

Progetto Finale Data Analysis: Education

Emanuele Terragni

In questo progetto finale verrà affrontata l'analisi completa di un dataset relativo ad una piattaforma online progettata a scopo educativo.

All'interno di essa gli studenti possono affrontare dei corsi formativi di differenti aree e livelli, sono presenti anche i dati generali degli studenti (come nazionalità e device tramite il quale han seguito il corso). è presente anche un sistema di rating (da 1 a 5) per ogni classe.

Gli obiettivi dell'analisi sono quelli di osservare la distribuzione degli studenti e capire se si possono individuare dei modelli di "studente tipo", cercare eventuali trend temporali e proporre un'eventuale strategia per il futuro della piattaforma.

A livello tecnico è stata svolta un'analisi dei dati tramite Python e delle Dashboard utilizzando Tableau

Dataset:

https://drive.google.com/file/d/1H5BfxPE0e8V8pXHqcplbuQ0eY9_WLQ22/view?usp=drive_link

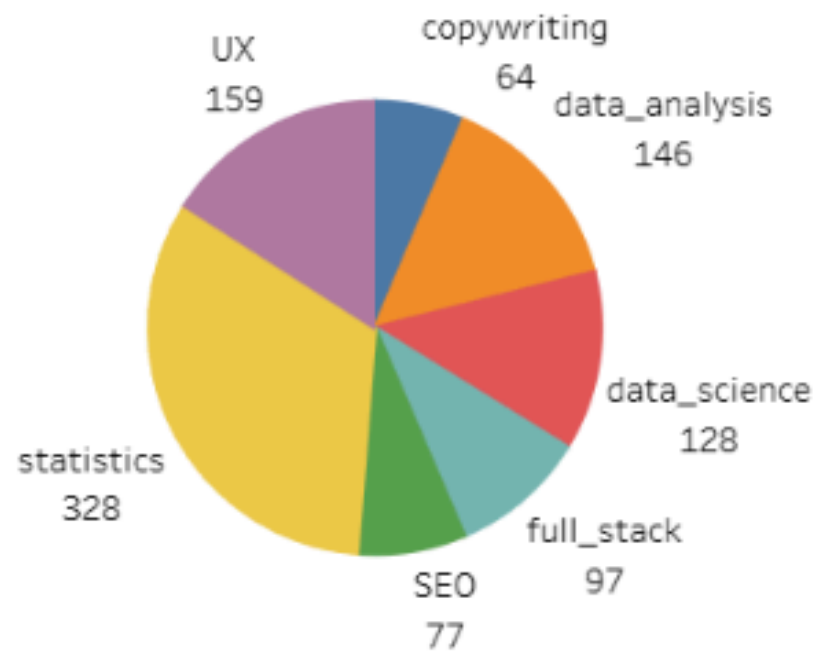
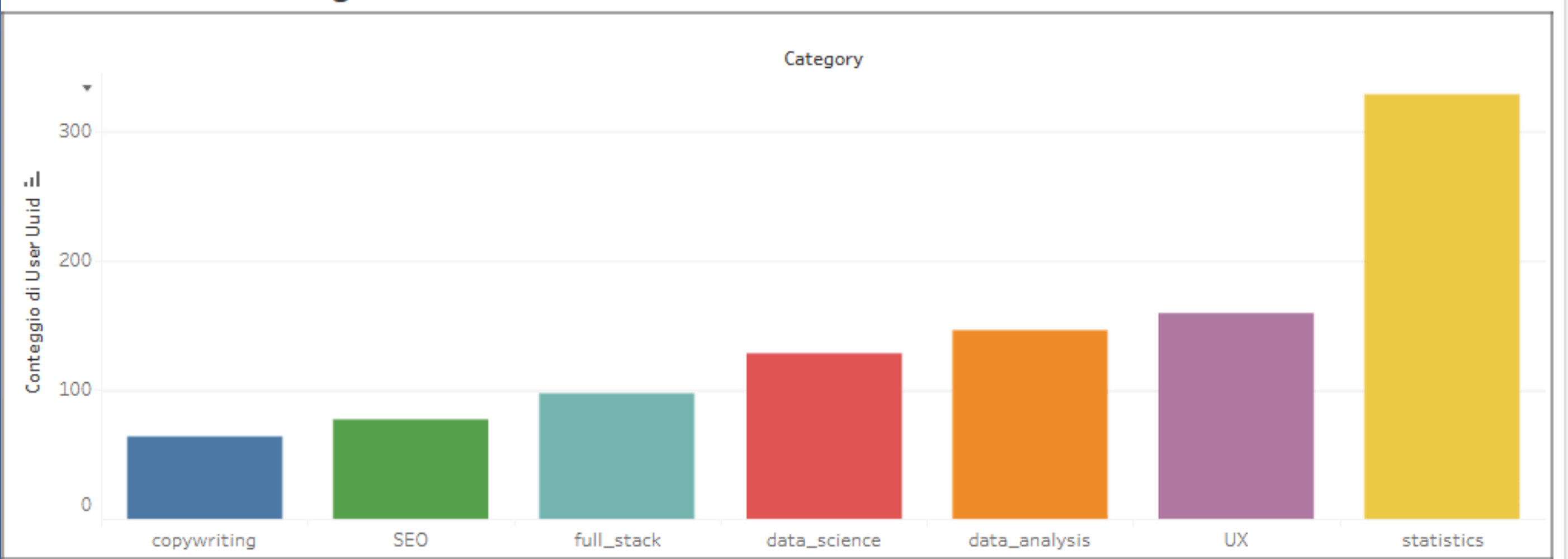
Link Python (JupyterNotebook):

https://drive.google.com/file/d/1LoJf6m7Mh2JVKyS1KlckFGheVoYmKoJW/view?usp=drive_link

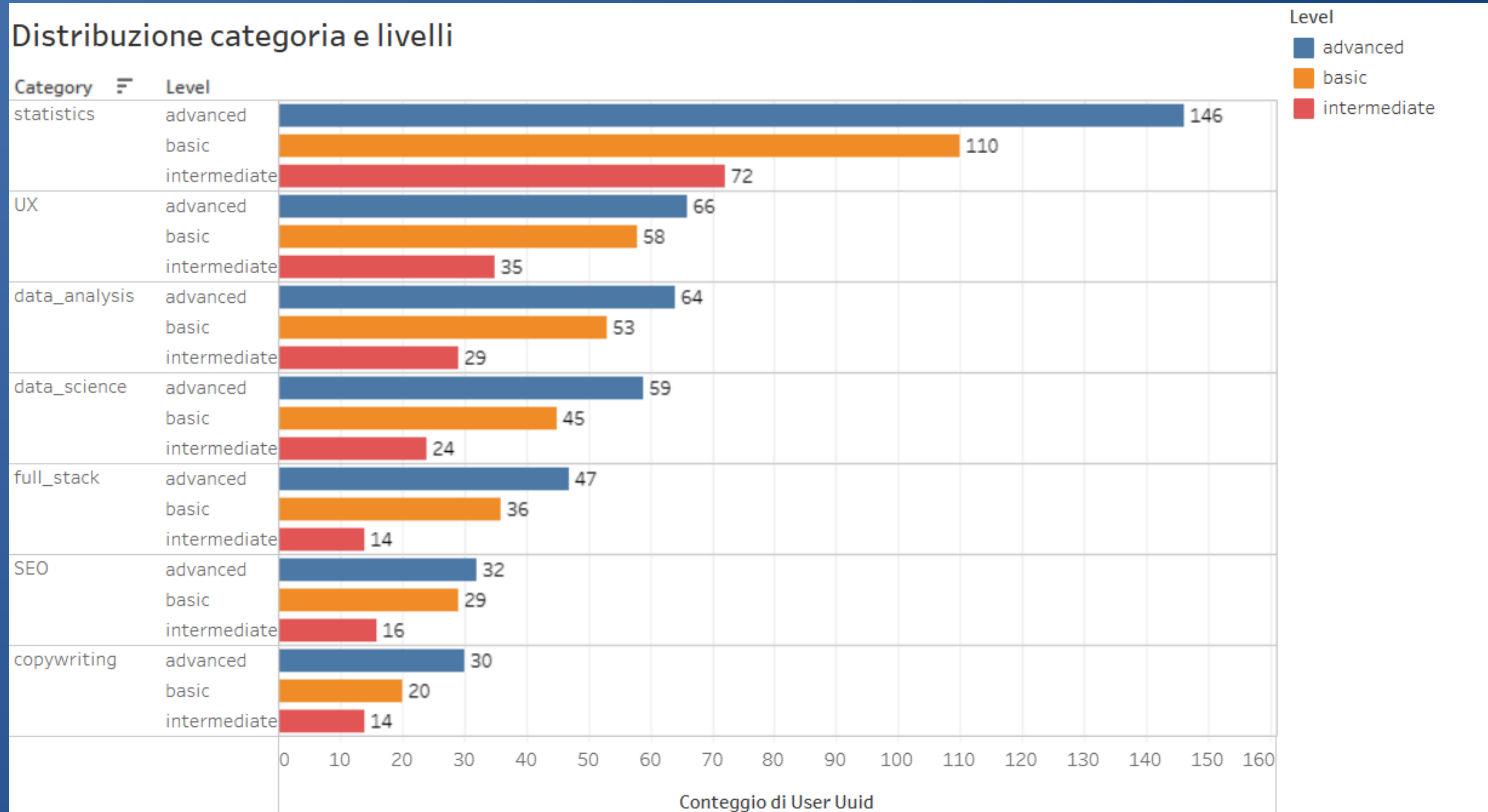
Link Tableau:

https://public.tableau.com/views/Progettofinale_17370467539980/Foglio1?:language=it-IT&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link

Come prima cosa ho provato ad individuare la distribuzione delle categorie:



Qui invece misto la categorie e il livello:

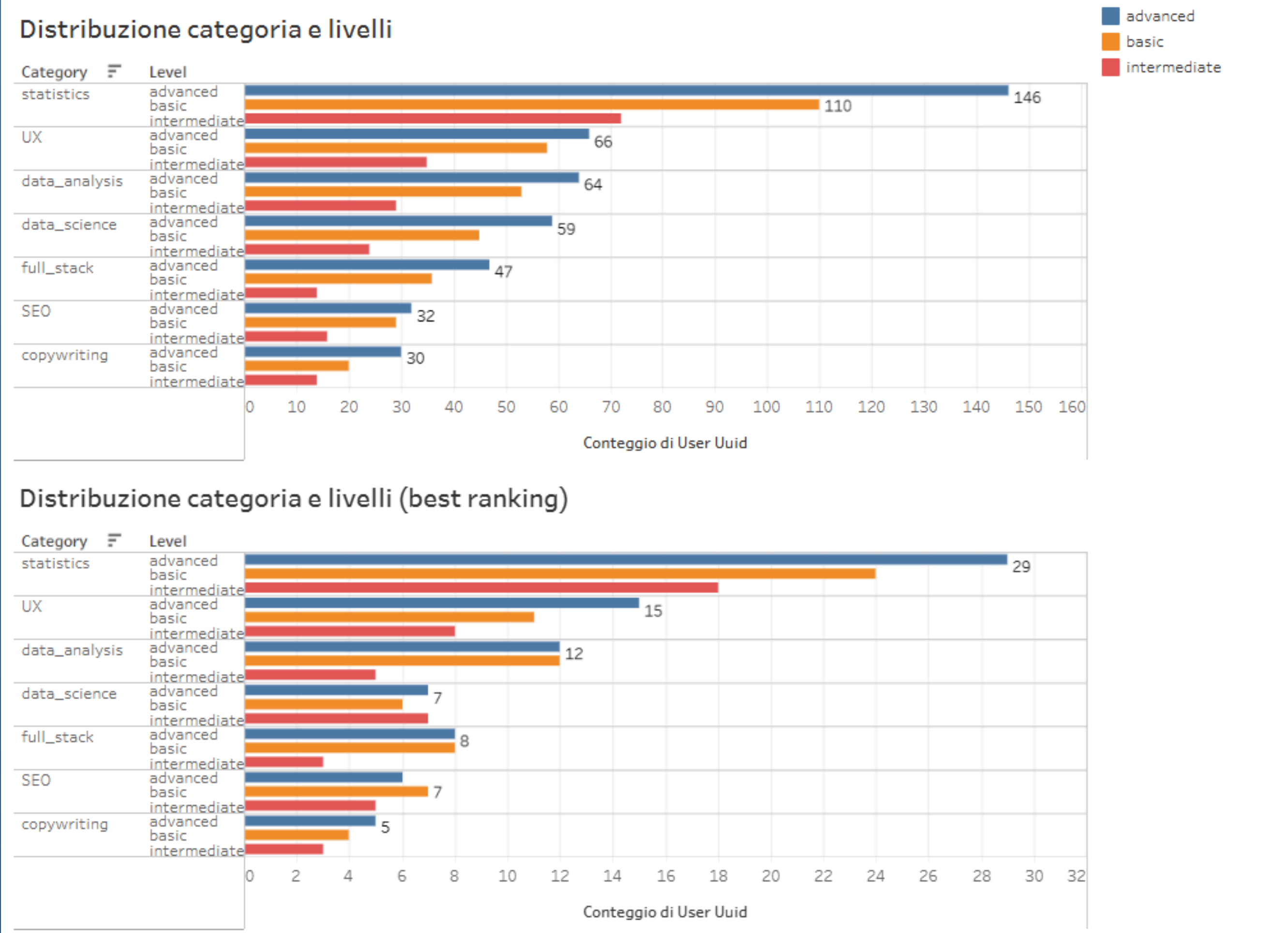


Come si può osservare nelle due Dashboard, la categoria maggiormente scelta è “Statistics” con molto distacco sulle categorie successive, le quali sono molto vicine come a quantità di osservazioni.

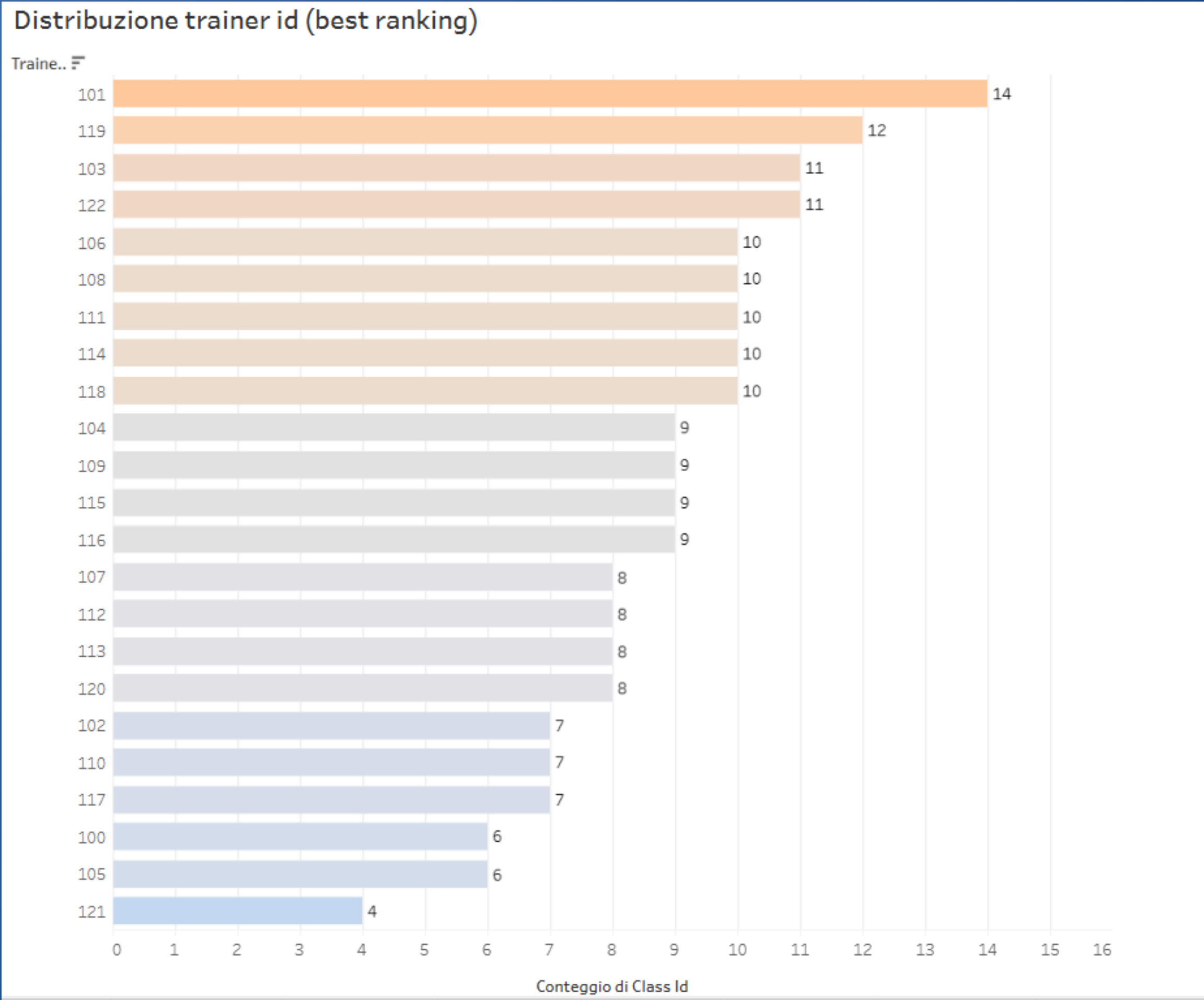
Per quanto riguarda i livelli, in tutte le categorie, il pattern è molto simile, con il livello “egemone” è quello avanzato

Ora, utilizzando la stessa tipologia di dashboard, provo a mettere a confronto solo le osservazioni che hanno ottenuto il massimo della valutazione (5 stelle)

Sia per le osservazioni generali, sia quelle a 5 stelle seguono lo stesso pattern



Prendendo sempre solo le osservazioni da 5 stelle possiamo osservare una panoramica degli insegnanti

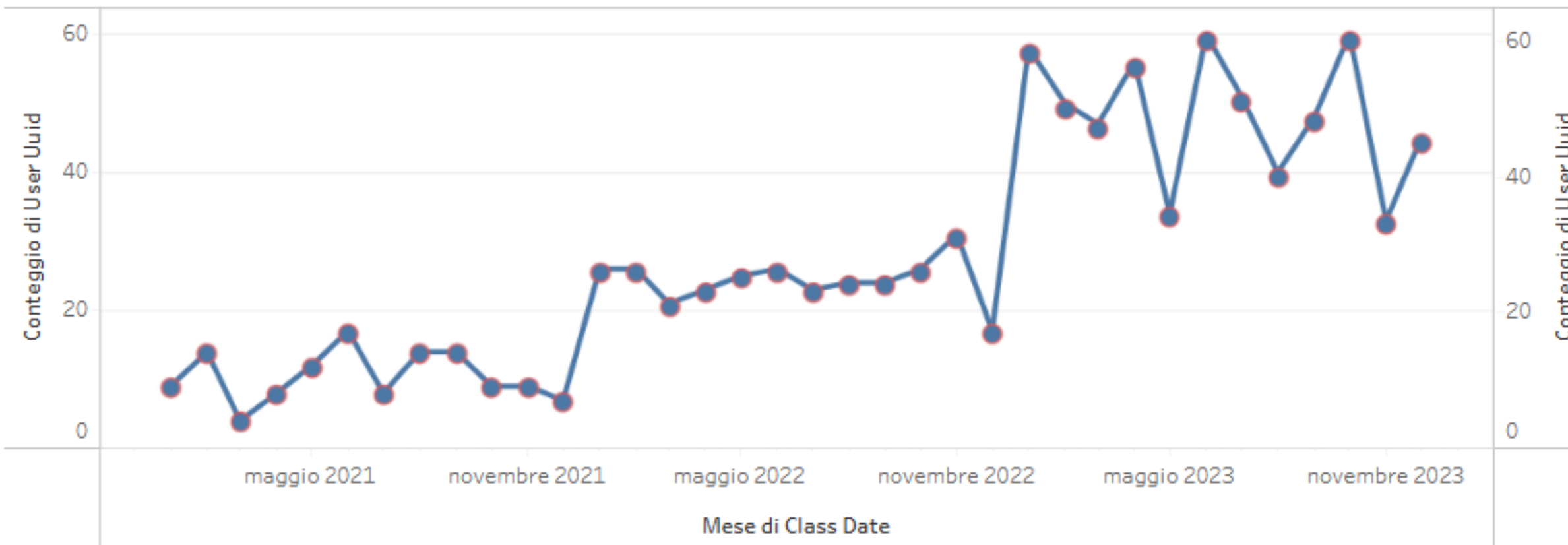


**La Dashboard precedente mostra, tra le osservazioni con 5 stelle di ranking, la distribuzione dei corsi per insegnante.
Si può ricavare, in questo modo quali siano gli insegnanti “più apprezzati”.**

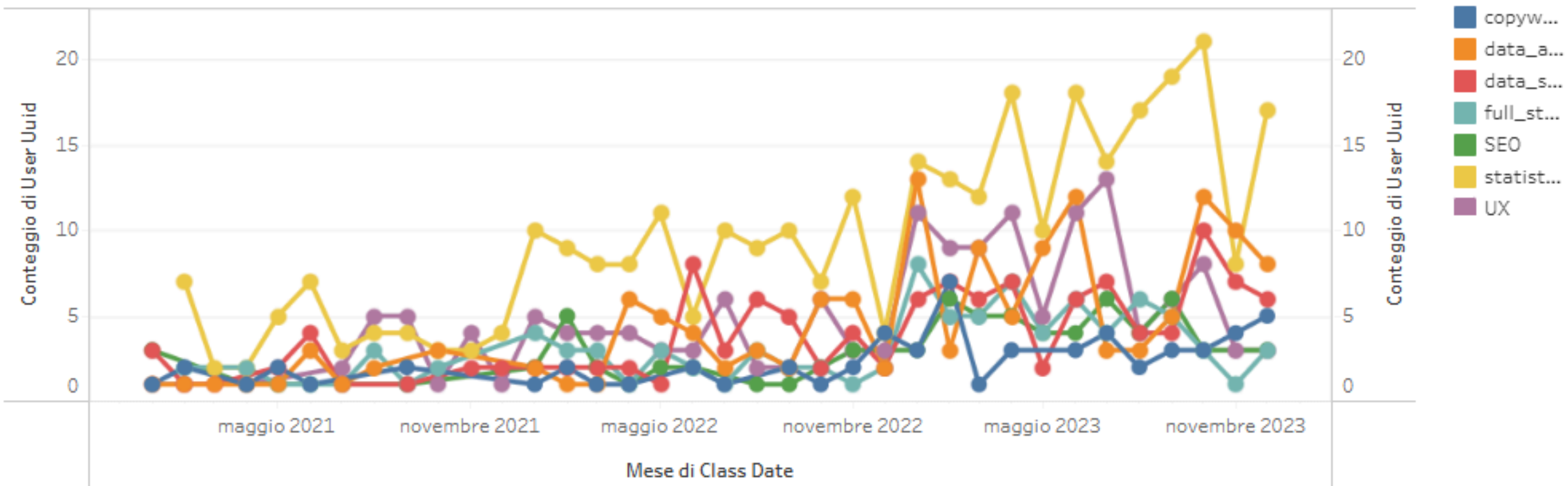
Un'altra questione importante può essere quella di vedere se esistono vari trend temporali, analizzando perciò, le iscrizioni alle classi

Andamento iscrizioni ai corsi

Iscrizioni per mese



Iscrizioni per mese (categoria)



Sono presenti due grafici a linee, il primo prende in esame le iscrizioni ai corsi in generale, mostrando la variazione nel corso degli anni presenti del dataset.


il secondo è suddiviso per categorie, in modo da poter controllare se ci sono dei trend, stagionali o non, relativi a particolari categorie

Guardando l'andamento delle iscrizioni ai corsi, si possono notare degli andamenti non regolari:

Nel 2021, si notano delle diminuzioni a marzo, luglio e nei mesi finali dell'anno. L'anno successivo vi è più stabile, salvo a dicembre dove c'è un calo evidente.

Il 2023 ha dei cali evidenti a maggio, agosto e novembre, classificandosi come l'anno più soggetto a variazioni. C'è da considerare, però, che c'è una crescita continua nel numero di iscrizioni e che il 2023 è di gran lunga quello con pi partecipazioni ai corsi.

Ora si può dare uno sguardo ai device tramite i quali gli studenti seguono i corsi:

Categorie e Device			
Category 	Platform	Conteggio di User Uuid	Utenti per categoria
statistics	mobile	86,0	328,0
	pc	106,0	328,0
	tv	136,0	328,0
UX	mobile	29,0	159,0
	pc	74,0	159,0
	tv	56,0	159,0
data_analysis	mobile	29,0	146,0
	pc	51,0	146,0
	tv	66,0	146,0
data_science	mobile	33,0	128,0
	pc	44,0	128,0
	tv	51,0	128,0
full_stack	mobile	24,0	97,0
	pc	32,0	97,0
	tv	41,0	97,0
SEO	mobile	20,0	77,0
	pc	19,0	77,0
	tv	38,0	77,0
copywriting	mobile	7,0	64,0
	pc	25,0	64,0
	tv	32,0	64,0

Nella tabella precedente si posson vedere le distribuzioni dei vari Device e le categorie.

Come già visto in precedenza, 'Statistics' è la categoria più presente, mentre per quanto riguarda i device in tutte le categorie (ad eccezione id UX), la maggior parte degli studenti utilizza la TV.

Considerazioni Finali:

Guardando il quadro completo si può affermare che la piattaforma è in netta crescita, guardando alle osservazioni del 2023. 'Statistics' è la categoria più di successo e il device più utilizzato è la TV.

Considerando anche solo le osservazioni con 5 stelle possiamo avere una panoramica sugli insegnanti e i corsi che hanno avuto più successo e partire da essi per preparare una strategia futura.