

The background is a light gray with various colorful abstract shapes and patterns. There are pink and red stars, a yellow zigzag line, a blue wavy line, a green oval pattern, and a red zigzag line. The title 'The Joy Portfolio' is written in a bold, rounded, pink font. The names 'Amelie, Arian, Rithvik, Sid' are written in a smaller, pink font below the title.

# The Joy Portfolio

Amelie, Arian, Rithvik, Sid

The background is a light gray with several abstract, colorful shapes. There is a pink shape in the top left, a large light orange shape in the top right, a yellow shape in the bottom left, a green shape in the bottom left, and a dark blue shape in the bottom right. A red star is located in the top right corner. A white jagged line is on the right side.

**What factors affect  
a country's  
happiness?**

# Our Idea



Factors and expenditures most  
contributive to a country's happiness

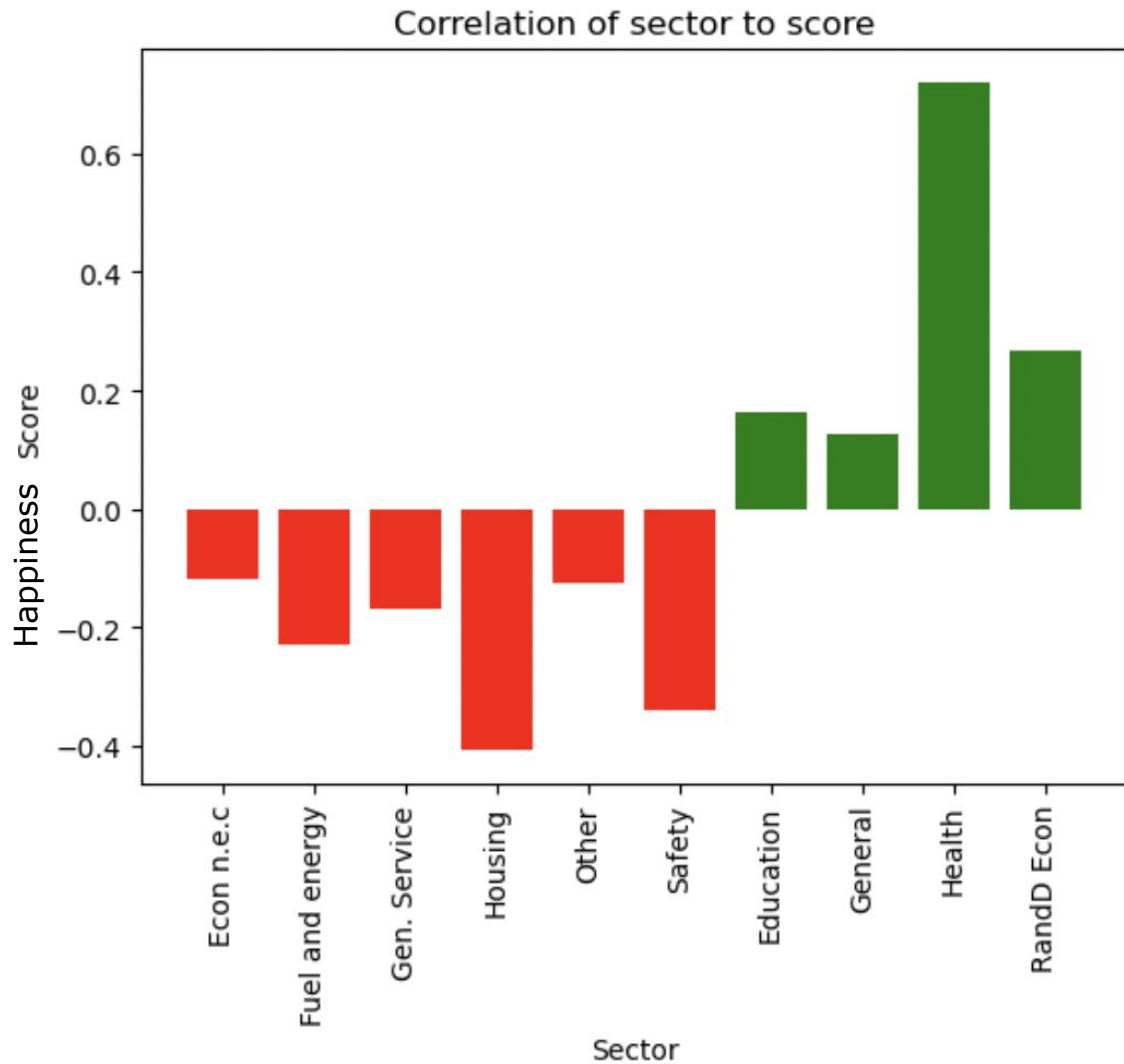
# Exploratory Data Analysis



Country	Econ n.e.c	Education	Energy	General Public Services	Health	Housing	Public order	Happiness Score
Canada	1	13	1	15	23	1	5	7
Paraguay	0	17	0	8	17	2	13	6

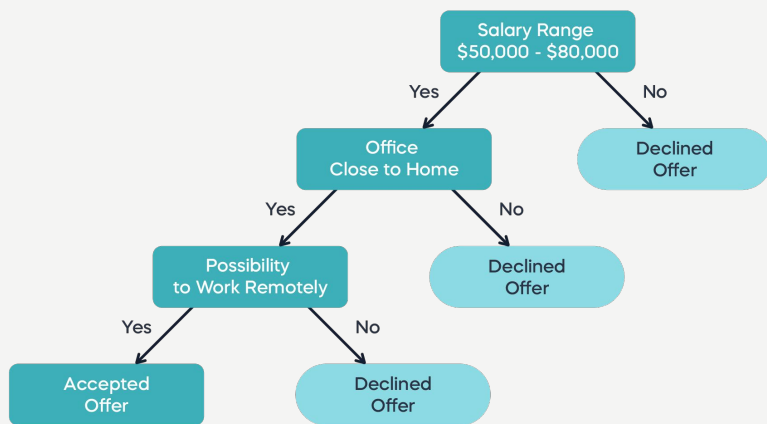


# Exploratory Data Analysis



# Why Regression?

- Sort and visualize complicated messy data
- Possible regressors (XGBoost, Decision Tree, Lasso)
- Accuracy vs Analysis



Decision tree

The diagram shows the linear regression equation  $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$  with labels for each term.  $Y_i$  is labeled "Dependent Variable".  $\beta_0$  is labeled "Population Y intercept".  $\beta_1$  is labeled "Population Slope Coefficient".  $X_i$  is labeled "Independent Variable".  $\epsilon_i$  is labeled "Random Error term". A bracket under  $\beta_0 + \beta_1 X_i$  is labeled "Linear component". A bracket under  $\epsilon_i$  is labeled "Random Error component".

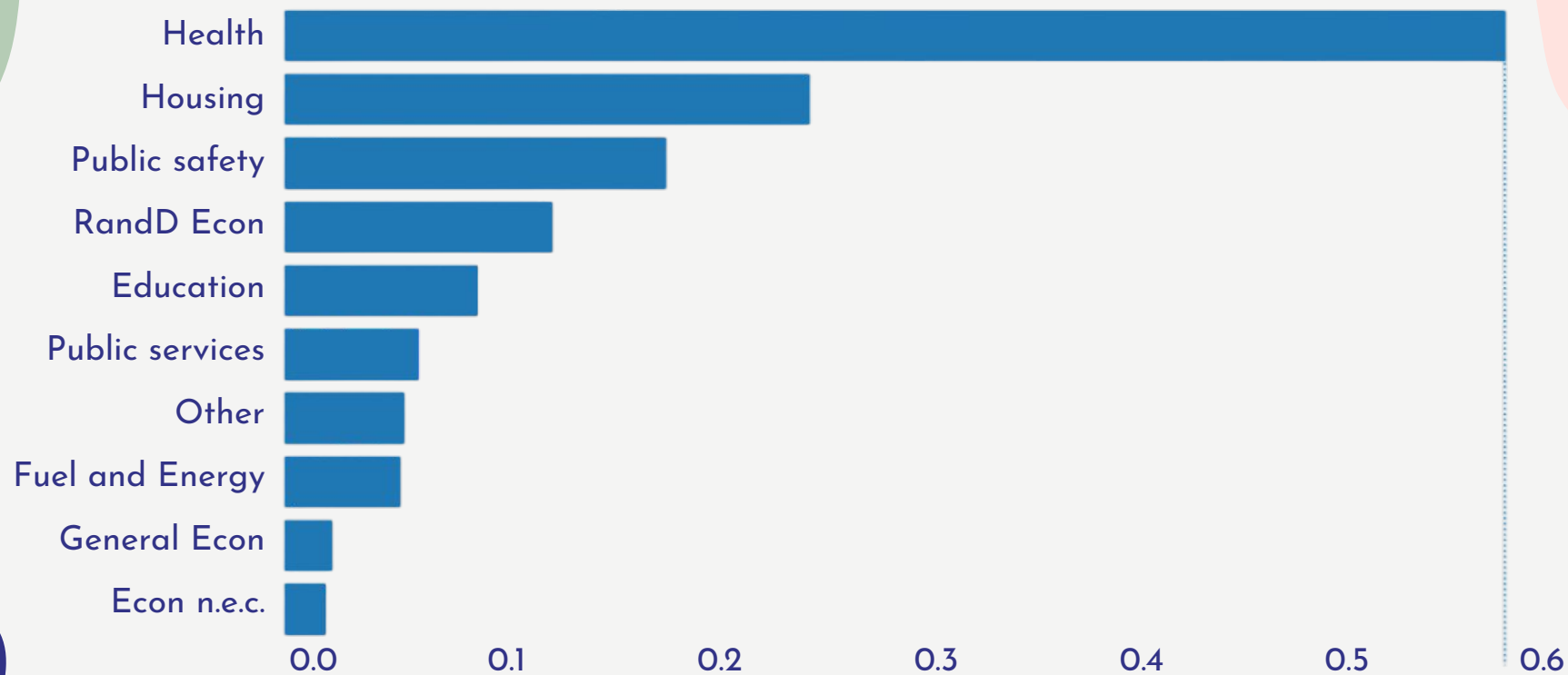
$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$



Linear component

Random Error component

Linear Regression (Lasso)

## Sectors of expenditure that affect happiness



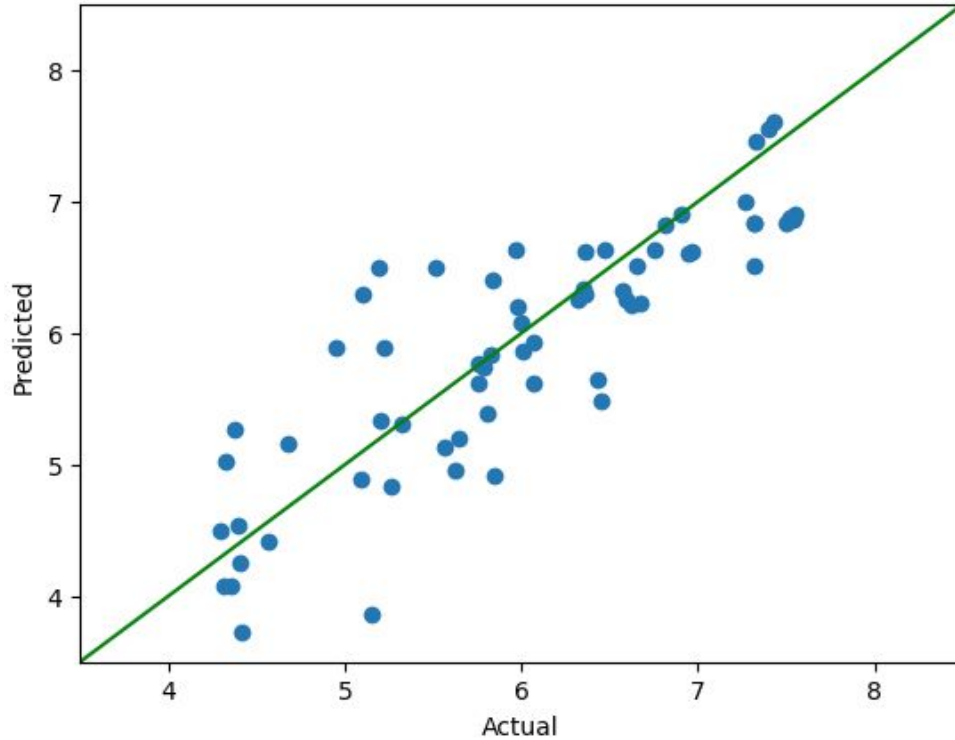

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \varepsilon$$

RandD Economic Affairs	+0.37
Health	+0.25
General Economic Commercial and Labour Affairs	+0.11
Education	+0.10
General Public Services	+0.07
Fuel and Energy	+0.05
Other Industries	+0.03
Public Order and Safety	+0.02
Housing and Community Amenities	-0.05





# Model Evaluation



$$R = 0.862$$

$$R^2 = 0.742$$

# Why?



## Policy Evaluation

Governments can evaluate policies to allocate resources effectively, and understand what to prioritize.



## Economic Planning

Governments can plan for sustainable economic development, that's aligned with factors that increase happiness.



## Happiness Benefits

Happy societies are more cohesive and less prone to unrest, including crime. They host motivated citizens who are physically and mentally healthier.



What  
can we  
do about  
it?



**Thank You!**