

# Happiness Before And After COVID-19



Alicia Zhou Zhu

Digital Humanities 100  
<link to course catalog>

---

# Overview

**Research Question: How has happiness changed since COVID-19? Compare it to the previous years.**

## Presentation topics:

1. Description
2. Data Set
3. Methods & Techniques
4. Sub-Questions / Hypothesis
5. Graph Analysis



Happiness is the state of the mind



Happiness Tokyo Japan Black And White Street



We chose happiness

# Description

The dataset being used is called "World Happiness Report". It is a landmark survey that reviews the state of global happiness and show how personal and national variations contribute to one's happiness. This dataset was acquired through the Gallup World Poll. The columns demonstrate happiness scores through six factors:

1. economic production,
2. social support,
3. life expectancy,
4. freedom,
5. absence of corruption, and
6. generosity.

The reasoning behind using this dataset is to measure the impacts of COVID-19 in terms of happiness levels throughout the world.

# Data Set : World Happiness Report

# World Happiness Report (2005 - 2020)

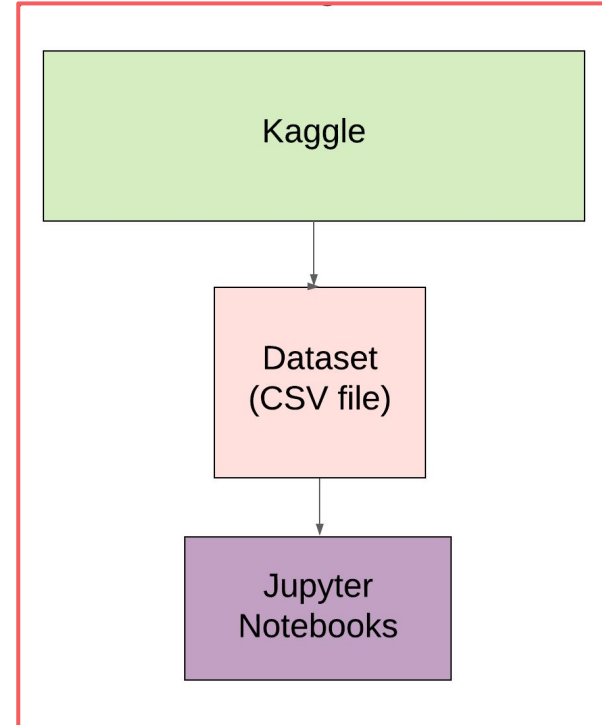
[illegible]

# World Happiness Report (2021)

[illegible]

# Methods & Techniques

- Kaggle
- Jupyter Notebook
- Libraries:
  - Pandas
  - Matplotlib
  - Numpy
  - Seaborn



# Sub-Questions / Hypothesis

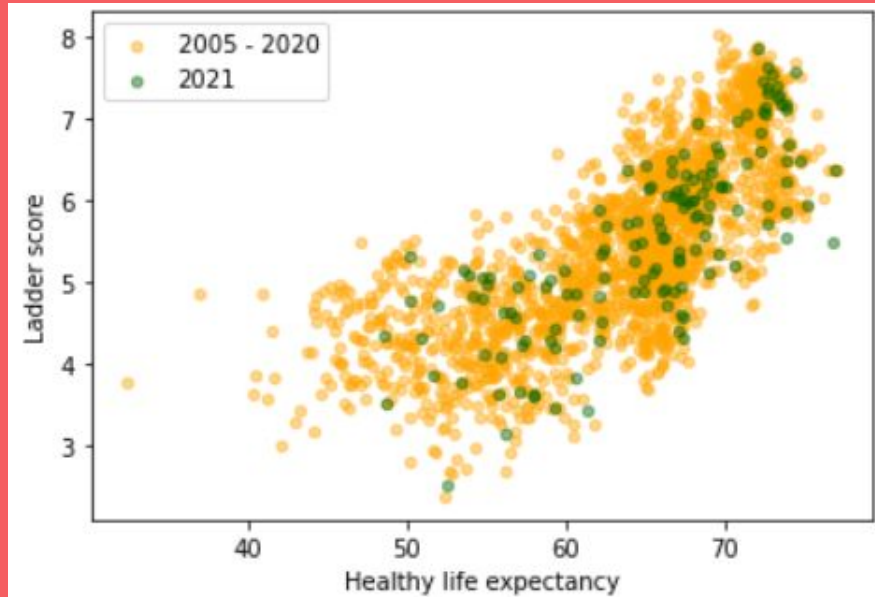
## Sub-Questions:

1. What does happiness per year look like?
2. What are the happiness levels per region in comparison to before and after COVID-19?
3. What are the differing factors between a happy and a less happy region?
4. Which World Happiness Report defines the situation of COVID-19?

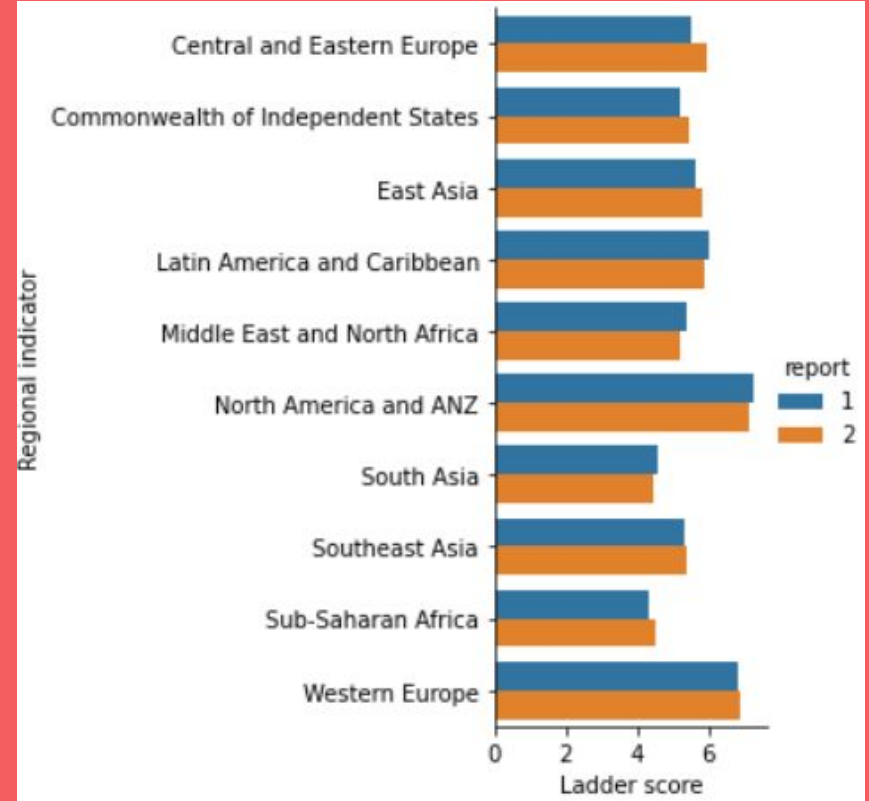
## Potential outcome:

- Happiness levels should decrease

# Correlation between Life Satisfaction and Life Expectancy



# Life Satisfaction sorted by region between years 2005 - 2021





## Before COVID-19:

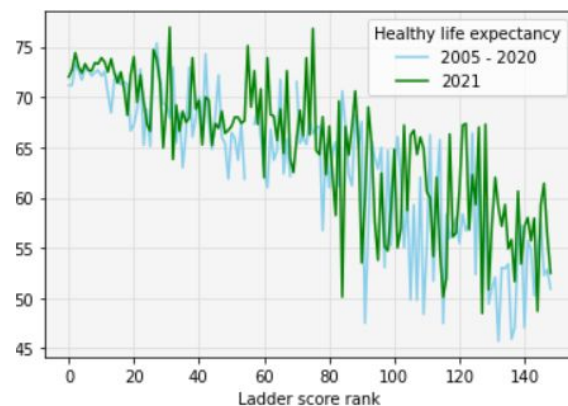
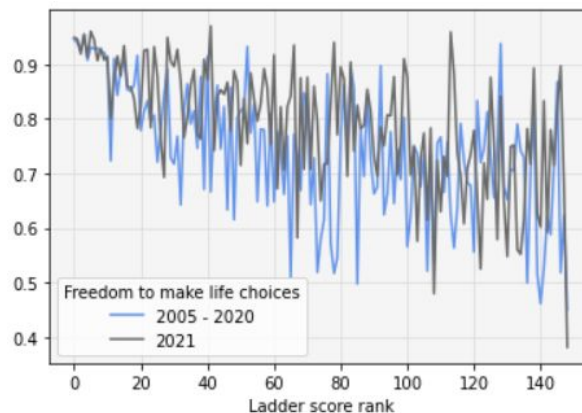
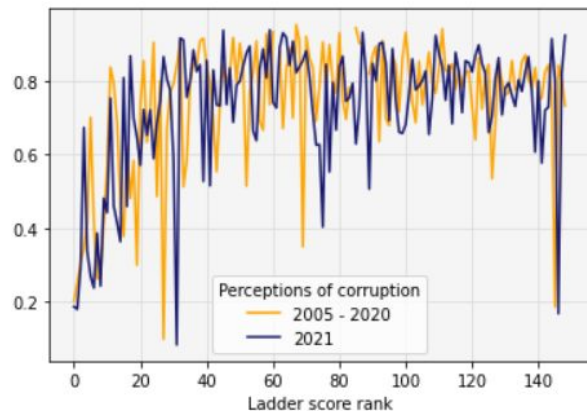
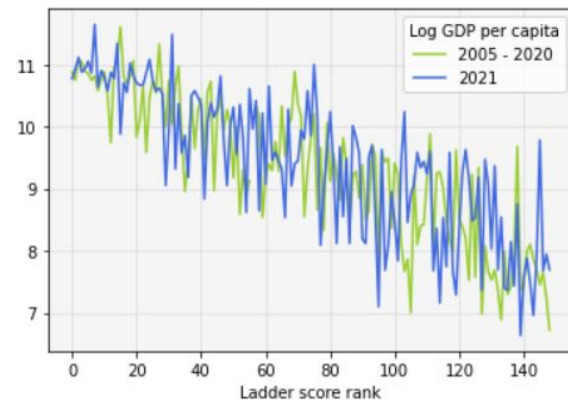
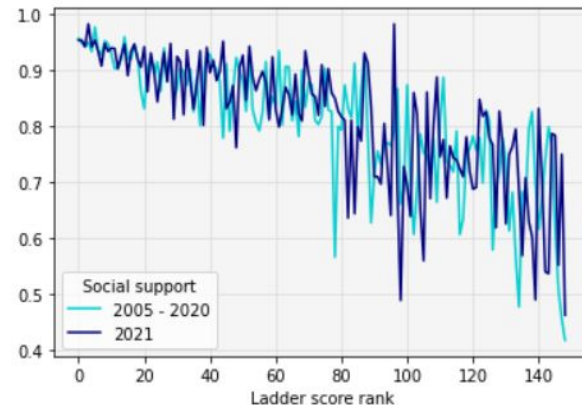
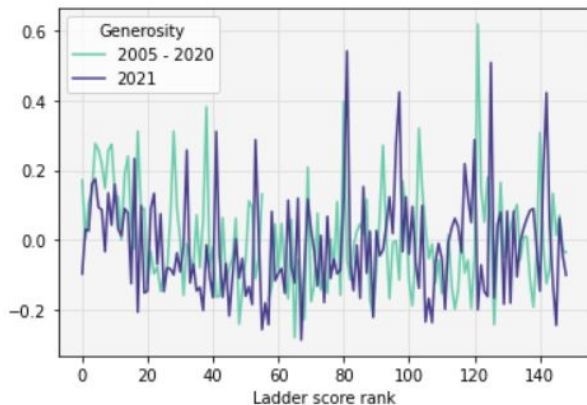


## After COVID-19:





# The Six Main Factors:

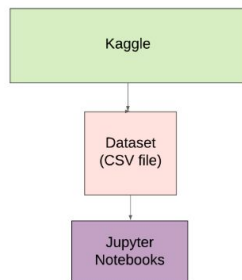


# Happiness Before And After COVID-19 : Storyboard

## Descriptions:

The dataset being used is called "World Happiness Report". It is a landmark survey that reviews the state of global happiness and show how personal and national variations contribute to one's happiness. This dataset was acquired through the Gallup World Poll. The columns demonstrate happiness scores through six factors: economic production, social support, life expectancy, freedom, absence of corruption, and generosity.

The reasoning behind using this dataset is to measure the impacts of COVID-19 in terms of happiness levels throughout the world.



## Methods & Tools:

Used Kaggle as the tool to obtain the dataset. To do data analysis, I have decided to use Jupyter Notebook as my means of creating graphs. This is because it can be run locally and it includes the packages I am going to use such as "pandas", "matplotlib", and "numpy".

Pandas will be used to import the dataset (CSV file) and for data analysis. Matplotlib has the potential of being used for graphing the necessary data charts. Numpy is for the data manipulation.

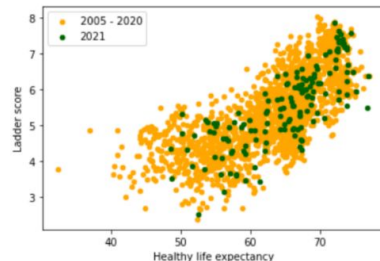
## Research Question:

How is the happiness after the spread of COVID-19 compared to the previous years?

## Sub-Questions:

- 1) What does happiness per year look like?
- 2) What are the happiness levels per region in comparison to before and after COVID-19?
- 3) What are the differing factors between a happy and a less happy region?
- 4) Which World Happiness Report defines the situation of COVID-19?

## Correlation between Life Satisfaction and Life Expectancy

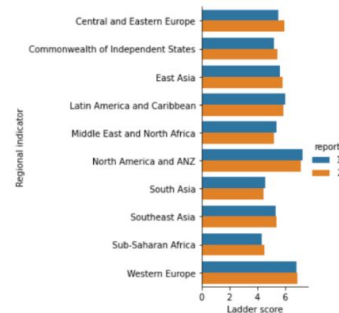


It shows what happiness per year in accordance to healthy life expectancy. The blue dots represent their life satisfaction rating all the countries throughout the years from 2005 to 2020. While, the green dots are data from 2021 to see a contrast between years. In this plot, there is also a strong positive correlation with a healthy life and life satisfaction. It seems that people tend to be happier when they are healthy and feel as though they have a longer lifespan. Thus, a possible factor that affects happiness is life expectancy.

## Potential Results:

The hypothetical outcome is that the happiness levels should have decreased world-wide due to the pandemic. The comparison of the tables in other years to the one relevant to the pandemic should display lower happiness levels and other categories that correlate with unhappiness.

## Life Satisfaction sorted by region between years 2005 - 2021



Report 1 represents the ladder scores of years 2005 to 2020. Report 2 represents the ladder scores of year 2021. The blue bar or report 1 represents the life satisfaction before COVID-19 happened. While, the orange bar or report 2 represents the life satisfaction after the pandemic.

The visualizations demonstrates that there was no significant change between the mean ladder score of the years 2005-2020 compared to 2021. If anything, some regions demonstrate that their ladder score in the year of 2021 increased. And, the correlation graph shows higher scores in 2021.

## Interpreting the results:

From the visualizations, happiness before and after has not changed in a drastic manner whether positive or negative. In some regions, they have become happier while others have lowered slightly. However, overall there is no significant change on the happiness after the spread of COVID-19.

## Conclusions

Worldwide issues such as pandemics are directly correlated to the happiness levels of the world. With understanding how the factors affect the happiness levels of each country, legislations could determine what could be improved through these factors. However, analyzing the life satisfaction rating alone has produced no concrete evidence of affected factors.

## Next Steps / Future Work

The dataset "World Happiness Report" can potentially be used for more than analyzing life satisfaction. It has the potential of be used to further research how economics of different countries affect life expectancy. Through a common pattern from the research, it could lead to understanding why some countries underperform and determine if it could be a causation to life satisfaction scores from each country

## Work Cited / References

Dataset:  
<https://www.kaggle.com/ajaypalsinghlo/world-happiness-report-2021>  
Helpful Links:  
<https://github.com/AshQTan/DH100>  
<https://worldhappiness.report/ed/2021/>  
[https://pandas.pydata.org/pandas-docs/stable/user\\_guide/index.html](https://pandas.pydata.org/pandas-docs/stable/user_guide/index.html)

## Code:

[https://github.com/azhouzhu/DH100\\_Summer2021](https://github.com/azhouzhu/DH100_Summer2021)

# Jupyter Notebook

## Jupyter Notebook code in Google Collab

### Jupyter Notebook Code

Correlation between Life Satisfaction and Life Expectancy

```
ax = report1.plot.scatter(x = 'Healthy life expectancy', y = 'Ladder score', color = 'Orange', label = '2005 - 2020')
report2.plot.scatter(x = 'Healthy life expectancy', y = 'Ladder score', color = 'DarkGreen', ax = ax, label = '2021')
```

Life Satisfaction sorted by region between years 2005 - 2021

```
r1_region_ladder = report1_meanByRegion[['Regional indicator', 'Ladder score']]
r2_region_ladder = report2_meanByRegion[['Regional indicator', 'Ladder score']]
r1_r2_meanByRegion = pd.concat([r1_region_ladder, r2_region_ladder])
r1_r2_meanByRegion['report'] = (len(r1_region_ladder)*(1,) + len(r2_region_ladder)*(2,))
r1_r2_meanByRegion.reset_index(inplace = True)
r1_r2_meanByRegion
sns.factorplot(x='Ladder score', y='Regional indicator', hue='report', kind='bar', data=r1_r2_meanByRegion)
```

Life Satisfaction histogram comparison (2005 - 2021)

```
ladScore1_arr = ladScore1.to_numpy()
ladScore2_arr = ladScore2.to_numpy()

plt.hist([ladScore1_arr, ladScore2_arr], bins = 15, label = ['2005 - 2021', '2021'], color = ['skyblue', 'steelblue'])
plt.legend(loc='upper left')
```

**For more code, refer to my GitHub. (page 1)**

# GitHub Repository










## DH 100 Repository | Summer 2021



[azhouzhu / DH100\\_Summer2021](#) PrivateUnwatch 2 Star 1 Fork 0

[Code](#) [Issues](#) [Pull requests](#) [Actions](#) [Projects](#) [Security](#) [Insights](#) [Settings](#)

[main](#) [1 branch](#) [0 tags](#)

[Go to file](#) [Add file](#) [Code](#)

 <b>azhouzhu</b> Update README.md <span>fb85c4f 2 days ago</span> <a href="#">23 commits</a>
 <code>.gitattributes</code> Initial commit <span>23 days ago</span>
 DH 100 Data Poster.pdf Upload final poster <span>2 days ago</span>
 DH 100 Data Storyboard v3.0 - Page ... Update storyboard <span>17 days ago</span>
 DH 100 Data Storyboard v3.0 - Page ... png <span>17 days ago</span>
 DH100 project.ipynb Upload final poster <span>2 days ago</span>
 README.md Update README.md <span>2 days ago</span>
 world-happiness-report-2021.csv Add files via upload <span>23 days ago</span>
 world-happiness-report.csv Add files via upload <span>23 days ago</span>

 README.md 

## DH 100 Portfolio

DH100 Theory and Method: The emphasis of this introductory digital humanities course is on the theoretical and

### About

DH100 Theory and Method: The emphasis of this introductory digital humanities course is on the theoretical and methodological considerations involved with data-intensive research within the context of the humanities. We will explore the ways the union of data analysis and humanities research is both productive and in tension.

[Readme](#)

### Releases

No releases published  
[Create a new release](#)

### Packages

No packages published  
[Publish your first package](#)



# Poster (P.1 Preview)

# Happiness Before And After COVID-19 : Poster

Course Title: DH 100 Theory and Methods  
Instructor: Dr. Adam Anderson  
Student: Alicia Zhou Zhu  
Date: 06/18/2021



Happiness is the state of the mind



Happiness Tokyo Japan Black And White Street Photography



We chose happiness

## World Happiness Report (2005 - 2020)

	Country name	year	Ladder score	Log GDP per capita	Social support	Healthy life expectancy	Freedom to make life choices	Generosity	Perceptions of corruption	Regional indicator
0	Afghanistan	2008	3.724	7.370	0.451	50.80	0.718	0.168	0.882	South Asia
1	Afghanistan	2009	4.402	7.540	0.552	51.20	0.679	0.190	0.850	South Asia
2	Afghanistan	2010	4.758	7.647	0.539	51.60	0.600	0.121	0.707	South Asia
3	Afghanistan	2011	3.832	7.620	0.521	51.92	0.496	0.162	0.731	South Asia
4	Afghanistan	2012	3.783	7.705	0.521	52.24	0.531	0.236	0.776	South Asia

## Descriptions

The World Happiness Report relies on data collected from the Gallup World Poll surveys, an annual survey that started in 2005 and covers supposedly 99 per cent of the world's population. In this project, I will be working with the datasets spanning from 2005 to 2021. 2021 is the most recent one with the happenings of COVID-19, while the previous years, 2005 to 2020 are before.

Our main outcome is dependent on the comparison of the main life evaluation question asked in the poll, the so-called *Cantril ladder* (in the dataset it is represented as Ladder score). The basis of the Cantril ladder basically tells respondents to imagine themselves on a ladder with steps numbered from zero (the bottom) to ten (the top) where zero is the worst and ten being the best possible life. As one can tell, Cantril ladder is an evaluative measure of subjective well-being.

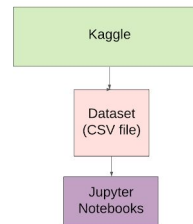
In addition to the Ladder score, there are other factors to explain life evaluations. The columns demonstrate happiness scores through six factors: economic production (log GDP per capita), social support, life expectancy, freedom, absence of corruption, and generosity. They have been deemed as contributing variables to the impact of happiness and of importance within research literature, as it explains national-level differences in life evaluations.

Economic production column is self-explanatory, it is the log GDP per capita. Social support is based on the national average response (Yes = 1, No = 0) to the question: "If you were in trouble, do you have relatives or friends you can count on to help you whenever you need them, or not?". Life expectancy is constructed based on data from the World Health Organization (WHO) and World Development Indicators (WDI). Freedom is determined by the freedom to make life choices. It takes the national average of the binary responses to the question: "Are you satisfied or dissatisfied with your freedom to choose what you do with your life?". Generosity is the residual of regressing the national average responses to the question "Have you donated money to a charity in the past month?" on GDP per capita. Absence of corruption are the average of binary answers to two questions: "Is corruption widespread throughout the government or not?" and "Is corruption widespread within businesses or not?" Where data for government corruption are missing, the perception of business corruption is used as the overall corruption-perception measure.

## Methods & Tools

Kaggle was used as the tool to obtain the dataset. Kaggle is an online crowd-sourced platform for data scientists to find and publish data sets, work with others, and to learn and explore predictive analytics problems.

To do the data analysis, I have decided to use Jupyter Notebook as my means of creating graphs through code. Jupyter Notebook is a server-client application that allows editing and running Notebook documents. The Notebook documents contain computer code (like Python) and rich text elements (such as paragraphs, equations, links, etc.). It can be executed locally (requiring no internet access) and it includes the packages I am going to use such as "pandas", "matplotlib", "numpy", and "seaborn". Pandas is used to import the dataset (CSV file) and for data analysis. Matplotlib and Seaborn are being used for graphing the necessary data charts. Numpy is for data manipulation and other miscellaneous usage.



## Research Question:

**How is happiness after the spread of COVID-19 compared to the previous years?**

## Sub-Questions:

- 1) What does happiness per year look like?
- 2) What are the happiness levels per region in comparison to before and after COVID-19?
- 3) What are the differing factors between a happy and a less happy country?
- 4) How does the world's subjective well-being fare after the COVID-19 pandemic (in comparison to before)?

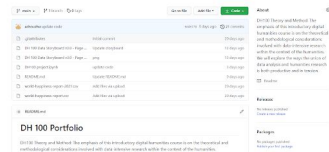
## Potential Results

The hypothetical outcome is that the happiness levels should have decreased worldwide due to the pandemic. The comparison of the tables in other years to the one relevant to the pandemic should display lower happiness levels and other categories that correlate with unhappiness.

## GitHub Repository

Within my GitHub repository, all the datasets and code used to produce the graphs will be provided. There will also be additional resources included in the repository to better understand my project such as presentation slides and a video walkthrough..

Click here or copy and paste the link below:  
[https://github.com/azhouzhu/DH100\\_Summer2021](https://github.com/azhouzhu/DH100_Summer2021)



# Video

<Upload playable youtube version of video>

# Interpretations & Conclusions / Takeaways

**Interpreting the results:** From the visualizations, happiness before and after has not changed in a drastic manner whether positive or negative. In some regions, they have become happier while others have lowered slightly. However, overall there is no significant change on the happiness after the spread of COVID-19. The factors to happiness have also not demonstrated major changes in the before and after data. The factors, however, help explain the ladder score average of each country and how they are associated.

**Conclusions:** From my research, worldwide issues such as pandemics are not directly associated to the happiness levels of the world. Analyzing the life satisfaction rating (ladder score) by country has produced no concrete evidence of happiness being affected by COVID-19. If there is significant changes, it is not within the factors I have chosen to study.

**Next Steps / Future Work:** The dataset "World Happiness Report" can potentially be used for more than analyzing life satisfaction. It has the potential to be used to further research how economics of different countries affect life expectancy. Through a common pattern from the research, it could lead to understanding why some countries underperform and determine if it could be a causation to life satisfaction scores from each country.



# Works Cited & Acknowledgements

- **Dataset:**

- [World Happiness Report 2005 - 2021](#)

- **Helpful Links:**

- [DH100 GitHub Repository demo](#)
- [World Happiness Report analysis](#)
- [Pandas User Guide](#)
- [EDA - World Happiness data - Deals multiple data](#)
- [Practical Statistics & Visualization With Python & Plotly](#)
- [World Happiness Report Analysis 2018](#)
- [Analysis Between World Happiness and City Happiness 2020](#)