JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

GGS Gießener Graduiertenzentrum Sozial-, Wirtschafts- und Rechtswissenschaften

Online-Workshop

Data Science Hands-on: SQL and Data Visualization for Research and Business Intelligence

Organizational details

Instructor:	Dr Tobias Keller
Date:	July 9 and 10, 2020, 9.00 - 17.00 h
Location:	Webex Conference Room (more information follows registration)
ECTS:	2
Max. participants:	12

Objectives

Learn how to manage data using MySQL, how to visualize data using Tableau, and how to bring it all together in your research project and building your first Business Intelligence application.

MySQL is an open-source relational database management system which allows data management using the Structured Query Language (SQL), the most widely used database language.

Tableau is a popular, proprietary data visualization software that offers a free (but fully functional) academic version for (doctoral) students.

Content and methods

Data management and visualization are essential to both empirical research and data science in practice. This 2-day-course covers the basics of working with relational database management systems using SQL and a state-of-the-art data visualization software, Tableau. Participants will learn by examples and exercises from the instructor's experience in research and practice.

Exercises will make up about 40% of the course time. Students will complete those using their own computers. Please see the preparation requirements below for a list of software to be installed. The exercises will be based on exemplary datasets that will be provided to the participants before the course.

Structure

1. Data Science and Business Intelligence by Example
   - The Data Scientist's Role
   - Business Intelligence Basics
   - Examples of Data Science & Business Intelligence from Practice and Research
2. Data Management using MySQL
   - The MySQL Project and the SQL Language
   - Creating tables and importing data
   - Retrieving data (SELECT)
   - Filtering and sorting data (WHERE , ORDER BY)
   - Changing data (UPDATE)
   - Inserting data (INSERT)
   - Aggregations, summary statistics (GROUP BY)
   - Joining tables
   - Using views
   - Optional: Using subqueries
3. Data Visualization using Tableau
   - Connecting with your data
   - Dimensions vs. Measures
   - Visualizing data in Tableau worksheets
     - Using different types of charts. i.a.:
       - bar charts
       - geographic maps
       - scatter plots
     - Pagination and filtering
     - Formatting
     - Working with statistics and trend lines
     - Sharing worksheets (i.a. Tableau Reader, PDF, Powerpoint)
   - Building interactive dashboards
   - Calculations
4. Synthesis: Building a Business Intelligence Application using MySQL and Tableau

To gain the ECTS credit points participants have to:

- Setup the MySQL Community Server on their computer
  - Download here: https://dev.mysql.com/downloads/mysql/
  - Windows users, please follow the instructions in the manual provided with the downloads.
  - Setup a user account in your local MySQL Server
  - Create a new, empty database called "worldbank"
- Install Tableau Desktop and, if they have a student ID, apply for a free one-year license
  - https://www.tableau.com/academic

- Download and extract the course material and exemplary datasets that will be provided in time before the course.
- Familiarize yourself with the exemplary datasets from the course material. They have originally been downloaded from the World Bank Open Data Platform (http://data.worldbank.org/data-catalog/world-development-indicators).
- In order to better understand the research questions we will be tackling, take a look at the World Development Indicators Report, downloadable here: https://openknowledge.worldbank.org/bitstream/handle/10986/26447/WDI-2017-web.pdf

Target group

Doctoral candidates or postdoctoral researchers doing empirical research or intending to work as data scientists.

Course language

English (German, if only German participants)

Please note: As this is not an English language course proficiency in English at the C1 level of competency is required.

About the Instructor

Dr. Tobias Keller has been working as Data Scientist at ZERO.ONE.DATA, the big data startup of Deutsche Bahn AG since 2016. He consults on and applies machine learning and statistics for artificial intelligence systems in a Big Data environment. Furthermore, he teaches data science at Deutsche Bahn, in the doctoral education programs at the Justus-Liebig-Universität Gießen, and in the doctoral program and masters program at the WHU – Otto Beisheim School of Management. His research interests include machine learning and artificial intelligence, finance and accounting, strategic management.

Registration

By June 29, 2020 via e-mail at info@ggs.uni-giessen.de.