

World Happiness Report

2019 - 2021

**Clovis Lacerda
Kinjal Maru
Matthew Granson**

Motivation and Interest

We felt it would be interesting to look at the relative happiness scores from countries around the world and assess the impact Covid might have had on reported happiness measures.

While it would be informative to look back to 2008, we feel it would be better to examine and compare measures from just the past three years to present a targeted analysis regarding Covid's impact on the world specifically.

Questions and Data

Datasets were sourced from Kaggle and Data.world.

Raw response data was generated from the Gallup World Poll which is distributed and sampled before March of each year.

For this case study we wanted to look at a number of dependent variables for comparison:

- “Happiness” or ladder score
- Life Expectancy
- Social Support

Operational Definitions: Happiness

Happiness or Ladder Score

- Ladder Score and Subjective Well Being (SWB)
- National average response to questions of life evaluation
- Coded as 0 - 10 with ten being the highest, or best, quality of life

“Please imagine a ladder, with steps numbered from 0 at the bottom to 10 at the top. The top of the ladder represents the best possible life for you and the bottom of the ladder represents the worst possible life for you. On which step of the ladder would you say you personally feel you stand at this time?”

English wording of the Gallup World Poll

Operational Definitions: Life Expectancy

(Healthy) Life Expectancy:

- Scores extracted from the World Health Organization's (WHO) Global Health Observatory data repository
- Contains an estimate of how long a person of a specific region/country will live starting from birth



Operational Definitions: Social Support

Social Support:

- This question examines an individual's accessibility to social support/resources
- It is coded as a binary response (true or false, 1 or 0)



“If you were in trouble, do you have relatives or friends you can count on to help you whenever you need them, or not?”

English wording of the Gallup World Poll

Data Conversion and Cleanup

PostgreSQL Methodology:

- Used PgAdmin to create two tables.
- Inserted values from Excel input files.
- Unioned the two tables - necessary fields and new fields.
- Used SQLAlchemy in python to call the query from SQL.
- Created a dataframe for downstream visualizations.

Data Visualization

Used Pandas to take segments of the complete data frames to create graphs based on certain variables, grouped by region and year.

```
# Group regions and average ladder score for each region, split into 3 years, then merged into 1 data frame
region_ladder = HappinessPD1.groupby(["Regional_Indicator", "year"])["Life_Ladder"].mean()
region_ladder_df = pd.DataFrame(region_ladder)
region_ladder_df = region_ladder_df.sort_values(["Regional_Indicator", "year"], ascending=True)
region_ladder_df
```

```
pivot = region_ladder_df.pivot_table(index="Regional_Indicator", columns="year", values="Life_Ladder")
```

	year	2019	2020	2021
Regional_Indicator				
Central and Eastern Europe		5.885063	6.076688	5.984765
Commonwealth of Independent States		5.530750	5.641571	5.467000
East Asia		5.785667	5.956500	5.810333
Latin America and Caribbean		6.117579	5.660182	5.908050

```
pivot.plot(kind="bar", figsize=(15,8), xlabel="Region", ylabel="Ladder Score", title="Happiness By Country Per Year")
plt.show()
```

Analysis: Happiness or Ladder Score

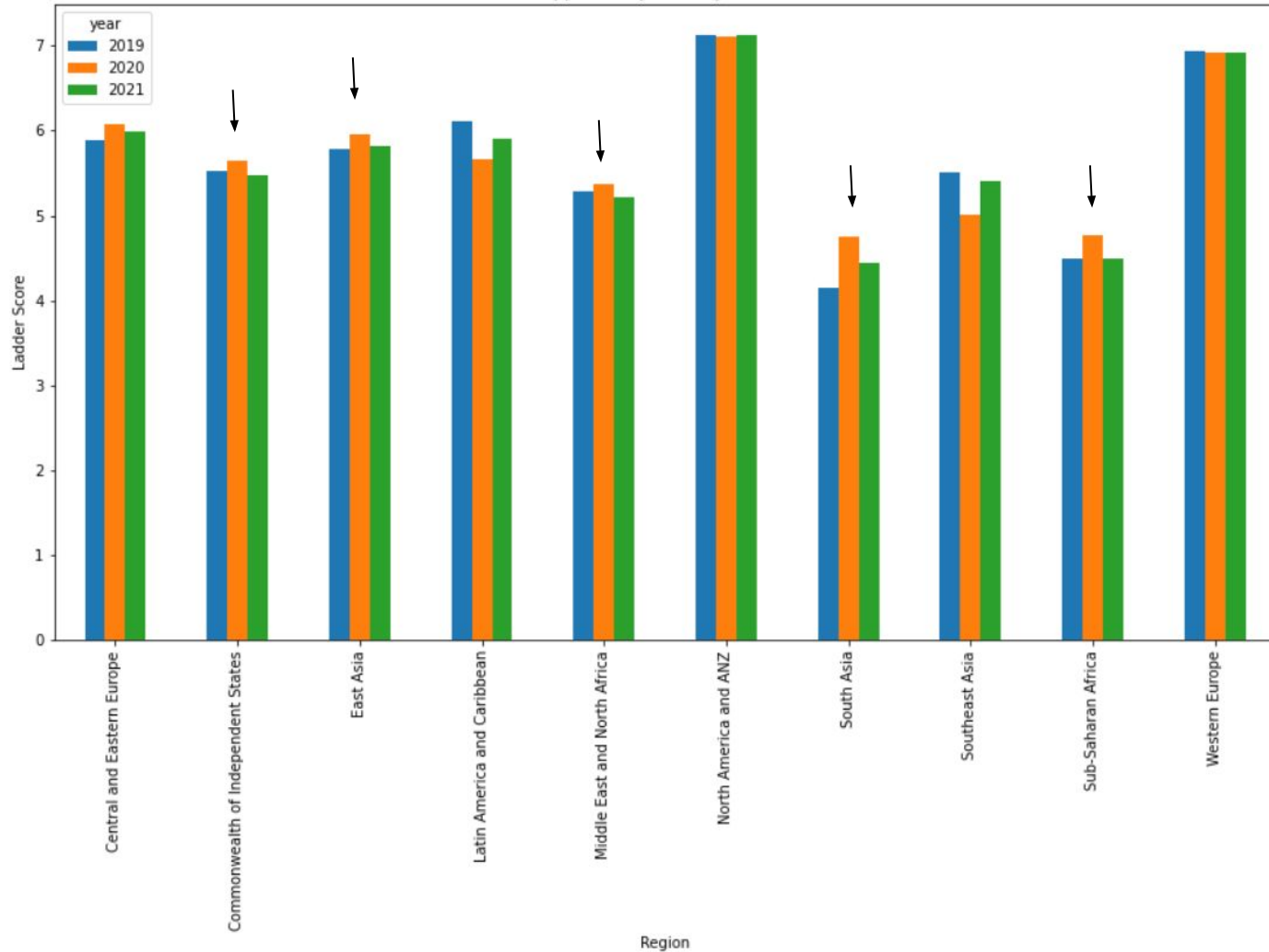
		Life_Ladder
Regional_Indicator	year	
Central and Eastern Europe	2019	5.885063
	2020	6.076688
	2021	5.984765
Commonwealth of Independent States	2019	5.530750
	2020	5.641571
	2021	5.467000
East Asia	2019	5.785667
	2020	5.956500
	2021	5.810333
Latin America and Caribbean	2019	6.117579
	2020	5.660182
	2021	5.908050
Middle East and North Africa	2019	5.288625
	2020	5.363455
	2021	5.219765
North America and ANZ	2019	7.123000
	2020	7.111750
	2021	7.128500
South Asia	2019	4.140500
	2020	4.752500
	2021	4.441857
Southeast Asia	2019	5.504333
	2020	5.011400
	2021	5.407556
Sub-Saharan Africa	2019	4.499314
	2020	4.761571
	2021	4.494472
Western Europe	2019	6.940905
	2020	6.919421
	2021	6.914905

Analysis: Happiness Ladder Score

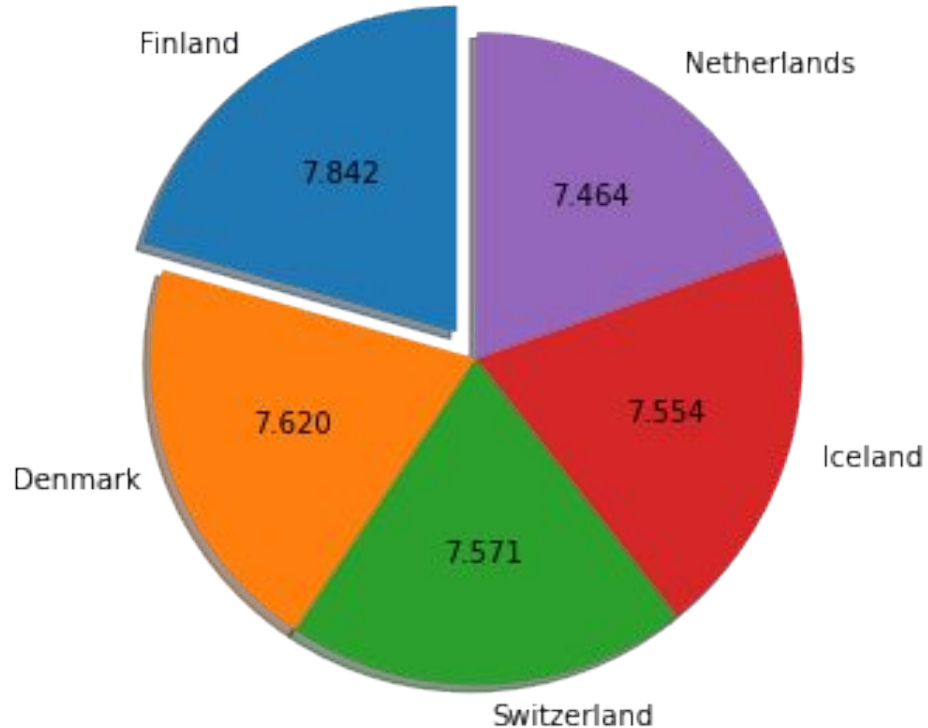
Most noteworthy observations:

- ↓ ● Central and Eastern Europe
 - Decrease from 2020 to 2021 (6.1 to 5.9)
- ↓ ● Sub Saharan Africa
 - Decrease from 2019 to 2021 (4.8 to 4.4)
- ↓ ● South Asia
 - Decrease from 2020 to 2021 (4.8 to 4.4)
- ↓ ● Middle East and North Africa
 - Decrease from 2020 to 2021 (5.3 to 5.2)
- ↓ ● East Asia
 - Decrease from 2020 to 2021 (6.0 to 5.8)

Happiness By Country Per Year



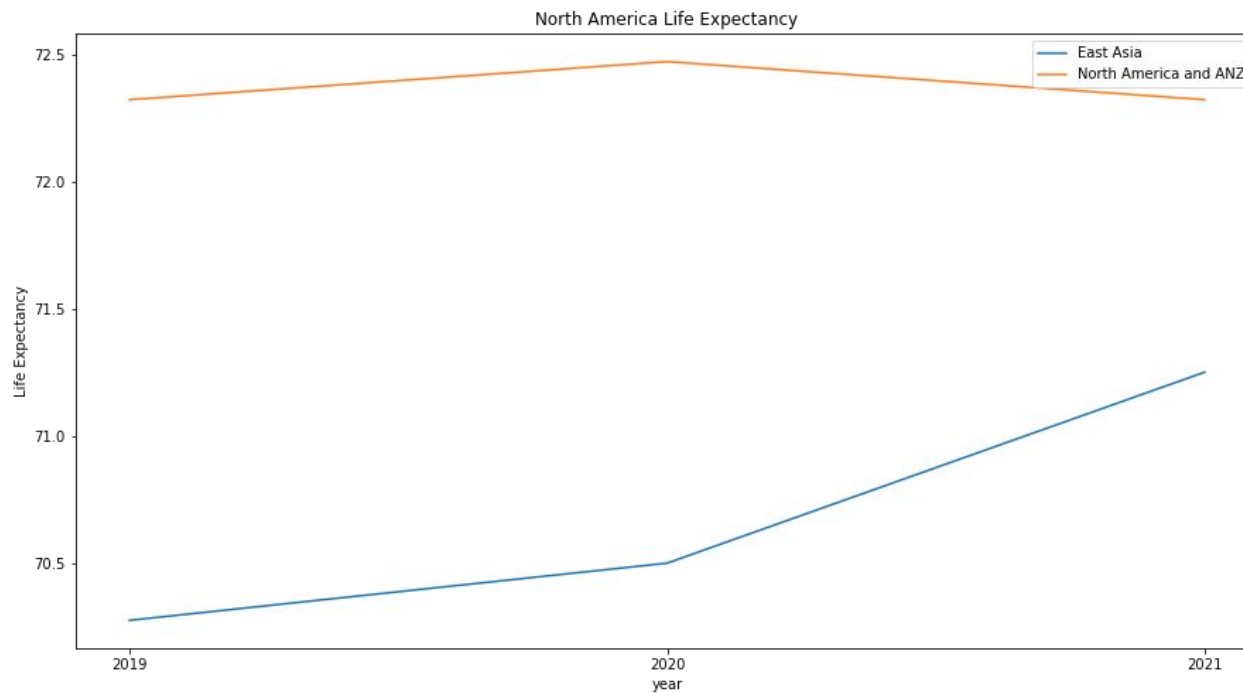
Top 5 Happiest Countries By Ladder Score

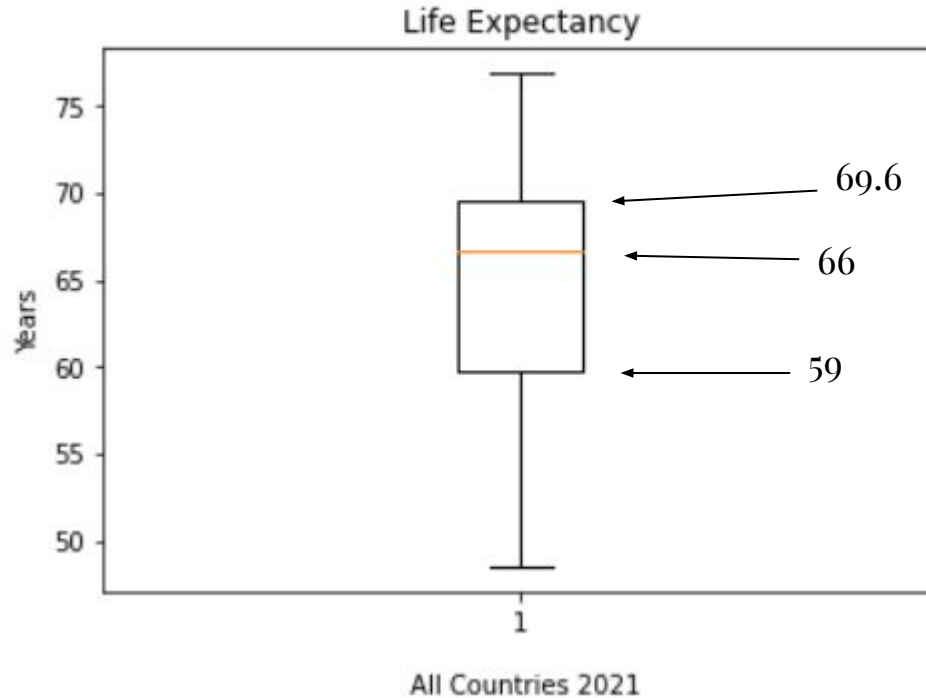


Analysis: Life Expectancy

	year	2019	2020	2021
Regional_Indicator				
Central and Eastern Europe	68.484933	69.044000	68.338412	
Commonwealth of Independent States	65.083333	65.142857	65.009500	
East Asia	70.275000	70.500000	71.252167	
Latin America and Caribbean	67.710526	67.800000	67.076050	
Middle East and North Africa	66.200000	66.990909	65.609118	
North America and ANZ	72.325000	72.475000	72.325000	
South Asia	61.433333	63.100000	62.681000	
Southeast Asia	64.944444	62.240000	64.888444	
Sub-Saharan Africa	56.127714	57.064286	55.886472	
Western Europe	72.990000	73.189474	73.033095	

Analysis: Life Expectancy





The lower quartile of ratings is:
59.802

The upper quartile of ratings is: 69.6

The interquartile range of ratings is:
9.797999999999995

The the median of ratings is: 66.603

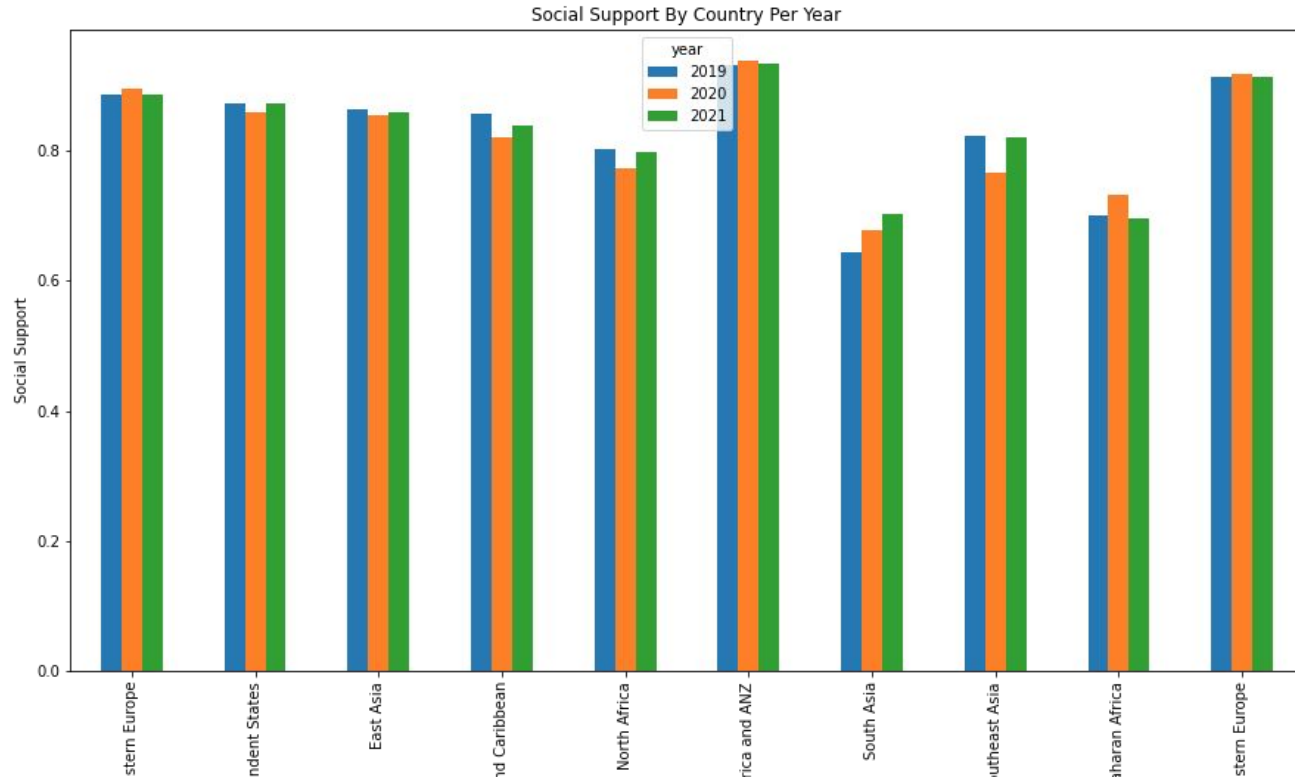
Values below 45.105000000000004
could be outliers.

Values above 84.296999999999998
could be outliers.

Analysis: Social Support

Regional_Indicator		Social support	
		year	
Central and Eastern Europe	2019		0.885813
	2020		0.895938
	2021		0.887412
Commonwealth of Independent States	2019		0.872333
	2020		0.860286
	2021		0.872500
East Asia	2019		0.863000
	2020		0.855833
	2021		0.860500
Latin America and Caribbean	2019		0.856105
	2020		0.820818
	2021		0.839500
Middle East and North Africa	2019		0.802437
	2020		0.772727
	2021		0.797647
North America and ANZ	2019		0.931000
	2020		0.939250
	2021		0.933500
South Asia	2019		0.643000
	2020		0.678000
	2021		0.703429
Southeast Asia	2019		0.824000
	2020		0.765600
	2021		0.820333
Sub-Saharan Africa	2019		0.700200
	2020		0.733429
	2021		0.696750
Western Europe	2019		0.913762
	2020		0.919316
	2021		0.914476

Analysis: Social Support



Discussion

Originally we had planned to look at data ranging back to 2008, but there were several issues:

- The dataset was expansive and contained 140+ countries
- For better organization, we decided to look at the world regions rather than every country
- To better capture the effect of the pandemic, we limited it to only include the years 2019, 2020, and 2021.

Conclusion Happiness

Based on the data:

- A majority of world region scores displayed a noticeable decrease in happiness when transitioning from 2020 to 2021
- However, some regions displayed an increase in happiness score.

Regions may be skewed depending on the size of each country within them.

Regions with larger countries/denser populations will outweigh smaller countries

Sociopolitical elements such as war and terrorism may heavily influence average reported score of a region.

Conclusion Life Expectancy

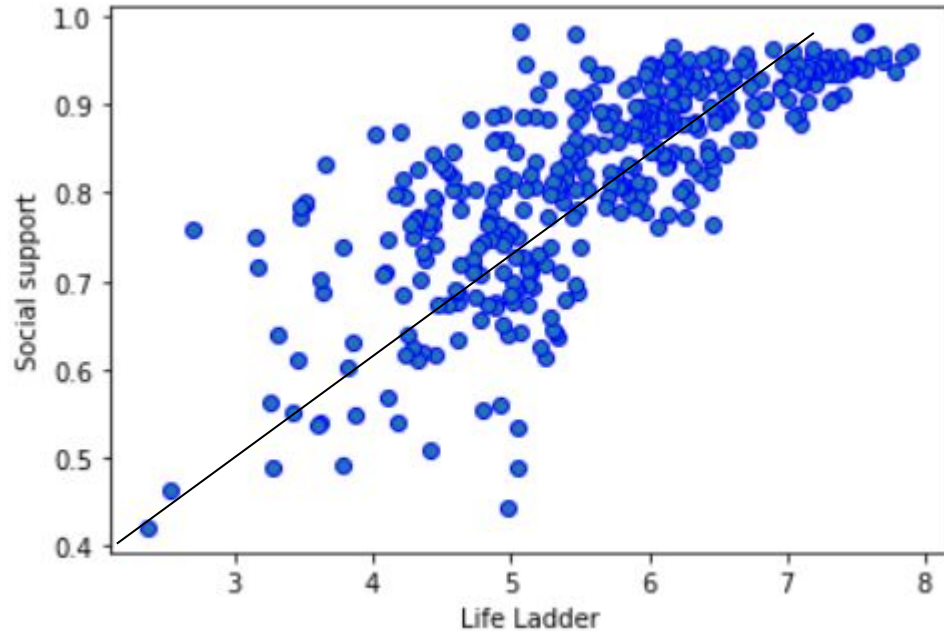
Based on the data:

- The reported life expectancy for every region displayed a noticeable dip between 2020 and 2021
- While there may be other elements to consider for this decline, with high confidence, it seems safe to assume this was a direct result from Covid 19

Conclusion Social Support

Based on the data:

- To show the importance of social support, here is a scatterplot showing the relationship between “happiness” and social support.
- Note the strong upward trend.



Questions?

References

<https://www.kaggle.com/ajaypalsinghlo/world-happiness-report-2021?select=world-happiness-report.csv>

<https://data.world/promptcloud/world-happiness-report-2019>

<https://happiness-report.s3.amazonaws.com/2021/Appendix1WHR2021C2.pdf>