30 credits - Population predictions in Swedish municipalities using data mining, machine learning and statistical modelling

Background

BizOne has worked with various Swedish municipalities who face the challenge of predicting short-term and long-term population changes. Generally today's predictions are done by manually analyzing aggregate data and current trends such as immigration flows to form an idea of how the population will evolve. The predictions are used in the planning of property development and preparing for the future need of education and welfare etc. Historically predictions have often been inaccurate, sometimes leading to the wrong decisions, lack of necessary investment and failed budgets. A substantial improvement of population predictions could have a highly positive impact on individual municipalities planning and economy, on their citizens' quality of life and in the end on the Swedish society as a whole.

Target

The purpose of the thesis project is to create a statistical model and a functional algorithm for predicting population growth in different municipalities across Sweden, based on currently available data.

Assignment

The assignment will consist of answering the key questions below and putting the insights to work in a statistical model/algorithm.

- What are the variables that best predict population growth in a Swedish municipality, based on historical data?
- What is the effect of uncertainty about future values of highly unpredictable variables, such as immigration, on the precision/accuracy of the models used
- What is the feasibility of predicting population evolvement with the currently existing data sets, that is; On what level of detail is it possible to make statistically relevant predictions, for example township, residential district, block?

The given insights will form hypotheses to be tested and refined by statistical modelling and simulation using data mining and machine learning. This in the end should lead to a robust

statistical model and a functional algorithm, as well as recommendations for future data gathering that could improve the model.

BizOne wishes to work closely together with the students throughout the whole process. Population data for certain collaborating municipalities as well as overall Swedish population data will be provided by BizOne, but the students are of course welcome to suggest or integrate other data sources as found necessary during the project.

Education: Relevant Master's degree in Computer Science and/or Statistics
Number of students: 1-2
Start date: January 2017
Estimated time needed: 20 weeks

Contact persons and supervisors
Supervisor at BizOne AB: Andreas Granström
Supervisor at Uppsala University: Michael Ashcroft