TUM

Linking local government data with RecordLinkage

Project description

- Datasets often lack common identifiers. Therefore, one challenge is to best-possibly link records to make use of information stored in different datasets.
- A common approach is to use the respective entity names as the basis for this linkage, and to apply linking methods exploiting machine learning and stochastic approaches.
- The goal of the IDP is to link two datasets on U.S. local governments using the python package "RecordLinkage".
- In particular, you will link a large-scale Census dataset to a dataset on local government bonds.
- The outcome of the project is an optimal crosswalk between the two datasets.

What we are looking for

- Advanced python skills
- Ideally experience with the RecordLinkage package
- Interest to learn something about finance, in particular about corporate datasets
- 1-2 persons

What we offer

- Knowledge in (corporate) finance
- Kick-off session including introduction to relevant finance and/or business topics
- Experience with IDPs
- Open dialogue and support
- Access to prime capital markets databases

Interested?

Please send an e-mail with CV, academic transcript and your preference for this project to lisa.knauer@tum.de.

Questions?

In case of any (e.g. topic related) questions, please contact Lisa Knauer (lisa.knauer@tum or call +49 89 289 25196).