Sanremo

**Traccia**

Immagine che contiene testo, schermata, Carattere

Il contenuto generato dall'IA potrebbe non essere corretto.

Nel processo di Exploratory Data Analysis (EDA), ho inizialmente eseguito una pulizia approfondita dei dataset forniti, correggendo tipografie, eliminando irregolarità e gestendo valori nulli. Successivamente, ho uniformato i file per garantire coerenza nei dati.

In particolare, ho trasformato il formato della durata dei brani in minuti standardizzati, aggiornando i dati di conseguenza. È importante notare che la piattaforma Looker non supporta l'importazione diretta del formato in minuti, pertanto i dati relativi alla durata dei brani nelle tabelle devono essere mentalmente convertiti in questa unità.

In un secondo momento, ho arricchito i dataset disponibili attraverso un processo di web scraping sulla pagina Wikipedia, integrando informazioni aggiuntive. Tramite un algoritmo in Python, ho classificato gli artisti in tre categorie: solisti femminili, solisti maschili e gruppi. Ho scelto di non distinguere tra gruppi consolidati e quelli formati esclusivamente per la partecipazione al Festival.

Infine, ho caricato i dataset su Looker e ho sviluppato il report, corredando ogni grafico e tabella con una descrizione esplicativa per facilitare l'interpretazione dei dati.