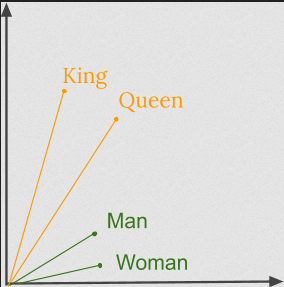
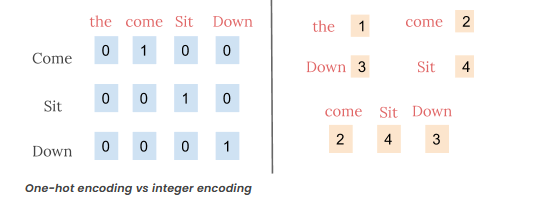
## **Word embedding**

In natural language processing, word embedding is a term used for the representation of words that are close in meaning are grouped near to one another in vector space for text analysis, typically in the form of a real-valued vector that encodes the meaning of the word such that the words that are closer in the vector space are expected to be similar in meaning.  
 

2 approaches to represent a word to machines.

1. Take all the words and find all the unique words and assign them to a number. That bag of words will be our dictionary. When a particular word comes up in the review, That one word will be “1” and the rest “0”. This is known as one-hot encoding and is a very inefficient approach.
2. Another approach is to do integer encoding by rank ordering each unique word by a number/word and every time that words come up in the review, I will have that number in an integer in an array and end up with a variable lengths of arrays that corresponds to the word of the review.



For more: <https://www.mygreatlearning.com/blog/word-embedding/>

**However, the integer-encoding is arbitrary as it does not capture any relationship between words.** Because there is no relationship between the similarity of any two words and the similarity of their encodings, this feature-weight combination is not meaningful.

Thus by using word embeddings, words that are close in meaning are grouped near to one another in vector space. Word embeddings learn relationships. Vector differences between a pair of words can be added to another word vector to find the analogous word. For example, “man” -“woman” + “queen” ≈ “king”.

## **What is word2Vec**

Word2vec is a method to efficiently create word embeddings by using a two-layer neural network.   
The input of word2vec is a text corpus and its output is a set of vectors known as feature vectors that represent words in that corpus. While Word2vec is not a deep neural network, it turns text into a numerical form that deep neural networks can understand.

The Word2Vec objective function causes the words that have a similar context to have similar embeddings. Thus in this vector space, these words are really close. Mathematically, the cosine of the angle (Q) between such vectors should be close to 1, i.e. angle close to 0.