



BUSINESS CASES WITH DATA SCIENCE

SYLLABUS
2021-2021



<p>INSTRUCTOR INFORMATION</p>	<p>FERNANDO LUCAS BAÇÃO 2º floor, room 10 Tel: 21 3870413 (ext. 222) bacao@novaims.unl.pt http://www.novaims.unl.pt/fbacao</p> <p>JOÃO PEDRO FONSECA jpfonseca@novaims.unl.pt</p> <p>HUGO SAISSE MENTZINGEN DA SILVA hsilva@novaims.unl.pt</p>
<p>SCHEDULE</p>	<p>Sessions – Tuesdays and Wednesdays from 14h00 – 17h00</p>
<p>OFFICE HOURS:</p>	<p>Mondays from 12h00 – 14h00 (schedule appointment by email), 2nd Floor, Room 10</p>
<p>CONTACT</p>	<p>Email: bacao@novaims.unl.pt; jpfonseca@novaims.unl.pt; hsilva@novaims.unl.pt. Moodle: announcements; Business Case X Forum, forum on Moodle for each business case;</p>
<p>DESCRIPTION AND OBJECTIVES</p>	<p>Using a case-based learning approach, the Business Cases with Data Science course addresses the ways in which enterprises such as businesses, non-profits, and governments can use data to gain insights, improve the decision-making process and leverage the informational resources available in operations, marketing, finance, and strategic planning among other functions. The students will use the knowledge and skills they developed during the courses of the first semester to come up with relevant and intelligent solutions to real world business problems, through the use of analytical models. During the course the students will have the opportunity to use different Python-based analytical tools, appropriate to the different business problems proposed.</p>

	<p>The fundamental objective of the course is to help the students bridge the gap between understanding the analytic tools and being able to apply them appropriately in a specific business context. Through hands-on projects the students will be exposed to a number of real-world business problems, where they should be able to provide relevant analytical solutions. Additionally, the students will also develop communication and teamwork skills, which are critical for the course. Additionally, the course will have a large number of guests, mostly alumni of the master, which will share their experience in the job market and on how to develop a career in data science.</p>
LEARNING OUTCOMES	<ul style="list-style-type: none"> • Understand what a business case is and why to use it; • Identify the typical components of a business case; • Model business cases in accordance to the CRISP-DM process model; • Identify and implement the most adequate analytical models to different business cases; • Interpret model results from both a data science and a business perspective; • Make data-driven decisions to optimize business processes; • Improve communication skills, both oral and written; • Improve teamwork skills.
COURSE SUCCESS	<p>In this course success depends on a number of factors:</p> <ul style="list-style-type: none"> • Deep understanding of the topics learnt on the courses of machine learning and data mining; • Knowledge of the analytic tools; • Class attendance; • Work during the semester on the assigned projects and deliver them on time; • Read the suggested references.
ORGANIZATION (GROUPS)	<p>Most of this course is based on teamwork, thus it is of crucial importance that the students define their groups before the first business case is presented. It is also important to choose wisely the group members as once the group is submitted there will be no opportunities to change the group composition. The guidelines for the groups are:</p> <ul style="list-style-type: none"> • Students must organize themselves into groups of

	<p>4 students (give a name to your group);</p> <ul style="list-style-type: none"> • Students should assume the role of consultants providing a service to the company; • The instructors assume the role of the company project stakeholder and business expert; • Students must submit the groups until the 22nd of February otherwise they will be randomly allocated.
<p>PROJECTS DELIVERABLES</p>	<ul style="list-style-type: none"> • Source Code: <ul style="list-style-type: none"> ○ Python (Jupyter Notebook and/or Python scripts); ○ Code should be properly commented/documented to facilitate comprehension of what it is intended to do; • Report: <ul style="list-style-type: none"> ○ Should consider the following topics, but should not replicate what is on source code: <ul style="list-style-type: none"> ▪ Problems and their solutions according to CRISP-DM phases; ▪ Interpret results and their implications to business, including data-driven decisions to optimize the business processes; ▪ Theoretical considerations about deployment and maintenance plans; ▪ Considerations for future model improvement; ▪ Should be written in the provided template and not exceed 5 pages, excluding references and any appendixes; • Presentation: <ul style="list-style-type: none"> ○ PowerPoint, Prezi, or any other presentation tool; ○ The presentation should be designed as a presentation to be delivered to the company board of directors to obtain the “go ahead” for deployment; ○ Should not exceed 10 minutes, with the exception of the first business case in which the presentation should not take longer than 5 minutes. It is important to highlight that there is a penalty for the presentations that exceed the time limit; • All deliverables should be submitted through Moodle

	until 23h59 of the day before the business case presentation date
STRUCTURE OF THE CLASSES	<p>The content of the course is composed of 5 business cases. The typical rotation of cases, in the first business case, will be:</p> <p>Week 1:</p> <ul style="list-style-type: none"> • the instructors or a company presents the business case to the students, explaining the business context and objectives of the company; • practical session <p>Week 2:</p> <ul style="list-style-type: none"> • the students pitch their work (5 minutes), explaining what they did, the models used, and the results achieved in the project. <p>In business cases 2, 3, 4 and 5 there is an additional week of work and two presentation sessions:</p> <p>Week 1:</p> <ul style="list-style-type: none"> • the instructors or a company presents the business case to the students, explaining the business context and objectives of the company; • practical session <p>Week 2:</p> <ul style="list-style-type: none"> • this session will be used to work and, eventually, consult with the instructors about questions related with the project. This week is meant to alleviate the student's stress and give additional time to complete the project; • practical session <p>Week 3:</p> <ul style="list-style-type: none"> • the students pitch their work, explaining what they did, the models used, and the results achieved in the project. <p>The business cases are:</p> <p style="text-align: center;">BC 1. Segmentation - Online Wine Store BC 2. Classification - Hotel</p>

	BC 3. Recommendation System - Delivery App BC 4. Forecasting – Supermarket BC 5. Crypto Dashboard	
CONTENTS	<ol style="list-style-type: none"> 1. CP1. Introduction to Business Cases with Data Science course; 2. CP2. Introduction to Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology; 3. CP3. Example of a business case and how to solve it; 4. CP4. Online Wine Store – who are my customers? 5. CP5. Hotel – are cancellations hurting my business? 6. CP6. Delivery App – can I optimize the app? 7. CP7. Supermarket – how many oranges am I going to sell? 8. CP8. Crypto Dashboard – exploring the crypto trading universe 	
BIBLIOGRAPHY	References: <ul style="list-style-type: none"> □ Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0: Step-by-step data mining guide. Retrieved from https://the-modeling-agency.com/crisp-dm.pdf □ Provost, F., and Fawcett, T. (2013). Data Science for Business . Sebastopol, CA: O'Reilly □ Guidici, P., and Figini, S. (2009). Applied Data Mining for Business and Industry . UK: Wiley □ Shmueli, G., Bruce, P. C., Gedeck, P., Patel, N. R. (2019). Data Mining for Business Analytics: Concepts, Techniques, and Applications in Python. Hoboken, NJ: Wiley 	
EVALUATION	75% Group projects (BC 1: 10%; BC 2-4: 15%; BC 5: 20%) 25% Final exam (individual open-book)	
CALENDAR	15/02	Course overview; Business cases; CRISP-DM; Deliverables
	16/02	Practical Sessions (4x)
CASE PRESENTATION 1	22/02	Looking for a data scientist Guest <ul style="list-style-type: none"> • Francesco Costigliola (Business

	<p>Analytics & Data Science Manager @EDP Comercial)</p> <ul style="list-style-type: none"> • Sandra Catarino (Business Analytics & Data Science Manager @Novo Banco) • David Morais (Head of Analytics (Data and Models Department) @Cofidis) <p>Case 1: Segmentation</p>
	23/02 Practical Sessions (4x)
	01/03 Carnival Break
	02/02 Case 1: presentations and discussions (5 min)
CASE PRESENTATION 2	08/03 <p>Guests:</p> <ul style="list-style-type: none"> • Ivo Bernardo (Partner & Data Scientist @DareData Engineering Udemy Bestseller Instructor and Teacher) – A career in Data Science: building, teaching and writing; • Monika Brown (Data Science Lead @Aprawa) - Working remotely as a data scientist <p>Case 2: Classification</p>
	09/03 Practical Sessions (4x)
	15/03 <p>Guests:</p> <ul style="list-style-type: none"> • Georgios Douzas (Machine Learning Researcher @NOVA IMS) – Machine learning models: from experiments to production • Max Maukner (Data Scientist Analyst Engineer @BMW Group) – leveraging analytics in automotive industry supply chain <p>Case 2: general support</p>
	16/03 Practical Sessions (4x)
	22/03 Case 2: presentations and discussions
	23/03 Case 2: presentations and discussions
CASE PRESENTATION 3	29/03 <p>Guests:</p> <ul style="list-style-type: none"> • Paulo Lapa (Data Scientist @Talkdesk TDX) – working at a

	<p>Portuguese unicorn</p> <ul style="list-style-type: none"> • Rene Rauch (Siri - AI/ML @Apple) – working at Apple as a machine learning researcher <p>Case 3: Recommender System</p>
30/03	Practical Sessions (4x)
12/04	<p>Guests:</p> <ul style="list-style-type: none"> • Vitor Manita (Machine Learning Engineer @FARFETCH) - Managing consultants in a Data Science Project • Jan-Benedikt Jagusch (Data Engineer @QuantCo) – Consulting in Data Science <p>Case 3: general support</p>
13/03	Practical Sessions (4x)
19/04	Case 3: presentations and discussions
20/04	Case 3: presentations and discussions
CASE PRESENTATION 4	26/04
	<p>Guests:</p> <ul style="list-style-type: none"> • Pedro Fonseca, Bruno Alho, Mariana Martins, Miguel Zina - How to win a data challenge <p>Case 4: Forecasting</p>
27/04	Practical Sessions (4x)
03/05	<p>Guests:</p> <ul style="list-style-type: none"> • tba • tba <p>Case 4: general support</p>
04/05	Practical Sessions (4x)
10/05	Case 4: presentations and discussions
11/05	Case 4: presentations and discussions
CASE PRESENTATION 5	17/05
	<p>Guests</p> <ul style="list-style-type: none"> • tba • tba <p>Case 5: Dashboard</p>
18/05	Practical Sessions (4x)
24/05	<p>Guests:</p> <ul style="list-style-type: none"> • tba

	<ul style="list-style-type: none">• tba Case 5: general support
25/05	Practical Sessions
31/05	Case 5: presentations and discussions
01/06	Case 5: presentations and discussions