# Tutorial 6 Part 1

1. Consider given below, Co-occurrence vectors for four words in the Wikipedia corpus, showing six of the dimensions (hand-picked for pedagogical purposes).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **aadvark** | **…** | **computer** | **data** | **result** | **pie** | **sugar** | **…** |
| **Cherry** | 0 | … | 2 | 8 | 9 | 442 | 25 | … |
| **Strawberry** | 0 | … | 0 | 0 | 1 | 60 | 19 | … |
| **Digital** | 0 | … | 1670 | 1683 | 85 | 5 | 4 | … |
| **information** | 0 | … | 3325 | 3982 | 378 | 5 | 13 | … |

Compute the cosine similarity vectors of the given pairs:

1. Cherry, information
2. Digital, information

# Answer Key

1.

Src: Section 6.4: [6.pdf (stanford.edu)](https://web.stanford.edu/~jurafsky/slp3/6.pdf)