

# World Happiness Report: 2015 & 2016

Marlene Aviles

Final Project

Data Exploration and Analysis



# About World Happiness Report

The World Happiness Report is a landmark survey of the state of global happiness that ranks 156 countries by how happy their citizens perceive themselves to be. [1]

- People in a survey were asked to rank the extent to which perception of the factors contributes to happiness scores.
- The elements used include factors such as freedom, generosity, economy, and family.
- The years and 2015 and 2016 were compared.



# Variables

- *Country: Name of country*
- *Region : Region the country pertains to*
- *Happiness Rank : Rank of the country based on the Happiness Score*
- *Happiness Score : A metric measured in 2015 by asking the sampled people: "How would you rate your happiness on a scale of 0 to 10 where 10 is the happiest."*
- *Lower Confidence Interval: the lower confidence limit of an estimate of a parameter \*(2016 only)*
- *Upper Confidence Interval: the upper confidence limit of an estimate of a parameter \*(2016 only)*
- *Economy (GDP per Capita): real GDP per capita*
- *Family : social support*
- *Health (Life Expectancy): healthy life expectancy*
- *Freedom : freedom to make life choices*
- *Trust (Government Corruption): perceptions of corruption*
- *Generosity : perceptions of generosity*
- *Dystopia : each country is compared against a hypothetical nation that represents the lowest national averages for each key variable and is, along with residual error, used as a regression benchmark*



[This Photo](#) by Unknown Author is licensed under [CC BY-NC](#)



[This Photo](#) by Unknown Author is licensed under [CC BY](#)



[This Photo](#) by Unknown Author is licensed under [CC BY-NC](#)



## Hypothesis:

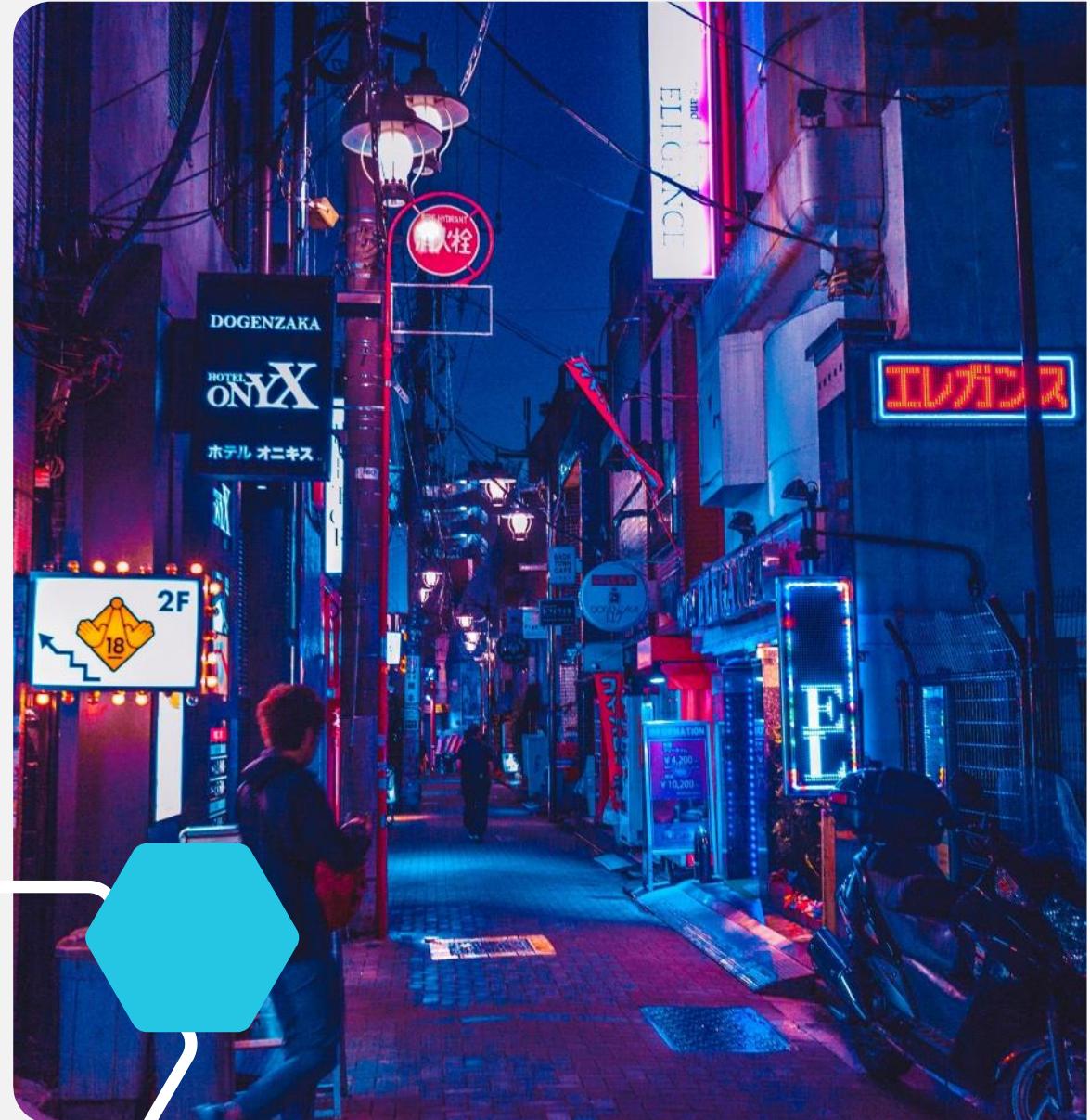
*Regions with greater freedom to make individual life choices will tend to have a higher perception of happiness.*

# Analysis

To address the hypothesis, the datasets were analyzed using several Exploratory Data Analysis (EDA) methods.

## Methods Applied:

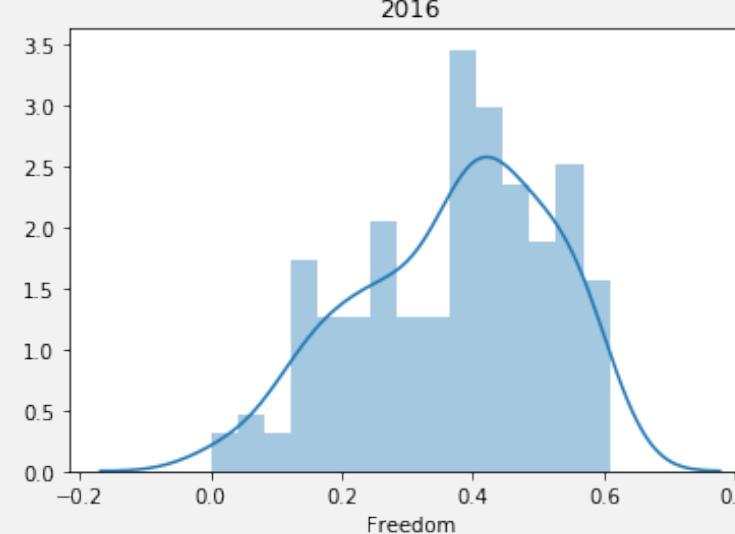
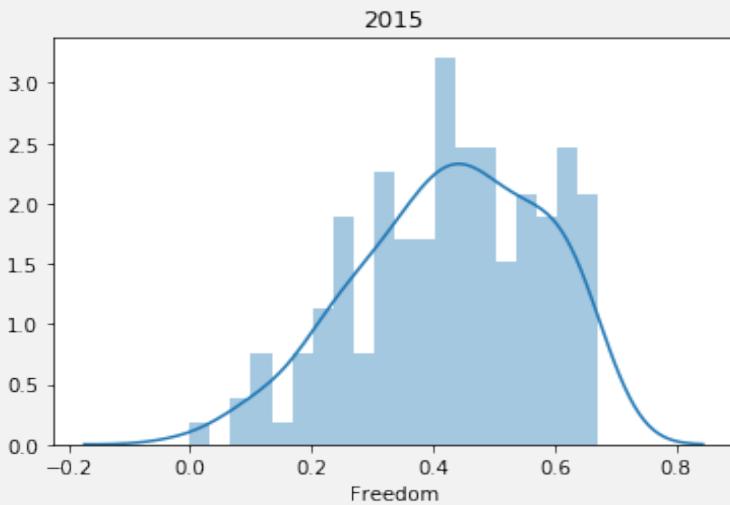
- Summary statistics analyzed (e.g. Mean, standard deviation, etc.)
- Plots (histograms & scatterplots) were created for the variables analyzed
- Probability Mass Function & Cumulative Distribution Function plotted
- Correlation and correlation test applied
- Regression analysis



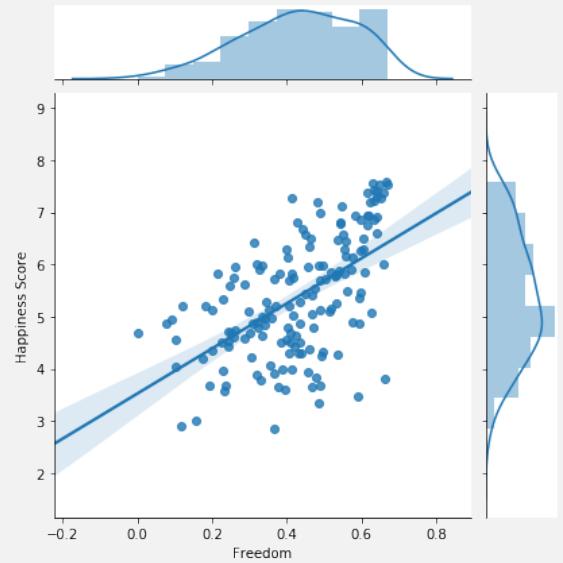
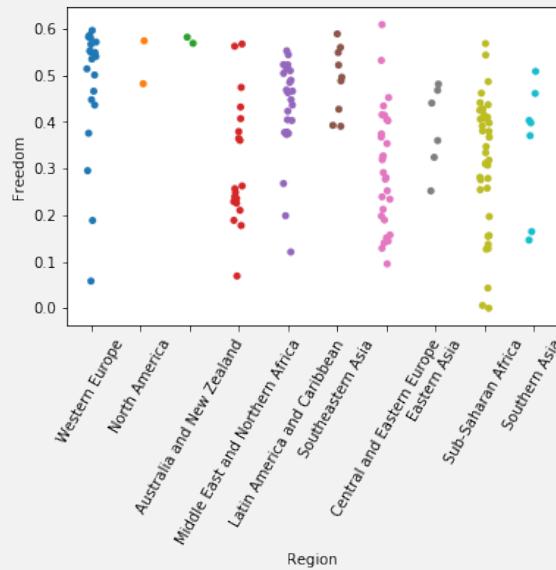
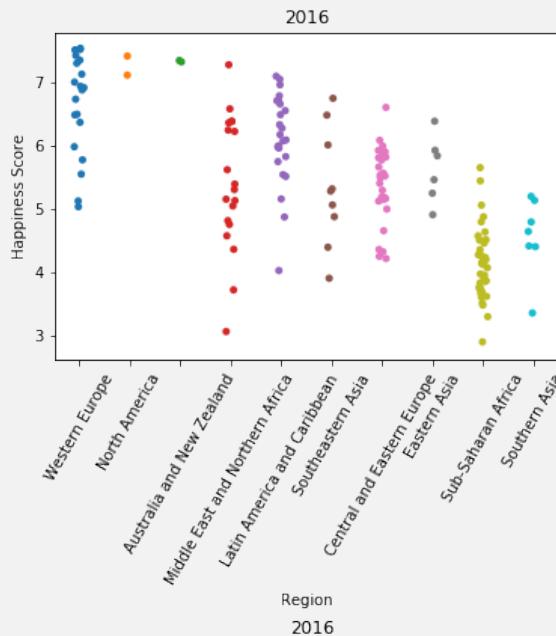
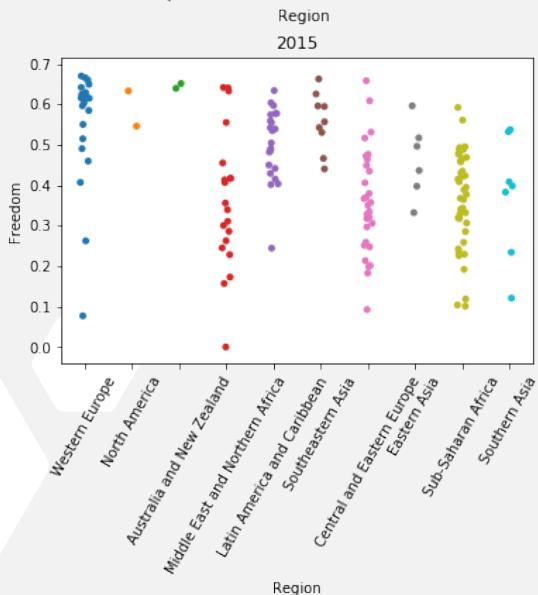
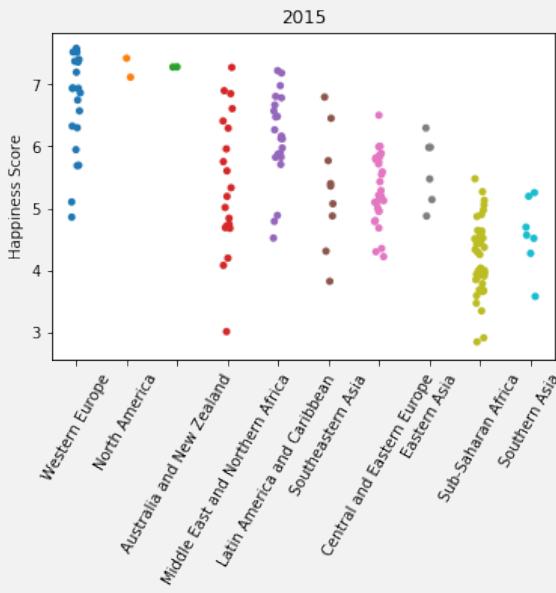
# Exploratory Data Analysis

1. Histograms
2. Scatter Plots
3. Regression Analysis

# Histograms of Happiness Score & Freedom



# Scatter Plots & Regression Analysis



# Results

- Although it was noted that *freedom* does have a *positive correlation*, other factors were found to have a *more* significant impact on the happiness score.
- Pearson's correlation was measured, and the *top three variables* with the *highest correlation* to happiness scores in the years 2015 and 2016 were found to be *economy, family, and health*.
- Regression analysis was applied to the 2015 dataset. This analysis was conducted on the *happiness scores* with *freedom* as a predictor, EDA showed a *positive linear relationship*.
- It was also noted that the scatter plots support that *regions* with the *highest* sense of *freedom* tend to have a *higher happiness score*, as reflected in the data collected from western Europe. The *mean* score for both years remained *consistent* at a 5.38. On the other hand, the sense of *freedom* did *decrease* from 0.43 in 2015 to 0.37 in the year 2016.



# Observations

- What do you feel was missed during the analysis?

*Additional demographic information, such as the socio-economic status of those individuals surveyed, could have been included in this analysis to provide a better understanding to see if their current situation has any weight on the selection of the scores.*

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

*The inclusion of a variable that measures the relationship between a person's work environment and their happiness score could assist this analysis.*

- Were there any assumptions made you felt were incorrect?

*It was assumed that the level of sense of freedom a person has would have the most significant effect on their happiness score.*

## Sources Cited

1. Gallup (n.d.) Tracking the World's Happiness.  
Retrieved from <https://www.gallup.com/analytics/247355/gallup-world-happiness-report.aspx>
2. United Nations Sustainable Development Solutions Network. (2017). World Happiness Report. [Data file].  
Retrieved from <https://www.kaggle.com/unssdsn/world-happiness>





# Thank You

 Marlene Aviles

 Maviles@my365.bellevue.edu