



## Course and Examination Fact Sheet: Spring Semester 2022

### 8,273: Multivariate Statistical Analysis with Python

ECTS credits: 4

#### Overview examination/s

(binding regulations see below)

Decentral - Oral examination (individual) (60%)

Examination time: term time

Decentral - examination paper written at home (in groups - all given the same grades) (40%)

Examination time: term time

#### Attached courses

Timetable -- Language -- Lecturer

[8,273,1.00 Multivariate Statistical Analysis with Python](#) -- Englisch -- [Fengler Matthias Reginald](#)

#### Course information

#### Course prerequisites

Initial programming knowledge, not necessarily in Python, solid knowledge and interest in statistics, econometrics, and data analytics. Prior attendance of "Data Analytics I: Predictive Econometrics" may be advantageous.

#### Learning objectives

- Students will understand the challenges of multivariate data.
- Students will know the basic statistical techniques of data visualization, hypothesis testing, cluster analysis, dimensionality reduction, and discriminant analysis.
- Students will know how to apply and interpret these methods for multivariate data analysis.
- Students will know how to take advantage of Python for multivariate data analysis.

#### Course content

##### Short summary

The class introduces students to the most important methods in multivariate statistics. Thereby, students will learn how to use Python for multivariate data analysis.

##### Detailed Description

Increasingly larger, richer and more versatile datasets are nowadays available for economic and financial research - a phenomenon usually referred to as the "Big Data paradigm". Corporates, banks, and policymakers aim at extracting knowledge from such data sets for various purposes, e.g., for understanding their costumers, improving their product lines, offering new products and services, and for policy decisions and policy implementations. Most of the ideas and methods used for such "fancy and modern" analyses are, in fact, methods coming from multivariate statistics.

The class covers the key methods from multivariate statistics in order to extract insights from multivariate data, in particular variable selection, factor space reduction, factor modelling, and cluster analysis. The objective is to combine conceptual/theoretical material with the practical application using the open source programming language Python.

A typical application could be this: suppose you like to use indicators of economic activity to predict or describe some variable of interest, but you have hundreds of them. How can you find a small set of relevant predictors for your purpose? Is there a way to comprise all variables in a single index such as to distil the essential information from all data? Can you distinguish certain



groups among these variables, which may help understand their economic content?

The class connects to and expands on topics being part of "Data Analytics I: Predictive Econometrics" and naturally complements classes that more explicitly focus on the challenges of big data, such as big data analytics and machine learning. The course is recommended for students wishing to develop a profile in data science.

## Topics

1. Descriptive techniques for multivariate data
2. Multivariate random variables and distributions
3. Theory of Estimation and Hypothesis testing
4. Principal Component Analysis
5. Factor Models
6. Cluster analysis
7. Discriminant analysis

## Course structure and indications of the learning and teaching design

Weekly lecture with theoretical and practical exercises.

## Course literature

W. Härdle, L. Simar (2019): Applied Multivariate Statistical Analysis, 5th edition, Springer-Verlag

## Additional course information

In the case of the President's Board having to implement new directives due to the SARS-CoV-2 pandemic in SpS2022, the course information listed above will be changed as follows:

- The course is conducted online via the platform zoom;
- The recordings of the course are available for 30 days;
- The lecturer informs via StudyNet or e-mail on the changed implementation modalities of the course;
- Otherwise there are no changes necessary to the course information.

The examination information listed below would be changed as follows:

- The oral examinations are conducted online and are not recorded.
- Otherwise there are no changes necessary to the examination information.

## Examination information

### Examination sub part/s

#### 1. Examination sub part (1/2)

##### Examination time and form

Decentral - Oral examination (individual) (60%)

Examination time: term time

##### Remark

--

##### Examination-aid rule

Extended Closed Book

The use of aids is limited; any additional aids permitted are exhaustively listed under "Supplementary aids". Basically, the following is applicable:

- At such examinations, all the pocket calculators of the Texas Instruments TI-30 series and mono- or bilingual dictionaries



(no subject-specific dictionaries) without hand-written notes are admissible. Any other pocket calculator models and any electronic dictionaries are inadmissible.

- In addition, any type of communication, as well as any electronic devices that can be programmed and are capable of communication such as notebooks, tablets, mobile telephones and others, are inadmissible.
- Students are themselves responsible for the procurement of examination aids.

#### Supplementary aids

none

#### Examination languages

Question language: English

Answer language: English

---

## 2. Examination sub part (2/2)

#### Examination time and form

Decentral - examination paper written at home (in groups - all given the same grades) (40%)

Examination time: term time

#### Remark

Assignments (2-3 students collaborate)

#### Examination-aid rule

Term papers

Written work must be written without outside help according to the known citation standards, and a declaration of authorship must be attached, which is available as a template on the StudentWeb.

Documentation (quotations, bibliography, etc.) must be carried out universally and consistently according to the requirements of the chosen/specified citation standard such as e.g. APA or MLA.

The legal standard is recommended for legal work (cf. by way of example: FORSTMOSER, P., OGOREK R., SCHINDLER B., Juristisches Arbeiten: Eine Anleitung für Studierende (the latest edition in each case), or according to the recommendations of the Law School).

The reference sources of information (paraphrases, quotations, etc.) that has been taken over literally or in the sense of the original text must be integrated into the text in accordance with the requirements of the citation standard used. Informative and bibliographical notes must be included as footnotes (recommendations and standards e.g. in METZGER, C., Lern- und Arbeitsstrategien (latest edition)).

For all written work at the University of St.Gallen, the indication of page numbers is mandatory, regardless of the standard chosen. Where page numbers are missing in sources, the precise designation must be made differently: chapter or section title, section number, article, etc.

#### Supplementary aids

none

#### Examination languages

Question language: English

Answer language: English

---

## Examination content

The oral examination (60%) covers the following topics:

1. Descriptive techniques for multivariate data
2. Multivariate random variables and distributions
3. Theory of Estimation and Hypothesis testing
4. Principal Component Analysis
5. Factor Models



6. 6. Cluster analysis
7. 7. Discriminant analysis

The examination paper (40%) consists in a group project on a specific assignment.

## Examination relevant literature

W. Härdle, L. Simar (2019): Applied Multivariate Statistical Analysis, Springer-Verlag

### Please note

Please note that only this fact sheet and the examination schedule published at the time of bidding are binding and takes precedence over other information, such as information on StudyNet (Canvas), on lecturers' websites and information in lectures etc.

Any references and links to third-party content within the fact sheet are only of a supplementary, informative nature and lie outside the area of responsibility of the University of St.Gallen.

Documents and materials are only relevant for central examinations if they are available by the end of the lecture period (CW21) at the latest. In the case of centrally organised mid-term examinations, the documents and materials up to CW 12 are relevant for testing.

Binding nature of the fact sheets:

- Course information as well as examination date (organised centrally/decentrally) and form of examination: from bidding start in CW 04 (Thursday, 27 January 2022);
- Examination information (regulations on aids, examination contents, examination literature) for decentralised examinations: in CW 12 (Monday, 21 March 2022);
- Examination information (regulations on aids, examination contents, examination literature) for centrally organised mid-term examinations: in CW 12 (Monday, 21 March 2022);
- Examination information (regulations on aids, examination contents, examination literature) for centrally organised examinations: two weeks before the end of the registration period in CW 15 (Monday, 11 April 2022).