*An Analysis of Happiness Pre and Post-COVID-19*

PSTAT 100 Course Project

Maya Hetz | Amir Voloshin | Sarah Feuer

# Data Description

The World Happiness Report 2023 aims to measure global happiness by how happy their citizens perceive themselves to be. About 1,000 people from over 150 countries are asked annually a series of questions related to happiness, and The World Happiness Report 2023 quantifies this data - weighted by population - by country from 2008-present. The observational units are countries, and some variables observed on each unit are year, life ladder, log GDP per capita, life expectancy, positive effect, and negative effect. Life Ladder is an averaged index to the question: "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?" Positive effect is an average index of three positive emotions: laugh, enjoyment and doing interesting things, on a particular day i.e. "Did you smile or laugh a lot yesterday?" Similarly, negative effect averaged the three emotions: worry, sadness and anger with similar sampling questions. The focus of this write-up uses these three variables mainly as measures of happiness.

Our project specifically focuses on three countries: the United States, Israel, and China. We chose these countries for various reasons. Our question of interest relates to the recent global crisis, the COVID-19 pandemic. The US is a good candidate to answer our question of interest because as its citizens we are familiar with its internal policies and regulations and this can be helpful in our data analysis. Additionally, Israel is another leading first-world country that our group members personally connect to so we thought it would add another interesting and personal perspective to our investigation. Lastly, China is a country who took drastic measures during that era, so it will be interesting to see how those extreme conditions affected national happiness.

## Question of Interest

How did the COVID-19 pandemic affect overall happiness in the United States, Israel, and China? What similarities and differences are there?

## ▾ Packages & Additional Utilities

```
1 # Run this to acces your drive for the Data
2 from google.colab import drive
3 drive.mount('/content/drive')
```

```
Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_rem
```

```
1 # For Rendering at the end
2 #%%shell
3 #jupyter nbconvert --to html /content/mp2_ncca.ipynb
```

```
1 import pandas as pd
2 import numpy as np
3 import altair as alt
4 import statsmodels.api as sm
5 import warnings
6 warnings.filterwarnings('ignore')
7
8 # We don't need this for google colab, just adding incase we render through Jypter Hub
9 #alt.data_transformers.disable_max_rows()
10 #alt.renderers.enable('mimetype')
```

## ▾ Data

```
1 happy = pd.read_csv('/content/drive/MyDrive/Final Project - 100/whr-2023.csv')
2 happy.head()
```

|   | Country name | year | Life Ladder | Log GDP per capita | Social support | Healthy life expectancy at birth | Freedom to make life choices | Generosity | Percept corrup |
|---|---|---|---|---|---|---|---|---|---|
| **0** | Afghanistan | 2008 | 3.724 | 7.350 | 0.451 | 50.5 | 0.718 | 0.168 | |
| **1** | Afghanistan | 2009 | 4.402 | 7.509 | 0.552 | 50.8 | 0.679 | 0.191 | |
| **2** | Afghanistan | 2010 | 4.758 | 7.614 | 0.539 | 51.1 | 0.600 | 0.121 | |
| **3** | Afghanistan | 2011 | 3.832 | 7.581 | 0.521 | 51.4 | 0.496 | 0.164 | |

## ▾ Comparing Life Ladder, Positive affect, and Negative affect over time in each Country

### ▾ Israel

```
1  # Selecting Israel's Data
2  happy_israel = happy[happy['Country name'] == 'Israel']
3
4  # To add a vertical line at year 2020
5  vertical_line = alt.Chart({'values': [{'x': 2020}]}).mark_rule(opacity = 0.5, color = 'red').encode(x = 'x:Q',)
6
7  # Plot figures
8  life_ladder = alt.Chart(happy_israel).mark_line(point={"fill": "salmon"}, color = 'salmon').encode(
9      x = alt.X('year', title = 'Year', scale = alt.Scale(domain = (2000, 2025))),
10     y = alt.Y('Life Ladder', scale = alt.Scale(domain = (6, 8))),
11 ).properties(title = 'Life Ladder vs Year in Israel', width = 250, height = 250) + vertical_line
12
13 positive = alt.Chart(happy_israel).mark_line(point={"fill": "blue"}, color='blue').encode(
14     x = alt.X('year', title = 'Year', scale = alt.Scale(domain = (2000, 2025))),
15     y = alt.Y('Positive affect', title='Positive Affect', scale = alt.Scale(domain = (0, 1))),
16 ).properties(title = 'Positive affect vs Year in Israel', width = 250, height = 250)  + vertical_line
17
18 negative = alt.Chart(happy_israel).mark_line(point={"fill": "green"}, color = 'green').encode(
19     x = alt.X('year', title = 'Year', scale = alt.Scale(domain = (2000, 2025))),
20     y = alt.Y('Negative affect', title = 'Negative Affect', scale = alt.Scale(domain = (0, 1))),
21 ).properties(title = 'Negative affect vs Year in Israel', width = 250, height = 250)  + vertical_line
22
23 # Viewing
24 # pos_neg = alt.layer(positive, negative)
25 alt.hconcat(life_ladder , positive, negative)
```
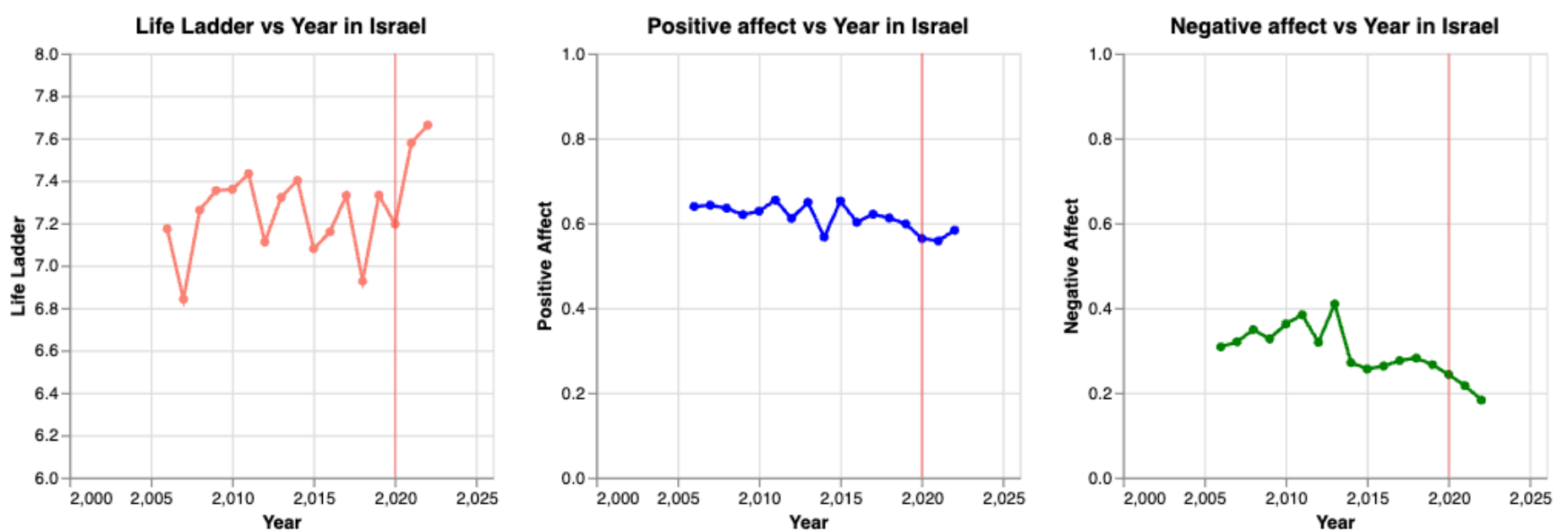


### ▾ United States

```
1  # Selecting United States Data
2  happy_us = happy[happy['Country name'] == 'United States']
3
4  # Plot figures
5  life_ladder_us = alt.Chart(happy_us).mark_line(point={"fill": "salmon"}, color = 'salmon').encode(
6      x = alt.X('year', title = 'Year', scale = alt.Scale(domain = (2000, 2025))),
7      y = alt.Y('Life Ladder', scale = alt.Scale(domain = (6, 8))),
8  ).properties(title = 'Life Ladder vs Year in the US', width = 250, height = 250) + vertical_line
9
10 positive_us = alt.Chart(happy_us).mark_line(point={"fill": "blue"}, color='blue').encode(
```

```
11      x = alt.X('year', title = 'Year', scale = alt.Scale(domain = (2000, 2025))),
12      y = alt.Y('Positive affect', title='Positive Affect', scale = alt.Scale(domain = (0, 1))),
13 ).properties(title = 'Positive affect vs Year in the US', width = 250, height = 250)  + vertical_line
14
15 negative_us = alt.Chart(happy_us).mark_line(point={"fill": "green"}, color = 'green').encode(
16      x = alt.X('year', title = 'Year', scale = alt.Scale(domain = (2000, 2025))),
17      y = alt.Y('Negative affect', title='Negative Affect', scale = alt.Scale(domain = (0, 1))),
18 ).properties(title = 'Negative affect vs Year in the US', width = 250, height = 250)  + vertical_line
19
20 # Viewing
21 # pos_neg = alt.layer(positive, negative)
22 alt.hconcat(life_ladder_us , positive_us, negative_us)
```
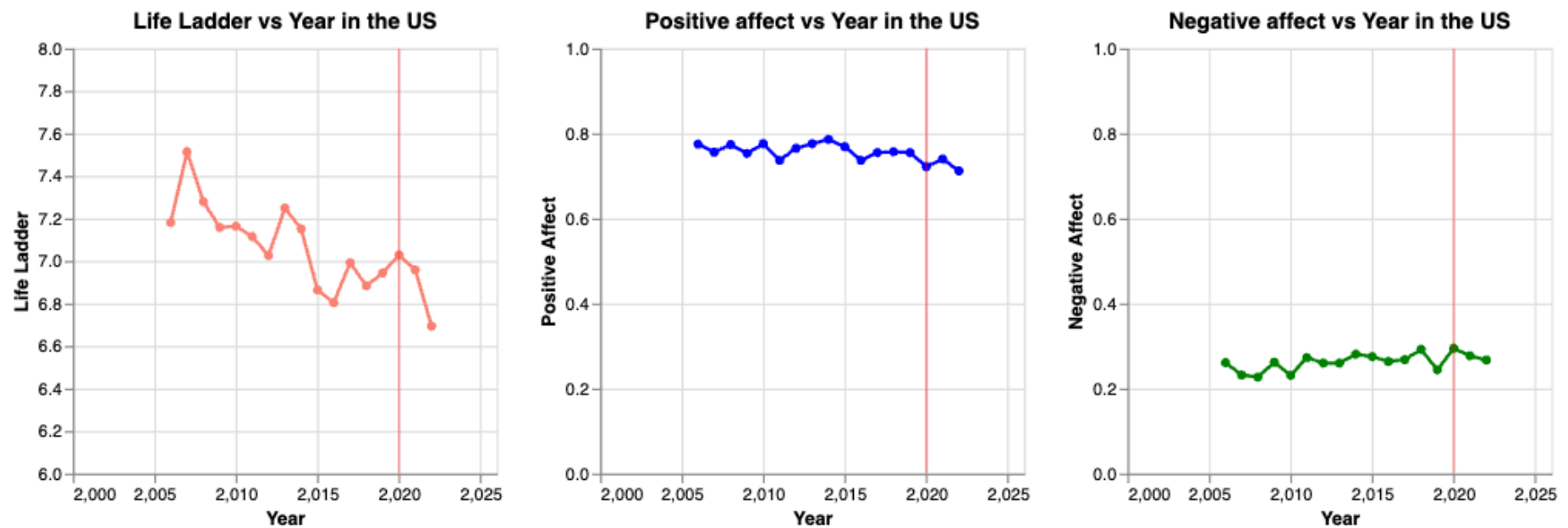


## China

```
 1 # Selecting China's Data
 2 happy_china = happy[happy['Country name'] == 'China']
 3
 4 # Plot figures
 5 life_ladder_china = alt.Chart(happy_china).mark_line(point={"fill": "salmon"}, color = 'salmon').encode(
 6      x = alt.X('year', title = 'Year', scale = alt.Scale(domain = (2005, 2022))),
 7      y = alt.Y('Life Ladder', scale = alt.Scale(domain = (3, 10))),
 8 ).properties(title = 'Life Ladder vs Year in China', width = 250, height = 250) + vertical_line
 9
10 positive_china = alt.Chart(happy_china).mark_line(point={"fill": "blue"}, color='blue').encode(
11      x = alt.X('year', title = 'Year', scale = alt.Scale(domain = (2005, 2022))),
12      y = alt.Y('Positive affect', title='Positive Affect', scale = alt.Scale(domain = (0, 1))),
13 ).properties(title = 'Positive affect vs Year in China', width = 250, height = 250)  + vertical_line
14
15 negative_china = alt.Chart(happy_china).mark_line(point={"fill": "green"}, color = 'green').encode(
16      x = alt.X('year', title = 'Year', scale = alt.Scale(domain = (2005, 2022))),
17      y = alt.Y('Negative affect', title='Negative Affect', scale = alt.Scale(domain = (0, 1))),
18 ).properties(title = 'Negative affect vs Year in China', width = 250, height = 250)  + vertical_line
19
20 # Viewing
21 # pos_neg = alt.layer(positive, negative)
22 alt.hconcat(life_ladder_china, positive_china, negative_china)
```
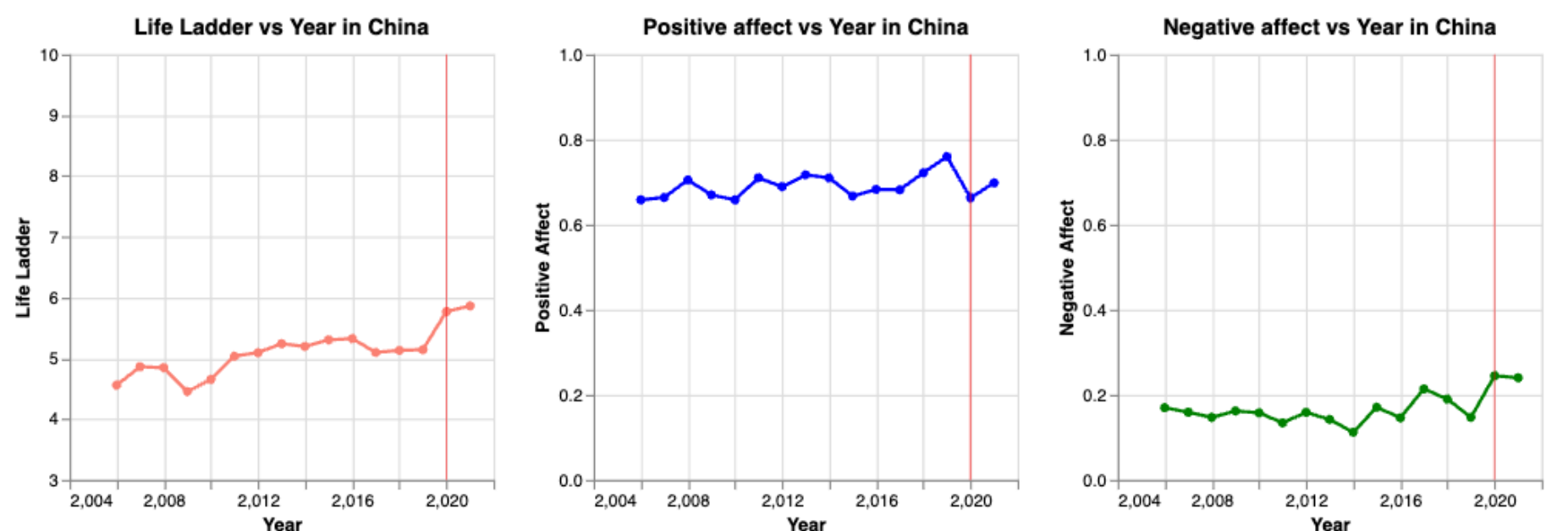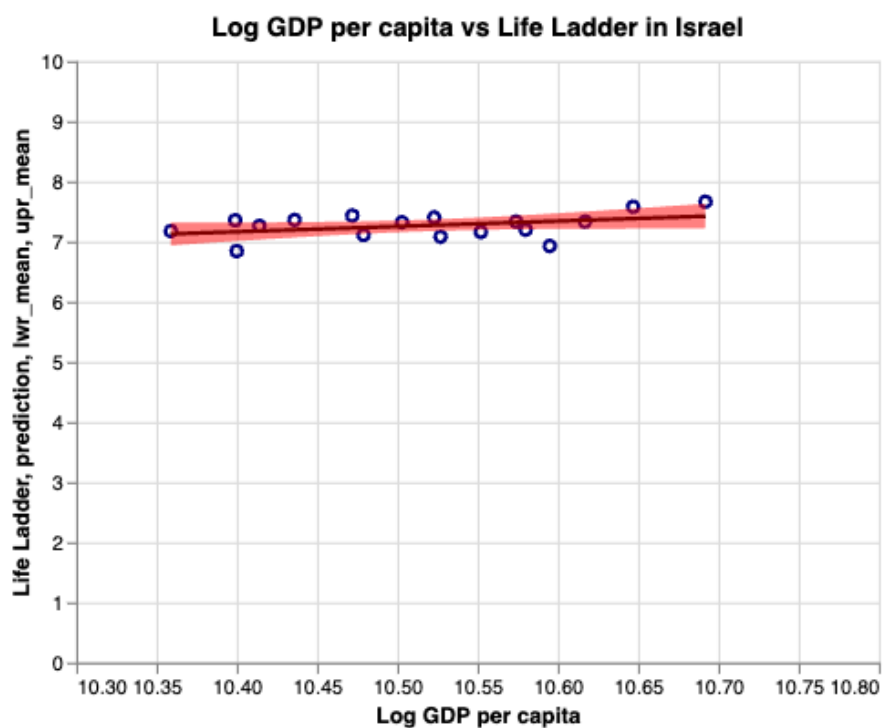


## Predicting Life Ladder from log GDP in Israel

```
 1  # retrieve response
 2  y = happy_israel['Life Ladder']
 3
 4  # construct explanatory variable matrix
 5  x = sm.tools.add_constant(happy_israel['Log GDP per capita'])
 6
 7  # fit model
 8  slr = sm.OLS(endog = y, exog = x)
 9  rslt = slr.fit()
10
11  # Adding fitted values to dataframe
12  happy_israel['fitted_vals'] = rslt.fittedvalues
13
14  # Predicting
15  preds = rslt.get_prediction(x)
16
17  # Adding confidence interval and predictions to dataframe
18  happy_israel['lwr_mean'] = preds.predicted_mean - 2*preds.se_mean
19  happy_israel['upr_mean'] = preds.predicted_mean + 2*preds.se_mean
20  happy_israel['prediction'] = preds.predicted_mean
21
22  # constructing linear reg plot
23  slr_line = alt.Chart(happy_israel).mark_line(color = 'black').encode(
24      x = 'Log GDP per capita',
25      y = 'prediction'
26  )
27
28  # constructing a confidence band
29  band = alt.Chart(happy_israel).mark_area(opacity = 0.5, color = 'red').encode(
30      x = 'Log GDP per capita',
31      y = 'lwr_mean',
32      y2 = 'upr_mean'
33  )
34
35  # Log GDP per capita vs Life Ladder
36  plot = alt.Chart(happy_israel).mark_point(opacity = 1, color = 'navy').encode(
37      x = alt.X('Log GDP per capita', scale = alt.Scale(domain = (10.3, 10.8))),
38      y = alt.Y('Life Ladder', scale = alt.Scale(domain = (0, 10)))
39  ).properties(title = 'Log GDP per capita vs Life Ladder in Israel')
40
41  # layering
42  plot + slr_line + band
```



The black line in the plot above is the linear regression model which was trained on log GDP per capita, and predicts the Life Ladder. The red highlighting is the confidence interval. Given this model was trained on Israels data and is predicting those exact points there is a very high confidence on the prediction of the points and hence the small confidence interval. This is an indication that a linear model is very suitable for the data we have.

## ▾ Predicting China's Life Ladder based on log GDP

```
 1  # Adding constants to China explanatory variable
 2  x_us = sm.tools.add_constant(happy_us['Log GDP per capita'])
```

```
 3
 4 # Predicting
 5 preds_us = rslt.get_prediction(x_us)
 6
 7 # Adding confidence interval and predictions to dataframe
 8 happy_us['lwr_mean'] = preds_us.predicted_mean - 2*preds.se_mean
 9 happy_us['upr_mean'] = preds_us.predicted_mean + 2*preds.se_mean
10 happy_us['prediction'] = preds_us.predicted_mean
11
12 # constructing linear reg plot
13 slr_line_us = alt.Chart(happy_us).mark_line(color = 'black').encode(
14     x = 'Log GDP per capita',
15     y = 'prediction'
16 )
17
18 # Constructing a confidence band
19 band_us = alt.Chart(happy_us).mark_area(opacity = 0.5, color = 'red').encode(
20     x = 'Log GDP per capita',
21     y = 'lwr_mean',
22     y2 = 'upr_mean'
23 )
24
25 # Log GDP per capita vs Life Ladder
26 plot_us = alt.Chart(happy_us).mark_point(opacity = 1, color = 'navy').encode(
27     x = alt.X('Log GDP per capita', scale = alt.Scale(domain = (10.8, 11.2))),
28     y = alt.Y('Life Ladder', scale = alt.Scale(domain = (5, 10)))
29 ).properties(title = 'Log GDP per capita vs Life Ladder in the US')
30
31 # layering
32 plot_us + slr_line_us + band_us
```



Log GDP per capita vs Life Ladder in the US
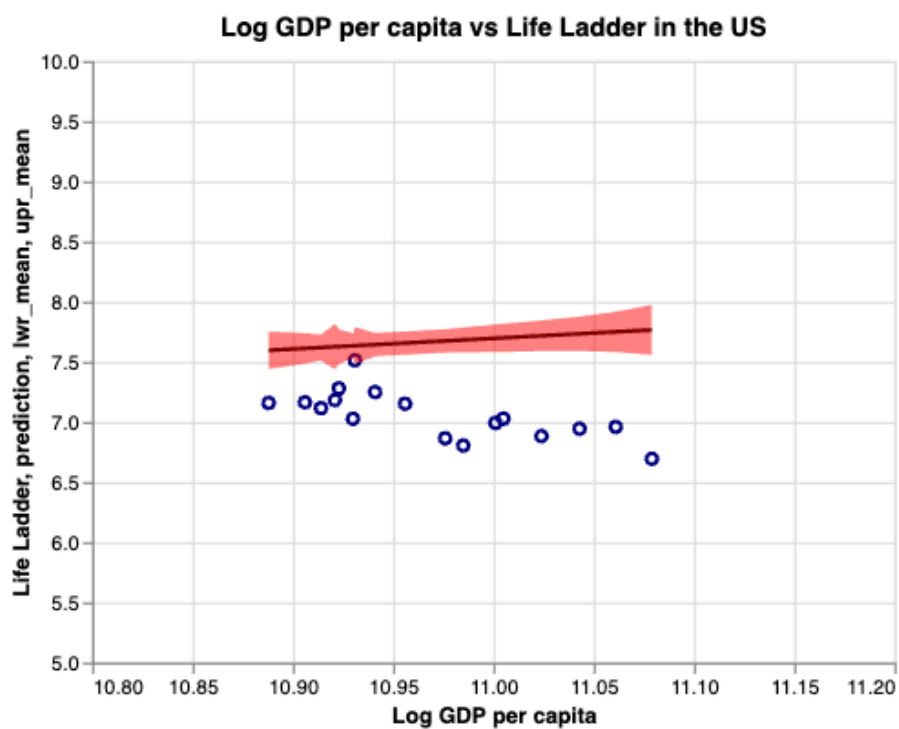
In the plot above we predicted the life ladder in the United States based on the Log GDP per capita. The black line is the linear regression model and it is obvious the model is not very accurate. This is likely because the linear regression model used was trained on Israel's data rather than the United States. This result tells us that a model trained on one countries data can be extremely inaccurate in predicting data points for another country. Furthermore, the wider confidence interval band in red indicates a less confident model as opposed to the previous model.

## ▾ Data Analysis

**Introduction**

In this section, we will analyze the impact of the COVID-19 pandemic on overall happiness in the United States, Israel, and China. By examining the happiness scores and emotional experiences before, during, and after the pandemic, we aim to identify any similarities and differences in the effects of the pandemic on happiness across these countries.

**Happiness Scores Analysis (Life Ladder)**

The analysis of happiness scores before and during the pandemic reveals interesting insights about the impact of COVID-19 on overall happiness.

In the United States, the happiness scores remained relatively stable from 2006 to 2019. However, a slight decline in happiness scores can be observed in 2020, coinciding with the outbreak and subsequent effects of the COVID-19 pandemic. Although the scores recovered slightly in

2021 and 2022, they did not reach the pre-pandemic levels. This suggests that the pandemic had a negative impact on overall happiness in the United States. This trend is clearlu visible in the graphics provided in the United States section above.

Israel, on the other hand, experienced a different trend. The happiness scores in Israel remained consistently high throughout the years, including during the COVID-19 pandemic. There was no significant deviation from the pre-pandemic levels, indicating that the pandemic did not substantially affect overall happiness in Israel. The line charts show this interesting pattern by remaining relatively stable over time, or at least having only subtle changes from year to year.

China's happiness scores also remained relatively stable throughout the pandemic period. Although there was a slight decrease in 2020, the scores quickly rebounded in 2021 and 2022, suggesting resilience and a relatively minor impact of the pandemic on overall happiness in China. This was a pretty interesting finding considering what our perception of China was in relation to its strict COVID restrictions.

Graphics predicting each countries future life ladder scores based on log GDP can provide further interesting insight about what the post-pandemic world's happiness levels could look like.

**Emotional Experiences Analysis (Positive and Negative Affect)**

Examining the positive and negative affect during the pandemic period provides further insights into the specific emotional experiences within each country.

In the United States, both positive and negative affect showed noticeable changes during the pandemic. The positive affect scores experienced a decline in 2020, indicating a reduction in positive emotions among the population. Similarly, the negative affect scores showed an increase, suggesting a rise in negative emotional experiences during the pandemic.

Israel, however, displayed a different pattern. The positive affect scores remained consistently high, with no significant changes observed during the pandemic. Likewise, the negative affect scores remained low, indicating a limited presence of negative emotions throughout the pandemic period.

China's positive affect scores showed a slight decline in 2020, indicating a temporary decrease in positive emotions during the pandemic. However, the scores quickly recovered in 2021 and 2022. The negative affect scores remained consistently low, suggesting that negative emotional experiences were limited in China during the pandemic.

Again, all of these findings are visible in the line charts provided in each country's section in this report. The clear presenations make the trends described above very easy to see.

**Similarities and Differences**

Analyzing the impact of the COVID-19 pandemic on overall happiness in the United States, Israel, and China reveals both similarities and differences among these countries.

Similarities: The United States and Israel both experienced a decline in happiness scores during the pandemic, however to slightly different extents. The positive affect scores in the United States and China showed a decline in 2020, indicating a temporary decrease in positive emotions during that wild year. However, China, Israel, and the United States all showed resilience and a relatively quick recovery in happiness scores and both positive and negative affect almost immediately after the initial impact of the pandemic.

Differences: Israel's happiness scores remained consistently high, suggesting a minimal impact of the pandemic on overall happiness as measured by the World Happiness Report 2023. The United States experienced a more significant decline in happiness scores as well as a noticeable increase in negative affect during the pandemic. China displayed a relatively minor impact on happiness scores, with a only temporary decline in positive affect during the pandemic.

**Limitations**

A quick note on how this report and its findings could potentially be limited and what factors could lead to the trends we are viewing is neccessary before wrapping up with a conclusion. It is important to acknowledge that the analysis of the COVID-19 pandemic's impact on happiness is complex, influenced by various factors beyond the scope of this analysis. Factors such as specific government policies, healthcare systems, socioeconomic conditions, and cultural differences can contribute to the differences we see among the countries. Also, perhaps the way that the questions for each variable are phrased could be affecting how country citizens are responding to them. Maybe these questions aren't going deep enough into emotional issues around the global crisis so the true trends lie further within. Future research could go much deeper into the specific mechanisms through which the pandemic influenced happiness in each country. Exploring additional variables could definitely provide a more comprehensive understanding of the impact of the pandemic on overall happiness.

**Conclusion**

In conclusion, the analysis reveals that the COVID-19 pandemic had varying effects on overall happiness in the United States, Israel, and China. While the United States experienced a decline in happiness scores and an increase in negative affect, Israel and China displayed more resilience and minor fluctuations in happiness. These findings highlight the diverse impacts of the pandemic on national happiness levels and call for further investigation into the underlying factors contributing to these trends.

▾ Summary of Findings

An added layer of interest to the initial question is that the US, Israel, and China are all first world, relatively wealthy countries who act as major international figures. So, the respective effects of COVID-19 on national happiness can shed light on how powerful goverments with abundant resources handled the pandemic. In Israel, while the year 2020 saw a slight relative dip in life quality and positive emotion, there was an increase in both variables in the proceeding years 2021 and 2022. Addiotnally, the national feelings of negative emotions has been on a steady decline since 2019, and the pandemic did not seem to effect that trend.

Next, looking at the United States, 2020 saw a slight increase in reported life quality, and a sharp decline from 2021-2022, with 2022 having the all-time lowest life quality since the beginning of the reported data. Though, it should also be noted that life quality has been on an overall decline since the beginning of the reported data. Negative effect did not have a strong change in either direction during the pandemic.

Lastly, China had an increase in life quality during the pandemic, with the highest reported happiness over the sampled years being in the year 2021. Though, positive emotions had a decrease, and negative emotions had an increase in pandemic years.

Through comparing life quality, positive effect, and negative effect from pre-pandemic years to pandemic years, the data indicates that the US had the strongest adverse repsonse to the pandemic, followed by China, then Israel . It should also be noted that other factors could have contributed to the rise or fall of national happiness, such as political circumstances unrelated to the pandemic, but the data can certainly suggest trends on how COVID-19 effected national happiness.

## References

https://worldhappiness.report/