**Program: MBA**

**Specialization: Digital Transformation & Data Science**

**Syllabus for Module: International Immersion (Selected Topics in Digital Transformation & Data Science)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Date:**  July 7-8, 2025 | **Lecturer:**  Ted Kwartler | **ECTS:**  2,5 | **Language:**  English |

**Course Description**

|  |
| --- |
| This intensive seminar on Data Science and Business Analytics, targeted towards Executive MBA students at WU EA, will delve into the analysis of data and data-driven business decisions. The course supports learning a comprehensive data science workflow using R; commencing with problem understanding, continuing through the stages of obtaining data as well as cleaning & munging, exploratory data analysis (EDA), model development, and ultimately, model testing. By the end of this short-course, students will have a growing understanding of data used to create insights in management, thus bridging the gap between data science and strategic decision-making.  Many of the course materials and assignments are shared with three learning levels. There is a *basic*, *advanced* and *master* level to some of the material. This is meant to ensure students of varying ability are challenged. In terms of grading, no student will be at an advantage or disadvantage for choosing a more difficult path. Grades are earned by the effort, intellect, correctness, applied knowledge regardless of the difficulty path.  Choose the **basic** path if  1. you have never programmed in R before.  2. you expect to manage or direct analytical groups with the goal of effective communication and collaboration  Choose the **advanced** path if  1. you have some R programming experience, in a professional setting or if applicable have a solid grasp of the previous session material  2. your goal is to become a Data Scientist yourself  Choose the **master** path if  1.    you are proficient in R and have already used it for data analytics and/or machine learning purposes.  Course material will be on our class repository here:  <https://github.com/kwartler/Vienna_25> |

**Objectives and Learning Outcomes**

|  |
| --- |
| **By the end of the module students will have learned:**   * Introductory R syntax * Framing analytics and data science problems * Analytics workflow: Obtain, Scrub, Explore, Model, Interpret   **Knowledge and understanding**  After completing this module the students will have:   * Cleaned realistic data * Created visualizations and summary statistics to explore data * Fit a basic machine learning model * Management presentation of an analytical project   **Cognitive and subject specific skills**  After completing this module the students will have the skills:   * To accomplish basic tasks with R * Basic computer science knowledge such as logical operations, loops, and functions   **Key skills**  After completing this module the students will:   * Confidently frame analytical projects * Have basic coding skills to accomplish some of the methods needed to complete analytical projects |

**Course Structure and Teaching Methods**

|  |
| --- |
| Each morning, the course will be delivered with a series of code examples and explanations where students can learn R syntax and appropriate steps of a data science project. Students will then divide into groups to accomplish the case with support from the instructor in the afternoon. At the conclusion of each session, groups will be randomly selected to present their code, and a business presentation. |

**Transversal Themes and Current Discourse**

|  |
| --- |
| **Contemporary issues highlighted in this module:**   * Responsible AI – the appropriateness of protected classes in/out of data IE gender, race etc. * Emerging AI regulatory landscape – EU AIA, NIST, AI healthcare regulations |

**Reading Material**

|  |
| --- |
| **Required readings**   * Chapter 1 of the free book: <https://intro2r.com/chap1.html> * Chapter 2 of the free book: <https://intro2r.com/basics_r.html> * Chapter 3 of the free book: <https://intro2r.com/data_r.html> * Advanced & Masterclass: Chapter 5 of the free book: <https://intro2r.com/graphics_r.html> * Masterclass: Chapter on git & version control: https://intro2r.com/github\_r.html   **Additional readings**   * Ethics Related: PDF of IBM Watson Oncology [here](https://github.com/kwartler/Harvard_DataMining_Business_Student/blob/master/EthicsArticles/IBM%20pitched%20Watson%20as%20a%20revolution%20in%20cancer%20care.pdf) regarding AI in healthcare. * Regulation & Compliance Related:   + FDA [article](https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device) on AI in medical devices   + EU AIA Overview [Summary](https://artificialintelligenceact.eu/high-level-summary/)   **Recommended links and websites**   * Basic R syntax [cheatsheet](https://iqss.github.io/dss-workshops/R/Rintro/base-r-cheat-sheet.pdf) * Modeling “[cheatsheet](https://bcheggeseth.github.io/253_fall_2021/r-cheatsheet.html)” |

**Guidelines Regarding AI Tools**

|  |
| --- |
| AI tools, specifically large language models, ChatGPT and others, are encouraged and allowed for any assignment with proper citation. |

**Pre-Module**

|  |
| --- |
| **Assignment 1 (individual work):**  **Reading assignments/material**   * Basic & Advanced: Install R, R-studio & Jan AI (<https://github.com/isaka/JanAI>) * Masterclass: Install R, R-Studio, JanAI and install git and clone the class repo * Read Chapter1,2,3 from above * Review the FDA [article](https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device) on AI in medical devices * Review the EU AIA Overview [Summary](https://artificialintelligenceact.eu/high-level-summary/) * [JanAI Explanation](https://www.youtube.com/watch?v=QpMQgJL4AZA&ab_channel=DevelopersDigest) video * [R & Git Integration](https://www.youtube.com/watch?v=bUoN85QvC10&ab_channel=RiffomonasProject) Video   **Assignment 2 (individual work):**   * Write 1 paragraph summary of FDA article * Write 1 paragraph summary of EU AIA overview * Basic: Complete and submit the basic R code “[Intro\_to\_R\_Homework.R”](https://github.com/kwartler/Vienna_25/blob/main/Summer/Intro_to_R_Homework.R) from the class repository * Advanced: Using <http://universities.hipolabs.com/search?country=Austria> perform a GET request and create a geo-located map * Masterclass: Find any data related API, pull some data from it, perform EDA, partition it and build a regression or logistic regression model   **Deadline for pre-module assignment 2 and process of delivery**  June 30, 2025 at 9:00am (CET) – upload on Moodle  *All student work is checked for both plagiarism and the use of AI upon submission.*  *Please submit your assignment as a PDF file onto Moodle, indicate your names in the document and name your document “SURNAME / GROUP NAME\_PRE”.* |

**Core-Module**

|  |
| --- |
| **Reading assignments/material**  Chapter 5 & 9 from above  Masterclass: video for Jan AI explanation; git R studio integration |

**Distinguished Guest Speakers**

|  |
| --- |
| Faculty instructions: In line with our commitment to promoting diverse perspectives and enriching the learning experience, we encourage you to invite guest speakers to your class - physically and/or virtually. If you do, kindly share their names and schedules for seamless integration into your teaching plan. Our team is committed to accommodating your needs.  Name:Jared Bowns, https://www.linkedin.com/in/jbowns/  Expertise and background: Head of Data & AI Practice at Elyxor  Topic: Advice from the front lines of Business Data & AI  Date and time: TBD (working on aligning schedules) |

**Detailed Schedule**

**Day 1**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | 08:30 – 10:30 | Environment set-up, R, R studio, Git (masterclass only),  Jan AI  (masterclass only)  Some R Coding examples | | 10:30 – 10:45 | *Coffee break* | | 10:45 – 12:15 | Framing an analytics project  Obtain data, clean/scrub data for modeling | | 12:15 – 13:15 | *Lunch time* | | 13:15 – 13:45 | Explore data – summary statistics, visualizations | | 13:45 – 14:30 | Sales Challenge Data Group EDA & Visuals Work | | 14:30 – 15:00 | Each group will make a business presentation from their EDA *remaining time some groups will be asked to make a presentation to the class* | | 15:00 – 15:15 | 2 groups will be selected to present their findings | | 15:15 – 15:30 | *Coffee break* | | 15:30 – 17:00 | Start case data performing Obtain, data cleaning and exploration steps to  create a modeling matrix. | |

**Day 2**

|  |  |
| --- | --- |
| 08:30 – 10:30 | Two modeling methods, and evaluation |
| 10:30 – 10:45 | *Coffee break* |
| 10:45 – 12:15 | Improving model performance examples (ensemble, more observations, more features, feature engineering) |
| 12:15 – 13:15 | *Lunch time* |
| 13:15 – 15:15 | Group work to model, evaluate, interpret results then create slides for the case  Masterclass: Create a git repo for your group work  Masterclass: Using JanAI to extract data from documents for use in modeling |
| 15:15 – 15:30 | *Coffee break* |
| 15:30 – 17:00 | Group presentations including code step through & slides explaining the modeling steps, results and implications |

**Post-Module**

|  |
| --- |
| **Assignment (individual work):**  Please follow the Case Study instructions in the class repository related to EDA and modeling of Diabetes data. In the folder there is a CSV for data, and three scripts, ranging from Basic, Advanced and Expert. Please follow the instructions for your track and submit the reuiqred files.  **Deadline for post-module assignment and process of delivery**  August 4, 2025 @ 9:00am (CET) – upload on Moodle  *All student work is checked for both plagiarism and the use of AI upon submission.*  *Please submit your assignment as a PDF file onto Moodle, indicate your names in the document and name your document “SURNAME / GROUP NAME\_POST”.* |

**Assessment**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | **Pre-module:** |  | 30 % Intro to R Homework | | **Core-module:** |  | 35 % Group Work: Diabetes Data  Obtain Scrub Explore Model Interpret & slides | | **Post-module:** |  | 30 % Individual work: New data set script:  Obtain Scrub Explore Model Interpret [no slides] | | **Peer assessment:** |  | 5 % | |  |  | =100% | |

**Instructor and Contact Information**

|  |
| --- |
| **Ted Kwartler**  Organization: Accenture and Harvard Extension School  Position: Managing Director (North America Responsible AI Lead) and Adjunct Professor  E-mail: edwardkwartler@fas.harvard.edu |

**Short CV of Instructor and Contact Information**

|  |
| --- |
| Ted Kwartler is a distinguished data science and AI professional with roles in both academia and the corporate sector. He is an Adjunct Lecturer at Harvard University Extension School and Global School in Empirical Research Methods at St. Gallen University. After earning his Master’s of Business Administration from the University of Notre Dame, he served as an operations leader at Amazon.com, an AVP of Data Science Innovation at Liberty Mutual Insurance, followed by roles as the VP of Trusted AI and Chief Technology Officer for Generative AI at DataRobot. He's currently the Managing Director at Accenture charged with leading the North American Responsible AI practice.  In addition to his corporate successes, Kwartler is recognized for his robust contributions to academia, developing innovative courses encompassing machine learning, natural language processing, and data-driven business decisions at renowned universities. His impact also extends to DataCamp.com, where his courses were taken by more than 32,000 in subjects including Text Mining, Sentiment Analysis, Statistics and Data-Driven Business Decisions. As a published author, Kwartler has penned important works on text mining, sports business analytics, and customer insight. In 2022, he was awarded Hult University's Faculty of the Year, and severed on the editorial board of the Applied Marketing Analytics Journal from 2017 to 2024. |