

Università degli Studi di Milano Bicocca



# SOCIAL MEDIA ANALYSIS

MATTEO CARCANO: 873258  
MARCO DONZELLA: 829358  
ANDREA MARINONI: 799690





# Tabella dei Contenuti

Punti chiave della presentazione

Introduzione

I dati

Social Network Analysis - Canale Youtube

Social Network Analysis - Python videos

Social Content Analysis

Conclusioni



# Introduzione



**Obiettivo:** studio delle interazioni che gli utenti hanno con il canale Youtube *Luke Barousse*

**Metodo:**

- Analisi dei commenti relativi ai video del canale
- Analisi dei commenti relativi una serie di video legati da un macro-argomento (Python)
- Studio del sentimento legato ai commenti

**Linguaggio di programmazione utilizzato:** Python

# Luke Barousse

"What's up, Data Nerds! I'm Luke, a data analyst, and I make videos about tech and skills for data science. If you have a question about becoming a data analyst, drop a comment in any video. No need to email, as I answer all comments!"



Luke Barousse è un *data scientist* statunitense emergente, che nell'agosto 2020 ha aperto il proprio canale Youtube.

Il canale conta, ormai, più di 130000 iscritti.

# I dati



I dati sono stati raccolti utilizzando le API Key concesse da Youtube. In questo modo è stato possibile scaricare i commenti più recenti legati ai video del canale.

Il dataset finale risulta caratterizzato da 4975 record (i commenti) e 11 variabili, riguardanti il commento, come ad esempio l'autore, il video, l'url, ecc.

Nel dataset sono presenti 65 video e 4017 utenti diversi che hanno commentato tali video.



Youtube API

Download dei  
commenti.



Pandas

Gestione del dataset.

# Social Network Analysis



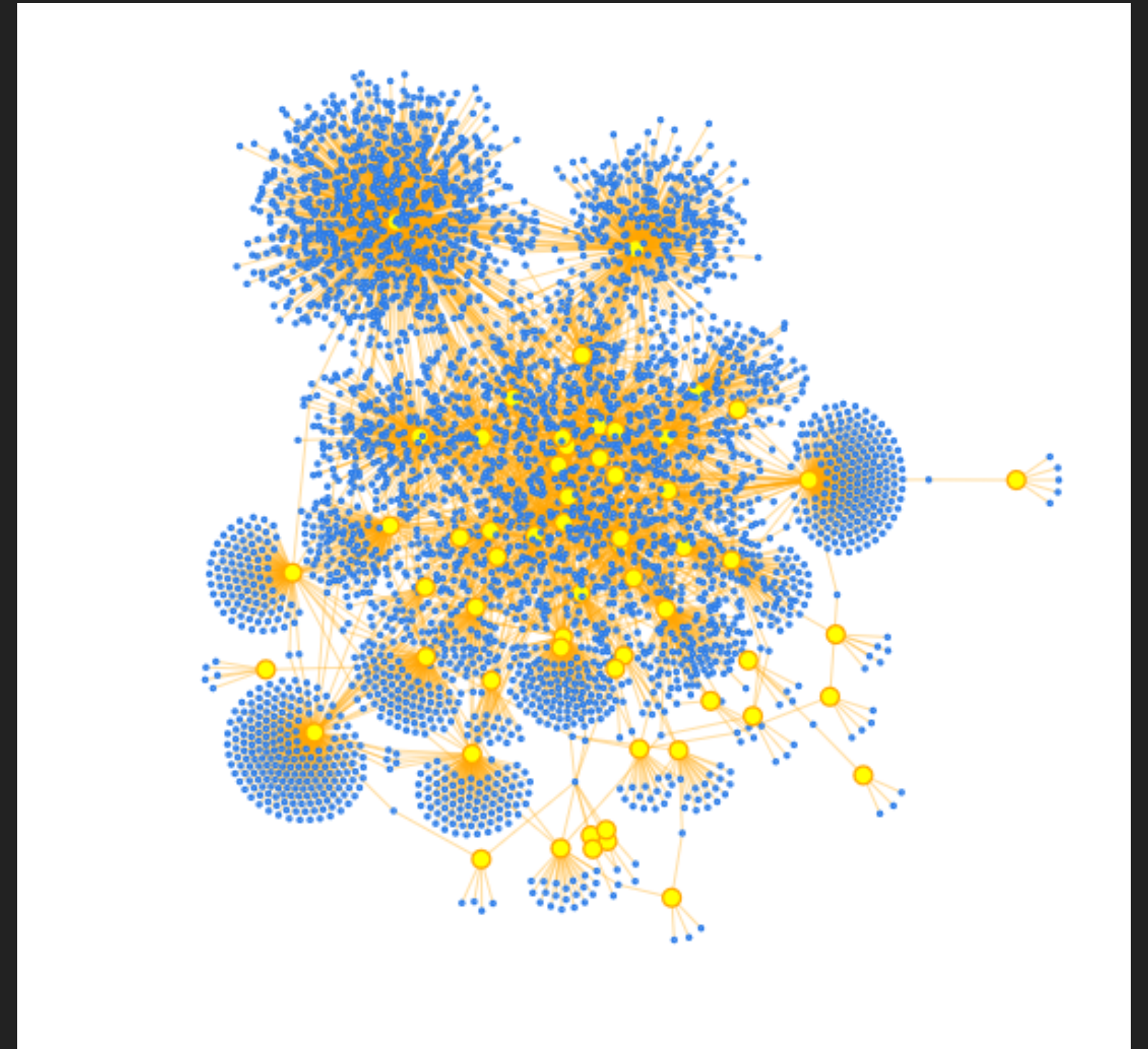
# Social Network Analysis - Canale Youtube

...

Si è costruito un grafo dove gli autori dei commenti e i video rappresentano i nodi e sono stati collegati tra di loro.

Lo si è effettuato per uno scopo prettamente esplorativo.

Si nota come numerosi utenti interagiscono con più video.

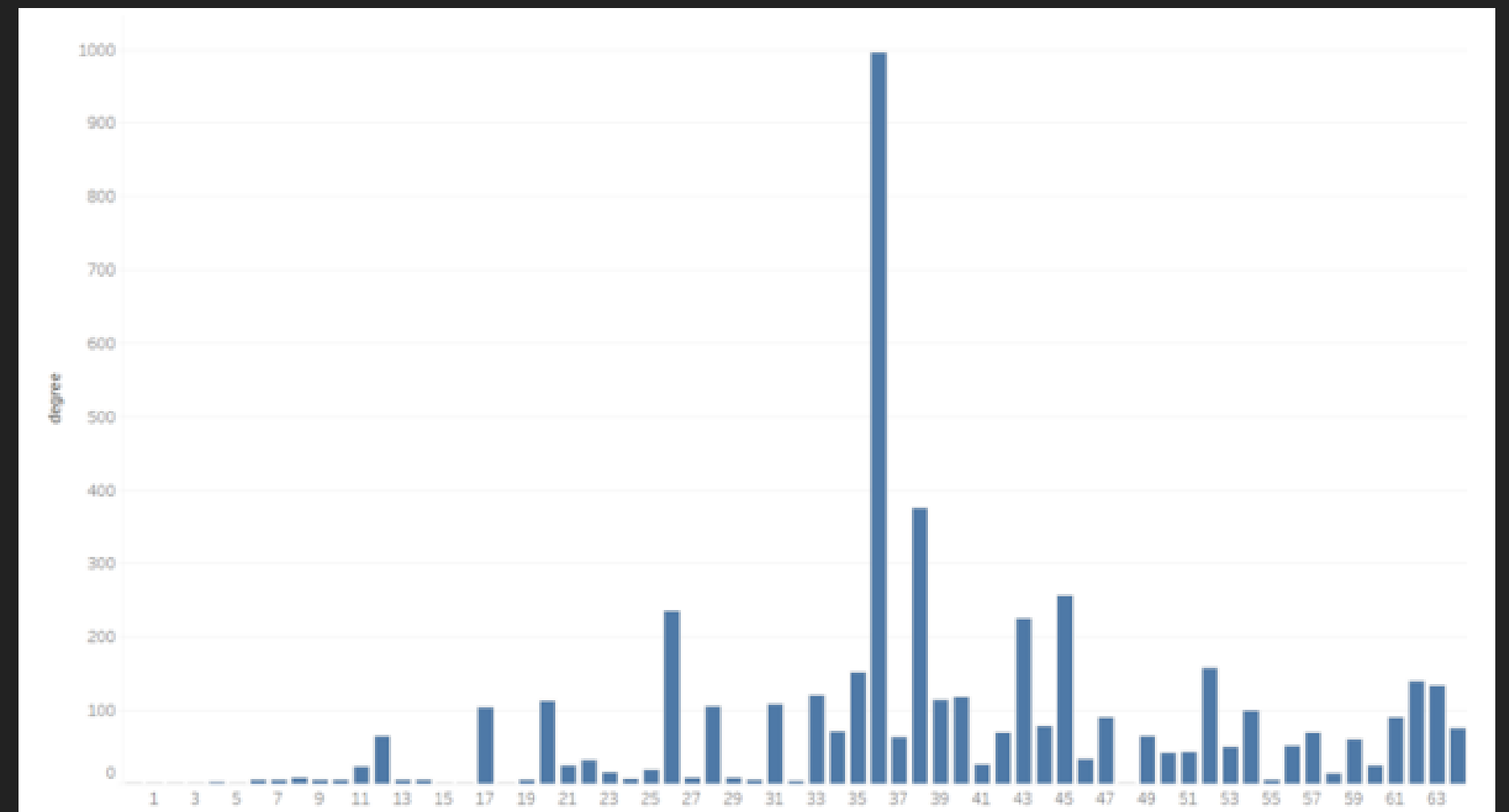




# Analisi esplorative dei video

...

Video	Grado
Become a DATA ANALYST with NO degree?!? The Google Data Analytics Professional Certificate	996
Google vs IBM Data Analyst Certificate - BEST Certificate for Data Analysts	376
Get a JOB w/ Google Data Analytics Certificate?!? (ft. Certificate Holders)	257
Windows on the M1 Mac - What are your options?	236
STOP using Spreadsheets for Everything!	226
How To Use Tableau Desktop Controls - Tableau Tutorial P.2	2
Calculated Fields in Tableau (Formulas & IF Statements) - Tableau Tutorial P.6	2
Top Non-technical Skills for Business Intelligence	1
What is Business Intelligence (BI)? #shorts	1
Conditional Format Tables in Tableau (Like Excel!) - Tableau Tutorial P.5	1





# Social Network Analysis - Python Videos



Per motivi computazionali e di interesse, si è ridotto il campo di analisi alle playlist python basics e python setup.

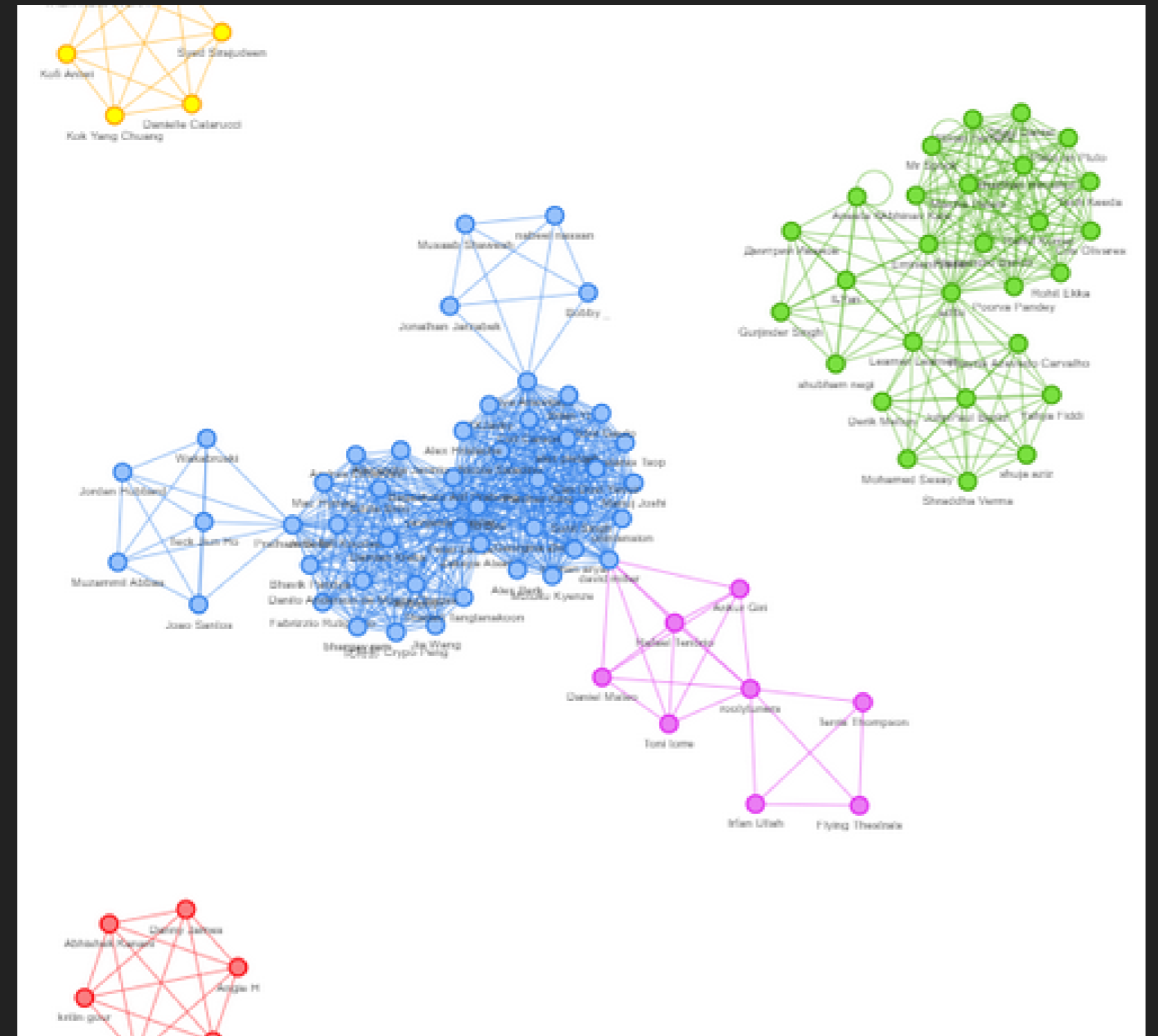
Si sono impostati gli utenti come nodi e i video come archi.

Il grafo ricavato risulta composto da 96 nodi e 702 archi.



# Community Detection

Si osserva la presenza di  
5 community,  
corrispondenti a  
particolari aspetti della  
playlist



# Modularity and Assortativity

- Il valore del coefficiente di modularità della rete è pari a 0.477. La composizione randomica della rete è quindi esclusa, e il numero di archi presenti nella rete è maggiore del numero di archi attesi.
- Il valore del coefficiente di assortatività della rete è pari a 0.482. Si evince che la rete sia assortativa e quindi i gradi dei nodi della rete sono correlati tra loro positivamente.



# Social Content Analysis



# Sentiment Analysis

## Vader

è un modello utilizzato per l'analisi del sentiment testuale sensibile sia alla polarità in termini di positiva/negativa, sia all'intensità delle emozioni.

## Textblob

é un modello utilizzato per l'analisi dei sentiment testuali.

il sentiment finale viene calcolato da operazioni di pooling come l'utilizzo di una media di tutti i sentimenti

# Vader



L'analisi Vader prende una stringa e restituisce un dizionario di punteggi in ciascuna delle quattro categorie:

- **negativo**  $\Rightarrow$  calcola la percentuale che un testo sia negativo da 0 a 1
- **neutro**  $\Rightarrow$  calcola la percentuale che un testo sia neutro da 0 a 1
- **positivo**  $\Rightarrow$  calcola la percentuale che un testo sia positivo da 0 a 1
- **composto**  $\Rightarrow$  calcolo normalizzato dei 3 indici sopra, range compreso tra -1 (negativo) e +1 (positivo)



# Text-Blob



TextBlob restituisce la polarità e la soggettività di una frase.

La polarità è compresa tra  $[-1, 1]$ , -1 definisce un sentimento negativo e 1 definisce un sentimento positivo. Le parole di negazione sono comunque valutate e comprese ed invertono la polarità.

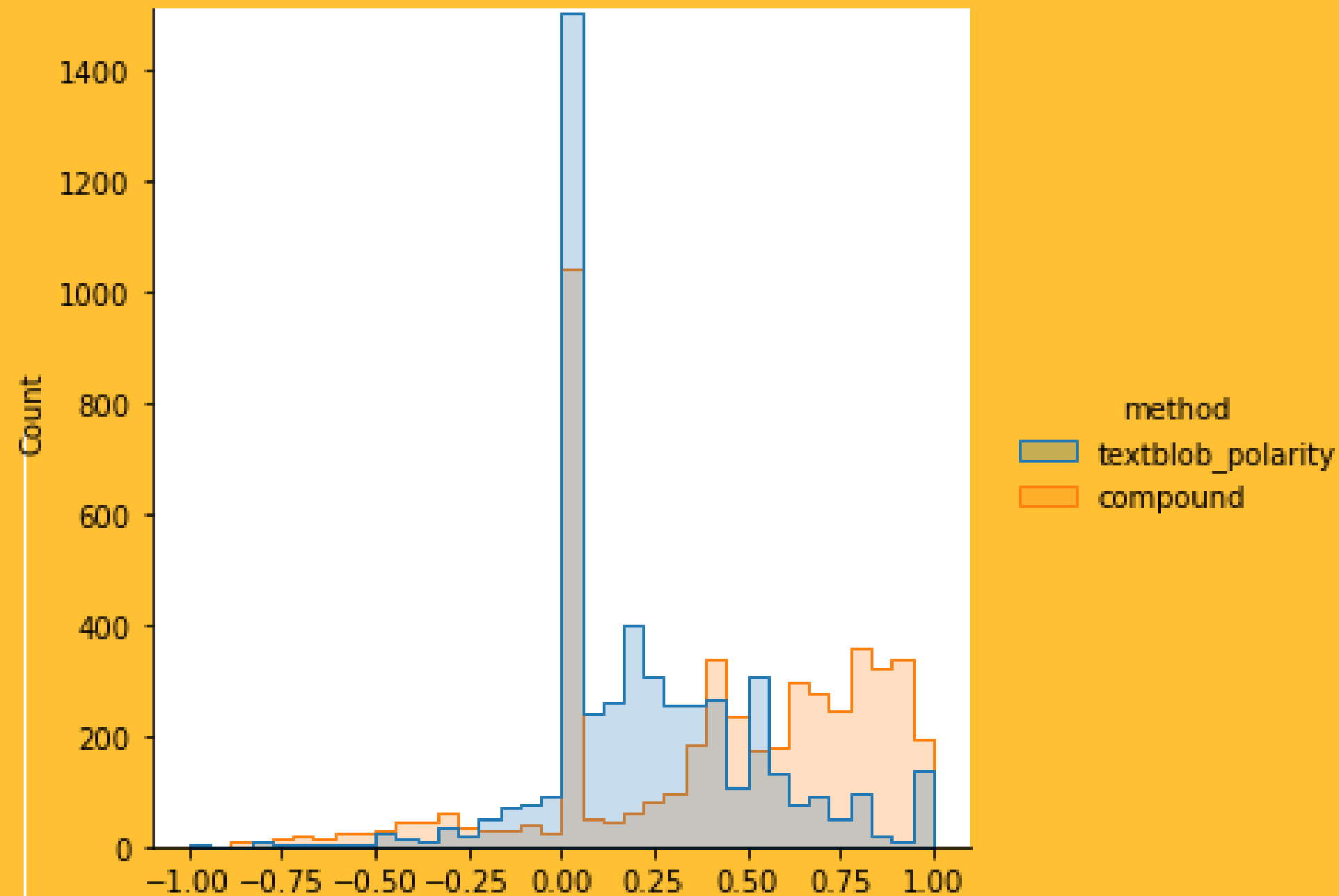
- 





# Confronto Text-Blob/ Vader

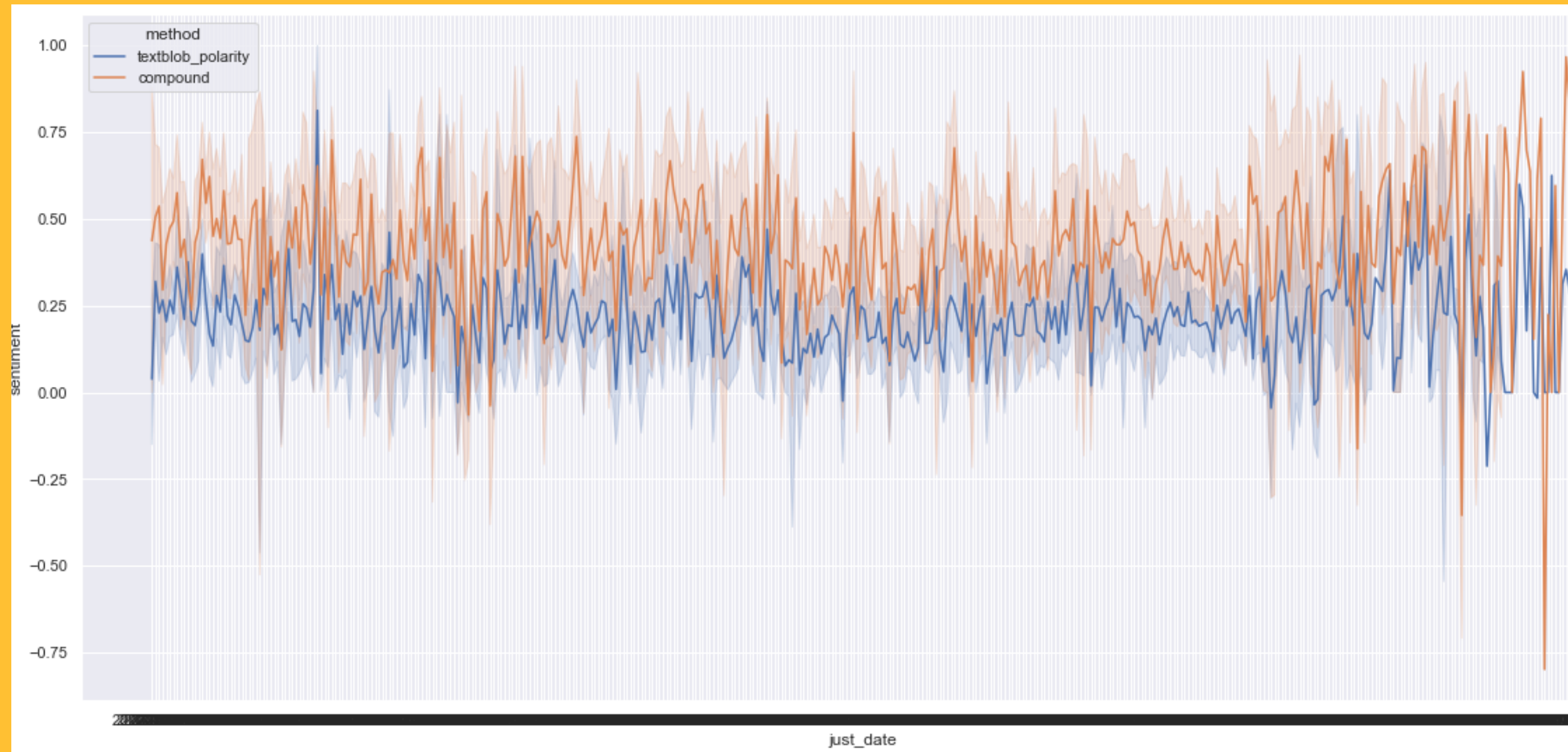
SOCIAL MEDIA REPORT



- Presenza di un grande accumulo di commenti con sentiment 0.0
- è evidente come la maggior parte dei commenti risultano essere positivi

# Andamento Text-Blob/ Vader

SOCIAL MEDIA REPORT



Le serie storiche mostrano andamenti costanti nel tempo ad eccezione dell'ultimo periodo. Si evidenzia infatti un aumento della variazione, simbolo che i nuovi contenuti proposti nel canale polarizzano di più gli utenti

# Migliori Video

SOCIAL MEDIA REPORT

	compound
vid_title	
Dimensions Vs Measures (Blue Vs Green Data) - Tableau Tutorial P.3	0.842350
Python Functions for Data Science / Data Analysis - P.5	0.838717
M1 vs Intel Mac for Python #shorts	0.830200
Create Stacked Bar Chart (and any other visuals EASILY!) w/ Show Me! - Tableau Tutorial P.4	0.808750
Parameters (Create & Use in Calculated Fields and/or Visuals) - Tableau Tutorial P.7	0.768333
...	...
Install VS Code with Python for Data Science / Data Analysis - P.3	0.331717
Python for M1 Mac vs Intel (SPOILER: M1 is 2x faster)	0.323996
Become a DATA ANALYST with NO degree?!? The Google Data Analytics Professional Certificate	0.308248
Install Python for Data Science on Mac & Windows (PC) with Anaconda - P.1	0.139120
M1 Chip is as FAST as M1 Max!!! 🤖 (13" Mac Air Vs. 14" Mac Pro) #shorts	0.103780

Questa esplorazione può aiutare il proprietario del canale a meglio comprendere i propri contenuti video e capire quali hanno avuto più successo e quali invece sono stati ritenuti scadenti dalle community del suo canale

# Emotion Analysis

SOCIAL MEDIA REPORT

L'Emotion Analysis è il processo di identificazione e analisi delle emozioni celate e nascoste espresse nei dati testuali.

L'analisi emotiva utilizza un sistema complesso per comprendere le risposte dei consumatori, analizza in modo approfondito una gamma completa di emozioni e sensibilità umane.

Anche scoprire una risposta confusa può rivelare che i tuoi contenuti sono troppo complicati e devi fornirne di più chiari.

	emotion	comment
0	{'anticipation': 1, 'joy': 1, 'positive': 1, '...	What is the average salary?
1	{'positive': 3, 'trust': 2, 'anticipation': 1}	Hi. What do you recommend For stata ana spss? ...
2	{'positive': 14, 'anticipation': 8, 'trust': 1...	to talk about my self , i started university w...
3	{'positive': 2, 'anticipation': 1}	#1 skill is learning how to identify what the ...
4	{'positive': 7, 'joy': 4, 'trust': 5, 'anticip...	Buying an exorbitantly priced Mac to end up ru...



# Conclusioni

## Social Network Analysis

- Presenza di 5 comunità
- La modularità trovata esclude una possibile composizione randomica della rete
- La rete risulta quindi assortativa e perciò i gradi dei nodi della rete sono correlati tra loro positivamente.

## Social Content Analysis

- Il canale di Luke Barousse produce contenuti apprezzati
- I contenuti stanno diventando col tempo polarizzanti