

Task 2: Shared Vocabulary Analysis (Swahili)

| Word | Total Count |
|----------|---|
| na | 1146 |
| ni | Top 20 Words Common to All 5 Categories (Positive) 1079 |
| ya | 975 |
| hawa | 736 |
| wa | 580 |
| kwa | 508 |
| pussy | 442 |
| kama | 426 |
| kutombwa | 384 |
| la | 371 |
| tu | 309 |
| wewe | 279 |
| matako | 225 |
| mjinga | 216 |
| rt | 213 |
| watu | 202 |
| za | 196 |
| majembe | 194 |
| sana | 180 |
| sisi | 168 |

| Word | Count Untagged |
|--------------|--|
| ya | 1239 |
| na | Top 20 Words in UNTAGGED Samples (All Categories = 0) 1108 |
| ni | 830 |
| wa | 701 |
| kwa | 658 |
| pussy | 525 |
| kutombwa | 477 |
| la | 474 |
| kama | 387 |
| bitch | 298 |
| wewe | 258 |
| kwenye | 227 |
| kenya | 225 |
| jembe | 214 |
| tu | 210 |
| kenyadecides | 210 |
| kuwa | 202 |
| hii | 201 |
| iebc | 200 |
| za | 198 |