

# Aakash Rami

## BIA 660

### Assignment 6

#### I) Topic names along with heavily weighted word in each topic

('Topic0', '0.007\*"movie" + 0.006\*"song" + 0.006\*"album" + 0.005\*"released" + 0.004\*"series" + 0.004\*"music" + 0.004\*"singer" + 0.003\*"film" + 0.003\*"chart" + 0.003\*"band"')

('Topic1', '0.007\*"de" + 0.005\*"bishop" + 0.004\*"catholic" + 0.003\*"prince" + 0.003\*"french" + 0.003\*"king" + 0.003\*"al" + 0.003\*"france" + 0.003\*"el" + 0.003\*"px"'),

('Topic2', '0.034\*"manga" + 0.019\*"anime" + 0.017\*"wwe" + 0.012\*"nxt" + 0.010\*"apps" + 0.010\*"fame" + 0.009\*"inducted" + 0.009\*"sitcom" + 0.008\*"voiced" + 0.007\*"vowel"'),

('Topic3', '0.008\*"game" + 0.004\*"dinosaur" + 0.004\*"pancreatic" + 0.004\*"computer" + 0.003\*"sauropod" + 0.003\*"windows" + 0.003\*"charts" + 0.002\*"system" + 0.002\*"video" + 0.002\*"use"')

('Topic4', '0.031\*"taluk" + 0.016\*"pakistan" + 0.012\*"merit" + 0.011\*"qualification" + 0.009\*"sharma" + 0.008\*"lahore" + 0.008\*"punjab" + 0.007\*"adviser" + 0.007\*"hemorrhage" + 0.007\*"raj"'),

('Topic5', '0.023\*"hurricane" + 0.018\*"storm" + 0.012\*"tehran" + 0.008\*"administrator" + 0.008\*"iceland" + 0.008\*"tropical" + 0.008\*"mep" + 0.007\*"damage" + 0.007\*"baden" + 0.006\*"eliminated"')

('Topic6', '0.009\*"nordic" + 0.008\*"governors" + 0.007\*"columnist" + 0.006\*"prasad" + 0.006\*"al" + 0.006\*"delegates" + 0.005\*"nominations" + 0.005\*"belarusian" + 0.005\*"streaming" + 0.005\*"typeface"')

('Topic7', '0.008\*"complications" + 0.006\*"brazilian" + 0.005\*"cardiac" + 0.004\*"paulo" + 0.004\*"dementia" + 0.004\*"são" + 0.004\*"disease" + 0.004\*"diagnosed" + 0.004\*"brazil" + 0.003\*"liver"')

('Topic8', '0.010\*"music" + 0.009\*"guitar" + 0.007\*"productions" + 0.007\*"leukemia" + 0.006\*"composer" + 0.006\*"orchestra" + 0.005\*"congestive" + 0.005\*"bass" + 0.005\*"loop" + 0.004\*"alonso"')

('Topic9', '0.056\*"county" + 0.020\*"census" + 0.013\*"lake" + 0.012\*"population" + 0.011\*"seat" + 0.011\*"park" + 0.009\*"area" + 0.008\*"river" + 0.007\*"communities" + 0.007\*"idaho"')

('Topic10', '0.021\*"river" + 0.020\*"district" + 0.016\*"arrondissement" + 0.016\*"municipality" + 0.015\*"saint" + 0.014\*"department" + 0.014\*"province" + 0.013\*"population" + 0.011\*"de" + 0.010\*"sur"')

('Topic11', '0.017\*"unincorporated" + 0.011\*"pitcher" + 0.008\*"queensland"  
+ 0.007\*"endorsed" + 0.005\*"baseman" + 0.005\*"dish" + 0.005\*"conservatory"  
+ 0.004\*"cuisine" + 0.004\*"spanned" + 0.004\*"edmonton"')

('Topic12', '0.013\*"professionally" + 0.012\*"mrt" + 0.012\*"constituency"  
+ 0.009\*"tag" + 0.009\*"ordained" + 0.007\*"singapore" + 0.007\*"nanjing" + 0.  
.007\*"coached" + 0.006\*"freddy" + 0.006\*"appointment"')

('Topic13', '0.005\*"age" + 0.005\*"february" + 0.004\*"december" + 0.004\*"a  
ged" + 0.004\*"august" + 0.004\*"july" + 0.004\*"september" + 0.004\*"actor" +  
0.004\*"november" + 0.004\*"june"')

('Topic14', '0.006\*"species" + 0.004\*"water" + 0.003\*"found" + 0.003\*"gen  
us" + 0.002\*"para" + 0.002\*"long" + 0.002\*"red" + 0.002\*"males" + 0.002\*"b  
irds" + 0.002\*"body"')

('Topic15', '0.007\*"indian" + 0.007\*"india" + 0.006\*"hockey" + 0.005\*"kera  
la" + 0.005\*"singh" + 0.004\*"played" + 0.004\*"pradesh" + 0.004\*"nhl" + 0.0  
04\*"league" + 0.003\*"bangladesh"')

('Topic16', '0.022\*"bar" + 0.015\*"cup" + 0.014\*"text" + 0.012\*"league" + 0  
.012\*"goals" + 0.011\*"uefa" + 0.011\*"till" + 0.010\*"club" + 0.009\*"footbal  
l" + 0.009\*"fontsize"')

('Topic17', '0.010\*"emeritus" + 0.007\*"japanese" + 0.006\*"japan" + 0.005\*"  
sulawesi" + 0.005\*"zee" + 0.005\*"hebrew" + 0.004\*"thai" + 0.004\*"jerusalem  
" + 0.004\*"pageant" + 0.004\*"rabbi"')

('Topic18', '0.020\*"arrondissements" + 0.012\*"franchise" + 0.008\*"star" +  
0.007\*"space" + 0.007\*"galaxy" + 0.007\*"nasa" + 0.007\*"haute" + 0.006\*"cer  
tifications" + 0.006\*"duo" + 0.006\*"constellation"')

('Topic19', '0.005\*"party" + 0.004\*"president" + 0.004\*"minister" + 0.004\*  
"politician" + 0.004\*"university" + 0.003\*"member" + 0.003\*"served" + 0.00  
3\*"election" + 0.002\*"age" + 0.002\*"government"')

## II) Random Articles and list of topics assigned to each article

[https://en.wikipedia.org/wiki/House\\_of\\_Travel\\_New\\_Zealand](https://en.wikipedia.org/wiki/House_of_Travel_New_Zealand)

(Topic0, 0.12835535)  
(Topic1, 0.086408444)  
(Topic7, 0.02821285)  
(Topic9, 0.21741326)  
(Topic13, 0.079468556)  
(Topic15, 0.034921773)  
(Topic17, 0.016076958)  
(Topic19, 0.392534)

[https://en.wikipedia.org/wiki/Robert\\_M.\\_Scotten](https://en.wikipedia.org/wiki/Robert_M._Scotten)

```
[ (Topic1, 0.17825942)
  (Topic3, 0.045222808)
  (Topic15, 0.02922789)
  (Topic17, 0.028153796)
  (Topic19, 0.6993992) ]
```

[https://en.wikipedia.org/wiki/Vaishno\\_Devi](https://en.wikipedia.org/wiki/Vaishno_Devi)

```
[ (Topic0, 0.12531745)
  (Topic1, 0.091022365)
  (Topic7, 0.025777908)
  (Topic13, 0.16631003)
  (Topic15, 0.023228452)
  (Topic19, 0.5418987) ]
```

### **III) Summarizing the results**

I obtained links to three Wikipedia pages: The House of New Zealand, Robert M. Schotten and Vaishno Devi.

House of Travel page has highest correlation with Topic1 (0.086408444);

Robert M. Schotten page has highest correlation with Topic19(0.6993992)

Vaishno Devi page has highest correlation with Topic19(0.5418987).

Vaishno Devi and Robert Schotten have correlevance with the same topic, and highly relevant words.