

## Syllabus of an educational component of a degree programme

Name of unit conducting a component	<b>Doctoral School of Social Sciences</b>
Name of an educational component	<b>Introduction to Data Science</b>
Language of education	English
Goals of education	<p><b>Understanding of:</b></p> <ul style="list-style-type: none"> <li>- The role of science in business</li> <li>- Applications of scientific approach to solve real life problems</li> <li>- Skillset required to become Data Scientist</li> </ul> <p><b>Practical knowledge:</b></p> <ul style="list-style-type: none"> <li>- Ability to use advanced Data Science Platform, tool</li> <li>- Building simple Data Science project workflow</li> <li>- Gathering requirements, collecting data required to solve a problem</li> <li>- Visualizing results, drawing conclusions</li> </ul>
Learning outcomes of an educational component	Theoretical & Practical knowledge on Data Science Techniques and how to use them in real life data analysis and real-life business cases with Advanced Data Science Tools.
Verification methods and assessment criteria of learning outcomes obtained by students	Each student needs to complete both: tasks assigned to them during our meetings; comprehensive project build from a scratch (defining goals, exploring data, testing analytical model, building visuals, writing analytical report).
Type of an educational component (obligatory/optional)	Optional.
Year of study	2 <sup>nd</sup> , 3 <sup>rd</sup>
Semester	Summer
Name and surname of the coordinator of a component and/or person/s conducting a component	Ph.D. Krzysztof Tomanek
Name and surname of person/s conducting an examination or granting credit in the case when this spósis other person than conducting a component	Ph.D. Krzysztof Tomanek
Manner of completion	Distance learning - exercises via e-learning platform.  Face to Face meetings.

Preliminary and additional requirements	Obligatory – completion of: Statistics (basic level), Research techniques for Social Sciences  Preferable – completed: Text Mining, Data Visualization Methods
Type and number of hours of courses requiring direct participation of academic staff and students, if in a given component such courses are included	30
Number of ECTS credits assigned to a component	3
Balance of ECTS credits	3
Applied teaching methods	Interactive workshops with Data Science Tool installed on student's computers.  Conversations, problem solving, explorations through data analysis techniques, with mixed-method approached between qualitative and quantitative analytics.
Form and conditions of passing a component, including conditions of allowing to take an examination, as well as form and conditions of passing each type of courses included in a given component	Each student needs to complete Data Science Project that will be presented during the course final meeting.
Content of an educational module (with division into forms of courses completion)	<ol style="list-style-type: none"> <li>1. Data Science: key concepts and key terms, skills, areas of application, problems solving.</li> <li>2. Practical examples of Data Science approach to real life problems: business cases.</li> <li>3. Main concept for Text Mining Analytics: <ol style="list-style-type: none"> <li>a. Text cleansing techniques</li> <li>b. Text exploration &amp; classification</li> <li>c. Using dictionaries to classify texts</li> <li>d. Sentiment analytics</li> <li>e. Topic modeling</li> <li>f. Using AI for text analytics.</li> </ol> </li> <li>4. Data explorations techniques and what comes out of it? Bias, regularities, anomalies, all the rest.</li> <li>5. Why do we use classification methods and what are misclassifications? Decisions Trees in life cases.</li> <li>6. Why do we cluster? How are clusters used in practice?</li> <li>7. What is Machine Learning? Why is it in use?</li> <li>8. ML techniques, models. <ol style="list-style-type: none"> <li>a. Simple examples of ML usage in practice</li> </ol> </li> </ol>

	<ul style="list-style-type: none"> <li>b. Not so simple examples of ML usage</li> <li>c. Practical exercise with building ML model and analytical path.</li> </ul> <p>9. Visualize any data but how? Practical lessons on data visual techniques.</p> <p>10. Bad report &amp; good report.</p>
<p>List of basic as well as supplementary literature, knowledge of which is required in order to pass a given component</p>	<p>The Art of Data Science: Roger Peng, Elizabeth Matusi  <a href="http://bedford-computing.co.uk/learning/wp-content/uploads/2016/09/artofdatascience.pdf">http://bedford-computing.co.uk/learning/wp-content/uploads/2016/09/artofdatascience.pdf</a></p> <p>Numsense Data Science for Layman (no math) Annalyn Ng &amp; Kenneth Soo  <a href="https://www.amazon.com/dp/B01N29ZEM6?tag=algobbeans0e-20">https://www.amazon.com/dp/B01N29ZEM6?tag=algobbeans0e-20</a></p> <p>The elements of data analytics style - Jeff Leek  <a href="https://www.freetechbooks.com/the-elements-of-data-analytic-style-t1257.html">https://www.freetechbooks.com/the-elements-of-data-analytic-style-t1257.html</a></p> <p>Understanding Machine Learning Theory Algorithms - Shalev-Shwartz &amp; Ben-David  <a href="https://www.cs.huji.ac.il/w~shais/UnderstandingMachineLearning/understanding-machine-learning-theory-algorithms.pdf">https://www.cs.huji.ac.il/w~shais/UnderstandingMachineLearning/understanding-machine-learning-theory-algorithms.pdf</a></p> <p>The Big Book of Data Science Use Cases. The Big Book of Data Science Use Cases  <a href="#">The Big Book of Data Science Use Cases - Databricks</a></p> <p>Intelligent Techniques for Data Science, Rajendra Akerkar · Priti Srinivas Sajja. (1.2 History of Data Science)</p> <p>Mining of Massive Datasets  <a href="http://infolab.stanford.edu/~ullman/mmds/book.pdf">http://infolab.stanford.edu/~ullman/mmds/book.pdf</a></p> <p>DATA SCIENCE HANDBOOK. ADVICE AND INSIGHTS FROM 25 AMAZING DATA SCIENTISTS, Foreword by JAKE KLAMKA. DJ Patil, Hilary Mason, Pete Skomoroch, Riley Newman, Jonathan Goldman, Michael Hochster, George Roumeliotis, Kevin Novak, Jace Kohlmeier, Chris Moody, Erich Owens, Luis Sanchez, Eithon Cadag, Sean Gourley, Clare Corthell, Diane Wu, Joe Blitzstein, Josh Wills, Bradley Voytek, Michelangelo D'Agostino, Mike Dewar, Kunal Punera, William Chen, John Foreman, Drew Conway</p>