Guia docente					
Datos Identificativos					2022/23
Asignatura (*)	Ciencia de Datos			Código	710G04026
Titulación	Grao en Xestión Dixital de Información e Documentación				
Descriptores					
Ciclo	Periodo	Curso		Tipo	Créditos
Grado	2º cuatrimestre	Segundo		Formación básica	6
Idioma	CastellanoGallego				
Modalidad docente	Presencial				
Prerrequisitos					
Departamento	Matemáticas				
Coordinador/a	López Igrexas, Macías		Correo electrónico	macias.lopez@udc.es	
Profesorado	González Rueda, Ángel Manuel López Igrexas, Macías		Correo electrónico	angel.manuel.rueda@udc.es macias.lopez@udc.es	
Web	https://estudos.udc.es/gl/study/start/710G04V01				
Descripción general	Esta materia introduce y describe una serie de conceptos estadísticos centrados en la ciencia de datos. En concreto, se comenzará con una introducción a las técnicas de muestreo y diseño de encuestas, pasando por la inferencia estadística así como las principales técnicas de análisis multivariante. Adicionalmente, se introducirán distintas herramientas computacionales relacionadas con el software estadístico R para la generación de informes. Se empleará un enfoque principalmente aplicado, tratando de presentar todos los conceptos dentro del contexto de la gestión de la información y documentación.				

Competencias del título	
Código	Competencias del título
A1	CE1 - Conocer y comprender los principios teóricos y metodológicos de la gestión de información y la documentación para aplicarlos en su actividad profesional
A8	CE8 - Dominar los diferentes métodos de representación de los datos, información y el conocimiento que garanticen su recuperación eficiente
A13	CE13 - Conocer y dominar las técnicas y normativas para la creación y autenticación, reunión, selección, organización, representación, preservación, recuperación, acceso, difusión e intercambio, y evaluación de los recursos y servicios de información
A20	CE20 - Dominar las bases para desarrollar actividades de investigación utilizando métodos y principios multidisciplinares
A21	CE21 - Poseer conocimientos de estadística y análisis cuantitativo de la información
A22	CE22 - Adquirir habilidades computacionales y de manejo de las nuevas TIC
B1	CB1 - Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación
B2	CB2 - Aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio
B3	CB3 - Ser capaz de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios
B4	CB4 - Saber comunicar sus conclusiones -y los conocimientos y razones últimas que las sustentan- a públicos especializados y no especializados de un modo claro y sin ambigüedades
B5	CB5 - Poseer las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo
B6	CG1 - Capacidad para la cooperación, el trabajo en equipo y el aprendizaje colaborativo
B7	CG2 - Capacidad de reflexión y razonamiento crítico
B8	CG3 - Capacidad de planificación, organización y gestión de recursos, información y operaciones
B9	CG4 - Capacidad de análisis, diagnóstico y toma de decisiones
B10	CG5 - Capacidad de trabajar en un contexto internacional y global

B11	CG6 - Capacidad de comprender la importancia, el valor y la función de la Gestión Digital de Información y Documentación en la actual sociedad de las TIC
C1	CT1 - Expresarse correctamente, tanto de forma oral como escrita, en las lenguas oficiales de la comunidad autónoma
C2	CT2 - Utilizar las herramientas básicas de las tecnologías de la información y las comunicaciones (TIC) necesarias para el ejercicio de su profesión y para el aprendizaje a lo largo de su vida
C3	CT3 - Desarrollarse para el ejercicio de una ciudadanía respetuosa con la cultura democrática, los derechos humanos y la perspectiva de género
C4	CT4 - Entender la importancia de la cultura emprendedora y conocer los medios al alcance de las personas emprendedoras
C5	CT5 - Adquirir habilidades para la vida y hábitos, rutinas y estilos de vida saludables
C6	CT6 - Desarrollar la capacidad de trabajar en equipos interdisciplinares o transdisciplinares, para ofrecer propuestas que contribuyan a un desarrollo sostenible ambiental, económico, político y social
C7	CT7 - Valorar la importancia que tiene la investigación, la innovación y el desarrollo tecnológico en el avance socioeconómico y cultural de la sociedad
C8	CT8 - Tener la capacidad de gestionar tiempos y recursos: desarrollar planes, priorizar actividades, identificar las críticas, establecer plazos y cumplirlos

Resultados de aprendizaje			
Resultados de aprendizaje	Competencias del título		
Conocimiento de las técnicas de inferencia básicas y adquisición de habilidades para la estimación e interpretación de intervalos de confianza y contraste de hipótesis de una y dos poblaciones.	A8 A13 A21	B1 B8 B9	
Conocer los principales tipos de muestreo y las herramientas básicas para el diseño de encuestas.	A1 A13 A20 A21	B2 B3 B4 B5 B9	
Capacidad para comparar dos o más poblaciones a partir de bases de datos de diferente grado de complejidad.	A1 A21	B1 B2 B3 B4 B5	
Conocimiento de las diferentes técnicas de análisis de datos multivariantes para la descripción y obtención de información relevante a partir de bases de datos complejas.	A1 A20 A21	B1 B2 B3 B4 B5	
Habilidad para utilizar las herramientas computacionales de análisis de datos multivariantes.	A22	B11	C2 C6 C8
Integrar os conocimientos estadísticos teóricos y prácticos como vía para del conocimiento y pensamiento reflexivo y totalizador.	A1 A13	B2 B3 B4 B5 B6 B7 B10	C4 C7 C8

Capacidad de análisis y de síntesis aplicada a la gestión y organización de la información.		B2 B3 B4 B5 B6 B7 B8 B9	C1 C3 C5
Adquisición de habilidades para la toma de decisiones a partir del análisis estadístico bases de datos complejas.	A21	B2 B3 B8 B9	C8

Contenidos	
Tema	**Subtema**
Los temas siguientes desarrollan los contenidos establecidos en la ficha de la Memoria de Verificación, siendo:	Introducción y conceptos estadísticos principales relativos a muestreo y diseño de encuestas. Introducción a la inferencia estadística y estimación puntual. Intervalos de confianza. Contraste de hipótesis. Análisis de la varianza (ANOVA). Modelos de regresión. Otras técnicas de análisis multivariante. Herramientas computacionales para la generación de informes estadísticos.
1. Muestreo y encuestas: introducción y conceptos principales.	Conceptos generales de muestreo estadístico y diseño de encuestas.
2. Introducción a la inferencia estadística y estimación puntual.	Conceptos generales. Muestreo. Estimación de parámetros. Propiedades de los estimadores. Estimación puntual: estimación puntual de la media, la varianza y de una proporción.
3. Intervalos de confianza	Concepto de Intervalo de confianza. Intervalo de confianza para una media, para una varianza, para una proporción y para la diferencia de dos medias.
4. Contraste de hipótesis	Conceptos generales. Contrastes de hipótesis para la media, la proporción y para la diferencia de dos medias. Contrastes de independencia.
5. Análisis de la varianza (ANOVA).	ANOVA gráfico. ANOVA de un factor. ANOVA de más de un factor.
6. Modelos de regresión.	Modelo de regresión lineal simple y múltiple. Otros modelos de regresión.
7. Otras técnicas de análisis multivariante: análisis de componentes principales, análisis factorial, análisis de correspondencias, escalado multidimensional.	Introducción a las técnicas multivariantes más usadas.
8. Herramientas computacionales para la generación de informes estadísticos.	Introducción a distintas herramientas del software estadístico R para la generación de informes: Rstudio, Rmarkdown, Gráficos con R, Htmlwidgets.

Planificación				
Metodologías / pruebas	**Competéncias**	**Horas presenciales**	**Horas no presenciales / trabajo autónomo**	**Horas totales**
Sesión magistral	A1 A8 A20 A21 B1 B3 B7 C4 C7	21	0	21
Prácticas a través de TIC	A13 A22 B11 C2	12	0	12
Estudio de casos	A1 A8 A21 B2 B3 B4 B5 B6 B7 B8 B9 C1 C8	7	7	14
Trabajos tutelados	B2 B4 B5 B6 B8 B9 B10 C1 C3 C5 C6 C8	1.02	100.98	102
Prueba objetiva	A21 B1 B2	1	0	1

Atención personalizada		0	0	0
(*)Los datos que aparecen en la tabla de planificación són de carácter orientativo, considerando la heterogeneidad de los alumnos				

Metodologías	
Metodologías	**Descripción**
Sesión magistral	Serán sesiones expositivas en las que se introducirán y describirán los diversos temas de la asignatura, mediante presentaciones (usando los adecuados medios audiovisuales) que incluirán teoría y ejemplos.
Prácticas a través de TIC	Se desarrollarán clases prácticas mediante software estadístico, en las que se introducirá su programación y aplicación a partir de casos reales y simulados.
Estudio de casos	Se aplicarán las técnicas estadísticas impartidas en la asignatura para la resolución de ejercicios y casos de estudio reales y simulados en el ámbito de la gestión digital de la información.
Trabajos tutelados	Se realizarán trabajos individuales y/o en grupo, tutelados por los docentes de la materia, en los que se abordará la resolución, mediante la aplicación de técnicas estadísticas y el software R, de ejercicios prácticos o de casos de estudio particulares relacionados con el ámbito de la comunicación y de las ciencias de la información. También se podrá realizar un estudio de revisión acerca de un tema concreto de la materia o del software utilizado. Los trabajos podrán ser propuestos por los docentes o por los propios alumnos (las propuestas se tendrán en cuenta o no siempre según el criterio del docente).
Prueba objetiva	Consistirá en una prueba tipo test sobre los contenidos impartidos en la asignatura.

Atención personalizada	
Metodologías	**Descripción**
Sesión magistral Prácticas a través de TIC	La atención personalizada se hará, globalmente, mediante tutorías personalizadas directas y virtuales, individuales y grupales

Evaluación			
Metodologías	**Competéncias**	**Descripción**	**Calificación**
Prácticas a través de TIC	A13 A22 B11 C2	Se valorará la asistencia y/o el desempeño de los alumnos en las clases prácticas con el software estadístico.	20
Prueba objetiva	A21 B1 B2	Prueba tipo test consistente en un número de preguntas entre 10 y 20 con 3 respuestas posibles.	40
Trabajos tutelados	B2 B4 B5 B6 B8 B9 B10 C1 C3 C5 C6 C8	Se realizarán trabajos individuales y/o en grupo, tutelados por los docentes de la materia, en los que se abordará la resolución, mediante la aplicación de técnicas estadísticas y el software R, de ejercicios prácticos o de casos de estudio particulares relacionados con el ámbito de la comunicación y de las ciencias de la información. También se podrá realizar un estudio de revisión acerca de un tema concreto de la materia o del software utilizado. Los trabajos podrán ser propuestos por los docentes o por los propios alumnos (las propuestas serán tenidas en cuenta o no siempre según el criterio del docente).	40

Observaciones evaluación

Primera oportunidadSe realizará una prueba de respuesta múltiple de 10 a 20 preguntas que representa el 40% de la nota. Por otra parte, la evaluación continua constará de la asistencia y/o entrega de prácticas relacionadas con el aprendizaje y aplicación del software estadístico R para la resolución de problemas en el campo de la gestión digital de la información (20% de la nota global), además de la entrega de uno y/o varios trabajos de aplicación de la estadística para la resolución de casos de estudio en documentación digital (alternativamente podrán ser trabajos de revisión o ampliación de la materia) que representa el 40% de la nota total. Segunda oportunidad En la evaluación de la segunda oportunidad se seguirá el mismo criterio que en la primera. Convocatoria adelantada Todas las observaciones previas son aplicables a los estudiantes que soliciten la convocatoria adelantada. Calificación de no presentado En cualquiera de las dos oportunidades anuales figurará un NO PRESENTADO en aquellos casos en los que el alumnado no se presente al examen oficial. Alumnado con reconocimiento de dedicación a tiempo parcial y dispensa académica de exención de asistencia En el caso del alumnado con reconocimiento de dedicación a tiempo parcial y dispensa académica de exención de asistencia que decida no asistir a clases, este será evaluado en las dos oportunidades como el resto del alumnado que se encuentra en una situación similar. La realización fraudulenta de las pruebas o actividades de evaluación implicará directamente la calificación de suspenso "0" en la materia en la convocatoria correspondiente, invalidando así cualquier calificación obtenida en todas las actividades de evaluación de cara a la convocatoria extraordinaria.

Fuentes de información	
Básica	- Daniel Peña (2002). Análisis de datos multivariantes. S.A. MCGRAW-HILL / INTERAMERICANA DE ESPAÑA - Cao, R., Francisco, M., Naya, S., Presedo, M.A., Vázquez, M., Vilar, J.A. y Vilar, J.M. (2001). Introducción a la Estadística y sus aplicaciones. Pirámide - Egghe, L. y Rousseau, R. (1990). Introduction to Infometrics. Quantitative Methods in Library, Documentation and Information Science. Amsterdam: Elsevier - Everitt, B. y Hothorn, T. (2011). An Introduction to Applied Multivariate Analysis with R. Springer-Verlag New York
Complementária	- Daniel Zelterman (2015). Applied Multivariate Statistics with R. Springer International Publishing - Cástor Guisande, Antonio Vaamonde (2012). Gráficos estadísticos y mapas con R. Díaz de Santos - Vélez, R. & García, A. (1993). Principios de Inferencia Estadística. UNED

Recomendaciones
Asignaturas que se recomienda haber cursado previamente
Fundamentos Estadísticos/710G04040
Asignaturas que se recomienda cursar simultáneamente
Asignaturas que continúan el temario
Minería de Datos/710G04030
Otros comentarios
Para ayudar a conseguir un entorno inmediato sostenido y cumplir con el objetivo de la acción número 5: ?Docencia e investigación saludable y sostenible ambiental y social? del "Plan de Acción Green Campus Ferrol: La entrega de los trabajos documentales que se realicen en esta materia: ? Se solicitarán en formato virtual y/o soporte informático. ? Se realizará a través de Moodle, en formato digital sin necesidad de imprimirlos. ? En caso de ser necesario realizarlos en papel: - No se emplearán plásticos. - Se realizarán impresiones a doble cara. - Se empleará papel reciclado. - Se evitará la impresión de borradores. ? Se debe de hacer un uso sostenible de los recursos y la prevención de impactos negativos sobre el medio natural. ? Se trabajará para identificar y modificar perjuicios y #actitud sexistas, y se influirá en el entorno para modificarlos y fomentar valores de respeto e igualdad. ? Se deberán detectar situaciones de discriminación y se propondrán acciones y medidas para corregirlas.

(*) La Guía Docente es el documento donde se visualiza la propuesta académica de la UDC. Este documento es público y no se puede modificar, salvo cosas excepcionales bajo la revisión del órgano competente de acuerdo a la normativa vigente que establece el proceso de elaboración de guías