World Population Data

Analysis with Python

Elizabeth Asuncion

22 April 2023

Project Overview

Data set: World Population Data

- Life Expectancy, Birth Rate, Death Rate
- Height, Weight, BMI
- Quality of Life: Safety, Stability, Climate, etc.
- IQ & Education Expenditures
- Population Density

Goal: Find out what might make countries better

Tools: Python

- Pandas - Data Exploration and Manipulation
- Statistics - Dataset Description and Summary
- Seaborn - Data Storytelling and Visualization

Hypothesis 1

Higher country health scores are related to longer life and better quality of life.

Hypothesis 1 – Specifics

- Healthier people live longer: Health Score is positively correlated with Life Expectancy for both females and males.
- Countries with higher health scores tend to have lower death rates: Health Score is negatively correlated with death rate.
- Better quality of life goes hand in hand with healthier lifestyles: Health Score is positively correlated with quality of life indicators such as Rights, Stability, and Safety.

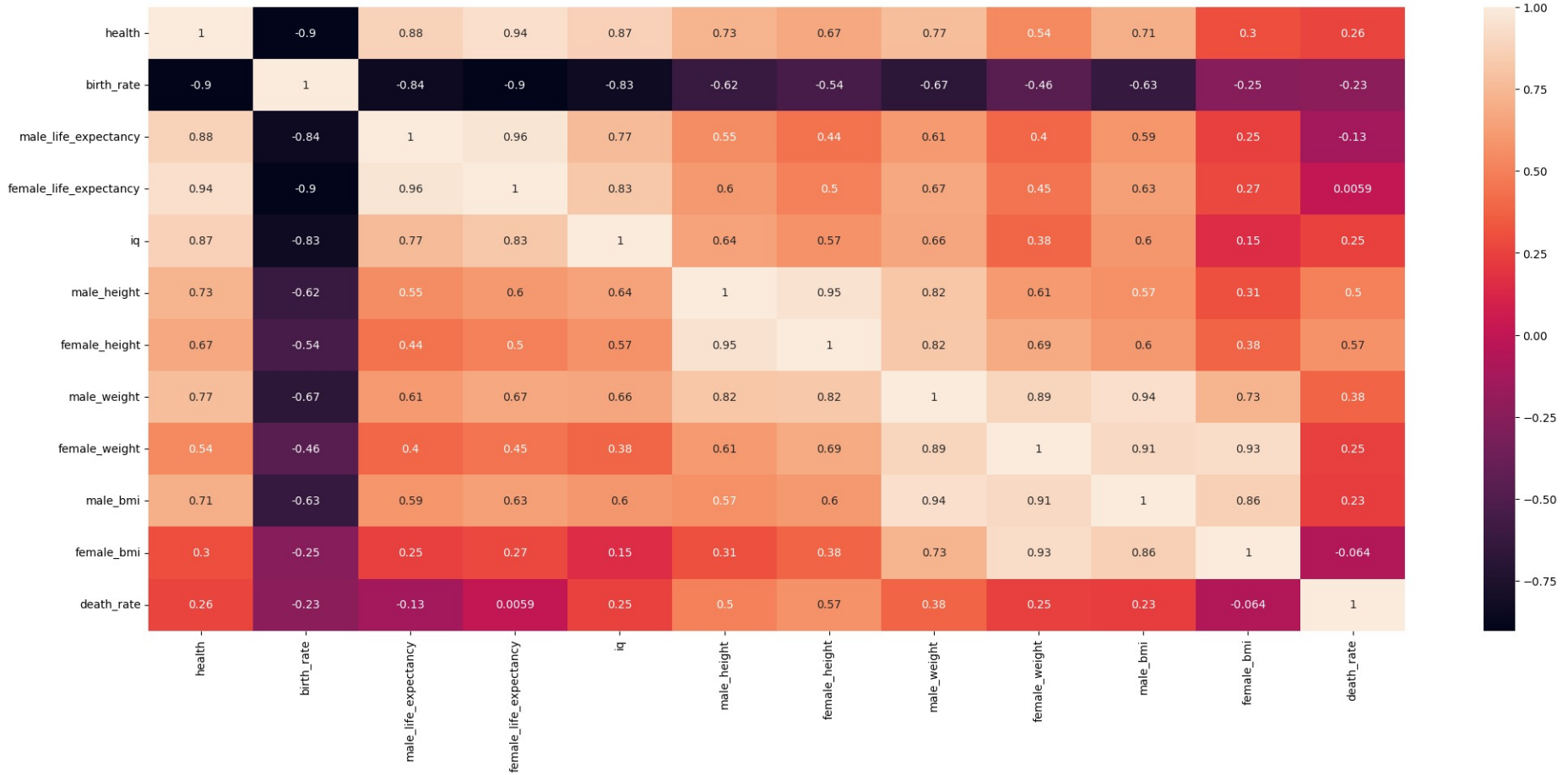
Correlation

Correlation Coefficients between Health Score and Other Variables

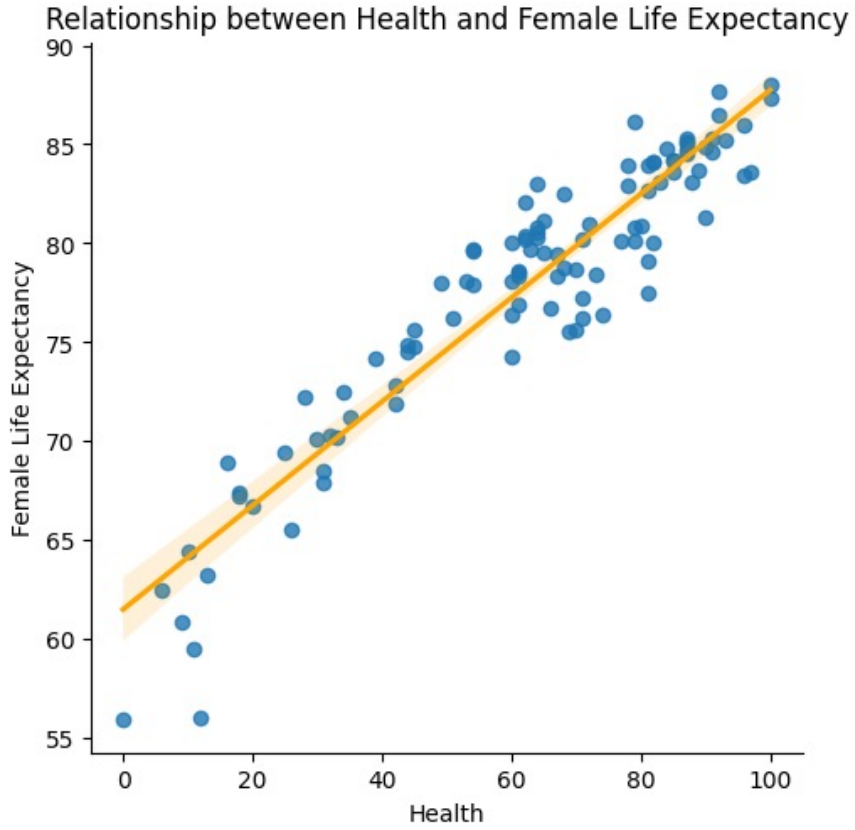
health	1.000000
female_life_expectancy	0.939634
male_life_expectancy	0.879709
iq	0.869019
male_weight	0.772444
rights	0.761785
male_height	0.727495
male_bmi	0.705656
stability	0.693995
female_height	0.666702
safety	0.665243
education_expenditure_per_inhabitant	0.604932
female_weight	0.537620
female_bmi	0.299138
death_rate	0.261467
pop_per_km_sq	0.065392
popularity	0.055298
area	0.035834
costs	-0.365072
climate	-0.435529
daily_max_temp	-0.657592
birth_rate	-0.901974

Heat Map

Correlations between Health-Related Metrics



Health Score & Life Expectancy



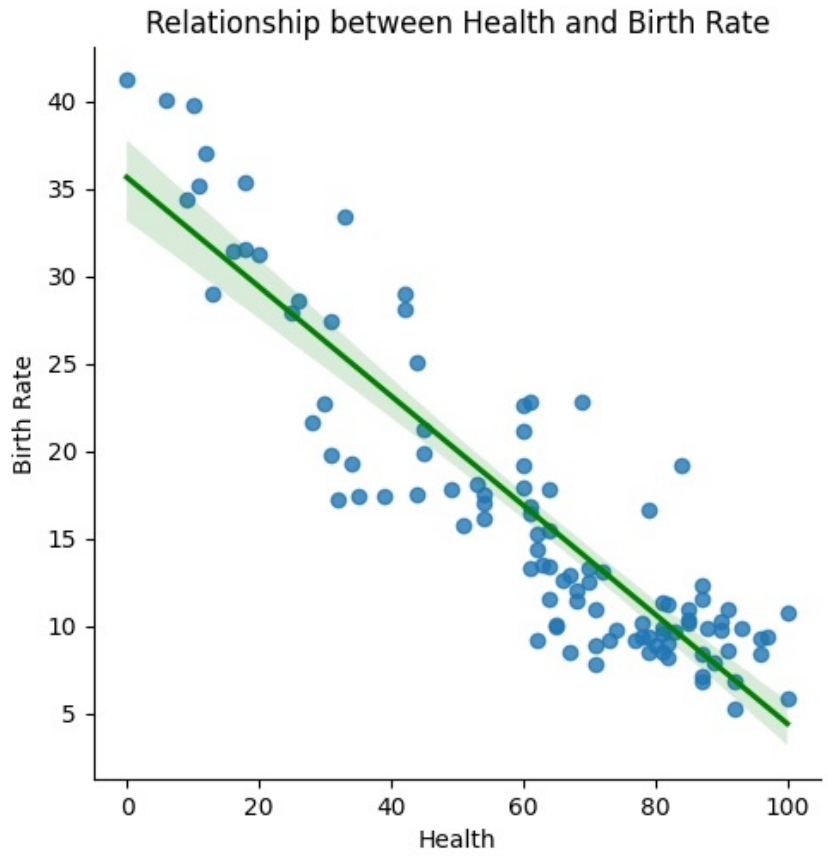
Positive Correlation

Hypothesis 1 – Specifics

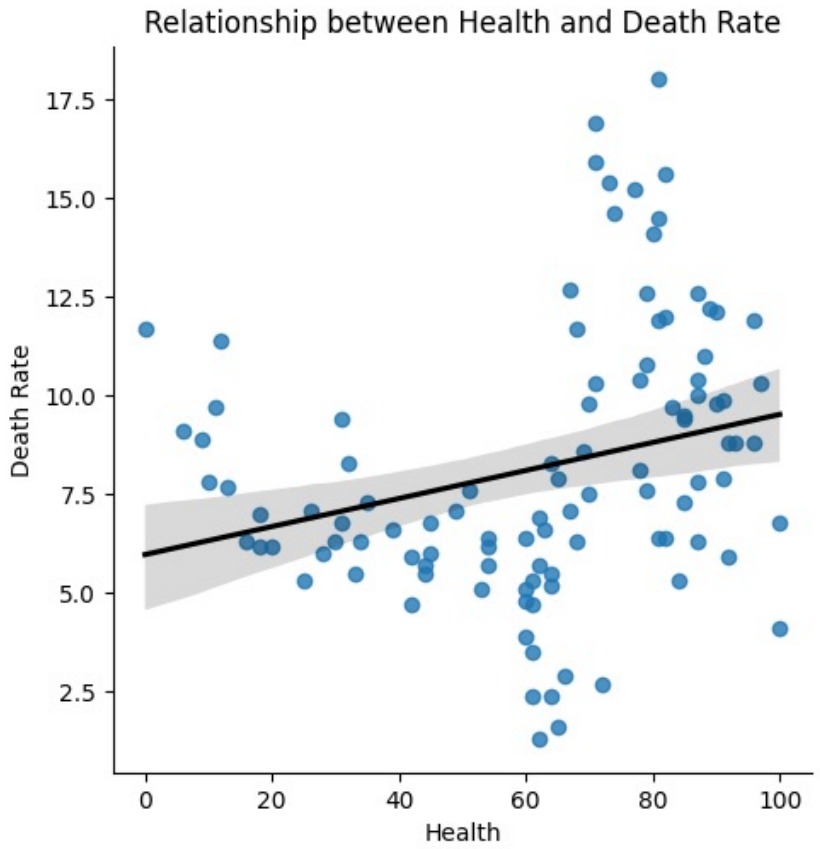
Healthier people live longer: Health Score is positively correlated with Life Expectancy for both females and males.

- Countries with higher health scores tend to have lower death rates: Health Score is negatively correlated with death rate.
- Better quality of life goes hand in hand with healthier lifestyles: Health Score is positively correlated with quality of life indicators such as Rights, Stability, and Safety.

Health Score & Birth/Death Rates



Negative Correlation



Weak Positive Correlation

Hypothesis 1 – Specifics

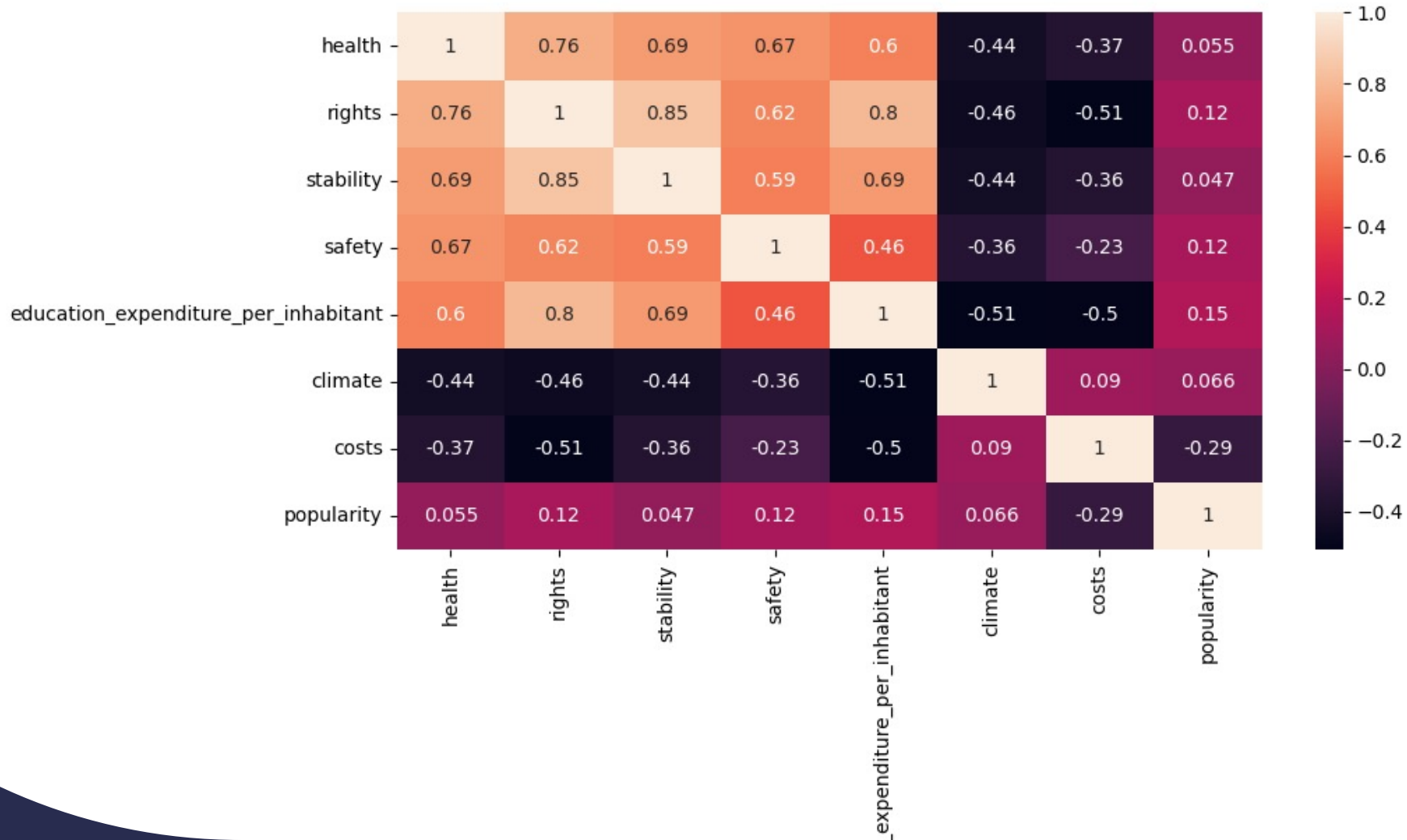
Healthier people live longer: Health Score is positively correlated with Life Expectancy for both females and males.

Countries with higher health scores tend to have lower death rates: Health Score is negatively correlated with death rate.

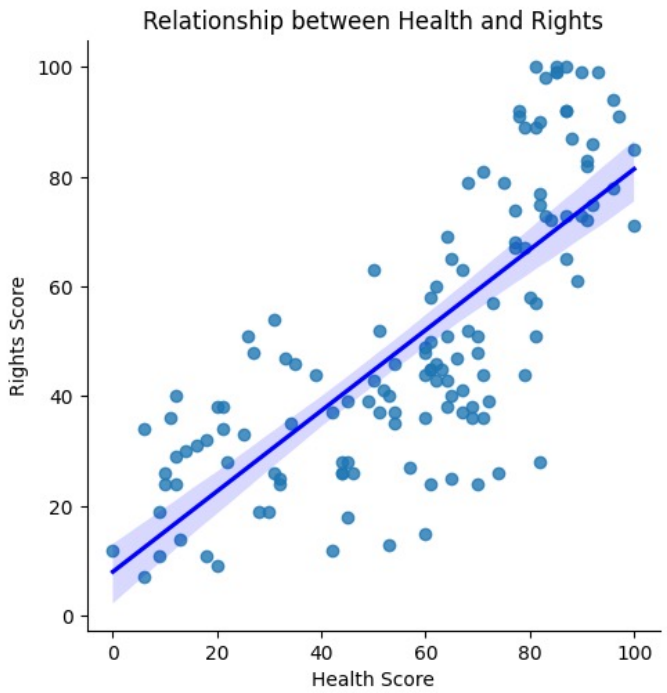
- Better quality of life goes hand in hand with healthier lifestyles: Health Score is positively correlated with quality of life indicators such as Rights, Stability, and Safety.

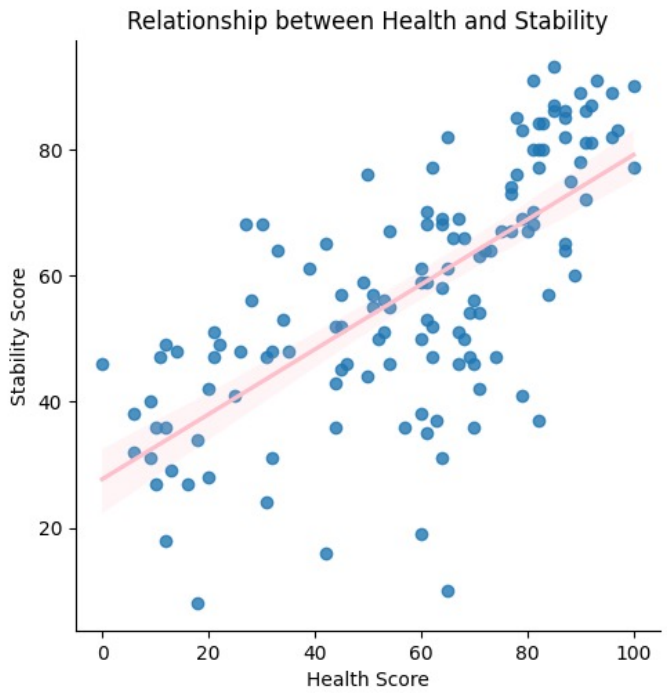
Heat Map

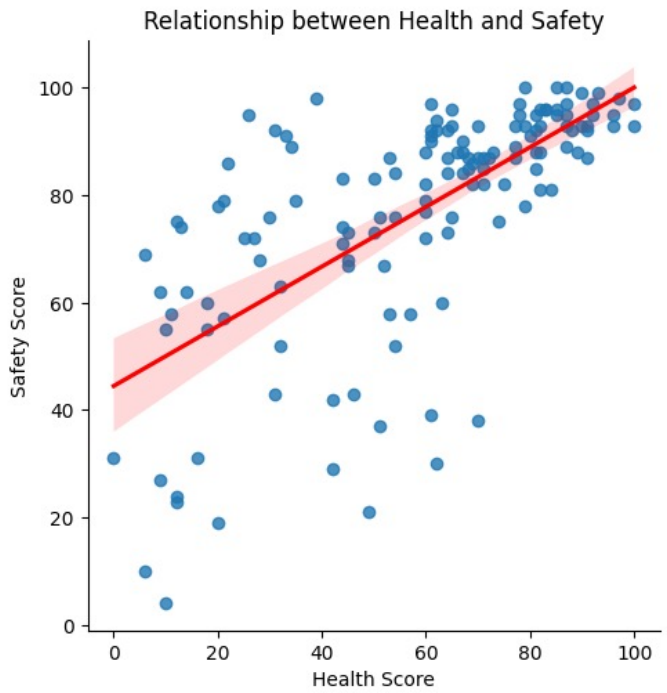
Correlations between Health Score and Quality of Life Metrics



Health & Quality of Life Indicators







Positive Correlations

Hypothesis 1 – Specifics

- ✓ Healthier people live longer: Health Score is positively correlated with Life Expectancy for both females and males.
- ✗ Countries with higher health scores tend to have lower death rates: Health Score is negatively correlated with death rate.
- ✓ Better quality of life goes hand in hand with healthier lifestyles: Health Score is positively correlated with quality of life indicators such as Rights, Stability, and Safety.

Conclusions and Insights

- Causality is difficult to determine: healthier people live longer
  - Does life expectancy lead to higher health scores?
  - Does higher health score lead to longer life?
  - Does better quality of life lead to better health?
  - Does higher health score lead to better quality of life?
- If you want to live longer and have a good quality of life, it’s better to live in countries with good health

Hypothesis 2

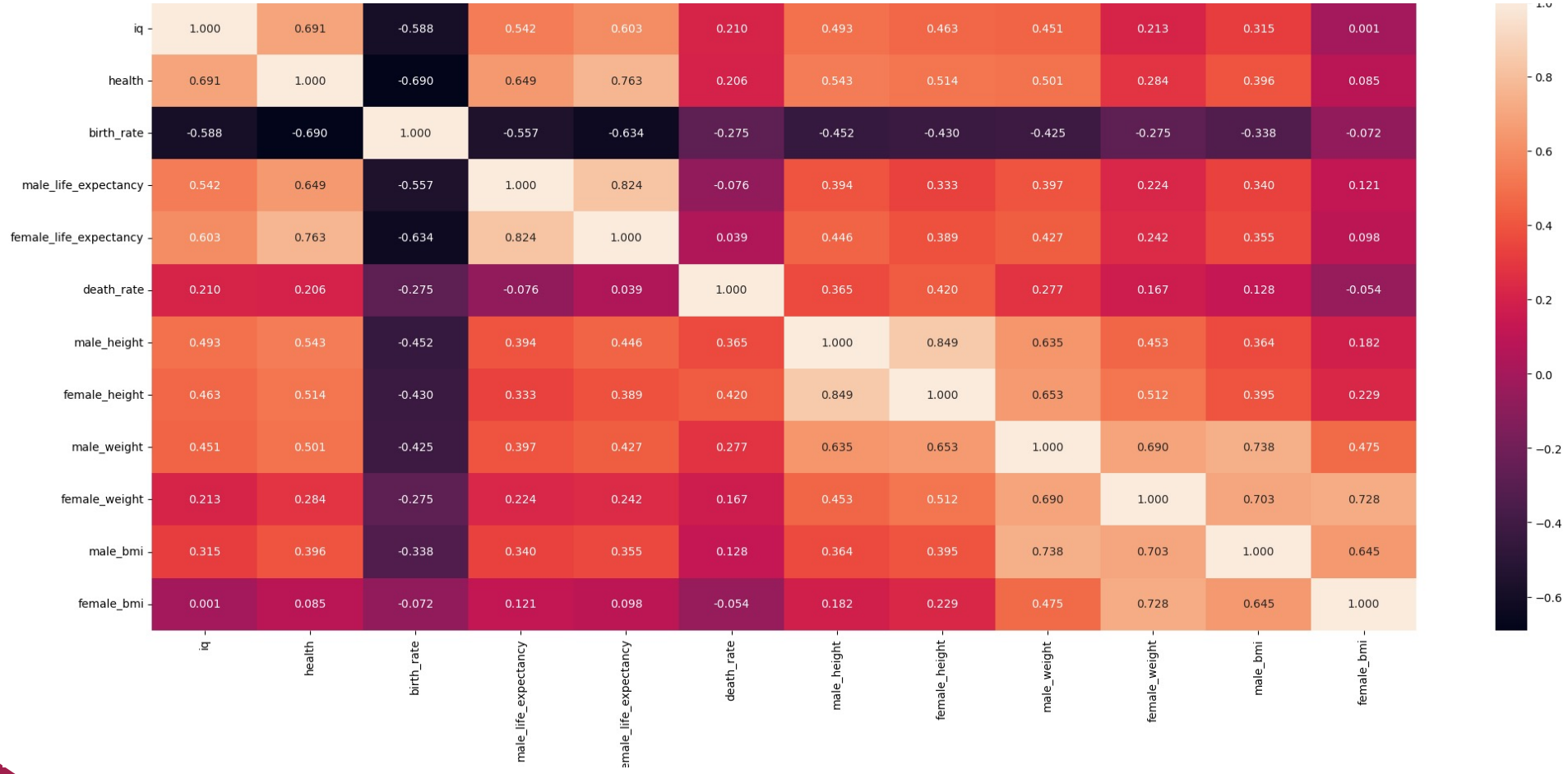
Higher IQ scores are related to longer life expectancy and better quality of life.

Hypothesis 2 – Specifics

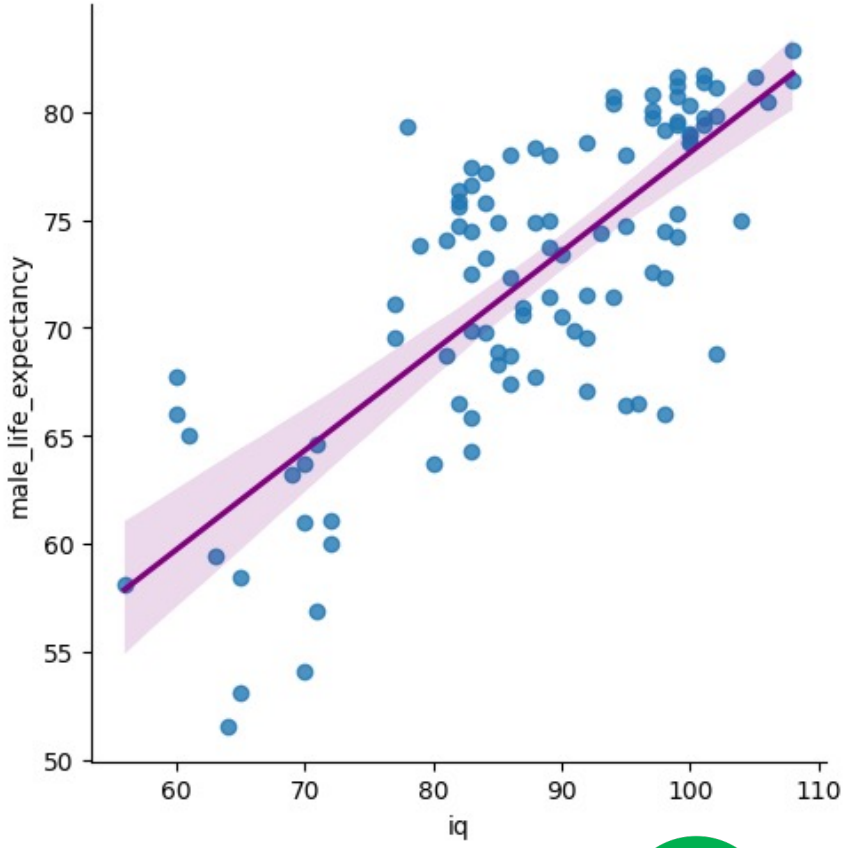
- People with higher IQ scores are healthier and live longer: IQ is positively correlated with Life Expectancy for both females and males.
- Countries with higher IQ scores tend to have lower death rates: IQ is negatively correlated with death rates.
- Countries with higher IQ scores have lower birth rates: IQ is negatively correlated with birth rate.
- People with higher IQ scores live in ountries with better quality of life : IQ is positively correlated with quality of life indicators such as Rights, Stability, and Safety.

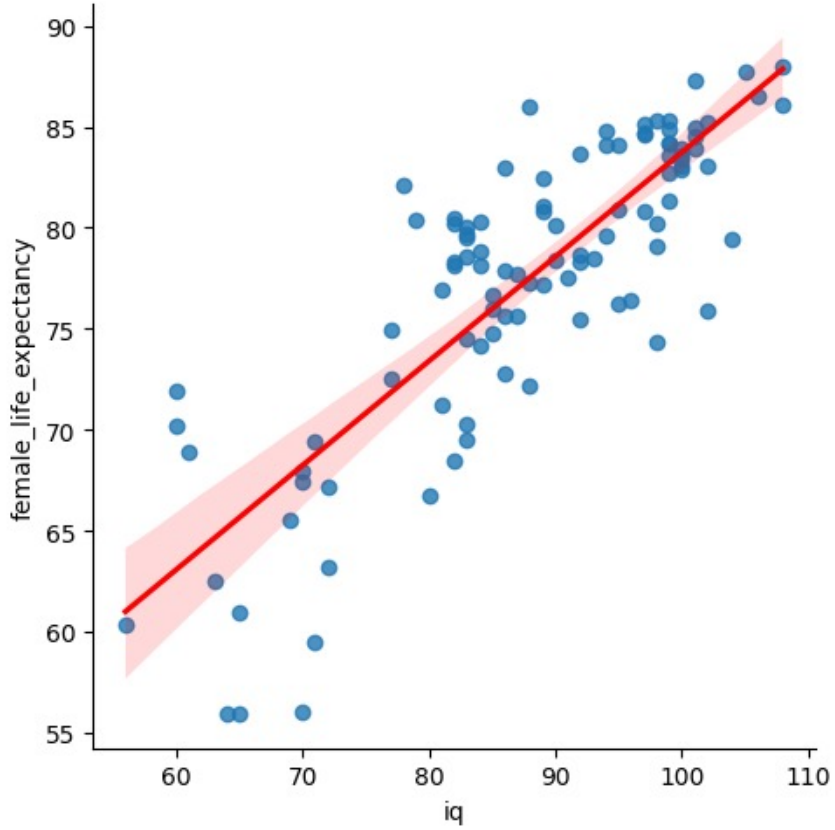
Heat Map

Correlation between IQ and Health-Related Metrics



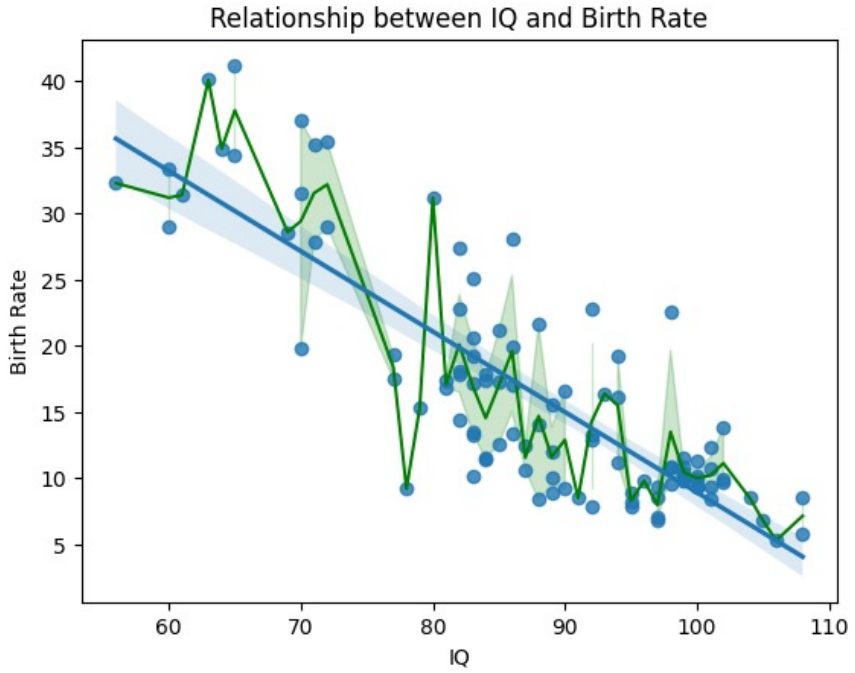
IQ & Life Expectancy

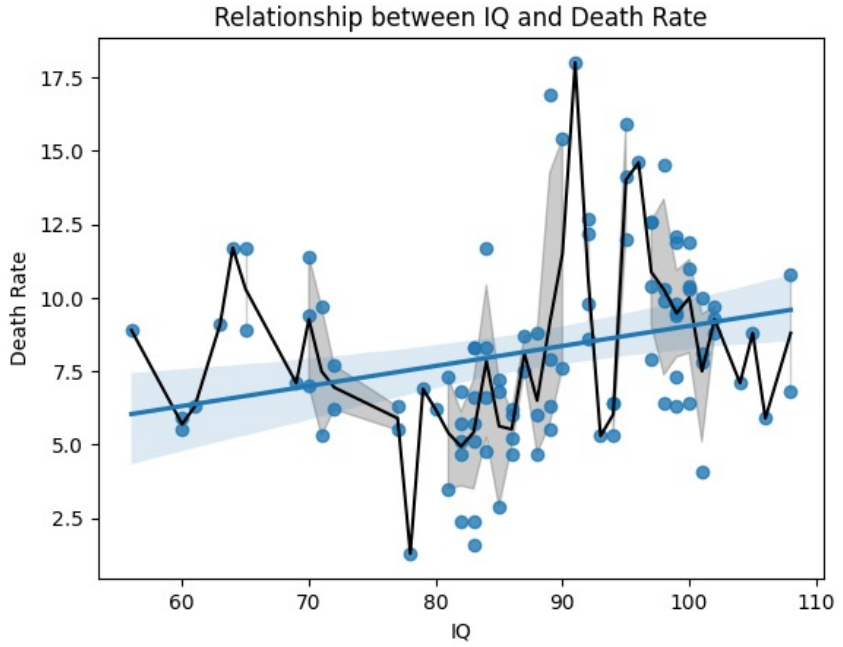




Positive Correlations

IQ & Birth/Death Rates



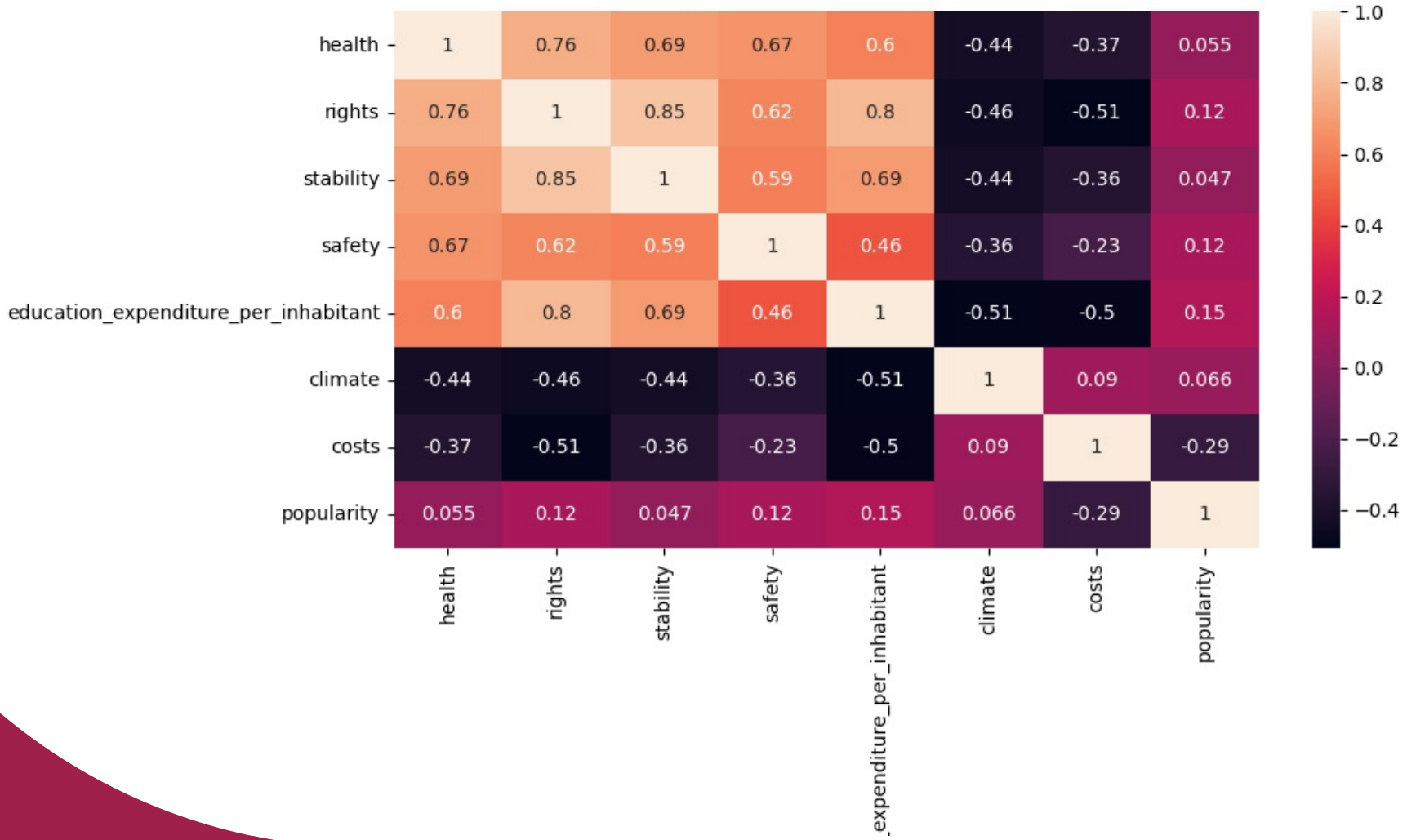


Negative Correlation

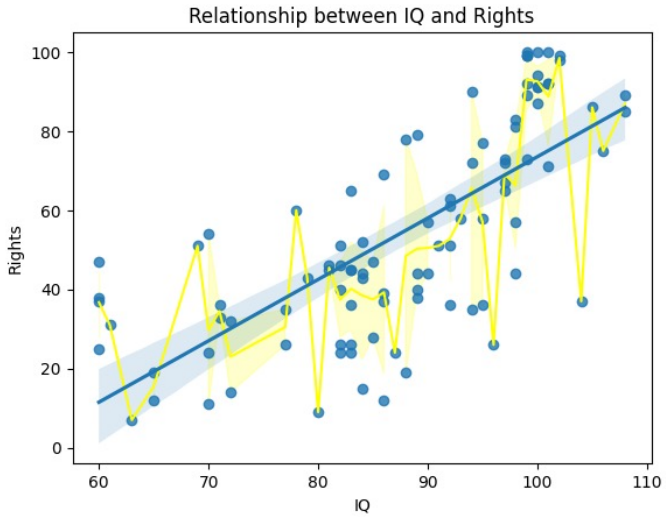
Weak Positive Correlation

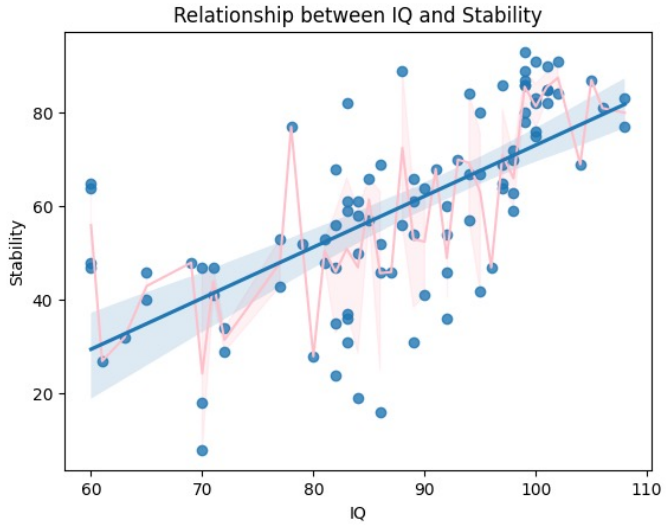
Heat Map

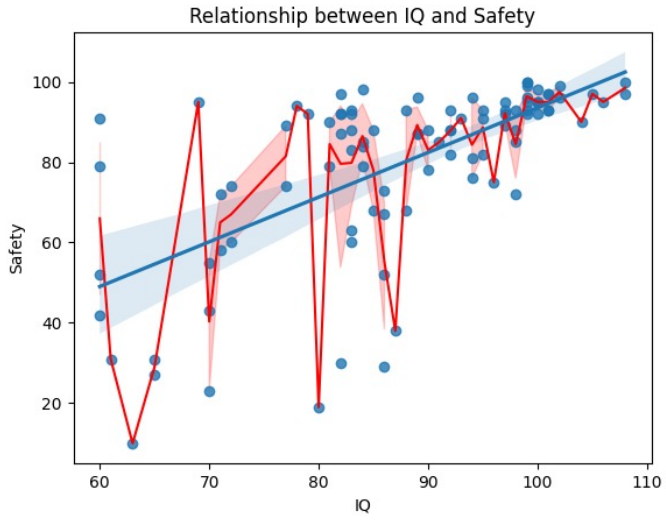
Correlations between IQ Score and Quality of Life Metrics



IQ & Quality of Life Indicators







Positive Correlations

Hypothesis 2 – Specifics

- People with higher IQ scores are healthier and live longer: IQ is positively correlated with Life Expectancy for both females and males.
- Countries with higher IQ scores tend to have lower death rates: IQ is negatively correlated with death rates.
- Countries with higher IQ scores have lower birth rates: IQ is negatively correlated with birth rate.
- People with higher IQ scores live in ountries with better quality of life : IQ is positively correlated with quality of life indicators such as Rights, Stability, and Safety.

Conclusions and Insights

- There may be a relationship between average IQ scores and health outcomes, as countries with higher average IQ scores may have better access to healthcare and make better health-related decisions.
- IQ is positively correlated with Life Expectancy for both females and males: countries with higher Life Expectancy may have a more educated and healthier population, which in turn may lead to higher average IQ scores.
- IQ is negatively correlated with death rate: This suggests that countries with higher death rates may have a lower average IQ, potentially due to factors such as lower access to healthcare or lower socio-economic status.

Data Limitations

- Missing values
- Data not checked for accuracy, completeness, uniqueness, and timeliness
- Figures not explained, for example:
  - What exactly did the Quality of Life metrics represent?
  - How were Quality of Life metrics calculated?
- Data origin sources unknown so possible biases not considered

Remaining Analyses and Next Steps

- Repeating the analysis with cleaned data
- Repeating the analysis with additional data (e.g. data related to economic development)
- Exploring reasons for the lack of statistical difference for certain metrics
- Repeating analyses with different segments of data
- Developing plan to determine causality
- Determining remaining analyses that would support causality claims


THANK YOU!