# Exploratory Analysis of the 2023 World Happiness Report

Talha Iqbal - 217967951

EECS3401

Dr. Ruba Alomari

3/24/2024

## Framing the problem and looking at the big picture

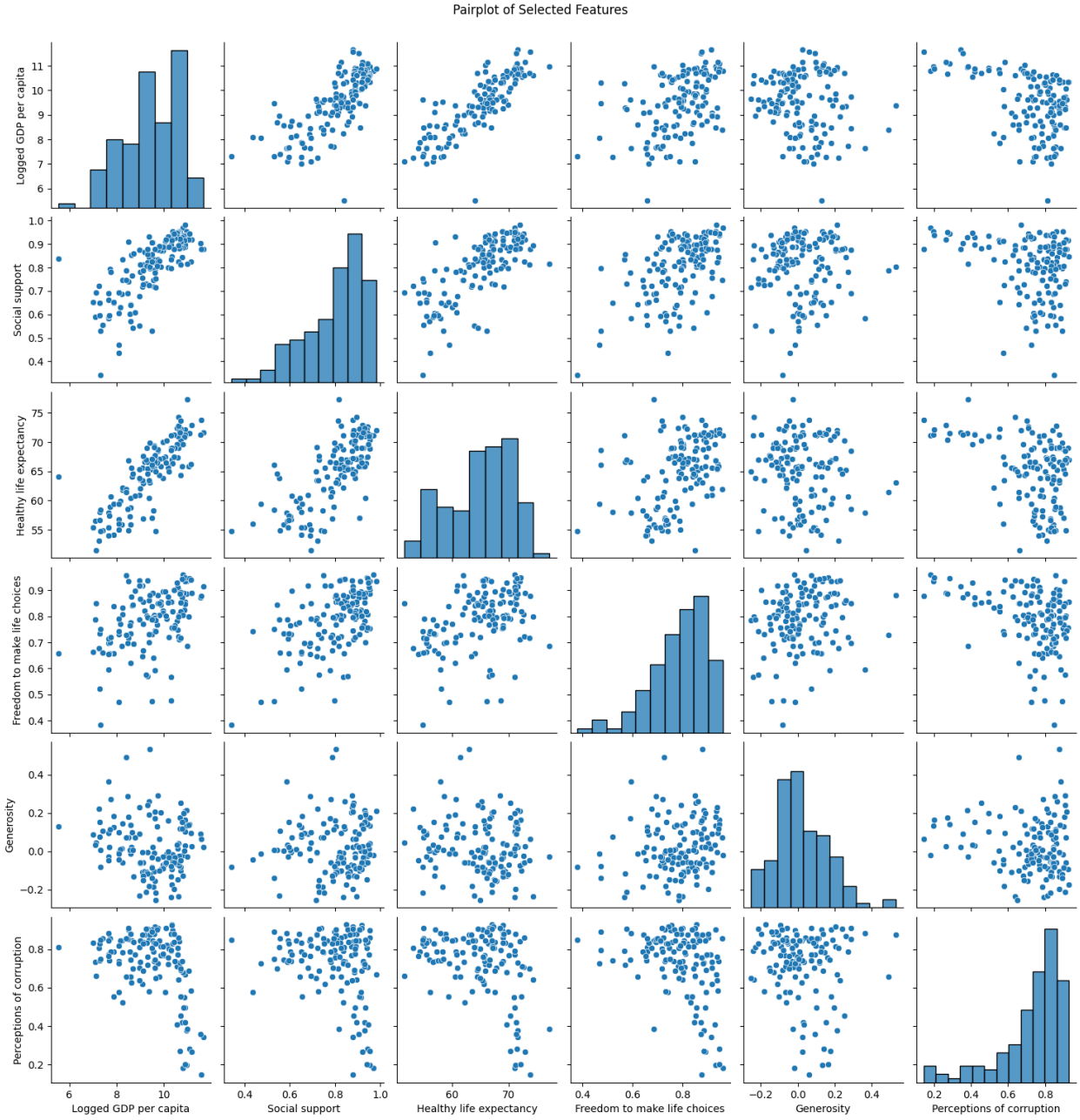
- This problem requires supervised learning.
- It is a regression task.
- Batch learning is required.

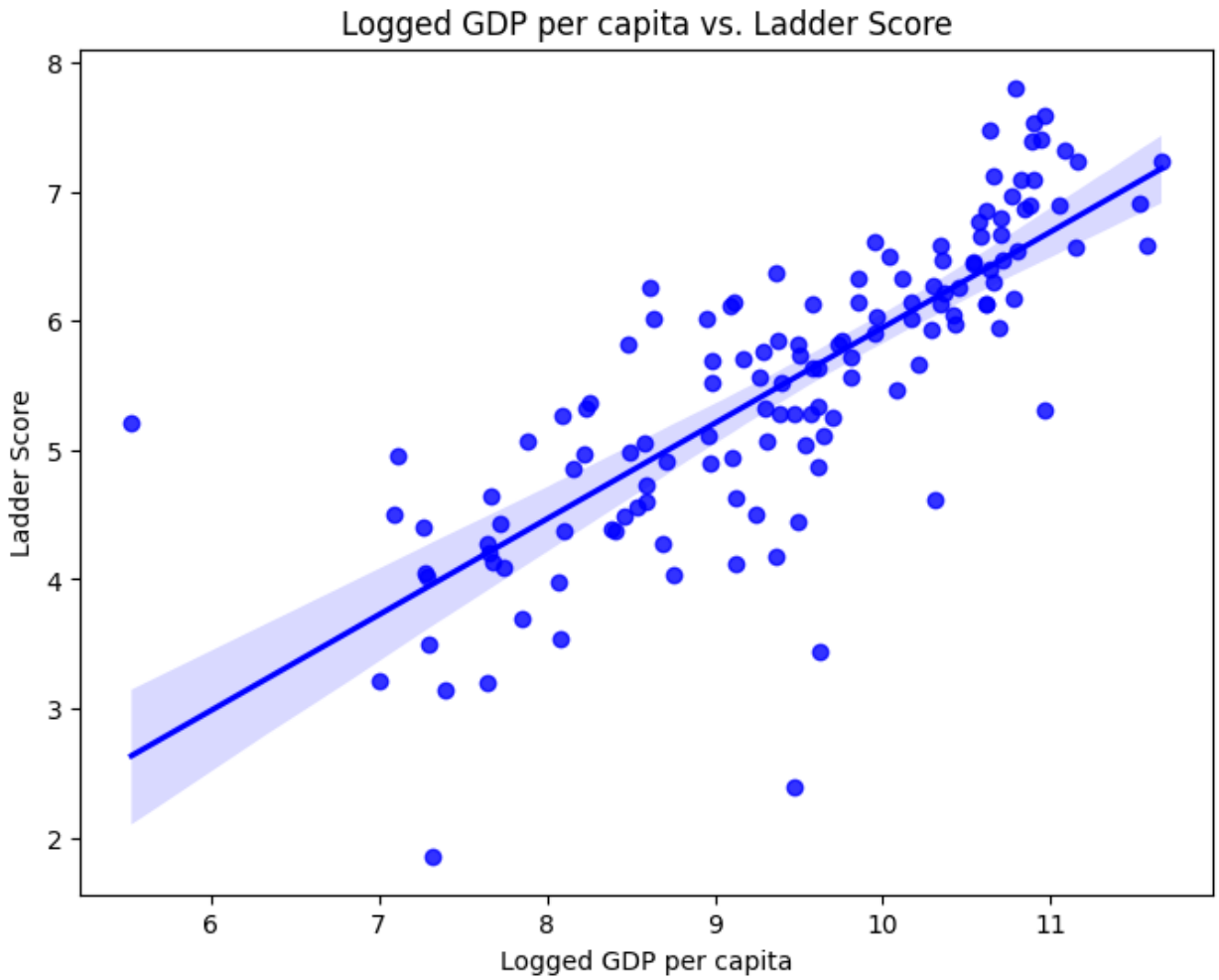
- The problem: Determine the most influential factors when it comes to a country's happiness and create a model that will predict a country's happiness given such factors.
- Benefit of solution: Understanding the factors driving happiness aids companies in directing Corporate Social Responsibility efforts. The insight of this project informs investments in social programs, and ethical business practices. By aligning CSR initiatives with happiness drivers, companies contribute to community well-being and enhance their societal impact and reputation.

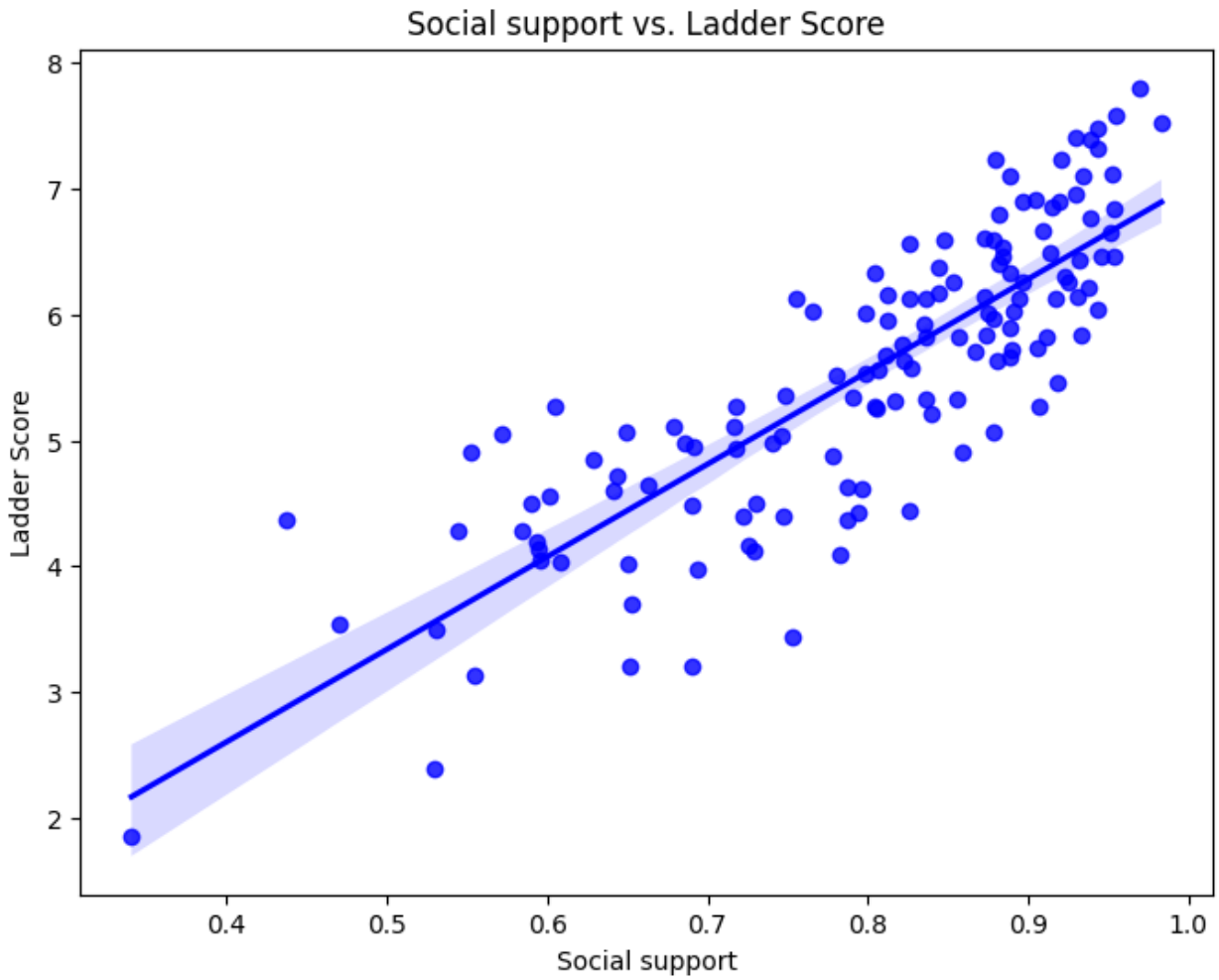
## Description of the dataset and 3+ graphs of EDA

The World Happiness Report is a landmark survey of the state of global happiness . 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.

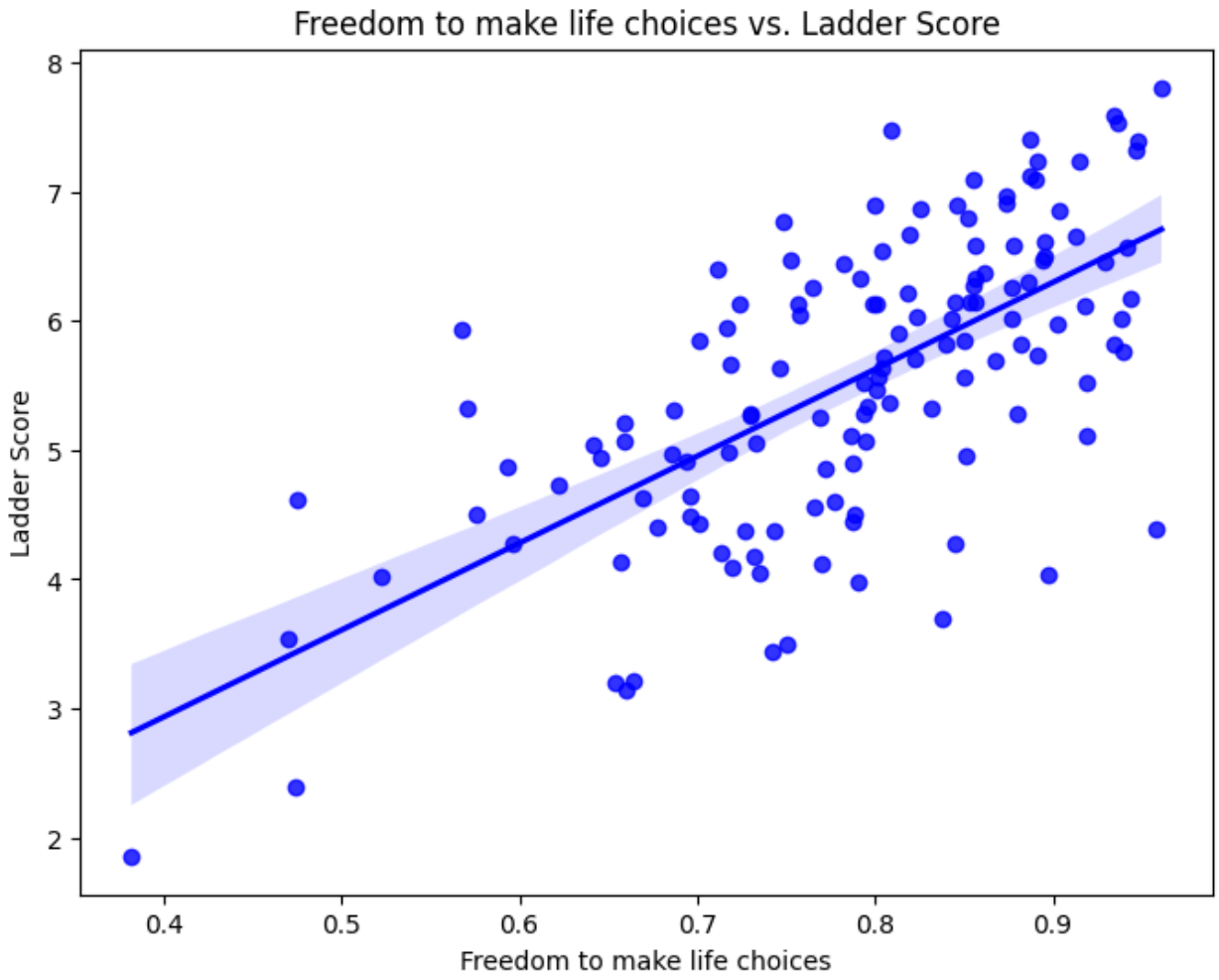

Histogram of Ladder Score
Frequency
Ladder Score
0
5
10
15
20
2
3
4
5
6
7
8

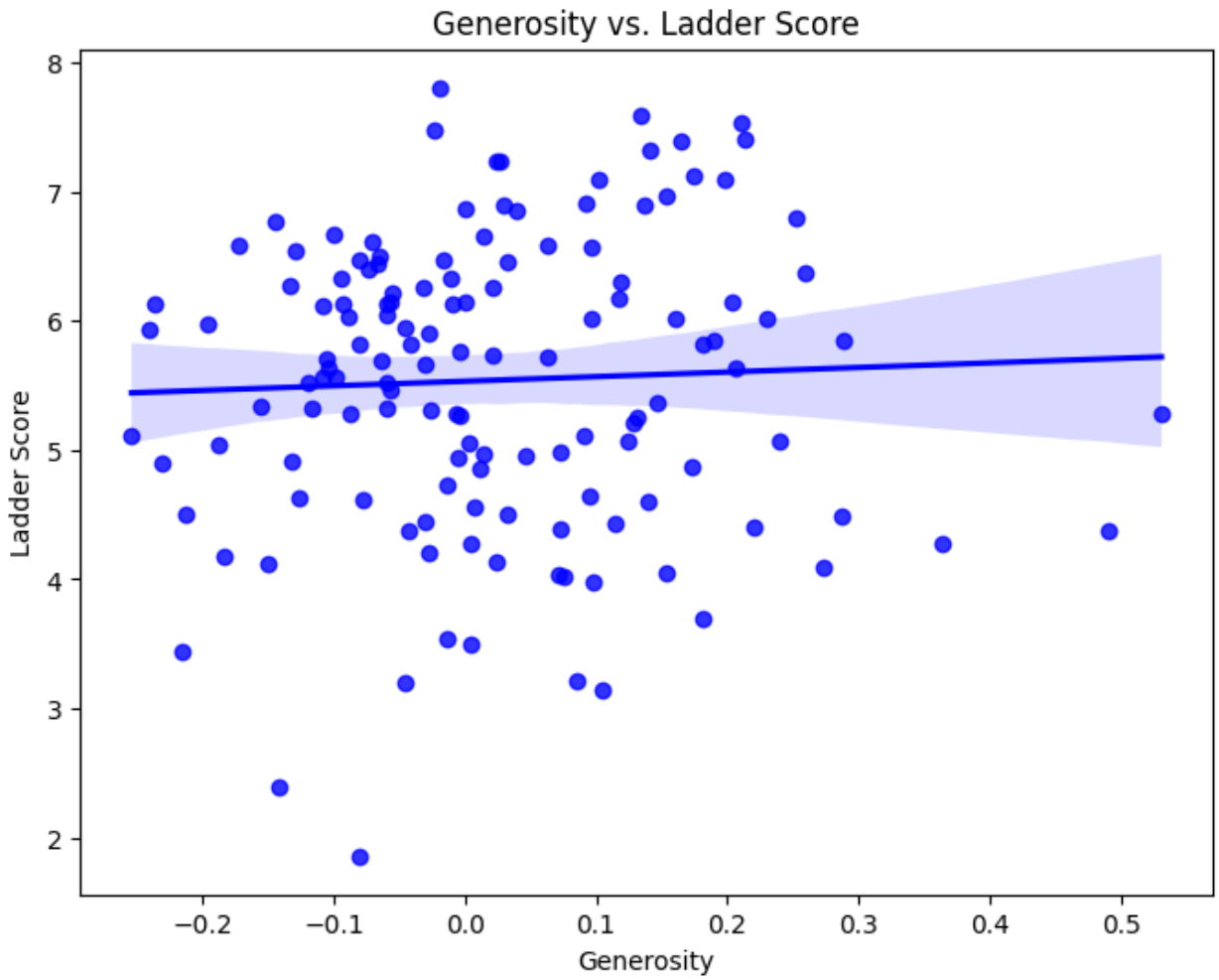

Pairplot of Selected Features
Logged GDP per capita
Social support
Healthy life expectancy
Freedom to make life choices
Generosity
Perceptions of corruption

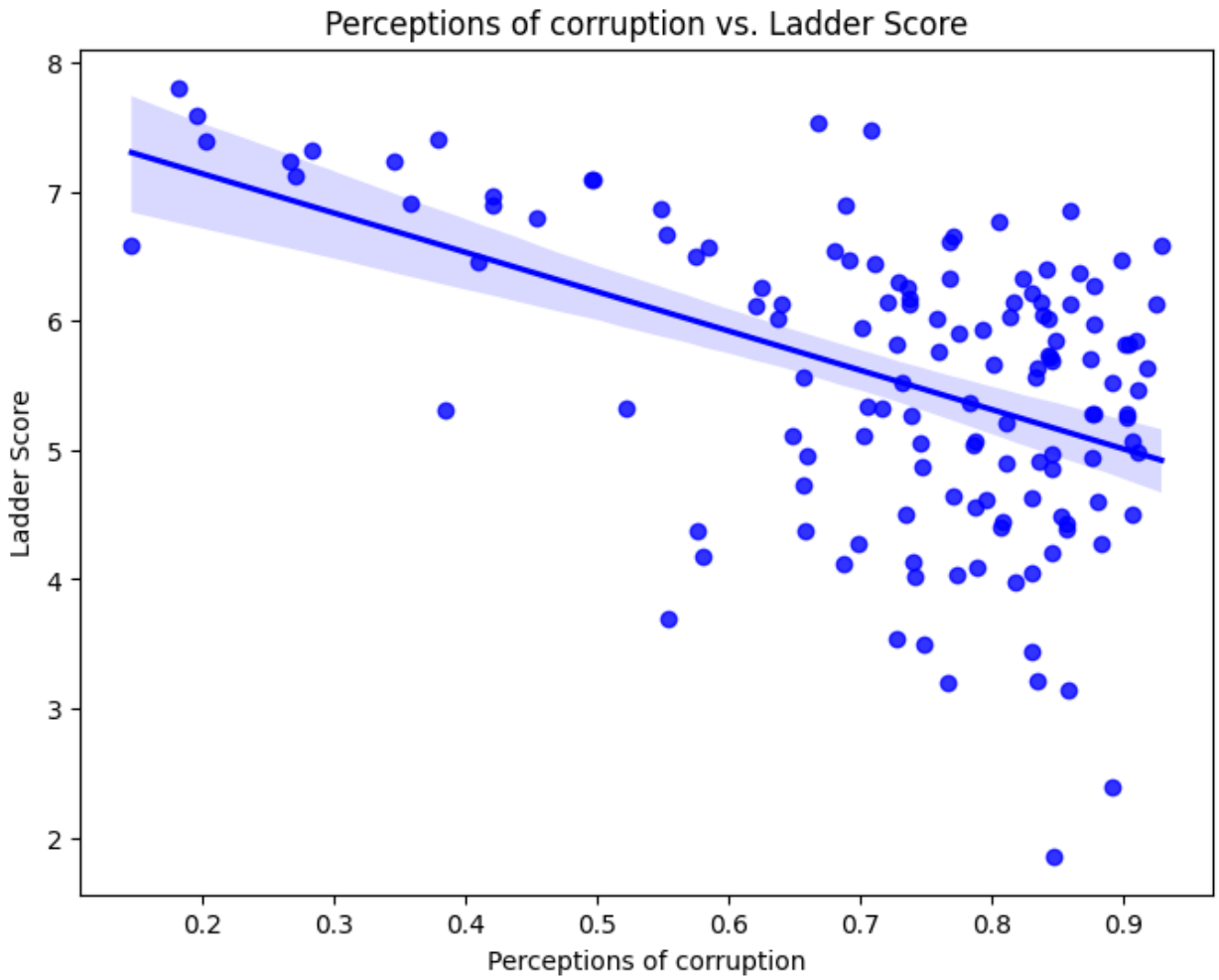

Logged GDP per capita vs. Ladder Score
Ladder Score
Logged GDP per capita
8
7
6
5
4
3
2
6
7
8
9
10
11

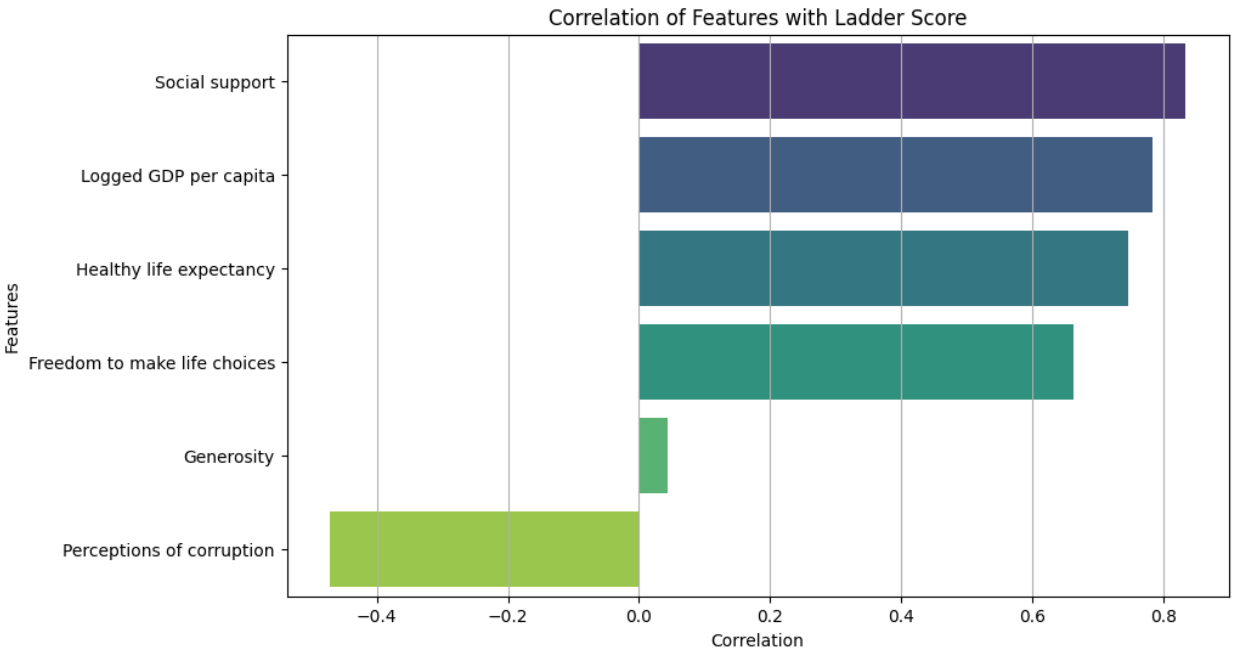

Social support vs. Ladder Score
Ladder Score
Social support
8
7
6
5
4
3
2
0.4
0.5
0.6
0.7
0.8
0.9
1.0


Healthy life expectancy vs. Ladder Score
Ladder Score
Healthy life expectancy
8
7
6
5
4
3
2
55
60
65
70
75


Freedom to make life choices vs. Ladder Score
Ladder Score
Freedom to make life choices
8
7
6
5
4
3
2
0.4
0.5
0.6
0.7
0.8
0.9


Generosity vs. Ladder Score
Ladder Score
Generosity
8
7
6
5
4
3
2
−0.2
−0.1
0.0
0.1
0.2
0.3
0.4
0.5


Perceptions of corruption vs. Ladder Score
Ladder Score
Perceptions of corruption
8
7
6
5
4
3
2
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9


Correlation of Features with Ladder Score
Features
Social support
Logged GDP per capita
Healthy life expectancy
Freedom to make life choices
Generosity
Perceptions of corruption
−0.4
−0.2
0.0
0.2
0.4
0.6
0.8
Correlation

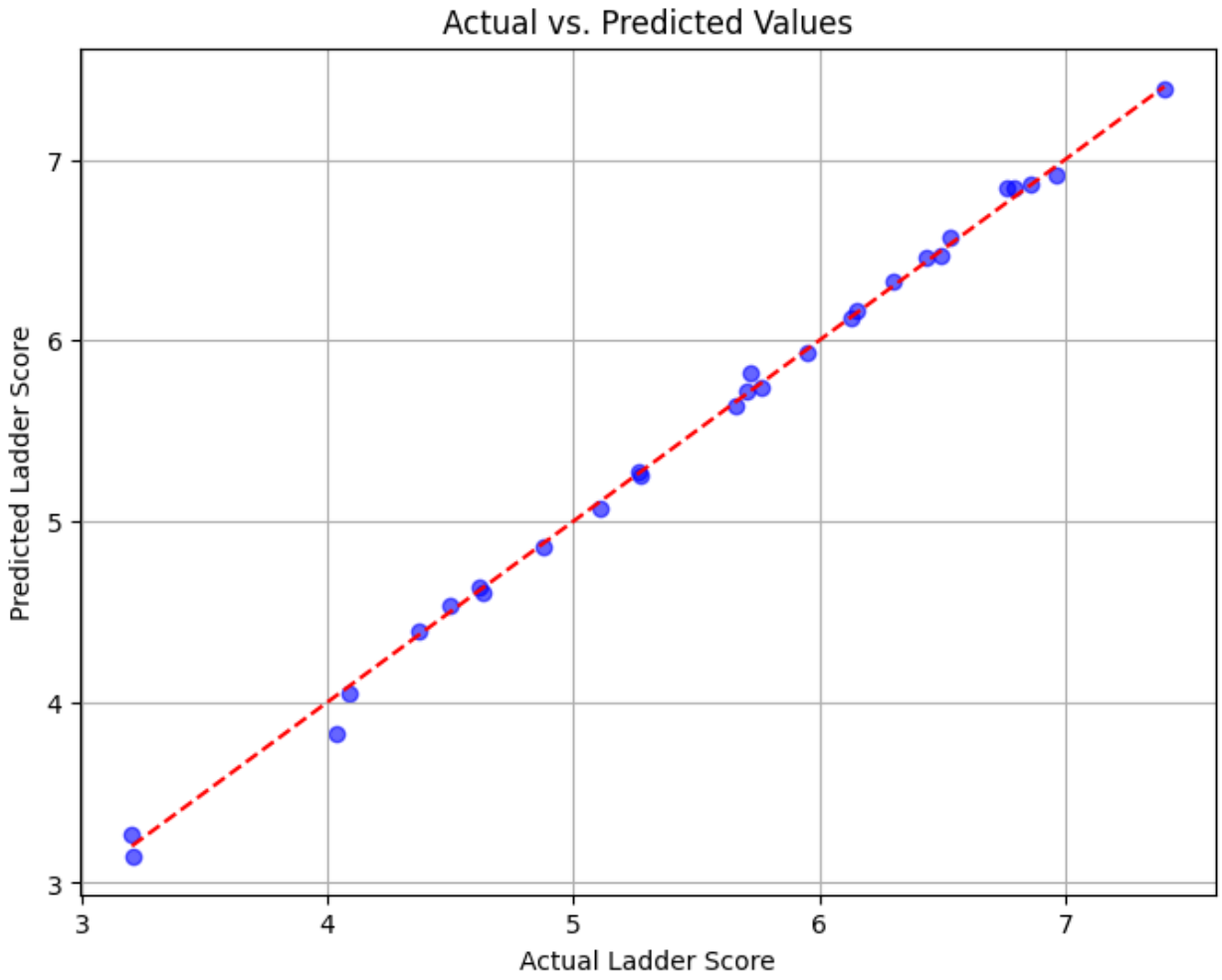
## Data cleaning and preprocessing

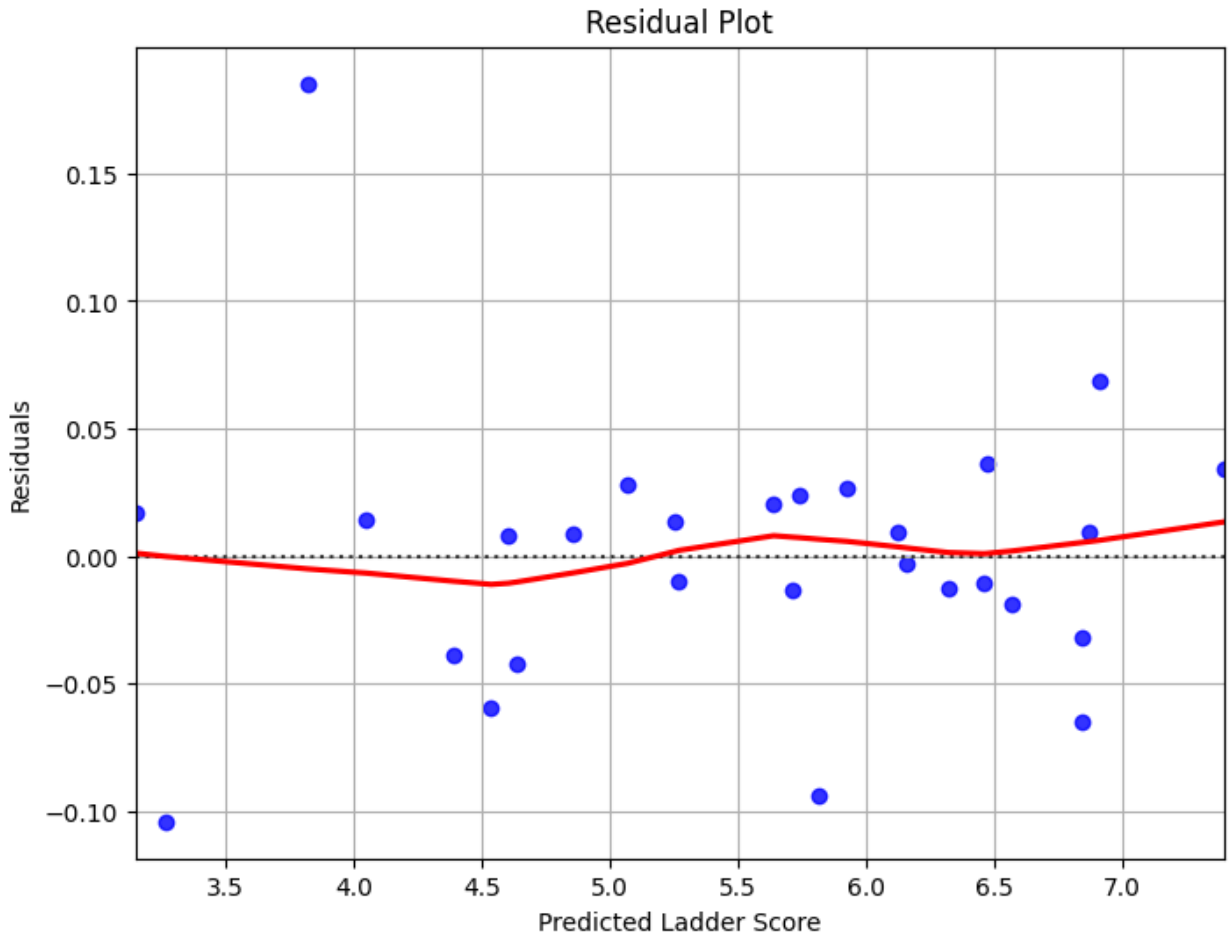
## Training and evaluation of three machine learning algorithms

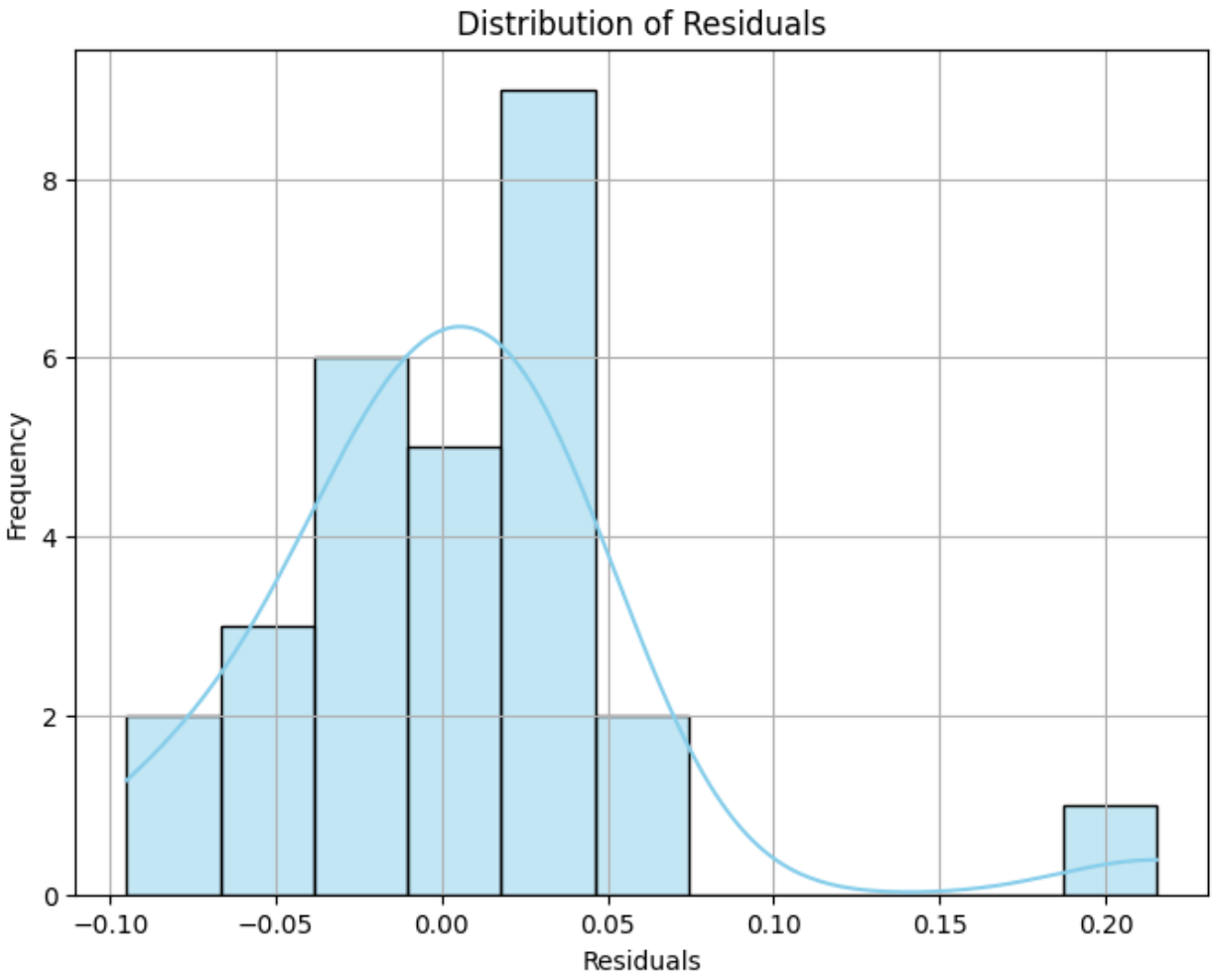
The 3 models where fairly close when it came to regression. However, linear regression ultimately won.

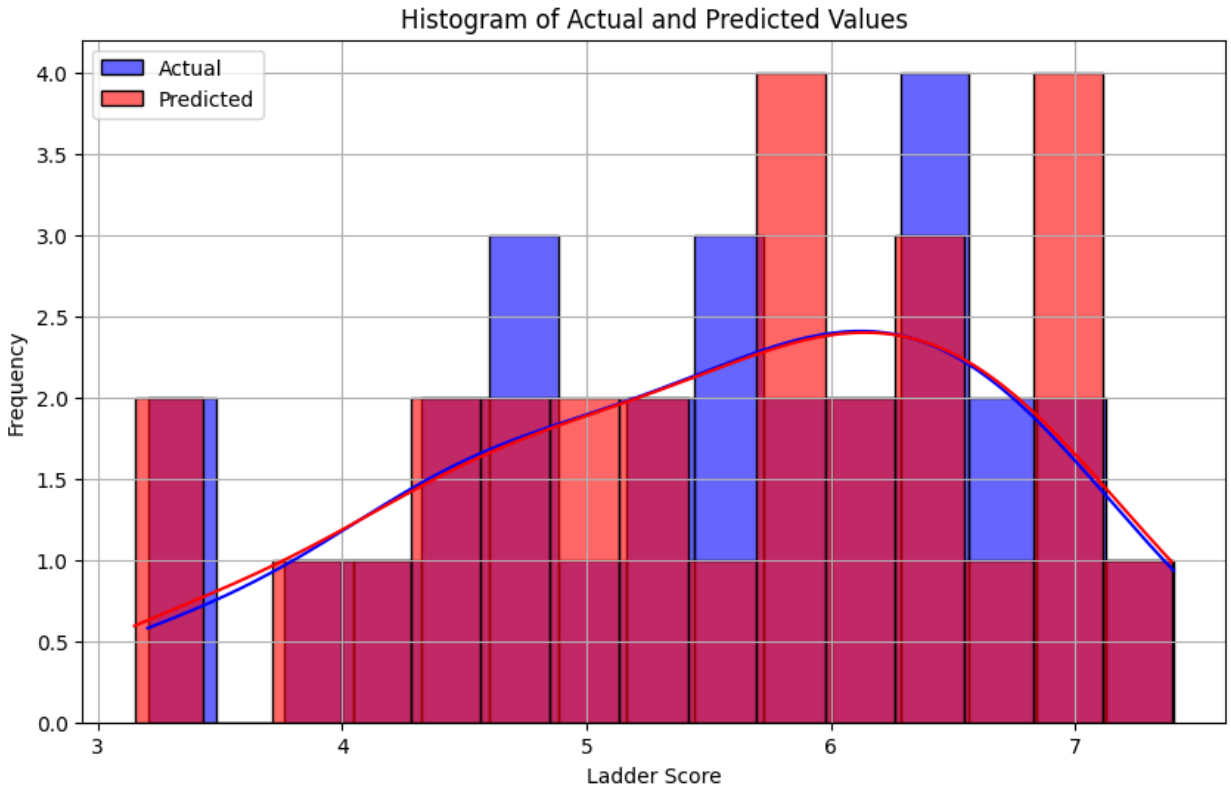
```
Performance Metrics:
                           Model  Mean Squared Error  R-squared
0              Linear Regression        1.461316e-07   1.000000
1        Random Forest Regression        2.356780e-03   0.998101
2  Gradient Boosting Regression        3.098134e-03   0.997504
```

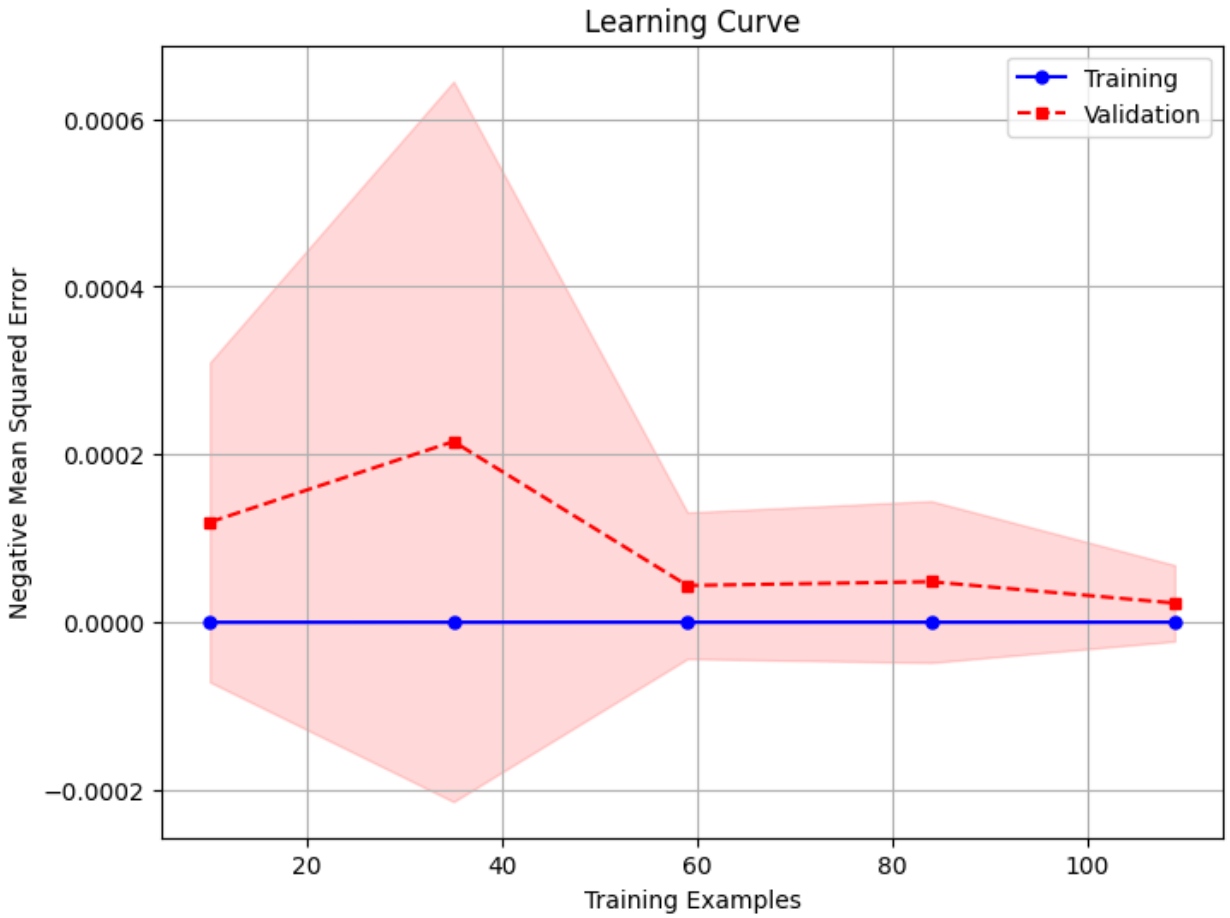
## 2+ graphs for the best performing algorithm




Residual Plot
Residuals
Predicted Ladder Score
0.15
0.10
0.05
0.00
−0.05
−0.10
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0


Distribution of Residuals
Frequency
Residuals
0
2
4
6
8
−0.10
−0.05
0.00
0.05
0.10
0.15
0.20


Histogram of Actual and Predicted Values
Actual
Predicted
Frequency
Ladder Score
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
3
4
5
6
7




## Limitations

These models are limited to taking only the 6 factors and cant take other factors such as crime rate.

## Next steps

Add various other factors by using combining datasets.

# Appendix 1

## Dataset link

https://www.kaggle.com/datasets/ajaypalsinghlo/world-happiness-report-2023?resource=download

## Github link

https://github.com/Iqbal-Talha/aiProject

# Appendix 2

Include your source code with proper comments and attribution to any code you have reused.