



UNIVERSITÀ DEGLI STUDI DI PALERMO

DEPARTMENT	Scienze Economiche, Aziendali e Statistiche		
ACADEMIC YEAR	2022/2023		
BACHELOR'S DEGREE (BSC)	STATISTICS FOR DATA ANALYSIS		
INTEGRATED COURSE	EXPLORATORY STATISTICS WITH LABORATORY - INTEGRATED COURSE		
CODE	22605		
MODULES	Yes		
NUMBER OF MODULES	2		
SCIENTIFIC SECTOR(S)	SECS-S/01		
HEAD PROFESSOR(S)	PLAIA ANTONELLA	Professore Ordinario	Univ. di PALERMO
OTHER PROFESSOR(S)	SCIANDRA MARIANGELA	Professore Associato	Univ. di PALERMO
	PLAIA ANTONELLA	Professore Ordinario	Univ. di PALERMO
CREDITS	12		
PROPAEDEUTICAL SUBJECTS			
MUTUALIZATION			
YEAR	1		
TERM (SEMESTER)	1° semester		
ATTENDANCE	Not mandatory		
EVALUATION	Out of 30		
TEACHER OFFICE HOURS	<p>PLAIA ANTONELLA Wednesday 15:00 - 17:00 La modalita, in studio o su Teams, va concordata col docente</p> <p>SCIANDRA MARIANGELA Wednesday 12:00 - 14:00 DSEAS 2 piano</p>		

DOCENTE: Prof.ssa ANTONELLA PLAIA

PREREQUISITES	Mathematics
LEARNING OUTCOMES	<p>Knowledge and understanding Knowledge of statistical methods and procedures for exploratory and descriptive analysis of univariate and bivariate data. General knowledge of the R programming environment; Knowledge of the main objects used in R; Knowledge of methods to build tables and graphs in R; Knowledge of commands to import data into R; Knowledge of commands for building user-defined functions.</p> <p>Ability to apply knowledge and understanding Being able to choose the appropriate statistical computer analysis solutions for the particular study. Ability to use the R programming environment in general terms; Ability to manipulate the main objects used in R; Ability to build tables and graphs in R; Ability to import data in R; Ability to construct user-defined functions.</p> <p>Making judgments To be able to: 1. translate into statistical terms a knowledge requirement arose in different application fields. 2. intervene with cleaning activities, reorganization, descriptive analysis and interpretation of data from different types of sources, 3. process and communicate consistently the results of their observations The student must be able to use the main R commands seen during the course. The student must also be able to set some elementary problems and statistical analysis to be solved using the R programming environment.</p> <p>Communication skills To be able to: 1. grasp and define the statistical target of a study involving parties not experts, justifying the choices of the tools used for the analysis, 2. communicating the results of analysis with appropriate language.</p> <p>Learning ability The student must have developed the learning skills necessary to continue the study of statistical, methodological and applied, by mastering the basic cultural and educational content necessary also to follow the updates of the discipline To Be able to use R both as statistical software and as a programming environment.</p>
ASSESSMENT METHODS	<p>Ongoing test, written test, oral interview. The exam is done through an oral interview, subject to passing the written test, and, for the winter session tests, to overcoming the ongoing evaluation. The Selection Committee is formed by the two chairs of the classes.</p> <p>Ongoing evaluation The ongoing evaluation (CB) will be held on the 21st of November 2022, it will last one hour and will focus on the part of the program of the first 6 weeks of the course. It will consist of problems to be solved in R and multiple-choice/open questions. The final mark will be in thirties, it is believed passed with a score of 18/30, and is valid only if the student completes the exam (written and oral) during the winter session.</p> <p>Written test The written exam for those who passed the ongoing test, and limited to the winter session, consists of one or more exercises to be solved in R and multiple-choice/open questions on subjects covered in the second part of the course and lasts one hour. The written test for those who did not pass the ongoing test and during all the other sessions lasts two hours and includes two or more exercises to be solved in R and multiple-choice/open questions focusing on the whole program of the course. It will not be allowed to use books or notes. Starting from the spring session, all the students will have to take the written test lasting 2 hours. The tests, well-defined, clear, different in difficulty and uniquely interpretable, allow students to formulate the answer independently, and they are structured to allow comparison with that provided by other students. The sufficiency threshold, which indicates the passing of the test, is in the proper use of terms relating only to the basic concepts of the course, and i) in the case of a practical question, with the application of appropriate statistical methodology although</p>

	<p>spoiled by mere miscalculation (provided consistent with the methodology itself); ii) in the case of a theoretical question, inconsistency (from a statistical point of view) of the answer, albeit not complete. The final mark will be in the thirties, it is believed to pass with a score of 18/30</p> <p>Oral exam The oral test is intended to deepen the written work and assessment learning of the student.</p> <p>FINAL EVALUATION The final evaluation of the examination will consider three aspects: i) mastery of the subjects; ii) the ability to the application of knowledge and iii) the properties of language, assessed as a whole in the written test (with a weight of 0.7, or 1 in the absence of ongoing test) and the ongoing evaluation (if present, with weight 0.3).</p> <p>The teacher could take into account the other factors (such as active participation in classes and exercises, or presence of some disables). If the student does not pass the examination, he can enrol for the next one.</p>
TEACHING METHODS	Lectures, problem classes, homework, lab

<p>MODULE EXPLORATORY STATISTICS</p> <p><i>Prof.ssa ANTONELLA PLAIA</i></p>	
SUGGESTED BIBLIOGRAPHY	
Cicchitelli, D'Urso, Minozzo, Statistica: Principi e Metodi, III ed., Pearson, 2017. Capp. 1-7, 9-11 Borra, Di Ciaccio, Statistica 3ª ed., McGraw-Hill. Capp 1-4, 6. Materiale didattico fornito dal docente, disponibile sul Portale Studenti.	
AMBIT	50244-Statistico - probabilistico
INDIVIDUAL STUDY (Hrs)	98
COURSE ACTIVITY (Hrs)	52
EDUCATIONAL OBJECTIVES OF THE MODULE	
The course aims to provide the basic elements useful for the study of collective phenomena. Indeed statistics analyzes quantitatively the phenomena whose study requires the observation of a series of individual events. In particular the methods of that part of the statistics, known as Descriptive Statistics, directed to the analysis of the characteristics of a collective observed in its entirety, will be introduced. At the end of the course the student will be 'able to identify the appropriate methods and techniques and apply them to the analysis, summarizing the results in the most effective way.	

SYLLABUS

Hrs	Frontal teaching
2	Objectives of the discipline and its subdivision
2	Statistical concepts and introductory terminology
2	Population; statistical units; characters and variables
4	Tables ; absolute, relative and cumulative frequencies.
2	Statistical plots
6	Central tendencies. Measures of variability and eterogeneity
2	Cross-tabulations and contingency tables. Conditional distributions
4	Marginal and conditional central tendencies and measures of variability
4	Statistical dependence. Mean dependence
4	Linear regression and correlation
Hrs	Practice
8	Distribution of a statistical variable. Statistical plots. Central tendencies and measures of variability
6	Distributions of pairs of variables
6	Association and correlation

MODULE "R" WORKSHOP

Prof.ssa MARIANGELA SCIANDRA

SUGGESTED BIBLIOGRAPHY

V. Muggeo e G. Ferrara: Il linguaggio R: concetti introduttivi ed esempi (II edizione). Pubblicato nel sito del CRAN (The Comprehensive R Archive Network). (Ultima versione disponibile sul sito).
 W. N. Venables, D. M. Smith and the R Core Team, An Introduction to R. Notes on R: A Programming Environment for Data Analysis and Graphics. Pubblicato nel sito del CRAN (The Comprehensive R Archive Network). (Consigliata qualsiasi edizione del testo).

AMBIT	50250-Statistico, statistico applicato, demografico
INDIVIDUAL STUDY (Hrs)	98
COURSE ACTIVITY (Hrs)	52

EDUCATIONAL OBJECTIVES OF THE MODULE

The course has as its fundamental objective to offer the student the tools to be able to use the programming environment R in a conscious way. At the end of the course the student must be able to solve problems and conduct elementary statistical analysis with the notions learned during the course.

SYLLABUS

Hrs	Frontal teaching
2	Course presentation. Main differences between free software and commercial software for statistical analysis
2	Introduction to the R programming environment
6	Introduction to RStudio. The basic object of R: the vector. Types of vectors. Assignment. Operations on vectors. Factors.
2	Array e matrici: R function to build matrices
4	Lists and data-frame. The attach() and detach() functions. Data-frame available in R.
2	Reading data file
3	Construction of simple and double frequency distributions. Calculation of absolute and relative frequencies. Conditional and marginal distributions.
4	Graphs: high-level and low-level functions. Main graphical representations. Use of the par() function.
3	Statistical analysis: main functions for calculating descriptive statistics. Probability distributions.
4	Elements R programming . User-defined functions.
Hrs	Practice
6	Application of major commands in R to produce statistics.
10	Applications related to the construction of frequency distributions and graphs.
4	Applications related to the construction of user-defined code and functions.