Analytics Lab 2016

ABOUT

Each year, ~15 million babies are born preterm, ~150 million children have stunted growth, and >600 million children have impaired neurocognitive development. These easily preventable outcomes are linked to diseases and malnutrition in many low-income countries.

The Gates Foundation launched a global health initiative named Healthy Birth, Growth, and Development (‘HBGD’) to address these complicated and interrelated outcomes. The team has created an integrated knowledge base consisting of existing maternal and child health data from various clinical and population survey studies in different countries. At present, their knowledge base includes some 1,700 demographic, clinical, and socioeconomic covariates from more than 8 million children.

PROJECT SCOPE

Our task was to analyse the dataset provided by The Gates Foundation and generate actionable insights that could be used to form relevant interventions to improve children’s health.

We also dedicated a large amount of time to back-filling missing data to help drive more statistically significant outcomes.

DATA

The dataset was aggregated from 27 longitudinal studies measuring child growth between birth and age 16 across 12 countries. It was a multi-variate time series dataset with ~800,000 data points and 46 independent variables, comprising:

- Continuous variables, such as birth length and mother’s weight at birth
- Ordinal variables, such as the APGAR1 score
- Nominal variables, such as the feeding mode

Studies were focused on lower-income developing countries, and large amounts of data had missing values.

KEY FINDINGS / RECOMMENDATIONS

1. PAST Z-SCORES ARE RELIABLE INDICATORS OF THE FUTURE

We analysed children’s weight and height z-scores based on WHO standards and found that we could accurately predict a child’s future z-score based on historic values using a linear regression model.

Actionable insights for The Gates Foundation

I. The model could be used by clinics and hospitals to prioritize cases for the doctors so they can spend more time analyzing cases where our model predicted a higher probability of stunted growth

II. We can create charts based on our model that helps parents to quickly identify at-risk cases and report them to doctors proactively

2. WE CAN PREDICT MOVEMENTS IN Z-SCORES

We also analyzed various demographic factors for the children in our study group, and found that we could identify key drivers of significant movements in children’s z-scores and accurately predict such movements based on those factors.

Actionable insights for The Gates Foundation

I. Use identified demographic factors to help shape government policy. We discovered that children with significantly decreasing z-scores had almost double the breast-feeding duration than children with significantly increasing z-scores. Governments (or NGOs) could, for example, develop a welfare scheme for mothers with newborns to support reduced dependence on breastfeeding in later life

II. Other model outputs could be provided to doctors and other medical practitioners to take precautionary measures to avoid some of the extreme results that were predicted

David Rolls
MBA 2017

Angad Singh
SDM

Yunyue (Nikki) Li
MBAn

Ravi Subrahmanyan
Project Advisor