

# Understanding Greenhouse Gas Emissions through Rashomon Effect

Directed Reading Program Spring 2024

Presenter: Wenxin(Ariel) Xia

Mentor: Aparajithan Venkateswaran

Mentee: Wenxin(Ariel) Xia

# Contents

01

## Regression Models

introduction to linear regression and lasso model

02

## Datasets Selection

datasets merging process and meaning of variables

03

## Dataset in Models

the statistics generated from fitting the dataset into two models

04

## Models Comparison

compare and contrast the results generated from models

05

## Rashomon Effect

introduction to the Rashomon Effect and its effect in this research

06

## Future Work

further improvements that could be made in future

01

# Regression Models

# Linear Regression & Lasso

## Linear Regression Model

**Definition:** Linear regression is a statistical method that models the relationship between a dependent variable and one or more independent variables using a linear equation.

**Equation:**  $Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \dots + \beta_nX_n + \epsilon$

**Objective:** Minimize the sum of squared residuals (differences between observed and predicted values).

## Lasso Model

**Definition:** Lasso is a regression analysis method that performs both variable selection and regularization to enhance the prediction accuracy and interpretability of the statistical model.

**Equation:**  $Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \dots + \beta_nX_n + \epsilon$

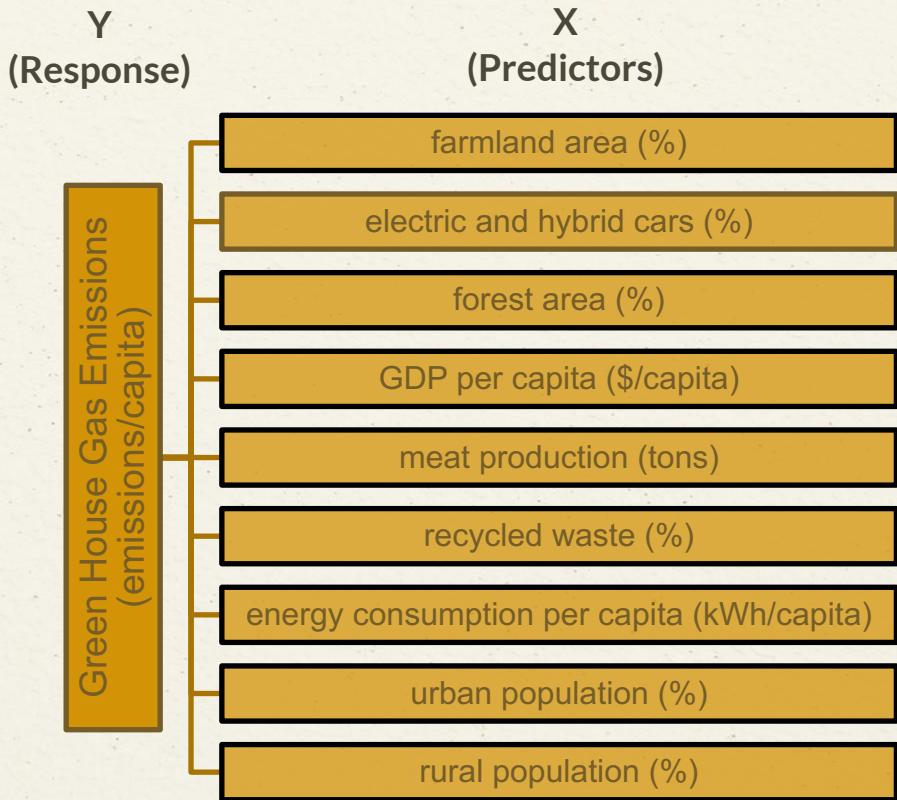
**Objective:** Minimize the sum of squared residuals plus the penalty term ( $\lambda \sum_{j=0}^n |\beta_j|$ ) to shrink coefficients of less important variables to zero.

02

## Datasets Selection

# Datasets

- 1) Agricultural land (% of land area). **World Bank**.
- 2) Share of car sales that are battery electric or plug-in hybrid (% of electric/hybrid cars). **Our World in Data**.
- 3) Forest area (% of land area). **World Bank**.
- 4) GDP per capita (current US\$). **World Bank**.
- 5) Meat production. **Our World in Data**.
- 6) Municipal waste recycling rate (% of recycled waste).  
**Our World in Data**.
- 7) Energy (kWh/capita). **Our World in Data**.
- 8) Urbanization (total population). **Our World in Data**.
- 9) CO2 and Greenhouse Gas Emissions (emissions/capita). **Our World in Data**.



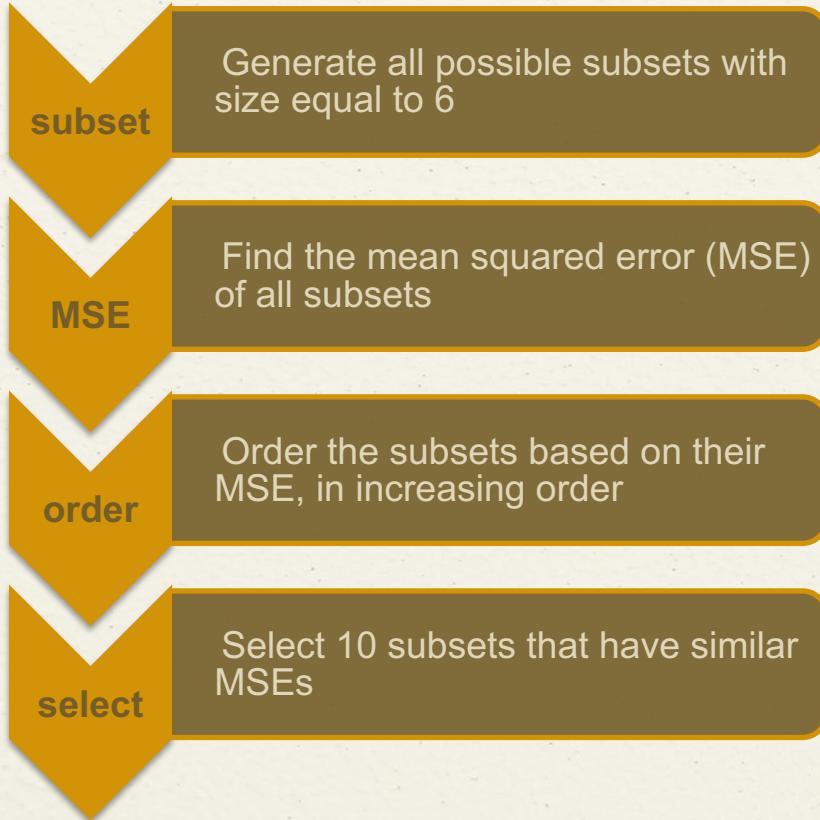
03

# Datasets In Models

# Lasso

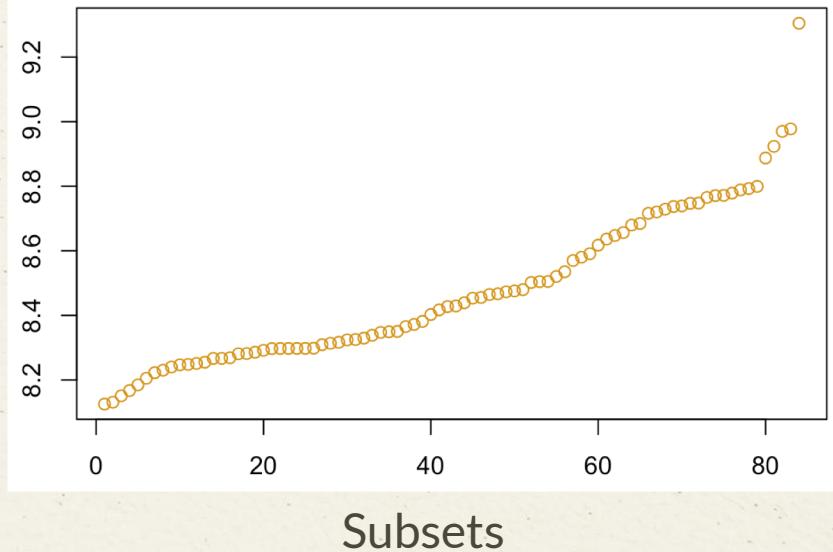
<b>predictor</b>	<b>Recycled waste (%)</b>	<b>Forest (%)</b>	<b>Farmland (%)</b>
Beta value	5.575313e-02	-1.161841e-02	5.507615e-05
<u>Selected Predictors:</u>			
1) Recycled waste	<b>Energy (kWh/capita)</b>	<b>GDP (\$/capita)</b>	<b>Meat (tons)</b>
2) Forest	7.756624e-06	-5.905040e-06	4.451071e-08
3) Farmland			
4) Energy/Capita	<b>E&amp;h car (%)</b>	<b>Urban (%)</b>	<b>Rural (%)</b>
5) GDP/Capita			
6) Meat Production	≈ 0	≈ 0	≈ 0

# Linear Regression



84 Subsets from 9 predictors

MSE Values of 84 Subsets



# Linear Regression – Subsets

Subset	Farmland (%)	E&h car (%)	Forest (%)	GDP (\$/c)	Meat (T)	Recycled waste (%)	Energy (kWh/c)	Urban (%)	Rural (%)	MSE
1		✓	✓	✓	✓	✓	✓			8.125
2			✓	✓	✓	✓	✓	✓		8.131
3			✓	✓	✓	✓	✓		✓	8.151
4	✓		✓	✓	✓	✓	✓			8.167
5			✓		✓	✓	✓	✓	✓	8.185
6		✓	✓		✓	✓	✓	✓		8.205
7		✓	✓		✓	✓	✓		✓	8.223
8	✓		✓		✓	✓	✓	✓		8.230
9	✓	✓	✓		✓	✓	✓			8.240
10				✓	✓	✓	✓	✓	✓	8.247

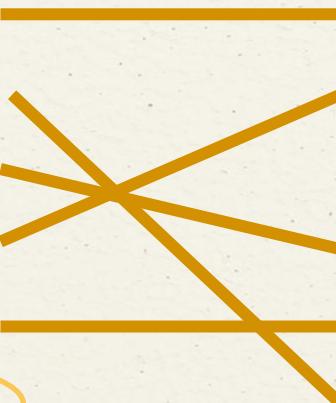
# 04

# Models Comparison

# Linear Regression vs. Lasso

## Linear Regression:

- 1) Recycled waste
- 2) Meat Production
- 3) Energy/Capita
- 4) Forest
- 5) GDP/Capita
- 6) Urban Population



## Lasso:

- 1) Recycled waste
- 2) Forest
- 3) Farmland
- 4) Energy/Capita
- 5) GDP/Capita
- 6) Meat Production



Rashomon  
Effect

05

# Rashomon Effect

# What is Rashomon Effect?

- Definition: the phenomenon where different people have contradictory accounts of the same event.
- Origin: film Rashomon (1950) directed by Akira Kurosawa. In the film, multiple characters provide conflicting versions of a crime they witnessed.



# Our Research & Rashomon Effect

- Linear regression vs. Lasso
- Implication: importance of model selection
- Challenges: unable to explain everything using one model; choosing model

06

# Future Work

# Future Work

- **Model Fitting**
  - Current: some predictors in linear regression model do not have clear ranking
  - Future: find numerical values that could put predictors in sequence
- **Visualization**
  - Current: visualization in console
  - Future: create a Shiny app which everyone could use to visualize the data of a range of years and selected countries
- **Dataset Processing**
  - Current: convert “NA” values to 0
  - Future: convert “NA” values to the value of same country and the closest year

# Citations

Posteritati. (1951). \*Rashomon (Original) [Polish A1]\* [Movie poster]. Retrieved from [https://posteritati.com/poster/23320/rashomon-original-1951-polish-a1-movie-poster](https://posteritati.com/poster/23320/rashomon-original-1951-polish-a1-movie-poster)

RStudio Team. (2022). \*RStudio: Integrated Development for R\*. RStudio, PBC, Boston, MA. Retrieved from [http://www.rstudio.com/](http://www.rstudio.com/)

Datasets:

Our World in Data. (2024). \*Our World in Data\*. Retrieved from [https://ourworldindata.org](https://ourworldindata.org)

The World Bank. (2024). \*The World Bank\*. Retrieved from [https://www.worldbank.org](https://www.worldbank.org)

# Thanks for Watching!

Presenter: Wenxin(Ariel) Xia

Mentor: Aparajithan Venkateswaran

Mentee: Wenxin(Ariel) Xia

CREDITS: This presentation template was created by [Slidesgo](#), and includes icons by [Flaticon](#), and infographics & images by [Freepik](#)