**Project Title: Analyzing Happiness Factors of a Country**

**Introduction**

Happiness of a country depends on various factors that cumulatively affect the happiness score.

The World Happiness 2017 was first released at the United Nations. The happiness indicators are factored in when making decisions. The Happiness scores are calculated based on the Economy, Health, Family, Freedom, Generosity, Trust on Government, Freedom of expression etc.

This project analyzes the Happiness dataset of 155 countries for the year 2017.

**Statistical / Hypothetical Question**

Following hypothesis was tested.

1. GDP is higher for happiest countries than the happy and rest of the countries.

2. Health and life expectancy plays a critical role in the happiness score of a country

**Outcome of your EDA**

We do see some relation between some parameters/variables on the Happiness Scores.

The variables that were analyzed were -

* **Economy (GDP per Capita)**
* **Family**
* **Health (Life Expectancy)**
* **Freedom**
* **Trust (Government Corruption)**
* **Generosity**

The data set provides the extend to which the above factors affect the overall Happiness scores.

For my analysis, I have considered the 2017-year data set using which I was able to analyze, test and support the hypothesis.

**What do you feel was missed during the analysis?**

The Dystopia Residual value which is an imaginary benchmark against which the country is measured. My analysis was not able to utilize this value.

**Were there any variables you felt could have helped in the analysis?**

The data set could have included the Employment metrics, educational opportunities, cultural and spiritual influences, living standards to say the least that contribute to the happiness of the country.

**Were there any assumptions made you felt were incorrect?**

The dataset size is very small with only 155 countries and providing the happiness score at a country level. If the report was divided into states/regions within the country, it would provide deeper understanding of which of the different variables positively and negatively affect the happiness scores of a country.

**What challenges did you face, what did you not fully understand?**

Hypothesis testing and Regression testing are challenging areas for me. I was unable to conduct the regression analysis. The Hypothesis testing can be conducted using various methods that I still do not have clear and complete understanding of. I am still learning how to analyze the data as there are multiple factors that could influence the outcome of test results.

I intend to continue gaining experience with different data sets.

**References:**

Data set from -

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

The World Happiness Report is a publication of the Sustainable Development Solutions Network, powered by data from the Gallup World Poll and Lloyd’s Register Foundation, who provided access to the World Risk Poll.

The Report is supported by The Ernesto Illy Foundation, illycaffè, Davines Group, The Blue Chip Foundation, The William, Jeff, and Jennifer Gross Family Foundation, The Happier Way Foundation, Indeed, and Unilever’s largest ice cream brand, Wall’s.

<https://worldhappiness.report/ed/2017/>

Think Stats, Allen B Downey