

Project Proposal

Title: Happiness And Covid

The aim of this project is to find the impact of COVID mortality on happiness. The World Happiness Report (WHR) conducts a yearly global survey and collects the data to report how people evaluate their own lives in over 150 countries worldwide. Through the last 10 years WHR delivered the happiness outcomes based on the averaged scores collected for each country by accessing their subjective well-being divided into 6 categories (Income, Social Support, Life Expectancy, Freedom, Generosity, Corruption).

From 2020 until today, the pandemic had a tremendous impact on human health and freedom affecting our overall happiness. The World Health Organization (WHO) monitors and daily updates the number of people affected by COVID across the whole world.

In this project, we aim to combine both sources of data and find how the last two years of living in a pandemic shaped the outcomes of World Happiness. We used Jupyter notebook python scripts to transform and analyse data from the data sources via file "Covid_Happiness_Analysis.ipynb" where we use data in the "Resources" folder.

We have uploaded our final data "Analysis_Covid_Happiness_Data.csv" into PostgreSQL relational Database called "HappinessCovid_db" using "Schemas_HappinessCovid.sql" to create table import the final data into the database. We also used "Query_HappinessCovid.sql" file to run a query and double check if the data has been imported correctly. The plotly scatter output to show the correlation of the data in "Output" folder.

Findings

Per the scatter plots, we found that there are no correlations between the percentage of COVID deaths (Percentage_deaths) and the percentage change in happiness (Happiness_Difference).

Through the last three years (2020-2022), mainly Norden countries reported the highest scores in happiness survey, Finland, Denmark, Switzerland, Iceland, Norway and Sweden found at the top of the ranking. In the last two years a decrease in happiness scores was registered for 54, while 66 countries reported higher content with their lives regardless of the pandemic.

Although the happiest 5 countries scored less in the survey in the last two years, they maintained their high positions. The number of COVID-caused deaths in these countries accounted for 0.05 - 0.1% of the total population (<10M), which is relatively low when compared to more populated countries which lost nearly 1% of their population. These findings suggest that Happy Countries' Social Support (including medical assistance) may play important roles in

well-being and reported happiness. Switzerland, although not part of Nordic countries, is also renowned for their socialist policies as well as one of the highest GDP per capita in the world.

Extract

This data has been downloaded and saved in the “Resources” folder. The sources are as follows:

- (1) World Happiness - 3 csv files holding the outcomes for surveys collected in 153 countries. (<https://worldhappiness.report/ed/2022/#appendices-and-data>)
- (2) World Health Organization - csv file containing data on COVID mortality collected from all 195 countries between 01/01/2020 - today (May 2022). (<https://covid19.who.int/data>)
- (3) World Bank - csv file containing the total population for each country collected in 2020. (<https://data.worldbank.org/indicator/SP.POP.TOTL>)

WHO data was used up to 31/03/2022, as to match the timeline of the World Happiness yearly survey. All data files used were originally in .csv format.

Transform

Data transformation was done using Python through Jupyter Notebook with Pandas, Matplot, Numpy and Plotly modules.

Happiness Data

- Filtered the columns ['Country name' : 'Country', 'Ladder score'] Ladder score = Overall happiness
- Removed NaN values
- Combined each year 2020, 2021, 2022 into one table by merging on ['Country']
- Cleaned up the country names to match the names in other data using *country_converter*
- Calculated the percentage change of reported happiness between 2020 (pre-COVID) and 2022 (during COVID pandemic)

Covid Death Data

- Filtered the data copying two columns into a new DataFrame: ['Country', 'Cumulative_deaths']
- Removed NaN values
- Calculated the total number of deaths per each year: 31/12/202X (2020, 2021, 2022)
- Combined the data for three years into one dataframe by merging on ['Country']
- Cleaned up the country names to match the names in other data using *country_converter*
- Left-joined the Happiness data with COVID data on 'Country' to eliminate countries not found in the Happiness data

- Calculated the total number of COVID deaths between 2020-2022 to reflect total deaths for each country

World Bank Population Data

- Filtered the data by copying two columns into a new DataFrame: ['Country', 'Population_2020']
- Cleaned up the country names to match the names in other data using *country_converter*

Final Analysis Data

- Join all the above data together based on country names
- Analyse the data by calculating the following:
 - Average Happiness from 2020 to 2022
 - Total deaths from 2020 to 2022 per population
 - Percentage changes in happiness from 2020 and 2022

Load

The aim of this project is to see the impact of covid in a world happiness index.

We have five tables: 2020-world happiness, 2021-world happiness, 2022-world happiness, WHO-World-Covid-gobal-data and World-Bank where we extract happiness scores from world happiness tables, population from world-Bank and death from WHO-Covid-gobal-data to analyse and reach the final data "Analysis_Covid_Happiness_Data.csv" and upload to PostgreSQL Database called "HappinessCovid_db".

The table below describes the final data in the database.

Column Name	Description
Country	Country that participated in the happiness survey
Happiness_Difference	Percentage difference between the happiness in 2022 and 2020
Percentage_deaths	Percentage of total deaths cumulative from 2020 - 2022 compared to the population of the country
Average_Happiness	Average of happiness score across three years (2020 to 2022)