Brian Avila

**Global Happiness Report Data Source**

**Data Sourcing:** This is external data retrieved from Kaggle. The World Happiness Report utilizes data from the Gallup World Poll to generate happiness scores and rankings, considering factors such as economic production, social support, life expectancy, freedom, absence of corruption, and generosity.

**Data Collection**: The method for this type of data collection is

administrative as it is part of Government records collection. The data is

collected manually as it requires input from a qualified professional. There

might be time lag as results can take time to obtain.

**Data Contents**: The data set consists of most countries worldwide. The variables used in this data are country, region, year, Rank, Happiness Score, GDP\_per Capita, Family, Life\_Expectancy, Freedom, Government\_Corruption, Generosity, and Dystopia Residual.

**Data Relevance:** We can assume high relevance and accuracy of this data.

The World Happiness Report is a widely recognized survey that assesses global happiness, gaining popularity among governments and organizations for informing policy decisions. Experts from various fields, including economics, psychology, and public policy, highlight the effective use of happiness indicators to gauge national progress.

**Data Limitations**: the possible limitations of this particular data set can be

due to manual input errors, missing data or inaccurate counts.

**Data Profile**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| variables | data type |  |  |  |
| Country | qualitative | nominal | structured | time invariant |
| Year | quantitative | cont | structured | time invariant |
| Rank | quantitative | discrete | structured | time variant |
| Happiness Score | quantitative | discrete | structured | time variant |
| GDP\_Capita | quantitative | discrete | structured | time variant |
| Family | quantitative | discrete | structured | time variant |
| Life\_Expectancy | quantitative | discrete | structured | time variant |
| Freedom | quantitative | discrete | structured | time variant |
| Gov\_Corruption | quantitative | discrete | structured | time variant |
| Generosity | quantitative | discrete | structured | time variant |
| Dystopia | quantitative | discrete | structured | time variant |
| Region | qualitative | nominal | structured | time invariant |

**Data Cleaning Measures**

**Data Wrangling**

|  |  |  |
| --- | --- | --- |
| Columns dropped | Columns Renamed | reason |
| St\_Error |  | Unnecessary for analysis |
| Lower Confidence Interval |  | empty |
| Upper Confidence Interval |  | empty |
|  | Score to Happiness Score | More descriptive |
|  | Dystopia\_Residual to Dystopia | More concise |

A screenshot of a computer

Description automatically generated

Questions to explore on the Analysis:

1. Which top 5 countries have highest happiness scores?
2. Which top 5 companies rank the lowest?
3. Does Freedom correlate with high/low happiness score?
4. Does GDP\_Capita correlate with high/low happiness score?
5. What is the major contributor to high happiness score?