

The background of the image is a dark, atmospheric concert scene. A band is performing on a stage illuminated by bright yellow and white spotlights. In the foreground, the silhouettes of many audience members' hands are raised in the air, creating a sense of energy and participation. The overall mood is vibrant and celebratory.

Spotify Explorative Analysis

Levratti Mattia – matr. 139022
A.A. 2020/2021



Introduzione

- Analisi esplorativa del dataset ‘Spotify Dataset 1922– 2021 ~ 600k Tracks’
(<https://www.kaggle.com/yamaerenay/spotify-dataset-19212020-160k-tracks>)
- Descrizione del panorama attuale
- Analisi delle Features di questo dataset
- Potenziali consigli per musicisti interessati a pubblicare
- BIAS: brani prevalentemente presi dal mercato US



Descrizione del dataset

- Chiave primaria: id
- Features quantitative:
 - Acousticness
 - Danceability
 - Energy
 - Duration_ms
 - Instrumentalness
 - Value
 - Popularity
 - Tempo
 - Liveness
 - Loudness
 - Speechiness
- Features qualitative:
 - Key
 - Timesignature
 - Artists
 - Id_artists
 - Release_date
 - Name
 - Mode
 - Explicit

Analisi delle features



Brani pubblicati e durata dei brani

Pubblicazioni:

- Andamento globalmente crescente, calo negli anni 2000 (problema del dataset)
- Mediana di 5591 brani, picco massimo di 13937 brani
- Picco nel 2020 (Covid?)

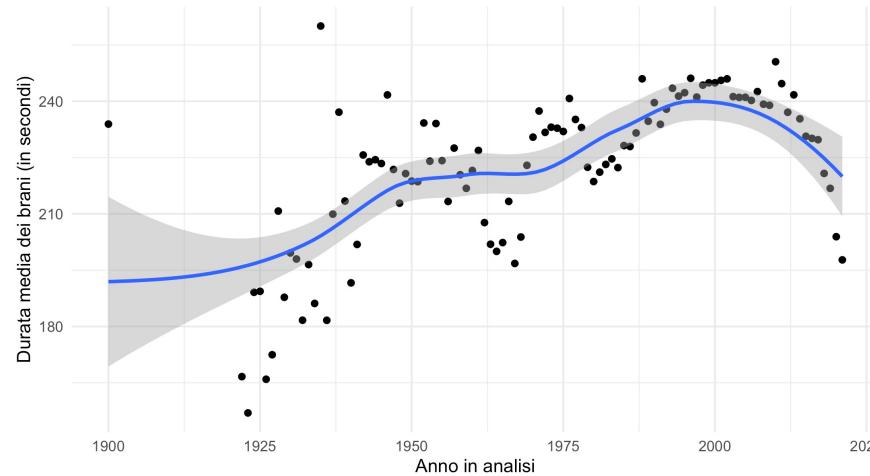
Durata:

- Andamento crescente fino al 2000, poi decrescente
- Mediana di 226 secondi, picco di 260 secondi (nel 1935)



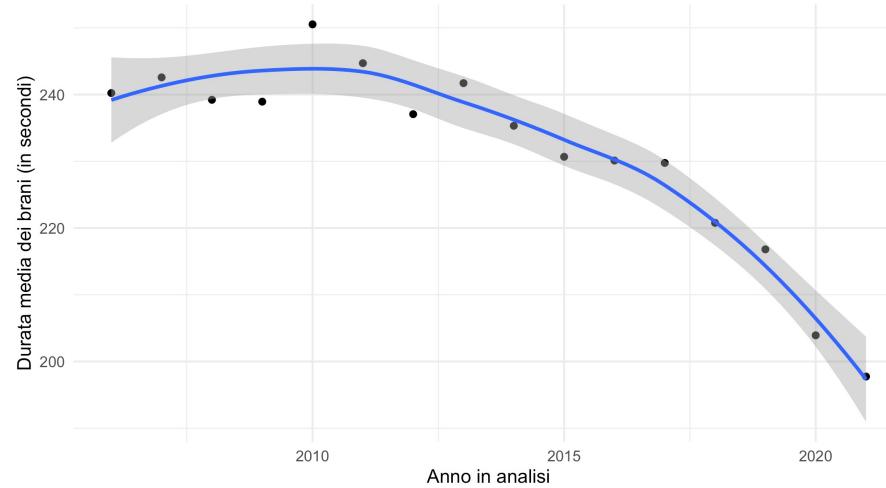
Durata media dei brani nel tempo

In aumento fino agli anni '90 e poi in decrescita



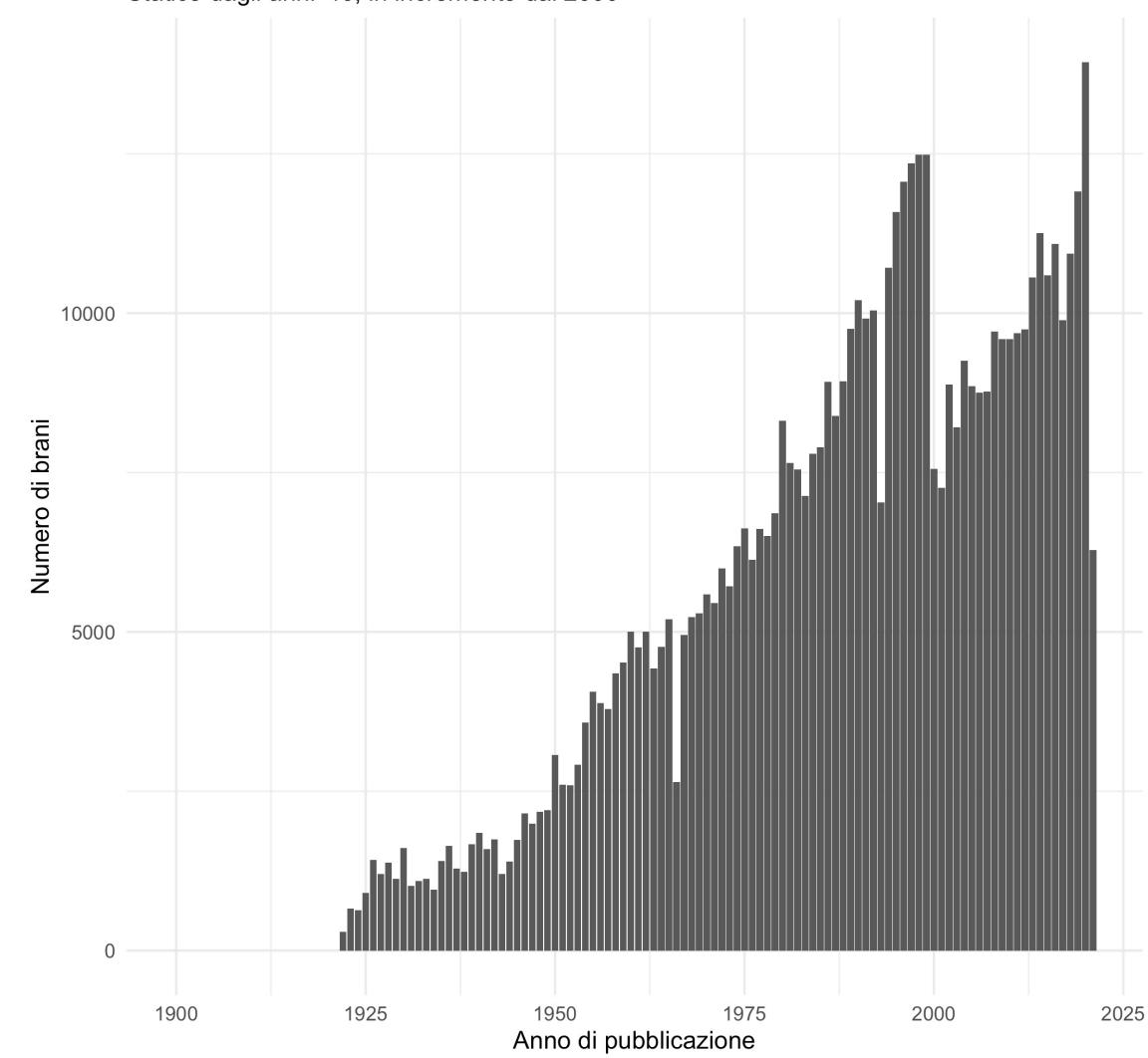
Durata media dei brani nel tempo

Focus sull'inversione di tendenza iniziata dal 2011



Numero di brani pubblicati nel tempo

Statico dagli anni '40, in incremento dal 2000





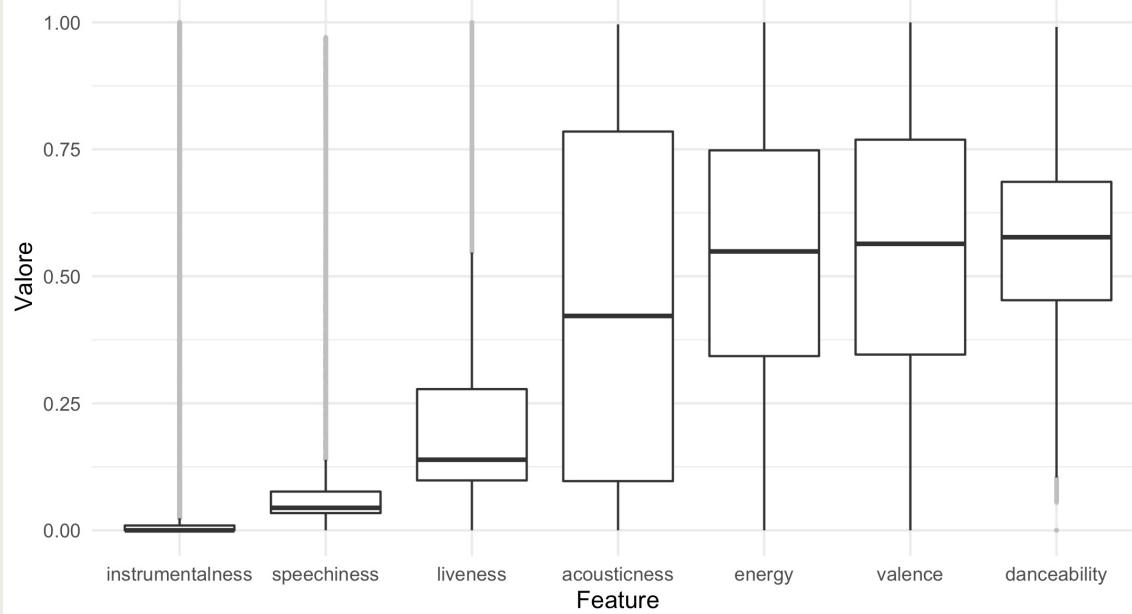
Distribuzione delle features musicali

- Instrumentalness, speechiness e liveness assumono valori tendenzialmente bassi
- Energy, valence e danceability assumono valori tendenzialmente alti
- Acousticness assume il più ampio range di valori (pressoché tutti)



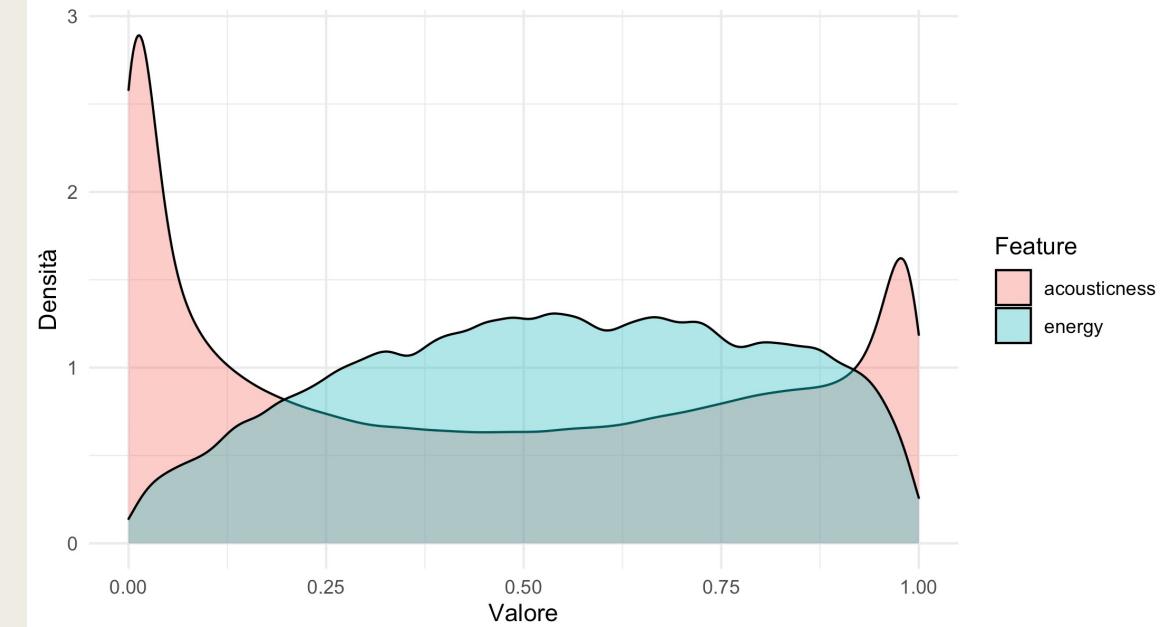
Visualizzazione dei valori assunti dalle varie features

Analisi visiva mediante boxplot



Confronto distribuzione di acousticness ed energy

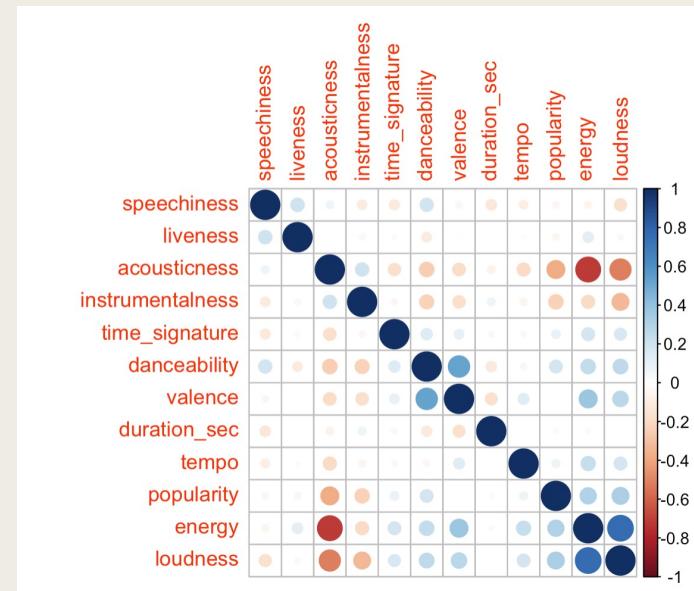
Distribuzione nettamente differente





Correlazione tra le features

- Analisi tramite coefficiente di correlazione di Pearson
 - Forte correlazione diretta tra Energy - Loudness ed Valence – Danceability
 - Forte correlazione inversa tra Energy - Acousticness ed Acousticness - Loudness





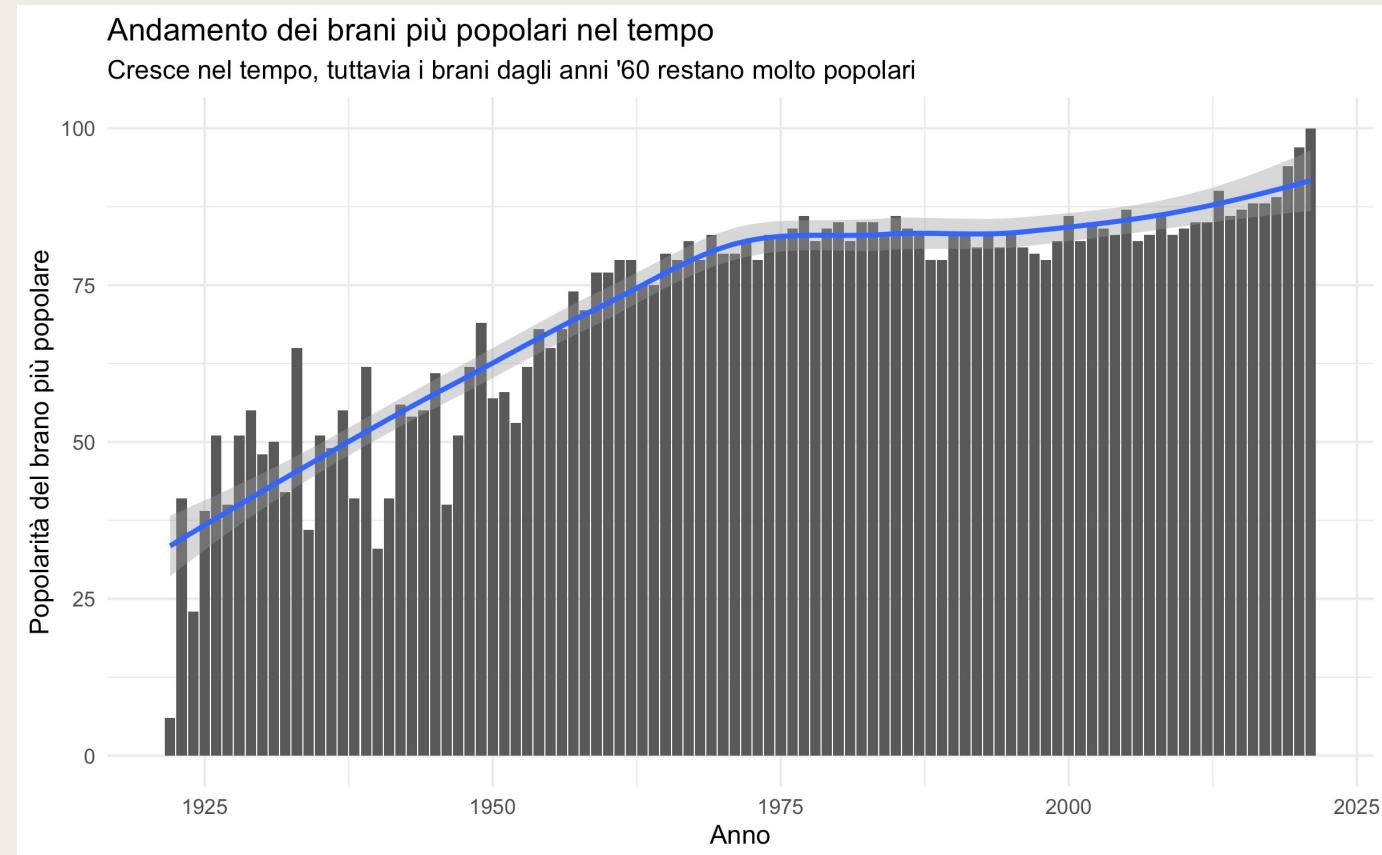
Analisi compositiva: Toni, modi, metrica

- Modo maggiore molto più utilizzato del modo minore
- Utilizzate maggiormente le tonalità con meno alterazioni in chiave (Do, Sol, Re, La, Fa)
- Utilizzati maggiormente tempi binari e quaternari (feature non completamente attendibile)

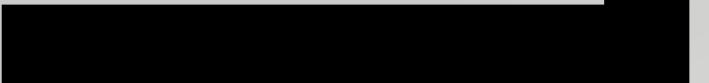


Popolarità dei brani

- Feature non totalmente attendibile in quanto legata al periodo storico in cui si effettua l'analisi
- Andamento crescente della popolarità dei brani, con flessione della crescita attorno agli anni '70

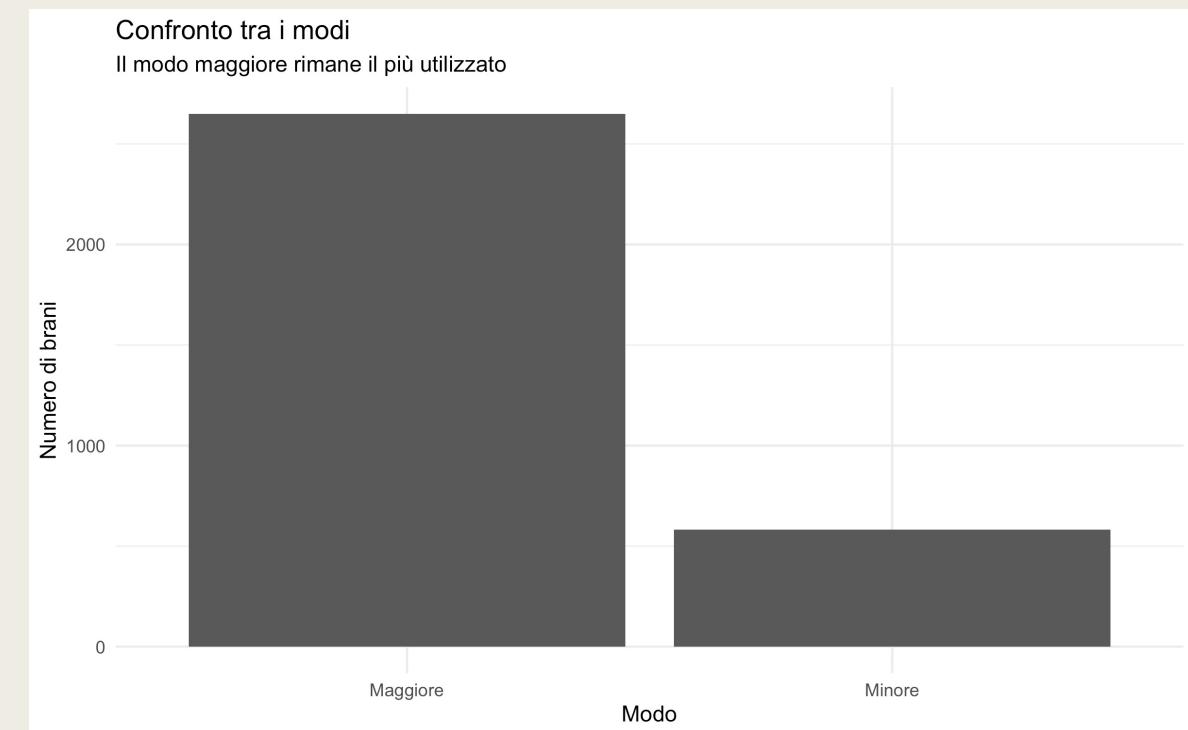
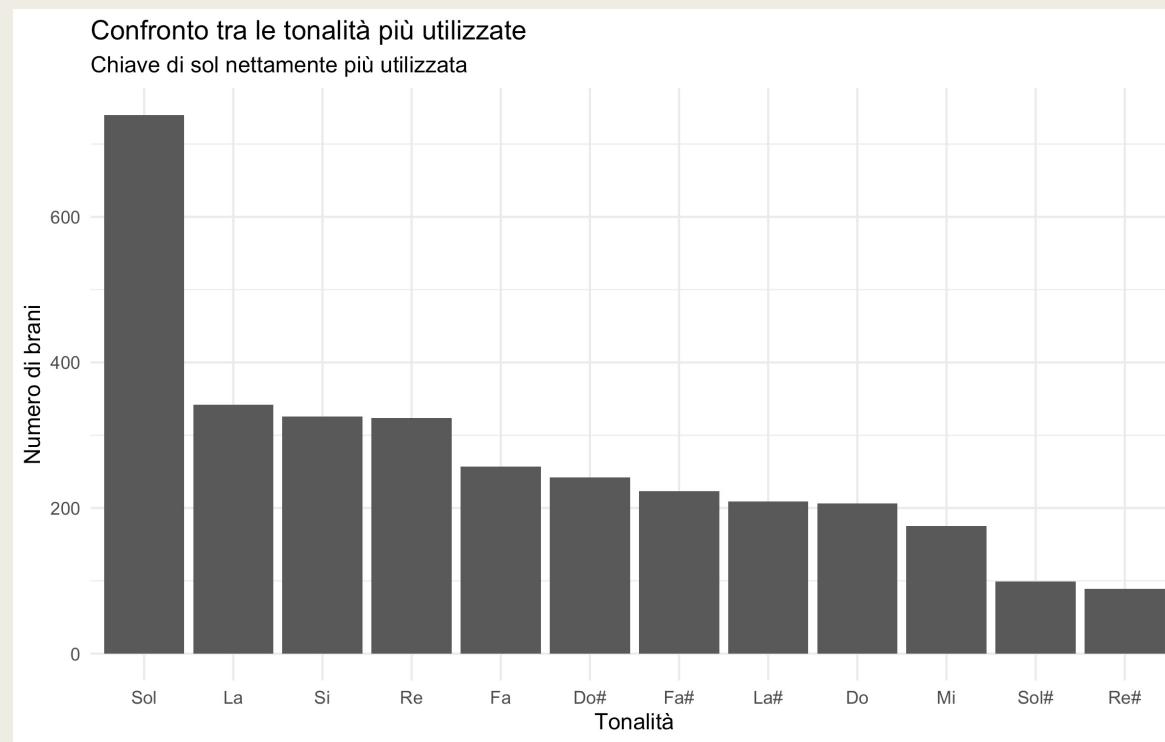


Analisi dei generi musicali





- Generi più popolari: sottogeneri afferenti al pop, indie, rap
- Tonalità più utilizzate mediamente nei generi più popolari: sol, la, si, re, fa
- Modo maggiore sempre il più utilizzato



Analisi dei trend musicali nel tempo



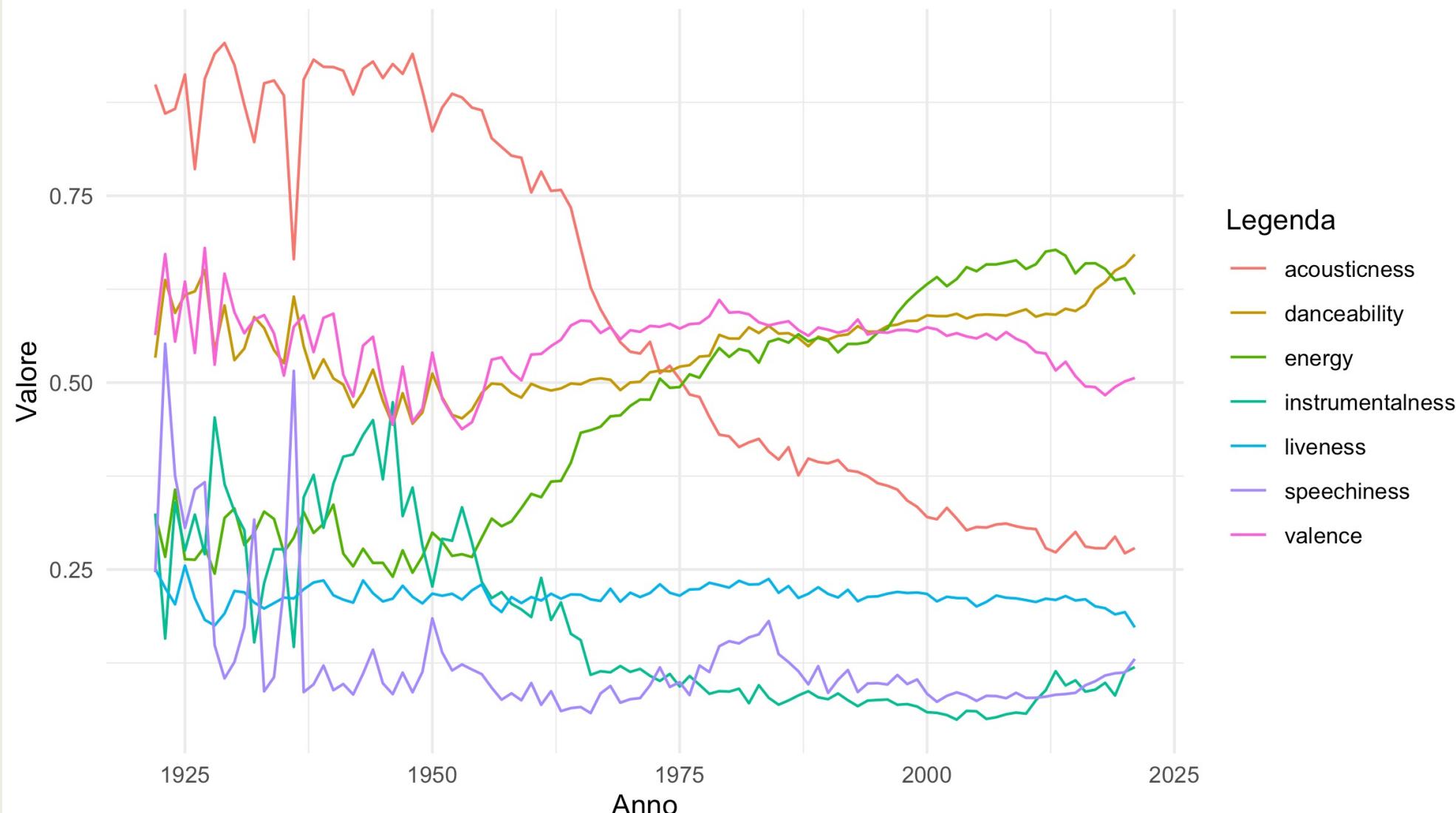


- Fino agli anni '50 andamento irregolare dei dati
- Stabilizzazione sopra allo 0.5 di
 - *Energy*
 - *Danceability*
 - *Valence*
- Stabilizzazione sotto allo 0.5 di
 - *Acousticness*
 - *Liveness*
 - *Instrumentalness*
 - *Speechiness*
- Andamento nettamente contrastante di Acousticness ed Energy



Andamento delle features nel tempo

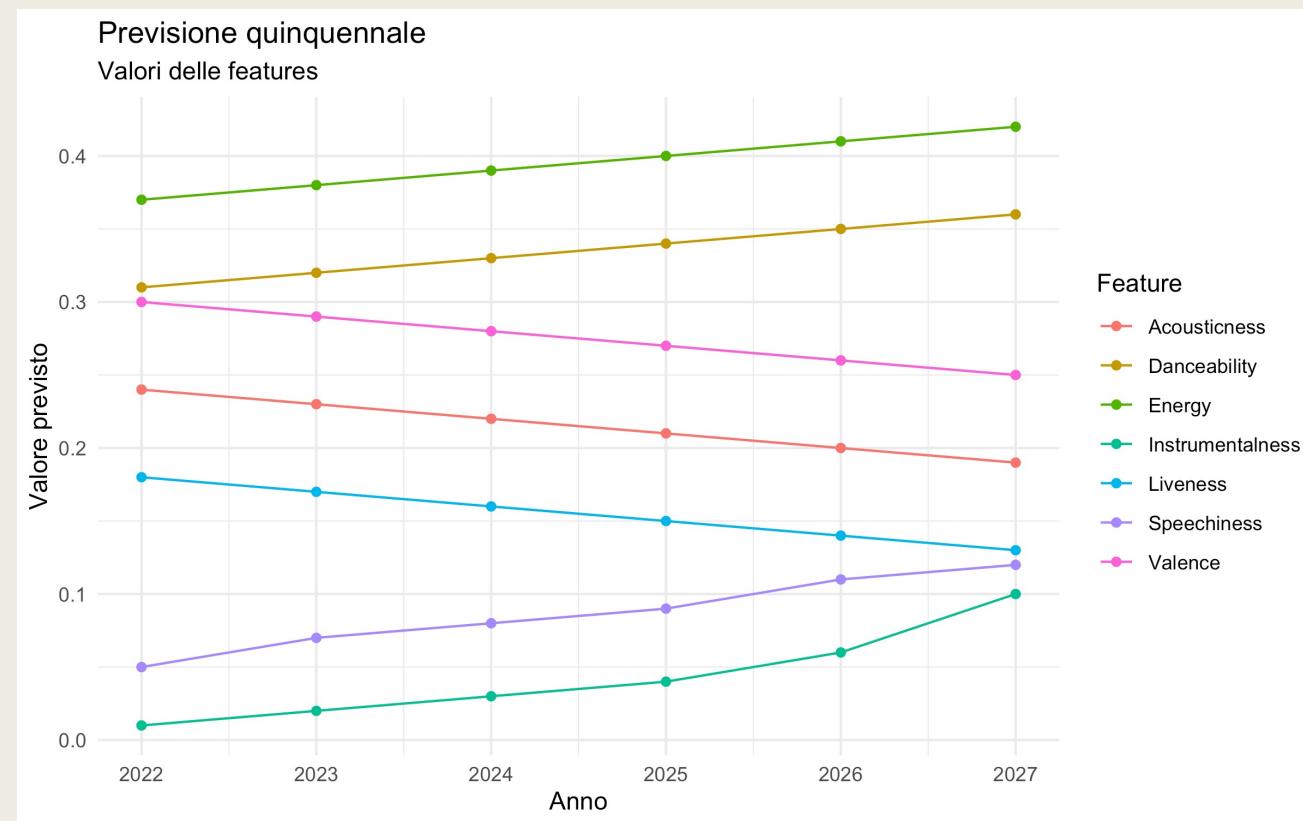
Interessante andamento di Acousticness ed Energy



Potenziali sviluppi
futuri



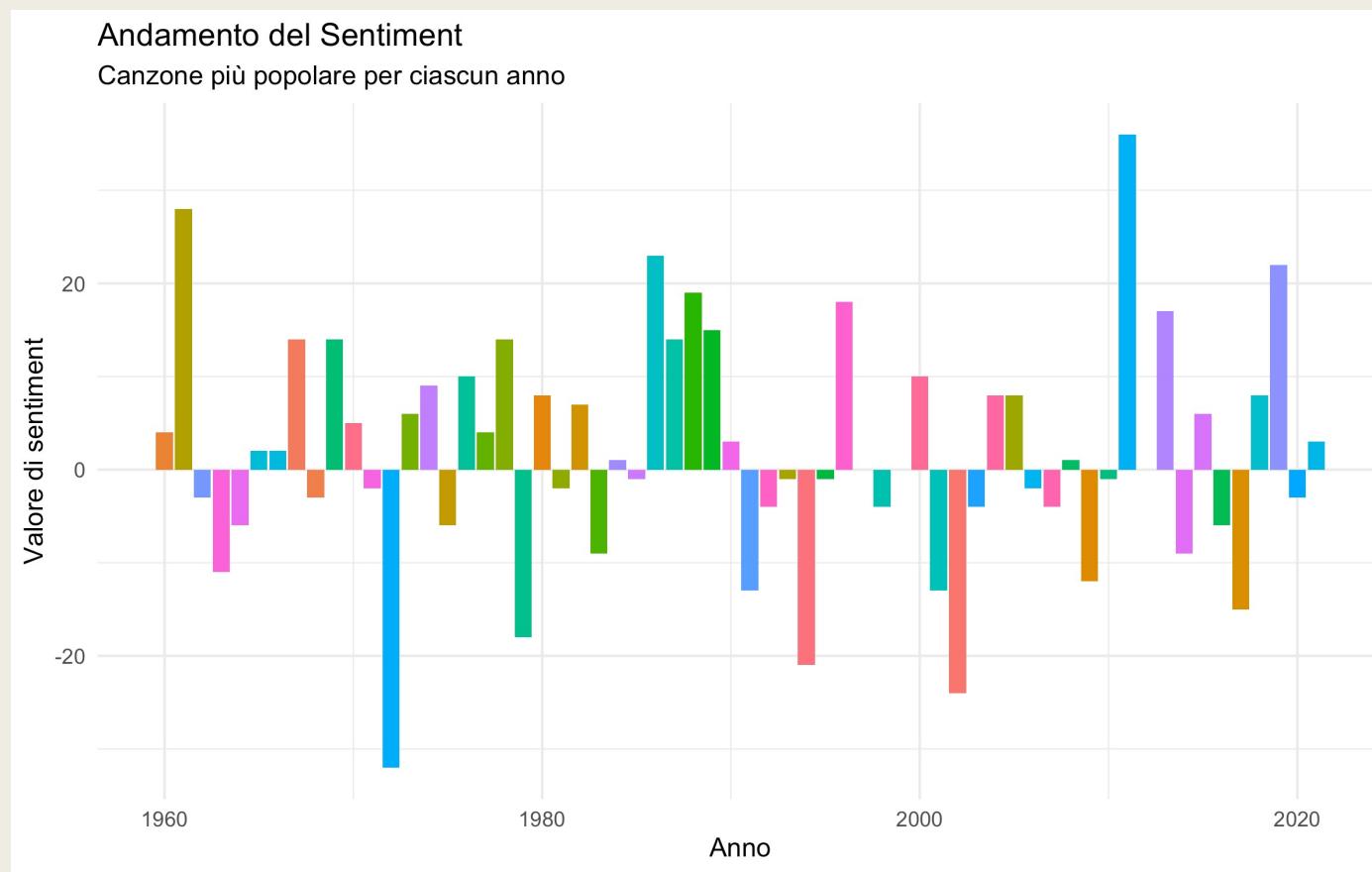
- Previsione effettuata tramite regressione polinomiale di grado 2
- In crescita Energy, Danceability, Speechiness e Instrumentalness
- In decrescita Valence, Acousticness e Liveness



Sentiment Analysis



- Eseguita a livello dei brani più popolari per ciascun anno
- Metodologia BING
- Non si rilevano particolari pattern



Conclusioni



Come dev'essere il mio brano?

- Durata leggermente inferiore ai 200 secondi
- Tonalità semplici (Do, Sol), modo maggiore e metriche binarie (2/4, 4/4)
- Energico, ballabile, con un volume sostenuto
- Non troppo strumentale o parlato



... infatti Bohemian Rhapsody

(Queen, 1975, prima in classifica in UK, Belgio, Nuova Zelanda, Paesi Bassi,
eletta Canzone del secolo in UK nel 2000, Grammy Hall of Fame)

- Ha una metrica completamente irregolare
- Usa tonalità con alterazioni costantemente presenti
- Varia su almeno quattro generi musicali, con parti totalmente diverse tra loro
- Dura circa 3 volte di più rispetto alla media