```
%% Project 2

% An exploration of themes of life, death and love across
% Grateful Dead studio albums
% Life: Life, lives, live
% Death: Death, dead, decease
% Love: Lover,


% Create table of songs and album info
song_info=readtable("GD_songs.csv")
```

song_info = 99×4 table

	Title	Filename	Album	Year
1	'100,000 Tons of Steel'	'100,000_Tons_of_Steel.txt'	'In The Dark'	1987
2	'Alabama Getaway'	'Alabama_Getaway.txt'	'Go To Heaven'	1980
3	'All New Minglewood Blues'	'All_New_Minglewood_Blues.txt'	'Shakedown Street'	1978
4	'Alligator'	'Alligator.txt'	'Anthem Of The Sun'	1968
5	'Althea'	'Althea.txt'	'Go To Heaven'	1980
6	'Attics Of My Life'	'Attics_of_My_Life.txt'	'American Beauty'	1970
7	'Beat it on Down The Line'	'Beat_it_on_Down_The_Line.txt'	'The Grateful Dead'	1967
8	'Black Muddy River'	'Black_Muddy_River.txt'	'In The Dark'	1987
9	'Black Peter'	'Black_Peter.txt'	'Workingman's Dead'	1970
10	'Blow Away'	'Blow_Away.txt'	'Built To Last'	1989
11	'Blues For Allah'	'Blues_For_Allah.txt'	'Blues For Allah'	1975
12	'Born Cross-Eyed'	'Born_Cross-Eyed.txt'	'Anthem Of The Sun'	1968
13	'Box of Rain'	'Box_of_Rain.txt'	'American Beauty'	1970
14	'Brokedown Palace'	'Brokedown_Palace.txt'	'American Beauty'	1970
15	'Built To Last'	'Built_to_Last.txt'	'Built To Last'	1989
16	'Candyman'	'Candyman.txt'	'American Beauty'	1970
17	'Casey Jones'	'Casey_Jones.txt'	'Workingman's Dead'	1970
18	'Caution'	'Caution.txt'	'Anthem Of The Sun'	1968
19	'China Cat Sunflower'	'China_Cat_Sunflower.txt'	'Aoxomoxoa'	1969
20	'China Doll'	'China_Doll.txt'	'Grateful Dead From The Mars Hotel'	1974
21	'Cold Rain and Snow'	'Cold_Rain_and_Snow.txt'	'The Grateful Dead'	1967
22	'Cosmic Charlie'	'Cosmic_Charlie.txt'	'Aoxomoxoa'	1969
23	'Crazy Fingers'	'Crazy_Fingers.txt'	'Blues For Allah'	1975
24	'Cream Puff War'	'Cream_Puff_War.txt'	'The Grateful Dead'	1967

	Title	Filename	Album	Year
25	'Cumberland Blues'	'Cumberland_Blues.txt'	'Workingman's Dead'	1970
26	'Dancing in the Streets'	'Dancing_in_the_Streets.txt'	'Terrapin Station'	1977
27	'Dire Wolf'	'Dire_Wolf.txt'	'Workingman's Dead'	1970
28	'Doin' That Rag'	'Doin'_That_Rag.txt'	'Aoxomoxoa'	1969
29	'Don't Ease Me In'	'Don't_Ease_Me_In.txt'	'Go To Heaven'	1980
30	'Dupree's Diamond Blues'	'Dupree's_Diamond_Blues.txt'	'Aoxomoxoa'	1969
31	'Easy To Love You'	'Easy_To_Love_You.txt'	'Go To Heaven'	1980
32	'Easy Wind'	'Easy_Wind.txt'	'Workingman's Dead'	1970
33	'Estimated Prophet'	'Estimated_Prophet.txt'	'Terrapin Station'	1977
34	'Eyes Of The World'	'Eyes_Of_The_World.txt'	'Wake Of The Flood'	1973
35	'Far From Me'	'Far_From_Me.txt'	'Go To Heaven'	1980
36	'Feel Like A Stranger'	'Feel_Like_A_Stranger.txt'	'Go To Heaven'	1980
37	'Fire On The Mountain'	'Fire_On_The_Mountain.txt'	'Shakedown Street'	1978
38	'Foolish Heart'	'Foolish_Heart.txt'	'Built To Last'	1989
39	'France'	'France.txt'	'Shakedown Street'	1978
40	'Franklin's Tower'	'Franklin's_Tower.txt'	'Blues For Allah'	1975
41	'Friend of the Devil'	'Friend_of_the_Devil.txt'	'American Beauty'	1970
42	'From the Heart of Me'	'From_the_Heart_of_Me.txt'	'Shakedown Street'	1978
43	'Good Lovin''	'Good_Lovin'.txt'	'Shakedown Street'	1978
44	'Good Morning Little Schoolgirl'	'Good_Morning_Little_Schoolgirl.txt'	'The Grateful Dead'	1967
45	'Hell in a Bucket'	'Hell_in_a_Bucket.txt'	'In The Dark'	1987
46	'Help on the Way'	'Help_on_the_Way.txt'	'Blues For Allah'	1975
47	'Here Comes Sunshine'	'Here_Comes_Sunshine.txt'	'Wake Of The Flood'	1973
48	'High Time'	'High_Time.txt'	'Workingman's Dead'	1970
49	'I Need A Miracle'	'I_Need_A_Miracle.txt'	'Shakedown Street'	1978
50	'I Will Take You Home'	'I_Will_Take_You_Home.txt'	'Built To Last'	1989
51	'If I Had The World To Give'	'If_I_Had_The_World_To_Give.txt'	'Shakedown Street'	1978
52	'Just A Little Light'	'Just_a_Little_Light.txt'	'Built To Last'	1989
53	'Let Me Sing Your Blues Away'	'Let_Me_Sing_Your_Blues_Away.txt'	'Wake Of The Flood'	1973
54	'Loose Lucy'	'Loose_Lucy.txt'	'Grateful Dead From The Mars Hotel'	1974
55	'Lost Sailor'	'Lost_Sailor.txt'	'Go To Heaven'	1980
56	'Mississippi Halfstep Uptown Toodeloo'	'Mississippi_Halfstep_Uptown_Toodeloo.txt'	'Wake Of The Flood'	1973

	Title	Filename	Album	Year
57	'Money, Money'	'Money,_Money.txt'	'Grateful Dead From The Mars Hotel'	1974
58	'Morning Dew'	'Morning_Dew.txt'	'The Grateful Dead'	1967
59	'Mountains of the Moon'	'Mountains_of_the_Moon.txt'	'Aoxomoxoa'	1969
60	'New Minglewood Blues'	'New_Minglewood_Blues.txt'	'The Grateful Dead'	1967
61	'New Potato Caboose'	'New_Potato_Caboose.txt'	'Anthem Of The Sun'	1968
62	'New Speedway Boogie'	'New_Speedway_Boogie.txt'	'Workingman's Dead'	1970
63	'Operator'	'Operator.txt'	'American Beauty'	1970
64	'Passenger'	'Passenger.txt'	'Terrapin Station'	1977
65	'Picasso Moon'	'Picasso_Moon.txt'	'Built To Last'	1989
66	'Pride of Cucamonga'	'Pride_of_Cucamonga.txt'	'Grateful Dead From The Mars Hotel'	1974
67	'Ripple'	'Ripple.txt'	'American Beauty'	1970
68	'Rosemary'	'Rosemary.txt'	'Aoxomoxoa'	1969
69	'Row Jimmy'	'Row_Jimmy.txt'	'Wake Of The Flood'	1973
70	'Saint of Circumstance'	'Saint_of_Circumstance.txt'	'Go To Heaven'	1980
71	'Saint Stephen'	'Saint_Stephen.txt'	'Aoxomoxoa'	1969
72	'Samson and Delilah'	'Samson_and_Delilah.txt'	'Terrapin Station'	1977
73	'Scarlet Begonias'	'Scarlet_Begonias.txt'	'Grateful Dead From The Mars Hotel'	1974
74	'Shakedown Street'	'Shakedown_Street.txt'	'Shakedown Street'	1978
75	'Ship of Fools'	'Ship_of_Fools.txt'	'Grateful Dead From The Mars Hotel'	1974
76	'Sitting on Top of the World'	'Sitting_on_Top_of_the_World.txt'	'The Grateful Dead'	1967
77	'Stagger Lee'	'Stagger_Lee.txt'	'Shakedown Street'	1978
78	'Standing On The Moon'	'Standing_on_the_Moon.txt'	'Built To Last'	1989
79	'Stella Blue'	'Stella_Blue.txt'	'Wake Of The Flood'	1973
80	'Sugar Magnolia'	'Sugar_Magnolia.txt'	'American Beauty'	1970
81	'Sunrise'	'Sunrise.txt'	'Terrapin Station'	1977
82	'Terrapin Station - Lady With a Fan'	'Terrapin_Station_-_Lady_With_a_Fan.txt'	'Terrapin Station'	1977
83	'That's It For The Other One'	'That's_It_For_The_Other_One.txt'	'Anthem Of The Sun'	1968
84	'The Golden Road (to unlimited devotion)'	'The_Golden_Road_(to_unlimited_devotion).txt'	'The Grateful Dead'	1967
85	'The Music Never Stopped'	'The_Music_Never_Stopped.txt'	'Blues For Allah'	1975
86	'Throwing Stones'	'Throwing_Stones.txt'	'In The Dark'	1987

	Title	Filename	Album	Year
87	'Till The Morning Comes'	'Till_The_Morning_Comes.txt'	'American Beauty'	1970
88	'Touch Of Grey'	'Touch_of_Grey.txt'	'In The Dark'	1987
89	'Truckin''	'Truckin'.txt'	'American Beauty'	1970
90	'U.S. Blues'	'U.S._Blues.txt'	'Grateful Dead From The Mars Hotel'	1974
91	'Unbroken Chain'	'Unbroken_Chain.txt'	'Grateful Dead From The Mars Hotel'	1974
92	'Uncle John's Band'	'Uncle_John's_Band.txt'	'Workingman's Dead'	1970
93	'Victim Or The Crime'	'Victim_or_the_Crime.txt'	'Built To Last'	1989
94	'Viola Lee Blues'	'Viola_Lee_Blues.txt'	'The Grateful Dead'	1967
95	'Weather Report Suite (Part I)'	'Weather_Report_Suite_(Part_I).txt'	'Wake Of The Flood'	1973
96	'Weather Report Suite (Part II) - Let It Grow'	'Weather_Report_Suite_(Part_II)_-_Let_It_Grow.txt'	'Wake Of The Flood'	1973
97	'West L.A. Fadeaway'	'West_L.A._Fadeaway.txt'	'In The Dark'	1987
98	'What's Become of the Baby'	'What's_Become_of_the_Baby.txt'	'Aoxomoxoa'	1969
99	'When Push Comes to Shove'	'When_Push_Comes_to_Shove.txt'	'In The Dark'	1987

```matlab
allwords=strings(0,1); % Initialize
allLyrics={1,height(song_info)}; % Initialize

for i=1:height(song_info)

    filepath=fullfile(strcat("lyrics\",song_info{i,"Filename"})); % filepath
to .txt file of lyrics for song i

    allLyrics{i} = fileToArray(filepath); % Add lyrics of song i to cell array of
lyrics
    allwords=[allwords;allLyrics{i}]; % Add lyrics of song i to cumulative vector
of all words for all songs
end

uniqueWords=unique(allwords); % vector of unique words accross all songs

% Create Matrix A for LSM
for i=1:length(uniqueWords)
    word=uniqueWords(i);
    for j=1:height(song_info)
        song=allLyrics{j};
        A(i,j)=nnz(strcmp(song,word)); % number of times word(i) appears in
song(j)
    end
end
```

```matlab
% Create table of word frequency per song from matrix A
words_Freq=array2table(A,"RowNames",uniqueWords,"VariableNames",song_info{:,"Title"}
);
words_Freq.sum=sum(A,2);
words_Freq=sortrows(words_Freq,"sum",'descend');


%%Query1 (life)
life=["live","lived","livin","living","life","lifetime","alive"];
cosDist_life = semanticAnalysis(A,life,uniqueWords)';
song_info.cosDist_life=cosDist_life; % Add results to table

%%Query2 (death)
death=["dead","drowned","murder","die","fatal","drown","dying","kill","killer","kill
ing"];
cosDist_death = semanticAnalysis(A,death,uniqueWords)';
song_info.cosDist_death=cosDist_death; % Add results to table

%%Query3 (love)
love=["love","loved","lover","lovers","love","lovin","loving","baby","heart","romanc
e","romancin","romantics"];
cosDist_love = semanticAnalysis(A,love,uniqueWords)';
song_info.cosDist_love=cosDist_love; % Add results to table

%%Query4 (Gambling)
gambling=["money","played","jack","lie","lies","whiskey","win","wine","fair","jail",
"shakedown","shots","fight","drink","drinkin","steal","gin","jailhouse","liquor","lo
sing","pistol"];
cosDist_gambling = semanticAnalysis(A,gambling,uniqueWords)';
song_info.cosDist_gambling=cosDist_gambling; % Add results to table

%%Query5 (Darkness)
darkness=["moon","night","dark","black","cry","broken","darkness","devil","lonely","
carrion","dungeon","forsaken","nightmare","nights","nighttime","poisoned","tears"];
cosDist_darkness = semanticAnalysis(A,darkness,uniqueWords)';
song_info.cosDist_darkness=cosDist_darkness; % Add results to table

%%Query6 (Light)
light=["mornin","morning","sun","light","dancing","dance","happy","sunshine","summer
","dawn","miracle","laughter","dove","fun"];
cosDist_light = semanticAnalysis(A,light,uniqueWords)';
song_info.cosDist_light=cosDist_light; % Add results to table

% Display table of Cos Dist for each song and theme
song_info(:,
["Title","Album","Year","cosDist_life","cosDist_death","cosDist_love","cosDist_gambl
ing","cosDist_darkness","cosDist_light"])
```

ans = 99×9 table

	Title	Album	Year	cosDist_life	cosDist_death
1	'100,000 Tons of Steel'	'In The Dark'	1987	0.9939	1.0000
2	'Alabama Getaway'	'Go To Heaven'	1980	0.9953	0.9773
3	'All New Minglewood Blues'	'Shakedown Street'	1978	0.9325	0.8842
4	'Alligator'	'Anthem Of The Sun'	1968	0.9977	0.9988
5	'Althea'	'Go To Heaven'	1980	0.9990	0.9871
6	'Attics Of My Life'	'American Beauty'	1970	0.9995	0.9965
7	'Beat it on Down The Line'	'The Grateful Dead'	1967	0.9948	0.9999
8	'Black Muddy River'	'In The Dark'	1987	0.9917	0.9999
9	'Black Peter'	'Workingman's Dead'	1970	0.9992	0.9971
10	'Blow Away'	'Built To Last'	1989	0.9947	0.9760
11	'Blues For Allah'	'Blues For Allah'	1975	0.9807	0.9513
12	'Born Cross-Eyed'	'Anthem Of The Sun'	1968	0.9996	0.9963
13	'Box of Rain'	'American Beauty'	1970	0.9956	0.9997
14	'Brokedown Palace'	'American Beauty'	1970	0.9934	1.0000
15	'Built To Last'	'Built To Last'	1989	0.9992	0.9879
16	'Candyman'	'American Beauty'	1970	0.9855	0.9591
17	'Casey Jones'	'Workingman's Dead'	1970	1.0000	0.9939
18	'Caution'	'Anthem Of The Sun'	1968	0.9713	0.9370
19	'China Cat Sunflower'	'Aoxomoxoa'	1969	0.9855	0.9985
20	'China Doll'	'Grateful Dead From The Mars Hotel'	1974	0.9999	0.9919
21	'Cold Rain and Snow'	'The Grateful Dead'	1967	0.9999	0.9951
22	'Cosmic Charlie'	'Aoxomoxoa'	1969	0.9996	0.9960
23	'Crazy Fingers'	'Blues For Allah'	1975	0.9957	0.9997
24	'Cream Puff War'	'The Grateful Dead'	1967	0.9884	0.9641
25	'Cumberland Blues'	'Workingman's Dead'	1970	0.9999	0.9917
26	'Dancing in the Streets'	'Terrapin Station'	1977	0.9998	0.9907
27	'Dire Wolf'	'Workingman's Dead'	1970	0.9998	0.9953
28	'Doin' That Rag'	'Aoxomoxoa'	1969	0.9972	0.9818
29	'Don't Ease Me In'	'Go To Heaven'	1980	1.0000	0.9942
30	'Dupree's Diamond Blues'	'Aoxomoxoa'	1969	0.9970	0.9812
31	'Easy To Love You'	'Go To Heaven'	1980	0.9884	0.9640
32	'Easy Wind'	'Workingman's Dead'	1970	0.9998	0.9905

	Title	Album	Year	cosDist_life	cosDist_death
33	'Estimated Prophet'	'Terrapin Station'	1977	0.9907	0.9998
34	'Eyes Of The World'	'Wake Of The Flood'	1973	0.9980	0.9986
35	'Far From Me'	'Go To Heaven'	1980	0.9943	0.9752
36	'Feel Like A Stranger'	'Go To Heaven'	1980	0.9989	0.9975
37	'Fire On The Mountain'	'Shakedown Street'	1978	1.0000	0.9926
38	'Foolish Heart'	'Built To Last'	1989	0.9945	0.9999
39	'France'	'Shakedown Street'	1978	0.9891	0.9652
40	'Franklin's Tower'	'Blues For Allah'	1975	0.9968	0.9994
41	'Friend of the Devil'	'American Beauty'	1970	0.9983	0.9984
42	'From the Heart of Me'	'Shakedown Street'	1978	0.9924	0.9715
43	'Good Lovin''	'Shakedown Street'	1978	0.2244	0.1095
44	'Good Morning Little Schoolgirl'	'The Grateful Dead'	1967	0.9392	0.8929
45	'Hell in a Bucket'	'In The Dark'	1987	0.9983	0.9983
46	'Help on the Way'	'Blues For Allah'	1975	0.9992	0.9971
47	'Here Comes Sunshine'	'Wake Of The Flood'	1973	0.9995	0.9893
48	'High Time'	'Workingman's Dead'	1970	0.9484	0.9051
49	'I Need A Miracle'	'Shakedown Street'	1978	0.9547	0.9136
50	'I Will Take You Home'	'Built To Last'	1989	0.9925	1.0000
51	'If I Had The World To Give'	'Shakedown Street'	1978	0.9517	0.9095
52	'Just A Little Light'	'Built To Last'	1989	1.0000	0.9936
53	'Let Me Sing Your Blues Away'	'Wake Of The Flood'	1973	0.9982	0.9846
54	'Loose Lucy'	'Grateful Dead From The Mars Hotel'	1974	0.9998	0.9906
55	'Lost Sailor'	'Go To Heaven'	1980	0.9845	0.9574
56	'Mississippi Halfstep Uptown Toodeloo'	'Wake Of The Flood'	1973	0.9999	0.9947
57	'Money, Money'	'Grateful Dead From The Mars Hotel'	1974	1.0000	0.9940
58	'Morning Dew'	'The Grateful Dead'	1967	1.0000	0.9923
59	'Mountains of the Moon'	'Aoxomoxoa'	1969	0.9995	0.9889
60	'New Minglewood Blues'	'The Grateful Dead'	1967	0.9923	0.9711
61	'New Potato Caboose'	'Anthem Of The Sun'	1968	1.0000	0.9923
62	'New Speedway Boogie'	'Workingman's Dead'	1970	0.9968	0.9807
63	'Operator'	'American Beauty'	1970	0.9988	0.9977
64	'Passenger'	'Terrapin Station'	1977	0.9998	0.9904

	Title	Album	Year	cosDist_life	cosDist_death
65	'Picasso Moon'	'Built To Last'	1989	0.9993	0.9968
66	'Pride of Cucamonga'	'Grateful Dead From The Mars Hotel'	1974	0.9981	0.9842
67	'Ripple'	'American Beauty'	1970	0.9819	0.9973
68	'Rosemary'	'Aoxomoxoa'	1969	0.9946	0.9999
69	'Row Jimmy'	'Wake Of The Flood'	1973	0.9863	0.9988
70	'Saint of Circumstance'	'Go To Heaven'	1980	0.9997	0.9958
71	'Saint Stephen'	'Aoxomoxoa'	1969	0.9902	0.9672
72	'Samson and Delilah'	'Terrapin Station'	1977	0.9998	0.9909
73	'Scarlet Begonias'	'Grateful Dead From The Mars Hotel'	1974	0.9984	0.9982
74	'Shakedown Street'	'Shakedown Street'	1978	0.9943	0.9752
75	'Ship of Fools'	'Grateful Dead From The Mars Hotel'	1974	1.0000	0.9921
76	'Sitting on Top of the World'	'The Grateful Dead'	1967	0.9797	0.9498
77	'Stagger Lee'	'Shakedown Street'	1978	0.9991	0.9874
78	'Standing On The Moon'	'Built To Last'	1989	0.9968	0.9993
79	'Stella Blue'	'Wake Of The Flood'	1973	0.9948	0.9999
80	'Sugar Magnolia'	'American Beauty'	1970	0.9917	0.9700
81	'Sunrise'	'Terrapin Station'	1977	0.9921	1.0000
82	'Terrapin Station - Lady With a Fan'	'Terrapin Station'	1977	0.9984	0.9982
83	'That's It For The Other One'	'Anthem Of The Sun'	1968	0.9989	0.9975
84	'The Golden Road (to unlimited devotion)'	'The Grateful Dead'	1967	0.9925	0.9716
85	'The Music Never Stopped'	'Blues For Allah'	1975	0.9975	0.9826
86	'Throwing Stones'	'In The Dark'	1987	0.9993	0.9882
87	'Till The Morning Comes'	'American Beauty'	1970	0.9998	0.9954
88	'Touch Of Grey'	'In The Dark'	1987	0.9998	0.9953
89	'Truckin''	'American Beauty'	1970	0.9978	0.9832
90	'U.S. Blues'	'Grateful Dead From The Mars Hotel'	1974	0.9974	0.9990
91	'Unbroken Chain'	'Grateful Dead From The Mars Hotel'	1974	0.9918	0.9999
92	'Uncle John's Band'	'Workingman's Dead'	1970	0.9994	0.9888
93	'Victim Or The Crime'	'Built To Last'	1989	0.9998	0.9910

	Title	Album	Year	cosDist_life	cosDist_death
94	'Viola Lee Blues'	'The Grateful Dead'	1967	0.8566	0.7909
95	'Weather Report Suite (Part I)'	'Wake Of The Flood'	1973	0.9984	0.9850
96	'Weather Report Suite (Part II) - Let It Grow'	'Wake Of The Flood'	1973	0.9834	0.9978
97	'West L.A. Fadeaway'	'In The Dark'	1987	0.9836	0.9559
98	'What's Become of the Baby'	'Aoxomoxoa'	1969	0.9982	0.9844
99	'When Push Comes to Shove'	'In The Dark'	1987	0.9980	0.9986

```matlab
% Calculate sum of cos Dist for each theme per album
% Rank albums per theme

% Normalize by # of songs in album?

albums=unique(song_info(:,"Album"));
albums.cd_life(:)=0;
albums.cd_death(:)=0;
albums.cd_love(:)=0;
albums.cd_gambling(:)=0;
albums.cd_darkness(:)=0;
albums.cd_light(:)=0;

for a = 1:height(albums)
    album_name=albums{a,"Album"}{:};

    songs_album_a=song_info(string(song_info.Album) == album_name,:);

    albums{a,"cd_life"}=sum(songs_album_a.cosDist_life);
    albums{a,"cd_death"}=sum(songs_album_a.cosDist_death);
    albums{a,"cd_love"}=sum(songs_album_a.cosDist_love);
    albums{a,"cd_gambling"}=sum(songs_album_a.cosDist_gambling);
    albums{a,"cd_darkness"}=sum(songs_album_a.cosDist_darkness);
    albums{a,"cd_light"}=sum(songs_album_a.cosDist_light);

end

albums
```

albums = 13×7 table

...

	Album	cd_life	cd_death	cd_love	cd_gambling
1	'American Beauty'	9.9422	9.8972	4.9606	9.9335
2	'Anthem Of The Sun'	4.9675	4.9218	2.6488	4.9501
3	'Aoxomoxoa'	7.9618	7.8980	4.1763	7.9392

	Album	cd_life	cd_death	cd_love	cd_gambling
4	'Blues For Allah'	4.9699	4.9300	2.6077	4.9558
5	'Built To Last'	7.9769	7.9446	3.9513	7.9721
6	'Go To Heaven'	7.9601	7.8485	4.5283	7.9106
7	'Grateful Dead From The Mars Hotel'	7.9854	7.9500	3.9781	7.9788
8	'In The Dark'	6.9645	6.9364	3.4500	6.9603
9	'Shakedown Street'	8.0382	7.7087	6.1700	7.8661
10	'Terrapin Station'	5.9805	5.9701	2.8611	5.9847
11	'The Grateful Dead'	8.7434	8.5277	5.6605	8.6365
12	'Wake Of The Flood'	7.9587	7.9486	3.7792	7.9663
13	'Workingman's Dead'	7.9433	7.8430	4.4375	7.9001

```matlab
% Normalize?
% Change / optimize rank selection in LSA
% Add ranking of albums instead of just of cos dist?
% Reduce themes?

% Conclusions
%- Songs that have strong themes of life also have strong themes of death
% - The Grateful dead did not write many love songs
% - Shakedown Street is an outlier
% - Some albums, American Beauty, contain deeper and more diverse themes
% than others, Workingman's dead.




%---------------------- Functions -------------------------

function words = fileToArray(filename)
%% Returns cleaned array of words in song specified by "string"

fid=fopen(filename); % open file
words=textscan(fid, '%s'); % read words into cell array
fclose(fid); % close file
words=words{:}; % create string array from cell array
words=lower(words); % convert all characters to lowercase
words=erasePunctuation(words); % remove punctuation characters from strings
words=rmmissing(words); % remove empty elements

ignore=[
  "the";
  "i";
  "you";
```

```matlab
    "and";
    "to";
    "a";
    "of";
    "in";
    "on";
    "it";
    "that";
    "for";
    "is";
    "its";
    "this";
    "im";
    "as"
];


for i=1:length(ignore) % remove words specified in ignore array
    words(strcmp(words,ignore(i)))=[];
end

end

function cosDist = semanticAnalysis(A,query,words)

[m,n]=size(A); % number of words, number of songs

rank=2; % number of singular values to keep in reduced system

[U,S,V]=svd(A);

U1=U(:,[1,rank]);
S1=S([1,rank],[1,rank]);
V1=V(:,[1,rank]);

for i=1:m
    wordsProj(i,:)=U1(i,:)*S1;
end

for i=1:n
    docsProj(:,i)=S1*(V1(i,:))';
end


% Query

q_num=zeros(1,rank);

for i=1:length(query)
    index=find(strcmp(words,query(i)));
```

```
        q_num=q_num+wordsProj(index,:);
    end

    q=q_num/length(query);

    for i=1:n
        cosDist(i) = (dot(q',docsProj(:,i)))/(norm(q)*norm(docsProj(:,i)));
    end

end
```