

## Syllabus

### Course description

<b>Course title</b>	<b>Data Science for Social Sciences</b>
<b>Course code</b>	27276
<b>Scientific sector</b>	SECS-S/01
<b>Degree</b>	Bachelor in Economics, Politics and Ethics
<b>Semester</b>	2nd semester, a.y. 2024/2025
<b>Year</b>	2
<b>Credits</b>	8
<b>Modular</b>	No

<b>Total lecturing hours</b>	48 (42 Prof. Di Lascio + 6 Prof. Carlotti)
<b>Total lab hours</b>	--
<b>Total exercise hours</b>	12
<b>Attendance</b>	Highly recommended, but not required
<b>Prerequisites</b>	It is highly recommended both Mathematics for EPE and Statistics for EPE.
<b>Course page</b>	<a href="https://www.unibz.it/en/faculties/economics-management/bachelor-economics-politics-ethics/">https://www.unibz.it/en/faculties/economics-management/bachelor-economics-politics-ethics/</a>

<b>Specific educational objectives</b>	<p>The course is a complementary subject belonging to the scientific area of Statistic-Mathematic and covers statistical methods for data science in the context of social sciences.</p> <p>The course aims to provide students with the advanced methods of statistics devoted to data analysis in economics and social sciences. The theoretical aspects are complemented by modern data analysis with R and Python.</p> <p>By the end of the study of this course, the student should be able to</p> <ul style="list-style-type: none"> <li>- reduce the dimensionality of a data sets;</li> <li>- uncover structure within the data;</li> <li>- analyze a time series by applying the appropriate model;</li> <li>- validate a statistical model and provide a more accurate measure of its real-world effectiveness;</li> <li>- analyze and retrieve textual data;</li> <li>- use the software R and the programming language Python to perform advanced data analysis.</li> </ul>
--	---

<b>Lecturer</b>	<p>F. Marta L. Di Lascio  Office E510a  e-mail: <a href="mailto:marta.dilascio@unibz.it">marta.dilascio@unibz.it</a>  Tel: 0471/013285  <a href="https://www.unibz.it/en/faculties/economics-management/academic-staff/person/32845-francesca-">https://www.unibz.it/en/faculties/economics-management/academic-staff/person/32845-francesca-</a></p>
-----------------	---

	<a href="mailto:marta-lilja-di-lascio">marta-lilja-di-lascio</a> Benedetta Carlotti Office I2.01 E-mail <a href="mailto:benedetta.carlotti@unibz.it">benedetta.carlotti@unibz.it</a> Lecturer's page <a href="https://www.unibz.it/it/faculties/economics-management/academic-staff/person/45699-benedetta-carlotti">https://www.unibz.it/it/faculties/economics-management/academic-staff/person/45699-benedetta-carlotti</a>
<b>Scientific sector of the lecturer</b>	SECS-S/01 (Prof. Di Lascio) SPS/04 (Prof. Carlotti)
<b>Teaching language</b>	Italian
<b>Office hours</b>	21 hours Prof. Di Lascio + 3 hours Prof. Carlotti Cockpit – students' zone – individual timetable Webpage: <a href="#">Timetables / Free University of Bozen-Bolzano (unibz.it)</a>
<b>Lecturing assistant</b>	Somenzi Damiano Marino <a href="mailto:Damiano.Somenzi@unibz.it">Damiano.Somenzi@unibz.it</a>
<b>Office hours</b>	By appointment, please send an e-mail and refer to the lecturer's timetable
<b>List of topics covered</b>	<ul style="list-style-type: none"> <li>• Dimensionality reduction techniques: principal component analysis and factor analysis</li> <li>• Unsupervised learning: distance-based and model-based clustering</li> <li>• Time series analysis, modeling and forecasting</li> <li>• Model validation and re-sampling</li> <li>• Quantitative textual data analysis</li> <li>• Applications with the software R and the programming language Python</li> </ul>
<b>Teaching format</b>	Frontal lectures and lab sessions.

<b>Learning outcomes</b>	<p>Knowledge and understanding of multivariate data and time series data and several techniques to analyze them and get information on the phenomena of interest.</p> <p>Applying knowledge and understanding of advanced quantitative methods to describe and analyze economic and social phenomena through statistical software.</p> <p>Making judgments on models and statistical tools useful for advanced data analysis.</p> <p>Communication skills in presenting in a correct and concise way methods and results of a statistical analysis.</p> <p>Learning skills of different statistical methods useful in data science.</p>
--------------------------	---

<b>Assessment</b>	<u>Attending and non-attending students</u>
-------------------	---

	Written exam and lab: written exam with review questions, theoretical questions, conducting experiments and evaluating results.
<b>Assessment language</b>	Italian
<b>Evaluation criteria and criteria for awarding marks</b>	<p><u>Attending and non-attending students</u>  100% written exam consisting of theoretical questions and data analysis tasks.  Criteria for written exam: correctness and clarity of answers, knowledge and understanding of statistical methods, ability to interpret outputs and to correctly use formal code.</p>
<b>Required readings</b>	<ul style="list-style-type: none"> <li>- Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, Introduzione all'apprendimento statistico. Con applicazioni in R, Piccin-Nuova Libreria, 2020, ISBN: 978-88-299-3094-4. (Chapters 5, 10)</li> <li>- Tommaso Di Fonzo, Francesco Lisi, Serie storiche economiche. Analisi statistiche e applicazioni, Carocci, 2015, Ed. VIII, ISBN: 978-88-430-3423-9. (Chapters 1, 5, 6-7)</li> <li>- Alan Agresti, Maria Kateri, Statistica per data scientists. Con R e Python, Egea, 2022, ISBN: 9788823823426.</li> <li>- Selection of papers provided by the lecturer.</li> <li>- Lecture notes and case studies to analyze.</li> </ul>
<b>Supplementary readings</b>	- Further readings will be announced during the course.

## Syllabus

### Descrizione del corso

<b>Titolo del corso</b>	<b>Data Science per le Scienze sociali</b>
<b>Codice del corso</b>	27276
<b>Settore scientifico disciplinare del corso</b>	SECS-S/01
<b>Corso di studio</b>	Corso di laurea in Economia, Politica ed Etica
<b>Semestre</b>	2° semestre, a.a. 2024/2025
<b>Anno del corso</b>	2
<b>Crediti formativi</b>	8
<b>Modulare</b>	No

<b>Numero totale di ore di lezione</b>	48 (42 Prof.ssa Di Lascio + 6 Prof.ssa Carlotti)
<b>Numero totale di ore di laboratorio</b>	--
<b>Numero totale di ore di esercitazioni</b>	12
<b>Frequenza</b>	Fortemente consigliata ma non obbligatoria
<b>Corsi propedeutici</b>	È altamente consigliato aver seguito e superato sia il corso di Matematica per EPE che quello di Statistica per EPE.
<b>Sito web del corso</b>	<a href="https://www.unibz.it/it/faculties/economics-management/bachelor-economics-politics-ethics/">https://www.unibz.it/it/faculties/economics-management/bachelor-economics-politics-ethics/</a>

<b>Obiettivi formativi specifici del corso</b>	<p>Il corso è un'attività affine all'area scientifica di Statistica-Matematica e copre i metodi statistici per la scienza dei dati nel contesto delle scienze sociali.</p> <p>Il corso si propone di fornire agli studenti i metodi avanzati di statistica dedicati all'analisi dei dati in economia e nelle scienze sociali. Gli aspetti teorici sono integrati dalla moderna analisi dei dati con R e Python.</p> <p>Al termine dello studio di questo corso, lo studente dovrebbe essere in grado di</p> <ul style="list-style-type: none"> <li>- ridurre la dimensionalità di un insieme di dati</li> <li>- indagare e capire strutture sottostanti i dati;</li> <li>- analizzare una serie temporale applicando modelli appropriati;</li> <li>- validare un modello statistico e fornire una misura più accurata della sua affidabilità;</li> <li>- analizzare e raccogliere dati testuali;</li> <li>- utilizzare i linguaggi di programmazione R e Python per eseguire analisi avanzate dei dati.</li> </ul>
--	--

<b>Docente</b>	F. Marta L. Di Lascio Office E510a e-mail: <a href="mailto:marta.dilascio@unibz.it">marta.dilascio@unibz.it</a>
----------------	---

	<p>Tel: 0471/013285  <a href="https://www.unibz.it/en/faculties/economics-management/academic-staff/person/32845-francesca-marta-lilja-di-lascio">https://www.unibz.it/en/faculties/economics-management/academic-staff/person/32845-francesca-marta-lilja-di-lascio</a></p> <p>Benedetta Carlotti  Office I2.01  E-mail <a href="mailto:benedetta.carlotti@unibz.it">benedetta.carlotti@unibz.it</a>  Lecturer's page  <a href="https://www.unibz.it/it/faculties/economics-management/academic-staff/person/45699-benedetta-carlotti">https://www.unibz.it/it/faculties/economics-management/academic-staff/person/45699-benedetta-carlotti</a></p>
<b>Settore scientifico disciplinare del docente</b>	SECS-S/01 (Prof.ssa Di Lascio) SPS/04 (Prof.ssa Carlotti)
<b>Lingua ufficiale del corso</b>	Italiano
<b>Orario di ricevimento</b>	21 ore Prof.ssa Di Lascio + 3 ore Prof.ssa Carlotti Cockpit – area studenti – calendario online Webpage: <a href="#">Timetables / Free University of Bozen-Bolzano (unibz.it)</a>
<b>Collaboratore didattico</b>	Somenzi Damiano Marino <a href="mailto:Damiano.Somenzi@unibz.it">Damiano.Somenzi@unibz.it</a>
<b>Orario di ricevimento</b>	Su appuntamento, inviando un'e-mail e consultando il calendario online del docente.
<b>Lista degli argomenti trattati</b>	<ul style="list-style-type: none"> <li>- Tecniche di riduzione della dimensionalità: analisi delle componenti principali e analisi dei fattori</li> <li>- Apprendimento non supervisionato: metodi di clustering basati sul concetto di distanza e metodi di clustering basati sui modelli</li> <li>- Analisi, modellizzazione e previsione delle serie temporali</li> <li>- Convalida dei modelli e ricampionamento</li> <li>- Analisi quantitativa dei dati testuali</li> <li>- Applicazioni con il software R e il linguaggio di programmazione Python</li> </ul>
<b>Attività didattiche previste</b>	Lezioni frontali e sessioni di laboratorio.

<b>Risultati di apprendimento attesi</b>	<p>Conoscenza e comprensione dei dati multivariati e delle serie temporali e delle varie tecniche per analizzarli e ottenere informazioni sui fenomeni di interesse.</p> <p>Capacità di applicare la conoscenza e la comprensione di metodi quantitativi avanzati per descrivere e analizzare fenomeni economici e sociali attraverso software statistici.</p> <p>Autonomia di giudizio su modelli e strumenti statistici utili per l'analisi avanzata dei dati.</p> <p>Abilità comunicative nel presentare in modo corretto e conciso i metodi e i risultati di un'analisi statistica.</p>
--	---

	Capacità di apprendimento di diversi metodi statistici utili nella scienza dei dati.
<b>Metodo d'esame</b>	<u>Studenti frequentanti e non frequentanti</u> Esame scritto e laboratorio: esame scritto con domande di verifica, domande teoriche, conduzione di esperimenti e valutazione dei risultati.
<b>Lingua dell'esame</b>	Italiano
<b>Criteri di misurazione e criteri di attribuzione del voto</b>	<u>Studenti frequentanti e non frequentanti</u> Esame scritto al 100% composto da domande teoriche e domande relative all'analisi dei dati. Criteri per l'esame scritto: correttezza e chiarezza delle risposte, conoscenza e comprensione dei metodi statistici, capacità di interpretare i risultati e di utilizzare correttamente il codice formale.
<b>Bibliografia fondamentale</b>	<ul style="list-style-type: none"> <li>- Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, Introduzione all'apprendimento statistico. Con applicazioni in R, Piccin-Nuova Libreria, 2020, ISBN: 978-88-299-3094-4. (Chapters 5, 10)</li> <li>- Tommaso Di Fonzo, Francesco Lisi, Serie storiche economiche. Analisi statistiche e applicazioni, Carocci, 2015, Ed. VIII, ISBN: 978-88-430-3423-9. (Chapters 1, 5, 6-7)</li> <li>- Alan Agresti, Maria Kateri, Statistica per data scientists. Con R e Python, Egea, 2022, ISBN: 9788823823426.</li> <li>- Selezione di papers forniti dal docente.</li> <li>- Verranno forniti appunti delle lezioni ed casi di studio da analizzare.</li> </ul>
<b>Bibliografia consigliata</b>	- Ulteriori letture saranno annunciate durante il corso.