

FICHA TÉCNICA DE LA ASIGNATURA

Datos de la asignatura	
Nombre completo	Optativas complementarias: Advanced Data Analytics for Business
Código	E000011607
Impartido en	Máster Universitario en Administración de Empresas (MBA) [Primer Curso]
Nivel	Postgrado Oficial Master
Cuatrimestre	Semestral
Créditos	3,0 ECTS
Carácter	Optativa
Departamento / Área	Departamento de Métodos Cuantitativos
Responsable	Alejandro Rodríguez Gallego
Horario de tutorías	Upon request

Datos del profesorado	
Profesor	
Nombre	Alejandro Rodríguez Gallego
Departamento / Área	Departamento de Gestión Financiera
Despacho	Alberto Aguilera 23 shared office space, 4th floor, east side.
Correo electrónico	argallego@comillas.edu

DATOS ESPECÍFICOS DE LA ASIGNATURA

Contextualización de la asignatura
<p>Aportación al perfil profesional de la titulación</p> <p>This course introduces the participants to the rapidly evolving Data Science ecosystem and its multiple applications in numerous business settings. As data is reshaping the business environment, executives must identify, evaluate, and face new challenges. These unprecedented opportunities led to the emergence of new products and business models and the revision of traditional ones.</p> <p>Students will acquire a deeper understanding of the core concepts and technologies used to extract information from data (structured or otherwise) by delving into their characteristics, benefits, and disadvantages.</p> <p>This course also bridges the gap between the managerial and technical standpoints by offering a hands-on approach and exposure to technical matters. Upon conclusion, participants will be ready to propose business solutions supported by data and communicate efficiently with technical staff.</p> <p>Course objectives</p> <ul style="list-style-type: none"> Understand the importance and value of data.



- Understand the people, processes, and technologies involved in extracting knowledge from data.
- Propose data-driven solutions.
- Manage and communicate efficiently with technical teams.
- Demystify technical topics.

Prerequisitos

- Attending the course "Data and Information" is recommended.
- Even though no programming skills are required, a predisposition towards technical matters is expected due to the hands-on nature of the subject.

Competencias - Objetivos

Competencias

GENERAL

CG01	Capacidades cognitivas de análisis y síntesis aplicadas a situaciones de negocios y problemáticas organizativas de gestión.	
	RA01	Describe, relaciona e interpreta situaciones y planteamiento de nivel medio.
	RA02	Selecciona los elementos más significativos y sus relaciones en las situaciones planteadas.
	RA03	Identifica las carencias de información y establece relaciones con elementos externos a la situación planteada.
	RA04	Es capaz de resumir y estructurar la información empleando los conceptos adecuados.

CG02	Gestión de la información y de datos como elementos clave para la toma de decisiones y la identificación, formulación y resolución de problemas empresariales.	
	RA01	Busca, conoce, sintetiza y utiliza adecuadamente datos primarios y secundarios procedentes de diversas fuentes.
	RA02	Conoce y usa Internet para buscar y manejar información, textos y datos.
	RA03	Discierne el valor y la utilidad de diferentes fuentes y tipos de información.

CG03	Resolución de problemas y toma de decisiones en los niveles estratégico, táctico y operativo de una organización empresarial, teniendo en cuenta la interrelación entre las diferentes áreas funcionales y de negocio.	
	RA01	Identifica y define adecuadamente el problema y sus posibles causas.
	RA02	Plantea posibles soluciones pertinentes y diseña un plan de acción para su aplicación.
	RA03	Identifica problemas antes de que su efecto se haga evidente.
	RA04	Dispone de la capacidad para tomar decisiones de una forma autónoma.



CG03	Resolución de problemas y toma de decisiones en los niveles estratégico, táctico y operativo de una organización empresarial, teniendo en cuenta la interrelación entre las diferentes áreas funcionales y de negocio.	
	RA05	Reconoce y busca alternativas a las dificultades de decisión en situaciones reales.
	RA06	Es capaz de ponderar diferentes factores (económicos, sociales y técnicos, entre otros) en el proceso de toma de decisiones y determinar su impacto.

CG09	Capacidad de aprendizaje autónomo para seguir formándose para aprender a aprender las habilidades cognitivas y los conocimientos relevantes aplicados a la actividad profesional y empresarial.	
	RA01	Realiza sus trabajos y su actividad necesitando sólo unas indicaciones iniciales y un seguimiento básico.
	RA02	Busca y encuentra recursos adecuados para sostener sus actuaciones y realizar sus trabajos.
	RA03	Amplía y profundiza en la realización de sus trabajos.

SPECIFIC

CE09	Ser capaz de analizar problemas de la empresa y su entorno mediante el conocimiento de los datos y de la información; su naturaleza, recolección, almacenaje, modelización y extracción y mediante el uso de métodos cuantitativos, distinguir técnicas de análisis y de modelización apropiadas y aplicar las mismas a casos prácticos de predicción y simulación en la gestión empresarial.	
	RA01	Aplica la abstracción y la simplificación para modelar en términos matemáticos el problema al que se enfrenta.
	RA02	Conoce los instrumentos matemáticos necesarios para la modelización.
	RA03	Integra el análisis gráfico, verbal y los datos cuantitativos y cualitativos al estudio de cuestiones decisionales relacionadas con el problema.
	RA04	Es capaz de interpretar los resultados obtenidos y usarlos para tomar decisiones de manera razonada.

BLOQUES TEMÁTICOS Y CONTENIDOS

Contenidos – Bloques Temáticos

1. Software development (for Data Analysis)

- Introduction
- Project management: agile, waterfall, CRISP-DM.
- Dev Ops. Git.
- Data life cycle: gathering, identifying, storing, analysis, visualization, actionable results.
- New roles & leadership: data scientist, ML engineer, Chief Data Officer...

2. [Workshop: Introduction to Python](#)

3. Data sources

- File formats
- Web services



- Databases
 - Types: SQL, no-SQL, in-memory.
 - Distributed databases. CAP theorem.
 - Normalization.
 - ERD diagram.
 - OLAP cube. Star or snowflake schema. Fact table. Dimension tables. Slice, dice, drill up/down, pivot.
 - Backup strategies.
 - Data warehouse vs. Data Lake.
 - ETL, ELT.
- Web scraping.
- Data encryption: at rest, in transit.
- Data anonymization: differential privacy.

4. **Workshop: Data manipulation in Python.**

5. **Workshop: data sources.**

6. **Workshop: SQL.**

7. **The Cloud**

- On-premise, private cloud, public cloud.
- IaaS, PaaS, SaaS, PaaS
- Vendors (AWS, Google, Microsoft, Oracle...).
- Virtualization

8. **Machine learning**

- Learning types: supervised, unsupervised, reinforced, deep.
- Tools: preprocessing, dimensionality reduction, model selection, classification, regression, clustering, others.
- Generalization: Underfitting/overfitting.
- Training/test split. Cross validation. Regularization.
- Metrics: classification (confusion matrix), regression, similarity.
- Ensembles: bias-variance trade-off. Boosting, bagging.
- Natural Language Processing.

9. **Workshop: Orange data mining.**

10. **Workshop: Machine Learning in Python.**

11. **Workshop: Visualization in Python.**

METODOLOGÍA DOCENTE

Aspectos metodológicos generales de la asignatura

Metodología Presencial: Actividades

- Teaching lectures to introduce the basic of each topic.
- Development of a model example by the teacher.
- Guided practice of cases applying the concepts learnt.
- Oral presentation of the applied practical cases done in group.

Metodología No presencial: Actividades

- Tutored personal work.
- Individual and group practice work.

EVALUACIÓN Y CRITERIOS DE CALIFICACIÓN

Evaluation Criteria	0 - 10	Evaluation Criteria Weight
Attendance & active participation in class	0 - 10	20%
Public oral presentation	0 - 10	20%
Cases & exercises	0 - 10	50%
Self & group assessment	0 - 10	10%

Calificaciones

To pass the course, the student should obtain at least the following:

- Attendance requirements: 80%
- Cases and exercises: 5/10
- Oral presentation: 5/10

Total grade weighting all components equal or superior to 5/10.

Those students who do not pass the subject will be able to submit the cases and exercises and repeat the oral presentation.

Students with an attendance waiver

In order to pass the module, these students will have to submit the cases and exercises and repeat the oral presentation. Total grade weighting all components equal or superior to 5/10.

BIBLIOGRAFÍA Y RECURSOS

Bibliografía Básica

- Class notes
- Handouts & slides
- External resources

Bibliografía Complementaria

BUSINESS & DATA SCIENCE

- Provost, F., Fawcett, T. (2013) *Data Science for Business: What you need to know about data mining and data-analytic thinking*. Ed. O'Reilly Media. ISBN 9781449361327
- Shmueli et al. (2017) *Data Mining for Business Analytics: Concepts, Techniques, and Applications in R*. Ed. Wiley.
- Wisniewsky, M., Shafti, F. (2020) *Quantitative Analysis for Decision Makers, 7th Edition* (formerly known as Quantitative Methods for Decision Makers), 7th Edition. Ed. Pearson. ISBN: 9781292276663



COMILLAS

UNIVERSIDAD PONTIFICIA

ICAI

ICADE

CIHS

GUÍA DOCENTE

2021 - 2022

BUSINESS & BIG DATA

- Marr, B. B. (2015) *Big Data: Using Smart Big Data, Analytics and Metrics to Make Better Decisions and Improve Performance*. Ed. John Wiley & Sons. ISBN 9781118965832.
- Marr, B. B. (2016) *Big Data: how 45 successful companies used Big Data analytics to deliver extraordinary results*. Ed Willey. ISBN 9781119231387.

TECHNICAL

- Bishop, C. (2009) *Pattern Recognition and Machine Learning*. Ed. Springer. ISBN 9781493938438.
- EMC Education Services (2015) *Data Science & Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data*. Ed. Wiley (2015). ISBN:9781118876138
- James, Gareth, et al. (2013) *An introduction to statistical learning*. Vol. 112. New York. Ed. Springer.
- Murphy, K. (2012) *Machine Learning, a probabilistic perspective*. The MIT Press. ISBN 9780262018029.