

PCA Analysis Results

Principal Component	Explained Variance	Explained Variance (%)	Eigenvalue
PC1	0.0353	3.53%	0.0183
PC2	0.0321	3.21%	0.0167
PC3	0.0208	2.08%	0.0108
PC4	0.0171	1.71%	0.0089
PC5	0.0154	1.54%	0.0080
Total explained variance (first 10 PCs): 0.1808 (18.08%)			

Table 1: Explained variance and eigenvalues for the first five principal components.

Top 10 selected concepts (highest cosine similarity):

Rank	Concept	Similarity
1	someone	0.6816
2	inside	0.6798
3	somebody	0.6705
4	humans	0.6684
5	jealousy	0.6483
6	folks	0.6480
7	desperate	0.6443
8	face-off	0.6414
9	robbery	0.6413
10	stranger	0.6388

Top 5 Positive/Negative Concepts for Each PC (1–5):

PC	Top 5 Positive Concepts	Top 5 Negative Concepts
PC1	newark 0.1924	mantra -0.1340
	america 0.1851	bollywood -0.1015
	york 0.1850	love-song -0.0783
	american 0.1848	karma -0.0695
	hurricane 0.1845	love -0.0677
PC2	suicide 0.2036	contemporaries -0.0860
	death 0.1973	professionalism -0.0819
	snake 0.1951	mid-sixties -0.0802
	stupid 0.1911	contemporary -0.0765
	raccoon 0.1780	mid-twenties -0.0755
PC3	combat 0.1720	fiancee -0.2790
	weaponry 0.1616	girlfriend -0.2664
	samurai 0.1550	bride-to-be -0.2624
	bombardment 0.1422	fiance -0.2505
	wartime 0.1395	bride -0.2503
PC4	infancy 0.2108	gangster -0.2546
	childhood 0.1634	actor -0.2402
	autumn 0.1632	gunman -0.2182
	orphanage 0.1621	bollywood -0.2137
	upbringing 0.1575	murderer -0.1924
PC5	vocalist 0.1941	crime -0.2339
	motherhood 0.1816	criminal -0.2015
	actress 0.1788	extortion -0.1916
	teens 0.1653	bribery -0.1876
	co-ed 0.1635	businessman -0.1833

Table 2: Top 5 most positive and negative concept words for each of the first five principal components.