Internship Province of Noord-Holland Datalab:
Business Park Dynamics

Project description

There are currently hundreds of thousands of businesses situated in the Province of Noord-Holland (PNH), all of them depending on a suitable location that facilitates most of their current and future needs. These businesses are quite diverse in nature and each individual business can also change over time, for instance by growing or scaling up. This results in interesting dynamics involving a lot of movements across the province of businesses switching one location for another. At PNH we would like to better grasp these dynamics. In particular, we are interested in the business parks, as these are the districts that can be invested in, transformed, and made future proof. This way PNH can play an active role guiding the involved municipalities.

For the current project we will make use of the LISA administration as the main data source. This administration holds information about all businesses of PNH, including location information, number of employees, and business sector. Other data sources could be linked with LISA in order to examine the attributes of the business parks as relevant factors or to analyze the relation with the economic developments of particular business sectors.

Modeling framework to expect

PNH would like to predict, in the near future, the number of businesses that move in and out each business park. This may either require a prediction model that generates predictions for each individual business, or estimates at a business park or regional level. Additionally, we would like to understand the movements' directions. Therefore we would like to incorporate the start and end position of movements within the predictions.
Together with a VU faculty member and a PNH staff member you will develop such models. Given the nature of the research question and the data, you will likely use a non-linear time series model for durations (time-to-stay in a business park) or for transition probabilities. Given the number of firms in the data, further algorithms may be used (like clustering) to reduce dimensionality and yield interpretable results. As such models typically go outside standard toolboxes, we require good programming skills and a drive to go beyond the standard models in your curriculum.

Work environment

The province of Noord-Holland is based in Haarlem (www.noord-holland.nl). Some of our main tasks are sustainability, spatial development, mobility, and the conservation of nature. You will be working with colleagues from the Datalab in a very open atmosphere. The Datalab consists of 6 people. We offer a cloud-based data science platform which you can use to work on your code (preferably in Python), store data, and run models. Depending on the Covid-19 situation the work will be mostly from home. There is an internship compensation of € 625,- gross per month, based on a 36-hour work week. Prior to the internship we ask you to provide a statement of conduct, and to sign an integrity statement.

Interested

Apply with a letter of motivation, grade list and CV to masterthesis.eds.sbe@vu.nl.