


LIFE EXPECTANCY INSIGHTS

Data Science Solutions

Dmitri Gourianov


AGENDA

- Overview
 - Methodology
 - Data
 - Analysis
 - Conclusion and Recommendations
 - Future work
- 
- Several white lines of varying lengths and slopes are positioned in the bottom right corner of the slide, creating a modern, abstract graphic element.

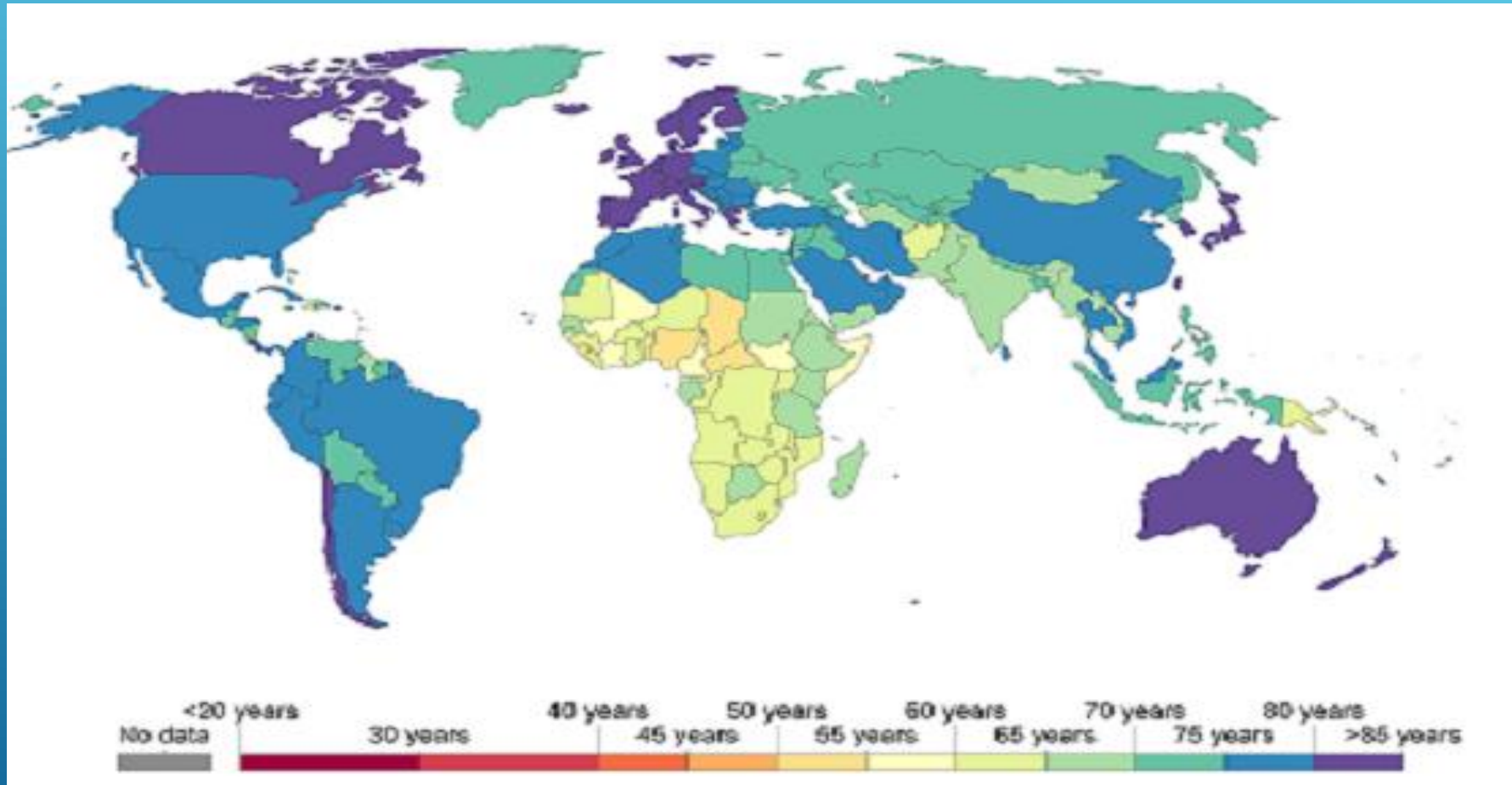
OVERVIEW

Issue: A decline in life expectancy across certain countries

Objective: To identify key determinants that can be addressed to increase life expectancy

Several white lines of varying lengths and angles are positioned in the bottom right corner of the slide, creating a modern, abstract graphic element.

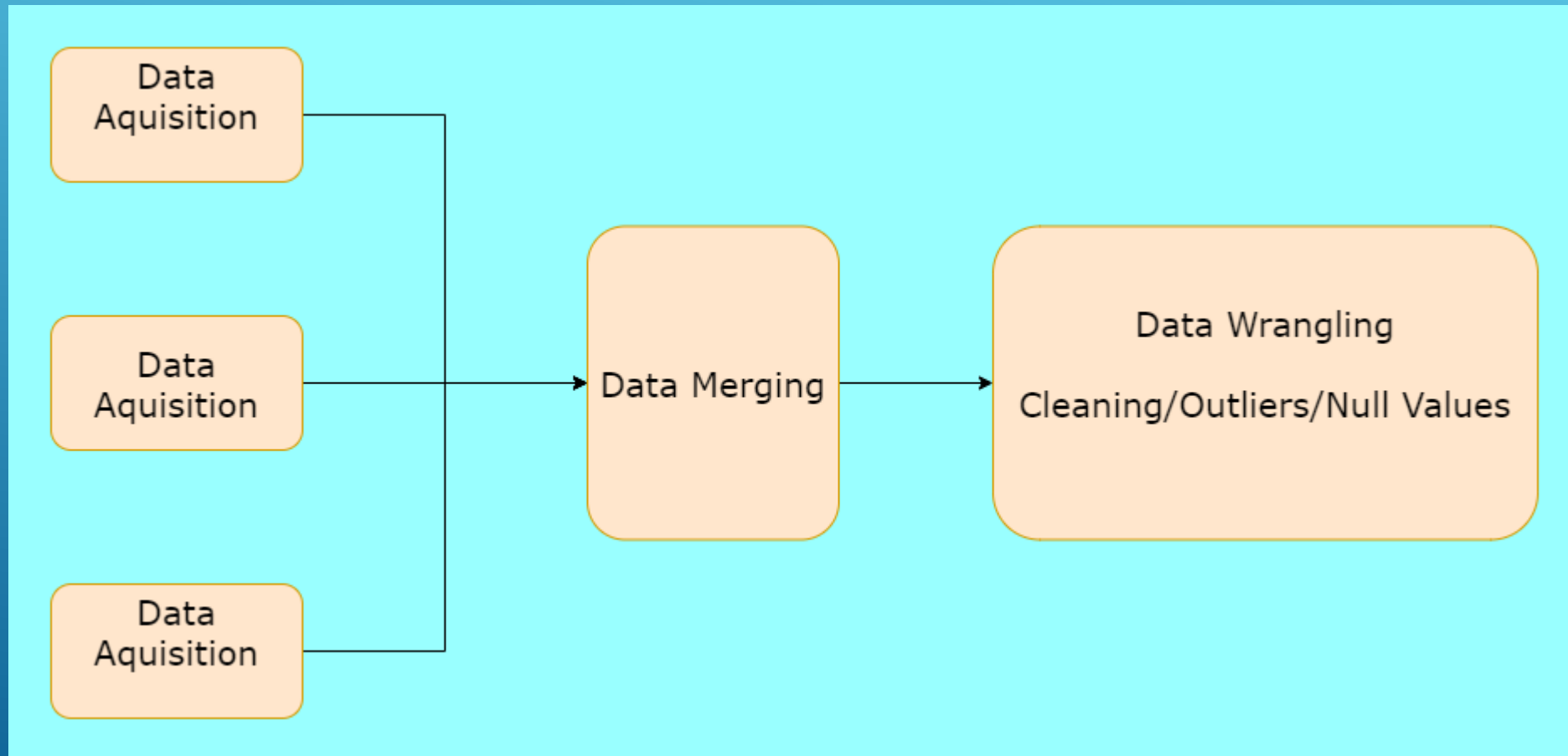
LIFE EXPECTANCY IN THE WORLD



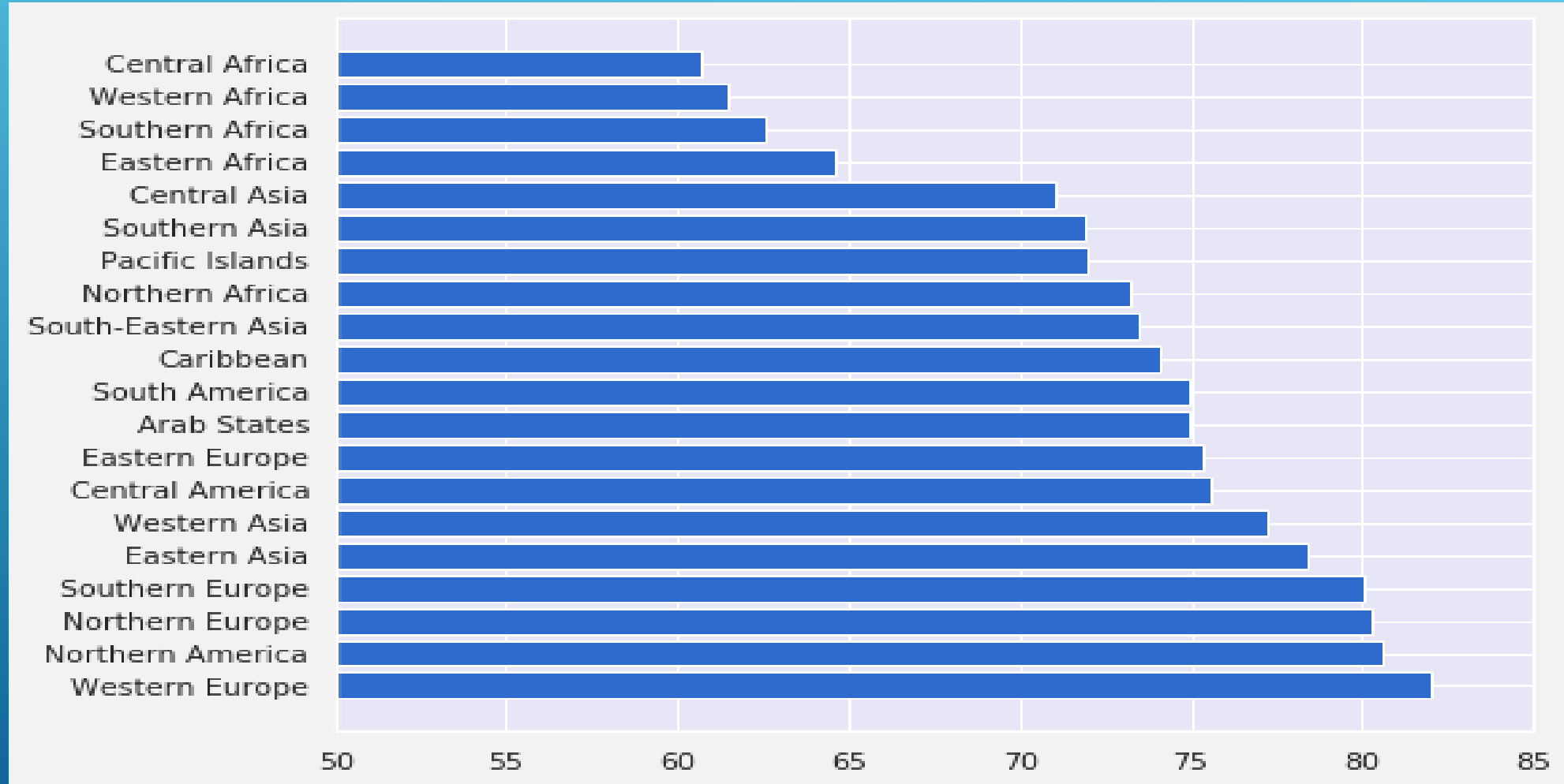
INDEPENDENT VARIABLES

- Geographical Subregion
 - GDP Per Capita
 - Health Expenditures Per Capita
 - Infant Mortality
 - Alcohol Consumption
 - Obesity
 - Mental Health Problems Index
 - Democracy Index
 - Peace Index
 - Environmental Index (EPI)
 - Smoking Rate
 - Importance of Religion Index
 - Road Fatalities
- 
- A series of three parallel white diagonal lines are positioned in the bottom right corner of the slide, extending from the bottom edge towards the right edge.

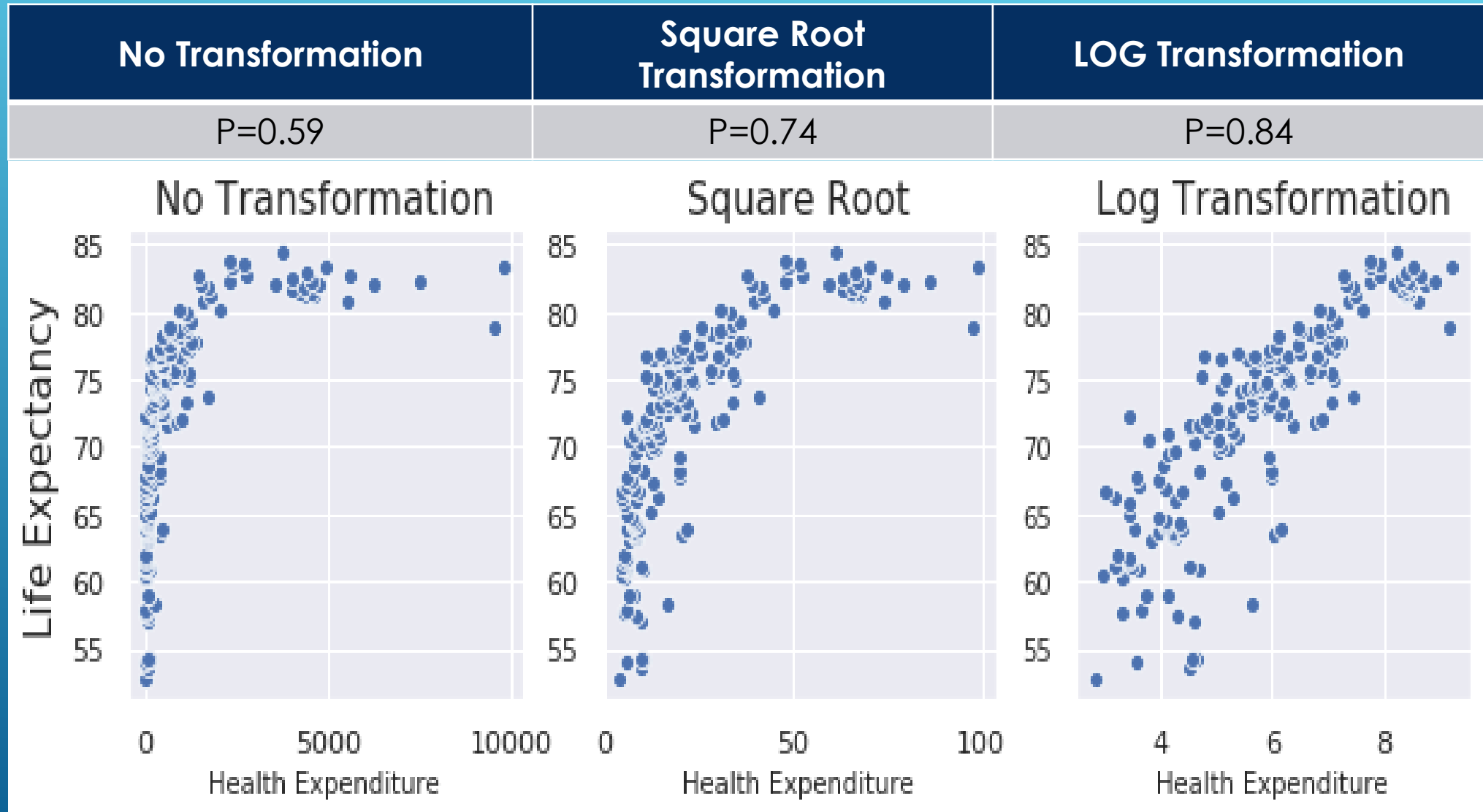
DATA PIPELINE



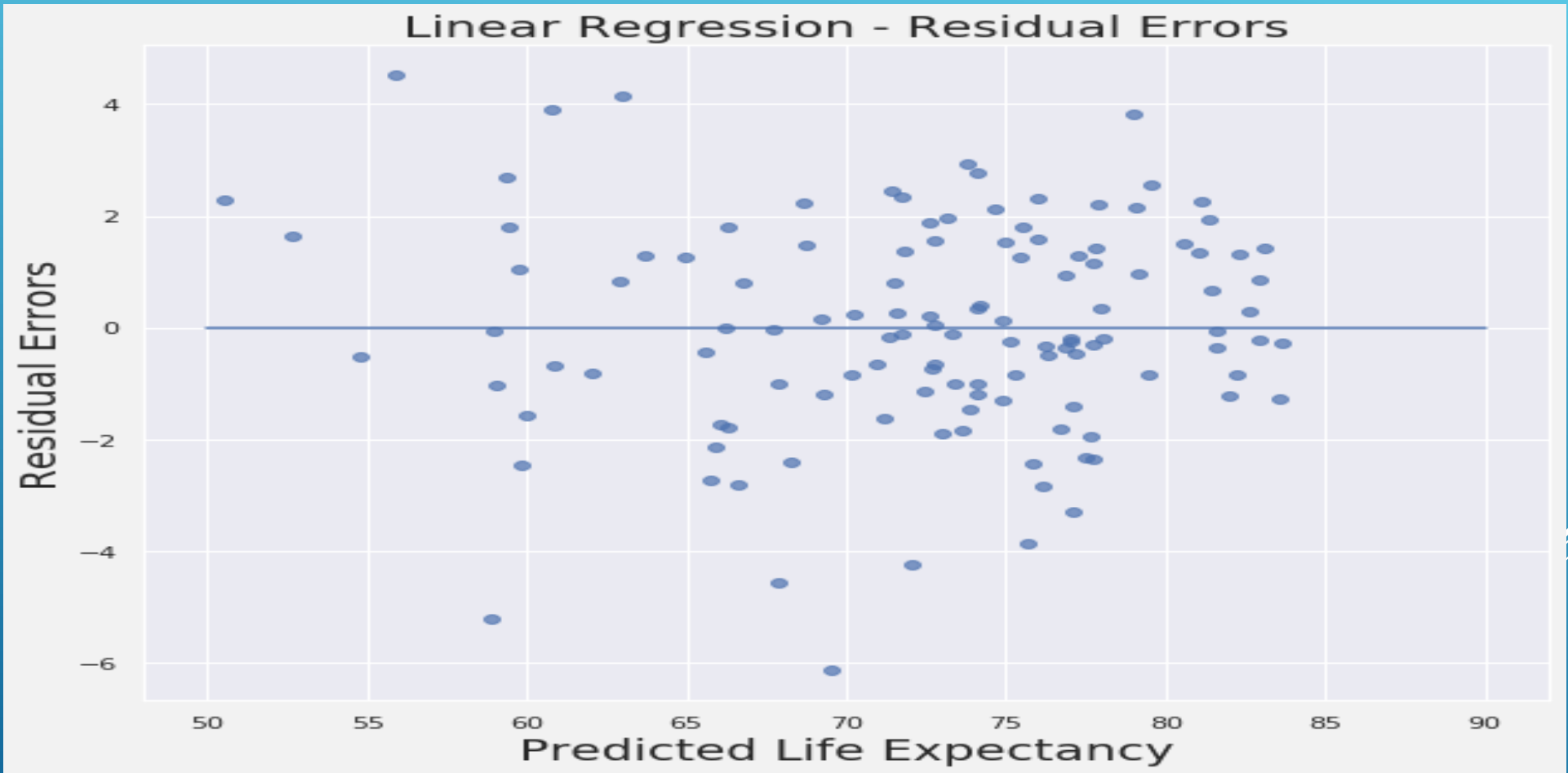
LIFE EXPECTANCY / SUB-REGIONS



FEATURE ENGINEERING



RESIDUAL ERRORS



CROSS-VALIDATION

MODELS	R2 SCORE
Linear Regression	0.911
Polynomial Regression	0.66
Ridge Regression	0.913
Lasso Regression	0.914
Random Forest	0.911

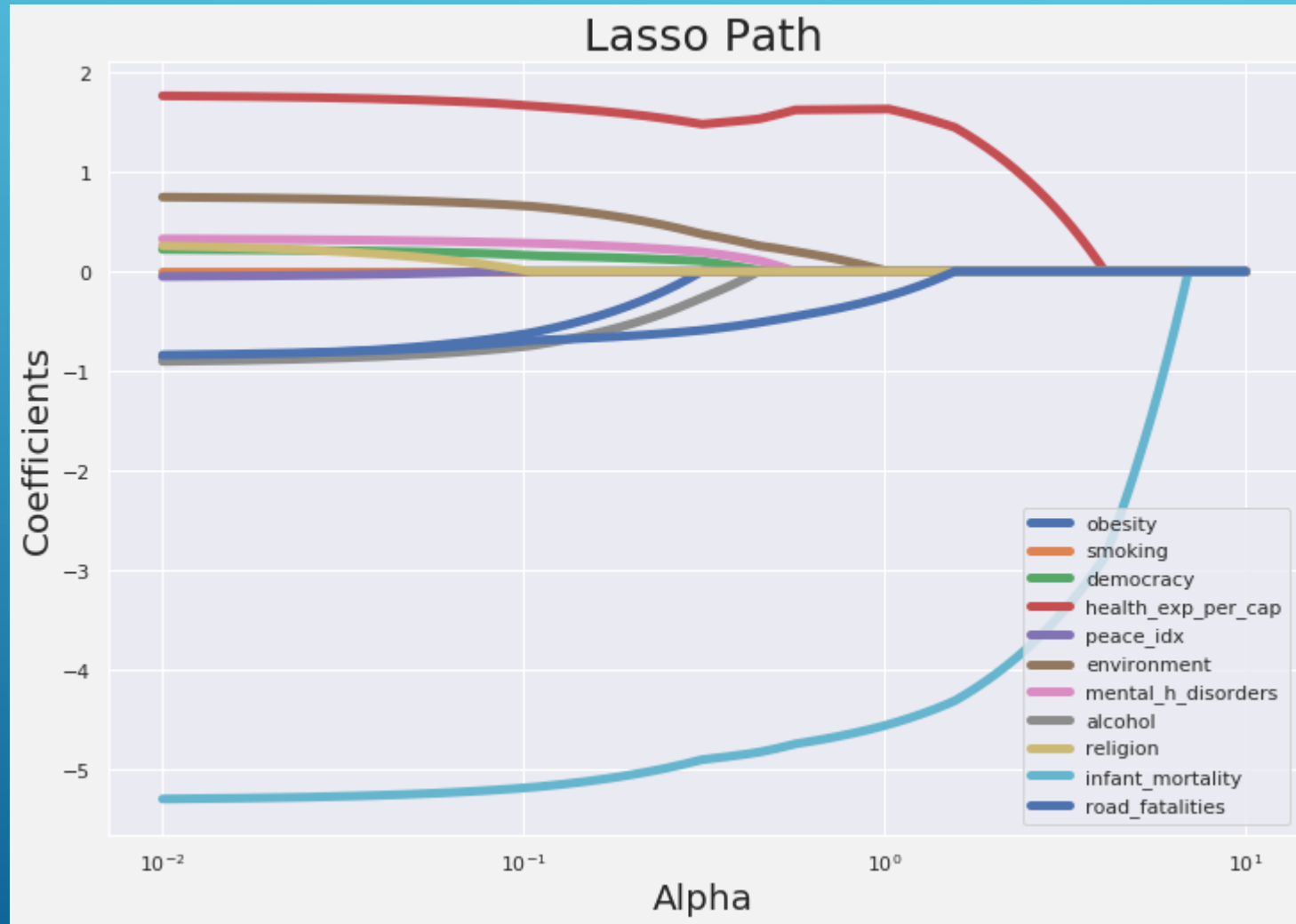


Test Score:
R2=0.90

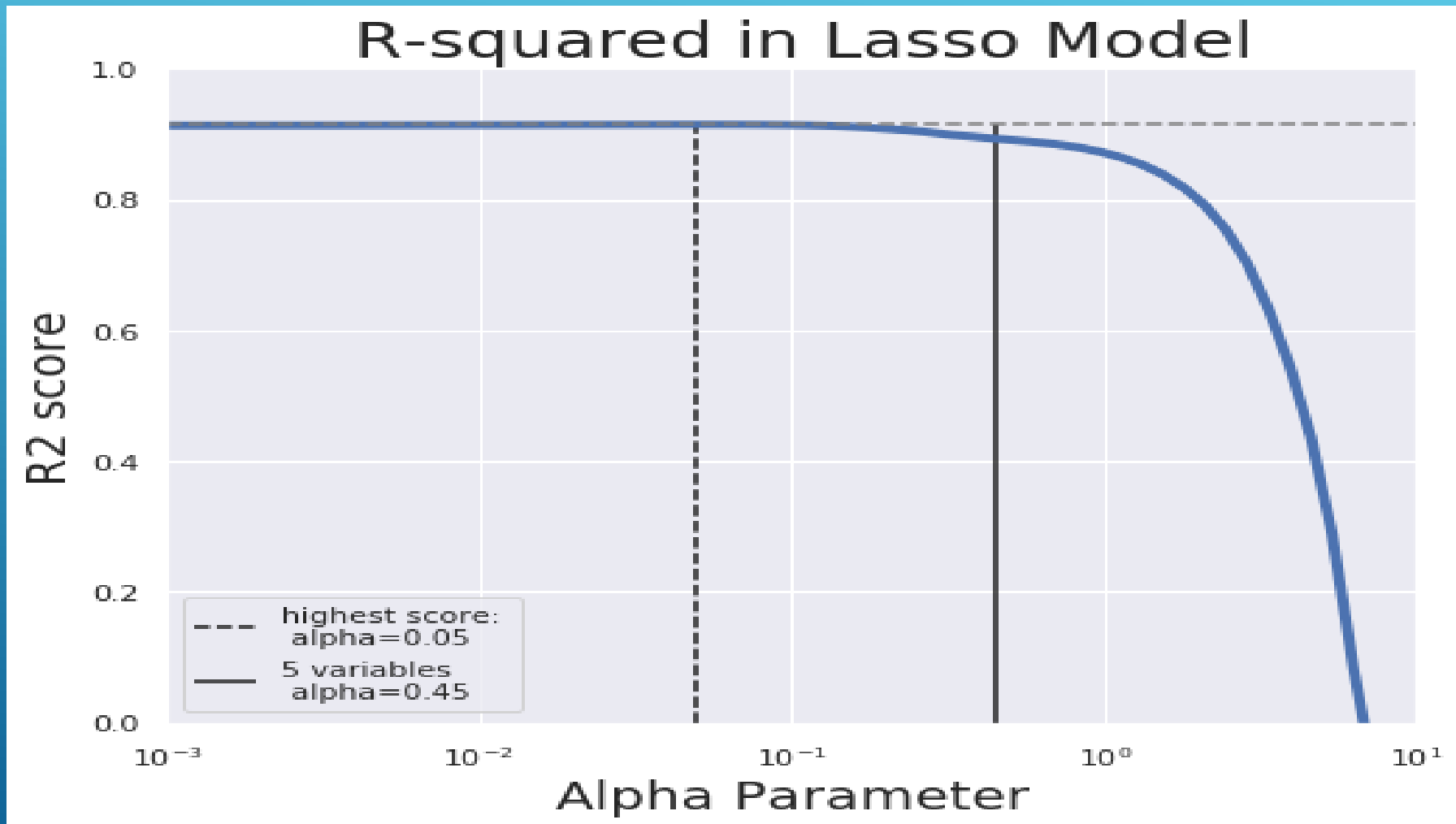
Feature Importance

Features	Coefficients
Infant Mortality	(4.83)
Health Expenditures Per Capita	1.68
Road Fatalities	(0.93)
Alcohol Consumption	(0.86)
Environmental Index	0.83
Obesity	(0.70)
Religion	0.31
Mental Health Disorders Rate	0.31
Democracy Index	0.3
Peace Index	(0.12)
Smoking Rate	0.11

LASSO REGULARIZATION



EFFECT OF ALPHA ON R2 SCORE



MOST IMPORTANT FEATURES


Lasso Feature Selection*	Random Forest Feature Importance**
Infant Mortality	Infant Mortality
Health Expenditure Per Capita	Health Expenditure Per Capita
Road Fatalities	Environmental Index
Environmental Index	Road Fatalities
Mental Health Disorders Rate	Democracy Index

*Red font color in Lasso results indicates negative effect on life expectancy

**Random Forest provides only absolute value of importance

CONCLUSIONS AND RECOMMENDATIONS

Most important features were identified and must be addressed to increase life expectancy:

- Implement Measures to Reduce Infant Mortality
 - Increase Health Expenditure per Capita
 - Implement Measures to Reduce Road Fatalities
 - Establish strict environmental regulation
- 
- A series of three parallel white diagonal lines extending from the bottom right towards the top right of the slide.

FUTURE WORK

- Analyze historical trends
 - Add new features
 - Analyze life expectancy by gender and race/ethnicity
- 
- A series of white lines of varying lengths and orientations are positioned in the bottom right corner of the slide, creating a modern, abstract graphic element.

APPENDIX



SOURCES

- ▶ Ourworldindata.com
 - ▶ Wikipedia.com
 - ▶ World health organization
- 
- Several white lines of varying lengths and slopes are positioned in the bottom right corner of the slide, creating a modern, abstract graphic element.

LEADERS AND LAGGARDS

HIGHEST LIFE EXPECTANCY		LOWEST LIFE EXPECTANCY	
COUNTRY	AGE	COUNTRY	AGE
Japan	84.5	C.A.R.	52.8
Singapore	83.8	Lesotho	53.7
Italy	83.6	Chad	54.0