



UNIVERSITÀ DEGLI STUDI DI PALERMO

DIPARTIMENTO	Scienze Economiche, Aziendali e Statistiche		
ANNO ACCADEMICO OFFERTA	2023/2024		
ANNO ACCADEMICO EROGAZIONE	2023/2024		
CORSO DILAUREA	STATISTICA PER L'ANALISI DEI DATI		
INSEGNAMENTO	STATISTICA ESPLORATIVA CON LABORATORIO C.I.		
CODICE INSEGNAMENTO	22605		
MODULI	Si		
NUMERO DI MODULI	2		
SETTORI SCIENTIFICO-DISCIPLINARI	SECS-S/01		
DOCENTE RESPONSABILE	PLAIA ANTONELLA	Professore Ordinario	Univ. di PALERMO
ALTRI DOCENTI	SCIANDRA	Professore Associato	Univ. di PALERMO
	MARIANGELA		
	PLAIA ANTONELLA	Professore Ordinario	Univ. di PALERMO
CFU	12		
PROPEDEUTICITA'			
MUTUAZIONI			
ANNO DI CORSO	1		
PERIODO DELLE LEZIONI	1° semestre		
MODALITA' DI FREQUENZA	Facoltativa		
TIPO DI VALUTAZIONE	Voto in trentesimi		
ORARIO DI RICEVIMENTO DEGLI STUDENTI	PLAIA ANTONELLA		
	Mercoledì	15:00 - 17:00	La modalita, in studio o su Teams, va concordata col docente
	SCIANDRA MARIANGELA		
	Mercoledì	12:00 - 14:00	DSEAS 2 piano

DOCENTE: Prof.ssa ANTONELLA PLAIA

PREREQUISITI	Conoscenze di base di matematica
RISULTATI DI APPRENDIMENTO ATTESI	Conoscenza e capacita' di comprensione Conoscenza dei metodi e delle procedure statistiche per analisi esplorative e descrittive dei dati univariati e bivariati. Conoscenza generale dell'ambiente di programmazione R; Conoscenza dei principali oggetti utilizzati in R; Conoscenza dei metodi per costruire tabelle e grafici in R; Conoscenza dei comandi per importare dati in R; Conoscenza dei comandi per la costruzione di funzioni definite dall'utente. Capacita' di applicare conoscenza e comprensione Essere in grado di scegliere soluzioni di analisi statistico-informatiche appropriate al particolare studio. Capacita' di utilizzare in termini generali l'ambiente di programmazione R; Capacita' di manipolare i principali oggetti utilizzati in R; Capacita' di costruire tabelle e grafici in R; Capacita' di importare dati in R; Capacita' di costruire funzioni definite dall'utente. Autonomia di giudizio Essere in grado di: 1. tradurre in termini statistici una esigenza conoscitiva sorta in differenti campi applicativi. 2. intervenire con attivita' di pulizia, riorganizzazione, analisi descrittiva e interpretazione, dei dati provenienti da fonti di tipo diverso, 3. elaborare e comunicare coerentemente i risultati delle proprie osservazioni Inoltre, lo studente deve essere in grado di utilizzare i principali comandi di R visti durante il corso e deve saper impostare alcuni elementari problemi e analisi statistiche da risolvere con l'utilizzo dell'ambiente di programmazione R. Abilita' comunicative Essere in grado di: 1. cogliere e di definire l'obiettivo statistico di uno studio con interlocutori non esperti, giustificando le scelte degli strumenti utilizzati per l'analisi, 2. comunicare i risultati delle analisi con linguaggio appropriato. Capacita' d'apprendimento Lo studente dovra' aver sviluppato le abilita' di apprendimento necessarie per continuare lo studio della statistica, metodologica e applicata, padroneggiando contenuti culturali e formativi di base necessari anche per seguire gli aggiornamenti della disciplina. Essere in grado di utilizzare R sia come software statistico che come ambiente di programmazione
VALUTAZIONE DELL'APPRENDIMENTO	Prova in itinere, prova scritta, prova orale. La verifica dell'apprendimento avviene attraverso un colloquio orale, subordinato al superamento della prova scritta e, limitatamente alla sessione invernale di esami, al superamento della prova in itinere. Durante il corso potranno essere organizzati lavori di gruppo, report di analisi, presentazioni autonome degli studenti anche con svolgimento di attività in modalità homework da discutere in aula: queste attività, potranno costituire parte della valutazione. La Commissione giudicatrice e' formata dai docenti titolari dei due moduli del CI. Prova in itinere La prova in itinere (al pc) si svolgera' il 21 novembre 2022, avra' la durata di 1 ora e vertera' sulla parte di programma svolta nelle prime 6 settimane di corso. Comprendera' esercizi da svolgere in R e domande a risposta multipla. La valutazione della prova in itinere e' in trentesimi. La prova si ritiene superata con un punteggio di almeno 18/30, ed e' valida solo se si completa l'esame (prova scritta e prova orale) entro la sessione invernale. Prova scritta La prova scritta per chi ha superato la prova in itinere, e limitatamente agli appelli della sessione invernale, e' costituita da uno o piu' esercizi da svolgere in R e domande a risposta multipla o aperte sugli argomenti trattati durante la seconda parte del corso e ha la durata di 1 ora. La prova scritta per chi non ha superato la prova in itinere e per tutti nelle altre sessioni di esame ha la durata di due ore e comprende 2 o piu' esercizi da svolgere in R e domande a risposta multipla o aperte su tutto il programma svolto. Non sara' consentito l'uso di libri o di appunti. Gli studenti che si presenteranno a partire dall'appello di aprile di esami dovranno sostenere la prova scritta della durata di 2 ore. I testi, ben definiti, chiari, di difficolta' differente e unicamente interpretabili,

	permettono allo studente di formulare autonomamente la risposta e sono strutturati in modo da consentire il confronto con quella fornita dagli altri studenti. La soglia di sufficienza, che indica il superamento della prova, consiste nell'uso adeguato dei termini relativi ai soli concetti base propri del corso in oggetto d'esame, e i) nel caso di quesito pratico, con l'applicazione della idonea metodologia statistica anche se viziato dal mero errore di calcolo (purche' coerente con la metodologia stessa); ii) nel caso di quesito teorico, nella coerenza (anche statistica) della risposta, seppur non esaustiva dell'argomento. La valutazione della prova scritta e' in trentesimi. La prova si ritiene superata con un punteggio di almeno 18/30. Prova orale La prova orale mira ad approfondire l'elaborato scritto e la valutazione dell'apprendimento dello studente. VALUTAZIONE FINALE La valutazione finale dell'esame prendera' in considerazione tre aspetti: i) la padronanza degli argomenti; ii) la capacita' di applicazione delle conoscenze e iii) la proprieta' di linguaggio, valutati nel complesso della prova scritta (peso pari a 0.7, o a 1 in assenza di prova in itinere) e dell'eventuale prova in itinere (peso pari a 0.3). Il docente potra' tenere conto dei fattori di contesto dell'esame (come ad esempio la partecipazione attiva durante le lezioni e le esercitazioni, oppure la presenza di qualche disabilita').
ORGANIZZAZIONE DELLA DIDATTICA	Lezioni frontali, esercitazioni in aula alla lavagna, esercitazioni a casa, laboratorio

MODULO STATISTICA ESPLORATIVA

Prof.ssa ANTONELLA PLAIA

TESTI CONSIGLIATI

Cicchitelli, D'Urso, Minozzo, Statistica: Principi e Metodi, IV ed., Pearson, 2022. Capp. 1-7, 9-11
 Borra, Di Ciaccio, Statistica 3^a ed., McGraw-Hill. Capp 1-4, 6.
 Materiale didattico fornito dal docente, disponibile sul Portale Studenti iscrivendosi all'insegnamento.

TIPO DI ATTIVITA'	A
AMBITO	50244-Statistico - probabilistico
NUMERO DI ORE RISERVATE ALLO STUDIO PERSONALE	94
NUMERO DI ORE RISERVATE ALLE ATTIVITA' DIDATTICHE ASSISTITE	56

OBIETTIVI FORMATIVI DEL MODULO

Il corso mira a fornire gli elementi base utili per lo studio dei fenomeni collettivi.
 La statistica infatti analizza in termini quantitativi i fenomeni il cui studio richiede l'osservazione di un insieme di manifestazioni individuali.
 In particolare verranno introdotti i metodi di quella parte della Statistica, nota come Statistica Descrittiva, rivolti all'analisi delle caratteristiche di un collettivo osservato nella sua totalità.
 Alla fine del corso l'alunno sarà in grado di individuare i metodi e le tecniche adeguati all'analisi e applicarli, sintetizzando i risultati nel modo più efficace.

PROGRAMMA

ORE	Lezioni
2	Obiettivi della disciplina e sua suddivisione.
2	Terminologia statistica e concetti introduttivi
2	Popolazione; unità statistiche; caratteri e variabili; modalità
4	Tabelle semplici; frequenze assolute, relative e cumulate.
2	Rappresentazioni grafiche
6	Misure di posizione. Misure di variabilità ed eterogeneità.
2	Distribuzioni di frequenza doppie. distribuzioni marginali e condizionate; frequenze assolute e relative.
4	Misure di posizione e di variabilità marginali e condizionate
4	Dipendenza in distribuzione Dipendenza in media
4	Regressione lineare semplice e correlazione
ORE	Esercitazioni
8	Distribuzione di un carattere e sua rappresentazione: le medie e gli indici di variabilità.
8	Distribuzione congiunta di due caratteri.
8	La dipendenza e l'interdipendenza statistica.

**MODULO
LABORATORIO DI R**

Prof.ssa MARIANGELA SCIANDRA

TESTI CONSIGLIATI

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).

TIPO DI ATTIVITA'	B
AMBITO	50250-Statistico, statistico applicato, demografico
NUMERO DI ORE RISERVATE ALLO STUDIO PERSONALE	94
NUMERO DI ORE RISERVATE ALLE ATTIVITA' DIDATTICHE ASSISTITE	56

OBIETTIVI FORMATIVI DEL MODULO

Il corso ha come obiettivo fondamentale di offrire allo studente gli strumenti per potere utilizzare in maniera consapevole l'ambiente di programmazione statistica R. Alla fine del corso lo studente dovrà essere in grado di risolvere problemi e condurre analisi statistiche elementari con le nozioni imparate durante il corso.

PROGRAMMA

ORE	Lezioni
2	Presentazione del corso. Principali differenze tra software libero e software commerciale
2	Introduzione all'ambiente di programmazione R
6	Introduzione a RStudio. L'oggetto base di R: il vettore. Tipi di vettori. Operazione di assegnazione. Operazioni sui vettori. Fattori.
2	Array e matrici: operazioni e funzioni per l'utilizzo delle matrici
4	Liste e data-frame. Funzioni attach() e detach(). Gestione del cammino di ricerca. Data-frame forniti da R.
2	Lettura di dati da file.
3	Costruzione di distribuzioni di frequenze semplici e doppie. Calcolo di frequenze assolute e relative. Distribuzioni condizionate e marginali
4	Grafici: funzioni di alto livello e di basso livello. Principali rappresentazioni grafiche. Utilizzo della funzione par().
3	Analisi di tipo statistico: principali funzioni per il calcolo di statistiche descrittive. Distribuzioni di probabilità.
4	Elementi di programmazione in R. Funzioni definite dall'utente.
ORE	Esercitazioni
6	Applicazione dei principali comandi in R per la produzione di statistiche.
8	Applicazioni relative alla costruzione di codice e funzioni definite dall'utente.
10	Applicazioni relative alla costruzione di distribuzioni di frequenze e di grafici.