



WORLD HAPPINESS ANALYSIS

Thang Tran

Sep. 2022

Executive Summary

The purpose of this project is to perform exploratory analysis on World Happiness Report datasets. The dataset represents survey results from many countries between 2015 to 2019. The data includes characteristics such as country, happiness score, GDP, and life expectancy. This dataset is open-source, and it is available on Kaggle. This dataset was selected because it'd interesting to investigate what factors are related to happiness levels and how they vary in different countries.

Data Sources

<https://www.kaggle.com/datasets/unsdsn/world-happiness>

<https://geojson-maps.ash.ms/>

<https://www.kaggle.com/datasets/unsdsn/world-happiness>

Limitation & Ethics

- Survey results are prone to human bias .
- Sample bias as the survey may not represent the entire population.

Data Wrangling & Consistency Checks

- Check the dimensions and descriptive statistics
- Rename columns
- Check for missing data
- Check for duplicates
- Check for mixed data types
- Drop columns
- Combine the datasets

Data Profile

Dataset	Before cleaning		After cleaning	
	Columns	Rows	Rows	Columns
2015 Data	158	12	158	9
2016 Data	157	13	157	9
2017 Data	155	12	155	9
2018 Data	156	9	155	9
2019 Data	156	9	156	9

Combined Data	n/a	n/a	781	9
---------------	-----	-----	-----	---

Column Details

Column	Description	Data Type	Time Variant
Country	name of country	object	no
Happiness Rank	country's rank based on happiness score	int64	no
Happiness Score	happiness level from 0(least) to 10 (greatest)	float64	no
GDP	country's gross domestic product	float64	no
Life Expectancy	country's life expectancy	float64	no
Freedom	people's freedom of choice	float64	no
Generosity	how nice people are in the country	float64	no
Corruption	government corruption level from 0(least) to 10(greatest)	float64	no
Year	year of survey result	int64	yes

Questions to Explore

• Which country is the happiest? Least happy?
• What factor(s) is strongly/weakly related to the happiness level?
• How does the happiness level change over the years?