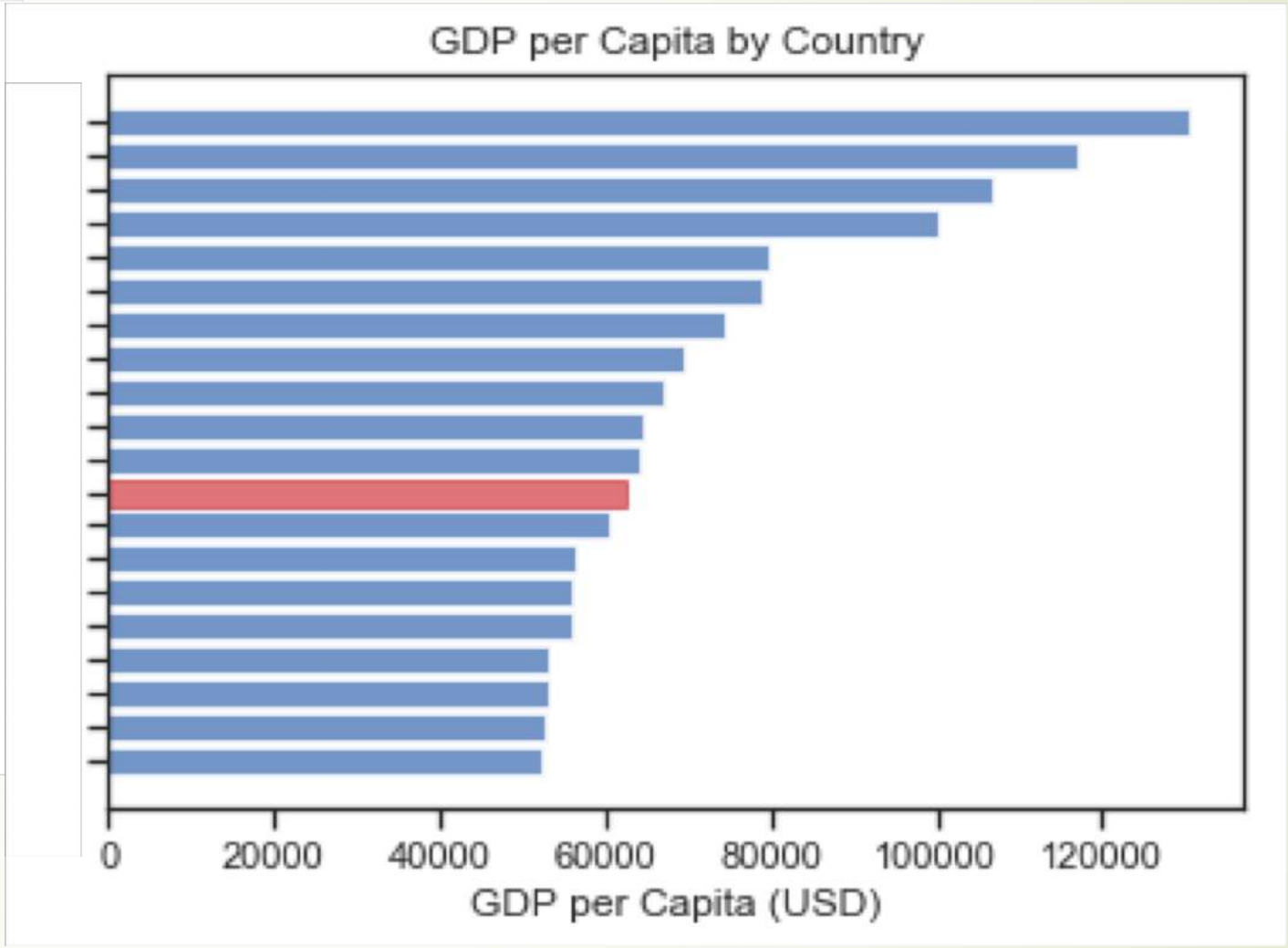
Modelling Real GDP per Capita

Data Exploration and Regression Analysis

GDP per Capita by Country/Territory

- 1. Qatar: $ 130,475
- 2. Macau : $ 116,808
- 3. Luxembourg : $ 106,704
- 4. Singapore : $ 100,345
- 5. Brunei : $ 79,530
- 6. Ireland : $ 78,785
- 7. Norway : $ 74,356
- 8. U.A.E.: $ 69,382
- 9. Kuwait : $ 67,000
- 10. Switzerland : $ 64,649
- .....
- 12. United States: $ 62,606



What factors predict gross domestic product (GDP) per capita at a country-level?

- Many factors to consider:
  - Global Economics
  - Government Policy
  - Public Health
  - Education
  - Natural Resources

Filtering Out the Noise

- How do we make sense of all the data?
- Generalize our phenomenon with a linear regression model
- Select for features with meaningful relationships
  - Linear Regression
  - Lasso (Least Absolute Shrinkage and Selection Operator) Regression
  - Ridge Regression
- Tune our parameters with feature engineering to optimize our model

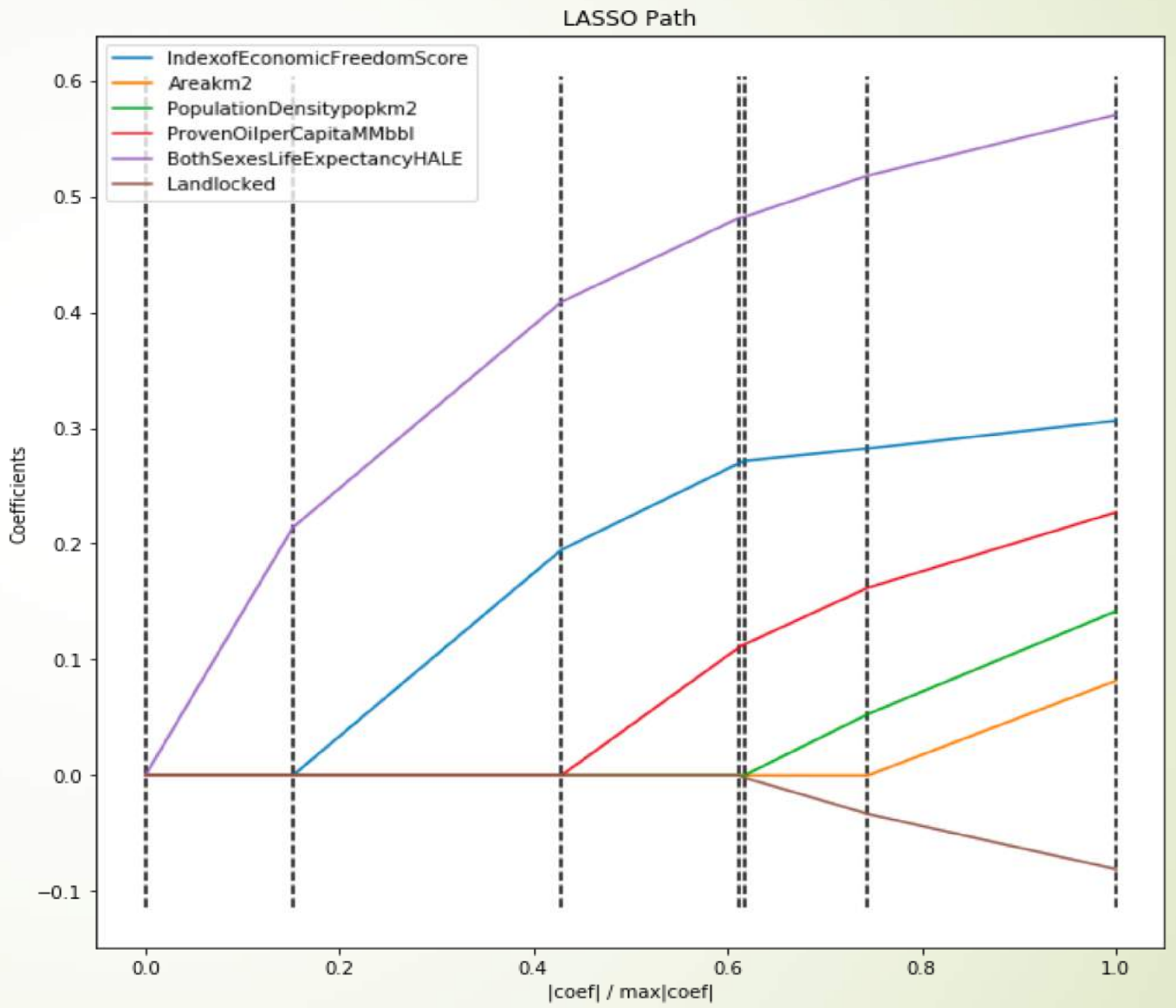
Optimizing Model Choices

- Train, Validate & Test our model:
  - This framework gives us an empirical way to select features and avoid over/under-fitting
- Train:
  - Linear Regression, Lasso Regression, Ridge Regression
- Validate:
  - Score each regression based on how well each model generalizes validation set
- Test:
  - Retrain best candidate on train & validation set and score final model on test set

Lasso Regularization

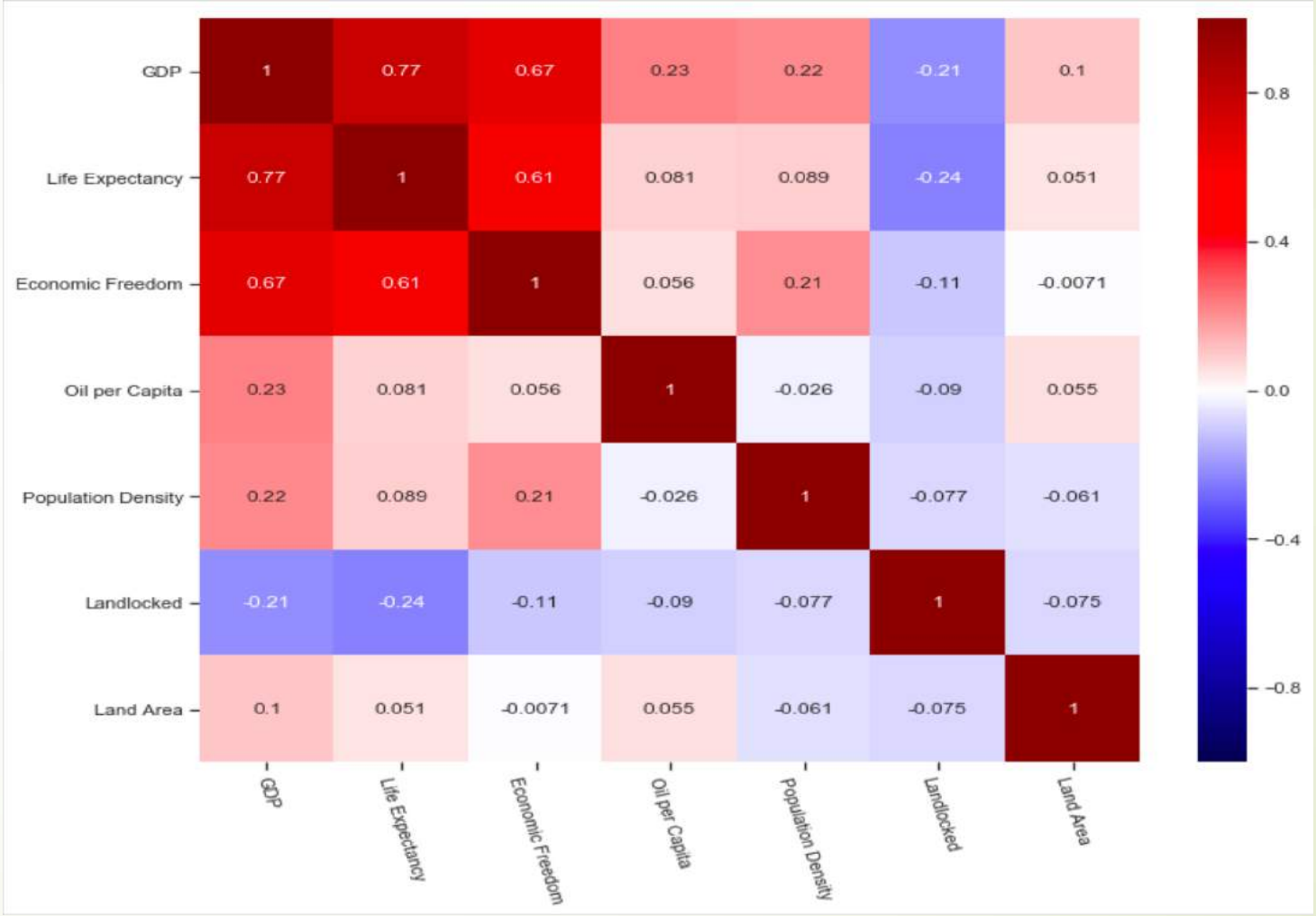
Linear Equation:

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 \ldots.$$



Correlation Heatmap

- Lasso Regression
- GDP Per Capita ($\hat{y}$):
  - Health Adjusted Life Expectancy (HALE)
  - > Economic Freedom
  - > Oil Reserves per Capita
  - > Population Density
  - > Coastline Nations
  - > Land Area



Key Takeaways

- Lasso Regression: $r^2 = 0.68$
  - Statistical measure of how close the data are to the fitted regression line
- Health adjusted life expectancy (HALE)
- Economic Freedom
  - Oil reserves per capita
  - > Population Density
  - > Coastline Nations
  - > Land Area