Project Proposal: ISyE 6414 (Dr. Yajun Mei) - Spring 2020

Samuel Garcia, Michael Szostak, Osman Ghandour, and Peter Williams

2020-03-01

(b) Project Description

Adolescent birth rate is a progress indicator for Millennium Development Goal target 5, which is to improve maternal health. It is generally accepted that socioeconomic factors influence adolescent birth rates. . . .

Adolescent birth rate (ABR) is the annual number of births to women 15 to 19 years of age per 1,000 women in that age group.

(c) Dataset Description: How and From Where

Adolescent birth rate: the annual number of births to women 15 to 19 years of age per 1,000 women in that age group

Adolescent birth rate data -> http://apps.who.int/gho/data/view.main.1630AG?lang=en

GDP per capita ->https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD?view=map

Education Index -> https://data.oecd.org/eduatt/adult-education-level.htm

Female labor force participation rate -> http://hdr.undp.org/en/content/labour-force-participation-rate-female-male-ratio

UN info about the adolescent birth rate:

http://mdgs.un.org/unsd/mi/wiki/5-4-Adolescent-birth-rate.ashx

Scientific Research Questions To Address (d)

Can we statistically confirm and identify which measures/indicators are best associated with adolescent birth rates?

Do factors such as development of a country (GDP per capita, Adult education level) affect ABR?

Do factors such as M/F ratio at different stages of life (Ratio of girls to boys in primary, secondary, and tertiary education) affect ABR?

Do health care factors affect ABR (Nurses and midwives (per 1000 population), Country's Current health expenditure (% of GDP) per year)?

Does migration to a country affect ABR?

Which factors seem statistically insignificant for association with ABR?

Do certain factors play a larger role depending on the region?

Proposed Statistical Methods and Models (e)

Multiple Linear Regression: Discover what variables contribute to describing the outcome of interest, along with their statistical significance. Purpose to describe the dataset intuitively, and identify statistical significance of particular explanatory variables on our outcome of interest.

Penalized Linear Regression Methods: Develop a predictive model designed to minimize out-of-sample error and risk, while also incorporating as many variables into a model as is sensible. The objective being to build a model that is focused on predicting future outcomes related to ABR for practitioners as opposed to describing the factors contributing to the response. The intention being to help practitioners to identify patients at higher risk?