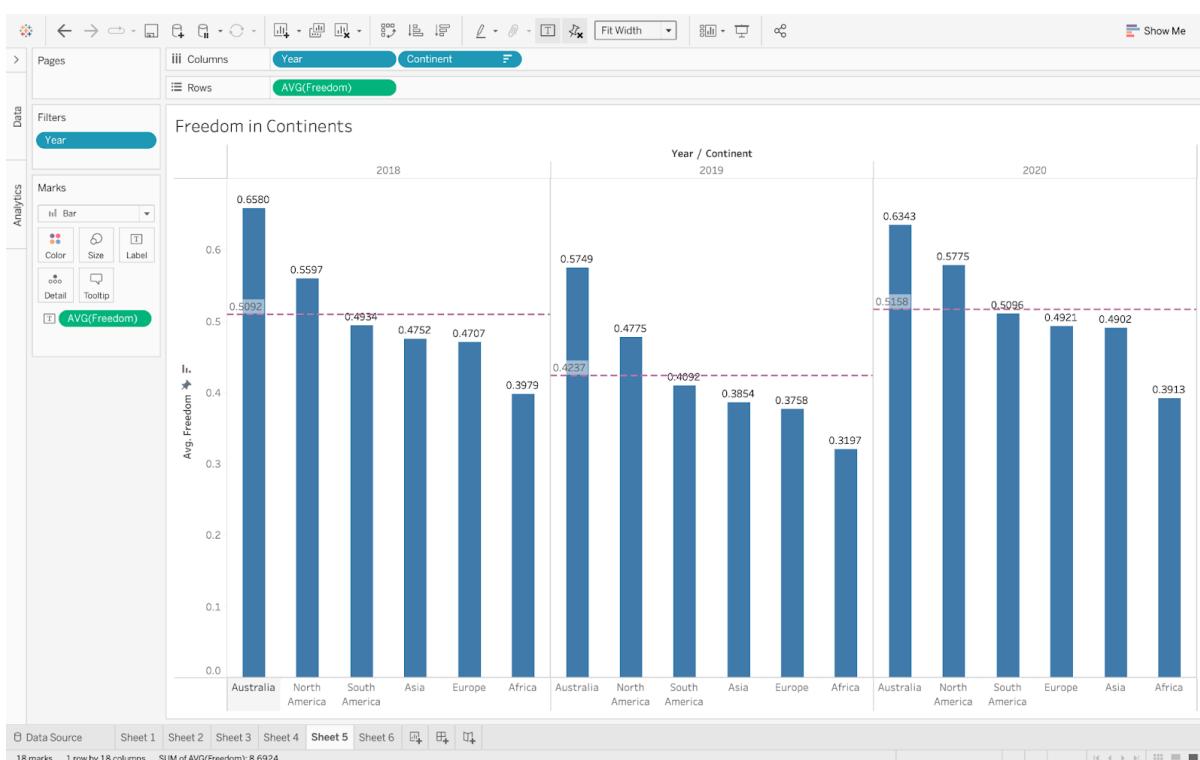
APPENDIX 1.5 - Correlation between freedom and happiness



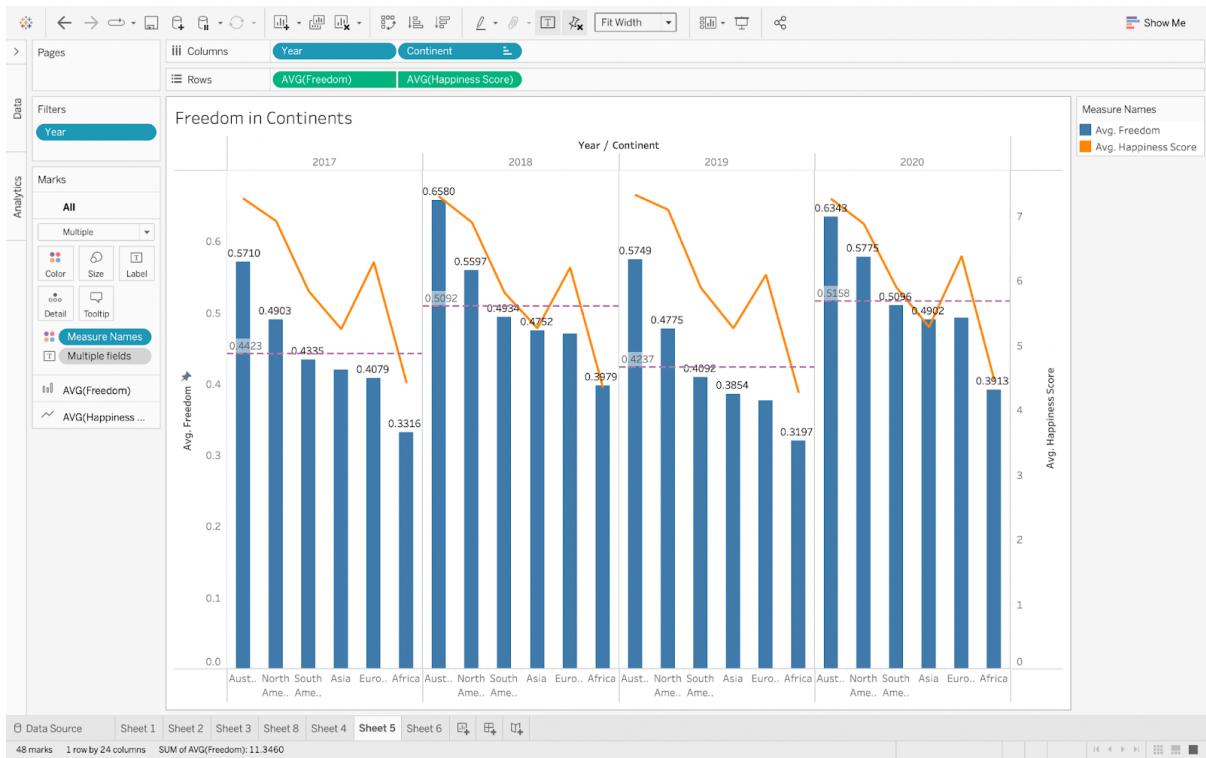
APPENDIX 1.6



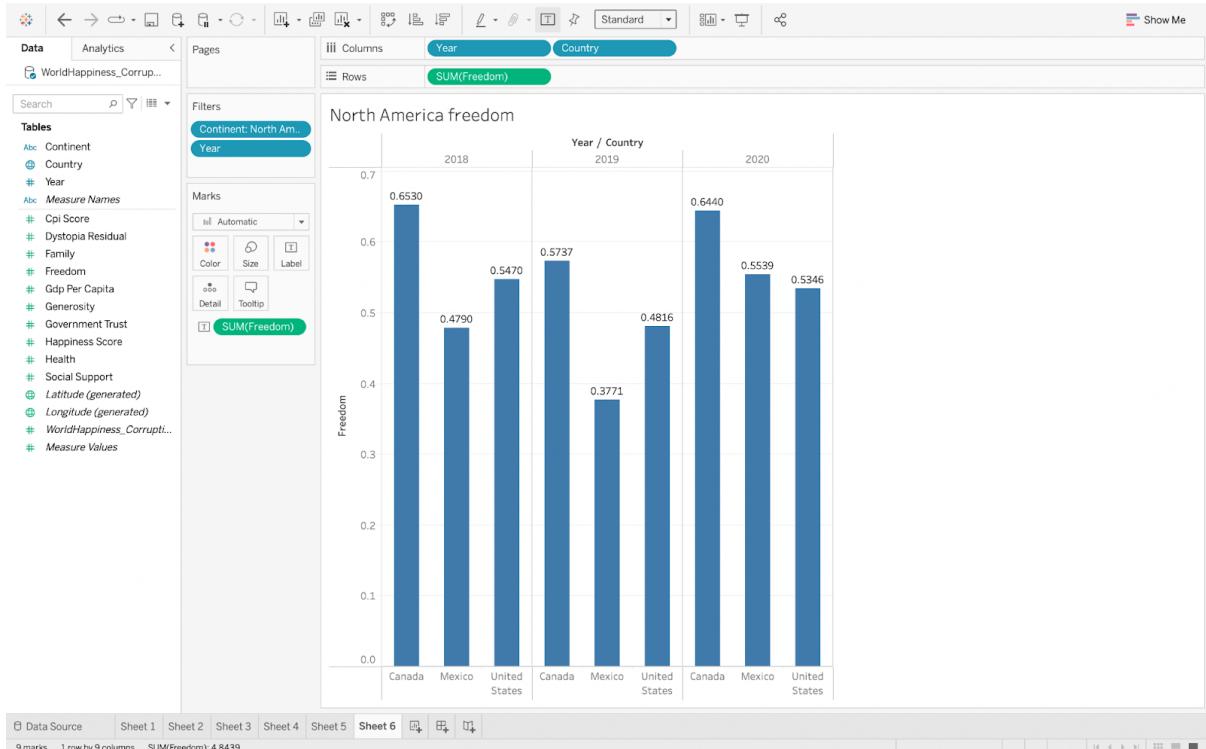
APPENDIX 1.7 - Freedom in each continent (years 2018-2020)



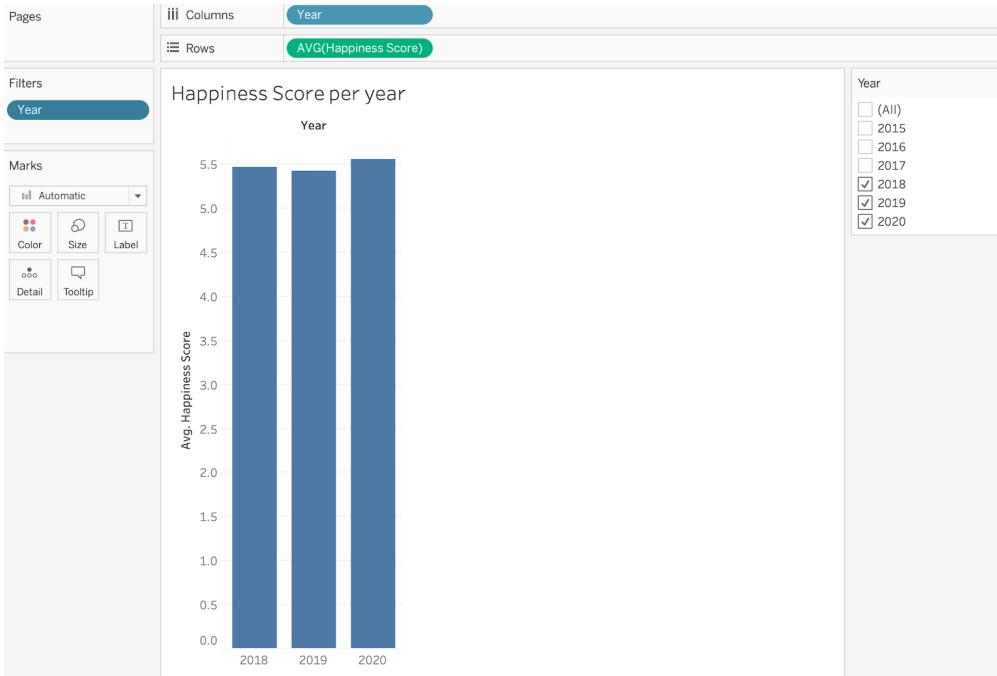
APPENDIX 1.8 -



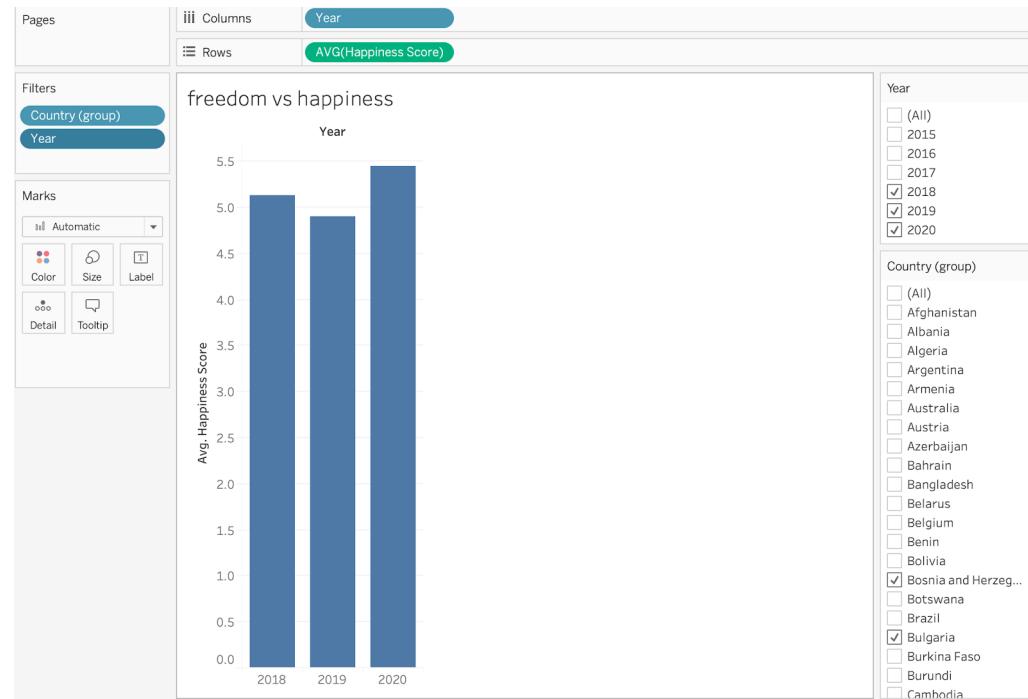
APPENDIX 1.9 - Freedom in North America (years 2018-2020)



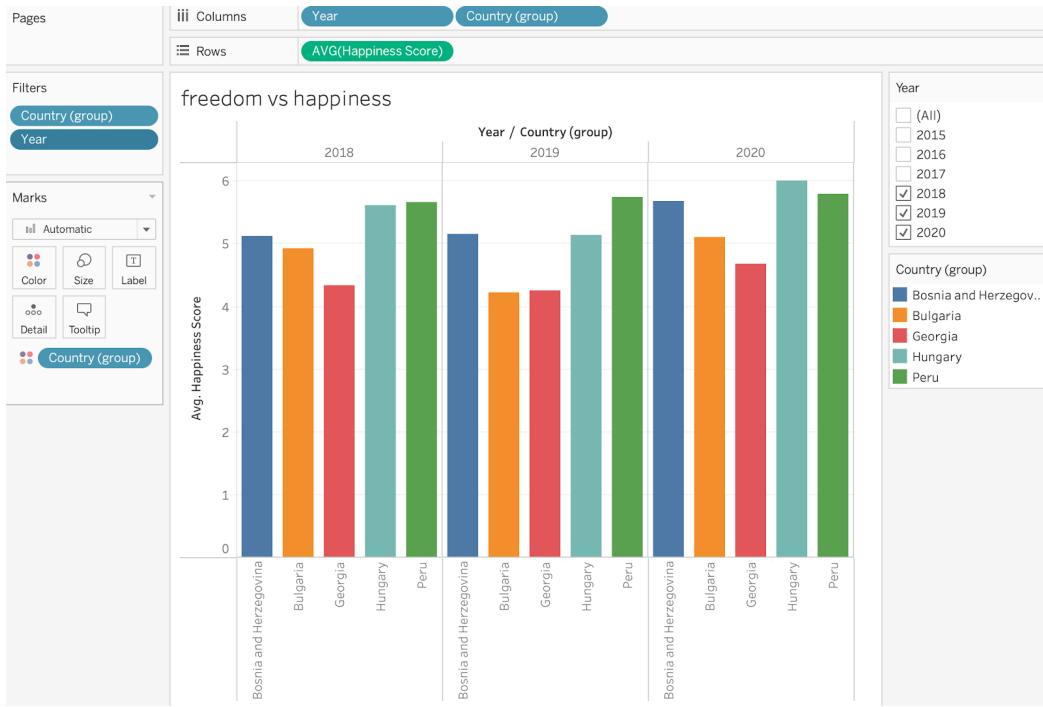
APPENDIX 1.10 - Average Happiness Score per year for all countries (years 2018-2020)



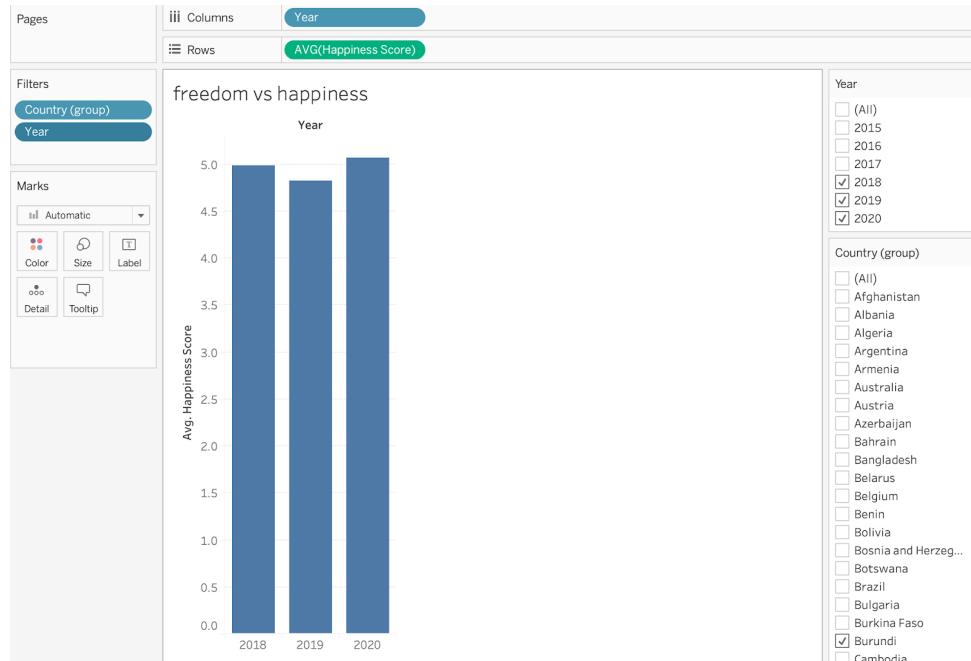
APPENDIX 1.11 - Average Happiness Score per year for top 5 deaths per capita due to Covid countries (years 2018-2020)



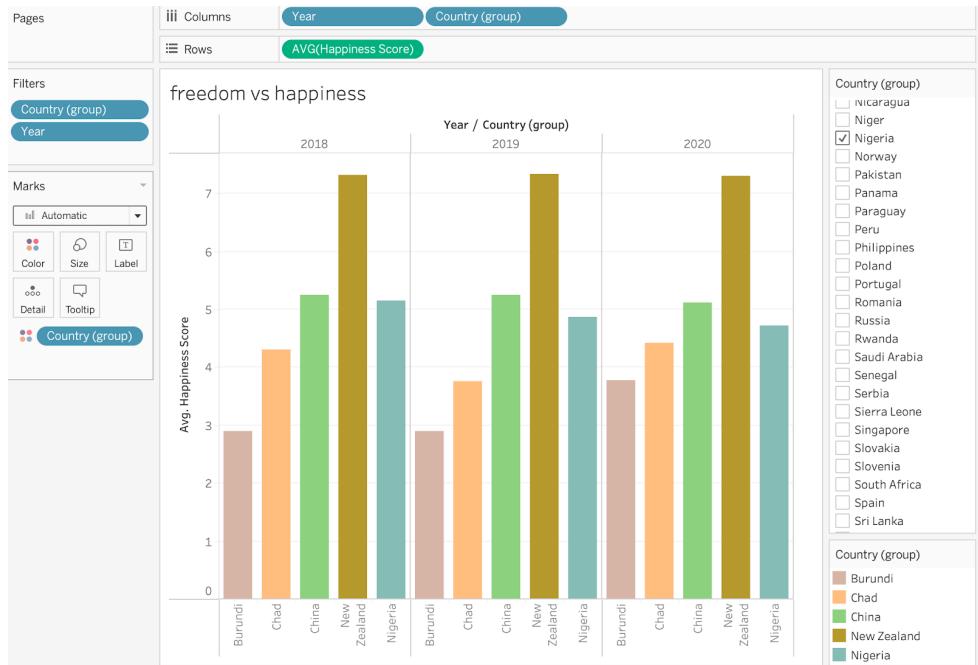
APPENDIX 1.12- Detail per country of the average Happiness Score per year for the top 5 deaths per capita due to Covid countries (years 2018-2020)



APPENDIX 1.13 - Average Happiness Score per year for the least 5 deaths per capita due to Covid countries (years 2018-2020)

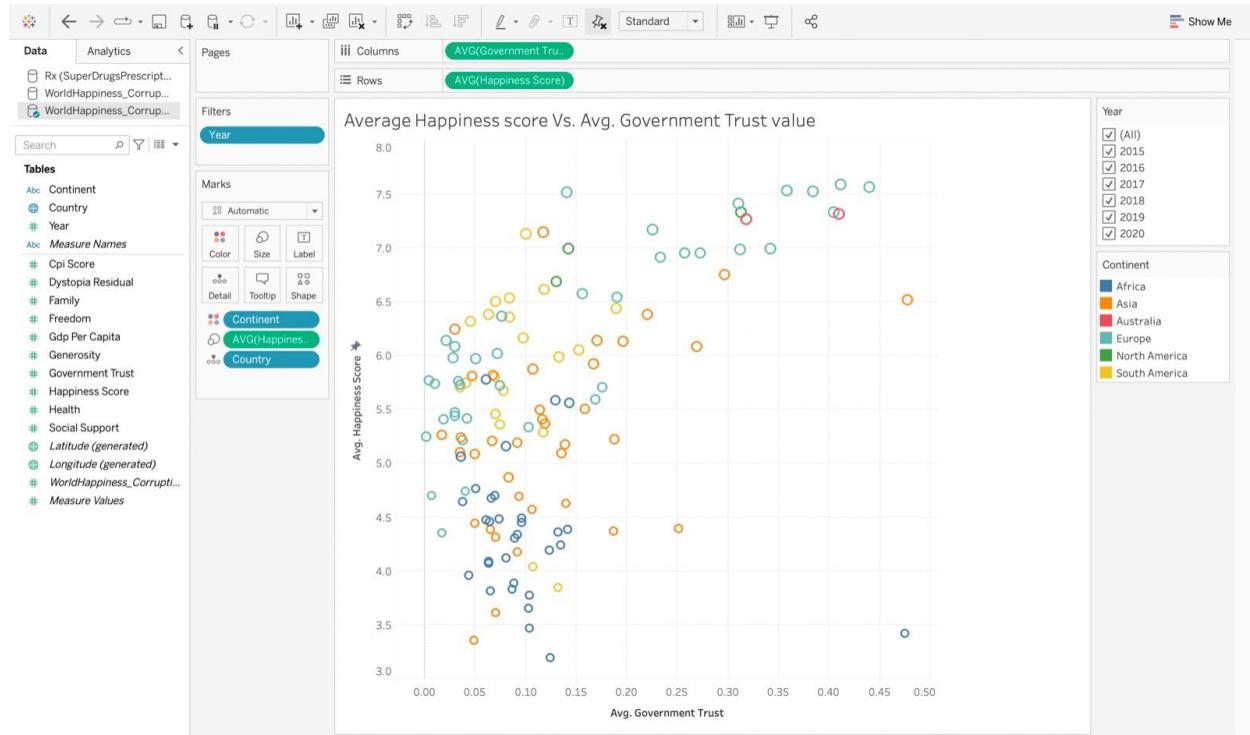


APPENDIX 1.14 - Detail per country the average Happiness Score per year for the least 5 deaths per capita due to Covid countries (years 2018-2020)

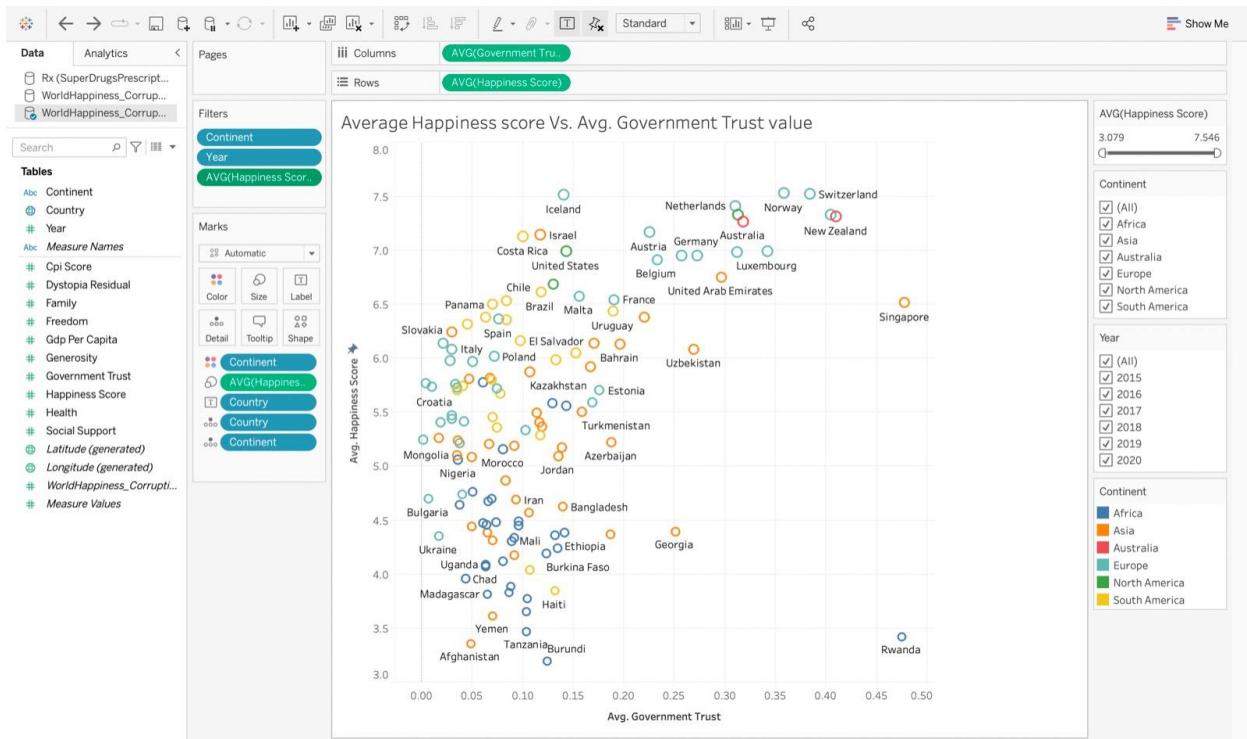


APPENDIX 2 - GOVERNMENT TRUST AND HAPPINESS SCORE

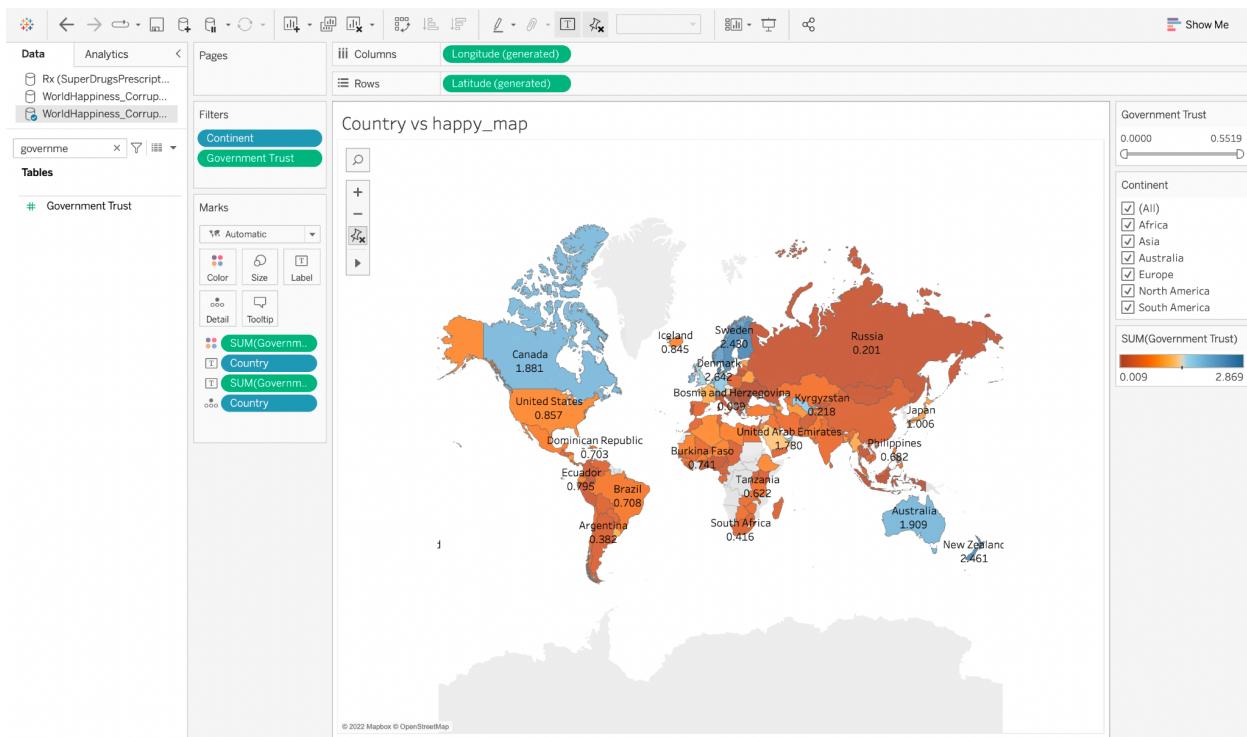
APPENDIX 2.1: INITIAL SCREENSHOT



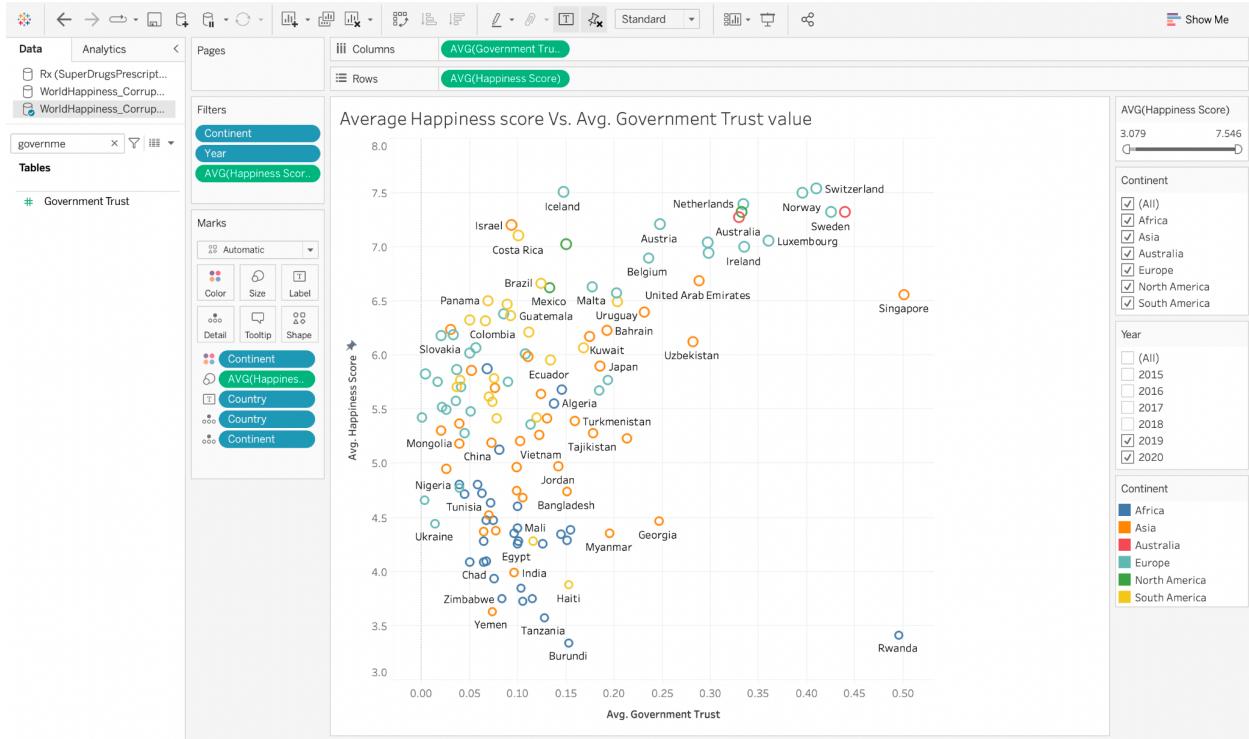
APPENDIX 2.2: FINAL SCREENSHOT



APPENDIX 2.3 - Freedom Map in 2018

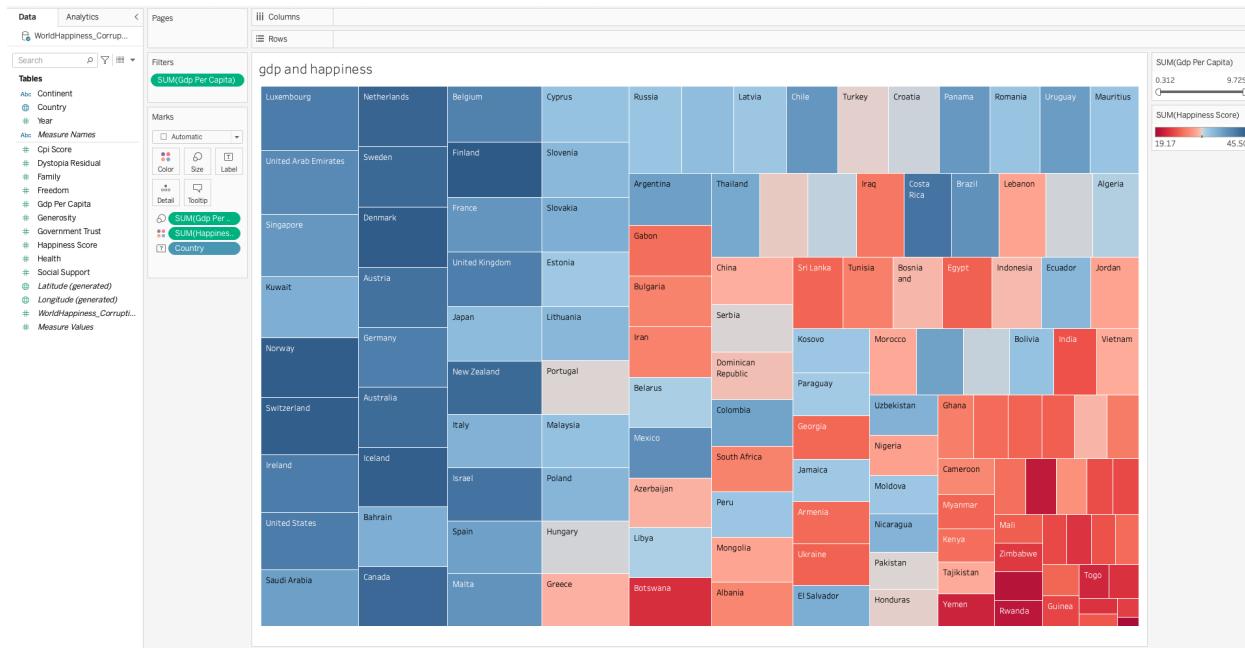


APPENDIX 2.4 - Relationship between Government trust vs happiness score after the advent of Covid in 2020.



APPENDIX 3 - GDP AND HAPPINESS SCORE

APPENDIX 3.1: INITIAL SCREENSHOT - GDP and Happiness scores of all countries (2015-2020)



APPENDIX 3.2: FINAL SCREENSHOT - GDP AND HAPPINESS ACROSS ALL CONTINENTS TO IDENTIFY OUTLIERS

