1

Salve Prof siamo federico ripoli e tullio pizzuti e vi presenteremo la nostra analisi su un dataset di brani estratti da spotify

2

Inizieremo parlando dell’idea del progetto, di come abbiamo strutturato l’applicazione web, daremo informazioni sul dataset e sulle query ed infine mostreremo una demo.

3

Iniziamo a parlare del progetto

4

Abbiamo realizzato un applicativo web che permette di analizzare artisti e brani in un range di date che vanno dal 1920 al 2020. Oltre alle operazioni classiche di filtraggio dati abbiamo anche la ricerca di duplicati e comparazione di artisti e generi.

5

Parliamo dell’applicativo

6

Come dbms nosql abbiamo utilizzato mongodb che è basato sui documenti. Per la costruzione delle interfacce web abbiamo utilizzato bootstrap, oltre a diversi plugin javascript. Come webserver abbiamo scelto nodeJS costruendo servizi rest e pagine compilate EJS; e poi abbiamo utilizzato mongoose come object modeling.

7

Qui possiamo vedere un esempio di pagine dell’applicativo che grazie a bootstrap è responsive.

8

Per quanto riguarda il dataset

9

È stato reperito su kaggle e l’autore lo ha estratto dalle Spotify Web API. Si divide in due file tracce e artisti.

10

Abbiamo effettuato delle trasformazioni sui dati delle collection. Per artists abbiamo trasformato genres da stringa ad array e followers da stringa ad intero.

Per tracks abbiamo trasformato artists e id\_artist in array e release date da stringa a data.

11

Abbiamo aggiunto diversi indici di tipo asc per migliore l’esecuzione delle query basate su quei campi. Come ad esempio nome followers per artisti e nome id per le tracce.

12

Parliamo ora delle query

13

Abbiamo diversi tipi di query:

* Quelle comuni come …
* Sugli artisti
* Sulle tracce
* E tre query di lookup

14

Abbiamo qui l’esempio della common query find che prende in input il modello di mongoose e va a filtrare i dati secondo diversi criteri passati in input. I criteri vengono generati dalla funzione createfilterstring. Viene utilizzata per le tabelle e restituisce il numero totale di record e i dati. È possibile poi ordinare per un campo ed impostare skip e limit che sono usati per la paginazione. A lato abbiamo un esempio di tabella.

15

Questa funzione per gli artisti permette di ragrappare secondo un campo e un numero di scaglioni. Viene calcolato il valore minimo e massimo su quel campo, generati gli scaglioni e successivamente la query effettua il raggruppamento attraverso la funzione bucket.

16

Qui abbiamo il grafico di esempio

17

Questa funzione va a raggruppare le tracce di un artista per anno oppure per mese/anno. Vengono poi mostrate secondo un grafico chiamato Polar Area.

18

La lookup mette in relazione le collection artists e tracce attraverso i campi id e id\_artists andando a contare il numero di tracce dell’artista. Questa query permette di filtrare sul nome dell’artista, il nome del brano ed impostare il limite e l’ordinamento.

19

Qui abbiamo l’esempio del’esecuzione che mostra anche il tempo di inizio e fine dell’esecuzione.

20

Adesso mostreremo la demo dell’applicativo

21

Grazie per l’attenzione