

# PLSC 503 – Spring 2018

## Regression, Conceptually

January 23, 2018

“Regression,” conceptually:

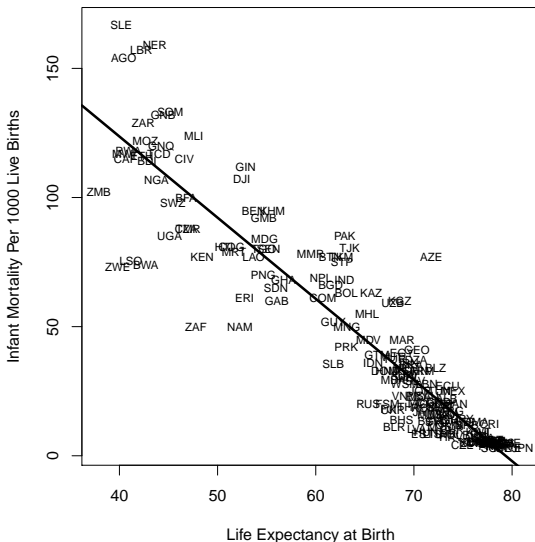
$$\Pr(Y|\mathbf{X}) = f(\mathbf{X})$$

Two important things:

- The distribution of  $Y$  is *conditional on all variables in  $\mathbf{X}$* , and
- The conditional distribution of  $Y$  is conditional on the *joint distribution* of the elements of  $\mathbf{X}$ .

→ Regression is hard...

Figure 1: Infant Mortality and Life Expectancy



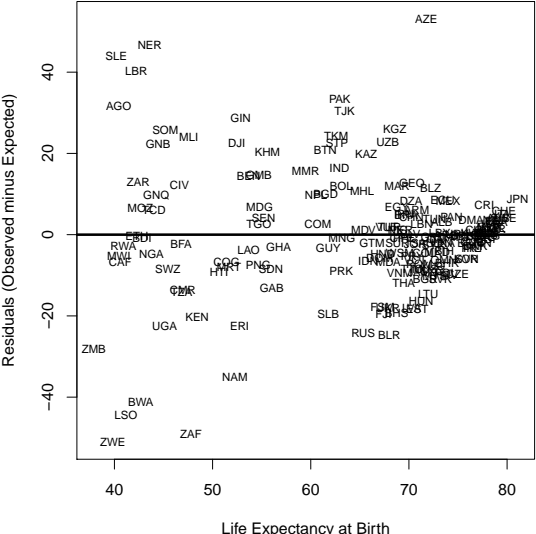


Figure 3: Infant Mortality and Fertility

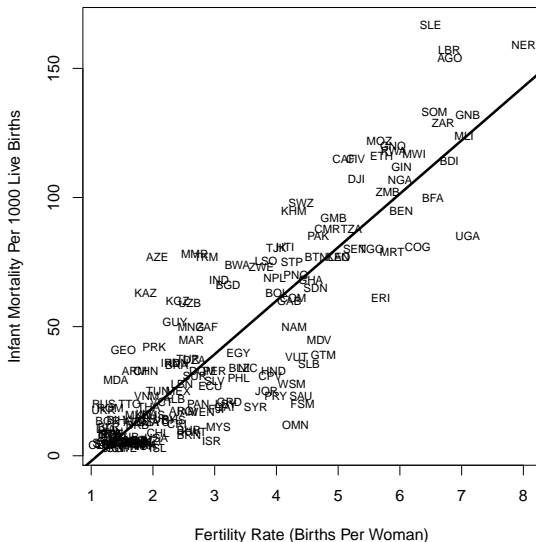


Figure 4: Infant Mortality and Wealth

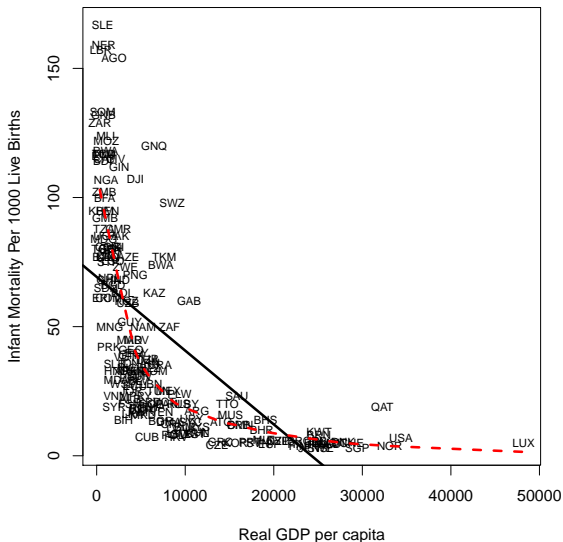


Figure 5: (Logged) Infant Mortality and (Logged) Wealth

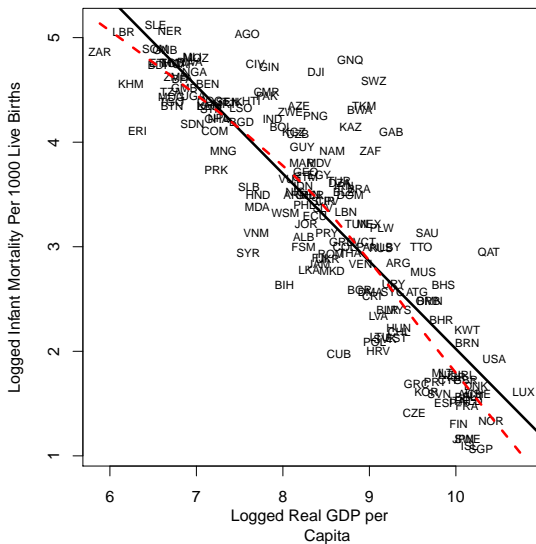


Figure 6: Infant Mortality and Democracy

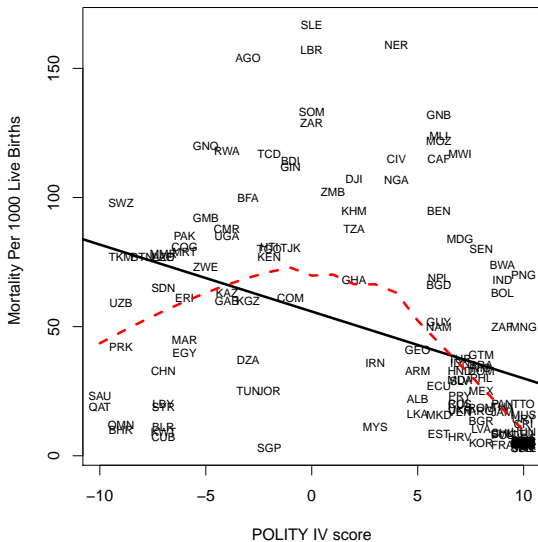




Figure 7: Infant Mortality, (Dichotomized) Wealth, and Democracy

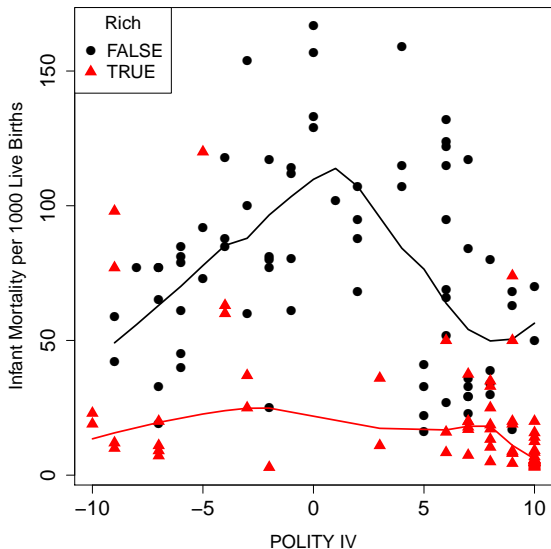


Figure 8: Measurement: National Health Indicators

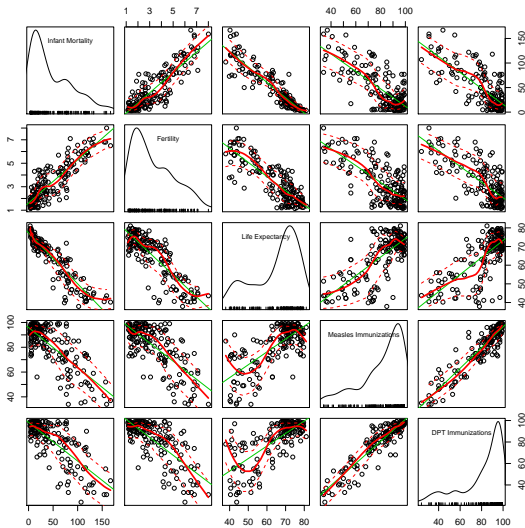
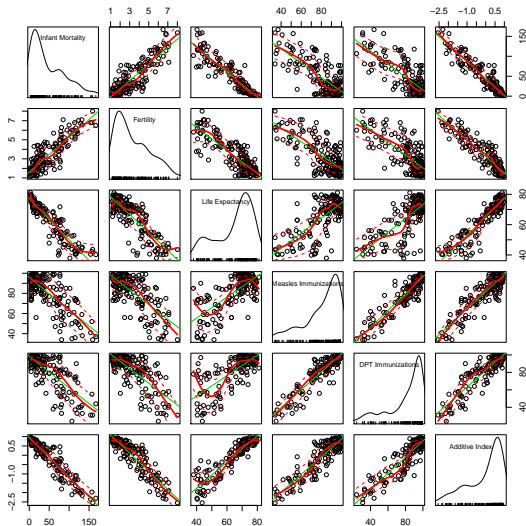


Figure 9: Measurement: National Health Indicators, Plus Additive Index



# Why regression?

	Description	Explanation	Prediction
<b>Task</b>	Summarize data	Correlation/causation	Forecast OOS / future data
<b>Emphasis</b>	Data	Theory / Hypotheses	Outcomes
<b>Focus</b>	Univariate	Multivariate	Multivariate
<b>Typical Application</b>	Summarize / "reduce" data	Discuss marginal associations between predictors and an outcome of interest	Optimize out-of-sample predictive power / minimize prediction error