

Understanding Medellín's Education System through Data Analysis and Modelling: On the Path to Innovation

“Medellin is an extremely beautiful city, on the way to become a major centre and very concerned with the role of science and technology in its development and future.”

Douglas Osheroff, Nobel Prize in Physics 1996

Project Background

Medellín is the second city of Colombia. It was once described by TIME as the most dangerous city in the world. The situation has now reversed, as a result of one of the most innovative projects ever undertaken to revive a city at the point of self-destruction. Medellín has transformed itself to become a world model of social inclusion, and a nucleus, for science and technology in Colombia. The city is an intellectual hub for Latin America, and is one of six cities that won the 2010 United Nations Human Settlement Program (UN-HABITAT) award.

Medellín has funded various social strategies in order to improve education around science, technology and innovation, demonstrating the city's vision and commitment to its long-term development. One of the goals that Medellín has is to improve the quality of teaching in order to improve education, and attain high educational standards. It is envisioned to redesign the education system, and the role of the teacher in the system to become an even more active, and fundamental part of the education system.

In order to achieve this, Parque Explora, together with the University of Antioquia, is performing an exhaustive quantitative and qualitative characterisation of Medellín's teachers. This involves a critical mobilisation for data collection, an appropriate statistical treatment and data analysis, together with system modelling, something the city hasn't done before. By this study it is intended to build a solid model that highlights the critical variables in education policy formulation, and the strategies that will tackle observed deficiencies in the system.

Project Outline

The aim of this project is to understand the dynamics, which affect the success of the education system in the city of Medellín, in order to help policymakers develop effective education policies.

A variety of quantitative and qualitative data has been collated, comprising of teacher demographics, quality of education and student performance. In the proposed project, this data will be analysed in order to develop data-driven public policy.

The data has been collated from a variety of organisations, which has meant that data sets have been analysed in isolation using standard statistical methods; hidden dynamic and connections between the different variables explored, have therefore been ignored. Consequently a systems understanding of educational performance has not yet developed; this project will therefore develop a complex systems understanding of educational performance.

Through data mining and classification, we expect to be able to build models that are able to identify connections between variables, capable of describing the inherent dynamics of the system. In particular, preliminary analysis of the data suggests that social network models are suitable to describe certain aspects of the data such as teachers' mobility. On the other hand agent-based models are more suitable to describe other dimensions, such as class size, investments and socio-geographic indicators.

The aim is that the developed models will be the basis upon which effective decisions can be made about developing a new education policy for Medellín. The models will explore some general rules that can be applied to public policy challenges of this type in any region, whilst incorporating the nuances that characterise different national contexts.

The collated data contains information on 10,000 teachers in the city including but not limited to gender, age, educational qualifications, current status, location etc. An initial statistic analysis will be performed in order to observe relations between educational deficiencies and performance, dominant populations and minorities, socioeconomic status, and violence and how these interrelationships affect each other.

Objectives

The specific objectives of the proposed project in the context of the MathSys CDT program for the period of mid-June to mid-September are:

- Familiarise with the data collated by the partner organisations in Medellín.
- Select data sets that will be subject to in-depth data mining processes.
- Explore different data mining and data analysis methods on the selected data to determine variables, co-dependencies, classifications, relevance, etc.
- Explore (social) network models and agent-based models that are suitable to describe the relationships amongst the variables and their dynamics.
- Select one particular model that would help to understand a relevant issue for policymaking and carry out a more exhaustive validation.
- Document the validation of the model.

We expect that the proposed objectives and the scope of the project are reasonable for the time window available and at the same time set the stage, and the path for a more comprehensive project that can lead to a doctoral thesis.

“A city is like rivers, sometimes goes up and sometimes goes down. Medellín was at the bottom now is at the top.” Rodolfo Llinas, Professor of Neuroscience

Polymaths Consulting

20 February 2015