



# UNIVERSITÀ DEGLI STUDI DI PALERMO

Dipartimento: Scienze Economiche, Aziendali e Statistiche

A.A. 2024/2025

## PIANO DI STUDI DEL CORSO DI LAUREA MAGISTRALE IN STATISTICA E DATA SCIENCE

- DATA SCIENCE AND STATISTICS (LM-DATA) -

### Caratteristiche

				
Classe di Laurea magistrale in Data science (LM Data)	2 ANNI	PALERMO	ACCESSO LIBERO	2316
Classe di Laurea magistrale in Scienze statistiche (LM-82 R)	2 ANNI	PALERMO	ACCESSO LIBERO	2316

### Obiettivi del Corso di Studi

Obiettivi specifici:

La Laurea Magistrale in Statistica e Data Science, coerentemente con quanto richiesto dagli obiettivi formativi delle due classi e con le indicazioni provenienti dalle indagini sulla collocazione nel mercato del lavoro dei laureati in discipline statistiche (come, ad esempio, dall'iniziativa interuniversitaria AlmaLaurea), si propone di ottenere una figura di laureato che, dotato di una solida preparazione superiore nell'ambito della matematica, della probabilità, della statistica e dell'informatica, sia capace di operare in vari settori di applicazione con autonomia e responsabilità e di inserirsi nel mercato del lavoro come esperto qualificato, in grado di produrre, gestire, analizzare e interpretare flussi informativi diversificati.

In particolare, il CdS fornisce gli strumenti per consentire una solida preparazione metodologica statistica e di interrogazione ed estrazione dei dati insieme con strumenti propri di alcuni contesti applicativi. L'attività didattica frontale sarà caratterizzata da una forte integrazione fra lezioni teoriche ed esercitazioni e laboratori. L'assetto formativo del CdS si caratterizza per: - la presenza di un pacchetto di insegnamenti comuni di livello avanzato di discipline matematiche, probabilistiche, statistiche ed informatiche in grado di assicurare allo studente l'approfondimento e l'acquisizione di conoscenze utili per i successivi ampliamenti metodologici e applicativi della statistica; - un buon grado di personalizzazione del Piano di Studi da parte dello studente, al quale sono consigliate alcune personalizzazioni con riferimento a competenze specifiche in alcuni contesti applicati; - l'attenzione alle metodologie didattiche, avendo cura che la solida formazione teorica, basata sulle lezioni frontali, sia integrata con laboratori applicativi il più possibile in ambienti eco-digitali nei quali saranno discussi casi e problemi reali e saranno approfonditi temi di rilevanza applicativa; - attenzione ai metodi, strumenti ed applicazioni di didattica innovativa (gamification, flipped classroom, etc.) e service learning (con specifico coinvolgimento degli studenti).

Si prevedono due curriculum anche nell'ipotesi di tenere conto della crescente domanda internazionale verso i percorsi formativi progettati. Nella programmazione proposta, le discipline sono quasi del tutto coperte da professori di ruolo di prima e seconda fascia e, per garantire massima aderenza ai contenuti ed obiettivi formativi, le discipline in gran parte sono attivate ex novo per il nuovo corso di studi, quindi con un'attenzione specifica a contenuti disegnati per il previsto profilo formativo in uscita e, per quelle mutate, previsto (ove necessario) un adeguamento della scheda di trasparenza (programma dei contenuti).

Sono previsti due curriculum. Il primo curriculum prevede un primo nucleo, a primo anno, di discipline statistiche avanzate e di data science e risponde all'esigenza di mantenere una filiera che, sulla base della laurea di primo livello in statistica (L41) già attiva in Ateneo da molti anni, sviluppi in modo coerente e specifico le competenze maggiormente relate alla tradizione disciplinare e contenutistica, teorica ed applicata, delle discipline statistiche. Il curriculum enfatizza il ruolo delle statistiche per dati complessi e per l'estrazione e la gestione di big data con particolare dettaglio alle modellazioni statistiche più avanzate e con specifiche applicazioni in campo sociale ed economico, questo anche grazie ad un ampio ricorso a tre gruppi opzionali (più gruppo a scelta dello studente) di materie in grado di garantire un sufficiente approfondimento tematico in relazione agli interessi dello studente.

In tal senso, il primo curriculum prevede una formazione orientata alle basi di processi stocastici e networks, con tecniche di data mining e supporto dei metodi e modelli statistici avanzati sia nel contesto parametrico che non parametrico e con specifica attenzione agli approcci bayesiani, senza trascurare gestione dati e di database, dunque le necessarie competenze giuridiche sulla tutela, gestione e conservazione del dato, nonché specifici ambiti applicativi orientati all'azienda ed al

Legenda: Per. = periodo o semestre, Val. = Valutazione (V=voto, G=giudizio), TAF= Tipologia Attività Formativa (A=base, B=caratterizzante, C=Affine, S=stages, D=a scelta, F=altre)

contesto economico e sociale di analisi. Questi aspetti possono essere ulteriormente specificati, a scelta dello studente, tramite opzionali (gruppo 1 e gruppo 2) che permettono approfondimenti di statistica genetica, metodi di valutazione e performance di servizi, campionamento ed indagini sociali, analisi di rischio finanziario, analisi economiche territoriali, od anche in termini di modellazione quali biostatistica, modelli per dati categoriali, elaborazione del linguaggio naturale e strumenti per internet. Le discipline sono attivate ex novo per garantire massima corrispondenza con le esigenze formative del CdS e massima coerenza all'interno del percorso di studio. Va evidenziato che lo studente pu' anche aggiungere tra le opzionali (gruppo 3) alcune discipline del secondo curriculum e, particolarmente, tecniche e metodi su alcuni temi di analisi come - ad esempio - cloud computing, high dimensional data, models for volatility in finance, ma anche discipline trasversali di piu' ampio respiro ed appartenenti all'area matematica, economica, aziendale, sociologica ed informatica. Come sempre, rimane l'ulteriore gruppo a scelta libera dello studente.

Il secondo curriculum offre contenuti parzialmente differenziati - anche con dedicate materie opzionali - in modo da garantire allo studente un percorso di studi in linea con aspetti specifici di analisi dati. Dunque, il secondo curriculum e' orientato maggiormente alla gestione del dato complesso, l'interrogazione ma anche la progettazione della struttura dati, ponendo enfasi sui metodi automatici piu' computer-side ed intensivi nel calcolo anche rispetto alla gestione di dati multidimensionali ad alta frequenza. Anche in questo curriculum la personalizzazione con materie opzionali (piu' gruppo a scelta dello studente) consente alcuni approfondimenti tematici, ma qui limitati a due gruppi piu' specifici con discipline tecniche volte anche a completare, all'occorrenza, le competenze statistico-metodologiche dello studente. Dunque, in questo profilo si propone dal primo anno un potenziamento maggiormente orientato alle competenze di statistical learning, data e text mining e calcolo avanzato, senza tralasciare gli aspetti applicativi in ambito economico e sociale, ma ora declinati in termini di approccio con dati complessi e piu' computer-intensive come gli approcci di modellazione in finanza oppure ottimizzazione ed applicazioni in contesti di dati multidimensionali ad altissima frequenza. I metodi piu' prettamente statistici vengono destinati ad un gruppo opzionale (gruppo 1) che seleziona alcuni specifici contenuti esclusivamente di approfondimento metodologico, mentre un altro gruppo opzionale (gruppo 2) individua discipline anche trasversali ma con alta qualificazione tecnica quale financial econometrics, risk management, survey sampling methods, marketing decisions ed ancora discipline di approfondimento matematico e informatico quali cybersecurity e internet of the things. Ancora una volta, l'ulteriore gruppo (gruppo 4) a scelta libera dello studente consente l'aggiunta libera di materie od altra attivita' formativa. Infine, l'erogazione in inglese di questo secondo curriculum appare utile per favorire l'accoglimento di qualificata domanda internazionale ed a completare un'offerta formativa ampia anche a vantaggio dello studente locale che avra' in sede opportunita' formative aggiuntive.

E' utile notare che i due curriculum prevedono uno stage di base estendibile anche tramite le scelte libere degli studenti che, dunque, possono raggiungere ampia copertura di CFU per tirocini garantendo massima flessibilita' anche in funzione di specifiche esigenze manifestate dallo studente o dall'azienda ospitante.

Entrambi i curriculum prevedono una distribuzione equa di CFU tra i 4 semestri dei due anni con una maggiore concentrazione delle attivita' formative nei primi 3 semestri, con l'ultimo semestre maggiormente dedicato ad attivita' di stage, tirocini e consulenze, nonche' attivita' a libera scelta dello studente e, comunque, organizzato in modo tale da lasciare allo studente maggiore liberta' anche per la predisposizione della tesi di laurea.

### **Sbocchi occupazionali**

Profilo:

Statistico Data Scientist

Funzioni:

Nell'ambito della funzione in un contesto di lavoro: analisi dei dati e supporto all'attivita' di gestione del rischio in ambito economico, finanziario e creditizio, analisi dei dati e supporto alla ricerca in ambito clinico, epidemiologico e biologico, nonche' nell'ambito delle funzioni proprie della statistica ufficiale tramite approcci statistici e procedure informatiche avanzati per estrarre e gestire il dato complesso.

Competenze:

Per le competenze associate alla funzione: progettazione e realizzazione di attivita' di valutazione per la gestione della qualita' e per la valutazione delle performance nell'ambito dei processi di prodotto e servizio; certificazione delle metodologie e delle tecniche statistiche applicate alle indagini; analisi di dati e formalizzazione di modelli matematico/statistici per indagare i fenomeni e per effettuare previsioni nei vari ambiti applicativi; progettazione, creazione e gestione di banche dati per finalita' di analisi statistica e del rischio; progettazione di piani di campionamento e di indagini statistiche demoscopiche complesse attinenti ai campi sociali ed economici; progettazione, analisi e verifica dei risultati di esperimenti e prove cliniche controllate; analisi di dati e formalizzazione di modelli matematico/statistici per indagare i fenomeni e per effettuare previsioni in ambito economico, finanziario, biologico, sanitario, epidemiologico. Costruzione di database complessi, interrogazione ed estrazione dell'informazione con sintesi statistica del dato.

Sbocchi:

Con riferimento agli sbocchi occupazionali: nelle pubbliche amministrazioni; negli uffici studi delle aziende che operano in campo economico, finanziario, e assicurativo; in uffici statistici di medio-grandi imprese; in uffici marketing di imprese di produzione e di distribuzione; in societa' di gestione di sistemi informativi; in societa' di consulenza statistica che svolgono attivita' di supporto esterno ad aziende private e pubbliche; in centri ed istituti di ricerca pubblici e privati; nelle aziende sanitarie, sia nel settore clinico che nel settore epidemiologico che in quello gestionale; negli assessorati alla sanita' nei settori valutazione ed epidemiologia; in uffici di progettazione e sperimentazione di aziende operanti nei settori biomedico, epidemiologico, biologico; in centri ed istituti di ricerca pubblici e privati; in societa' informatiche e di progettazione software nell'ambito dell'analisi dei dati.

### **Caratteristiche della prova finale**

Per conseguire la laurea lo/a studente/ssa deve aver acquisito 120 crediti formativi compresi quelli relativi alla prova finale.

Legenda: Per. = periodo o semestre, Val. = Valutazione (V=voto, G=giudizio), TAF= Tipologia Attività Formativa (A=base, B=caratterizzante, C=Affine, S=stages, D=a scelta, F=altre)

La prova finale, ovvero una tesi scritta che si caratterizza per la sua originalità, ha l'obiettivo di verificare il livello di maturità e la capacità critica del laureando, con riferimento agli apprendimenti e alle conoscenze acquisite, a completamento delle attività previste dell'ordinamento didattico. È ammessa la possibilità di svolgere la modalità "tesi in azienda" integrando tesi ed attività di stage nell'ambito e nelle modalità stabilite anche dai Regolamenti didattici di Ateneo.

Insegnamenti 1 ° anno	CFU	Sem.	Val.	Freq.	SSD	TAF
22216 - BIG DATA C.I.	12	Ann.	V			
- <i>BIG DATA TOOLS</i> <i>La Cascia(PO)</i>	6	1			ING-INF/05	B
- <i>BIG DATA ANALYTICS</i> <i>Pirrone(PO)</i>	6	2			ING-INF/05	B
23844 - DATA AND TEXT MINING C.I.	9	1	V			
- <i>DATA MINING</i> <i>Plaia(PO)</i>	6	1			SECS-S/01	B
- <i>TEXT MINING</i> <i>Plaia(PO)</i>	3	1			SECS-S/01	B
23843 - DATA PROTECTION <i>Piraino(PO)</i>	6	1	V		IUS/01	B
23845 - MACHINE LEARNING AND STATISTICAL MODELS C.I.	12	1	V			
- <i>STATISTICAL MACHINE LEARNING</i> <i>Sottile(RD)</i>	6	1			SECS-S/01	B
- <i>STATISTICAL MODELS</i> <i>Chiodi(PO)</i>	6	1			SECS-S/01	B
23161 - ADVANCED SURVIVAL ANALYSIS <i>Attanasio(PO)</i>	9	2	V		SECS-S/05	C
23815 - CLOUD COMPUTING <i>Gallo(PA)</i>	9	2	V		ING-INF/05	B
Gruppo di attiv. form. opzionali	6					B
	<b>63</b>					

Insegnamenti 2 ° anno	CFU	Sem.	Val.	Freq.	SSD	TAF
23953 - NETWORKS AND HIGH DIMENSIONAL DATA ANALYSIS C.I.	15	1	V			
- <i>NETWORKS AND OPTIMIZATION</i> <i>Tumminello(PO)</i>	9	1			SECS-S/06	C
- <i>STATISTICAL LEARNING FOR HIGH DIMENSIONAL DATA</i> <i>Augugliaro(PO)</i>	6	1			SECS-S/01	B
23941 - PERFORMANCE STATISTICS AND STATISTICAL MODELS IN FINANCE C.I.	9	1	V			
- <i>PERFORMANCE MANAGEMENT STATISTICS</i> <i>Cracolici(PO)</i>	3	1			SECS-S/03	C
- <i>STATISTICAL MODELS FOR VOLATILITY IN FINANCE</i> <i>Vassallo(PA)</i>	6	1			SECS-S/03	C
16047 - THESIS	12	2	G			E
Gruppo di attiv. form. opzionali II	6					C
Gruppo di attiv. form. opzionali III	6					F
Attiv. form. a scelta dello studente (consigliate)	9					D
	<b>57</b>					

## GRUPPI DI ATTIVITA' FORMATIVE OPZIONALI

Gruppo di attiv. form. opzionali	CFU	Sem.	Val.	Freq.	SSD	TAF
21225 - BAYESIAN STATISTICS <i>Abbruzzo(PA)</i>	6	2	V		SECS-S/01	B

Legenda: Per. = periodo o semestre, Val. = Valutazione (V=voto, G=giudizio), TAF= Tipologia Attività Formativa (A=base, B=caratterizzante, C=Affine, S=stages, D=a scelta, F=altre)

## GRUPPI DI ATTIVITA' FORMATIVE OPZIONALI

<b>Gruppo di attiv. form. opzionali</b>	<b>CFU</b>	<b>Sem.</b>	<b>Val.</b>	<b>Freq.</b>	<b>SSD</b>	<b>TAF</b>
22383 - MODELS FOR CATEGORICAL DATA <i>Sciandra(PA)</i>	6	2	V		SECS-S/01	B
21224 - NON PARAMETRIC STATISTICS <i>Muggeo(PO)</i>	6	2	V		SECS-S/01	B
16439 - STOCHASTIC PROCESSES <i>Adelfio(PO)</i>	6	1	V		SECS-S/01	B
<b>Gruppo di attiv. form. opzionali II</b>	<b>CFU</b>	<b>Sem.</b>	<b>Val.</b>	<b>Freq.</b>	<b>SSD</b>	<b>TAF</b>
19220 - CYBERSECURITY	6	1	V		ING-INF/05	C
23828 - DATA SCIENCE FOR MARKETING DECISIONS <i>Cracolici(PO)</i>	6	1	V		SECS-S/03	C
20526 - INTERNET OF THINGS	6	1	V		ING-INF/03	C
15507 - MATHEMATICS FOR ECONOMICS AND FINANCE <i>Consiglio(PO)</i>	6	2	V		SECS-S/06	C
15506 - RISK MANAGEMENT <i>Scannella(PO)</i>	6	1	V		SECS-P/11	C
23952 - SURVEY SAMPLING METHODS IN SOCIAL SCIENCES <i>Attanasio(PO)</i>	6	1	V		SECS-S/05	C
18122 - TOPICS IN MACRO AND FINANCIAL ECONOMETRICS <i>Cipollini(PO)</i>	6	1	V		SECS-P/05	C
<b>Gruppo di attiv. form. opzionali III</b>	<b>CFU</b>	<b>Sem.</b>	<b>Val.</b>	<b>Freq.</b>	<b>SSD</b>	<b>TAF</b>
06634 - STAGE	6	2	G			F
11033 - STAGE 3 CFU	3	2	G			F
23943 - STATISTICAL CONSULTANCY 2	6	2	G			F
23942 - STATISTICAL CONSULTANCY 3	3	2	G			F
<b>Attiv. form. a scelta dello studente (consigliate)</b>	<b>CFU</b>	<b>Sem.</b>	<b>Val.</b>	<b>Freq.</b>	<b>SSD</b>	<b>TAF</b>
08627 - STAGE	3	2	G		SECS-S/01	D
11028 - STAGE 6 CFU	6	2	G		SECS-S/01	D
11029 - STAGE 9 CFU	9	2	G		SECS-S/01	D
23944 - STATISTICAL CONSULTANCY	3	2	G		SECS-S/01	D
23945 - STATISTICAL CONSULTANCY 1	6	2	G		SECS-S/01	D

Legenda: Per. = periodo o semestre, Val. = Valutazione (V=voto, G=giudizio), TAF= Tipologia Attività Formativa (A=base, B=caratterizzante, C=Affine, S=stages, D=a scelta, F=altre)