Analyzing Dataset Child Mortality Using GBD Data 1970-2010

Dibbya Biswa and Su-Miao Lai
HMS 520 Final Project

Outline

- Background
- Research Aims
- Data Sources
- Descriptive Statistics
- Lessons we have learned
- Questions?

Background

- **SDG Target 3.2** | Newborn and child mortality: By 2030, end preventable deaths of newborns and children under 5 years of age, with all countries aiming to reduce neonatal mortality and under-5 mortality

Background

- Study population : Under 5 years old children
- Study location: South Asia countries
- Study metrics: Mortality rate
- The overall goal of this study is to evaluate whether the countries in South Asia are on track to meet SDG Target 3.2.

Research Aims

- Trends analysis for mortality rate from 1970-2010
  - Under-5 Mortality Rate
- Covariates for mortality rate in 2010
  - Water
  - Sanitation
  - sdi
- Regression model
  - Using data in 2010
  - Independent Variables: water, sanitation, sdi
  - Dependent Variable: Under-5 Mortality Rate

Data Sources: Exploring the data

- Covariates from GBD study
  - Education (years per capita) aggregated by age (15+) and sex
  - Sanitation
  - Water
  - sdi
- Country codes
  - Merged with GBD data

Data Sources: Exploring the data

- Child mortality rates
- Number of Deaths for each age group
- Covariate
  - sanditation
  - water
  - sdi

Descriptive Statistics

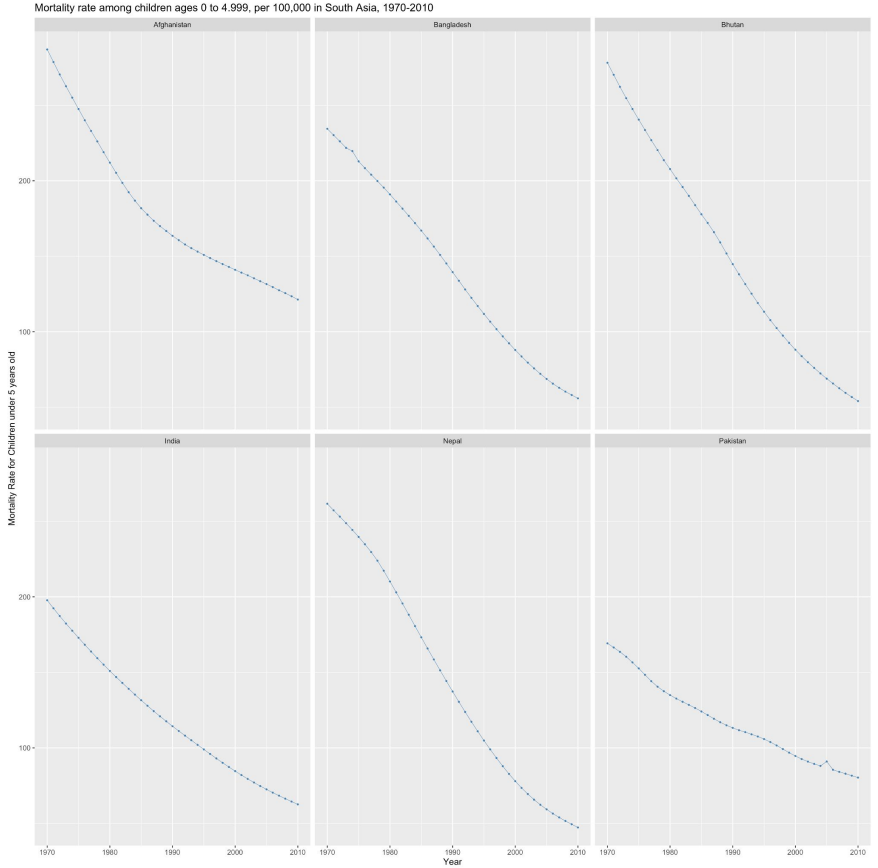
- Data visualization
  - Static graph: **ggplot2**
    - Time-series in South Asia
  - World Map**: ggplot2**
    - Under 5 mortality rate in South Asia
  - Interactive Map: **rbokeh**
    - Live map of under 5 deaths for world map
  - Web Apps: **shiny**
    - Web app displays the mortality rate for each country

Change of mortality rate

Over time (goem_plot)

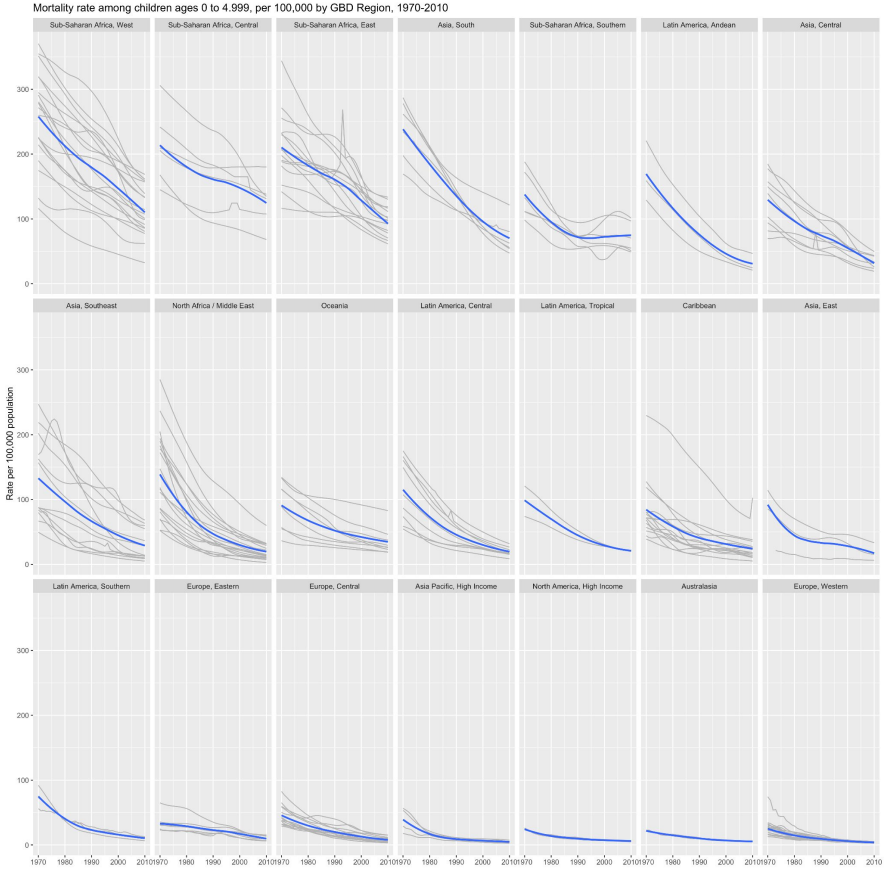
Compare between countries in South Asia region

- South Asia have extremely higher mortality rate, as expected.
  - But not expected to be in 300's level.
- Pakistan,Bangladesh and India have lower to start with in 1970.
- Bhutan, Nepal, and Afghanistan have high rate in 1970
  - But are decreasing over time.



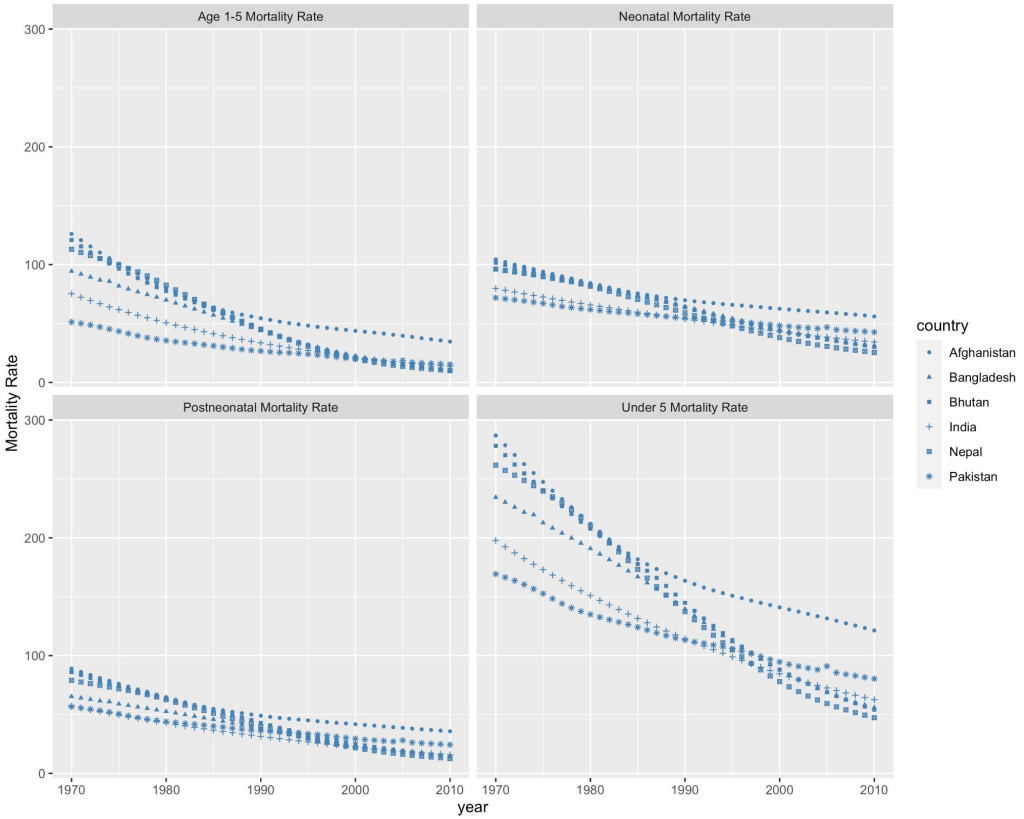
Data Visualization

- Under 5 Mortality rate trends
- GBD regions
  - Gray lines represent countries
  - Blue is estimate of all counties combined



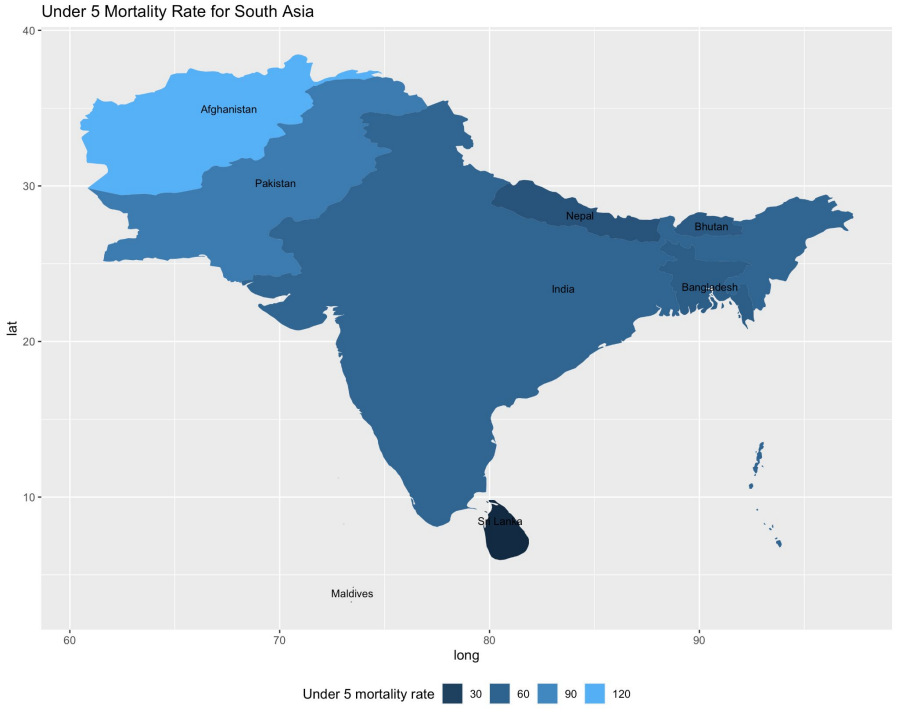
Data Visualization

- Mortality rates of South Asia
  - 4 age group over years
  - Not too surprising that Neonatal had higher value in 1970.
    - Maybe due to lack of malnutritions
  - Decreases as age group increase over the years



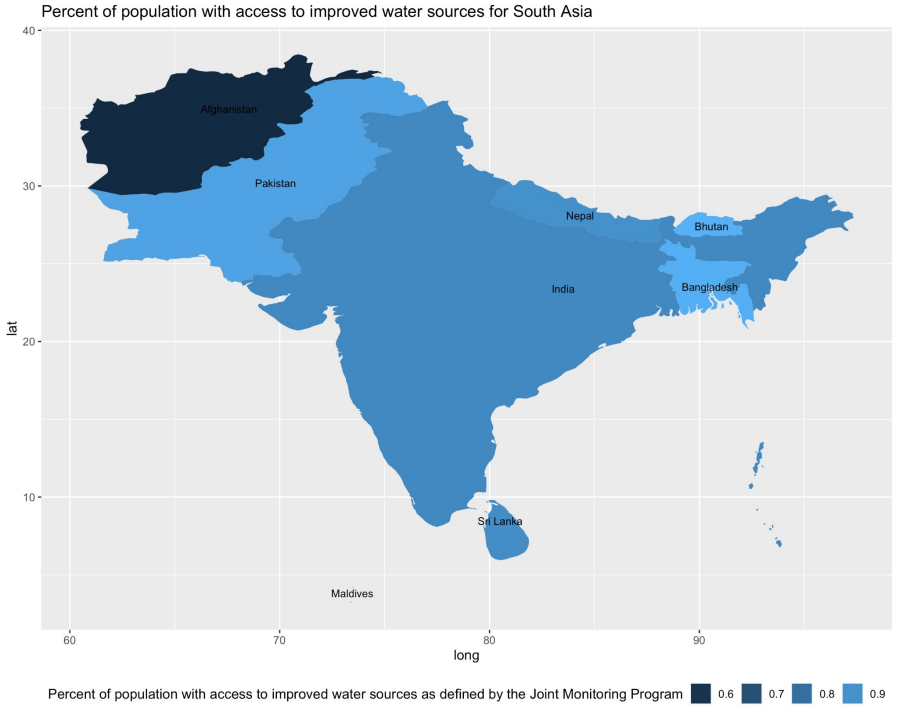
Data Visualization

- Under 5 Mortality Rate
  - ggplot2
    geom_polygon()



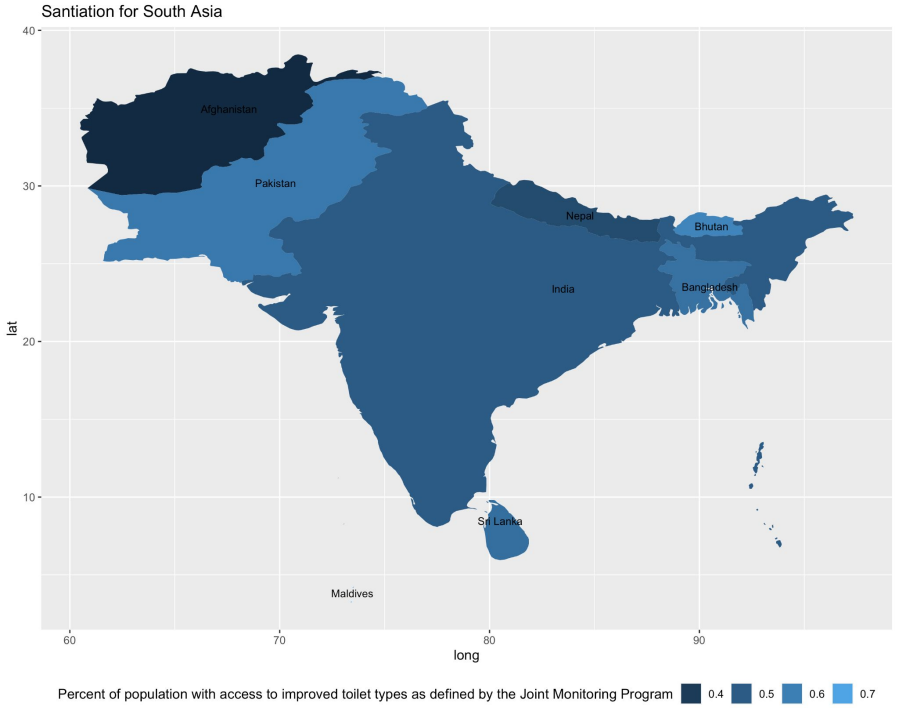
Data Visualization

- Water
  - ggplot2
    geom_polygon()



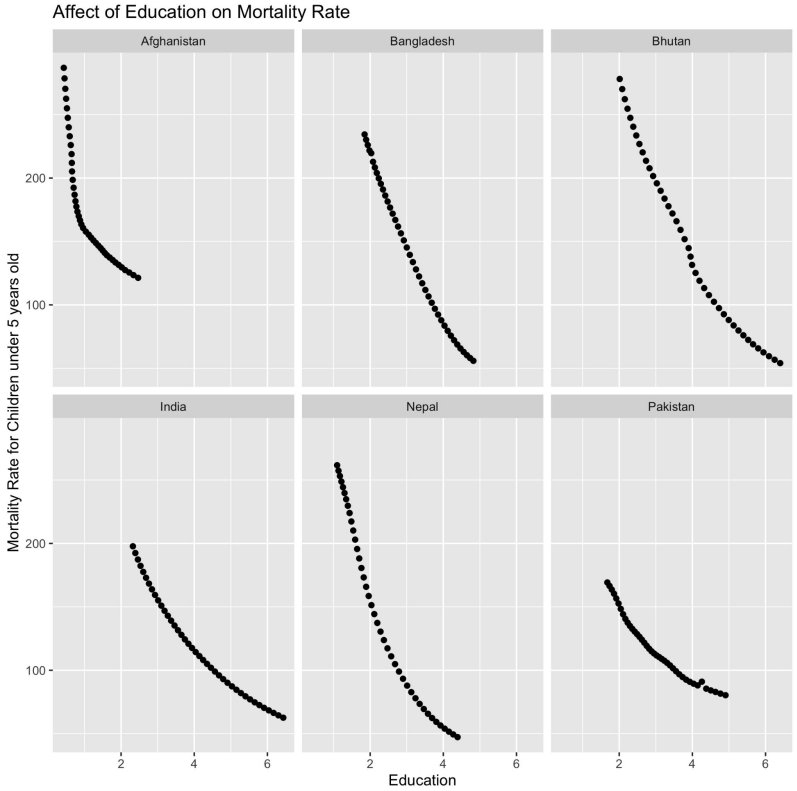
Data Visualization

- Sanitation
  - ggplot2
    geom_polygon()



Covariates association with mortality rate

- Effect of Education on mortality rate
- Less education has high mortality rate.
- Supering that education does show association with mortality.



Data Visualization

- Interactive Map
- Web app

Linear Regression Model

Under5MR ~ -89.909*sdi -85.474*water -42.847*sanitation

- The results show sdi, water and sanitation are statistical significantly associated with the under 5 mortality rate.

Lessons we have learned

- World map is a great tool to demonstrate the population health metrics at the country level.
- Shiny apps allow users to interact with static plots.
- Data Visualization is fun but data wrangling is time consuming and needs patience.
- Data visualization is great way to communicate with the audience. It’s important to keep in mind that the graphs should focus on conveying the information concisely.
- Study design is the key.

Questions?