

# World Happiness Trends

## **The Flying Tigers:**

**Kacee Kira Chien  
Ricky Lee  
Trevor Jones  
Vanessa Villagomez**

# World Happiness

## Outline

- Introduction
  - Data Collection
  - Data Exploration & Cleaning
  - Analysis Process
  - Conclusions & Implications
-

# **INTRODUCTION**

**Have you ever thought about world happiness trends?**

**We had no knowledge on the matter until discovering the data available in Kaggle.com. We were interested and determined to analyze World Happiness data available to discover the key factors that inspire happiness around the world.**

# QUESTIONS

- 1. What are the top 20 happiest countries in the world?**
- 2. What key factors inspire happiness in each region of the world?**
- 3. How does the happiness score differ per region?**
- 4. How have happiness factors changed over the years from 2015 through 2019?**

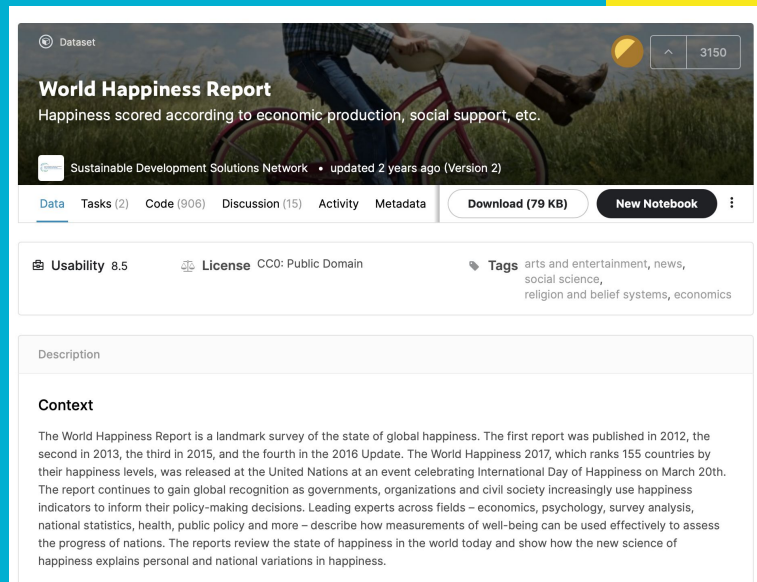
# **PRE-ANALYSIS HYPOTHESES**

- **In today's rapidly changing world, we would think that money might easily be one of the main reasons that contribute to people's happiness, we'll proceed with an analysis on the data to find out.**
- **Developed countries are happier.**

# DATA COLLECTION

## Kaggle Dataset(s): World Happiness Report

<https://www.kaggle.com/unsdsn/world-happiness?select=2015.csv>



The screenshot shows the Kaggle dataset page for the 'World Happiness Report'. The header features a banner image of a person riding a bicycle in a field, with the text 'World Happiness Report' and 'Happiness scored according to economic production, social support, etc.'. Below the banner, it indicates the dataset is from the 'Sustainable Development Solutions Network' and was updated 2 years ago (Version 2). The page has a navigation bar with tabs for 'Data', 'Tasks (2)', 'Code (906)', 'Discussion (15)', 'Activity', and 'Metadata'. There are buttons for 'Download (79 KB)' and 'New Notebook'. Below the navigation bar, the 'Usability' is 8.5, the 'License' is CC0: Public Domain, and the 'Tags' include 'arts and entertainment, news, social science, religion and belief systems, economics'. The 'Description' section is partially visible, starting with 'Context'.

Dataset

World Happiness Report

Happiness scored according to economic production, social support, etc.

Sustainable Development Solutions Network • updated 2 years ago (Version 2)

Data Tasks (2) Code (906) Discussion (15) Activity Metadata Download (79 KB) New Notebook

Usability 8.5 License CC0: Public Domain Tags arts and entertainment, news, social science, religion and belief systems, economics

Description

Context

The World Happiness Report is a landmark survey of the state of global happiness. The first report was published in 2012, the second in 2013, the third in 2015, and the fourth in the 2016 Update. The World Happiness 2017, which ranks 155 countries by their happiness levels, was released at the United Nations at an event celebrating International Day of Happiness on March 20th. The report continues to gain global recognition as governments, organizations and civil society increasingly use happiness indicators to inform their policy-making decisions. Leading experts across fields – economics, psychology, survey analysis, national statistics, health, public policy and more – describe how measurements of well-being can be used effectively to assess the progress of nations. The reports review the state of happiness in the world today and show how the new science of happiness explains personal and national variations in happiness.

# **DATA COLLECTION**

- **From the Gallup Global Research**
- **Groundbreaking benchmark from the United Nation's Sustainable Development Solutions Network**
- **Years 2015 - 2019**
- **Measures happiness score per country and region**
- **Happiness score estimates the extent to which each of six factors -- economy (GDP per capita), family (social support), health (life expectancy), freedom, trust (government corruption), and generosity -- contribute to making life evaluations higher in each country**
- **Over 150 Countries and 10 regions combined**

## **DATA EXPLORATION & CLEANING**

- **Step 1: Read in the CSV data for each year**
- **Step 2: Add a column denoting the year the data relates to**
- **Step 3: If needed, rename columns for the sake of consistency across years in the data**
- **Step 4: Drop all non-predictive variables that shouldn't be included in the analysis**
- **Step 5: Rearrange the data fields to be consistent year over year**
- **Step 6: Group each year's clean data set into a list, combined the data into a multi year data set, and create new variables**
- **Step 7: Export the data for the top 20 countries and for all countries**



## DATA EXPLORATION & CLEANING (Continued)

```
## 2018 Data Cleaning
# Import 2018 World Happiness Data
data_2018 = pd.read_csv('./Resources/2018_data.csv')

# Add column indicating the year
data_2018['Year'] = '2018'

# Rename data fields to match prior years
data_2018_rename = data_2018.rename(columns={'Overall rank': 'Happiness Rank',
                                             'Country or region': 'Country',
                                             'Score': 'Happiness Score',
                                             'GDP per capita': 'Economy (GDP per Capita)',
                                             'Social support': 'Family',
                                             'Healthy life expectancy': 'Health (Life Expectancy)',
                                             'Freedom to make life choices': 'Freedom',
                                             'Perceptions of corruption': 'Trust (Government Corruption)'})

# Rearrange the data fields
column_order = ['Year', 'Country', 'Happiness Rank', 'Happiness Score', 'Economy (GDP per Capita)', 'Family',
                'Health (Life Expectancy)', 'Freedom', 'Trust (Government Corruption)', 'Generosity', 'Dystopia Residual']
data_2018_clean = data_2018_rename.reindex(columns = column_order)

data_2018_clean
```

## DATA EXPLORATION & CLEANING (Continued)

*# Aggregate cleaned data sets from each year into one file*

```
data_list = [data_2015_clean, data_2016_clean, data_2017_clean, data_2018_clean, data_2019_clean]
multi_year_data = pd.concat(data_list)
multi_year_data
```

*# Group the data sets by country*

```
total_country_rank = multi_year_data.groupby('Country')['Happiness Rank'].sum()
```

*# Calculate the average happiness rank for each country over the 5 year period*

```
avg_country_rank = total_country_rank/5
```

```
data_rank = pd.merge(left=multi_year_data, right=avg_country_rank, on='Country')
data_rank
```

*# Rename columns and fill nulls*

```
data_rank_all_rename = data_rank.rename(columns={'Happiness Rank_x': 'Happiness Rank',
                                                  'Happiness Rank_y': 'Avg Happiness Rank'})
data_rank_all = data_rank_all_rename.ffill(axis=0)
```

```
data_rank_all
```

*# Export data for all years*

```
data_rank_all.to_csv('Resources/clean_data/data_country_all.csv', index=False)
```

*# Sort the data set and drop values below top 20*

```
data_rank_all_sort = data_rank_all.sort_values('Avg Happiness Rank')
```

```
data_rank_top = data_rank_all_sort.drop(range(100, len(data_rank_all_sort)))
```

```
data_rank_top
```

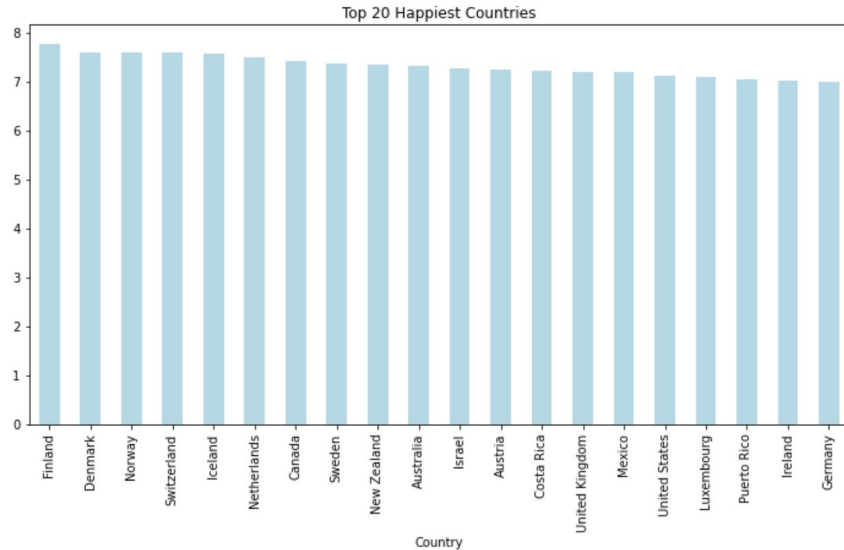
# ANALYSIS PROCESS

## Values Explained

- **Happiness Rank:** rank of any country in a particular year
- **Happiness Score:** sum of all numerical columns in the datasets
- **Each factor describes the extent to which it contributed in evaluating the happiness in each country/region:**
  - **Economy (GDP per Capita)** – “Gross domestic product (GDP) is among the most frequent indicators used to **monitor the health of a country's economy**. The calculation of a nation's GDP takes into account several distinct variables relating to this nation's market, including its investment and consumption.”
  - **Family (Social Support)** – “If you were in trouble, do you have relatives or friends you can count on to help you whenever you need them, or not?”
  - **Health (Life Expectancy)** – “The time series of healthy life expectancy at birth are constructed based on data from the World Health Organization (WHO) Global Health Observatory data repository.”
  - **Freedom** – Freedom to make life choices: “Are you satisfied or dissatisfied with your freedom to choose what you do with your life?”
  - **Trust (Government Corruption)** – “Is corruption widespread throughout the government or not?” & “Is corruption widespread within businesses or not?”
  - **Generosity** – “Have you donated money to a charity in the past month?”

Resource: [Social Environments for World Happiness](#) | [The World Happiness Report](#)

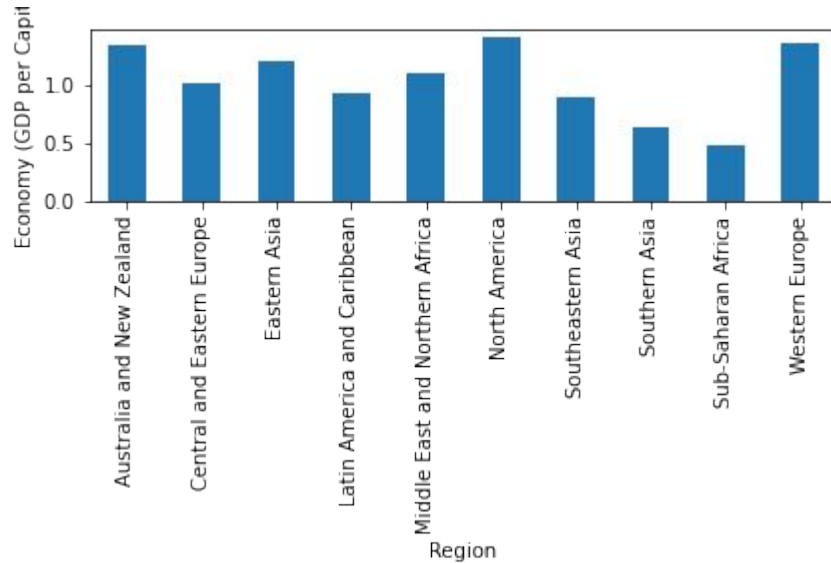
# ANALYSIS PROCESS



## Top 20 Happiest Countries:

From the bar chart, we can see the top 20 Happiest countries in the world. Our chart shows the average happiness rating in 5 years.

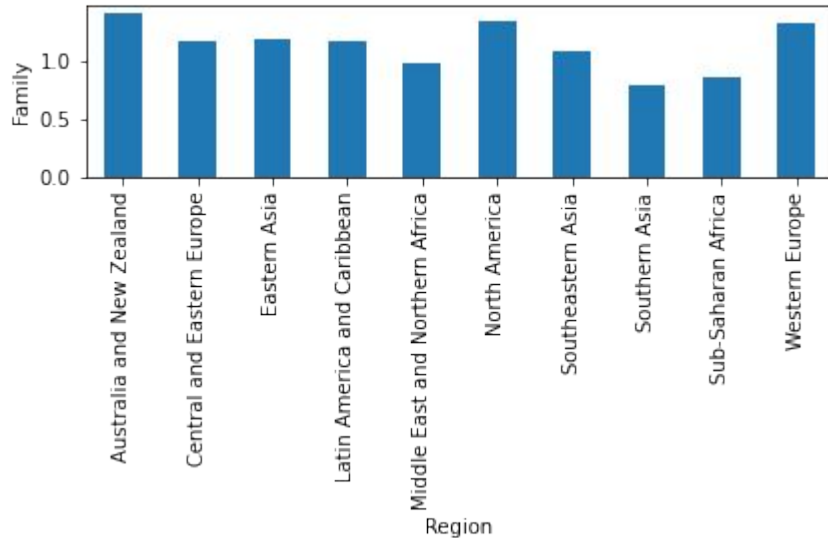
# ANALYSIS PROCESS



## 1. Economy Factor by Regions:

From the bar chart, we can see that between 2015 ~ 2019, North America has the highest GDP per capita, followed by Western Europe, and Australia & New Zealand. The lowest score is Sub-Saharan Africa.

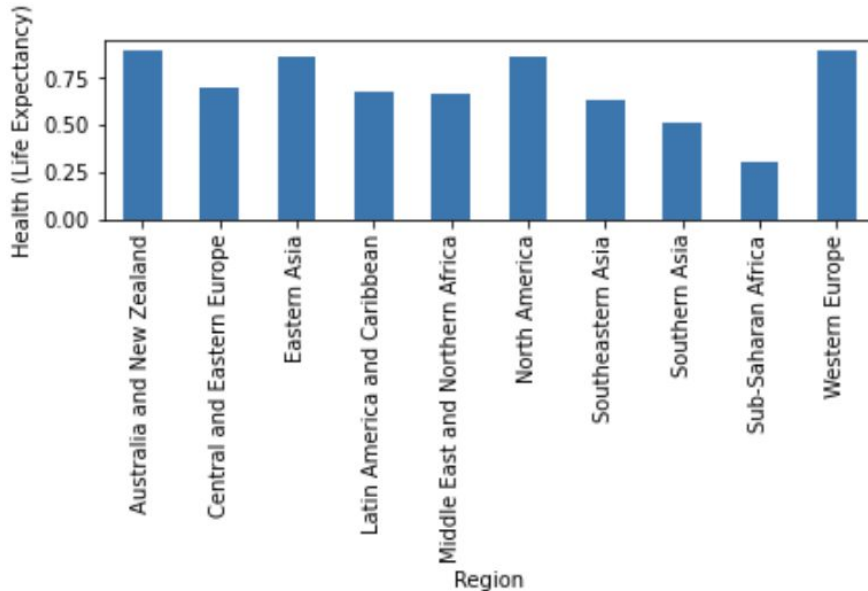
# ANALYSIS PROCESS



## 2. Family Factor by Regions:

From the bar chart, we can see that Australia & New Zealand, North America and Western Europe are the regions that have the highest scores related to family factors. The lowest score is Southern Asia.

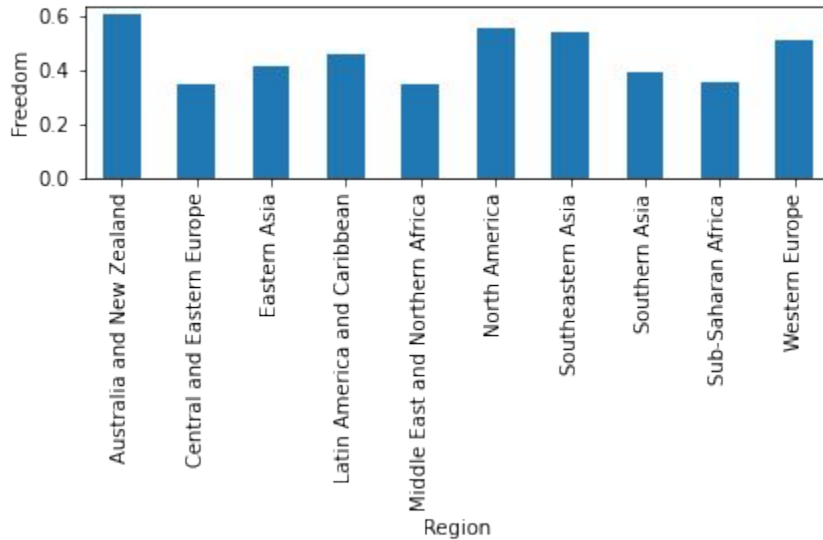
# ANALYSIS PROCESS



## 3. Health (Life Expectancy) by Regions:

The chart shows the relationship between Health and happiness by continents. Australia and New Zealand have the highest rating. Western Europe is second and Eastern Asia come third.

# ANALYSIS PROCESS

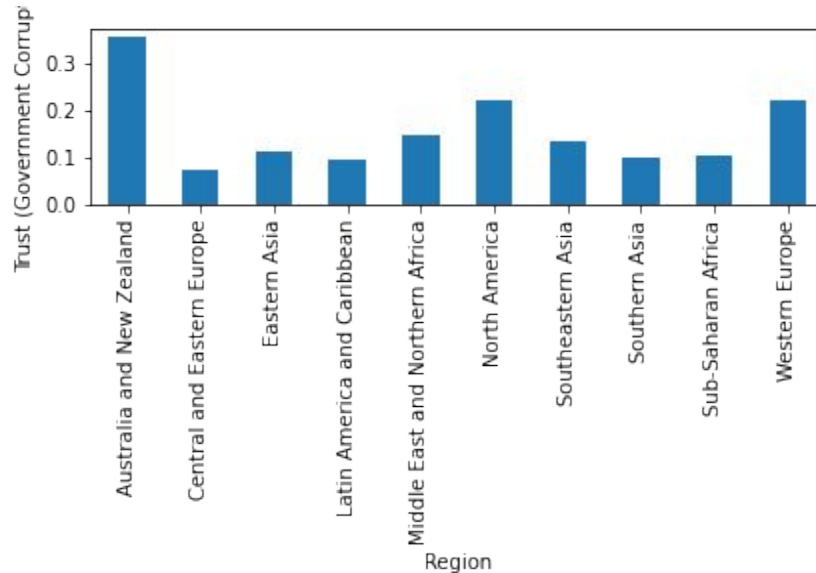


## 4. Freedom Factor by Regions:

From the bar chart, we can see that Australia & New Zealand has the highest score in Freedom, followed by North America and Southeastern Asia. The lowest in score is Middle East and Northern Africa.



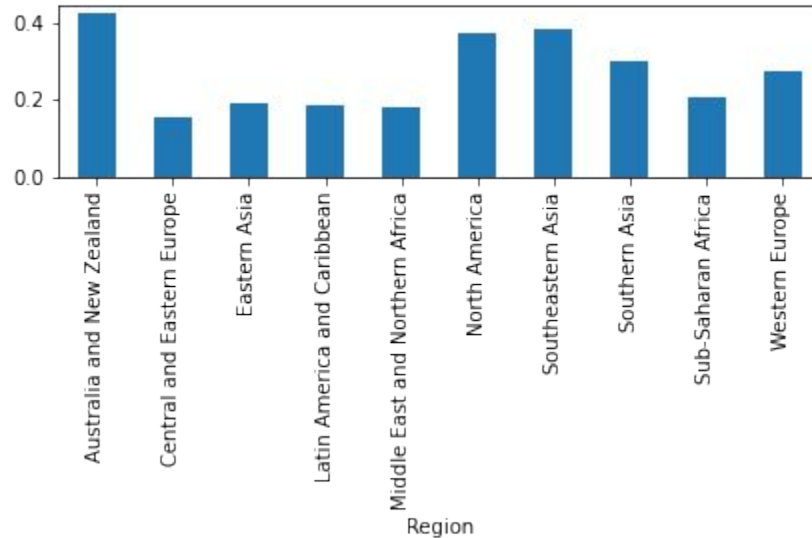
# ANALYSIS PROCESS



## 5. Trust(Government Corruption) by Regions:

The chart shows the relationship between trust and happiness by continents. Australia and New Zealand have the highest rating. The top 20 happiest countries all have relatively high rating of government trust.

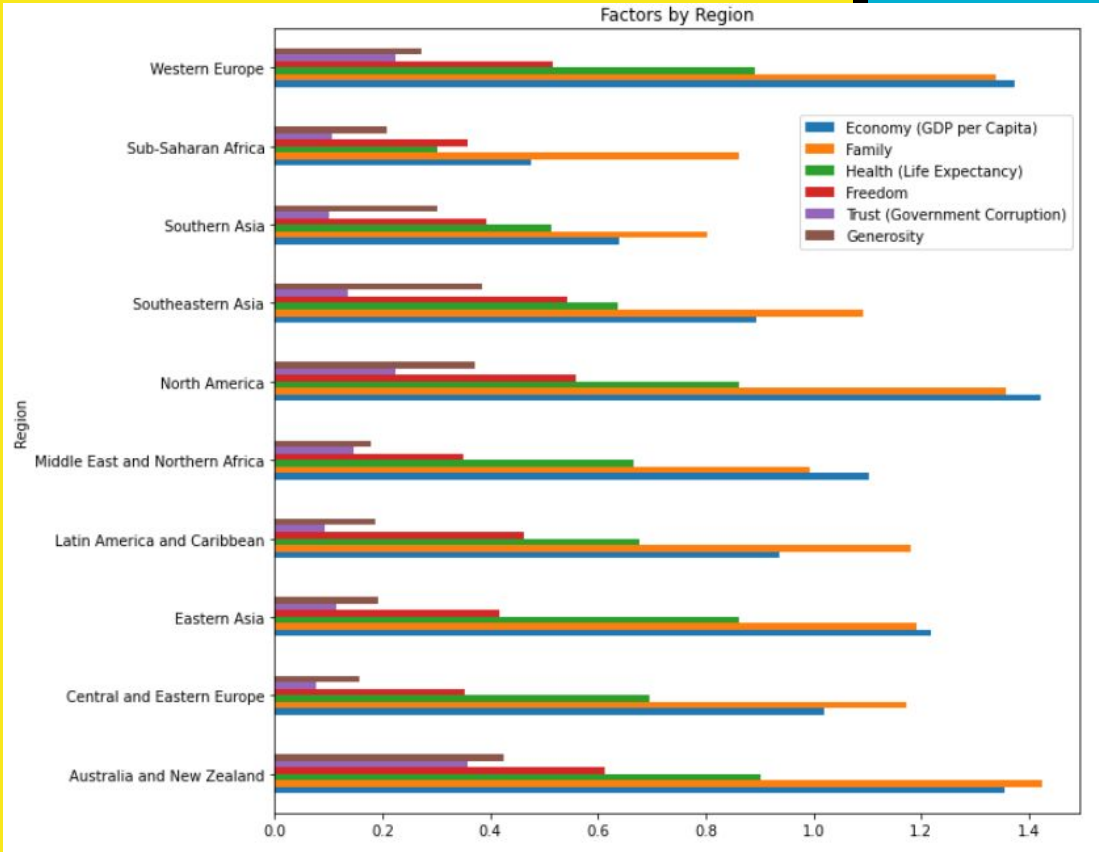
# ANALYSIS PROCESS



## 6. Generosity Factor by Regions:

From the bar chart, we can see that Australia and New Zealand has the highest score in Generosity, followed by Southeastern Asia, and North America. The lowest score is Central and Eastern Europe.

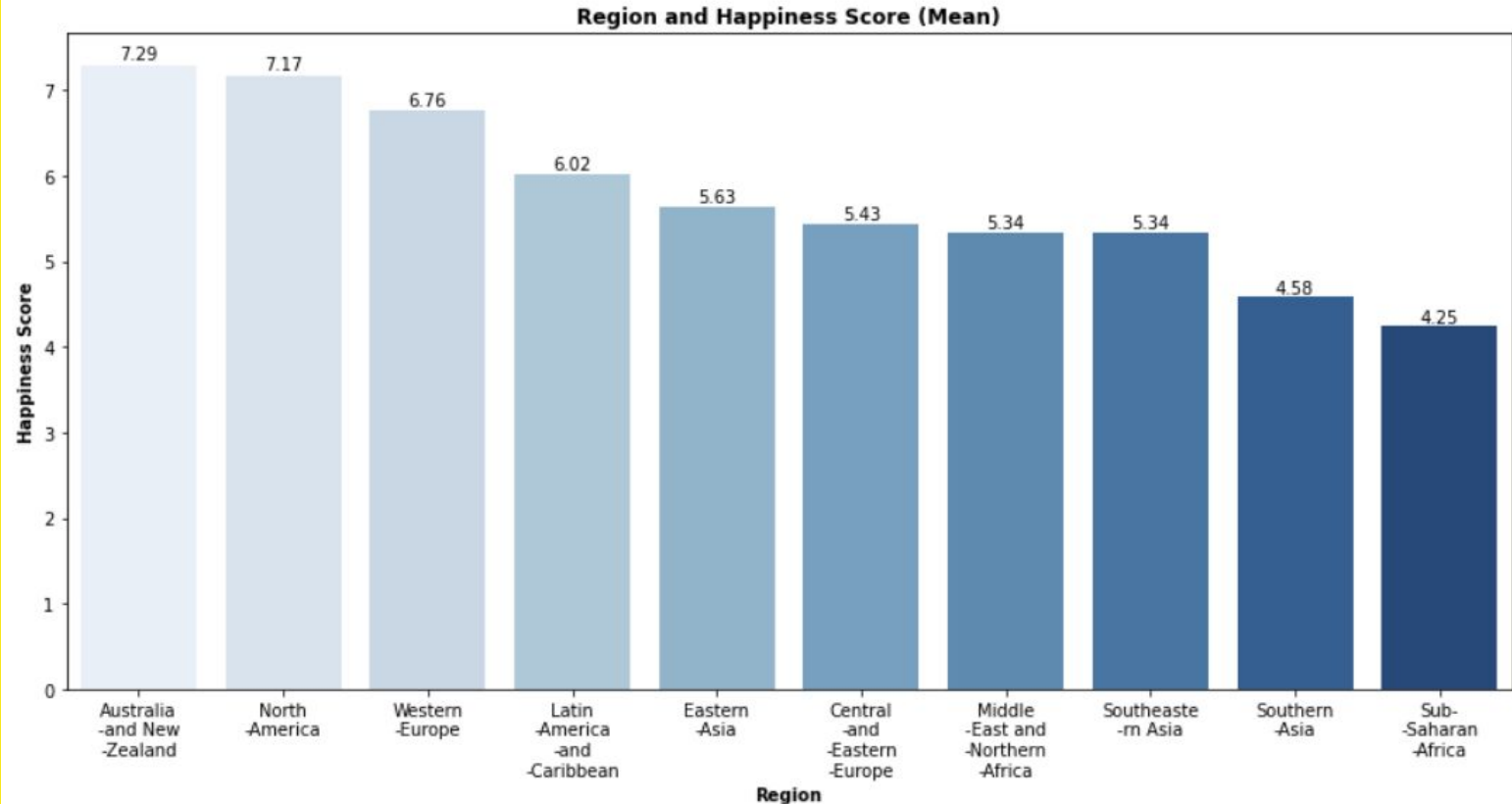
# ANALYSIS PROCESS



## Factors by Region:

According to the analysis and the chart, we can see that the GDP per Capita plays a biggest factors for most regions and the next is family. The least factor is Generosity and Trust.

# ANALYSIS PROCESS

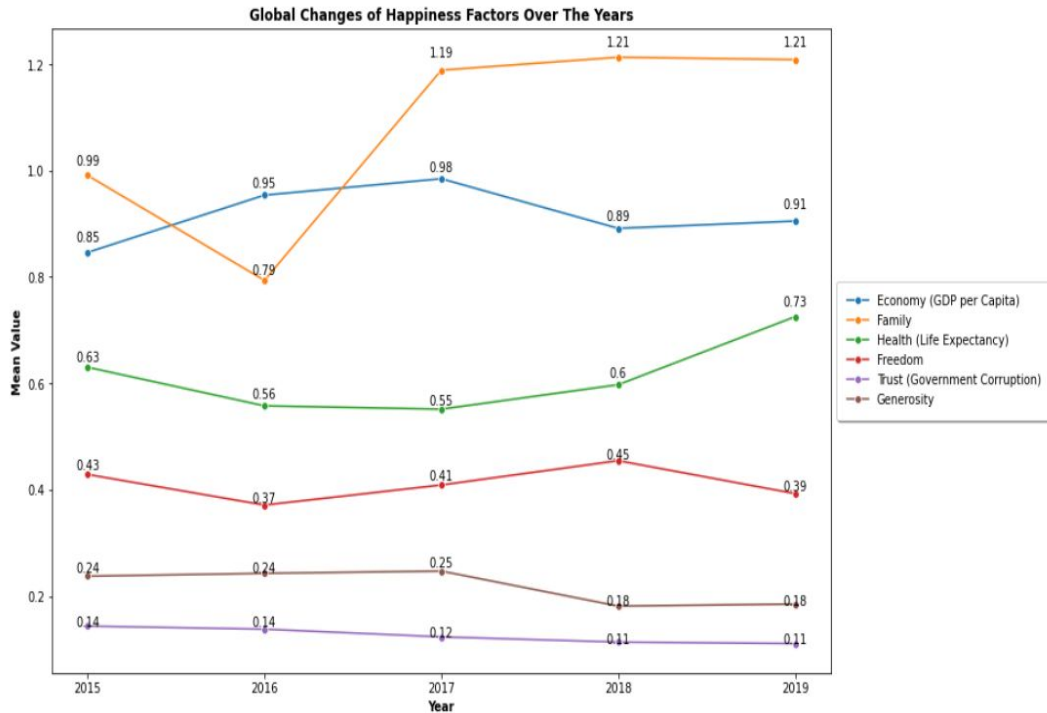


# ANALYSIS PROCESS

**How does the happiness score differ per region?**

- **By looking at the chart above, we determined the average happiness score per region from 2015 through 2019, the Australia and New Zealand region is the happiest with a mean score of 7.29, and North America comes second with a mean score of 7.17. The following regions came in order of highest to lowest: Western Europe (6.76), Latin America and Caribbean (6.02), Eastern Asia (5.63), Central and Eastern Europe (5.43), Middle East and North Africa (5.34), Southeastern Asia (5.34), Southern Asia (4.58), and Sub-Saharan Africa (4.25).**

# ANALYSIS PROCESS



## 7. Global Changes of happiness factors over the years:

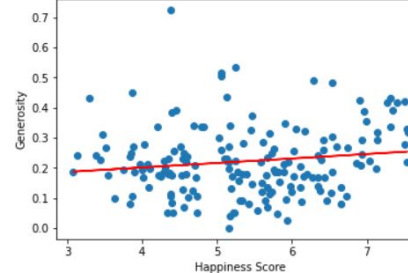
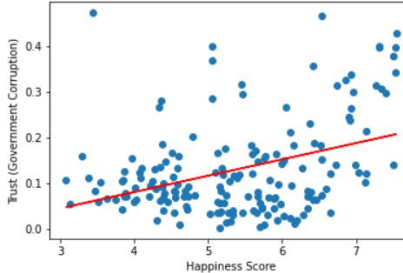
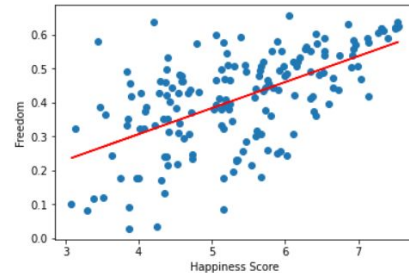
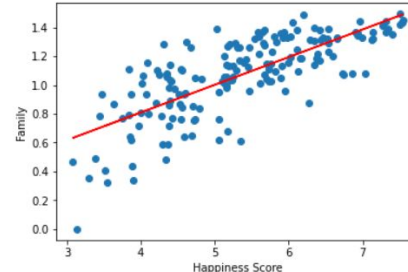
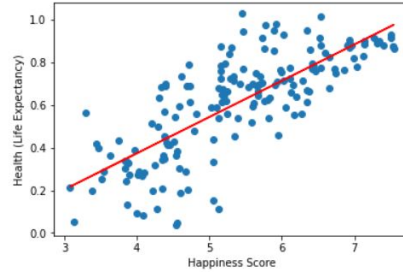
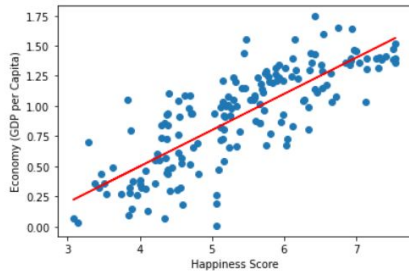
The chart shows how happiness factors were changing over the course of five years.

The impact of the factors do not change much over time. Except we can see that Family/Social Support increased its impact on the overall global happiness since 2016, surpassing GDP per Capita.

# CONCLUSIONS & IMPLICATIONS

## Scatter Plot Comparisons for Each Factor:

From the scatter plots below, we can see that most factors have positive correlation to the overall Happiness Score. Among all the factors, from the plots we can see that GDP has the highest correlation with Happiness Score, followed by Health (Life Expectancy), then Family (Social Support).



## R value rank:

1. **GDP per capita** = 0.645
2. **Health (Life Expectancy)** = 0.596
3. **Family** = 0.571
4. **Freedom** = 0.347
5. **Trust (Government Corruption)** = 0.142
6. **Generosity** = 0.020

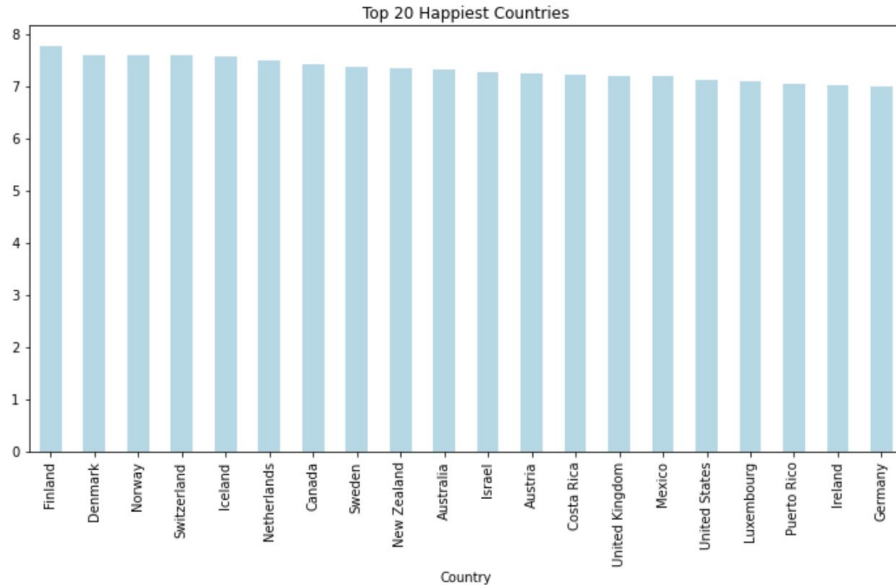
## **CONCLUSIONS & IMPLICATIONS**

**Based on the scatter plots and our findings, we can see that GDP has the highest correlation with the Happiness Score, and is one of the main reasons for people's happiness, but Health (Life Expectancy) and Family (Social Support) are almost as important factors as the GDP. So, we can say that money is important, but we also have to feel supported by family and society around, as well as remain healthy, in order to sustain the maximum degree of happiness.**



# CONCLUSIONS & IMPLICATIONS

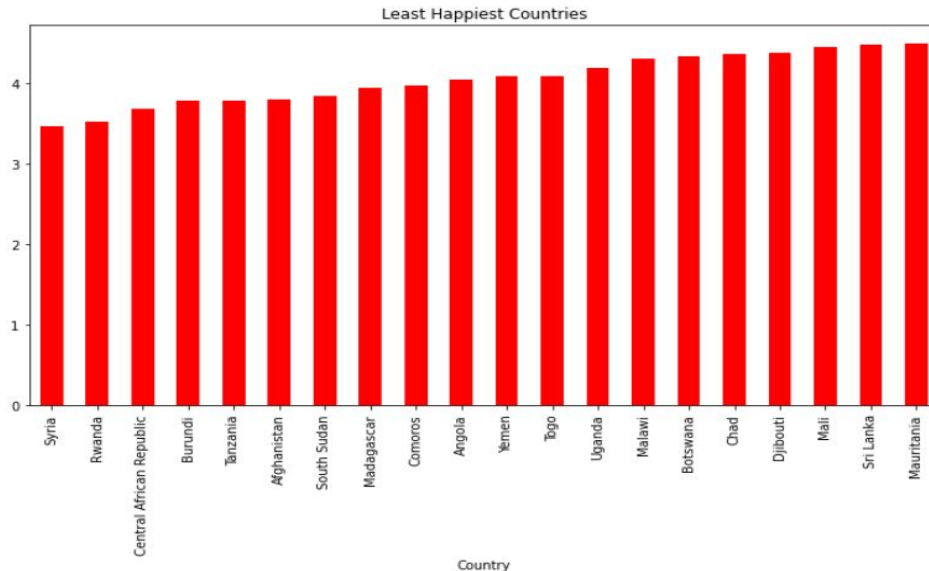
## Developed countries are happier.



Based on our analysis, developed countries project higher happiness scores than underdeveloped countries according to the top 20 countries, ranging above the 7.0 Happiness Score from years 2015 through 2019.

# CONCLUSIONS & IMPLICATIONS

Developing countries are not as happy.



Based on our analysis, developing countries project lower happiness scores than developed countries according to the bottom 20 countries, ranging below the 5.0 Happiness Score from years 2015 through 2019.

Q&A



**THANK YOU**