Teaching Guide				
Identifying Data				2024/25
Subject (*)	Fundamentals of Statistics		Code	710G04040
Study programme	Grao en Xestión Dixital de Información e Documentación			
Descriptors				
Cycle	Period	Year	Type	Credits
Graduate	2nd four-month period	First	Basic training	6
Language	SpanishGalician			
Teaching method	Face-to-face			
Prerequisites				
Department	Matemáticas			
Coordinador	Gómez Rodríguez, Marcos	E-mail	marcos.gomez.rodriguez@udc.es	
Lecturers	Gómez Rodríguez, Marcos	E-mail	marcos.gomez.rodriguez@udc.es	
Web	https://estudos.udc.es/gl/study/start/710G04V01			
General description	This subject introduces and describes the basic concepts of statistics, starting with the descriptive statistics of one or more variables, passing through probability theory, the concept of random variable and probability distributions. In addition, in view of their growing importance, time series (data characteristic of digitisation processes), bibliometric indicators, library quality indicators and index numbers will be introduced. The statistical software R and its various applications will also be introduced.			

Study programme competences / results	
Code	Study programme competences / results
A1	CE1 - Know and understand the theoretical and methodological principles of information and documentation management to apply them in their professional activity
A5	CE5 - Master the relevant sources of information that allow you to effectively meet the demands of users for both research and business
A8	CE8 - Master the different methods of representation of data, information and knowledge that ensure efficient recovery
A13	CE13 - Know and master the techniques and regulations for the creation and authentication, meeting, selection, organization, representation, preservation, recovery, access, dissemination and exchange, and evaluation of resources and information services
A21	CE21 - Possess knowledge of statistics and quantitative analysis of information
A22	CE22 - Acquire computational skills and management of new ICT
B1	CB1 - Possess and understand knowledge that provides a basis or opportunity to be original in the development and / or application of ideas, often in a research context
B2	CB2 - Apply the knowledge acquired and their ability to solve problems in new or unfamiliar environments within broader (or multidisciplinary) contexts related to their area of study
B3	CB3 - Be able to integrate knowledge and face the complexity of making judgments based on information that, being incomplete or limited, includes reflections on social and ethical responsibilities linked to the application of their knowledge and judgments
B4	CB4 - Know how to communicate their conclusions -and the knowledge and ultimate reasons that sustain them- to specialized and non-specialized audiences in a clear and unambiguous way
B5	CB5 - Possess the learning skills that allow them to continue studying in a way that will be largely self-directed or autonomous
B6	CG1 - Capacity for cooperation, teamwork and collaborative learning
B7	CG2 - Capacity for reflection and critical reasoning
B8	CG3 - Capacity for planning, organization and management of resources, information and operations
B9	CG4 - Capacity for analysis, diagnosis and decision making
B11	CG6 - Ability to understand the importance, value and function of the Digital Information and Documentation Management in the current ICT society
C1	CT1 - Express correctly, both orally and in writing, in the official languages ??of the autonomous community
C2	CT2 - Use the basic tools of information and communication technologies (ICT) necessary for the exercise of their profession and for learning throughout their lives
C3	CT3 - Develop oneself for the exercise of a citizenship that respects democratic culture, human rights and the gender perspective
C4	CT4 - Understand the importance of the entrepreneurial culture and know the means available to entrepreneurs

C5	CT5 - Acquire skills for life and habits, routines and healthy lifestyles
C6	CT6 - Develop the ability to work in interdisciplinary or transdisciplinary teams, to offer proposals that contribute to a sustainable environmental, economic, political and social development
C7	CT7 - Assess the importance of research, innovation and technological development in the socio-economic and cultural progress of society
C8	CT8 - Have the ability to manage time and resources: develop plans, prioritize activities, identify criticisms, establish deadlines and comply with them

Learning outcomes			
Learning outcomes	Study programme competences / results		
To know the most common descriptive measures and graphical representation of data.	A8 A13 A21	B1 B8 B9	
Ability to synthesise and analyse a set of data descriptively and graphically.	A8 A21 A22	B2 B3 B4 B5 B8 B9	
To know the concept of probability, rules of probabilistic calculation and the most common probabilistic models.	A1 A21	B1 B2 B3 B4 B5	
Ability to apply computer tools of statistical analysis for decision-making and for the development and exploitation of information systems.	A22	B11	C2 C6
To integrate theoretical and practical statistical knowledge as a pathway to knowledge and reflective and holistic thinking.	A1 A5	B2 B3 B4 B5 B6 B7	C4 C7 C8
Capacity for analysis and synthesis applied to the management and organisation of information.		B2 B3 B4 B5 B6 B7 B8 B9	C1 C3 C5

Contents	
Topic	Sub-topic

The following topics develop the contents established in the file of the Verification Report, being:	Sources of information and statistical methodology. Introduction to the use of statistical analysis programs (R software). Univariate descriptive statistics. Descriptive statistics of two variables. Analysis of dependence between variables and regression between statistical variables. Introduction to time series and index numbers. Methods for outlier detection (outliers). Probability. Random variables and probability distributions.
Introduction to statistics and sources of information	Introduction and purpose of statistics. Major concepts of statistics and data science. Data processing methodologies and main problems they address. Sources of statistical information. Organization of official statistics at national and international level (ONU agencies, Eurostat, INE, and IGE, among others). Main statistics in the socioeconomic field (demographic, social, employment, economic statistics). Digital information sources.
Descriptive statistics of a unidimensional variable.	General concepts. Frequencies tables. Measures of position, dispersion and shape. Graphical representation of unidimensional variables.
Descriptive statistics of more than one variable.	Descriptive statistics of two variables. Measures of position and dispersion. Graphical representation of multivariate data. Dependence relationship between qualitative variables. Dependence relationship between quantitative variables: simple linear regression. Other regression models. Descriptive methods of unsupervised classification by clusters and for the detection of outliers.
Introduction to the use of programs for the statistical analysis of data (R software).	Description of the statistical software R. Structure of R. Introduction to programming with R. Functions. Object definition. Assignment. Creating and importing databases. Obtaining graphs. Elaboration of reports.
Introduction to time series and index numbers.	Indicators in libraries and the field of documentation. Bibliometric indexes. Index numbers. Simple and composite index numbers. Introduction to descriptive analysis of time series.
Probability	Basic concepts. Operations with events. Laplace's rule. Properties of probability. Conditional probability. Product rule, rule of total probabilities, Bayes' rule. Applications to documentation problems.
Random variables	Definition of a random variable. Discrete random variables. Continuous random variables. Probability mass function. Density function. Distribution function. Mean, variance, calculation of probabilities and quantiles.
Probability distributions	Binomial distribution. Hypergeometric distribution. Negative binomial distribution. Normal distribution. Uniform distribution. Exponential distribution. Distributions associated to the normal distribution. Other distributions.

Planning				
Methodologies / tests	Competencies / Results	Teaching hours (in-person & virtual)	Student?s personal work hours	Total hours
ICT practicals	A13 A22 B11 C2	12	0	12
Guest lecture / keynote speech	A1 A5 A8 A21 B1 B3 B7 C4 C7	21	0	21
Supervised projects	B2 B4 B5 B6 B8 B9 C1 C3 C5 C6 C8	1.02	100.98	102
Objective test	A21 B1 B2	1	0	1
Case study	A1 A8 A21 B2 B3 B4 B5 B6 B7 B8 B9 C1 C8	7	7	14
Personalized attention		0		0
(*)The information in the planning table is for guidance only and does not take into account the heterogeneity of the students.				

Methodologies	
Methodologies	**Description**
ICT practicals	In the practical classes the student will be introduced to the handling of the statistical software R. Computational tools for the resolution of problems will be shown and applied through the statistical analysis of data, either from simulated or real data.
Guest lecture / keynote speech	Keynote speech will be given in which the teacher will explain, with the help of appropriate audiovisual media, the main contents of the subject.
Supervised projects	Students will be proposed to develop a group work (2 to 4 people) consisting of the application of statistical and computational tools shown in class to a particular case study, described by real or simulated data. You can also perform a work consisting of the description of a case study in the field of communications and information sciences, in which the resolution of a real problem is carried out based on the application of statistical techniques. A review study on a specific topic of the subject or the software used may also be carried out. The works can be proposed by the teachers or by the students themselves (the proposals will be taken into account or not always according to the teacher's criteria).
Objective test	It will consist of a multiple-choice test on the contents taught in the course.
Case study	The statistical techniques taught in the course will be applied to solve exercises and real and simulated case studies in the field of digital information management.

Personalized attention	
Methodologies	**Description**
ICT practicals Guest lecture / keynote speech	There will be keynote lectures in which the teacher will explain, with the help of appropriate audiovisual media, the main contents of the subject, promoting the debate in class. In the particular case of students with academic dispensation, you can perform face-to-face and virtual tutorials (email, video conference), which allow the student to satisfactorily follow the subject.

Assessment			
Methodologies	**Competencies / Results**	**Description**	**Qualification**
Supervised projects	B2 B4 B5 B6 B8 B9 C1 C3 C5 C6 C8	These works will be carried out in groups of 2 to 5 people, applying statistics to real or simulated data, reviewing a topic on statistics or data science or even regarding a specific application of statistics related to the field of communication and information scienses.	40
ICT practicals	A13 A22 B11 C2	The attendance and performance of the student in the practical classes will be evaluated, as well as the delivery of works related to the application of the statistical software R.	20
Objective test	A21 B1 B2	It will consist of 15 to 20 test questions with three possible answers.	40

Assessment comments

First chance evaluationThere will be a multiple-choice test of 10 to 20 questions that represents 40% of the grade. On the other hand, the continuous assessment will consist of attendance and / or delivery of practices related to learning and application of statistical software R for problem solving in the field of digital information management (20% of the overall grade), in addition to the delivery of one and / or several works of application of statistics for the resolution of case studies in digital documentation (alternatively may be tasks of revision or extension of the subject) which represents 40% of the final grade.Second chance evaluation The evaluation will be done following the same procedure as in the first opportunity. Early exam session All these remarks are applied to the early exam session. "No presentado" grade For any of the two opportunities to pass the subject, the "NO PRESENTADO" grade will be given to the students who did non take the objective final test. Students with recognition of part-time dedication and/or academic exemption of attendance In the case of students with recognition of part-time dedication and/or academic exemption of attendance that decides not to attend classes, they will be evaluated in the two opportunities as the rest of the students who are in a similar situation. Fraud in tests or evaluation activities will directly imply the failure grade "0" in the subject in the corresponding call, thus invalidating any grade obtained in all the evaluation activities for the extraordinary call.

Sources of information	
Basic	- Marín, J. (1999). Estadística Aplicada a las Ciencias de la Documentación. Murcia: Diego Marín Editor - Egghe, L. y Rousseau, R. (1990). Introduction to Infometrics. Quantitative Methods in Library, Documentation and Information Science.. Amsterdam: Elsevier. - Cao, R., Labora, A., Naya, S. e Ríos, M. (2001). Métodos estatísticos e numéricos. A Coruña: Baía - Moya, F., López, J. y García C. (1996). Técnicas Cuantitativas Aplicadas a la Biblioteconomía y Documentación. Madrid: Síntesis - 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 - Stephen, P. and Hornby, S. (1997). Simple statistics for library and information professionals.. London:Library Association Publishing - Judit Bar-Ilan (2008). Informetrics at the beginning of the 21st century?A review.. Journal of Informetrics - María Dolores Ugarte, Ana F. Militino, and Alan T. Arnholt (2012). Probability and Statistics with R. Springer - José María Sarabia Alegría , Faustino Prieto Mendoza , Vanesa Jordá Gil (2018). Prácticas de estadística con R. Pirámide
Complementary	- Gonick, L. e Smith, W. (2001). A estatística ¡en caricaturas!. Lugo:SGAPEIO - Judit Bar-Ilan (2008). Informetrics at the beginning of the 21st century?A review.. Journal of Informetrics - Cástor Guisande, Antonio Vaamonde (2012). Gráficos estadísticos y mapas con R. Díaz de Santos

Recommendations
Subjects that it is recommended to have taken before
Subjects that are recommended to be taken simultaneously
Subjects that continue the syllabus
Data Science/710G04026
Other comments

Para axudar a conseguir unha contorna inmediata sostida e cumprir co obxectivo da acción número 5: ?Docencia e investigación saudable e sustentable ambiental e social? do "Plan de Acción Green Campus Ferrol:

A entrega dos traballos documentais que se realicen nesta materia:

? Solicitaranse en formato virtual e/ou soporte informático.

? Realizarase a través de Moodle, en formato dixital sen necesidade de imprimilos.

? En caso de ser necesario realizalos en papel:

- Non se empregarán plásticos.

- Realizaranse impresións a dobre cara.

- Empregarase papel reciclado.

- Evitarase a impresión de borradores.

? Débese de facer un uso sustentable dos recursos e a prevención de impactos negativos sobre o medio natural.

? Traballarase para identificar e modificar prexuízos e actitudes sexistas, e influirase na contorna para modificalos e fomentar valores de respecto e igualdade.

? Deberanse detectar situacións de discriminación e propoñeranse accións e medidas para corrixilas.