Retrieving Wikipedia articles

9 questions

1.

Top word count words for Elton John



**(the, john, singer)**



**(england, awards, musician)**



**(the, in, and)**



**(his, the, since)**



**(rock, artists, best)**

2.

Top TF-IDF words for Elton John



**(furnish,elton,billboard)**



**(john,elton,fivedecade)**



**(the,of,has)**



**(awards,rock,john)**



**(elton,john,singer)**

3.

The cosine distance between 'Elton John's and 'Victoria Beckham's articles (represented with TF-IDF) falls within which range?



**0.1 to 0.29**



**0.3 to 0.49**



**0.5 to 0.69**



**0.7 to 0.89**



**0.9 to 1.0**

4.

The cosine distance between 'Elton John's and 'Paul McCartney's articles (represented with TF-IDF) falls within which range?



**0.1 to 0.29**



**0.3 to 0.49**



**0.5 to 0.69**



**0.7 to 0.89**



**0.9 to 1**

5.

Who is closer to 'Elton John', 'Victoria Beckham' or 'Paul McCartney'?



**Victoria Beckham**



**Paul McCartney**

6.

Who is the nearest neighbor to 'Elton John' using raw word counts?



**Billy Joel**



**Cliff Richard**



**Roger Daltrey**



**George Bush**

7.

Who is the nearest neighbor to 'Elton John' using TF-IDF?



**Roger Daltrey**



**Rod Stewart**



**Tommy Haas**



**Elvis Presley**

8.

Who is the nearest neighbor to 'Victoria Beckham' using raw word counts?



**Stephen Dow Beckham**



**Louis Molloy**



**Adrienne Corri**



**Mary Fitzgerald (artist)**

9.

Who is the nearest neighbor to 'Victoria Beckham' using TF-IDF?



**Mel B**



**Caroline Rush**



**David Beckham**



**Carrie Reichardt**