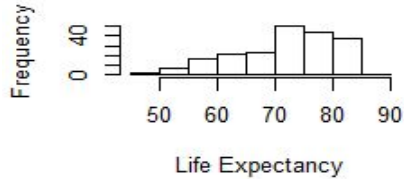


Life Expectancy Analysis and Prediction

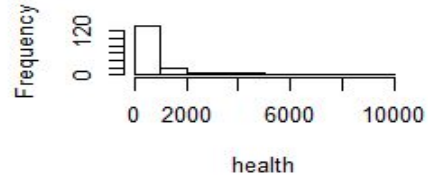
Xinyue (Lydia) Li, Yuqun (Azura) Tang, Xuehan Zhao

Introduction

Histogram of Life Expectancy



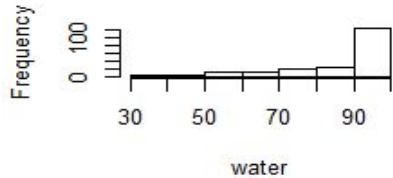
Histogram of health



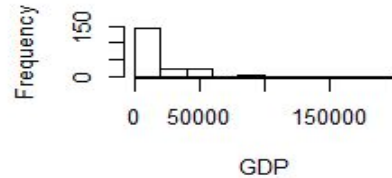
201 observations

- ❖ Dependent Variable
 - Life Expectancy
- ❖ Independent Variable
 - Health
 - Water
 - GDP
 - Measles
 - P.M. 2.5
 - Food
 - Region

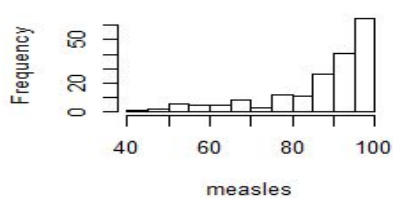
Histogram of water



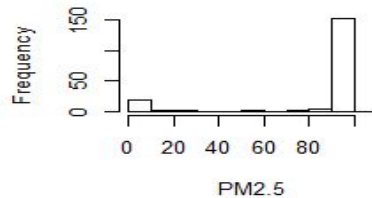
GDP



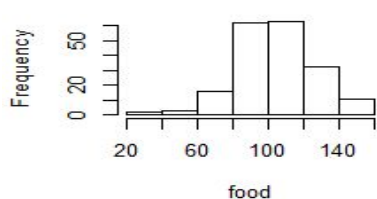
Histogram of measles



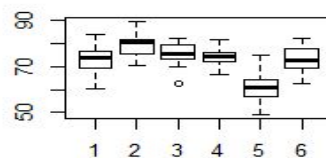
Histogram of PM2.5



Histogram of food



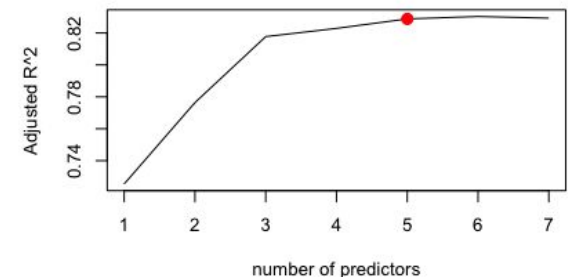
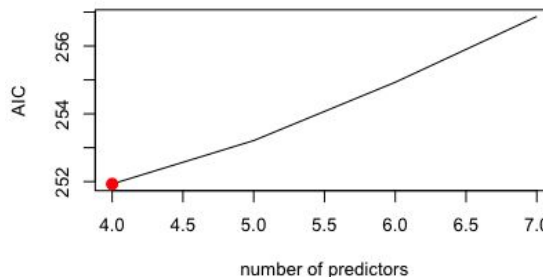
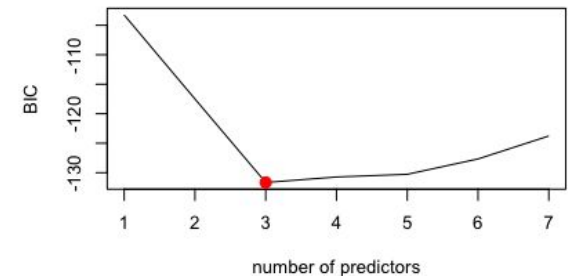
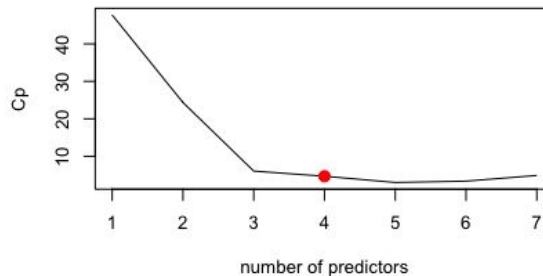
Region



	Health	Water	GDP	P.M 2.5	Food
Health	1.00	0.81	0.95	-0.31	-0.16
Water	0.81	1.00	0.78	-0.21	-0.15
GDP	0.95	0.78	1.00	-0.27	-0.19
P.M 2.5	-0.31	-0.21	-0.27	1.00	0.09
Food	-0.16	-0.15	-0.19	0.09	1.00

Statistical Analysis

- Multicollinearity:
 - VIF values > 5:
log(health), log(GDP)
- Model selection:
 - AIC, BIC, Cp, adjusted R-square
 - health, water, measles and region
- Check VIF again:
 - $VIF < 5$



Statistical Analysis

- Check interaction terms:
 - Full model VS. reduced model
 - ANOVA F test: $p\text{-value} = 0.8388 > 0.05$

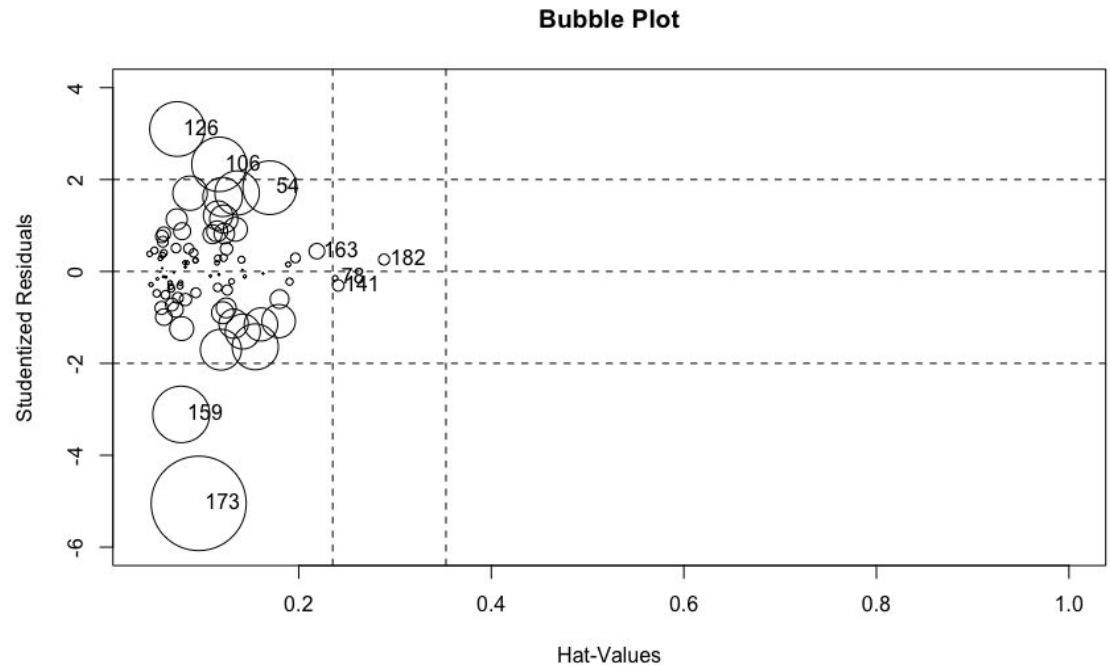
$$\begin{aligned} \text{Life Expectancy} = & \beta_0 + \beta_1 \times \log(\text{health}) + \beta_2 \times \frac{\text{water}^9 - 1}{9} + \gamma_1 M_1 + \gamma_2 M_2 + \delta_1 R_1 \\ & + \delta_2 R_2 + \delta_3 R_3 + \delta_4 R_4 + \delta_5 R_5 + \epsilon \end{aligned}$$

- Fit regression:
 - intercept
 - $\log(\text{health})$
 - water
 - Measles (75,100]
 - Africa

$$\begin{aligned} \text{Life Expectancy} = & 49.37 + 2.19 \log(\text{health}) + 6.22 * 10^{-17} \frac{\text{water}^9 - 1}{9} \\ & + \begin{cases} 0, & \text{measles} \in (0, 50] \\ 5.84, & \text{measles} \in (50, 75] \\ 7.61, & \text{measles} \in (75, 100] \end{cases} + \begin{cases} 0, & \text{Asia} \\ -1.51, & \text{Europe} \\ -1.13, & \text{North America} \\ -2.70, & \text{South America} \\ -6.02, & \text{Africa} \\ -1.60, & \text{Oceania} \end{cases} \end{aligned}$$

Statistical Analysis

- MSE: 14.03
- Prediction:
 - UK: 80.52 (81.1)
 - Thailand: 74.53 (74.4)
 - Ethiopia: 56.42 (64)
 - good job for predicting developed & developing countries
- Diagnostics:
 - Bubble plot
 - Obs. 173 (Swaziland)
 - Obs. 106 (Libya)



Conclusions & Discussion

- health expenditure per capita, % of population with access to improved water source, child immunization against measles, and the continent the country belongs to clearly have some impacts on the life expectancy at birth.
- Limitations:
 - some data still exhibit non-symmetric distributions, especially pm2.5
 - omitted variable bias
 - measurement errors
 - missing data: 201 → 173 observations
- Future studies:
 - more explanatory variables
 - panel dataset
 - deep analysis on one factor: e.g. health expenditure