

MASTER		Master en Data Science para Finanzas
ASIGNATURA	<i>Técnicas de análisis estadístico en Data Science (I)</i>	
Nº de ECTS	6	
Nº de horas docentes	Parte II: 30h	
Nº de horas actividades académicas dirigidas		
Profesor responsable de la asignatura	Juan Manuel López Zafra	
Curso académico	2016 / 2017	
Cuatrimestre	1º Cuatrimestre	

1.- DESCRIPCIÓN GENERAL DE LA ASIGNATURA Y OBJETIVOS DE DOCENCIA:

En Técnicas de análisis estadístico en Data Science (I) parte II se pretende un doble objetivo: mostrar al alumno las características del método de análisis en data science y adentrarle en los métodos de reducción de la dimensión y de la segmentación.

Técnicas de análisis estadístico en Data Science (I) parte II.

- Primera parte: Introducción al método en data science.
- Segunda parte: Métodos de reducción de la dimensión
 - Análisis de componentes principales
 - Escalas multidimensionales
 - Análisis factorial
 - Análisis de correspondencias
- Tercer parte: Métodos de agrupación
 - Análisis cluster: métodos jerárquicos y jerárquicos
 - Métodos no jerárquicos
 - Partitioning around Medoids
 - Agglomerative Nesting
 - Divisive Analysis

2.- FORMA DE EVALUACIÓN PREVISTA:

Participación y asistencia	10%
Actividades académicas dirigidas	40%
Prueba objetiva final	50%

PROGRAMA DETALLADO

Nº de sesión	Detalle del contenido docente: temas, casos prácticos, actividades académicas dirigidas que se verán en dicha sesión,...	Lecturas recomendadas o referencias bibliográficas relativas a los conceptos-temas desarrollados en la sesión
1	Qué es el data science	Stanton chapters 0 to 3 Booz Allen Hamilton full
2	Estructura de un proyecto de data science	Video. Lecture 7 - The Structure of a Data Science Project.

		https://es.coursera.org/learn/data-science-course/lecture/AQ7T2/the-structure-of-a-data-science-project Daroczi. Chaps 1 to 7
3	Gestión y etapas en el análisis de datos	The data science project lifecycle http://www.datasciencecentral.com/profiles/blogs/the-data-science-project-lifecycle Building a data science portfolio: Storytelling with data https://www.dataquest.io/blog/data-science-portfolio-project/
4	Análisis de datos exploratorio, modelización, interpretación y comunicación de resultados	10 Data Science Skills You Need to Improve Project Success. http://customerthink.com/10-data-science-skills-you-need-to-improve-project-success/ Communicating data science: An interview with a storytelling expert Tyler Byers http://blog.kaggle.com/2016/06/13/communicating-data-science-an-interview-with-a-storytelling-expert-tyler-byers/
5	Técnicas de reducción de la dimensión: el análisis de componentes principales (ACP). Características esenciales.	Zeltermán. Mayor. Dell.
6	Técnicas de reducción de la dimensión: ACP. Práctica	
7	Técnicas de reducción de la dimensión: las escalas multidimensionales (MDS). Características esenciales.	Borg et al. Dell.
8	Técnicas de reducción de la dimensión: MDS. Práctica	
9	Técnicas de reducción de la dimensión: el análisis factorial (ANFAC). Características esenciales.	Zeltermán. Dell.
10	Técnicas de reducción de la dimensión: ANFAC. Práctica	
11	Técnicas de reducción de la dimensión: el análisis de correspondencias (ANACOR). Características esenciales.	Greenacre. Dell.
12	Técnicas de reducción de la dimensión: ANACOR. Práctica	
13	Técnicas de reducción de la dimensión: Práctica global	

14	Técnicas de reducción de la dimensión: Práctica global	
15	Métodos de agrupación: el análisis cluster (AC). Características esenciales.	Everitt et al. Mayor. Dell.
16	Métodos de agrupación: AC. Métodos jerárquicos. Práctica	
17	Métodos de agrupación: AC. Métodos no jerárquicos. Práctica	
18	Métodos de agrupación: AC. Partitioning around Medoids (PAM). Práctica	
19	Métodos de agrupación: AC. Agglomerative Nesting (AGNES). Práctica	
20	Métodos de agrupación: AC. Divisive Analysis (DIANA)	

INFORMACION ADICIONAL

Bibliografía básica

- Análisis de datos con R. Sébastien Lê y Jérôme Pagès François Husson. Publisher: ESCUELA COLOMBIANA DE INGENIERIA. 2013. ISBN-10: 9588726050
- The R Book 2nd Edition. Michael J. Crawley. Publisher: Wiley. 2012. ISBN-10: 0470973927
- Applied Multivariate Statistics with R. Zelterman, Daniel. Springer International Publishing (2015) Hardcover ISBN: 9783319140926
- Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking. Provost y Fawcett. Publisher: O'Reilly Media (2013). ISBN-10: 1449361323

Bibliografía Complementaria

- Introduction to data science. Jeffrey M. Stanton, Syracuse University.
 Descargar gratuitamente aquí: <http://surface.syr.edu/istpub/165/>
- Tips for Building a Data Science Capability. Booz Allen Hamilton
 Descargar gratuitamente aquí:
<http://www.boozallen.com/consulting/strategic-innovation/nextgen-analytics-data-science/building-a-data-science-capability>
- Cluster Analysis 5th Edition. Brian S. Everitt, Sabine Landau, Morven Leese, Daniel Stahl. Wiley (2011)
- Statistics – Textbook. Dell. Disponible online en <https://documents.software.dell.com/statistics/current/textbook>
- Learning Predictive Analytics with R. Eric Mayor. Packt Publishing Limited (2015)
- Mastering Data Analysis with R. Daroczi. Packt Publishing Limited (2015)

	<ul style="list-style-type: none">• Correspondence Analysis in Practice, 2nd ed. Greenacre. Chapman and Hall/CRC, 2007• Applied Multidimensional Scaling. Borg, Groenen, Mair. Springer (2013)
Actividades Complementarias	
Localización del profesor	Despacho E7.7. jmlopezzafra@cunef.edu