



Convegno AlmaLaurea 2022
Integrazione dei dati e potere informativo
Dalla formazione al mondo del lavoro

**«Condividere i dati in una visione coordinata:
una miniera di opportunità»**

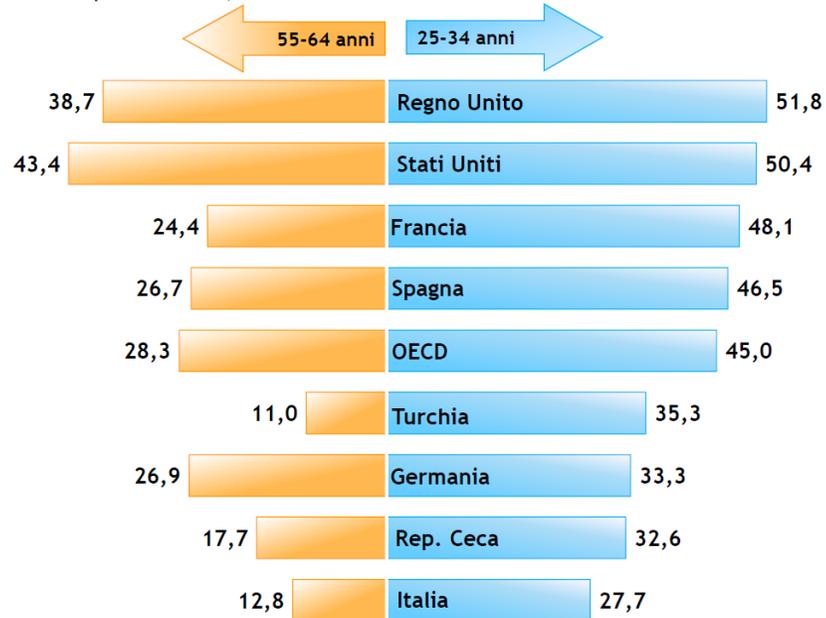
Stefania Mignani

Dipartimento di Scienze Statistiche, Università di Bologna

stefania.mignani@unibo.it

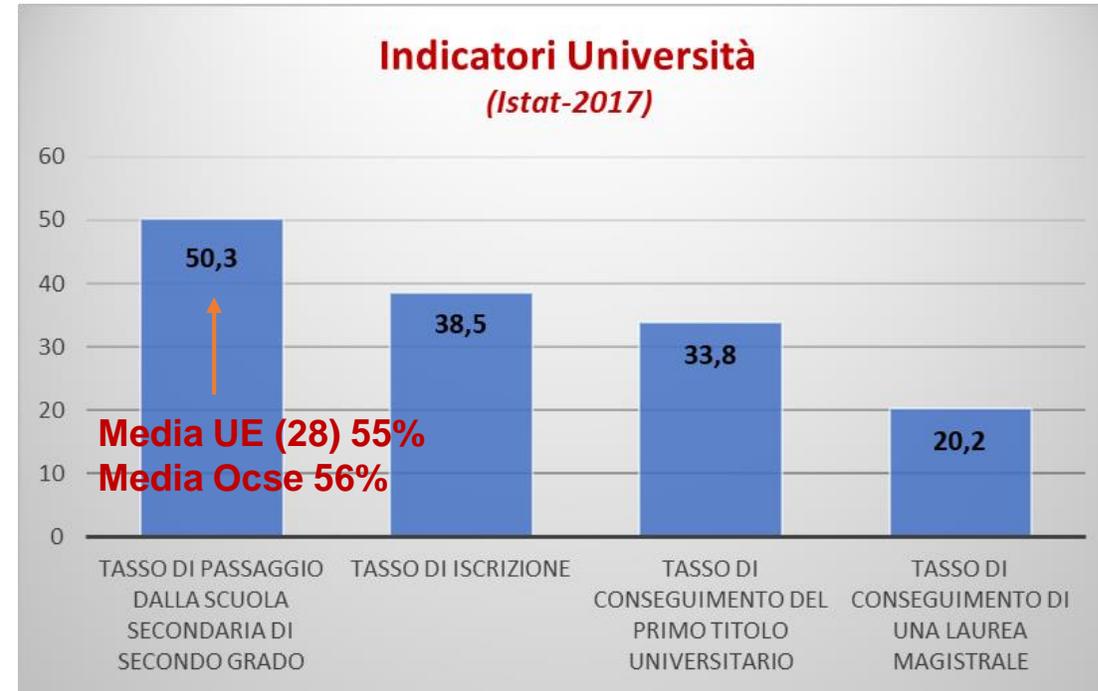


Figura 1.1 Popolazione di 25-34 e 55-64 anni con istruzione universitaria in alcuni Paesi OECD. Anno 2019 (valori percentuali)



Fonte: OECD, 2020.

Da Rapporto Almalaura 2021



- Circa 20% lasciano dopo un anno di corso → **1 su 5**
- Circa 40% abbandoni nei primi due anni
- solo il 22,3% dei giovani tra i 25 e i 34 anni porta a compimento gli studi accademici, contro il 35% del resto d'Europa



Punti per la discussione

1) Banche dati diverse Istat, Oecd, Eurostat, Miur, Almalaurea,...

→integrare questi dati

2) Come lavorare su questi dati possa aiutare a comprendere meglio il fenomeno dell'abbandono all'Università

Spiegazioni trovate nei dati ma come gestire questi dati??

- Continuità delle informazione sullo studente → Biografia formativa/educativa
- Dati di trend per il monitoraggio
- Problematiche: far parlare le banche dati, anonimizzazione, indicatori coerenti ed omogenei,...



Dallo scambio di dati alla cultura del dato: trasformare le informazioni in conoscenza utile per le decisioni

- Il sistema istruzione e formazione è **multidimensionale e altamente interconnesso**.
- **una sola fonte** di dati permette l'analisi di un solo aspetto di un problema, non riuscendo a far emergere connessioni/correlazioni e cause congiunte del fenomeno nella sua interezza.
- Incentivare forme di condivisione dei dati → la condivisione dei dati, la loro integrazione e il loro riutilizzo producono valore, capacità di scelta e di innovazione
- Estrarre informazione e valore dai dati analizzandoli secondo schemi di indagine **diversi da quelli per i quali sono stati raccolti**, rispondendo alle domande attraverso il riconoscimento di modelli emergenti dal comportamento dei dati, usando strumenti e algoritmi e forme di visualizzazione in pieno sviluppo.
- **Forme di innovazione guidata dai dati (approccio *data driven*)**



Alcuni punti centrali

- Fonti istituzionali e fonti non convenzionali
- Uscire dall'ottica che i dati sono di chi li produce

Gestione e raccordo

- Punto di vista informatico definito il codice di linkage (entity integrity)
- Punto di vista delle analisi statistiche **BIG-DATA**

Riservatezza dei dati

- anonimizzazione secondo un codice condiviso
- dati resi pubblici in forma aggregata **OPEN DATA**

Problematiche aperte:

- Dati raccolti da un punto di vista amministrativo invece usati a fini statistici → preparazione del dato e un coordinamento forte tra i «produttori» per **rendere omogeni alcuni indicatori** e permettere l'analisi congiunta (Esempio indicatore del back-ground socio economico della famiglia)



Alcuni punti centrali

- ogni base di dati deve essere memorizzata mediante un **supporto tecnico** adeguato;
- l'identificazione delle entità primarie dei vari sistemi informativi (studenti, scuole, ecc.) e la loro codifica in maniera coerente, al fine di **rendere possibile un collegamento tra i diversi archivi**;
- aumento della **qualità complessiva dei dati**, garantendo tempestività degli aggiornamenti, eliminazione delle ridondanze e riconciliazione dei dati presenti in diversi archivi;
- la possibilità di effettuare **analisi di tipo predittivo** che possano supportare in maniera sostanziale e tempestiva le politiche governative;
- il rilascio di parte dei dati e delle analisi in **modalità aperta** (e linked).
- l'accesso controllato al micro dato



Una possibile analisi trasversale Le carriere degli studenti → orientamento consapevole e abbandono

- PROBLEMA: numero dei laureati è basso.
- Si iscrivono in pochi e si laureano ancora meno
- Dispersione e drop-out

Il fenomeno noto come “dispersione universitaria”

- l’irregolarità generica nell’acquisizione dei crediti formativi,
- il prolungamento della permanenza nell’università (c.d. *fuoricorsismo*),
- la mancanza di linearità della carriera (ad es. passaggi di corso)
- l’abbandono vero e proprio del percorso di apprendimento che porta quindi all’uscita dal sistema universitario senza il conseguimento del titolo.



Orientamento in entrata all'Università

Orientamento in Ingresso all'Università: sistema di verifica delle conoscenze in ingresso, coordinato a livello nazionale per gruppi affini di corsi di studio, che comunica con le procedure d'ingresso ai corsi a numero programmato, e che fornisce ulteriori informazioni per l'orientamento.

→ **Orientamento** che si focalizza sulle competenze

- Voto di Diploma ottimo predittore ma non sembra bastare;
- studi hanno confermato l'esistenza di una forte correlazione fra risultati scadenti nella verifica delle conoscenze iniziali e risultati scadenti (o fallimentari) nel corso di studio;
- Analisi in corsi di studio di area scientifica, di Ingegneria, Architettura ed Economia (grazie al CISIA, a Con.Scienze e al Piano Lauree Scientifiche) dimostrandone la fattibilità e l'utilità
- Anche studi sugli OFA

- Integrare esiti scolastici e performance universitarie → Dati ANS, Miur, Università, CISIA, Invalsi,...

POTERE INFORMATIVO DEI DATI → Open Data → Open Science



Valore predittivo test TOLC

Analisi sui risultati di studenti iscritti ad una laurea di area economica

Anagrafe Nazionale Studenti (ANS)

Consorzio Interuniversitario Sistemi Integrati per l'Access (CISIA)

Coefficienti	OR	Stime	Errore standard
Intercetta	4.57	1.52	0.08
Genere=Maschio	0.85	-0.16	0.08
Voto diploma	1.05	0.05	0.000
Anno nascita < 1992	0.51	-0.67	0.09
Logica	1.08	0.08	0.04
Compr. Testo	1.03	0.03	0.04
Matematica	1.28	0.25	0.04
AL	0.74	-0.30	0.15
GE	0.82	-0.20	0.16
IP	0.24	-1.42	0.17
IT	0.47	-0.75	0.14
LC	0.90	-0.10	0.15
MP	0.84	-0.18	0.24
TC	0.63	-0.46	0.1

Totale studenti: 5848

Match: 3983 (68%)

Modello predittivo: l'aver acquisito più di 20 CFU entro la fine del primo anno

Massimo Attanasio, Marco Enea, & Vincenzo Giuseppe Genova (2020). Analisi della predittività dei test d'accesso per Economia. In Il test di ingresso: valenza predittiva e rilevanza nei percorsi di orientamento universitario. Le attività del CISIA al servizio degli Atenei



Analisi sui risultati degli OFA come fattore di rischio per il successo universitario

- Dati messi a disposizione dell'Ateneo, ma provenienti da diverse fonti:(segreterie studenti, Datawarehouse, coordinatori Corso di studio, Dipartimento).
- Obiettivo: capire se gli studenti con OFA sono coloro che hanno maggiori difficoltà nel superare brillantemente gli esami.
- Studente di successo" ha sostenuto in media un terzo dei crediti acquisibili nel primo semestre e conseguito un voto medio ponderato di circa 22=30.

		Estimate	Std. Error	z value	Pr(> z)
	β_0	-3.612	1.352	-2.671	0.008
Voto_maturità	β_1	0.069	0.018	3.883	0.000
Punteggio OFA	β_2	-1.463	0.379	-3.865	0.000

S. Bianconcini e C.Trivisano Dipartimento di Scienze statistiche nell'ambito del PLS



In conclusione...

POTERE INFORMATIVO DEI DATI →

Open Data →

Open Science