

**Laurea Magistrale in Data Science**

A.A 2025-2026

**VISUALIZZAZIONE DEI DATI, VISUAL ANALYTICS  
E REPORTING**

**Analisi degli indicatori  
ANVUR sui CdS**

Studente: Anisa Bakiu  
Matricola: 837632

# Riepilogo del lavoro

## Indicatori ANVUR dei Corsi di studio e Indicatori ANVUR AVA 3

Gli indicatori ANVUR del sistema universitario hanno la finalità di fornire una rappresentazione trasparente e comparabile delle università italiane. Essi consentono di monitorare l'evoluzione del sistema in termini di personale, offerta formativa e risultati della ricerca.

Il confronto degli indicatori presenti in questo dataset è effettuato tra i corsi di studio e quelli appartenenti alla stessa classe nel nostro Ateneo, nel Sud Italia e a livello nazionale.

I dati messi a confronto sono medi, ciò significa che per l'Italia e per il Sud, è disponibile un valore medio aggregato.

Dopo la fase di preprocessing del dataset, finalizzata alla selezione degli indicatori per le analisi, questi vengono confrontati mediante tecniche di visualizzazione, al fine di individuare gli aspetti dei nostri corsi che non seguono il trend nazionale oppure che presentano caratteristiche specifiche all'interno dell'Università degli Studi di Bari.

### Gli indicatori ANVUR per corso di studio

Fonte: <https://www.uniba.it/it/ateneo/ateneo-in-cifre/indicatori/indicatori-anvur-per-corso-di-studio>

# Obiettivi di analisi

Tramite queste analisi, il mio obiettivo è estrarre pattern e regole specifiche nascoste nei dati e portarle in superficie, senza ricorrere a tecniche complesse di machine learning, che richiederebbero maggior tempo e risorse per ottenere informazioni che possono invece essere rese trasparenti direttamente attraverso le tecniche di visualizzazione studiate.

Il dataset introdotto in precedenza è ricco di feature e comprende dati di diversa dimensionalità. Anche se alcune variabili sono fornite sotto forma di medie o aggregazioni di altri dati, esso rimane comunque caratterizzato da un'elevata varietà di variabili che possono essere combinate e messe in relazione tra loro.

Questo lavoro è stato guidato dalle regole e dai mantra più noti della visualizzazione dei dati, presentati e discussi durante le lezioni. Per maggiore precisione, di seguito ne riporto alcuni.

## Workflow:

Inizialmente è stata condotta un'analisi di pre-processing dei dati, finalizzata alla trasformazione delle informazioni in un formato coerente per le successive visualizzazioni in Tableau e D3.js.

Tableau è stato utilizzato per effettuare le visualizzazioni secondo il Visual Information-Seeking Mantra di Shneiderman.

Alla fine, viene aggiunto un confronto tra le tecniche di Tidy tree e Space tree effettuato in Observable.

## Tools:



## Observable

# Obiettivi di analisi

Partiamo dalle regole di **Tufte sull'eccellenza grafica**.

Secondo Edward Tufte, l'eccellenza grafica consiste nella comunicazione di idee complesse con chiarezza, precisione ed efficienza.

*Praticamente, un'eccellente rappresentazione grafica dovrebbe:*

1. Mostrare i dati, e possibilmente ridurre elementi decorativi non informativi
2. Indurre l'osservatore a riflettere sul contenuto sostanziale, piuttosto che sul design grafico o sulla tecnologia utilizzata
3. Evitare la distorsione del messaggio informativo contenuto nei dati
4. Presentare un elevato numero di valori in uno spazio ridotto, mantenendo la leggibilità
5. Rendere coerenti grandi insiemi di dati, facilitandone l'interpretazione complessiva
6. Incoraggiare il confronto visivo tra diverse parti dei dati
7. Rivelare i dati a diversi livelli di dettaglio, permettendo sia una visione d'insieme sia un'analisi approfondita
8. Servire uno scopo chiaro e ben definito, come descrizione, esplorazione, tabulazione o decorazione in casi limitati.
9. Essere strettamente integrata con le descrizioni statistiche e testuali del dataset analizzato.

Un'altra linea guida è il mantra di Shneiderman per la ricerca di informazioni:



1. La panoramica (Overview) è la forma iniziale del grafico. Il suo scopo non è mostrare tutto (cosa impossibile), ma fornire una visione "macro" di tutti i dati. La panoramica è una mappa che guida l'esplorazione del lettore.
2. Lo zoom e il filter, sono metodi per selezionare ciò che viene mostrato e concentrarsi su un argomento di interesse.
3. Dettagli su richiesta, consente al lettore di estrarre valori esatti dal grafico, anziché limitarsi ad approssimazioni visive.

# Pre-processing del dataset

Il dataset scelto presenta un'elevata complessità nella fase di pre-trattamento per diversi motivi. La sua alta dimensionalità, combinata con la presenza di particolari feature spesso apparentemente duplicate ma con sottili differenze tra loro, rende il compito di visualizzazione particolarmente complesso, sia per quanto riguarda la scelta delle tecniche più adatte, sia per la necessità di formulare ipotesi corrette e selezionare le variabili più appropriate al fine di comunicare in modo efficace il messaggio agli utenti.

Questo aspetto motiva anche la scelta di Python per la fase di preprocessing, poiché questo linguaggio di programmazione consente di effettuare operazioni di trasformazione, rotazione, suddivisione (splitting) e utilizzo di espressioni regolari in modo estremamente granulare e flessibile. Ciò rappresenta un vantaggio significativo rispetto a strumenti che offrono blocchi di operazioni di preprocessing pre-compilati, i quali nel caso di dataset reali, risultano spesso limitanti.

Il dataset inizialmente conteneva le seguenti colonne:

```
[ 'ID_ANNO_ACCADEMICO' ,  
  'NOME_CORSO' ,  
  'DIPARTIMENTO' ,  
  'CODICE' ,  
  'DESCRIZIONE' ,  
  'DURATA' ,  
  'CLASSE DI LAUREA' ,  
  'NUMERATORE_CDS' ,  
  'DENOMINATORE_CDS' ,  
  'INDICATORE_CDS' ,  
  'NUMERATORE_ITA' ,  
  'DENOMINATORE_ITA' ,  
  'INDICATORE_ITA' ,  
  'NUMERATORE_SUD' ,  
  'DENOMINATORE_SUD' ,  
  'INDICATORE_SUD' ] .
```

È da notare la presenza sia del numeratore sia del denominatore per ciascun indicatore. Per ridurre la ridondanza informativa, si è scelto di considerare esclusivamente la colonna “Indicatore”, che rappresenta il rapporto tra le due colonne. La colonna “Descrizione” è la più ricca di informazioni riguardanti diversi aspetti dei corsi. L’obiettivo è estrarre da questa colonna le feature che verranno successivamente analizzate per ciascuna classe di laurea. Tuttavia, trattandosi di una colonna particolarmente verbosa, è necessaria una fase preliminare di preprocessing, durante la quale i nomi delle feature vengono standardizzati tramite un mapping (si veda il Notebook).

# Pre-processing del dataset

Inoltre, il dataset fornisce informazioni relative al periodo compreso tra il 2020 e il 2024. Al fine di produrre visualizzazioni attuali e comparabili tra i diversi corsi, si è scelto di considerare esclusivamente l'anno più recente, ovvero il 2024.

## Trasformazione in formato Wide

Fino a questo punto, il dataset è strutturato in formato long. Per affrontare le problematiche legate all'elevata dimensionalità, è stata effettuata una trasformazione in formato wide. In pratica, è stata costruita una tabella pivot nella quale le feature estratte dalla colonna originale “Descrizione” vengono aggiunte come nuove colonne. A seguito di questa trasformazione, il dataset risulta composto da 405 righe e 52 colonne.

Non tutte le feature ottenute sono rilevanti ai fini dell'analisi, pertanto, viene considerato un sottoinsieme di variabili, selezionato in modo da essere rappresentativo di specifiche categorie di interesse.

## # Studenti

Feature	Spiegazione
Iscritti_totali	Iscritti totali in L, LM, LMCU

# Pre-processing del dataset

# Laureati

Feature	Spiegazione
Laureati_in_tempo	Laureati (L LM LMCU) entro la durata normale del corso
Laureati_totali	Laureati (L LM LMCU)

# Occupazione

Feature	Spiegazione
Percentuale_Occupazione_3anni_LM	Percentuale di Laureati occupati a tre anni dal Titolo (LM LMCU) - Laureati che dichiarano di svolgere un'attività lavorativa o di formazione retribuita (es. dottorato con borsa, specializzazione in medicina, ecc.)

# Pre-processing del dataset

# Prosecuzione studi

Feature	Spiegazione
Percentuale_Proseguono_II_anno	Percentuale di studenti che proseguono nel II anno nello stesso corso di studio**

# Docenti / qualità

Feature	Spiegazione
Ore_docenza_indeterrminato	Ore di docenza erogata da docenti assunti a tempo indeterminato sul totale delle ore di docenza erogata
Qualita_ricerca_QRDLM	Valori dell'indicatore di Qualità della ricerca dei docenti per le lauree magistrali (QRDLM) (valore di riferimento: 0,8)
Percentuale_Soddisfazione_studenti	Percentuale di laureandi complessivamente soddisfatti del CdS

# Pre-processing del dataset

Prima di proseguire con le analisi, è opportuno porsi una domanda:

“Ha senso includere nella stessa analisi comparativa sia le lauree scientifiche sia lauree appartenenti ad ambiti umanistici o economici, e trarre conclusioni in questo modo?”



La risposta è chiaramente negativa. Non è metodologicamente corretto presentare all’utente un confronto diretto tra lauree scientifiche (o secondo la definizione di origine americana, STEM) e lauree non-STEM, per poi derivarne conclusioni che non producono nuova conoscenza.

Ad esempio, ci si potrebbe aspettare che i corsi non-STEM presentino tassi di abbandono inferiori e una percentuale più elevata di studenti che completano il percorso di studi nei tempi previsti.

Tuttavia, ciò non implica una qualità superiore rispetto ai corsi STEM, che per loro natura sono generalmente più complessi e meno accessibili a una platea ampia di studenti.

Per questo motivo, è stata introdotta una nuova colonna denominata “Categoria\_laurea”, che consente di distinguere i corsi in base a una classificazione ufficiale pubblicata.

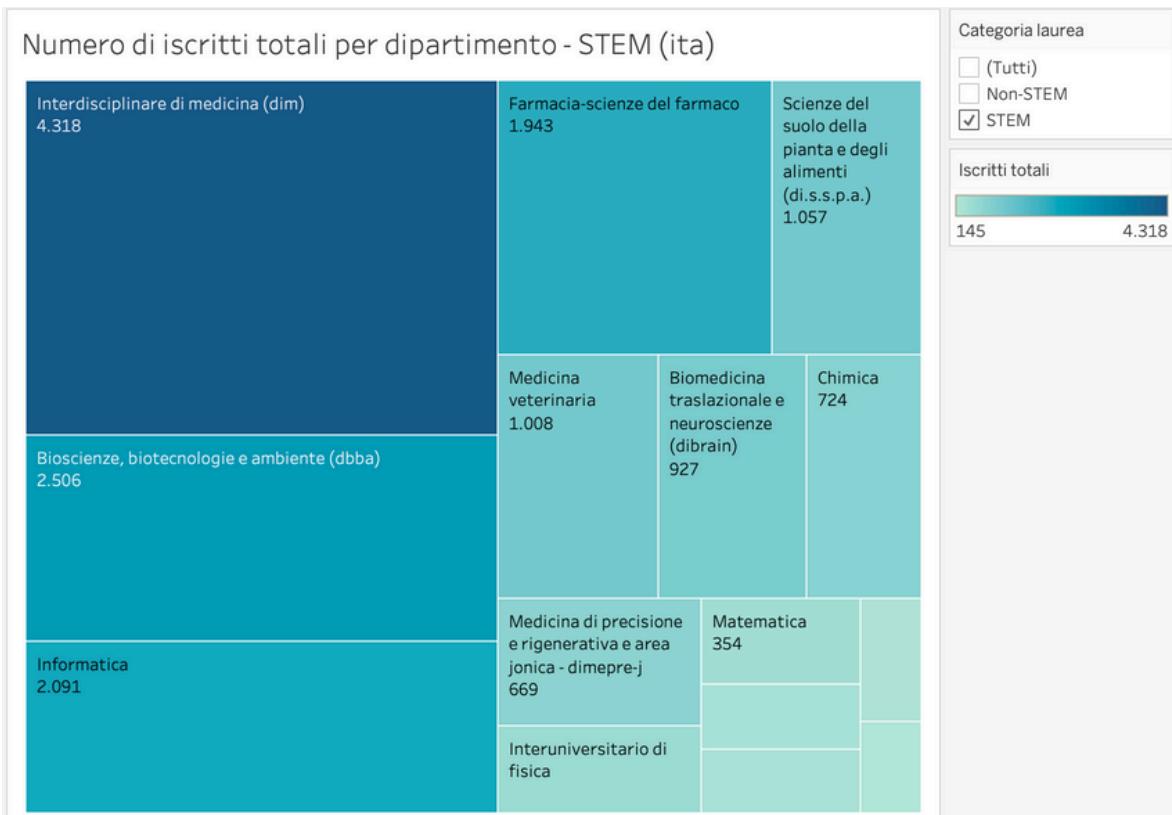
(si veda: <https://www.assolombarda.it/centro-studi/classi-steam>).

Un’ulteriore distinzione è stata effettuata tra corsi di laurea triennale e corsi di laurea magistrale.

# Visualizzazioni Tableau

## Categoria - STEM

RQ1: Come si presenta la situazione in Italia per quanto riguarda il numero totale di iscritti nei dipartimenti che offrono corsi STEM?



Per rispondere a questa domanda è stata adottata la tecnica della treemap. Questa visualizzazione consente di rappresentare la distribuzione degli iscritti tra i diversi dipartimenti, interpretando l'intera superficie come un'unità complessiva e ciascun rettangolo come una frazione proporzionale al numero di iscritti.

L'utilizzo del treemap è motivato dal framework di visualizzazione proposto da Shneiderman (1992), che introduce questa tecnica come metodo efficace per la rappresentazione di dati gerarchici mediante un approccio space-filling, consentendo una visione d'insieme immediata delle proporzioni relative tra gli elementi e supportando il principio di overview first, zoom and filter, details on demand.

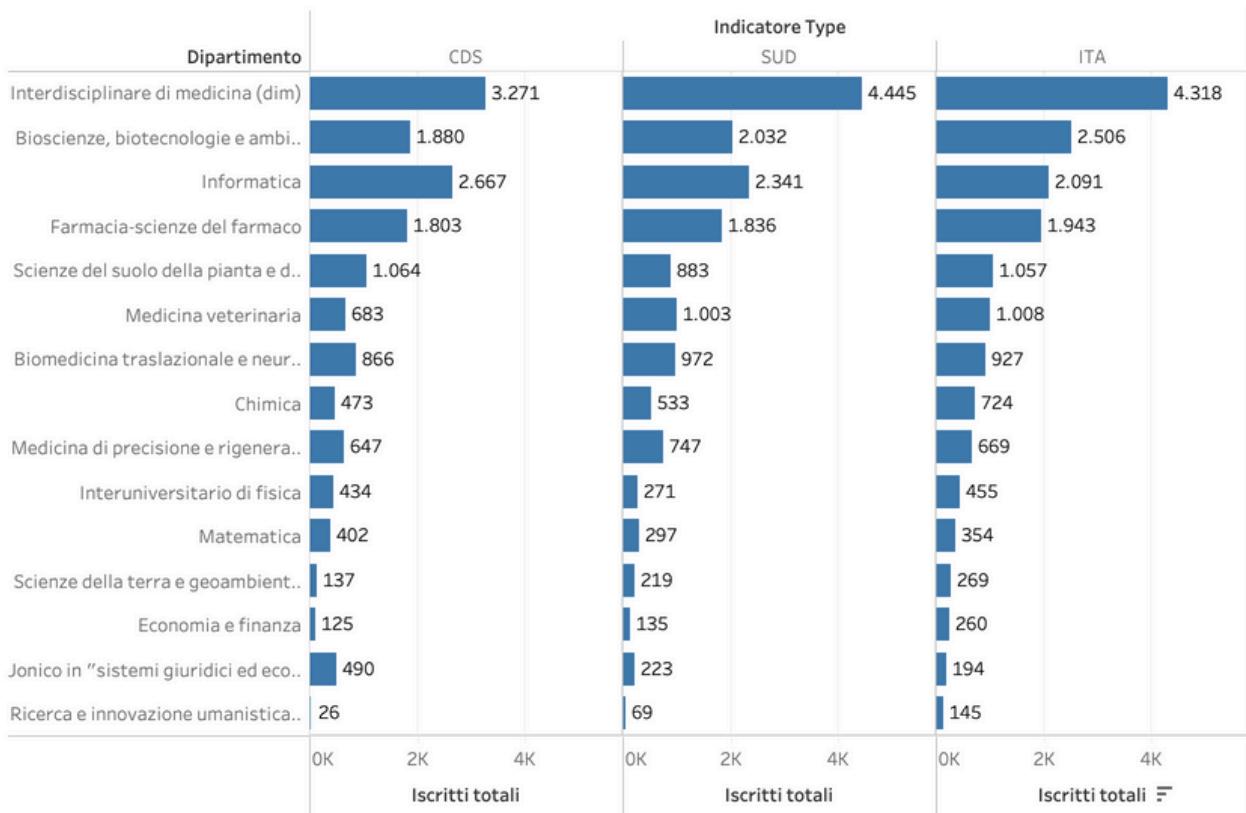
La dimensione dei rettangoli codifica il numero di iscritti, mentre l'intensità del colore fornisce un'ulteriore indicazione della concentrazione. A supporto della lettura, sono state inoltre aggiunte etichette numeriche per rendere esplicativi i valori rappresentati. Dalla visualizzazione emerge chiaramente una predominanza dei dipartimenti di area medica, bioscientifica e informatica, che concentrano la quota più elevata di iscritti rispetto agli altri ambiti STEM.

# Visualizzazioni Tableau

## Categoria - STEM

RQ2: Come cambia questa distribuzione nel Sud Italia e nel contesto di riferimento analizzato?

Confronto degli iscritti totali in base al tipo dell'Indicatore - STEM



Per rispondere a questa domanda sono stati utilizzati bar chart orizzontali, accompagnati da etichette numeriche, in modo da confrontare direttamente le differenze tra le tre realtà presenti nel database. Questa tipologia di grafico è stata scelta poiché permette una lettura immediata delle variazioni nei valori e soprattutto, delle istanze che non seguono il trend nazionale.

Dall'analisi emerge che il Dipartimento di Bioscienze e Biotecnologie e Economia e Finanza registra un numero di studenti iscritti inferiore ai valori di riferimento. Anche il Dipartimento di Scienze della Terra e Geoambientali registra quasi la metà degli iscritti rispetto al dato nazionale e si discosta significativamente anche dal dato del Sud. La stessa situazione si osserva per Medicina Veterinaria.

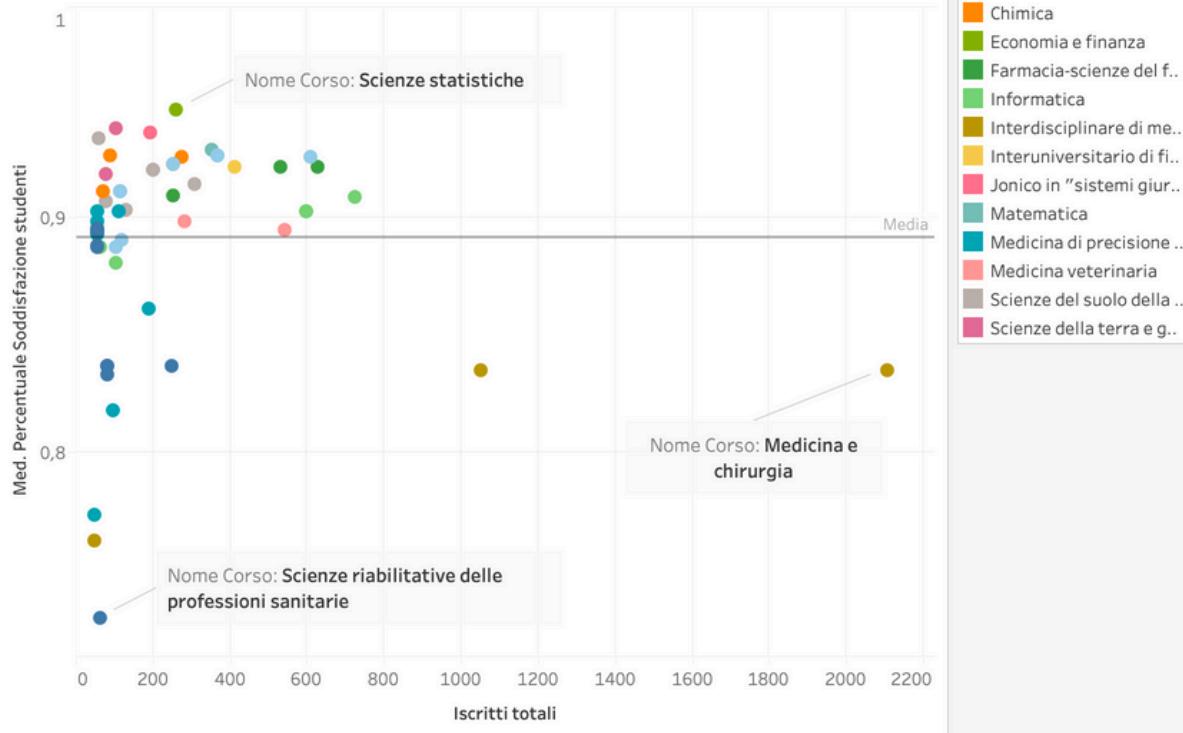
D'altra parte, i Dipartimenti di Informatica e il Dipartimento Jonico in Sistemi Giuridici ed Economici per la laurea in Scienze Marittime mostrano valori molto più alti rispetto ai valori di riferimento.

# Visualizzazioni Tableau

## Categoria - STEM

RQ3: Esiste una relazione tra il numero di iscritti e la percentuale di soddisfazione degli studenti?

Livello di soddisfazione degli studenti e numero totale degli iscritti - STEM (ita)



Dal punto di vista teorico, ci si potrebbe aspettare l'esistenza di una relazione tra queste due variabili. Lo scatter plot è utilizzato come strumento di analisi esplorativa dei dati secondo l'approccio dell'Exploratory Data Analysis proposto da Tukey (1977), in quanto consente di individuare relazioni, trend, cluster e valori anomali tra coppie di variabili quantitative.

Tuttavia, dall'osservazione dello scatter plot emerge che la soddisfazione degli studenti non rappresenta necessariamente un indicatore chiave. La distribuzione dei punti suggerisce infatti la presenza di altri fattori latenti che influenzano il fenomeno. In particolare, il corso di Medicina e Chirurgia, pur avendo un numero elevato di iscritti, non registra un livello alto di soddisfazione; Scienze Riabilitative delle Professioni Sanitarie mostrano sia la soddisfazione più bassa sia, che il numero più basso di iscritti; mentre Scienze Statistiche presentano la soddisfazione più alta tra i corsi analizzati.

# Visualizzazioni Tableau

## Categoria - STEM

RQ4: È possibile scendere maggiormente nel dettaglio e individuare quali dipartimenti presentano livelli di soddisfazione significativamente distanti dalla media nazionale (ITA) e da quella del Sud Italia (SUD)?

Dipartimento	Indicatore Type			MEDIA(Percentuale Sod... 0,7248 0,9660
	CDS	SUD	ITA	
Economia e finanza	0,9475	0,9660	0,9499	
Jonico in "sistemi giuridici ed economici del mediterraneo: società ambiente culture"	0,9402	0,9402	0,9391	
Matematica	0,9149	0,9344	0,9307	
Scienze della terra e geoambientali	0,9469	0,9190	0,9305	
Interuniversitario di fisica	0,9492	0,9149	0,9231	
Chimica	0,8704	0,9081	0,9230	
Farmacia-scienze del farmaco	0,9158	0,9174	0,9182	
Scienze del suolo della pianta e degli alimenti (di.s.s.p.a.)	0,9305	0,9353	0,9170	
Bioscienze, biotecnologie e ambiente (dbba)	0,9204	0,9238	0,9132	
Informatica	0,8052	0,8927	0,8981	
Medicina veterinaria	0,9063	0,9028	0,8964	
Medicina di precisione e rigenerativa e area jonica - dimepre-j	0,7285	0,8700	0,8676	
Biomedicina traslazionale e neuroscienze (dibrain)	0,7975	0,8523	0,8514	
Interdisciplinare di medicina (dim)	0,7248	0,8457	0,8320	

Per rispondere a questa domanda è stata utilizzata una tabella con codifica cromatica, in cui i valori sono evidenziati secondo una regola di colore. L'utilizzo della highlight table è giustificato dall'approccio alla visualizzazione quantitativa proposto da Few (2009) nel libro Now You See It, che definisce questa tecnica come un'estensione delle tabelle tradizionali basata sull'impiego del colore per evidenziare pattern, differenze e valori anomali all'interno dei dati.

Dalla visualizzazione emerge subito che, per i dipartimenti di Medicina di Precisione e per l'Interdisciplinare di Medicina, la percentuale di soddisfazione risulta inferiore rispetto ai valori di riferimento nazionali (ITA) e del Sud Italia (SUD). Con una differenza minore, lo stesso comportamento si osserva anche per i dipartimenti di Chimica e Informatica.

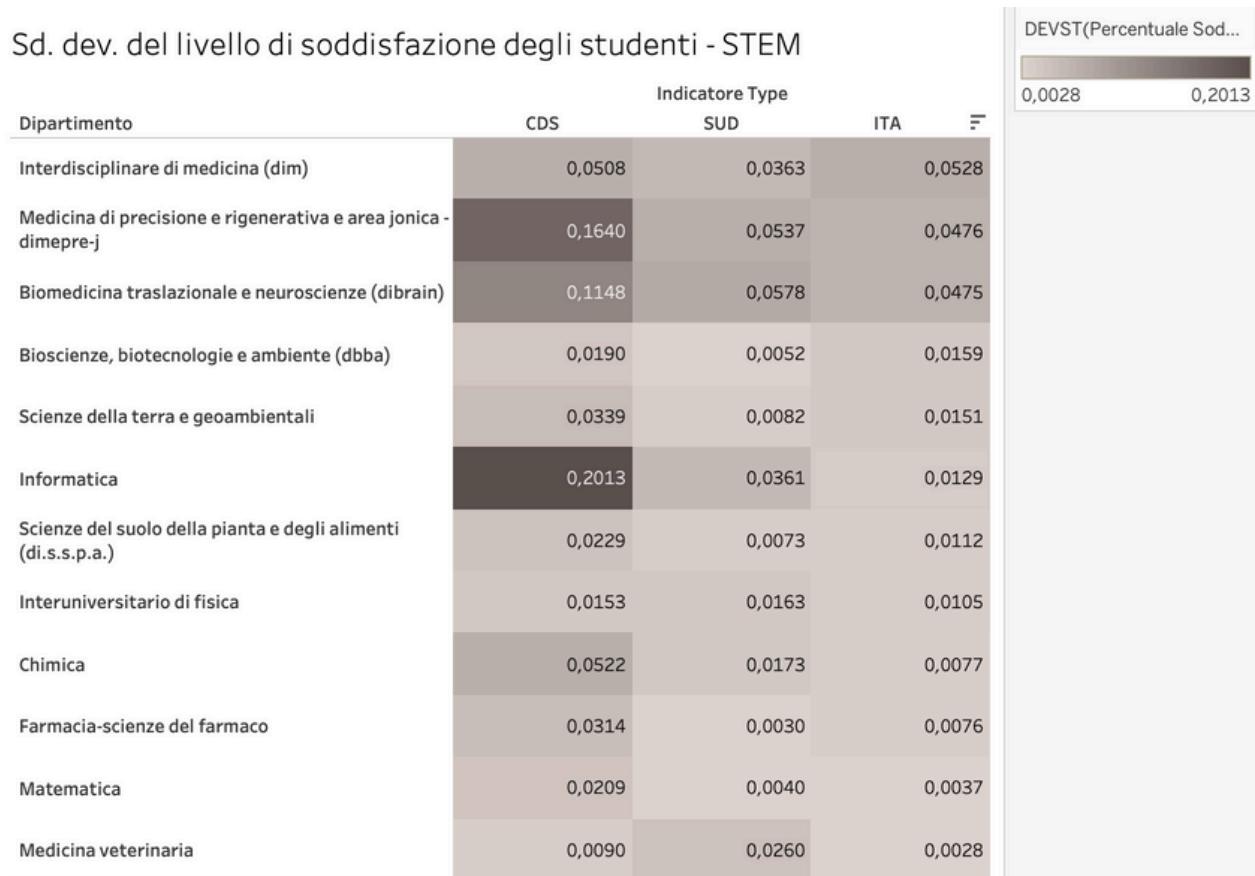
Per la maggior parte dei dipartimenti, le percentuali di soddisfazione risultano invece in linea con i valori di riferimento. Fanno eccezione il dipartimento di Fisica e quello di Scienze della Terra e Geoambientali, che mostrano una percentuale di soddisfazione chiaramente superiore sia alla media nazionale sia a quella del Sud Italia.

# Visualizzazioni Tableau

## Categoria - STEM

RQ5: Quanto omogenea è la soddisfazione tra i corsi in questi dipartimenti?

Sd. dev. del livello di soddisfazione degli studenti - STEM



La media rappresenta una buona pratica per valutare in termini generali, la qualità dell'offerta formativa; tuttavia, può mascherare differenze rilevanti tra i corsi appartenenti allo stesso dipartimento. Per questo motivo, è utile affiancare alla media una misura di dispersione, che evidenzi il grado di eterogeneità interna.

Dall'analisi emerge che il Dipartimento di Informatica presenta il valore più elevato di deviazione standard. In particolare, il corso di laurea magistrale in Data Science contribuisce in modo significativo a tale variabilità, registrando livelli di soddisfazione inferiori rispetto agli altri corsi dello stesso dipartimento. Sebbene l'eterogeneità e l'interdisciplinarietà del corso possano rendere più complessa la sua gestione, il fatto che la deviazione standard risulti significativamente più elevata rispetto ai valori di riferimento nazionali (ITA) e del Sud Italia (SUD) suggerisce che si potrebbero adottare buone pratiche già implementate in altri dipartimenti italiani che hanno affrontato problematiche simili.

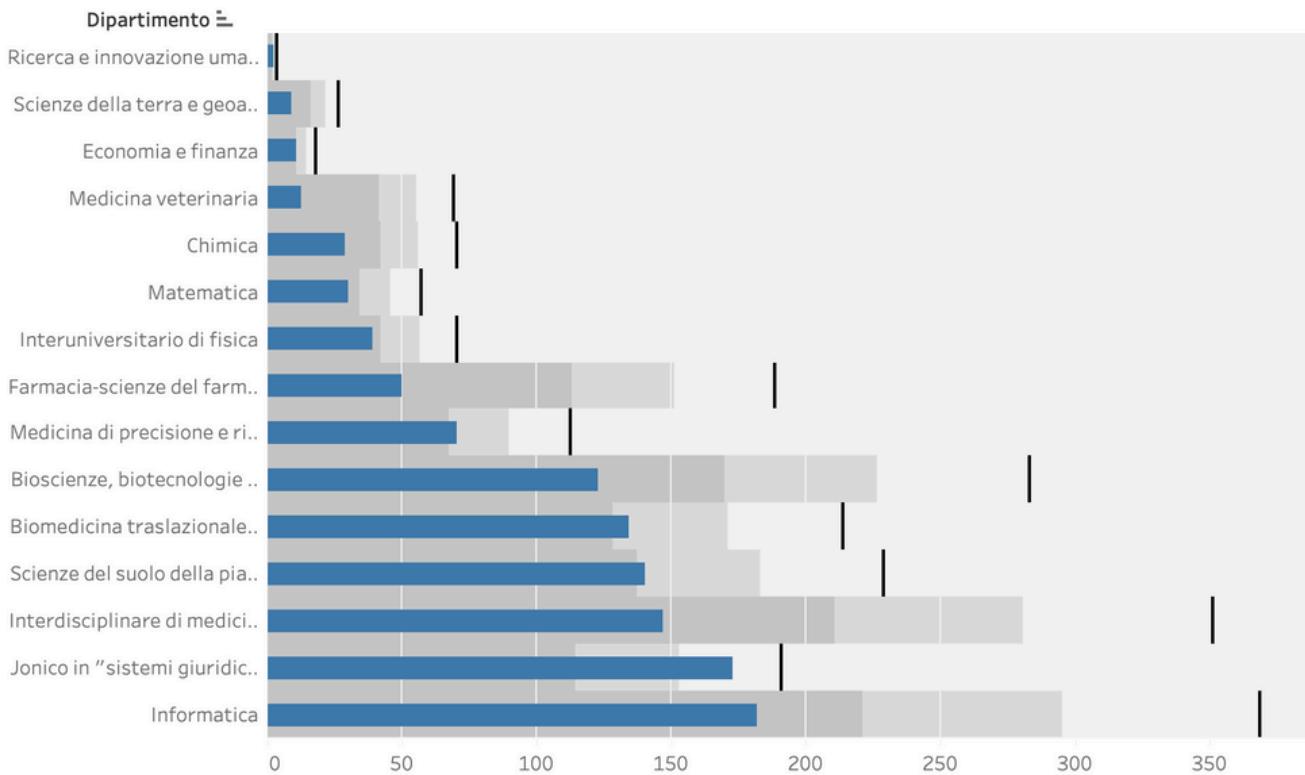
Valori di deviazione standard superiori ai valori di riferimento si osservano anche per i corsi afferenti al Dipartimento di Medicina di Precisione e Biomedicina.

# Visualizzazioni Tableau

## Categoria - STEM

RQ6: In quali dipartimenti i corsi presentano maggiori difficoltà nel conseguire la laurea nei tempi previsti, in rapporto al numero totale di laureati?

Confronto in media tra la numerosità di laureati totali e laureati in corso nelle discipline STEM (CdS)



Il bullet chart è la tecnica più adatta per rappresentare questa relazione, in quanto consente di visualizzare contemporaneamente diverse soglie di riferimento. Few (2006) propone il bullet chart come alternativa ai gauge tradizionali, evidenziando come questi ultimi privilegino l'aspetto decorativo a discapito dell'efficacia percettiva.

In questo grafico, la barra principale rappresenta il numero totale di laureati. Le barre in grigio chiaro e grigio scuro indicano rispettivamente l'80% e il 60% di tale quantità, mentre la barra colorata mostra il numero di laureati nei tempi previsti. Questo tipo di visualizzazione facilita il confronto tra i dipartimenti, evidenziando immediatamente quelli in cui la proporzione di laureati in tempo è più bassa.

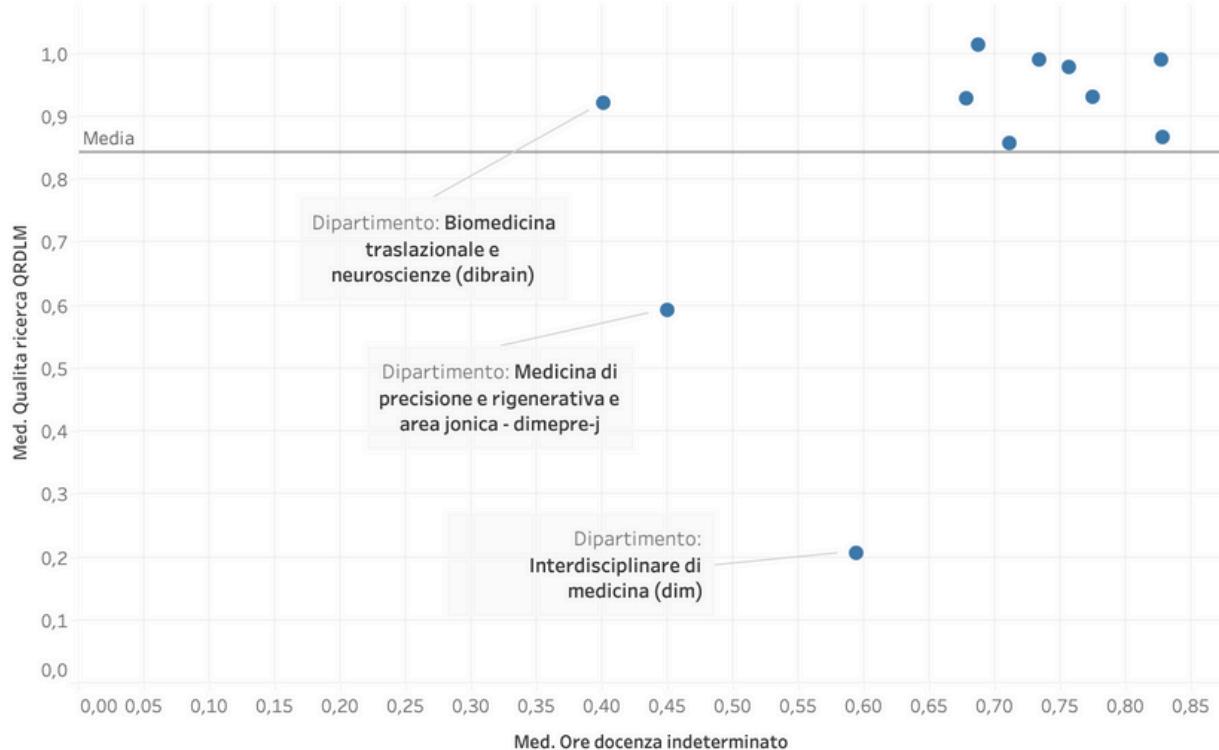
Ad esempio, si nota subito che per i dipartimenti di Chimica e Medicina Veterinaria, il numero di laureati nei tempi previsti è molto più basso rispetto al totale dei laureati. Senza questo riferimento, la situazione potrebbe sembrare simile a quella del dipartimento di Economia e Finanza, ma il valore effettivo è in realtà molto più vicino al numero totale di laureati. Preoccupante è anche la situazione dei dipartimenti di Informatica e Interdisciplinare di Medicina, che pur presentando un numero complessivo di laureati elevato, si collocano al di sotto del 60% di laureati nei tempi previsti.

# Visualizzazioni Tableau

## Categoria - STEM

RQ7: Nei dipartimenti con più ore erogate da docenti assunti a tempo indeterminato, come si presenta la qualità della ricerca?

Rapporto tra ore di docenza erogate da docenti assunti a tempo indeterminato e qualità della ricerca – STEM (CdS)



Dall'osservazione dei dati emergono subito due “gruppi”: gli outlier e il gruppo in alto a destra.

Il primo outlier è il Dipartimento di Medicina, che presenta un numero elevato di ore di docenza da parte di docenti a tempo indeterminato ma una bassa qualità della ricerca. L'altro outlier è il Dipartimento di Biomedicina, caratterizzato da una qualità della ricerca superiore alla media ma con un numero relativamente basso di ore di docenza da parte di tali docenti. Il Dipartimento di Medicina di precisione si colloca nel mezzo di questi punti.

Il gruppo in alto a destra, comprende invece tutti i dipartimenti con una qualità della ricerca maggiore della media e anche ore erogate da docenti a tempo indeterminato, ed è composto dai restanti dipartimenti, tra cui Matematica, Fisica, Chimica, Informatica, ecc.

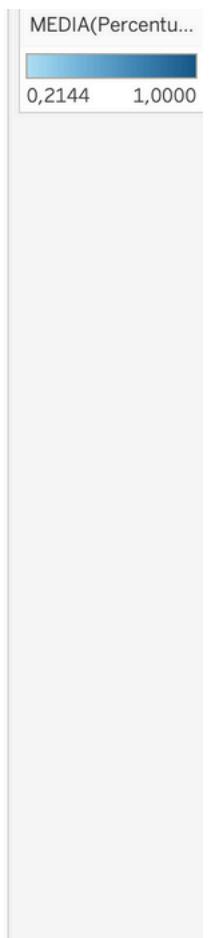
# Visualizzazioni Tableau

## Categoria - STEM

RQ8: Quali classi di laurea mostrano difficoltà nel passaggio al secondo anno?

Percentuale di studenti che proseguono il II anno - STEM (confronto completo)

Classe Di Laurea	Indicatore Type			
	CDS	SUD	ITA	EU
L-29 - scienze e tecnologie farmaceutiche	0,4029	0,4430	0,4780	
L-13 - scienze biologiche	0,5319	0,4264	0,4856	
L-2 - biotecnologie	0,2843	0,4061	0,4920	
L-32 - scienze e tecnologie per l'ambiente e la natura	0,4099	0,4066	0,5056	
L-27 - scienze e tecnologie chimiche	0,5809	0,4846	0,5549	
Lm-13, - farmacia e farmacia industriale	0,5266	0,4758	0,5693	
L-38 - scienze zootecniche e tecnologie delle produzioni animali	0,2144	0,5257	0,5796	
L-28 - scienze e tecnologie della navigazione	0,6307	0,5782	0,5855	
L-34 - scienze geologiche	0,5679	0,5635	0,5901	
L-35 - scienze matematiche	0,6437	0,6003	0,6236	
L-31 - scienze e tecnologie informatiche	0,5813	0,6062	0,6388	
L-25 - scienze e tecnologie agrarie e forestali	0,6280	0,6004	0,6415	
L-26 - scienze e tecnologie alimentari	0,6297	0,6100	0,6423	
L-30 - scienze e tecnologie fisiche	0,7040	0,6632	0,6920	
L/snt4 - professioni sanitarie della prevenzione	0,7209	0,6737	0,6998	
L-41 - statistica	0,6983	0,6540	0,7086	
L-sc,mat, - scienze dei materiali	0,3750	0,3750	0,7151	
L/snt3 - professioni sanitarie tecniche	0,6646	0,7933	0,7765	
Lm-46 - odontoiatria e protesi dentaria	0,6580	0,7705	0,8282	
L/snt2 - professioni sanitarie della riabilitazione	0,7893	0,8299	0,8293	
Lm-sc,mat, - scienze dei materiali	1,0000	1,0000	0,8606	
Lm-42 - medicina veterinaria	0,7087	0,7752	0,8702	
Lm-data - data science	0,8706	0,8510	0,8949	
Lm-10 - informatica	0,8901	0,8837	0,8902	



Per quanto riguarda le prime due lauree, L-29 e L-13, la situazione nella nostra università non presenta scostamenti significativi rispetto ai valori di riferimento nazionali (ITA) e del Sud Italia (SUD). Preoccupante risulta invece la situazione per la laurea in Biotecnologie, dove solo il 28% degli studenti prosegue nel secondo anno. Una situazione analoga si osserva per Scienze e Tecnologie delle Produzioni Animali, e Scienze dei materiali (Laurea triennale).

# Visualizzazioni Tableau

## Categoria - STEM

RQ9: Quali lauree magistrali hanno offerto maggiori opportunità di lavoro?

Percentuale di occupazione per le lauree magistrali - STEM  
(confronto completo)

Classe Di Laurea	Indicatore Type		
	CDS	SUD	ITA
Lm-66 - sicurezza informatica	1,0000	0,9129	0,9579
Lm-18 - informatica	0,9778	0,9563	0,9468
Lm-41 - medicina e chirurgia	0,9513	0,9441	0,9396
Lm/snt4 - scienze delle professioni sanitarie della prevenzione	0,9688	0,9231	0,9381
Lm/snt1 - scienze infermieristiche e ostetriche	0,8881	0,9396	0,9326
Lm-40 - matematica	0,8782	0,9115	0,9279
Lm-17 - fisica	0,9496	0,9185	0,9235
Lm-46 - odontoiatria e protesi dentaria	0,9114	0,8907	0,9158
Lm-54 - scienze chimiche	0,8439	0,8511	0,8963
Lm-13 - farmacia e farmacia industriale	0,8983	0,8684	0,8908
Lm-69 - scienze e tecnologie agrarie	0,9765	0,8379	0,8880
Lm-42 - medicina veterinaria	0,9058	0,8715	0,8808
Lm-69 - scienze e tecnologie agrarie; Lm-73 - scienze e tecnologie forestali ed ambie..	0,8921	0,8263	0,8764
Lm-70 - scienze e tecnologie alimentari	0,7545	0,7622	0,8671
Lm-74 - scienze e tecnologie geologiche; Lm-79 - scienze geofisiche	0,7305	0,7994	0,8670
Lm-7 - biotecnologie agrarie	0,8542	0,7866	0,8463
Lm-9 - biotecnologie mediche, veterinarie e farmaceutiche	0,8363	0,7930	0,8421
Lm-60 - scienze della natura; Lm-75 - scienze e tecnologie per l'ambiente e il territor..	0,8571	0,7745	0,8263
Lm-6 - biologia	0,8206	0,7502	0,8105

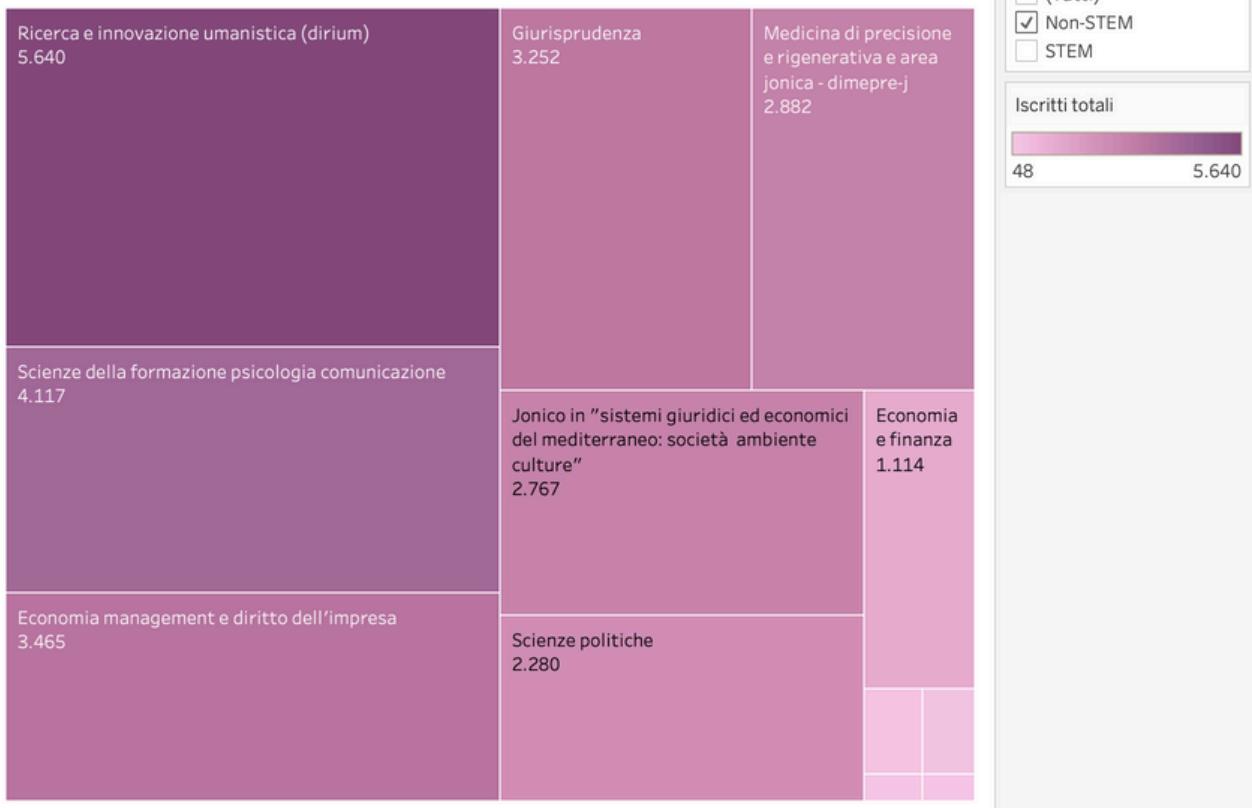
Da questa visualizzazione emerge chiaramente che gli studenti laureati in uno dei corsi del Dipartimento di Informatica hanno trovato lavoro quasi al 100% a tre anni dal conseguimento del titolo. Seguono, con percentuali leggermente inferiori, i laureati dei dipartimenti di Medicina e Chirurgia e Scienze Sanitarie. Si osserva inoltre che i laureati in Fisica e in Scienze e Tecnologie Agrarie riportano valori superiori alla media nazionale (ITA), mentre per Scienze e Tecnologie alimentari, Scienze geofisiche, la percentuale a Bari è circa il 11% inferiore rispetto alla media nazionale.

# Visualizzazioni Tableau

## Categoria - Non STEM

RQ1: Come si presenta la situazione in Italia per quanto riguarda il numero totale di iscritti nei dipartimenti che offrono corsi Non STEM?

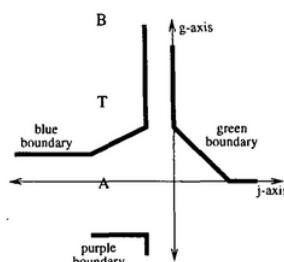
Numero di iscritti totali per dipartimento - Non Stem (ita)



Per quanto riguarda i corsi non-STEM, è stato scelto un colore differente per distinguerli visivamente dai corsi STEM. Per i corsi non-STEM è stato scelto il viola, un colore sufficientemente distinto dal blu nello spettro percettivo secondo i principi di selezione cromatica basata sulla percezione descritti da Healey (1996).

Dalla visualizzazione emerge che, a livello nazionale (ITA), i tre dipartimenti con il maggior numero di iscritti sono Ricerca e Innovazione Dirium, Scienze della Formazione ed Economia e Management.

[https://cs.brown.edu/research/vis/docs/pdf/Healey-1996-CEC.pdf?](https://cs.brown.edu/research/vis/docs/pdf/Healey-1996-CEC.pdf)

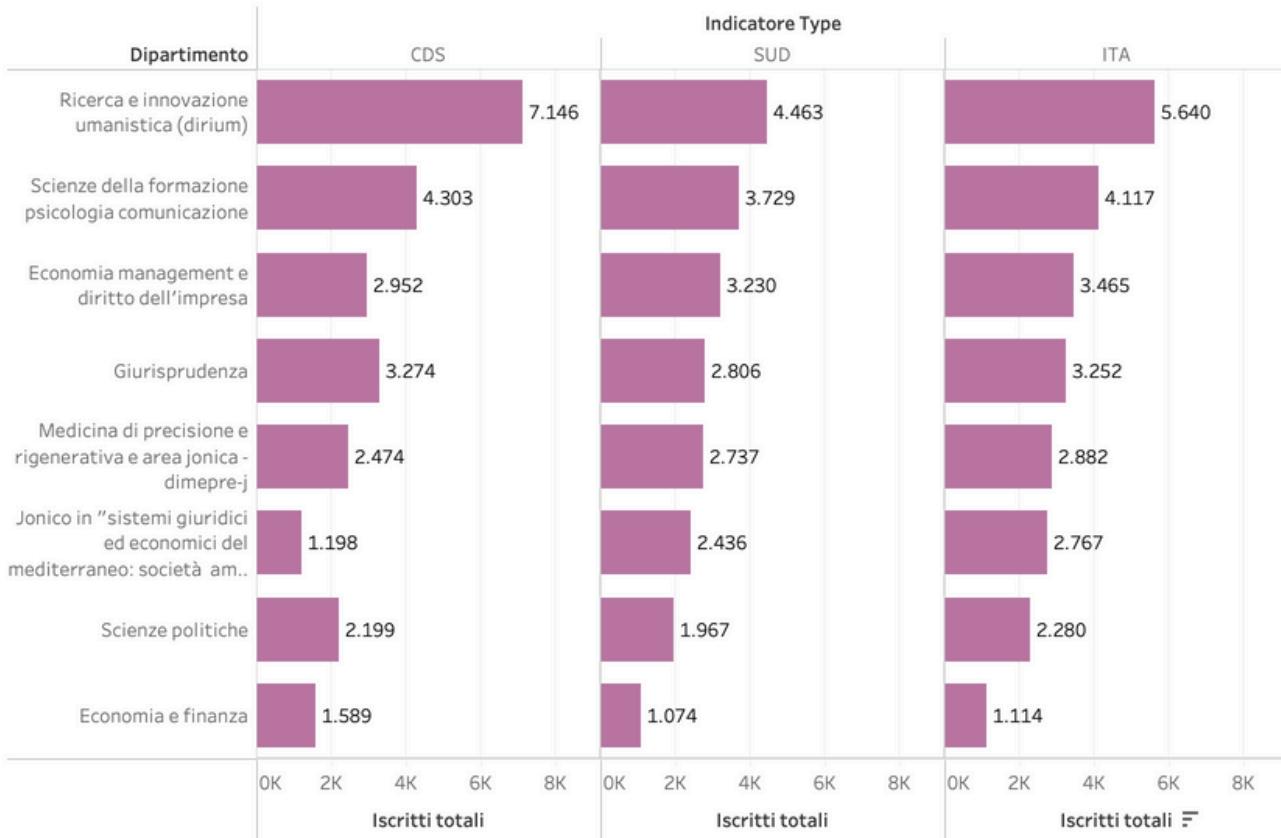


# Visualizzazioni Tableau

## Categoria - Non STEM

RQ2: Come cambia questa distribuzione nel Sud Italia e nel contesto di riferimento analizzato?

Confronto degli iscritti totali in base al tipo dell'Indicatore - Non STEM



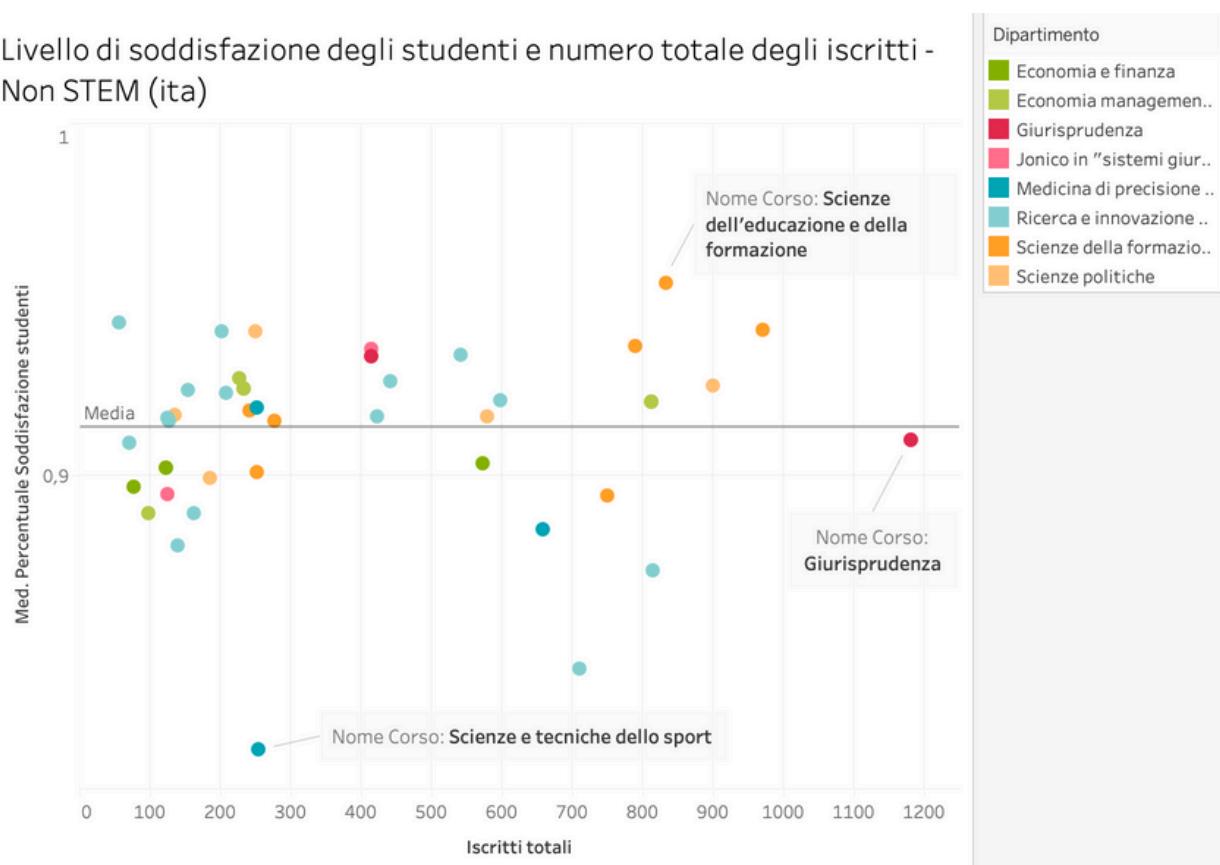
Si nota subito che il Dipartimento Dirium si colloca al primo posto per tutti e tre gli indicatori, ma all'UNIBA si registra un valore superiore di circa 2.000 studenti. Valori elevati si riscontrano anche per il Dipartimento di Scienze della Formazione. Il trend non viene invece seguito dal Dipartimento Jonico in Sistemi Economici e Giuridici, che presenta una differenza di oltre 1000 studenti in meno.

# Visualizzazioni Tableau

## Categoria - Non STEM

RQ3: Esiste una relazione tra il numero di iscritti e la percentuale di soddisfazione degli studenti?

Livello di soddisfazione degli studenti e numero totale degli iscritti - Non STEM (ita)



Dallo scatterplot emergono immediatamente due outlier: il corso di Scienze e tecniche dello sport, con pochi iscritti e una bassa percentuale di soddisfazione, e il corso di Giurisprudenza, con un numero elevato di iscritti e una percentuale di soddisfazione nella media.

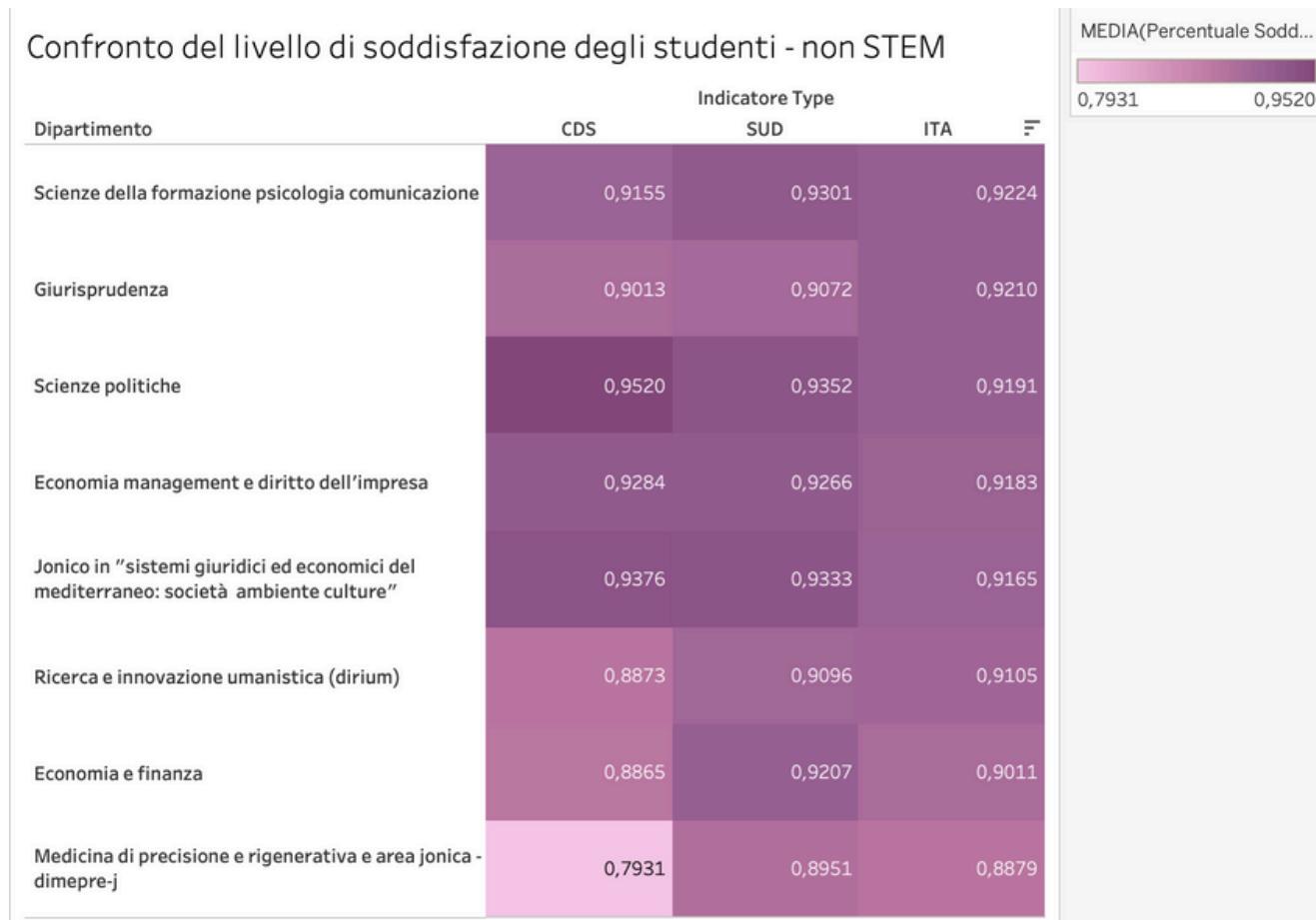
Si osserva inoltre che il corso di Scienze della Formazione presenta la soddisfazione più alta. È interessante notare come il Dipartimento Dirium abbia punti distribuiti in modo sparso su tutto il grafico.

Il Dipartimento di Medicina di Precisione mostra una percentuale di soddisfazione generalmente al di sotto della media, con un unico corso che supera la linea della media.

# Visualizzazioni Tableau

## Categoria - Non STEM

RQ4: È possibile scendere maggiormente nel dettaglio e individuare quali dipartimenti presentano livelli di soddisfazione significativamente distanti dalla media nazionale (ITA) e da quella del Sud Italia (SUD)?

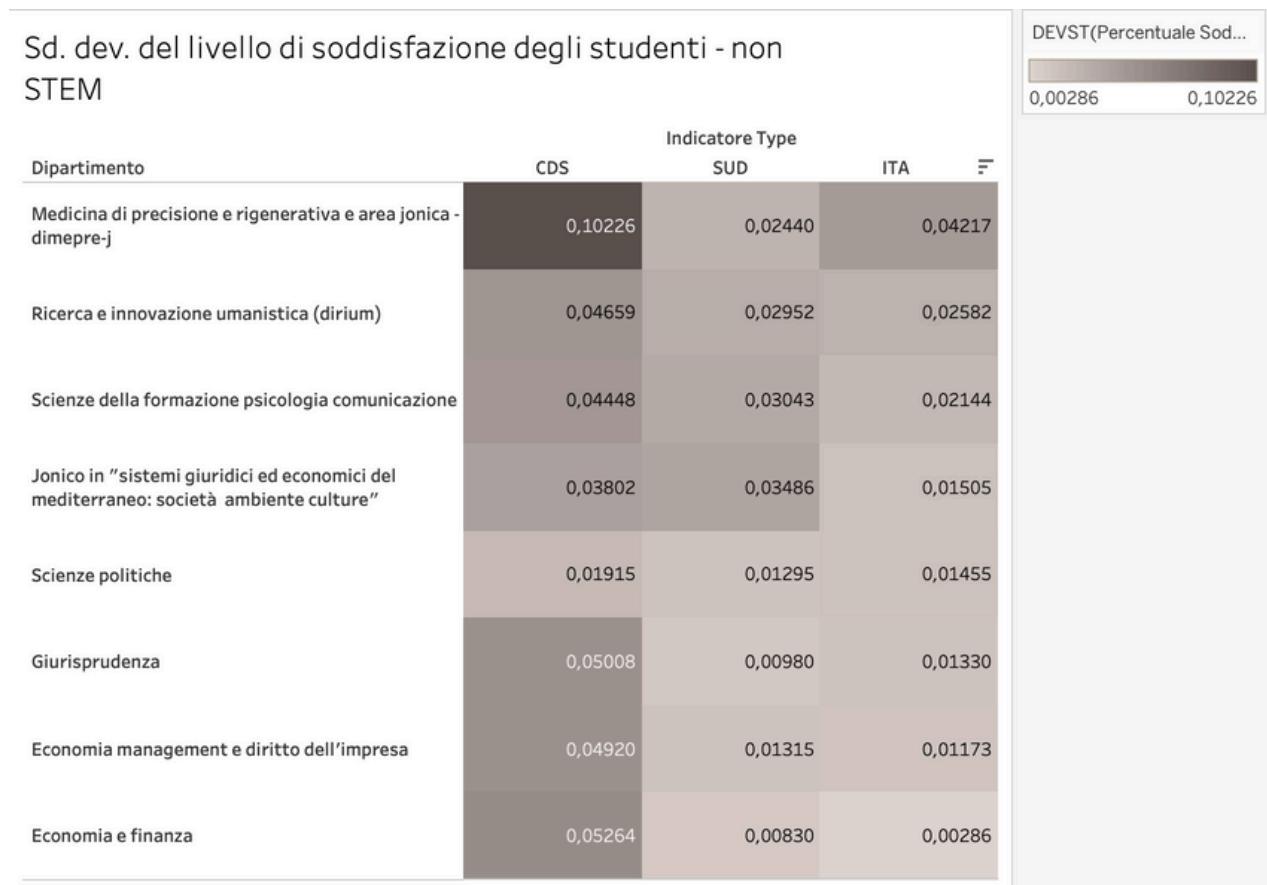


In relazione a questa domanda, si osserva che il Dipartimento di Scienze Politiche e il Dipartimento Jonico presentano valori superiori rispetto alla media nazionale (ITA). Al contrario, il Dipartimento di Economia e Finanza mostra valori inferiori sia alla media nazionale sia a quella del Sud Italia (SUD). Anche il Dipartimento di Medicina di Precisione riporta valori di soddisfazione inferiori rispetto alla media nazionale.

# Visualizzazioni Tableau

## Categoria - Non STEM

RQ5: Quanto omogenea è la soddisfazione tra i corsi in questi dipartimenti?



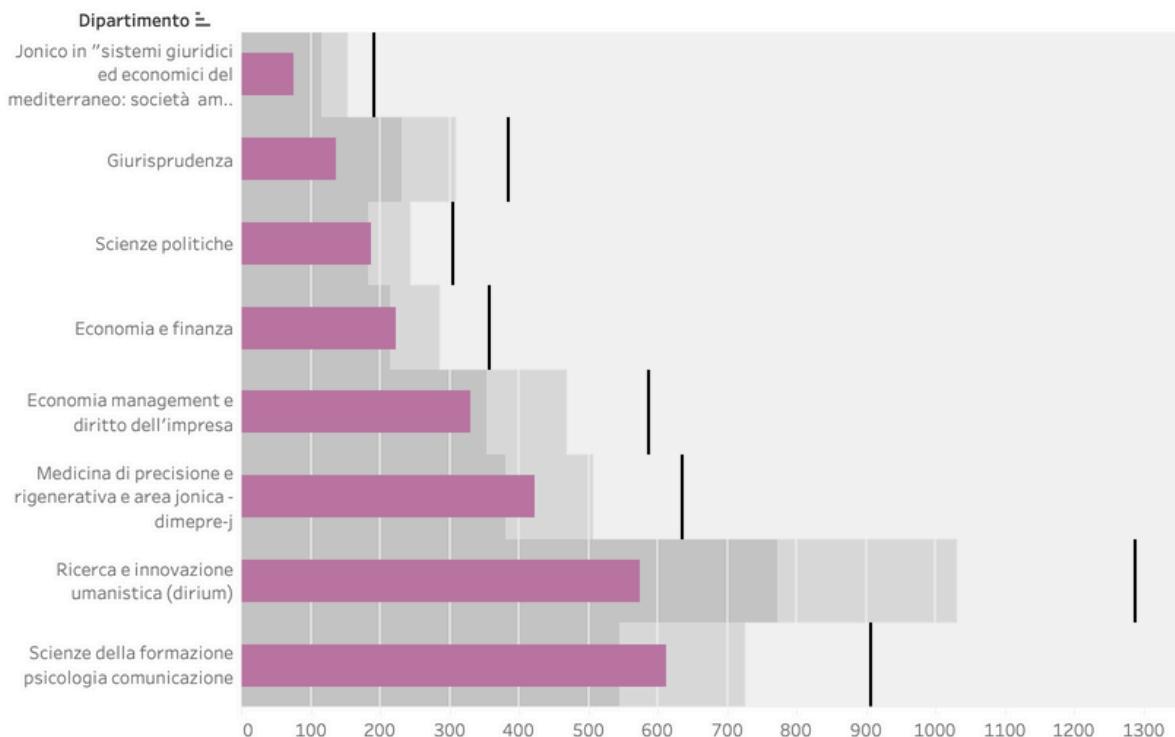
Il primo dipartimento che si distingue per lo scostamento nei livelli di soddisfazione è il Dipartimento di Medicina di Precisione, seguito da Giurisprudenza e Economia e Management, che presentano valori inferiori ma comunque circa 5 volte superiori rispetto alla media nazionale (ITA).

# Visualizzazioni Tableau

## Categoria - Non STEM

RQ6: In quali dipartimenti i corsi presentano maggiori difficoltà nel conseguire la laurea nei tempi previsti, in rapporto al numero totale di laureati?

Confronto tra laureati totali e laureati in corso nelle discipline - non-STEM (CdS)



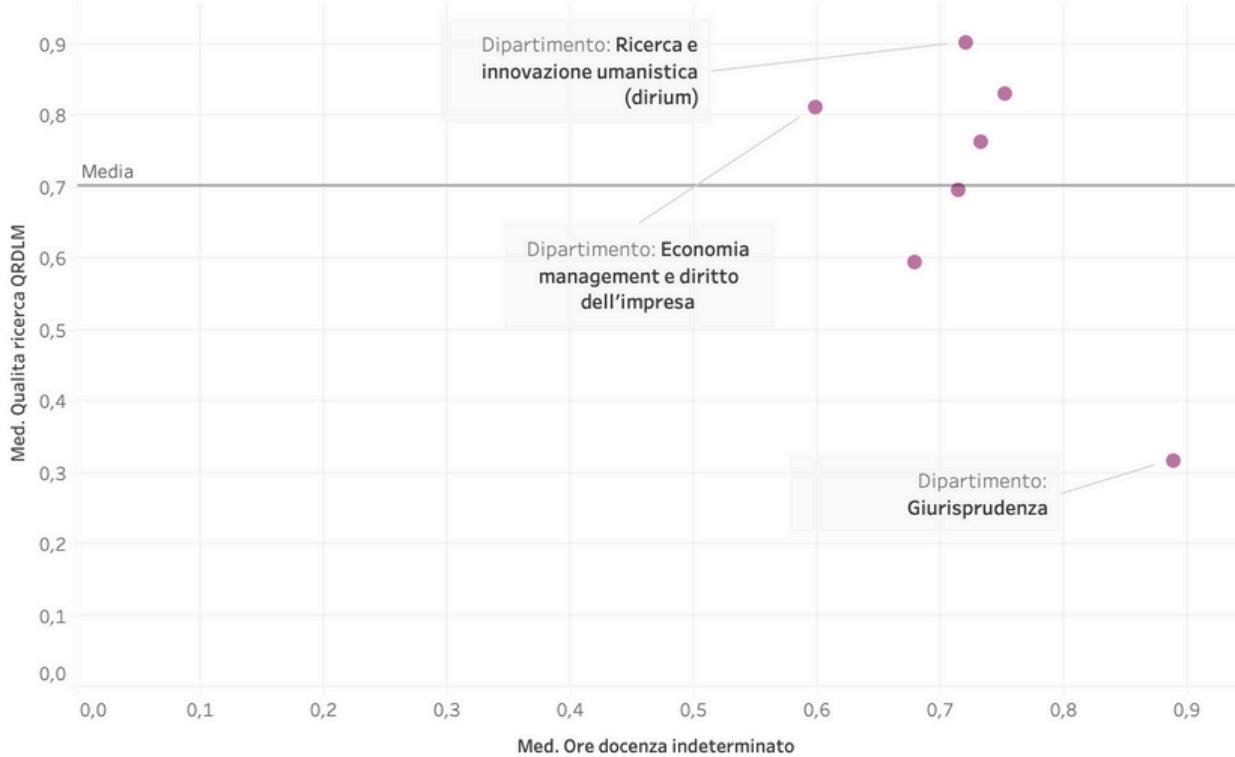
È interessante notare come il Dipartimento di Scienze della Formazione presenti quasi lo stesso numero di laureati in tempo rispetto al Dipartimento Dirium, sebbene per quest'ultimo tale numero rappresenti meno del 60% del totale dei laureati. Percentuali inferiori al 60% si osservano anche nei dipartimenti Jonico, Giurisprudenza, e seppur in misura minore, in Economia e Management e Diritto dell'Impresa.

# Visualizzazioni Tableau

## Categoria - Non STEM

RQ7: Nei dipartimenti con più ore erogate da docenti assunti a tempo indeterminato, come si presenta la qualità della ricerca?

Rapporto tra ore di docenza erogate da docenti assunti a tempo indeterminato e qualità della ricerca – Non STEM (CdS)



Il primo outlier di questo scatterplot, caratterizzato da un elevato numero di ore erogate da docenti a tempo indeterminato e da una bassa qualità della ricerca, è il Dipartimento di Giurisprudenza.

Il dipartimento con la qualità della ricerca maggiore è invece il Dipartimento Dirium. Tuttavia, la distribuzione dei dati in questo scatterplot è simile a quella dei dipartimenti STEM, nel senso che si osservano outlier e un gruppo in alto a destra; però in questo caso, il gruppo mostra una leggera correlazione positiva tra le ore di docenza e la qualità della ricerca.

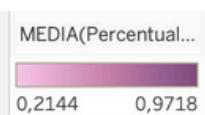
# Visualizzazioni Tableau

## Categoria - Non STEM

RQ8: Quali classi di laurea mostrano difficoltà nel passaggio al secondo anno?

Percentuale di studenti che proseguono il II anno - Non STEM (confronto completo)

Classe Di Laurea	Indicatore Type			
	CDS	SUD	ITA	ES
L-39 - servizio sociale; I-40 - sociologia	0,5953	0,6009	0,6151	
L-16 - scienze dell'amministrazione e dell'organizzazione	0,5071	0,5809	0,6528	
L-1 - beni culturali	0,5122	0,6306	0,6564	
L-5 - filosofia	0,6680	0,6387	0,6679	
L-42 - storia	0,6134	0,6449	0,6701	
L-11 - lingue e culture moderne	0,6730	0,6836	0,6750	
L-14 - scienze dei servizi giuridici	0,5032	0,6834	0,6859	
L-3 - discipline delle arti figurative, della musica, dello spettacolo e della moda	0,6873	0,6498	0,6961	
L-33 - scienze economiche	0,5765	0,6266	0,6967	
Lmg/01 - magistrali in giurisprudenza	0,6154	0,6325	0,7001	
L-10 - lettere	0,7593	0,7011	0,7024	
L-15 - scienze del turismo	0,8077	0,6365	0,7081	
L-19 - scienze dell'educazione e della formazione	0,8481	0,6866	0,7117	
L-22 - scienze delle attività motorie e sportive	0,7157	0,6873	0,7293	
L-36 - scienze politiche e delle relazioni internazionali	0,5586	0,6418	0,7335	
L-18 - scienze dell'economia e della gestione aziendale	0,5441	0,6544	0,7357	
L-12 - mediazione linguistica	0,6717	0,6921	0,7445	
L/snt1 - professioni sanitarie, infermieristiche e professione sanitaria ostetrica	0,8100	0,8303	0,7463	
L-p02 - professioni tecniche agrarie, alimentari e forestali	0,6121	0,6145	0,7507	
L-20 - scienze della comunicazione	0,6825	0,6795	0,7518	
L/gastr - scienze, culture e politiche della gastronomia	0,5455	0,6076	0,7531	
Lm/ds - scienze della difesa e della sicurezza	0,5286	0,7304	0,8217	
L-24 - scienze e tecniche psicologiche	0,8987	0,8274	0,8541	
Lm/03 - conoscenze e sostanza dei beni culturali	0,7076	0,7426	0,8578	



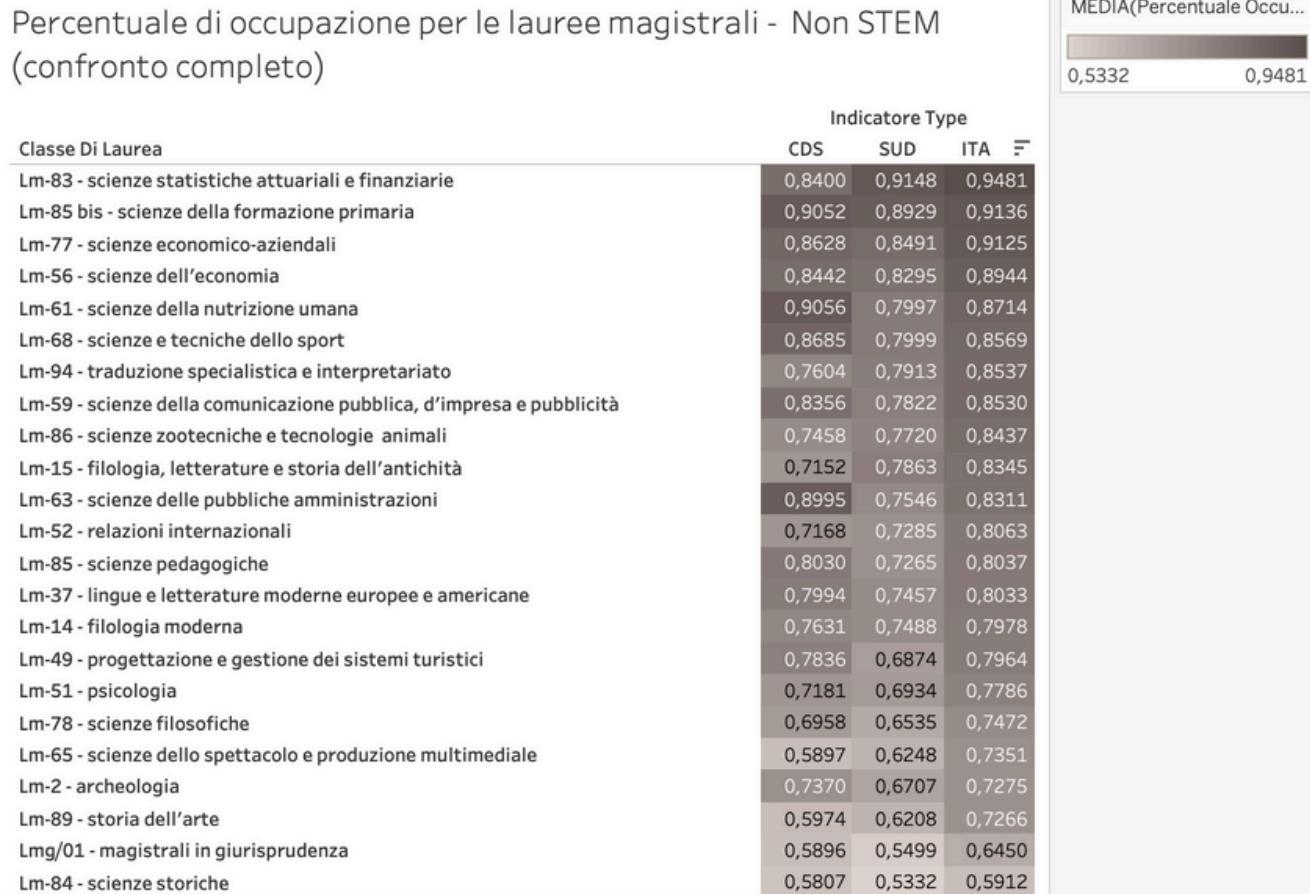
Da questa visualizzazione emerge che non si registrano situazioni drastiche, come percentuali di studenti che proseguono il secondo anno intorno al 20% nel caso STEM. Alcune classi di laurea riportano valori superiori alla media nazionale, ad esempio Scienze del Turismo, Scienze della Formazione e Professioni Sanitarie. Valori inferiori rispetto alla media si osservano invece per le classi di laurea in Scienze dell'Amministrazione e dell'Organizzazione, Beni Culturali, Scienze dei Servizi Giuridici e Scienze Economiche, Mediazione linguistica, Scienze della comunicazione, Scienze della difesa.

# Visualizzazioni Tableau

## Categoria - Non STEM

RQ9: Quali lauree magistrali hanno offerto maggiori opportunità di lavoro?

Percentuale di occupazione per le lauree magistrali - Non STEM  
(confronto completo)



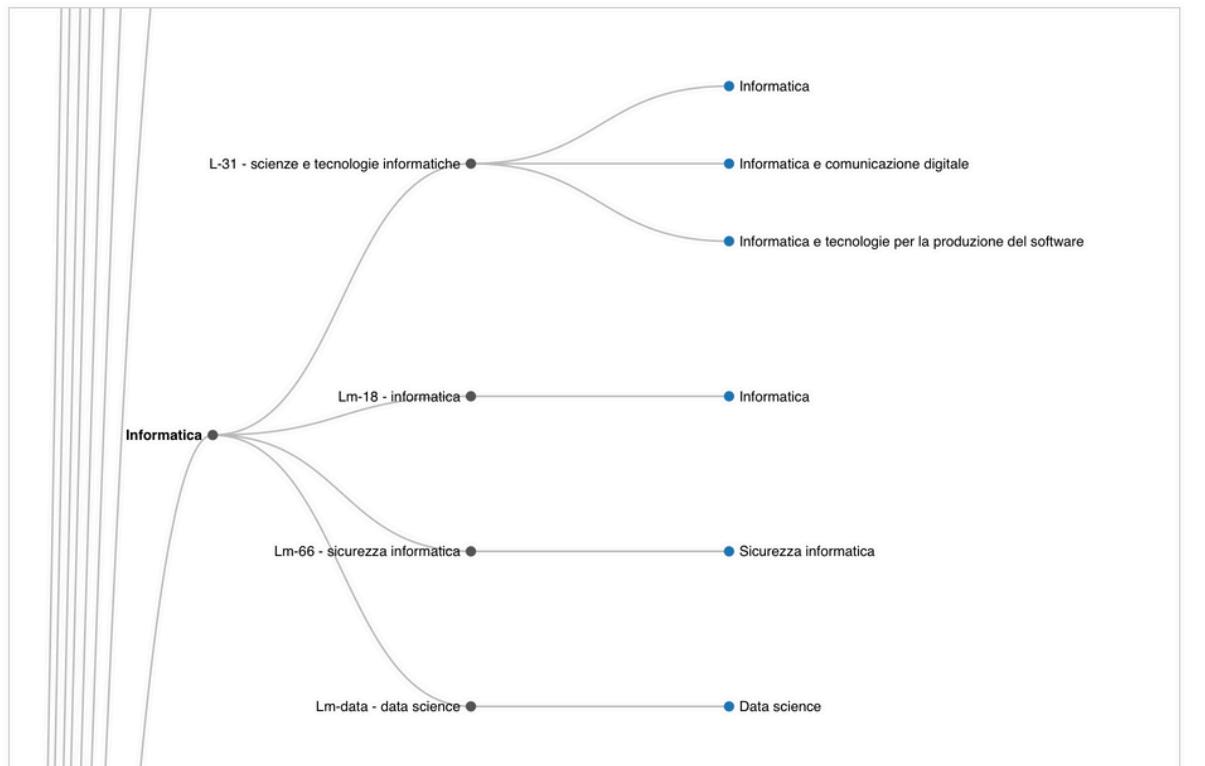
In Italia, le Scienze Statistiche, Attuariali e Finanziarie occupano il primo posto per percentuale di occupazione; nella realtà di Bari, invece, questa percentuale si discosta di circa il 10%. Valori significativamente inferiori si riscontrano anche per Scienze dell'Economia, Scienze dello Spettacolo, Storia dell'Arte e i corsi magistrali in Giurisprudenza.

Al contrario, valori più elevati si osservano per Scienze della Nutrizione, Scienze delle Pubbliche Amministrazioni e in misura minore per Archeologia.

# Visualizzazioni D3.js - Appendice

## Confronto Tidytree - Spacetree

Tidytree



Tidytree serve per disegnare diagrammi gerarchici nodo-collegamento. In questo caso è stato utilizzato per mostrare la struttura gerarchica dei dipartimenti e dei corsi, offrendo all'utente la possibilità di esplorare l'offerta formativa di ciascun dipartimento.

Tuttavia, questa rappresentazione non è ottimale per fornire informazioni immediate all'utente, poiché richiede di scorrere l'albero e numerosi dati che potrebbero non essere rilevanti.

Pertanto, la visualizzazione può essere migliorata utilizzando invece il SpaceTree. SpaceTree introduce il ridimensionamento dinamico dei rami dell'albero, adattandoli allo spazio disponibile sullo schermo, e include funzioni integrate di ricerca e filtro.

# Visualizzazioni D3.js - Appendice

## Confronto Tidytree - Spacetree

SpaceTree, anche questa tecnica sviluppata dal Human-Computer Interaction Lab dell'Università del Maryland, rappresenta le gerarchie come diagrammi node-link dinamici e interattivi. Quindi, la sua principale caratteristica è la possibilità di esplorare l'albero in maniera interattiva perché i rami possono essere espansi o compressi.

Questo rende SpaceTree più adatto a strutture complesse e profonde.

Tuttavia, questa interattività dipende dal coinvolgimento dell'utente, e la comprensione dipende dalla sua familiarità con i meccanismi di navigazione.

Il Tidy Tree layout, si concentra principalmente sull'ottimizzazione dello spazio e sull'estetica del layout. Questo approccio è efficace quando la gerarchia deve essere rappresentata in forma sintetica, senza la necessità di interazione dinamica.

Il principale svantaggio è che per alberi molto grandi (come in questo caso) possono risultare compressi e difficili da navigare.

Pertanto, SpaceTree è ideale quando l'utente deve navigare attivamente in gerarchie estese e complesse, mentre il Tidy Tree è preferibile per visualizzazioni statiche.

Entrambe le tecniche offrono vantaggi distinti, ma presentano anche limitazioni quindi la scelta dipende dal tipo di analisi e dal contesto.

The screenshot shows the HCIL Archive website with the title "HCIL Archive" and the subtitle "SpaceTree: a novel node-link tree browser". On the left, there is a navigation bar with links to "Participants", "Publications", and "Sponsors". Below this is a "Jump to:" dropdown menu. In the center, there is a large, complex tree diagram representing a dataset. The tree is highly branched and uses a dynamic layout where branches are compressed or expanded to fit the available space. To the right of the tree, there is a smaller, more compact version of the same tree diagram, likely representing a static tidytree layout. At the bottom, there is a "Project description:" section with a detailed explanation of the SpaceTree browser's features and principles.

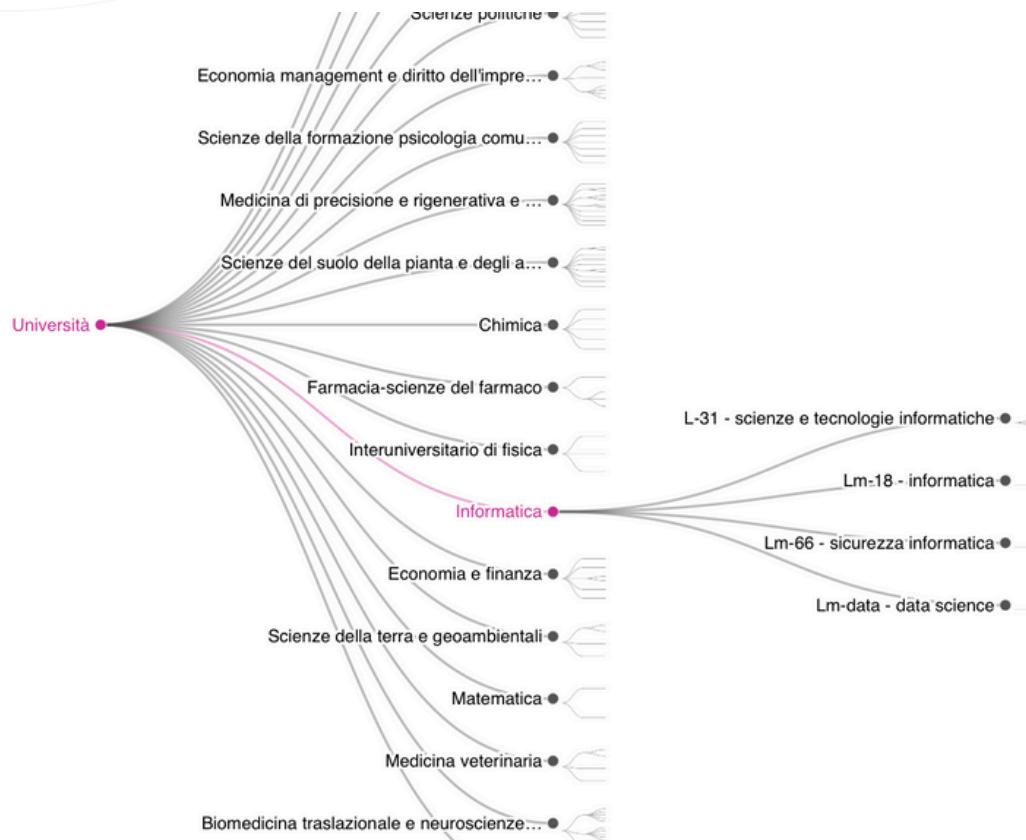
### Project description:

SpaceTree is a novel tree browser that builds on the conventional layout node link diagrams along a single preferred direction. It adds dynamic rescaling of branches of the tree to best fit the available screen space, optimized camera movement, and the use of preview icons summarizing the topology of the branches that cannot be expanded. In addition, it includes integrated search and filter functions. This paper reflects on the evolution of the design and highlights the principles that emerged from it. A controlled experiment showed benefits for navigation tasks to already previously visited nodes and estimation of overall tree topology.

[Participante:](#)

# Visualizzazioni D3.js - Appendice

## Confronto Tidytree - Spacetree



SpaceTree implementa una struttura dinamica e gerarchica, in cui il layout viene calcolato ricorsivamente e animato tramite transizioni SVG. Questo approccio consente di esplorare grandi alberi attraverso l'espansione e la contrazione dei rami, ma introduce complessità nella gestione degli eventi DOM.

In particolare, il package `@john-guerra/spacetree` incapsula il rendering all'interno di un wrapper JavaScript che gestisce internamente gli stati dei nodi e le transizioni. In Observable, questo wrapper non espone direttamente gli elementi SVG sottostanti. Da un punto di vista teorico, ciò implica che non è possibile agganciare meccanismi standard di hover o di drag/panning.

<https://www.cs.umd.edu/projects/hcil/spacetree/>

# Conclusioni

In questa sezione cercherò di riassumere le principali problematiche che ciascun dipartimento deve esplorare, al fine di migliorare il proprio posizionamento rispetto ai valori di riferimento.

Dalle analisi emerge che il **Dipartimento Interdisciplinare di Medicina** presenta un livello di soddisfazione inferiore di circa l'11% rispetto alla media nazionale, un numero di laureati totali relativamente basso rispetto a quelli in tempo e una qualità della ricerca bassa.

Nel caso del **Dipartimento di Bioscienze e Biotecnologie**, si osserva un numero di iscritti inferiore al dato nazionale, difficoltà per gli studenti nel passaggio al secondo anno e una discrepanza tra laureati totali e laureati in tempo.

Il **Dipartimento di Informatica** registra un livello di soddisfazione inferiore del 9% rispetto alla media nazionale e mostra un'alta eterogeneità dei livelli di soddisfazione tra i corsi, oltre a maggiori difficoltà per gli studenti nel laurearsi nei tempi previsti.

Situazioni analoghe si riscontrano nei dipartimenti di **Farmacia e Scienze del Farmaco, Medicina Veterinaria e Chimica**, caratterizzati da un numero di laureati totali molto inferiore rispetto a quelli in tempo.

Per il **Dipartimento di Scienze del Suolo** si evidenzia una percentuale di occupazione inferiore alla media nazionale.

Il **Dipartimento di Biomedicina** e quello di **Medicina di Precisione** presentano entrambe una forte eterogeneità dei livelli di soddisfazione tra i corsi e necessitano di interventi per uniformarla. In aggiunta, entrambi mostrano una qualità della ricerca relativamente bassa. Nel caso specifico di Medicina di Precisione, il livello di soddisfazione complessivo risulta inferiore di circa il 10% rispetto alla media nazionale.

Dall'analisi emerge anche che il **Dipartimento Interuniversitario di Fisica** e il **dipartimento di Matematica** presentano un numero di laureati totali basso rispetto ai laureati in tempo, con il dipartimento di Matematica che mostra inoltre una percentuale di occupazione più bassa rispetto alla media nazionale.

Il **Dipartimento di Scienze della Terra e Geoambientali** registra un numero di iscritti totali inferiore e una percentuale di occupazione anch'essa sotto la media nazionale. Analogamente, il corso di **Scienze Statistiche al dipartimento di Economia e Finanza** evidenzia un numero di iscritti inferiori ai valori nazionali.

Il dipartimento **Ricerca e Innovazione Umanistica Dirium** mostra difficoltà legate al completamento del percorso nei tempi previsti e al passaggio al secondo anno, e i laureati hanno poche possibilità di lavoro.

# Conclusioni

Nei corsi di **Scienze della Formazione, Psicologia e Comunicazione**, così come nei **dipartimenti di Economia e Management, Diritto dell'Impresa e Giurisprudenza**, emerge l'esigenza di rendere più omogenei i livelli di soddisfazione tra i corsi; per **Giurisprudenza** si aggiungono anche pochi laureati nei tempi previsti, bassa qualità della ricerca e necessità di maggiori opportunità occupazionali.

Il **Dipartimento Jonico in Sistemi Giuridici ed Economici** non segue il trend nazionale per gli iscritti totali nelle discipline non-STEM e presenta difficoltà nel passaggio al secondo anno. Situazioni analoghe si riscontrano nei dipartimenti **di Scienze Politiche ed Economia e Finanza**, dove gli studenti mostrano difficoltà nel proseguire gli studi e si rendono necessarie ulteriori possibilità di inserimento lavorativo.