DATA ANALYTICS OF WORLD HAPPINESS

Summary

An analysis of the World Happiness Report has been conducted using an open-source data set. The objective of the project includes studying the various factors which lead up to the calculation of a "Happiness Index" for each country and understanding its distribution throughout the world.

The process involves ingesting the data set in SQL.

Data Overview

Source: The dataset is an open-source dataset from a report published on world happiness, first in 2012 and thereafter every year. This report is an outcome of the survey results of the Gallup World Poll which takes a representative sample from each country and asks them questions in the form of a Cantril ladder, which is asking respondents to think of a ladder with the best possible life for them being a 10 and the worst possible life being a 0 and to rate their own current lives on that scale.

The dataset is present here for 2015-2017: https://www.kaggle.com/unsdsn/world-happiness/data

Data Description:

The datasets are identical except for the year they contain information and have the following columns:

- **Country:** Name of the country
- **Region:** Region of the world, the country belongs to
- **Happiness Rank:** Rank of the country according to happiness score
- **Happiness Score:** Metric measured as a combination of various factors
- **Economy (GDP per capita):** The extent to which GDP contributes to happiness
- **Family:** The extent to which Family contributes to happiness
- **Health (Life Expectancy):** The extent to which Life Expectancy contributes to happiness
- **Freedom:** The extent to which Freedom contributes to happiness
- **Trust:** The extent to which trust in government contributes to happiness
- **Generosity:** Generosity of the general public and its contribution to happiness
- **Dystopia Residual:** Contribution to Dystopia residual to happiness. Dystopia is an imaginary country that has the world's least happy people. The purpose of having this is to have a lower benchmark so that all countries do positively against it. This variable has no physical significance.

Detailed information about Data :

About Dataset

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.

Content

The happiness scores and rankings use data from the Gallup World Poll. The scores are based on answers to the main life evaluation question asked in the poll. This question, known as the Cantril ladder, asks respondents to think of a ladder with the best possible life for them being a 10 and the worst possible life being a 0 and to rate their own current lives on that scale. The scores are from nationally representative samples for the years 2013-2016 and use the Gallup weights to make the estimates representative.

The columns following the happiness score estimate the extent to which each of the six factors – economic production, social support, life expectancy, freedom, absence of corruption, and generosity – contribute to making life evaluations higher in each country than they are in Dystopia, a hypothetical country that has values equal to the world's lowest national averages for each of the six factors. They have no impact on the total score reported for each country, but they do explain why some countries rank higher than others.

Inspiration

What countries or regions rank the highest in overall happiness and each of the six factors contributing to happiness?

What is Dystopia?

Dystopia is an imaginary country that has the world's least-happy people. The purpose in establishing Dystopia is to have a benchmark against which all countries can be favorably compared (no country performs more poorly than Dystopia) in terms of each of the six key variables, thus allowing each sub-bar to be of positive width. The lowest scores observed for the six key variables, therefore, characterize Dystopia. Since life would be very unpleasant in a country with the world's lowest incomes, lowest life expectancy, lowest generosity, most corruption, least freedom and least social support, it is referred to as "Dystopia," in contrast to Utopia.

What are the residuals?

The residuals, or unexplained components, differ for each country, reflecting the extent to which the six variables either over- or under-explain average 2014-2016 life evaluations. These residuals have an average value of approximately zero over the whole set of countries. Figure 2.2 shows the average residual for each country when the equation in Table 2.1 is applied to the average 2014- 2016 data for the six variables in that country. We combine these residuals with the estimate for life evaluations in Dystopia so that the combined bar will always have positive values. As can be seen in Figure 2.2, although some life evaluation residuals are quite large, occasionally exceeding one point on the scale from 0 to 10, they are always much smaller than the calculated value in Dystopia, where the average life is rated at 1.85 on the 0 to 10 scale.

What do the columns succeeding the Happiness Score(like Family, Generosity, etc.) describe?

The following columns: GDP per Capita, Family, Life Expectancy, Freedom, Generosity, Trust Government Corruption describe the extent to which these factors contribute to evaluating the happiness in each country.

The Dystopia Residual metric actually is the Dystopia Happiness Score(1.85) + the Residual value or the unexplained value for each country as stated in the previous answer.

If you add all these factors up, you get the happiness score so it might be un-reliable to model them to predict Happiness Scores.

Experiment No. 09 and 10

You will get Happiness contest dataset 2019. This dataset contains three columns. **List of countries participated, with happiness rank and score.**

You need to insert this data based on schema to SQL table. When your data is ready, Write the query find solution of below problem statement.

Steps to Import the CSV Format Dataset into MySQL :

Link: https://drive.google.com/drive/folders/1HOuLjqKQP8bl0wIbNvkuAZ8JLQ1UZNYr?usp=share_link

IN THIS CODE THE TABLE NAME IS happiness_data_2019;

SO, CHANGE ACCORDINGLY.

```
USE HappinessDB;
Show Tables;

DESC happiness_data_2019;
```

1) Find top ten country with happiness rank and Happiness Score

```
SELECT * FROM happiness_data_2019
ORDER BY Score DESC
LIMIT 10;
```

2) Find happiness rank of India

```
SELECT Overall_Rank FROM happiness_data_2019
Where Country = 'India';
```

3) Find difference in happiness score of Ireland and Spain.

```
SELECT
  (SELECT Score FROM happiness_data_2019 WHERE Country = 'Ireland') -
  (SELECT Score FROM happiness_data_2019 WHERE Country = 'Spain')
  AS HappinessScoreDifference;
```

4) Find bottom five country with least happiness score.

```
SELECT * FROM happiness_data_2019
ORDER BY Score ASC
LIMIT 10;
```

5) Count the total number of countries with happiness score more than 5.25

```
SELECT COUNT(*) FROM happiness_data_2019
WHERE Score > 5.25;
```

6) Count the total number of countries with happiness score less than 5.25

```
SELECT COUNT(*) FROM happiness_data_2019
WHERE Score < 5.25;
```

7) Find the happiness rank of country Kenya.

```
SELECT Overall_Rank FROM happiness_data_2019
WHERE Country = 'Kenya';
```

8) Find the happiness score of the country Senegal

```
SELECT Score FROM happiness_data_2019
WHERE Country = 'Senegal';
```

9) Find the total number of country participated in Happiness contest.

```
SELECT COUNT(DISTINCT(Country)) FROM happiness_data_2019;
```

10) Find the average happiness score of the country participated in Happiness Contest.

```
SELECT AVG(Score) FROM happiness_data_2019;
```

SOURCE CODE :

```
USE HappinessDB;
Show Tables;

DESC happiness_data_2019;

SELECT * FROM happiness_data_2019
ORDER BY Score DESC
LIMIT 10;

SELECT Overall_Rank FROM happiness_data_2019
Where Country = 'India';

SELECT
   (SELECT Score FROM happiness_data_2019 WHERE Country = 'Ireland') -
   (SELECT Score FROM happiness_data_2019 WHERE Country = 'Spain')
   AS HappinessScoreDifference;

SELECT * FROM happiness_data_2019
ORDER BY Score ASC
LIMIT 10;

SELECT COUNT(*) FROM happiness_data_2019
WHERE Score > 5.25;
```

```
SELECT Overall_Rank FROM happiness_data_2019
WHERE Country = 'Kenya';

SELECT Score FROM happiness_data_2019
WHERE Country = 'Senegal';

SELECT COUNT(DISTINCT(Country)) FROM happiness_data_2019;

SELECT AVG(Score) FROM happiness_data_2019;
```