

## HW 2 Part 2

### Q1.1 part a and b)

```
[nltk_data] Downloading package punkt to /Users/Arman/nltk_data...
```

```
[nltk_data] Unzipping tokenizers/punkt.zip.
```

Token counts per document:

Jerry Decided To Buy a Gun.txt: 305 tokens

Rentals at the Oceanside Community.txt: 380 tokens

Gasoline Prices Hit Record High.txt: 303 tokens

Cloning Pets.txt: 266 tokens

Crazy Housing Prices.txt: 403 tokens

Man Injured at Fast Food Place.txt: 173 tokens

A Festival of Books.txt: 305 tokens

Food Fight Erupted in Prison.txt: 222 tokens

Better To Be Unlucky.txt: 359 tokens

Sara Went Shopping.txt: 165 tokens

Freeway Chase Ends at Newsstand.txt: 335 tokens

Trees Are a Threat.txt: 344 tokens

A Murder-Suicide.txt: 407 tokens

Happy and Unhappy Renters.txt: 317 tokens

Pulling Out Nine Tons of Trash.txt: 298 tokens

Total number of tokens in the entire collection: 4582

Number of unique tokens in the entire collection: 1427

### Q1.1 a and b)

Token counts per document (no stop words):

Jerry Decided To Buy a Gun.txt: 176 tokens

Rentals at the Oceanside Community.txt: 224 tokens

Gasoline Prices Hit Record High.txt: 176 tokens

Cloning Pets.txt: 158 tokens

Crazy Housing Prices.txt: 226 tokens

Man Injured at Fast Food Place.txt: 107 tokens

A Festival of Books.txt: 198 tokens

Food Fight Erupted in Prison.txt: 140 tokens

Better To Be Unlucky.txt: 218 tokens

Sara Went Shopping.txt: 112 tokens

Freeway Chase Ends at Newsstand.txt: 186 tokens

Trees Are a Threat.txt: 196 tokens

A Murder-Suicide.txt: 221 tokens

Happy and Unhappy Renters.txt: 185 tokens

Pulling Out Nine Tons of Trash.txt: 201 tokens

Total number of tokens in the entire collection (no stop words): 2724

Number of unique tokens in the entire collection (no stop words): 1265

### Q1.3)

	110	...	world	worms	\
Jerry Decided To Buy a Gun.txt	0.000000	...	0.000000	0.000000	
Rentals at the Oceanside Community.txt	0.000000	...	0.000000	0.061586	
Gasoline Prices Hit Record High.txt	0.000000	...	0.000000	0.000000	
Cloning Pets.txt	0.000000	...	0.000000	0.000000	
Crazy Housing Prices.txt	0.000000	...	0.000000	0.000000	
Man Injured at Fast Food Place.txt	0.000000	...	0.000000	0.000000	
A Festival of Books.txt	0.000000	...	0.000000	0.000000	
Food Fight Erupted in Prison.txt	0.000000	...	0.000000	0.000000	
Better To Be Unlucky.txt	0.056063	...	0.000000	0.000000	
Sara Went Shopping.txt	0.000000	...	0.000000	0.000000	
Freeway Chase Ends at Newsstand.txt	0.000000	...	0.000000	0.000000	
Trees Are a Threat.txt	0.000000	...	0.000000	0.000000	
A Murder-Suicide.txt	0.000000	...	0.066095	0.000000	
Happy and Unhappy Renters.txt	0.000000	...	0.065236	0.000000	
Pulling Out Nine Tons of Trash.txt	0.000000	...	0.000000	0.000000	

	worth	wrong	year	\
Jerry Decided To Buy a Gun.txt	0.000000	0.000000	0.000000	
Rentals at the Oceanside Community.txt	0.043261	0.000000	0.000000	
Gasoline Prices Hit Record High.txt	0.000000	0.000000	0.000000	
Cloning Pets.txt	0.000000	0.000000	0.105559	
Crazy Housing Prices.txt	0.054339	0.067171	0.034532	
Man Injured at Fast Food Place.txt	0.000000	0.000000	0.045375	
A Festival of Books.txt	0.000000	0.000000	0.099136	
Food Fight Erupted in Prison.txt	0.000000	0.000000	0.000000	
Better To Be Unlucky.txt	0.039381	0.000000	0.025027	
Sara Went Shopping.txt	0.000000	0.000000	0.000000	
Freeway Chase Ends at Newsstand.txt	0.000000	0.000000	0.028409	
Trees Are a Threat.txt	0.000000	0.000000	0.031423	
A Murder-Suicide.txt	0.000000	0.000000	0.101938	
Happy and Unhappy Renters.txt	0.000000	0.065236	0.033538	
Pulling Out Nine Tons of Trash.txt	0.059387	0.000000	0.037740	

	years	yelling	yellow	\
Jerry Decided To Buy a Gun.txt	0.060975	0.000000	0.000000	
Rentals at the Oceanside Community.txt	0.027493	0.000000	0.000000	
Gasoline Prices Hit Record High.txt	0.033227	0.000000	0.000000	
Cloning Pets.txt	0.000000	0.000000	0.000000	
Crazy Housing Prices.txt	0.069065	0.000000	0.000000	
Man Injured at Fast Food Place.txt	0.000000	0.000000	0.000000	
A Festival of Books.txt	0.066090	0.000000	0.000000	
Food Fight Erupted in Prison.txt	0.000000	0.000000	0.000000	
Better To Be Unlucky.txt	0.025027	0.000000	0.000000	
Sara Went Shopping.txt	0.077123	0.000000	0.000000	
Freeway Chase Ends at Newsstand.txt	0.000000	0.063639	0.000000	
Trees Are a Threat.txt	0.062845	0.000000	0.000000	
A Murder-Suicide.txt	0.101938	0.000000	0.000000	
Happy and Unhappy Renters.txt	0.000000	0.000000	0.000000	
Pulling Out Nine Tons of Trash.txt	0.037740	0.000000	0.084542	

TF-IDF Matrix:

	00	000	09 \
Jerry Decided To Buy a Gun.txt	0.000000	0.037550	0.000000
Rentals at the Oceanside Community.txt	0.000000	0.033862	0.000000
Gasoline Prices Hit Record High.txt	0.000000	0.000000	0.074431
Cloning Pets.txt	0.000000	0.130012	0.000000
Crazy Housing Prices.txt	0.000000	0.170128	0.000000
Man Injured at Fast Food Place.txt	0.101645	0.000000	0.000000
A Festival of Books.txt	0.000000	0.081401	0.000000
Food Fight Erupted in Prison.txt	0.000000	0.000000	0.000000
Better To Be Unlucky.txt	0.000000	0.123298	0.000000
Sara Went Shopping.txt	0.000000	0.000000	0.000000
Freeway Chase Ends at Newsstand.txt	0.000000	0.000000	0.000000
Trees Are a Threat.txt	0.000000	0.154807	0.000000
A Murder-Suicide.txt	0.000000	0.000000	0.000000
Happy and Unhappy Renters.txt	0.000000	0.000000	0.000000
Pulling Out Nine Tons of Trash.txt	0.000000	0.000000	0.000000

	10	1000	1000x \
Jerry Decided To Buy a Gun.txt	0.040511	0.000000	0.000000
Rentals at the Oceanside Community.txt	0.036532	0.000000	0.000000
Gasoline Prices Hit Record High.txt	0.044151	0.000000	0.000000
Cloning Pets.txt	0.000000	0.000000	0.000000
Crazy Housing Prices.txt	0.045886	0.000000	0.000000
Man Injured at Fast Food Place.txt	0.000000	0.000000	0.000000
A Festival of Books.txt	0.043910	0.000000	0.000000
Food Fight Erupted in Prison.txt	0.000000	0.000000	0.000000
Better To Be Unlucky.txt	0.000000	0.056063	0.112126
Sara Went Shopping.txt	0.051240	0.000000	0.000000
Freeway Chase Ends at Newsstand.txt	0.000000	0.000000	0.000000
Trees Are a Threat.txt	0.000000	0.000000	0.000000
A Murder-Suicide.txt	0.000000	0.000000	0.000000
Happy and Unhappy Renters.txt	0.000000	0.000000	0.000000
Pulling Out Nine Tons of Trash.txt	0.000000	0.000000	0.000000

	100x	10x	11 \
Jerry Decided To Buy a Gun.txt	0.000000	0.000000	0.000000
Rentals at the Oceanside Community.txt	0.000000	0.000000	0.000000
Gasoline Prices Hit Record High.txt	0.000000	0.000000	0.000000
Cloning Pets.txt	0.000000	0.000000	0.000000
Crazy Housing Prices.txt	0.000000	0.000000	0.000000
Man Injured at Fast Food Place.txt	0.000000	0.000000	0.000000
A Festival of Books.txt	0.000000	0.000000	0.000000
Food Fight Erupted in Prison.txt	0.000000	0.000000	0.000000
Better To Be Unlucky.txt	0.056063	0.056063	0.000000
Sara Went Shopping.txt	0.000000	0.000000	0.000000
Freeway Chase Ends at Newsstand.txt	0.000000	0.000000	0.000000
Trees Are a Threat.txt	0.000000	0.000000	0.070389
A Murder-Suicide.txt	0.000000	0.000000	0.000000
Happy and Unhappy Renters.txt	0.000000	0.000000	0.000000
Pulling Out Nine Tons of Trash.txt	0.000000	0.000000	0.000000

	yield	young
Jerry Decided To Buy a Gun.txt	0.000000	0.000000
Rentals at the Oceanside Community.txt	0.000000	0.000000
Gasoline Prices Hit Record High.txt	0.000000	0.064631
Cloning Pets.txt	0.000000	0.000000
Crazy Housing Prices.txt	0.000000	0.000000
Man Injured at Fast Food Place.txt	0.000000	0.088261
A Festival of Books.txt	0.000000	0.000000
Food Fight Erupted in Prison.txt	0.000000	0.000000
Better To Be Unlucky.txt	0.000000	0.000000
Sara Went Shopping.txt	0.000000	0.000000
Freeway Chase Ends at Newsstand.txt	0.063639	0.000000
Trees Are a Threat.txt	0.000000	0.000000
A Murder-Suicide.txt	0.000000	0.000000
Happy and Unhappy Renters.txt	0.000000	0.000000
Pulling Out Nine Tons of Trash.txt	0.000000	0.000000

Q1.4)

	Cloning Pets.txt \
Jerry Decided To Buy a Gun.txt	0.036922
Rentals at the Oceanside Community.txt	0.013216
Gasoline Prices Hit Record High.txt	0.019425
Cloning Pets.txt	1.000000
Crazy Housing Prices.txt	0.067994
Man Injured at Fast Food Place.txt	0.047735
A Festival of Books.txt	0.033434
Food Fight Erupted in Prison.txt	0.015805
Better To Be Unlucky.txt	0.044856
Sara Went Shopping.txt	0.027891
Freeway Chase Ends at Newsstand.txt	0.002999
Trees Are a Threat.txt	0.038140
A Murder-Suicide.txt	0.047519
Happy and Unhappy Renters.txt	0.019277
Pulling Out Nine Tons of Trash.txt	0.018124

	Crazy Housing Prices.txt \
Jerry Decided To Buy a Gun.txt	0.071246
Rentals at the Oceanside Community.txt	0.061602
Gasoline Prices Hit Record High.txt	0.083099
Cloning Pets.txt	0.067994
Crazy Housing Prices.txt	1.000000
Man Injured at Fast Food Place.txt	0.023678
A Festival of Books.txt	0.068660
Food Fight Erupted in Prison.txt	0.030588
Better To Be Unlucky.txt	0.060575
Sara Went Shopping.txt	0.036512
Freeway Chase Ends at Newsstand.txt	0.008403
Trees Are a Threat.txt	0.095101
A Murder-Suicide.txt	0.087660
Happy and Unhappy Renters.txt	0.081159
Pulling Out Nine Tons of Trash.txt	0.047411

	Man Injured at Fast Food Place.txt \
Jerry Decided To Buy a Gun.txt	0.069547
Rentals at the Oceanside Community.txt	0.014456
Gasoline Prices Hit Record High.txt	0.052478
Cloning Pets.txt	0.047735
Crazy Housing Prices.txt	0.023678
Man Injured at Fast Food Place.txt	1.000000
A Festival of Books.txt	0.030419
Food Fight Erupted in Prison.txt	0.007413
Better To Be Unlucky.txt	0.006747
Sara Went Shopping.txt	0.027431
Freeway Chase Ends at Newsstand.txt	0.013212
Trees Are a Threat.txt	0.010896
A Murder-Suicide.txt	0.040404
Happy and Unhappy Renters.txt	0.027890
Pulling Out Nine Tons of Trash.txt	0.011482

	Sara Went Shopping.txt \
Jerry Decided To Buy a Gun.txt	0.057951
Rentals at the Oceanside Community.txt	0.008003
Gasoline Prices Hit Record High.txt	0.054323
Cloning Pets.txt	0.027891
Crazy Housing Prices.txt	0.036512
Man Injured at Fast Food Place.txt	0.027431
A Festival of Books.txt	0.013832
Food Fight Erupted in Prison.txt	0.006738
Better To Be Unlucky.txt	0.021947
Sara Went Shopping.txt	1.000000
Freeway Chase Ends at Newsstand.txt	0.012001
Trees Are a Threat.txt	0.027941
A Murder-Suicide.txt	0.052147
Happy and Unhappy Renters.txt	0.021475
Pulling Out Nine Tons of Trash.txt	0.021339

	Freeway Chase Ends at Newsstand.txt \
Jerry Decided To Buy a Gun.txt	0.032660
Rentals at the Oceanside Community.txt	0.018965
Gasoline Prices Hit Record High.txt	0.027218
Cloning Pets.txt	0.002999
Crazy Housing Prices.txt	0.008403
Man Injured at Fast Food Place.txt	0.013212
A Festival of Books.txt	0.062476
Food Fight Erupted in Prison.txt	0.021445
Better To Be Unlucky.txt	0.037185
Sara Went Shopping.txt	0.012001
Freeway Chase Ends at Newsstand.txt	1.000000
Trees Are a Threat.txt	0.006608
A Murder-Suicide.txt	0.037314
Happy and Unhappy Renters.txt	0.030826
Pulling Out Nine Tons of Trash.txt	0.020122

	Trees Are a Threat.txt \
Jerry Decided To Buy a Gun.txt	0.018511
Rentals at the Oceanside Community.txt	0.061548
Gasoline Prices Hit Record High.txt	0.044802
Cloning Pets.txt	0.038140
Crazy Housing Prices.txt	0.095101
Man Injured at Fast Food Place.txt	0.010896
A Festival of Books.txt	0.050294
Food Fight Erupted in Prison.txt	0.026958
Better To Be Unlucky.txt	0.064875
Sara Went Shopping.txt	0.027941
Freeway Chase Ends at Newsstand.txt	0.006608
Trees Are a Threat.txt	1.000000
A Murder-Suicide.txt	0.038364
Happy and Unhappy Renters.txt	0.016480
Pulling Out Nine Tons of Trash.txt	0.036048

	A Festival of Books.txt \
Jerry Decided To Buy a Gun.txt	0.045236
Rentals at the Oceanside Community.txt	0.071991
Gasoline Prices Hit Record High.txt	0.047632
Cloning Pets.txt	0.033434
Crazy Housing Prices.txt	0.068660
Man Injured at Fast Food Place.txt	0.030419
A Festival of Books.txt	1.000000
Food Fight Erupted in Prison.txt	0.053347
Better To Be Unlucky.txt	0.033868
Sara Went Shopping.txt	0.013832
Freeway Chase Ends at Newsstand.txt	0.062476
Trees Are a Threat.txt	0.050294
A Murder-Suicide.txt	0.085042
Happy and Unhappy Renters.txt	0.041805
Pulling Out Nine Tons of Trash.txt	0.045189

	Food Fight Erupted in Prison.txt \
Jerry Decided To Buy a Gun.txt	0.020018
Rentals at the Oceanside Community.txt	0.033600
Gasoline Prices Hit Record High.txt	0.013885
Cloning Pets.txt	0.015805
Crazy Housing Prices.txt	0.030588
Man Injured at Fast Food Place.txt	0.007413
A Festival of Books.txt	0.053347
Food Fight Erupted in Prison.txt	1.000000
Better To Be Unlucky.txt	0.007922
Sara Went Shopping.txt	0.006738
Freeway Chase Ends at Newsstand.txt	0.021445
Trees Are a Threat.txt	0.026958
A Murder-Suicide.txt	0.070986
Happy and Unhappy Renters.txt	0.015118
Pulling Out Nine Tons of Trash.txt	0.025997

	Better To Be Unlucky.txt \
Jerry Decided To Buy a Gun.txt	0.032867
Rentals at the Oceanside Community.txt	0.012303
Gasoline Prices Hit Record High.txt	0.026767
Cloning Pets.txt	0.044856
Crazy Housing Prices.txt	0.060575
Man Injured at Fast Food Place.txt	0.006747
A Festival of Books.txt	0.033868
Food Fight Erupted in Prison.txt	0.007922
Better To Be Unlucky.txt	1.000000
Sara Went Shopping.txt	0.021947
Freeway Chase Ends at Newsstand.txt	0.037185
Trees Are a Threat.txt	0.064875
A Murder-Suicide.txt	0.036279
Happy and Unhappy Renters.txt	0.039435
Pulling Out Nine Tons of Trash.txt	0.039083

	Cloning Pets.txt \
Jerry Decided To Buy a Gun.txt	0.036922
Rentals at the Oceanside Community.txt	0.013216
Gasoline Prices Hit Record High.txt	0.019425
Cloning Pets.txt	1.000000
Crazy Housing Prices.txt	0.067994
Man Injured at Fast Food Place.txt	0.047735
A Festival of Books.txt	0.033434
Food Fight Erupted in Prison.txt	0.015805
Better To Be Unlucky.txt	0.044856
Sara Went Shopping.txt	0.027891
Freeway Chase Ends at Newsstand.txt	0.002999
Trees Are a Threat.txt	0.038140
A Murder-Suicide.txt	0.047519
Happy and Unhappy Renters.txt	0.019277
Pulling Out Nine Tons of Trash.txt	0.018124

	Crazy Housing Prices.txt \
Jerry Decided To Buy a Gun.txt	0.071246
Rentals at the Oceanside Community.txt	0.061602
Gasoline Prices Hit Record High.txt	0.083099
Cloning Pets.txt	0.067994
Crazy Housing Prices.txt	1.000000
Man Injured at Fast Food Place.txt	0.023678
A Festival of Books.txt	0.068660
Food Fight Erupted in Prison.txt	0.030588
Better To Be Unlucky.txt	0.060575
Sara Went Shopping.txt	0.036512
Freeway Chase Ends at Newsstand.txt	0.008403
Trees Are a Threat.txt	0.095101
A Murder-Suicide.txt	0.087660
Happy and Unhappy Renters.txt	0.081159
Pulling Out Nine Tons of Trash.txt	0.047411

	Man Injured at Fast Food Place.txt \
Jerry Decided To Buy a Gun.txt	0.069547
Rentals at the Oceanside Community.txt	0.014456
Gasoline Prices Hit Record High.txt	0.052478
Cloning Pets.txt	0.047735
Crazy Housing Prices.txt	0.023678
Man Injured at Fast Food Place.txt	1.000000
A Festival of Books.txt	0.030419
Food Fight Erupted in Prison.txt	0.007413
Better To Be Unlucky.txt	0.006747
Sara Went Shopping.txt	0.027431
Freeway Chase Ends at Newsstand.txt	0.013212
Trees Are a Threat.txt	0.010896
A Murder-Suicide.txt	0.040404
Happy and Unhappy Renters.txt	0.027890
Pulling Out Nine Tons of Trash.txt	0.011482

	A Murder-Suicide.txt \	
Jerry Decided To Buy a Gun.txt	0.080462	
Rentals at the Oceanside Community.txt	0.038473	
Gasoline Prices Hit Record High.txt	0.083924	
Cloning Pets.txt	0.047519	
Crazy Housing Prices.txt	0.087660	
Man Injured at Fast Food Place.txt	0.040404	
A Festival of Books.txt	0.085042	
Food Fight Erupted in Prison.txt	0.070986	
Better To Be Unlucky.txt	0.036279	
Sara Went Shopping.txt	0.052147	
Freeway Chase Ends at Newsstand.txt	0.037314	
Trees Are a Threat.txt	0.038364	
A Murder-Suicide.txt	1.000000	
Happy and Unhappy Renters.txt	0.081704	
Pulling Out Nine Tons of Trash.txt	0.055772	
	Happy and Unhappy Renters.txt \	
Jerry Decided To Buy a Gun.txt	0.053486	
Rentals at the Oceanside Community.txt	0.075391	
Gasoline Prices Hit Record High.txt	0.028659	
Cloning Pets.txt	0.019277	
Crazy Housing Prices.txt	0.081159	
Man Injured at Fast Food Place.txt	0.027890	
A Festival of Books.txt	0.041805	
Food Fight Erupted in Prison.txt	0.015118	
Better To Be Unlucky.txt	0.039435	
Sara Went Shopping.txt	0.021475	
Freeway Chase Ends at Newsstand.txt	0.030826	
Trees Are a Threat.txt	0.016480	
A Murder-Suicide.txt	0.081704	
Happy and Unhappy Renters.txt	1.000000	
Pulling Out Nine Tons of Trash.txt	0.047357	
	Pulling Out Nine Tons of Trash.txt	
Jerry Decided To Buy a Gun.txt	0.049805	
Rentals at the Oceanside Community.txt	0.054498	
Gasoline Prices Hit Record High.txt	0.024639	
Cloning Pets.txt	0.018124	
Crazy Housing Prices.txt	0.047411	
Man Injured at Fast Food Place.txt	0.011482	
A Festival of Books.txt	0.045189	
Food Fight Erupted in Prison.txt	0.025997	
Better To Be Unlucky.txt	0.039083	
Sara Went Shopping.txt	0.021339	
Freeway Chase Ends at Newsstand.txt	0.020122	
Trees Are a Threat.txt	0.036048	
A Murder-Suicide.txt	0.055772	
Happy and Unhappy Renters.txt	0.047357	
Pulling Out Nine Tons of Trash.txt	1.000000	



# Cosine Similarity Matrix:

	Jerry Decided To Buy a Gun.txt \
Jerry Decided To Buy a Gun.txt	1.000000
Rentals at the Oceanside Community.txt	0.011724
Gasoline Prices Hit Record High.txt	0.052392
Cloning Pets.txt	0.036922
Crazy Housing Prices.txt	0.071246
Man Injured at Fast Food Place.txt	0.069547
A Festival of Books.txt	0.045236
Food Fight Erupted in Prison.txt	0.020018
Better To Be Unlucky.txt	0.032867
Sara Went Shopping.txt	0.057951
Freeway Chase Ends at Newsstand.txt	0.032660
Trees Are a Threat.txt	0.018511
A Murder-Suicide.txt	0.080462
Happy and Unhappy Renters.txt	0.053486
Pulling Out Nine Tons of Trash.txt	0.049805

	Rentals at the Oceanside Community.txt \
Jerry Decided To Buy a Gun.txt	0.011724
Rentals at the Oceanside Community.txt	1.000000
Gasoline Prices Hit Record High.txt	0.035434
Cloning Pets.txt	0.013216
Crazy Housing Prices.txt	0.061602
Man Injured at Fast Food Place.txt	0.014456
A Festival of Books.txt	0.071991
Food Fight Erupted in Prison.txt	0.033600
Better To Be Unlucky.txt	0.012303
Sara Went Shopping.txt	0.008003
Freeway Chase Ends at Newsstand.txt	0.018965
Trees Are a Threat.txt	0.061548
A Murder-Suicide.txt	0.038473
Happy and Unhappy Renters.txt	0.075391
Pulling Out Nine Tons of Trash.txt	0.054498

	Gasoline Prices Hit Record High.txt \
Jerry Decided To Buy a Gun.txt	0.052392
Rentals at the Oceanside Community.txt	0.035434
Gasoline Prices Hit Record High.txt	1.000000
Cloning Pets.txt	0.019425
Crazy Housing Prices.txt	0.083099
Man Injured at Fast Food Place.txt	0.052478
A Festival of Books.txt	0.047632
Food Fight Erupted in Prison.txt	0.013885
Better To Be Unlucky.txt	0.026767
Sara Went Shopping.txt	0.054323
Freeway Chase Ends at Newsstand.txt	0.027218
Trees Are a Threat.txt	0.044802
A Murder-Suicide.txt	0.083924
Happy and Unhappy Renters.txt	0.028659
Pulling Out Nine Tons of Trash.txt	0.024639