UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Statistics for The Social Sciences

2627-1-E2006P006

Area di apprendimento

2: Aspetti psicosociali della comunicazione.

Obiettivi formativi

Conoscenza e comprensione

Il corso si occupa della formazione di opinioni e decisioni data driven. In particolare, considera i processi epistemologici che permettono di generare conoscenza dai dati. A questo scopo, il corso si focalizza innanzitutto sulla *statistical literacy* in uno scenario di rapida produzione di informazioni dove l'ipertrofia comunicativa richiede di orientarsi nell''asimmetria delle notizie', discriminando fra dati provvisti o meno di autorevolezza scientifica. Il corso familiarizza gli studenti e le studentesse con l'Istat, ente di statistica ufficiale intesa come bene pubblico al servizio dei cittadini. Di qui ci si estende ad altre fonti ufficiali, parallele e ai Big Data. La messa a punto della metadatazione diviene uno strumento cruciale per la lettura del dato, insieme al sistema di validazione della qualità dei dati, contestualizzandoli nei sistemi statistici sovranazionali, come Eurostat, OCSE, WHO.
Sul piano della metodologia statistica, l'indagine muove dai principali costrutti dello studio del comportamento, un collante interdisciplinare del corso di studi. A partire dalla concettualizzazione dei costrutti, si illustrano come siano 'operazionalizzati per le applicazione nell'indagine empirica, mediante la costruzione di variabili statistiche e misure relative. Il corso illustra l'analisi dei dati nelle due branche disciplinari della statistica descrittiva e dell'inferenza, sottolineando il duplice ruolo della prima *per se* nella comprensione di un collettivo e per la corretta applicazione dei metodi inferenziali.

Capacità di applicare conoscenza e comprensione

Concetti, definizioni, strumenti disciplinari verranno sistematicamente presentati in prospettiva interdisciplinare, così da maturare l'attitudine ad integrare i procedimenti statistici appresi nel corso con fonti informative di diversa

natura. Le competenze si sviluppano in un duplice contesto applicativo. In primo luogo, familiarizzare con l'analisi della raccolta dati e delle misure connesse, nella produzione ufficiale corrente, consente di una corretta lettura del dato, evitando approcci fuorvianti, grazie alla competenza acquisita nel riconosce e ricercare gli elementi chiave di contesti e metodi che hanno generato i dati, grazie all'esercizio di una corretta lettura, un'elaborazione congruente e un'interpretazione critica dei risultati.
In secondo luogo, attraverso i fondamenti metodologici della disciplina, studenti e studentesse sapranno organizzare i dati rilevati, grazie alla conoscenza delle scale di misura. L'operatività connessa ad esse consentirà di ricavare le corrette misure descrittive per l'analisi univariata e bivariata. Le sintesi descrittive saranno inoltre guida alle possibilità inferenziali. Le applicazioni con software open source potenzia la padronanza nell'uso e nel riconoscimento di significato delle differenti misure.

Autonomia di giudizio

La capacità di giudizio autonomo e di ragionamento critico nella comprensione della mole di dati che pervade la letteratura scientifica e l'informazione tout court rappresenta un volano ineludibile per orientarsi nella comunicazione odierna, che si nutre di molteplici linguaggi. Sviluppare sensibilità nel processo generativo dei dati consente a studentesse e studenti una loro corretta lettura, inquadradoli stessi nei contesti di origine, di cogliere le possibilità di rilevazione che ne delimitano il contenuto. L'osservazione diretta di informazioni quantitative da fonti diverse viene sperimentata durante il corso, esercitando il senso critico nell'evitare interpretazioni di superficie, affinando strategie e stili atti ad evitare banalizzazioni. Si estenderà questa prospettiva alle fonti parallele e poi alla comunicazione ad ampio raggio, decodificandone le strategie di presentazione.
Autonomia di giudizio e senso critico saranno espressamente promosse nelle elaborazioni di dati, nela costruzione di indici e indicatori, nella scelta delle misure statistiche più rappresentative declinandole sulle caratteristiche del fenomeno studiato. La modellazione statistica verrà introdotta tramite l'analisi bivariata per le diverse scale, riconoscendo le interpretazioni fallaci quali la correlazione spuria e il paradosso di Simpson. Tutto ciò sarà declinato sia nelle attività in aula o laboratorio informatico sia negli esami di profitto.

Abilità comunicative

L'alfabetizzazione statistica è cruciale nello spettro delle competenze dei laureati nelle scienze della comunicazione, sia in termini di comprensione piena della letteratura scientifica, che poggia ampiamente sulla validazione, sia per veicolare contenuti complessi, di necessità di con strumenti espressivi compositi. Si acquisirà l'abilità di esposizione di dati sintetici in termini nummerici e grafici esprimendone contemporaneamente il correlato empirico, le definizioni e i metadati che ne chiariranno la portata propria ingenue distorsioni interpretative. Si acquisirà dimestichezza nell'esprimere fenomeni statistici in formati grafici, tabellari o verbali, in relazione di finalità e audience. Le attività in aula e gi esercizi individuali, discussi e rivisti nei momenti dedicati, consentiranno di consolidare queste forme comunicative, che saranno oggetto di valutazione nelle prove d'esame.

Capacità di apprendimento

La struttura didattica del corso si inserisce nella prospettiva del corso di laurea, tesa a promuovere uno stile di apprendimento che integra le diverse discipline in una formazione culturale multidimensionale, tesa all'accrescimento delle conoscenze e delle competenze utili per affrontare la complessità della comunicazione contemporanea. . La prosecuzione della formazione accademica si indirizza ai sistemi informativi per l’informazione e alle discipline umanistiche evidence-based. L’autonomia e la sedimentazione delle capacità di apprendimento verranno seguite nello svolgimento di attività individuali e di gruppo, per essere quindi accertate nelle prove finali.

Contenuti sintetici

Il corso fornisce gli strumenti teorici ed analitici per l’accesso e la rielaborazione delle informazioni quantitative e quali-quantitative. I metodi di generazione dei dati sono affrontati in con riferimento alle statistiche ufficiali nazionali

come parte della rete Eurostat, quindi estese alle fonti parallele e ai Big Data. L’operazionalizzazione dei concetti e dei fenomeni in variabili, il riconoscimento dei relativi metodi di misura sono illustrati con riferimento a rilevazioni su temi multidisciplinari, privilegiando l'interdisciplinarietà nel corso di laurea.
Si promuove l’acquisizione autonoma e personale di adeguate chiavi di lettura e di linee guida nella comprensione dell’informazione e del ragionamento statistico-computazionale, sia nell’apprendimento di tecniche quantitative di base, sia nell’accesso ai risultati di indagini e alla loro disseminazione tramite i diversi media. La distinzione fra statistica descrittiva e statistica inferenziale include rudimenti di inferenza da esperimento statistico. Inoltre, si affrontano tecniche analitiche recenti quali dati testuali e reti sociali.

Programma esteso

- La terminologia statistica.
- Tipi di dati.
- Classificazione delle fonti statistiche.
- Le scale di misura.
- La raccolta dei dati.
- Le fonti statistiche ufficiali. Tipologie di indagini.
- La qualità dei dati: concetti e definizioni.
- Strategie di campionamento. Alcuni problemi di campionamento nella ricerca sociale.
- Aspetti di base della costruzione e somministrazione del questionario.
- Cenni alle tecniche di scaling.
- Visualizzazione dei dati.
- Misure di sintesi della distribuzione. Le medie di posizione e analitiche. La variabilità.
- La forma della distribuzione. Asimmetria e curtosi.
- Confronti fra grandezze. Rapporti e indici. Indicatori compositi.
- Elementi di probabilità. Distribuzione normale come distribuzione notevole.
- Verifica di ipotesi per l’eguaglianza delle medie.
- Analisi bivariata: contingenza, cograduazione e correlazione.
- La verifica di ipotesi nell’analisi bivariata.
- Cenni a text analitics e reti sociali.

Prerequisiti

Matematica: competenze algebriche comuni alle scuole secondarie di primo grado.
Informatica: si richiedono le competenze stabilite dal corso di Informatica (corso obbligatorio del primo semestre).
La revisione delle conoscenze di base, informatiche e matematiche, consentirà di attivare attività di recupero, ove necessario.

Metodi didattici

Insegnamento in blended learning con differenti modalità didattiche 8 CFU - 60 ore :

- 5 lezioni da 2 ore in presenza in modalità mista, erogativa per la presentazione di argomenti introduttivi e interattiva per v erificare la comprensione degli argomenti;
  -6 lezioni da 2 ore da remoto in modalità erogativa sui concetti più complessi della disciplina;
- 10 lezioni da 2 ore svolte in modalità interattiva da remoto, con applicazioni e con analisi dei siti di statistica

ufficiale;
- 9 esercitazioni con applicazioni informatiche da 2 ore ciascuna, per tre gruppi distinti di studentesse e studenti, in modalità interattiva in presenza in aula informatizzata.

Il corso si articola in lezioni teoriche sulla metodologia statistica, esercitazioni applicative ed attività on line.
La metodologia statistica si sviluppa secondo due moduli tematici.
Modulo 1: Fonti di dati, Statistica ufficiale, e Qualità dei dati
Il primo modulo tematico affronta le questioni metodologiche pertinenti i dati disponibili, a partire da quelli ufficiali. Vengono esplorati i siti dell'Istat nella rete dell'Eurostat, articolati in diversi ambiti disciplinari. Si risale alla documentazione on line relativa, con particolare attenzione alla metadatazione e alla qualità dei dati. Questo modulo si svolge in attività on line, nella forma di lezioni frontali in streaming per l'inquadramento metodologico, di attività hands-on con assistenza della docente e del/la tutor in merito all'esplorazione dei siti ufficiali.
Modulo 2: Statistica di base e analisi applicativa
Il secondo modulo riguarda le conoscenze teoriche, esposte privilegiando la comprensione dei concetti e la congruenza del ragionamento analitico, con particolare attenzione all'alle specificità dell'indagine psico-sociale. Le esercitazioni nelle aule informatizzate traducono sistematicamente i concetti metodologici nel correlato applicativo, così da contestualizzare le informazioni acquisite, mediante l'uso di analitici preferibilmente open source per il computo statistico. Le esercitazioni in aule informatizzate sono tutte frontali in presenza, così come le lezioni teoriche di inizio e fine corso, mentre le rimanenti lezioni sono frontali da remoto.
Attività online e lavoro di gruppo
Una sezione dell'attività on line consiste in attività di gruppo, esercizi e simulazioni ad accesso libero, predisposte sulla piattaforma didattica, atte a a compendiare i due moduli, collocando le misure apprese a livello metodologico nel quadro dell'informazione statistica fruibile correntemente. Sotto il profilo didattico, l'acquisizione di questa competenza valorizza la discussione dell'informazione numerica fornita dai media, anche nel confronto con le statistiche ufficiali. La discussione avviene in gruppi in base alle tematiche di interesse accademico.

Modalità di verifica dell'apprendimento

La prova finale è solo scritta e si articola in un project work e in un esame su piattaforma Moodle. Non sono previste prove *in itinere*.

Project work
Il project work consiste nell'analisi critica di un fenomeno o di un tema che possa essere studiato attraverso dati statistici prodotti dagli enti ufficiali competenti.
Gli studenti dovranno individuare e utilizzare fonti statistiche ufficiali, analizzandone i metodi di raccolta dei dati, i metadati e le definizioni adottate per delimitare e descrivere il fenomeno osservato. Le informazioni ricavate verranno poi confrontate criticamente con il modo in cui lo stesso tema è presentato nei media e nella comunicazione corrente.
Attraverso il progetto saranno valutate le capacità di:

- comprendere e applicare i principi fondamentali della statistica descrittiva;
- interpretare correttamente dati e informazioni quantitative alla luce dei metodi di rilevazione utilizzati;
- riconoscere punti di forza e limiti delle fonti informative;
- individuare le modalità più efficaci di rappresentazione e comunicazione dei dati;
- sviluppare una lettura critica dell'informazione statistica diffusa dai media.
  Oltre alla corretta applicazione degli aspetti metodologici, agli studenti è richiesta una riflessione originale sul fenomeno analizzato e una valutazione critica dell'utilizzo dei dati ufficiali nella comunicazione pubblica e giornalistica.
  Il project work contribuisce alla valutazione finale nella misura del 35% del voto complessivo.

Prova informatizzata su Moodle
La prova informatizzata, svolta sulla piattaforma Moodle, consiste nell'analisi statistica di un file di dati che simula

una rilevazione reale.
L'esame verifica la capacità di applicare i principali strumenti della statistica descrittiva e le nozioni introduttive di statistica inferenziale, utilizzando software per l'analisi dei dati.
La prova comprende:

- quesiti a risposta multipla;
- domande Vero/Falso;
- esercizi con risposta chiusa;
- produzione e interpretazione di grafici;
- analisi e interpretazione di risultati statistici.
  Nella parte finale sono previste alcune domande metodologiche a risposta multipla volte a verificare la capacità di individuare il metodo statistico più appropriato rispetto al problema proposto.
  L'obiettivo della prova è accertare la capacità di:
- riconoscere le caratteristiche delle variabili e delle misure statistiche;
- scegliere le procedure di analisi più adeguate;
- interpretare correttamente i risultati ottenuti;
- utilizzare strumenti statistico-informatici per la soluzione di problemi applicativi.
  La prova informatizzata contribuisce alla valutazione finale nella misura del 65% del voto complessivo.

Testi di riferimento

- Iezzi, Domenica Fioredistella, (2024) *Dai dati alla conoscenza. Statistica per le decisioni*. Roma: Carocci.
- Sitologia e materiale uteriore indicato dalla docente.

Gli studenti/le studentesse Erasmus possono contattare il/la docente per concordare la possibilità di studiare su una bibliografia in lingua inglese e/o la possibilità di sostenere l'esame in inglese. Il libro di riferimento, in lingua inglese, è: Howitt, D. Cramer, D. Introduction to statistics in psychology, 5th ed. Harlow : Pearson, 2011, disponibile nella Biblioteca di Ateneo.

Sustainable Development Goals

PARITÁ DI GENERE | PACE, GIUSTIZIA E ISTITUZIONI SOLIDE