

FICHA TÉCNICA DE LA ASIGNATURA

Datos de la asignatura	
Subject name	Introduction to Statistical Computing
Subject code	DTC-BA-514
Main program	Grado en Análisis de Negocios/Business Analytics
Involved programs	Grado en Ingeniería en Tecnologías de Telecom. y Grado en Análisis de Negocios/Business Analytics [Fifth year]
Quarter	Semestral
Credits	3,0 ECTS
Type	Obligatoria (Grado)
Department	Department of Telematics and Computer Sciences
Coordinator	Pizarroso Gonzalo, Jaime
Office hours	Enviar un correo para concertar cita.

Datos del profesorado	
Teacher	
Name	Jaime Pizarroso Gonzalo
Department	Department of Telematics and Computer Sciences
Office	D-507
E-Mail	jpizarroso@comillas.edu

DATOS ESPECÍFICOS DE LA ASIGNATURA

Contextualización de la asignatura
<p>Aportación al perfil profesional de la titulación</p> <p>This course aims to introduce students to statistical and probabilistic computing technologies available in the literature. Through a thorough review of different processing devices and their operations, students will acquire the skills to discern specific computing needs for a given problem, considering factors such as computational cost, execution time, and the suitability of the technology to be used. Subsequently, they will gain knowledge of programming best practices and code verification required in industrial software development. Finally, a review of the main descriptive and inferential statistical methods used for data analysis will provide students with the skills to analyze and extract information from large volumes of data.</p> <p>The course is divided into four blocks or topics. In the first block, students will study different electronic computing devices, namely, CPUs, GPUs, and TPUs. They will learn to differentiate them and understand their specific operation, enabling them to recommend their use based on the problem at hand. In the second block, the central theme is distributed programming. Students will study different programming paradigms and common software architectures, along with their main advantages and disadvantages. From a practical perspective, they will apply this knowledge and explore open-source literature tools and their applications to classic computing problems. In the third block, students will examine different code verification strategies, understanding the importance of systematically checking code functionality and industry requirements for code quality. Finally, in the fourth block, students will delve into probabilistic programming, process simulation, and uncertainty management in modeling and prediction. In this section, students will be able to apply</p>



uncertainty and confidence intervals to prediction models and understand how uncertainty is transmitted between processes.

Prerrequisitos

- Linear Algebra
- Basic programming in R/Python languages
- Understanding of Machine Learning and mathematical modeling
- Knowledge of descriptive statistics and probability theory

Competencias - Objetivos

Competencias

GENERALES

CG01	Capacidad de organización y planificación en la identificación de problemas en el contexto de datos masivos	
	RA1	Identificar y organizar temporalmente las tareas necesarias para la realización sus actividades de aprendizaje, cumpliendo con los plazos establecidos de entrega de actividades
	RA2	Identificar las fuentes de información más relevantes para cada problema
	RA3	Utilizar las herramientas digitales idóneas para la adquisición y gestión de la información y ser capaz de integrar esa información de una manera racional
CG02	Capacidad de análisis de datos masivos procedentes de diversas fuentes: texto, audio, numérica e imagen	
	RA1	Describir, relacionar e interpretar situaciones y planteamientos sencillos
	RA2	Seleccionar las fuentes de información digital más significativas
	RA3	Identificar las carencias de información de la situación planteada y establecer sus relaciones con elementos externos
CG03	Resolución de problemas y toma de decisiones en un entorno de datos masivos tanto cuantitativos como cualitativos	
	RA1	Resolver problemas y realizar trabajos aplicando una diversidad metodológica de análisis
	RA2	Identificar correctamente los conocimientos aplicables a cada situación
	RA3	Identificar la metodología más apropiada para la resolución de los problemas planteados, conociendo las herramientas software más relevantes y aportando una visión crítica.
CG07	Capacidad de liderazgo y trabajo en equipo, en la sociedad de la información	
	RA1	Participar de forma activa en el trabajo de grupo compartiendo información, conocimientos y experiencias
	RA2	Orientarse a la consecución de acuerdos y objetivos comunes
	RA3	Ser capaz de organizar y/o coordinar equipos de trabajo



CG08	Capacidad crítica y autocrítica en la sociedad de la información	
	RA1	Identificar los supuestos y las limitaciones de los modelos y metodologías a aplicar en cada situación
	RA2	Ejercer una valoración crítica de los modelos y metodologías a aplicar en cada situación
CG09	Compromiso ético en la sociedad de la información	
	RA1	Ser honesto en el desarrollo de las actividades académicas y en otros aspectos de la vida y no ser pasivo ante la deshonestidad de otros
	RA2	Conocer y desarrollar las bases de la ética profesional
CG11	Capacidad para aprender y trabajar autónomamente en la sociedad de la información	
	RA1	Organizar los materiales de la asignatura con el fin de articular su aprendizaje
	RA2	Poseer criterio para decidir la utilización de los medios humanos y otros recursos puestos a su disposición para el aprendizaje.
	RA3	Realizar sus trabajos y su actividad necesitando sólo unas indicaciones iniciales y un seguimiento básico
ESPECÍFICAS		
CE28	Conocer y comprender los paradigmas de programación más importantes y los fundamentos de la programación que permitan al alumno elaborar códigos sencillos orientados a la manipulación de información	
	RA1	Conocer y comprender las distintas aproximaciones a la programación y los lenguajes más utilizados en el contexto del análisis de información
	RA2	Ser capaz de elaborar programas sencillos con elementos como bucles o condicionales.
	RA3	Conocer y comprender el concepto de función y las capacidades de entrada y salida más comunes de los distintos lenguajes de programación.
CEO32	Capacidad de seleccionar y comprender las técnicas más adecuadas para el análisis crítico de problemas complejos de inteligencia y seguridad y otras cuestiones relevantes de la agenda internacional actual	
	RA1	Conoce los paquetes y lenguajes de programación estadísticos más utilizados y, en particular, el lenguaje de programación R
	RA2	Es capaz de manipular conjuntos de datos e importarlos y exportarlos en diferentes formatos. Asimismo, es capaz de obtener representaciones parciales de la información, generar resúmenes de los datos y presentarlos gráficamente
	RA3	Es capaz de programar e interpretar las medidas estadísticas más relevantes
CEO32	Saber manipular y analizar datos utilizando lenguajes de programación específicos para el análisis estadístico.	



	RA01	Conoce los paquetes y lenguajes de programación estadísticos más utilizados y, en particular, el lenguaje de programación R.
	RA02	Es capaz de manipular conjuntos de datos e importarlos y exportarlos en diferentes formatos. Asimismo, es capaz de obtener representaciones parciales de la información, generar resúmenes de los datos y presentarlos gráficamente.
	RA03	Es capaz de programar e interpretar las medidas estadísticas más relevantes.
CEO32	Saber manipular y analizar datos utilizando lenguajes de programación específicos para el análisis estadístico.	
	RA01	Conoce los paquetes y lenguajes de programación estadísticos más utilizados y, en particular, el lenguaje de programación R.
	RA02	Es capaz de manipular conjuntos de datos e importarlos y exportarlos en diferentes formatos. Asimismo, es capaz de obtener representaciones parciales de la información, generar resúmenes de los datos y presentarlos gráficamente.
	RA03	Es capaz de programar e interpretar las medidas estadísticas más relevantes.
CEO33	Conocer y usar técnicas de análisis de información para extraer el conocimiento para la toma de decisiones de negocio.	
	RA1	Conoce y sabe programar funciones elementales en VB para la manipulación de datos, rangos y celdas
	RA2	Comprende y es capaz de crear elementos funcionales en VB (botones, despletables, ¿) y su integración en hojas de cálculo Excel
CEO33	Conocer y comprender el concepto de macro Excel y su programación elemental en Visual Basic (VB	
	RA01	Conoce y sabe programar funciones elementales en VB para la manipulación de datos, rangos y celdas.
	RA02	Comprende y es capaz de crear elementos funcionales en VB (botones, despletables, ¿) y su integración en hojas de cálculo Excel.
CEO34	Entender la computación y las bases de datos distribuidas, con énfasis en la arquitectura Hadoop y las herramientas actuales más extendidas en el contexto de Big Data.	
	RA1	Conoce y aplica las técnicas de análisis y diseño orientado a objetos mediante el lenguaje de modelado UML
	RA2	Diseña programas basados en relaciones de herencia, sabiendo diferenciar entre clases, clases abstractas e interfaces
	RA3	Programa aplicaciones gráficas sencillas e interactivas
CEO34	Conocer y aplicar las técnicas de análisis y diseño orientado a objetos	



	RA01	Conoce y aplica las técnicas de análisis y diseño orientado a objetos mediante el lenguaje de modelado UML.
	RA02	Diseña programas basados en relaciones de herencia, sabiendo diferenciar entre clases, clases abstractas e interfaces.
	RA03	Programa aplicaciones gráficas sencillas e interactivas.
CEO35	Comprender y ser capaz de desarrollar herramientas de visualización avanzadas.	
	RA1	Conoce y comprende la problemática de la generación de números aleatorios y su utilización actual para la determinación de intervalos de confianza
	RA2	Conoce y comprende los métodos de optimización e inferencia y es capaz de resolver ese tipo problemas utilizando software estadístico
	RA3	Es capaz de generar sus propias librerías de funciones.
CEO35	Conocer y comprender herramientas avanzadas de análisis estadístico para la optimización y predicción.	
	RA01	Conoce y comprende la problemática de la generación de números aleatorios y su utilización actual para la determinación de intervalos de confianza.
	RA02	Conoce y comprende los métodos de optimización e inferencia y es capaz de resolver ese tipo problemas utilizando software estadístico.
	RA03	Es capaz de generar sus propias librerías de funciones.

BLOQUES TEMÁTICOS Y CONTENIDOS

Contenidos – Bloques Temáticos

Unit 0. Introduction to R programming
Unit 1. Statistical Analysis
Unit 2. Efficient Programming
Unit 3. Parallel Computing
Unit 4. GPU Computing
Unit 5. Testing

METODOLOGÍA DOCENTE

Aspectos metodológicos generales de la asignatura

Metodología Presencial: Actividades



- Exposition of the necessary theoretical concepts
- Execution of programming exercises
- Resolution of doubts regarding proposed exercises

CE28, CEO32, CEO34,
CEO32, CEO35

Metodología No presencial: Actividades

- Development of the code associated with the application
- Investigation of techniques presented in class
- Exercises and resolution of cases and problems

CG08, CG11, CEO34,
CEO35

RESUMEN HORAS DE TRABAJO DEL ALUMNO

CLASSROOM HOURS		
Lecciones de Carácter expositivo	Ejercicios y resolución de casos y de problemas	
10.00	20.00	
NON-PRESENTIAL HOURS		
Ejercicios y resolución de casos y de problemas	Estudios individual y/o en grupo, y lectura organizada	Trabajos monográficos y de investigación, individuales o colectivos
25.00	10.00	25.00
ECTS CREDITS: 3,0 (90,00 hours)		

EVALUACIÓN Y CRITERIOS DE CALIFICACIÓN

Evaluation activities	Evaluation criteria	Weight
Exam and final project.	A minimum grade of 5 is required to calculate the average with other grades (30% + 30%).	60 %
Short exams and/or small deliverables (part of the group work).	They can be oral questions or small quizzes.	10 %
Active participation of the student in practical sessions.	The practice should be partially started before each session.	10 %
Team practices and deliverables.	The practice should be partially started before each session.	20 %

Calificaciones

- Final exam: 30%
- Group work evaluation: 20%
- Active student participation: 20%

- Final project: 30%

BIBLIOGRAFÍA Y RECURSOS

Bibliografía Básica

- B. R. E. Ansell. Introduction to R – tidyverse.
- P. Araneda (2021). Introducción a las Bases de Datos
- C. Gillespie and R. Lovelace (2021). Efficient R programming.
- Lippman, S. (1997). Data structures and algorithms using C++ (1st ed.). Addison-Wesley Longman Publishing Co., Inc.
- Shwartz, A., & Ziv, Y. (2017). The nature of computation (1st ed.). Oxford University Press.
- R. G. Saavedra (2018). Introducción a la Computación con GPUs usando R.
- Kaeli, D., & Fox, P. (2015). GPU computing gems: Emerald edition (1st ed.). Morgan Kaufmann Publishers.
- Chai, J. (2019). High-performance computing with FPGAs (1st ed.). CRC Press.
- Matloff, N. (2011). The art of R programming (1st ed.). No Starch Press.
- Wickham, H., & Grolemund, G. (2017). R for data science: Import, tidy, transform, visualize, and model data (1st ed.). O'Reilly Media, Inc.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction (2nd ed.). Springer.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning: with applications in R (1st ed.). Springer.

Bibliografía Complementaria

- H. Wickham (2020). Advanced R. 2nd Edition.
- R. D. Peng (2020). R Programming for Data Science.
- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., ... Kudlur, M. (2016). TensorFlow: A system for large-scale machine learning. In 12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16) (pp. 265–283). {USENIX} Association.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning (1st ed.). Cambridge University Press.
- Xie, Y. (2016). Dynamic documents with R and knitr (2nd ed.). CRC Press.

In compliance with current regulations on the **protection of personal data**, we would like to inform you that you may consult the aspects related to privacy and data [that you have accepted on your registration form](#) by entering this website and clicking on "download"

<https://servicios.upcomillas.es/sedelectronica/inicio.aspx?csv=02E4557CAA66F4A81663AD10CED66792>