**Table of Contents**

[Introduction: 1](#__RefHeading___Toc1764_2688910746)

[Dataset: 1](#__RefHeading___Toc1766_2688910746)

[Research Questions: 1](#__RefHeading___Toc1768_2688910746)

[Formulation of Null and Alternative Hypotheses 1](#__RefHeading___Toc1770_2688910746)

[Visualization: 1](#__RefHeading___Toc1772_2688910746)

[3](#__RefHeading___Toc1774_2688910746)

[Analysis: 4](#__RefHeading___Toc1776_2688910746)

[Kendall rank correlation test: 5](#__RefHeading___Toc1778_2688910746)

[Significance level: 5](#__RefHeading___Toc1780_2688910746)

[Spearman rank correlation coefficient: 5](#__RefHeading___Toc1782_2688910746)

[Conclusion: 6](#__RefHeading___Toc1784_2688910746)

# Introduction:

This report contains all the information related to the World Happiness dataset from 2015 to 2019

## Dataset:

[The dataset is taken from](https://www.kaggle.com/jpmiller/publicassistance) <https://data.world/sventurini/world-happiness-2015-2019>

It covers gdp,life expectancy,generosity and coorruption and freedom to choices scores for all the countries from year 2015 to 2019

The dataset can benefit greatly from additional content. Economics, additional demographics, administrative costs and more.

## Research Question:

The research question for this study is “Is there a correlation between life expectancy and generosity?”

Here, we want to analyses if the generosity can impact the life expectancy of a person.

## Formulation of Null and Alternative Hypotheses

**Null hypothesis**

There is no correlation between life expectancy and generosity

H0:ρ=0

**Alternative Hypothesis:**

There is a correlation between life expectancy and generosity

H1: ρ≠ 0

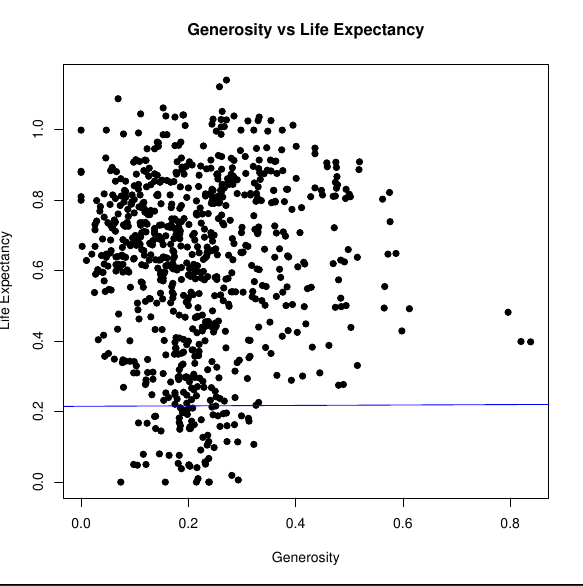
# Visualization:

Different visualization was used in this study to assess the true picture of data.

**Generosity vs Life Expectancy :**

The below scatter plot shows the relation between the generosity and life expectancy scores.

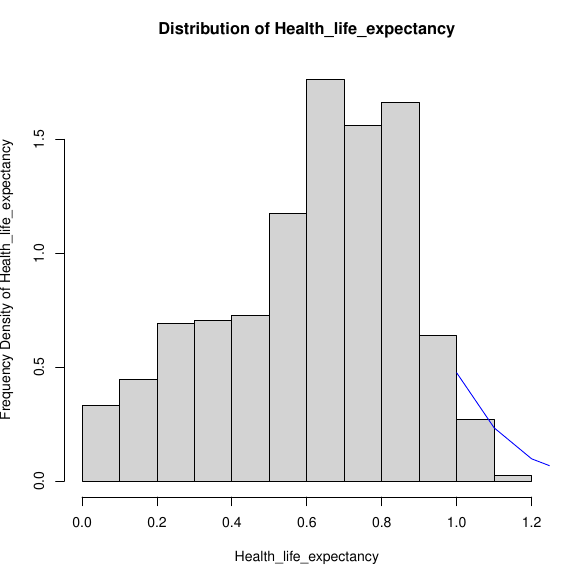
As both the variables are numeric we can unserstand the relation visually by drawing the scatter plot



In the above graph we can clearly observe all the points are scattered in a bubble shape and there is no clear linear regression line observed.

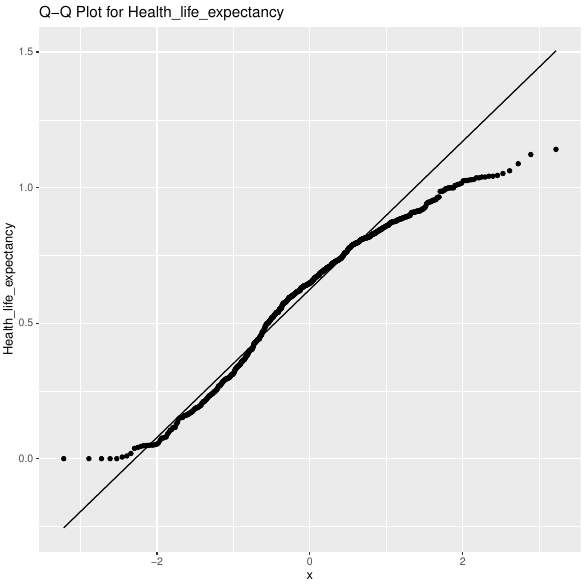
**Frequency Plot:**

From the below plot the frequency of health expectancy score showing a twisted bell curve to the right side. The normal line is not clearly observed as bell shaped curve.



# 

**Normal Q-Q Plot :**



Normal Q-Q Plot was examined to verify the assumption of normality. The above graph is showing that there is a linear trend

# Analysis:

From the Above Visualization we got mixed results about the normality of data from Frequency plot and Q-Q Plot. Shapiro Wilk Normality Test can be applied to check the normality distribution of data. Based on the results from the test appropriate Correlation test can be applied.

**Shapiro Wilk Normality Test:**

data: hapiness\_data$Generosity

W = 0.94599, p-value = 2.747e-16

data: hapiness\_data$Health\_life\_expectancy

W = 0.96648, p-value = 2.093e-12

From the Above two values we can see that the p-value is way less than 0.05 so, we can not say that the both the variables follow normal distribution.

As the data is not showing normality we cant apply the Pearson Correlation Test.

## **Kendall rank correlation test:**

The Kendall rank correlation coefficient or Kendall’s tau statistic is used to estimate a rank-based measure of association. This test may be used if the data do not necessarily come from a bi variate normal distribution

## **Significance level:**

ɑ=0.05

**Calculations:**

data: hapiness\_data$Generosity and hapiness\_data$Health\_life\_expectancy

z = 0.12568, p-value = 0.9

sample estimates:

tau

0.003008852

## **Spearman rank correlation coefficient:**

Spearman’s rho statistic is also used to estimate a rank-based measure of association. This test may be used if the data do not come from a bivariate normal distribution

**Calculations:**

data: hapiness\_data$Generosity and hapiness\_data$Health\_life\_expectancy

S = 78811525, p-value = 0.7551

sample estimates:

rho

0.01117045

# Conclusion:

As, the p-value is greater than significance level alpha=0.05 from the above tests, so we fail to reject H0 and conclude that there is no correlation between generosity and the life expectancy of a person.