

Report di analisi Festival di Sanremo

TRACCIA ESAME

The screenshot shows a slide from an EPICODE exercise. The top left features the EPICODE logo. The top right contains the text "Esercizio" and "Schema concettuale". The main content area contains the following text:

Vi sono stati forniti i dati relativi alle edizioni di Sanremo dal 1951 al 2023, organizzati in tre tabelle.

Effettuate un'analisi dei dati, scegliete una tematica che volete indagare (ad esempio, la differenza di genere nei vincitori, le ristampe delle canzoni, l'avvicendarsi dei presentatori... qualsiasi cosa vi appassiona va bene).

Effettuate tutto l'engineering che vi serve per poter generare delle visualizzazioni a sostegno dei vostri dati e generate un report su Looker Studio, prestando particolare attenzione allo storytelling: il report consegnato dovrà veicolare qualsiasi messaggio voi avrete scelto in modo chiaro e coerente.

La mia analisi si concentra maggiormente sui testi delle canzoni che hanno partecipato al Festival di Sanremo, sul lessico, la semantica, le parole più comuni e sulle tematiche maggiormente presenti.

L'analisi va a dare anche un'informazione circa gli autori più presenti alla stesura dei testi.

E continua, andando a fare una panoramica sull'influenza dei cantanti stranieri che hanno partecipato alle varie edizioni.

Un riepilogo circa le presenze di conduttori e dei vincitori, un breve sguardo anche dal punto di vista geografico, rispondendo alle domanda: quali sono le città dei vincitori e dei conduttori? quali le città più vincenti?

Concludendo con un'analisi sui generi sessuali dei partecipanti, non solo dei singoli ma anche sull'impatto dei gruppi di cantanti che hanno partecipato come duo, trio, quartetto, quintetto.

Spero che il report sia di vostro interesse. Buona visione.

Davide Buccino

INDICE

Pag. 1 - Traccia d'esame

Pag. 2 - Indice

Pag. 3 - Analisi della diversità del lessico nei testi

Pag. 4 - Analisi semantica dei testi

Pag. 5 - Frequenza parole più utilizzate

Pag. 6 - Focus sulle categorie tematiche dei testi

Pag. 7 - Partecipazione Autori alla stesura dei testi

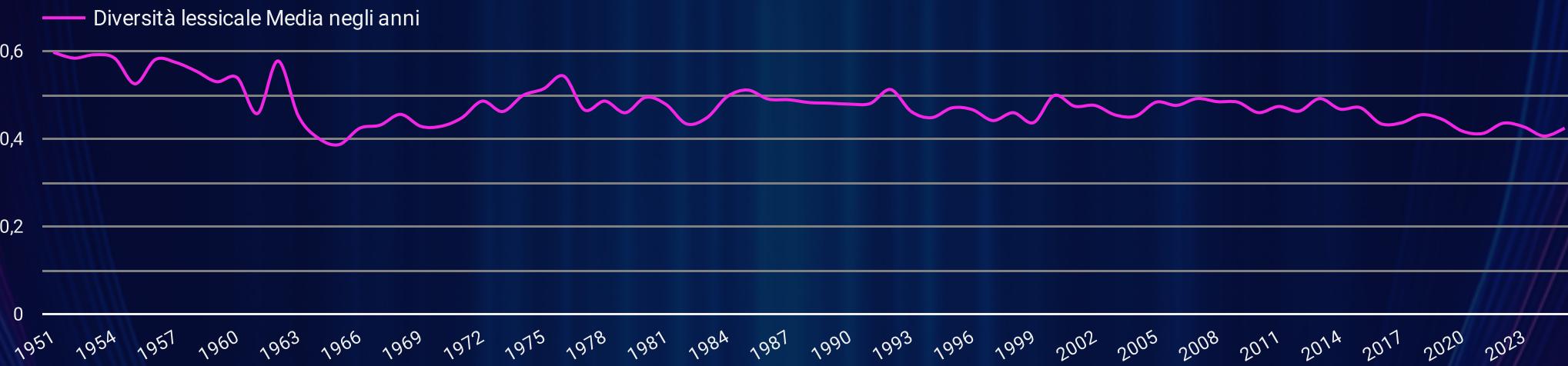
Pag. 8 - Influenza cantanti stranieri per Anno/Edizione

Pag. 9 - Analisi geografica Vincitori e Conduttori

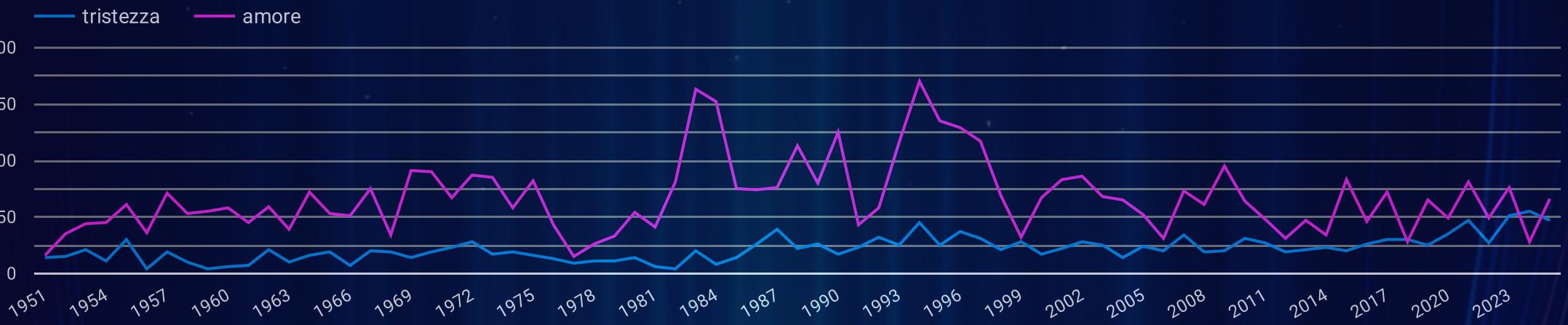
Pag. 10 -Analisi del Gender e dei gruppi

Pag. 11 - Conclusioni e Ringraziamenti

Il grafico a linee smussate mostra l'evoluzione della diversità lessicale nei testi delle canzoni del Festival di Sanremo dal 1951 al 2025, analizzando per ogni edizione il rapporto medio tra il numero totale di parole e quello delle parole uniche presenti nei brani. Nel corso del tempo si osserva una tendenza decrescente di questo rapporto: ciò suggerisce che, mentre in passato i testi utilizzavano un vocabolario più vario, nelle edizioni recenti risultano mediamente più ripetitivi, con un numero inferiore di parole uniche rispetto al totale. La tabella sottostante approfondisce ulteriormente questo fenomeno, riportando i valori esatti del rapporto di diversità lessicale per canzone in ciascuna edizione, offrendo una lettura più puntuale rispetto alla media annuale rappresentata nel grafico. È possibile utilizzare il filtro dinamico per focalizzarsi su un anno specifico e analizzarlo più nel dettaglio. I titoli delle canzoni non sono stati riportati per semplificare la lettura e concentrarsi esclusivamente sull'evoluzione lessicale nel tempo.



anno	parole_uniche	parole_totali	diversita_lessicale
1. 1960	67	82	0,82
2. 1962	59	73	0,81
3. 1976	92	114	0,81
4. 1951	73	91	0,8
5. 1962	81	103	0,79
6. 1984	107	137	0,78
7. 1985	88	113	0,78
8. 1962	83	109	0,76
9. 1957	85	112	0,76
10. 1956	82	108	0,76



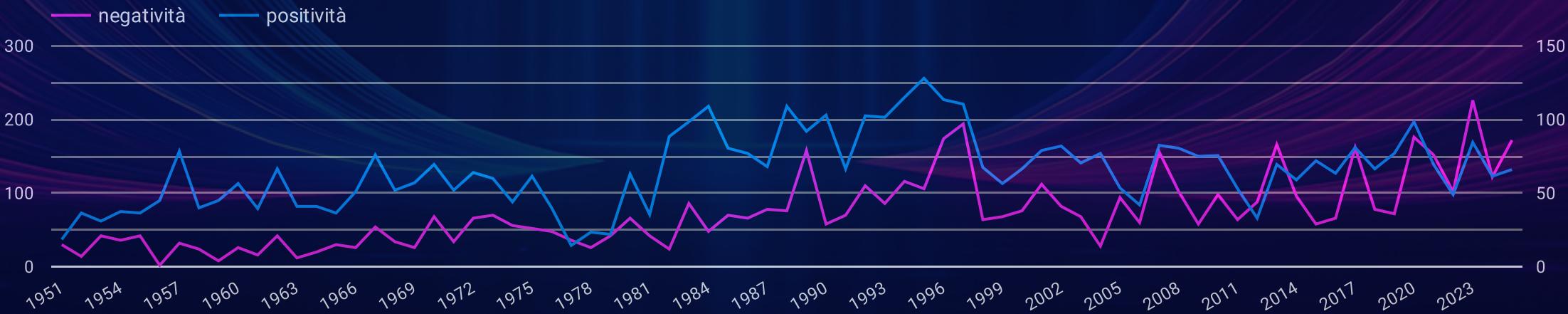
ANALISI SEMANTICA DEI TESTI

Questi due Grafici a Linee illustrano l'andamento nel tempo dell'utilizzo di determinate parole all'interno dei testi delle canzoni partecipanti al Festival di Sanremo.

Nel primo Grafico si osserva come i termini legati al tema dell'amore siano costantemente più presenti rispetto a quelli associati alla tristezza, con variazioni nei picchi di utilizzo a seconda dell'anno di riferimento.

Il secondo Grafico propone un'analisi simile, ma focalizzata sul confronto tra parole semanticamente riconducibili alla positività e alla negatività. Anche in questo caso, si osserva come i termini con connotazione "positiva" risultino, in generale, più frequenti rispetto a quelli "negativi", con un apparente cambio di tendenza nell'ultimo decennio. Questi andamenti offrono uno spunto interessante per comprendere l'evoluzione del lessico emotivo e valoriale nelle canzoni nel corso del tempo, riflettendo i cambiamenti culturali e sociali delle diverse epoche.

*Per un approfondimento circa le parole estratte dai testi ed utilizzate per i presenti grafici, si rimanda ai file di supporto e agli script di codice utilizzate e allegati al report



Parola	Periodo	Frequenza
Periodo	Parola	Frequenza
1976-2000	amore	
2001-2025	amore	
1951-1975	amore	
2001-2025	mai	
1976-2000	mai	
2001-2025	solo	
1976-2000	senza	
2001-2025	senza	
2001-2025	sempre	
2001-2025	ancora	
1976-2000	vita	
1976-2000	solo	
2001-2025	qui	

1 - 30 / 30 < >

Nello tabella di destra , la stessa analisi ma con un focus sui decenni, quindi con le parole più comuni.

Anche se simili, ci sono alcune aggiunte interessanti nei dati di alcuni decenni:

Parole nuove nella lista come: "mondo" , "voglio", "ogni" , "sole"; C'è una leggera chiusura verso l'esterno negli anni , verso il mondo, il desiderio, la totalità ("ogni", "voglio"), elementi più proattivi e meno malinconici sembrano essere caratteristica più del passato.

Oggi, invece, dominano ancora di più parole come: "mai", "solo", "senza", "niente", Parole intime, chiuse, interiori, legate alla mancanza, che descrivono un universo emotivo più ristretto, personale, spesso doloroso o introspettivo.

Possiamo concludere dicendo che Sanremo si fa più individuale e malinconico.

Questo potrebbe riflettere un cambiamento culturale, o semplicemente una fase in cui la musica sanremese privilegia emozioni chiuse, solitudini e assenze, piuttosto che slanci o speranze collettive.

Nello tabella di sinistra possiamo notare un elenco di parole, ovvero le parole presenti nei testi delle canzoni che hanno partecipato al festival di Sanremo più comuni per venticinquennio. Possiamo affermare come il tema centrale sia l'amore in tutte le sue sfumature lessico utilizzato è fortemente emotivo e soggettivo con parole più astratte come ("mai", "sempre", "senza") ; parole interiori come ("cuore", "dentro", "vita"); o parole relazionali e soggettive come ("solo", "sai", "qui"). Questo ci dice che la narrazione è quasi sempre in prima persona o in dialogo intimo ("io e te"), con forte introspezione.

Sanremo è da sempre dominato da testi che parlano d'amore, spesso in chiave malinconica, struggente o desiderante. Non solo "amore" in senso romantico, ma anche nelle sue varianti emotive: mancanza, attesa, separazione, ritorno, desiderio.

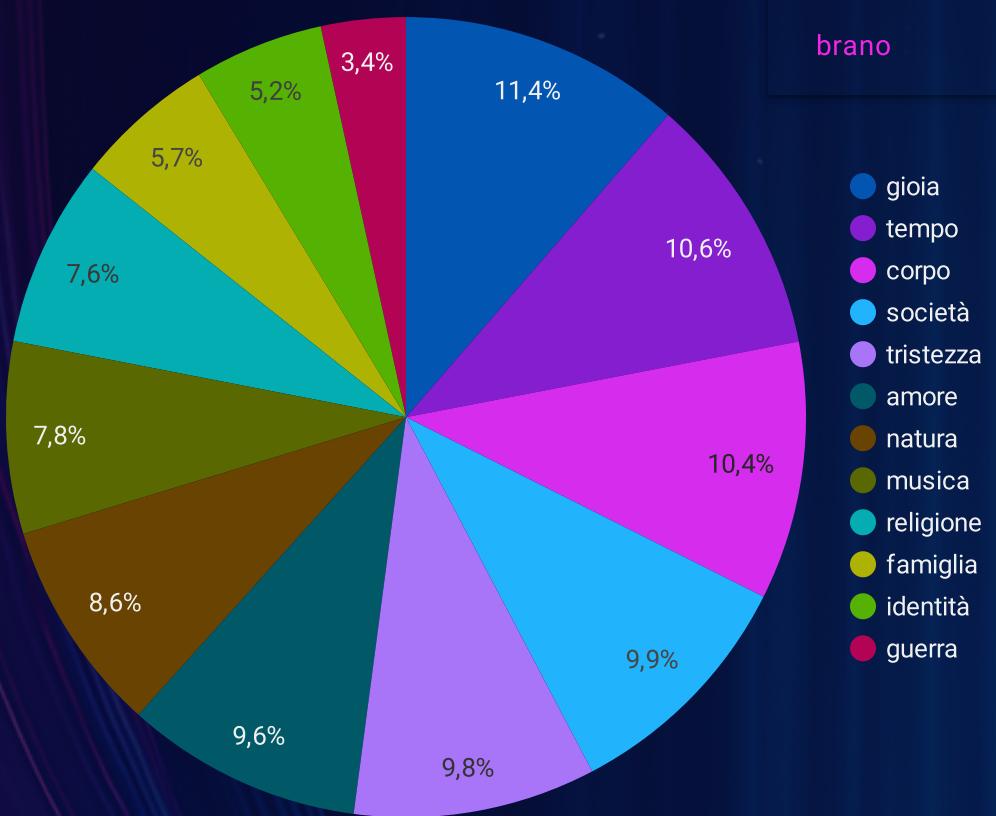
Parola	Periodo	Frequenza
Periodo	Parola	Frequenza
1981-1990	amore	
1991-2000	amore	
2001-2010	amore	
1991-2000	mai	
1971-1980	amore	
1961-1970	amore	
2011-2020	amore	
1991-2000	senza	
2011-2020	mai	
2001-2010	mai	
2011-2020	solo	
1981-1990	mai	
1991-2000	vita	
1991-2000	cuore	
1991-2000	sai	
2001-2010	dentro	

1 - 80 / 80 < >

anno

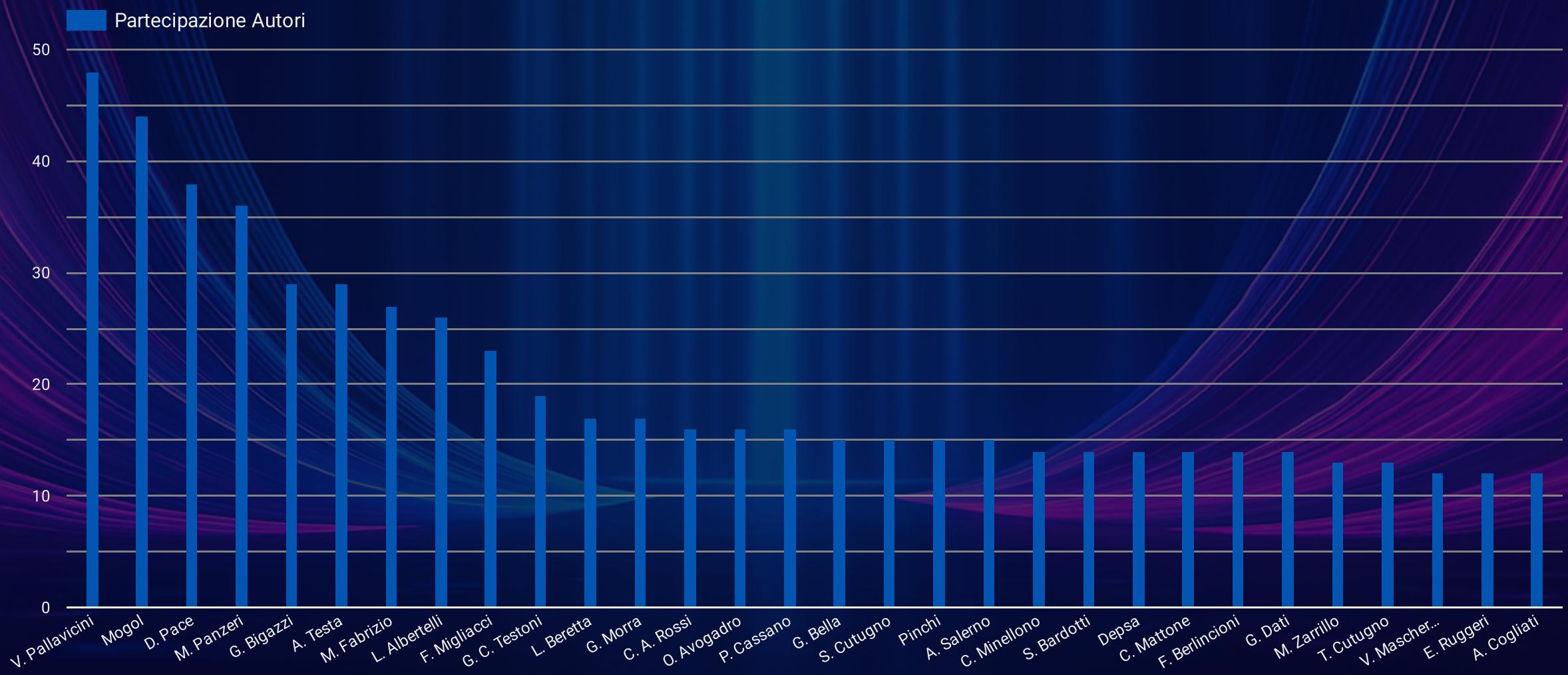
anno

brano



Per concludere il discorso iniziato nella slides precedenti, relativo all'analisi della semantica dei testi delle canzoni partecipanti al Festival di Sanremo nel corso degli anni, approfondiamo l'argomento osservando i grafici a disposizione. Le categorie semantiche analizzate sono quelle rappresentate nel grafico a torta, dove viene evidenziata la percentuale di impatto di ciascuna tematica, e nella tabella sottostante, che riporta il conteggio assoluto delle occorrenze. Le tematiche mostrate nei grafici corrispondono a insiemi di parole ricorrenti nei testi, raggruppate in base al loro significato. Questi insiemi tematici ci offrono uno spunto per comprendere come il contesto storico e culturale abbia influenzato, anno dopo anno, il lessico e i contenuti delle canzoni. È possibile esplorare ulteriormente i dati utilizzando i filtri dinamici relativi all'annata o al singolo brano, per ottenere una lettura più dettagliata e interattiva dell'evoluzione semantica nel tempo.

Nel Grafico sottostante possiamo notare il conteggio delle partecipazioni degli autori che hanno contribuito maggiormente alla stesura dei testi delle canzoni presentate negli anni al Festival di Sanremo.



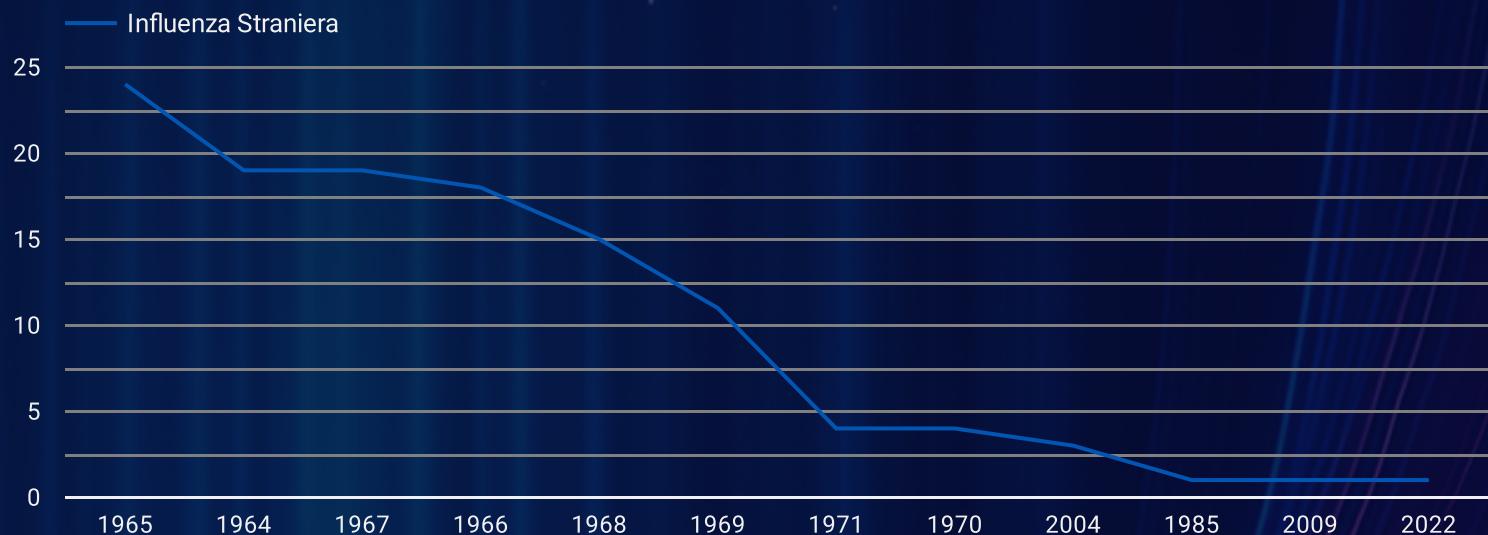
Di seguito un grafico dal quale possiamo evincere che l'influenza degli artisti stranieri a Sanremo ha vissuto un'evoluzione ben definita, legata a dinamiche artistiche, sociali e anche politiche.

Grazie ai seguenti controlli: per Anno, Influenza Straniera e Artista, è possibile entrare nel dettaglio del singolo Artista straniero e il duo, trio, quartetto di cui ha fatto parte, così da avere un elenco a discesa relativo ai nomi celebri che hanno influenzato il nostro Festival

Anno

Influenza Straniera

Artista



L'influenza degli artisti stranieri al Festival di Sanremo: un percorso a onda decrescente

Anni '60: l'apice dell'internazionalizzazione.

I picchi nel 1964, 1965, 1966, 1967 (con anche 25 artisti stranieri per edizione) raccontano un Sanremo profondamente internazionale. Questo coincide con: L'apertura della musica italiana al mondo angloamericano; L'adozione del meccanismo del doppio interprete (una versione italiana e una straniera della stessa canzone); Il desiderio di legittimazione artistica globale da parte del Festival. Sanremo in quel decennio cercava un dialogo forte con il mercato internazionale, e gli artisti stranieri erano una chiave di prestigio e visibilità.

Dal 1970 in poi: calo progressivo.

Dagli anni '70 in avanti, la presenza degli artisti stranieri declina gradualmente. Cambiano le regole, e Sanremo si orienta verso un'identità più nazionale e popolare. Diminuisce il peso della canzone italiana all'estero, e con esso anche l'interesse dei grandi nomi stranieri a partecipare.

Ritorno nei primi anni 2000: breve revival.

Intorno al 2004, si registra un nuovo picco di presenze estere. Questo periodo coincide con un tentativo del Festival di rinnovarsi e aprirsi nuovamente a linguaggi più internazionali, magari anche per competere con il crescente impatto dei talent show e dei media globali.

Ultimi anni: chiusura quasi totale.

Negli anni recenti, la presenza di artisti stranieri è sporadica e marginale (mai oltre le 5 partecipazioni). Sanremo si è ormai consolidato come evento centrato sulla scena italiana, con una forte componente nazional-popolare e un focus sul mercato locale, social e televisivo. Gli artisti stranieri, se presenti, assumono più un ruolo di ospiti o simboli internazionali che non di concorrenti in gara.

Conclusione

Il Festival di Sanremo ha attraversato fasi alterne di apertura e chiusura verso l'estero. Dall'apice cosmopolita degli anni '60 al ritorno all'identità italiana degli ultimi anni, l'andamento della presenza straniera riflette le trasformazioni del Festival e del Paese stesso, tra ambizioni globali e radicamento culturale.



A questo punto l'analisi si sposta su un piano geografico, quali sono le regioni e le rispettive città più vincenti nella storia del Festival?

Senza dubbio Centro e Nord d'Italia,
con Roma e Milano sono le vincitrici
indiscusse.

Ma anche il sud , grazie al contributo di regioni come la Campania e la Puglia tengono testa alle capolista

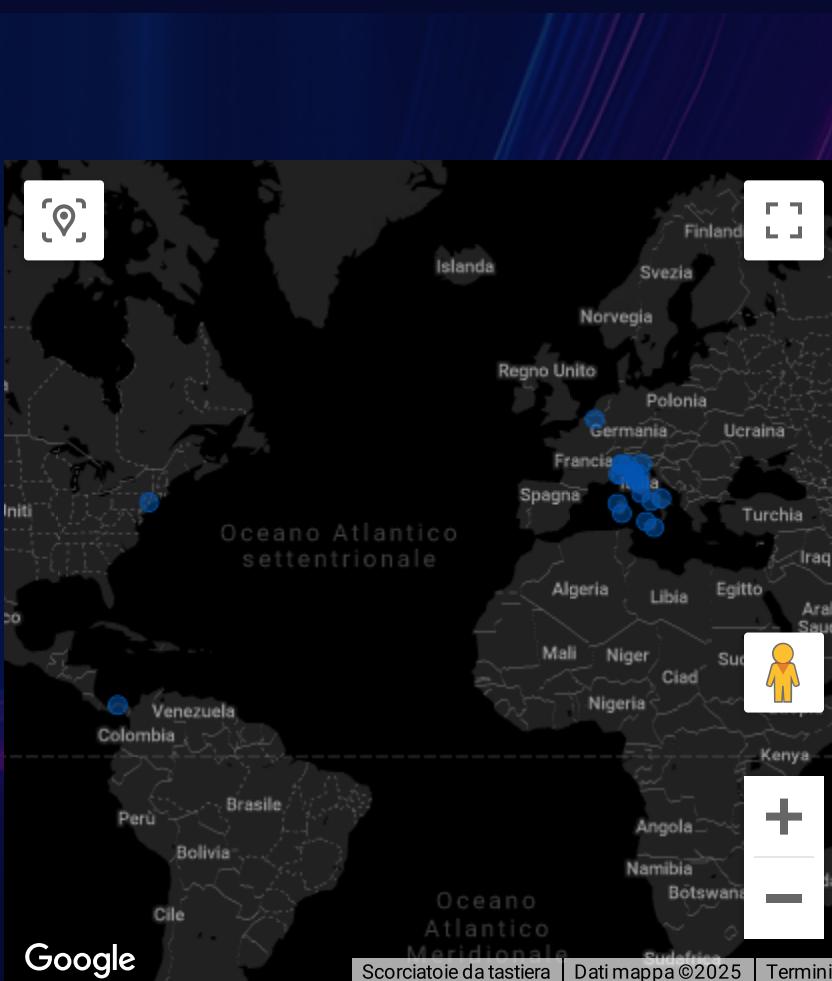
Se si ingrandisce il grafico a mappa notiamo che sono presenti anche vittorie di cantanti originari di altre zone del mondo

Puglia	Polignano a Mare	Domenico Modugno	4
Lazio	Roma	Claudio Villa	3
Emilia-Romagna	Sant'Agata Bolognese	Nilla Pizzi	2
Lazio	Roma	Bobby Solo	2
Liguria	Genova	Matia Bazar	2
Puglia	Bari	Anna Oxa	2
Lazio	Ronciglione (Viterbo)	Marco Mengoni	2
Lombardia	Milano	Mahmood	2
Campania	Capri (NA)	Peppino di Capri	2
Piemonte	Pontedera (PI)	Riccardo Fogli	1
Emilia-Romagna	Ligonchio (RE)	Iva Zanicchi	1
Toscana	Firenze	Homo Sapiens (gruppo musicale)	1
Lazio	Roma	Giorgia (cantante 1971)	1

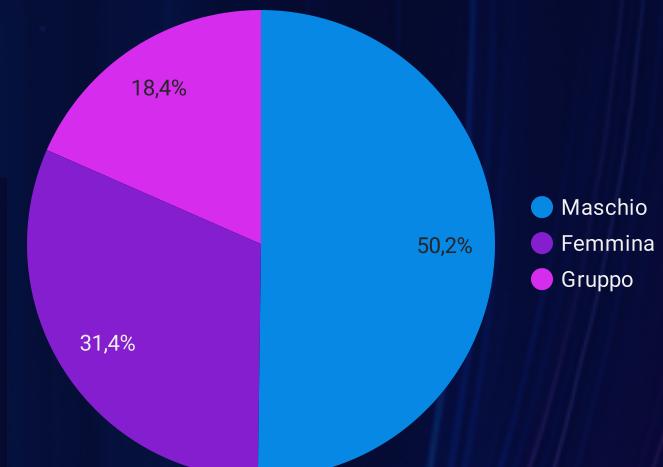
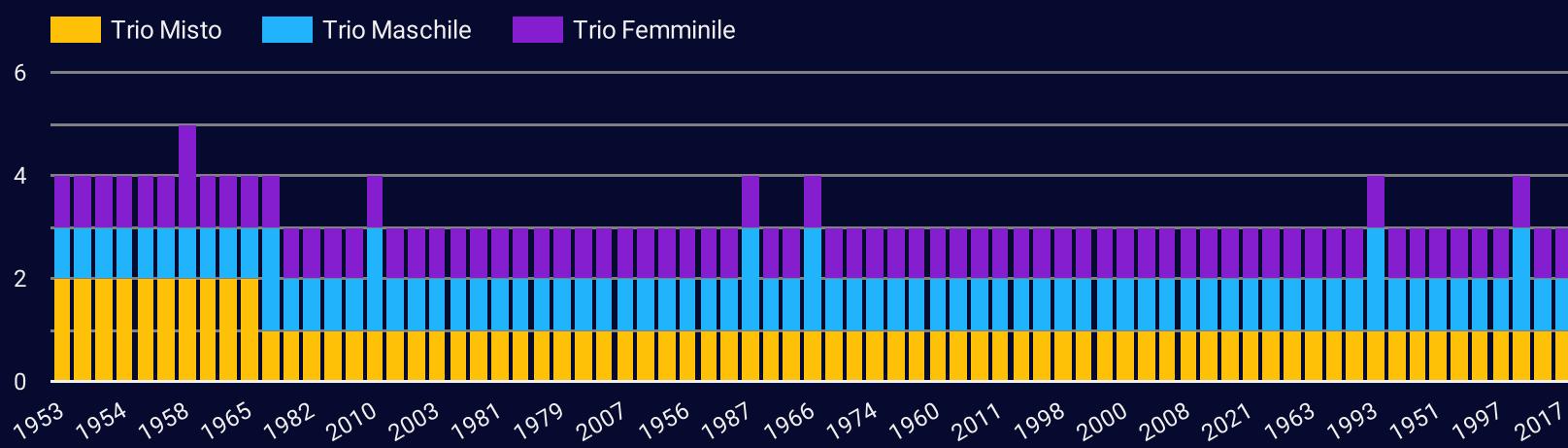
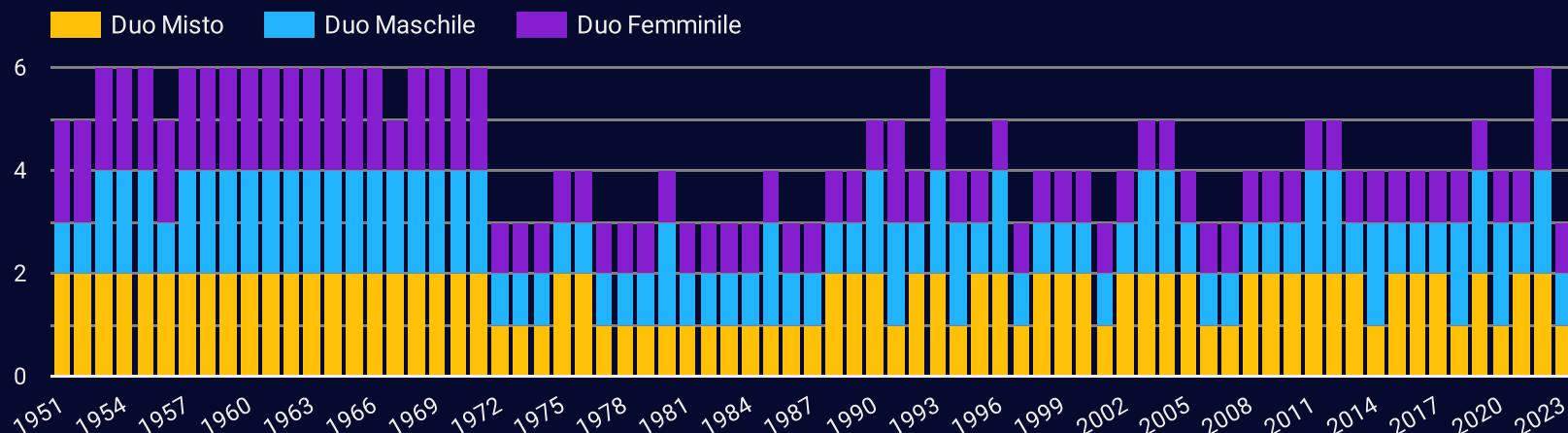
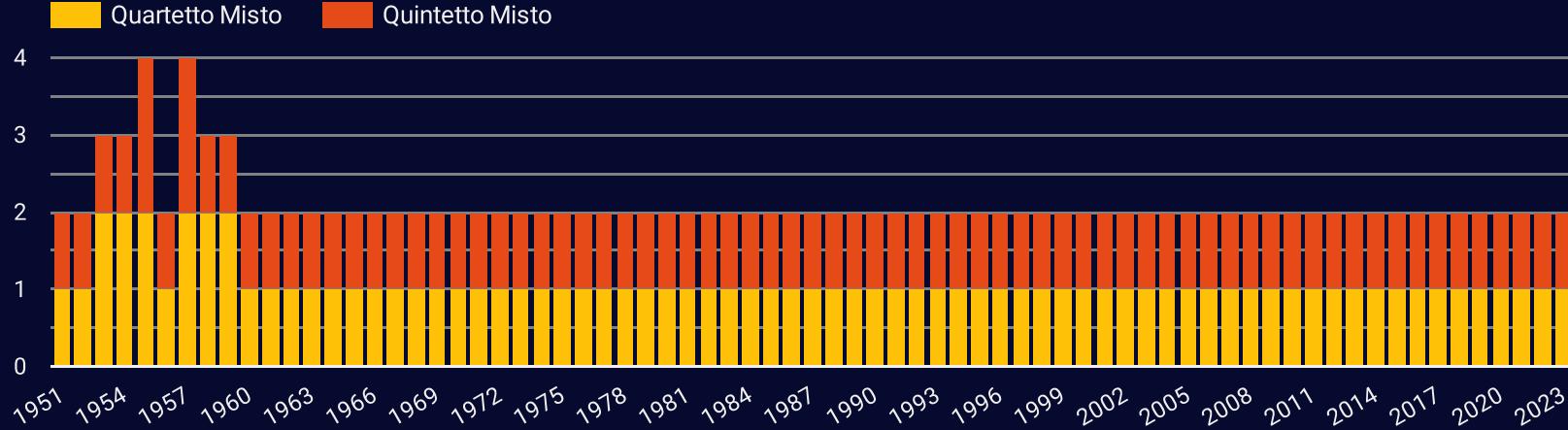
Concludiamo l'analisi geografica ma questa volta spostando l'attenzione sui conduttori del Festival

Anche in questo caso, spostandosi per il grafico a mappa possiamo notare che anche all'estero sono presenti le origini di alcuni dei nostri amati conduttori

Qui abbiamo un cambio di tendenza, i protagonisti sono principalmente del centro/sud d'Italia con una forte influenza americana data da Mike Bongiorno



Da quest'ultima analisi, risalta subito all'occhio come la distribuzione dei duetti, terzetti, quartetti e quintetti sia stata una prerogativa dei primi decenni del Festival. Dagli anni 50 agli anni 70 in particolare.



In alto una breve analisi dell'impatto dei generi sessuali sui partecipanti al Festival.

Con gli uomini maggiormente presenti rispetto alle donne

CONCLUSIONI

Partendo dall'utilizzo dei tre file Excel forniti da Epicode, ho optato per utilizzarne solo due in un primo momento, ovvero i file con le informazioni più generiche:
"dati-classifica-sanremo-1951-2023" e "dati-festival-sanremo-1951-2023".

Riguardo al file "dati-canzoni-spotify-sanremo-1951-2023" ho deciso di non utilizzarlo, nonostante avesse informazioni molto interessanti circa la durata e la popolarità dei brani.

Pertanto, il mio focus si è spostato principalmente su un discorso di analisi dei testi delle canzoni. Come ho fatto? la fonte dei miei dati sono state: il sito Le Parole di Sanremo = <https://www.leparoledisanremo.it/?s=>; il sito della SIAE (per un confronto dell'attendibilità del primo sito) ; ho estratto i testi con il supporto di Python.

Alternando l'utilizzo di strumenti quali: AI generative come ChatGPT - DeepSeeK - Copilot, Excel, Power Query, Python e siti web, ho potuto raccogliere e scremare moltissime informazioni al fine di generare dei file.

Ho cercato di sistemare i file e organizzarli in maniera da farli poi parlare con Looker Studio al meglio.

*Tutti i file e gli script Python utilizzati per questa analisi sono disponibili all'interno di una raccolta completa nel mio repository GitHub personale, per facilitare la comprensione e la consultazione del report.

Per semplificazione non allego in GitHub tutti i pdf relativi ai testi. A titolo informativo, comunico all'Utente in ascolto che con il file Python "ESTRAZIONE TESTI CON FOR.ipynb" presente nel repository, sarà automatica l'estrazione di tutti i brani semplicemente andando a modificare il range di anni nelle prime righe di codice.

LINK GITHUB: <https://github.com/DavideBuccino?tab=repositories>

In fine, un sentito ringraziamento a tutti i miei colleghi, con i quali ho avuto modo di confrontarmi, crescere professionalmente e condividere idee preziose durante tutto il percorso di lavoro su questo progetto