

# Master in Life Sciences

A cooperation between  
BFH, FHNW, HES-SO, ZFH

|  |   |
|--|---|
| <b>Module title</b>                      | <b>Modelling and Exploration of Multivariate Data</b>   |
| <b>Code</b>                              | D3  |
| <b>Degree Programme</b>                  | Master of Science in Life Sciences  |
| <b>Workload</b>                          | 3 ECTS (90 student working hours) <ul style="list-style-type: none"> <li>- Lessons contact (total 42 of which 28 central teaching): 32 h</li> <li>- Self-study: 58 h</li> </ul>   |
| <b>Module Coordinator</b>                | <p><b>Name:</b> Thomas Ott<br/> <b>Phone:</b> +41 (0)58 934 56 84<br/> <b>Email:</b> thomas.ott@zhaw.ch<br/> <b>Address:</b> ZHAW Life Sciences und Facility Management, Einsiedlerstrasse 31a, 8820 Wädenswil</p>  |
| <b>Lecturers</b>                         | <ul style="list-style-type: none"> <li>• Thomas Ott, ZHAW</li> <li>• Lorenzo Tanadini, BFH</li> </ul>   |
| <b>Entry requirements</b>                | Attending the module “Handling and Visualizing Data” is required. Prior to the module, additional mandatory preparatory reading, exercises and other material (videos, tests) will be made available to facilitate students preparation for the module. Students are advised to start five weeks before the module with the required preparatory work;  |
| <b>Learning outcomes and competences</b> | <p>After completing the module, students will be able to:</p> <ul style="list-style-type: none"> <li>• explore multivariate data by means of suitable visualisation and dimensionality reduction techniques</li> <li>• explore and describe the structure of multivariate data using clustering</li> <li>• explore and describe time series data on the basis of suitable visualisations and analysis methods analogue to multivariate data analysis</li> <li>• interpret, visualise and communicate the results of the analyses</li> <li>• use multiple regression models to answer research questions, understand their advantage over univariate methods; fit these models with R and quantify the fit of the model, describe the limitations of precision and reliability of inferential results; test the model assumptions; apply counter measures in case of problems with model assumptions</li> <li>• use elementary nonparametric regression methods to estimate the shape of not necessarily linear regression curves, understand the advantages and limitations of such flexible methods and apply related tools</li> <li>• perform elementary model selection and understand associated problems; test hypotheses, construct confidence and prediction intervals</li> <li>• identify typical pitfalls and amend these problems</li> <li>• understand typical statements in empirical research articles.</li> </ul> |
| <b>Module contents</b>                   | <p>This module introduces <u>exploratory methods</u> and <u>regression models</u> for data analysis.</p> <p><u>Exploratory part:</u></p> <ul style="list-style-type: none"> <li>• Basic plots to characterise and visually inspect multivariate data and time series data</li> <li>• Dimensionality reduction techniques (principal component analysis, multi-dimensional scaling)</li> </ul>   |

- Clustering methods (k-means clustering and related approaches, hierarchical clustering, evaluation methods)

Modelling part:

- Simple linear regression (including transformations)
- Nonparametric regression (regression splines, local regression) and quantile regression
- Multiple linear regression (including regression diagnostics)
- Model selection (linked to hypothesis tests and p values) and inference (especially confidence intervals, prediction intervals)
- Model diagnostics: assessment the validity of the model assumptions, reflect on the tools used to do this assessment
- Possible strengths and limitations of parametric models (link to the exploratory part)

Both parts:

- Interpretation and visualisation of the results using suitable graphical representations of the data and the results (e.g. 3D scatter plots with regression surface or biplots)
- Scientific reporting of the results, backtranslation from statistical methods to answer the original research questions to the data

**Teaching / learning methods**

This module has the following structure:

| Week              | <1  | 1   | 2  | 3  | 4  | 5  | 6  | 7  | >7  |
|-------------------|-----|-----|----|----|----|----|----|----|-----|
| <b>Central</b>    |     | 8L  |    | 8L |    | 8L |    | 4L |     |
| <b>Local</b>      |     | 2L  | 4L |    | 4L |    | 4L |    |     |
| <b>Self-study</b> | 10h | 38h |    |    |    |    |    |    | 10h |

In the spring semester, the schedule may be slightly amended due to holidays. Details will be communicated five weeks before the course starts.

In the weeks before module start, students are expected to do preparatory work to level prior knowledge.. The workload is expected to be roughly 10 hours.

The students receive preparatory and/or follow-up self-study work for each course day. The self-study consists e.g. of preparatory reading/videos, follow up exercises, examining case studies, etc.

Central teaching is a blend of classical lectures with more interactive teaching approaches such as large group activation methods and small group exercise sessions. Each course day contains an exercise session in which students work on the exercises, can ask questions and get individual support.

Local teaching consists of students actively solving exercises together with the local teachers. These exercises are meant to deepen the understanding of the material, give an opportunity to practice, provide extensions etc. The main type of tasks will be case

# Master in Life Sciences

A cooperation between  
BFH, FHNW, HES-SO, ZFH

|                                       |  |
|---------------------------------------|--|
|                                       | <p>studies which illustrate and exemplify the application of the material from central teaching to real life data sets and real problems.</p> <p>All the course contents come with comprehensive lecture notes and additional videos for an individual study and/or online learning.</p> |
| <b>Assessment of learning outcome</b> | Final written individualised exam (open book, using individual laptop computers to run data analyses) (100%)   |
| <b>Format</b>                         | 7-weeks  |
| <b>Timing of the module</b>           | For ZHAW and FHNW: Autumn semester, CW 45-51<br>For BFH and HES-SO: Spring semester, CW 15-21  |
| <b>Venue</b>                          | For ZHAW and FHNW: Olten<br>For BFH and HES-SO: Fribourg   |
| <b>Bibliography</b>                   | Material will be provided on Moodle.   |
| <b>Language</b>                       | English  |
| <b>Links to other modules</b>         | This module builds on module D1 "Handling and Visualising Data" and complements the module D2 "Design and Analysis of Experiments".  |
| <b>Comments</b>                       | Material treated during local teaching is relevant for the exam.<br>Students have to make sure that an updated version of R is installed. Details will be communicated in advance.   |
| <b>Last Update</b>                    | 17.9.2019  |