Prof. Melanie Krause, Ph.D.
Universitätsprofessorin für Nachhaltige Immobilien- und Stadtentwicklung
Universität Leipzig
Grimmaische Straße 12, Raum 308
04109 Leipzig
Email: melanie.krause@uni-leipzig.de

UNIVERSITÄT LEIPZIG

Big Data and Smart Cities

Master of Science in Management, Compulsory Course (for Students with Sustainability Focus)
Module Number 07-201-2101.SÜ01

Content of the Module

Data is everywhere - and the amount available is growing by the second. At the same time, our world is becoming more urbanized. More than half of the world's population is now living in cities, and urban development has to be managed in a sustainable way.
In this course, students will learn how to gather, process and analyze (big) data. They will get know statistical methods to draw insights from (big) data, discover patterns and relationships, and make predictions for urban life. In a nutshell, the question to be answered is: How we use (big) data in a responsible way to make cities more efficient and sustainable?

Goals of the Module

Students are able to reflect on the importance of data science to manage cities more efficiently and sustainably. They can independently analyze large amounts of data, from domains such as housing, transportation or pollution. They know methods of statistical data analysis and machine learning and can apply them to practical case studies, interpreting and questioning the results. They are able to present their findings in English and defend them argumentatively.

Module Format

Lecture "Big Data and Smart Cities" (2 SWS) = 30 hours of class attendance und 105 hours of independent study = 135 hours
Seminar with practical part "Big Data and Smart Cities" (4 SWS) = 60 hours of class attendance und 105 hours of independent study = 165 hours
Time and date: Monday, 9:15-10:45, 11:15-12:45, 15:15-16:45, Lecture Hall Seminarraum 12 - I131 I1.002

Grade

Students work together in groups to conduct their own urban data analysis project, applying machine learning methods, discussing and reflecting upon their results. The grade for the module is made up of a 30-minute presentation (30%) and a 12-page written paper (70%).
The **presentations** take place on 3 and 10 July.
Deadline for submitting the 12-page **seminar paper** of each group as a PDF by email to melanie.krause@uni-leipzig.de: **16.07.2023** (23:59)

Literature

No text book covers the material of the whole module. There are various textbooks on urban analytics and statistical learning which will be made reference to in the individual units.

Module Overview

1. 03.04.2023: **Getting Started**
    - Unit 1: Introduction
    - Unit 2: Course Overview and Projects
    - Unit 3: Data Types
2. 17.04.2023: **Describing Data**
    - Unit 4: Data Analysis Methods
    - Unit 5: Data Types in Python
    - Unit 6: Descriptive Statistics in Python
3. 24.04.2023: **Statistical Data Analysis, Modern Data and Exercises**
    - Unit 7: Hypothesis Tests
    - Unit 8: Data from the Internet (Google Trends API, Webscraping)
    - Unit 9: Exercises on Data Analysis in Python
4. 08.05.2023: **Linear Regression Models**
    - Unit 10: Linear Regressions
    - Unit 11: Case Studies - Linear Regressions in Urban Data Analysis
    - Unit 12: Linear Regression in Python
5. 15.05.2023: **Moving on to Machine Learning**
    - Unit 13: K-Nearest Neighbors and Model Complexity
    - Unit 14: Big Data Regularization Methods (Ridge and Lasso)

- Unit 15: Regularization Methods in Python + *First Project Support Session*

6. 22.05.2023: **Tree-Based Machine Learning Methods**
   - Unit 16: Decision Trees
   - Unit 17: Random Forests
   - Unit 18: Gradient Boosted Machines
7. 05.06.2023: **Data Pre-Processing and Neural Networks**
   - Unit 19: Data Pre-Processing
   - Unit 20: Neural Networks
   - Unit 21: Neural Networks in Python
8. 12.06.2023: **Data Sampling and Interpretation of Results**
   - Unit 22: Data Sampling
   - Unit 23: Interpretation of ML Results
   - *Project Support Session*
9. 19.06.2023: **Case Studies and Further Applications**
   - Unit 24: Case Study - Prediction Wildfires with ML
   - Unit 25: Neural Networks in Image Classification
   - *Project Support Session (including Academic English)*
10. 26.06.2023: **Text-Based Machine Learning Methods**
    - Unit 26: Natural Language Processing I
    - Unit 27: Natural Language Processing II
    - *Project Support Session*
11. 03.07.2023: **Course Wrap-Up and Student Presentations**
    - Unit 28: New Research Frontiers
    - Student Presentations
    - Student Presentations
12. 10.07.2023: **Student Presentations**
    - Student Presentations
    - Student Presentations
    - Student Presentations

This version is of 13 March 2023