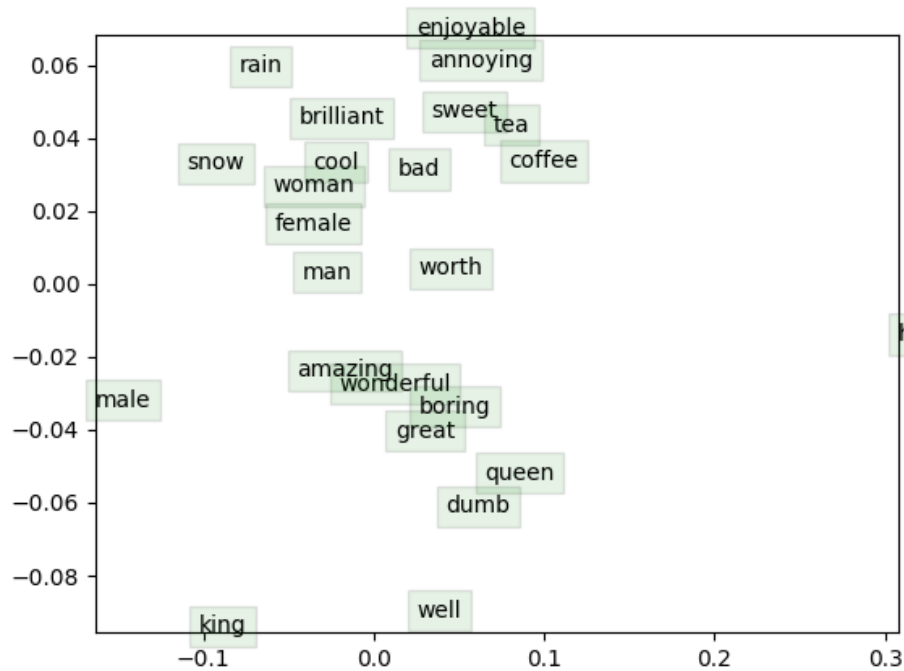


# Report

## 1. word vector plot



The above is the vector plot for the

```
visualizeWords = [  
    "great", "cool", "brilliant", "wonderful", "well", "amazing",  
    "worth", "sweet", "enjoyable", "boring", "bad", "dumb",  
    "annoying", "female", "male", "queen", "king", "man", "woman", "rain", "snow",  
    "hail", "coffee", "tea"]
```

Firstly, we find that:

$$\text{king} - \text{male} = \text{queen} - \text{female}.$$

Secondly, we find similar functional words are near to each other, for instance:

- 'tea' and 'coffee' are near to each other because they are both beverages.
- 'amazing', 'wonderful', 'boring', 'great' are near to each other because they are all adj. that some experience or object.
- 'rain', 'snow' is near to each other because they are both words about weather.

This phenomenon is reasonable. Because the vector of the word represents the position of each word in the document. The words near to each other means that they are more likely to

appear in the similar position in the document. So, they act as the same function in each document.

## 2. Knn example: this part is the result from the run.py

In this part, we use the given visualization words as examples. We randomly pick 10 words from the visualizeWords list. And print out its k nearest neighbors. And explain some result.

The output of the code:

```
sanity check: cost at convergence should be around or below 10
training took 0 seconds
The target word is :female
The top 10 neighbors words are:
['illuminates', 'skateboarding', 'dancing', 'fear-inducing', 'hand', 'kidman', 'dogs', 'wades', 'soderbergh', 'clear']

The target word is :worth
The top 10 neighbors words are:
['creatively', 'fascinate', 'painting', 'dilbert', 'devolves', 'lived-in', 'dating', 'laid', 'dollars', 'revulsion']

The target word is :brilliant
The top 10 neighbors words are:
['undermines', 'stephen', 'liberating', 'ugly-duckling', 'derisions', 'devilishly', 'drawn', 'halfway', 'missing', 'liberally']

The target word is :wonderful
The top 10 neighbors words are:
['inclination', 'artifice', 'certain', 'repulsive', 'choppy', 'decision', 'gag', 'hell-jacket', 'unforced', 'stagy']

The target word is :well
The top 10 neighbors words are:
['labored', 'countenance', 'hades', 'available', 'entirety', 'salacious', 'personifying', 'hypocrisy', 'inappropriate', 'again']

The target word is :hail
The top 10 neighbors words are:
['skullduggery', 'riveted', '88', 'scum', 'fighter', 'fish', 'overblown', 'cub', 'unrepentant', 'fringe']

The target word is :cool
The top 10 neighbors words are:
['rollicking', 'schmaltzy', 'fast-edit', 'hospitals', 'cynicism', 'mainstream', 'wall', 'clinch', 'right', 'intelligence']
```

```
The target word is :man
The top 10 neighbors words are:
['subjects', 'whole', 'stylized', 'below', 'loud', 'monotonous', 'robust', 'who-wrote-shake
speare', 'snappy', 'point']

The target word is :snow
The top 10 neighbors words are:
['compensate', 'presents', 'squandering', 'savvy', 'snatch', 'viewfinder', 'bluffs', 'textb
ook', 'binging', 'free']

The target word is :snow
The top 10 neighbors words are:
['compensate', 'presents', 'squandering', 'savvy', 'snatch', 'viewfinder', 'bluffs', 'textb
ook', 'binging', 'free']

The target word is :dumb
The top 10 neighbors words are:
['overinflated', 'notably', 'foreboding', 'sardonic', 'quench', 'haunting', 'rice', 'guide
', 'uptight', 'workers']

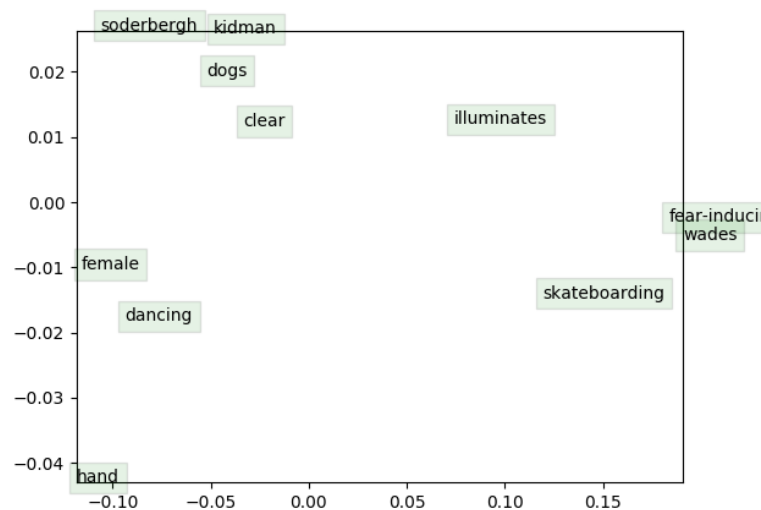
[Finished in 83.055s]
```

Some explanation for the result:

The target word is: female

The top 10 neighbors words are:

['illuminates', 'skateboarding', 'dancing', 'fear-inducing', 'hand', 'kidman', 'dogs', 'wades',  
'soderbergh', 'clear']

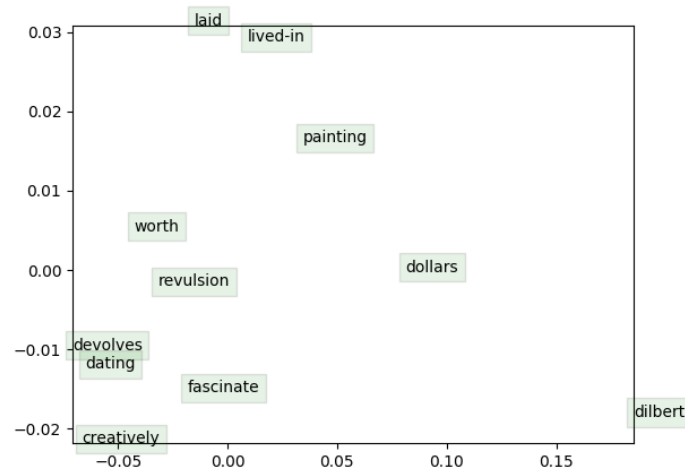


The neighbors of the word 'female' contains 'dancing', 'wades'. Because these words usually appear in the same position in a document.

The target word is :worth

The top 10 neighbors words are:

['creatively', 'fascinate', 'painting', 'dilbert', 'devolves', 'lived-in', 'dating', 'laid', 'dollars', 'revulsion']

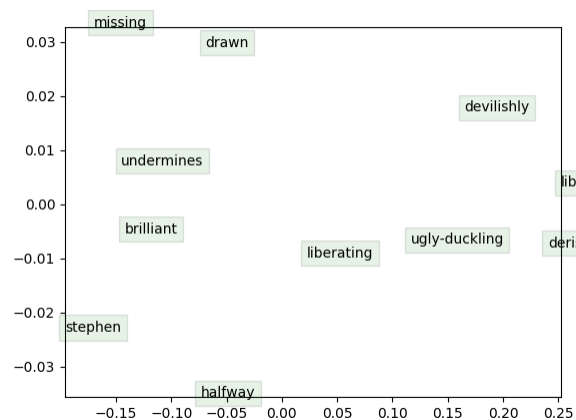


The neighbors of 'worth' contains 'dollars'. These two words both have the meaning of value, and they usually appear in the same position in a document.

The target word is :brilliant

The top 10 neighbors words are:

['undermines', 'stephen', 'liberating', 'ugly-duckling', 'derisions', 'devilishly', 'drawn', 'halfway', 'missing', 'liberally']

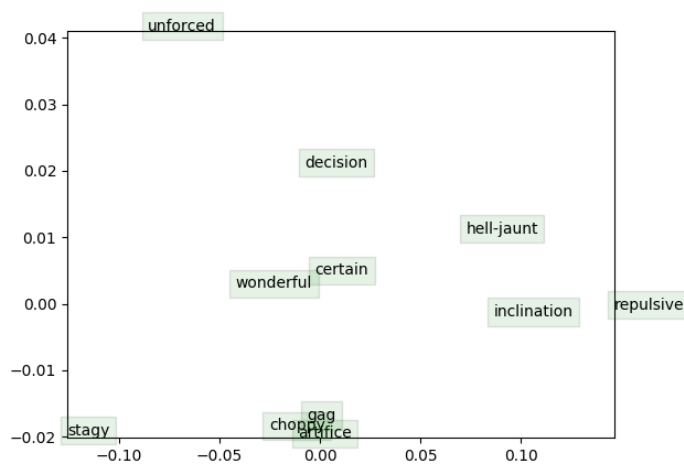


The result of target word 'brilliant' is not very reasonable.

The target word is :wonderful

The top 10 neighbors words are:

['inclination', 'artifice', 'certain', 'repulsive', 'choppy', 'decision', 'gag', 'hell-jaunt', 'unforced', 'stagy']

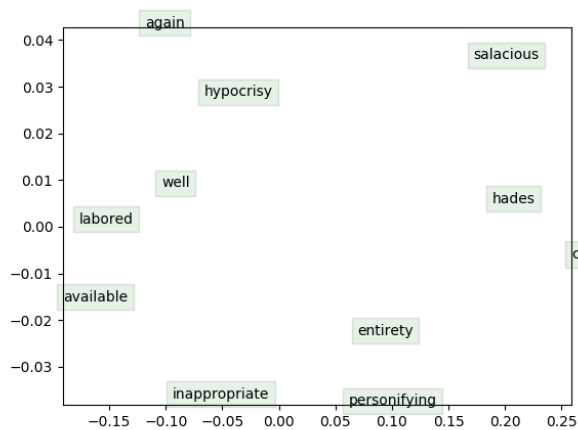


The result of target word 'wonderful' is not very reasonable.

The target word is :well

The top 10 neighbors words are:

['labored', 'countenance', 'hades', 'available', 'entirety', 'salacious', 'personifying', 'hypocrisy', 'inappropriate', 'again']

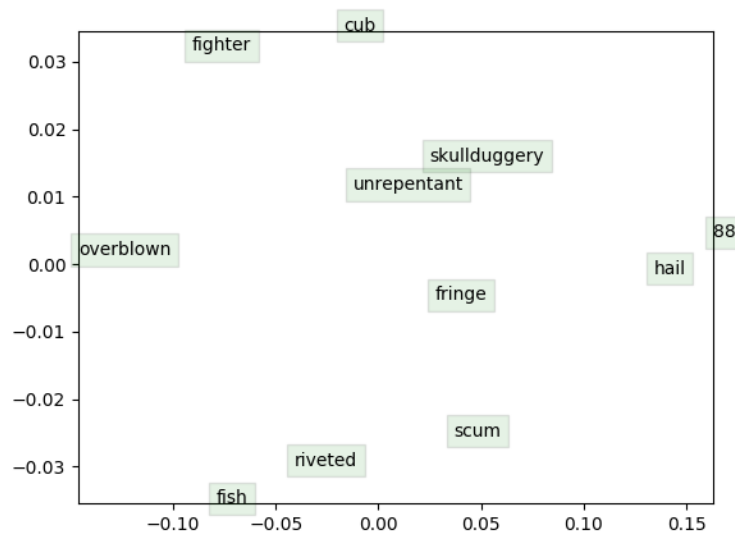


The result of target word 'well' is not very reasonable.

The target word is :hail

The top 10 neighbors words are:

['skullduggery', 'riveted', '88', 'scum', 'fighter', 'fish', 'overblown', 'cub', 'unrepentant', 'fringe']

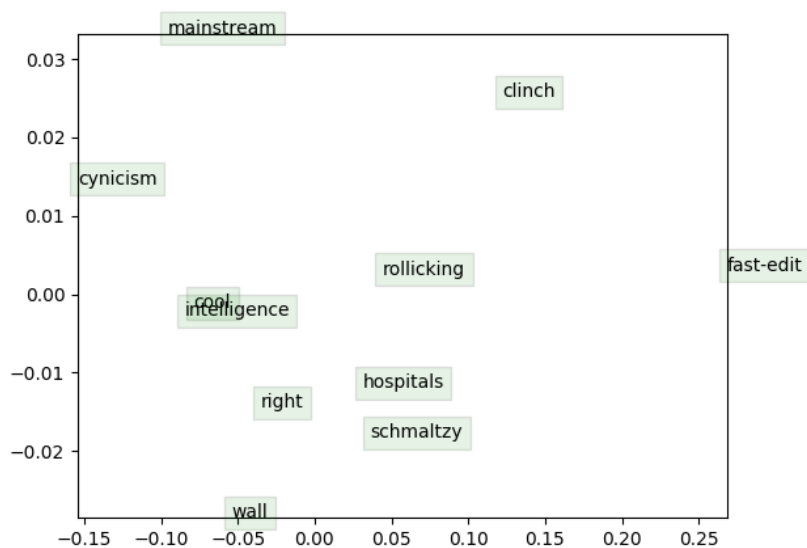


The neighbors of 'hail' contains 'skullduggery', 'delves', 'ennui-hobbled'. These words represent negative or scare meanings. It is reasonable that these words appear in the same position in a document.

The target word is :cool

The top 10 neighbors words are:

['rollicking', 'schmaltzy', 'fast-edit', 'hospitals', 'cynicism', 'mainstream', 'wall', 'clinch', 'right', 'intelligence']

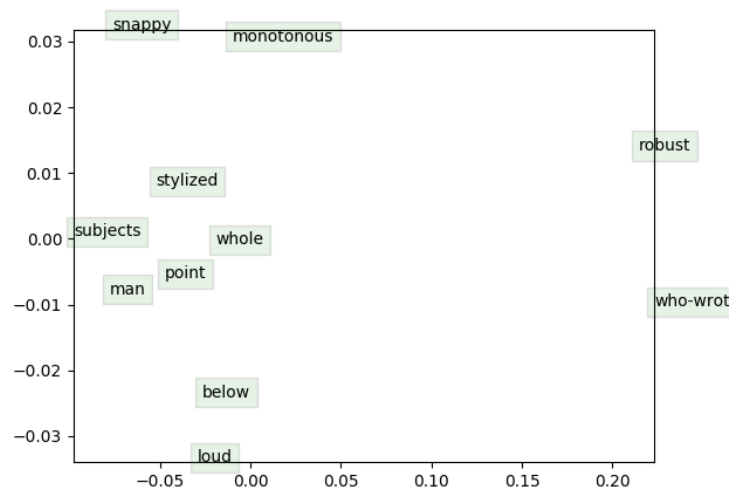


The neighbors of target word 'cool' contains 'right', 'intelligence'. These word have the same meaning in some cases. So it is reasonable that they appear in the same position in a document.

The target word is :man

The top 10 neighbors words are:

['subjects', 'whole', 'stylized', 'below', 'loud', 'monotonous', 'robust', 'who-wrote-shakespeare', 'snappy', 'point']

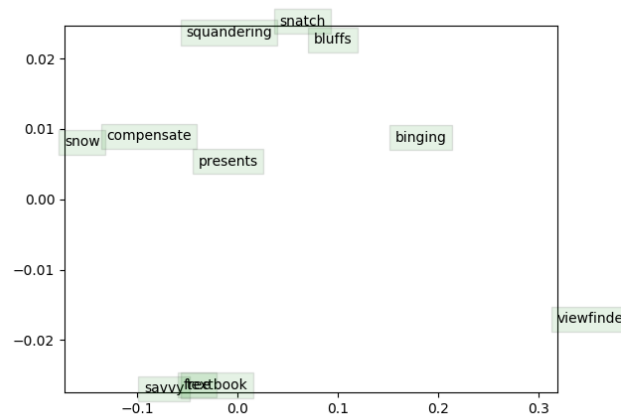


The neighbors of 'man' contains 'robust'. This word always used to describe a man. So, it is reasonable that they appear in the same position in a document.

The target word is :snow

The top 10 neighbors words are:

['compensate', 'presents', 'squandering', 'savvy', 'snatch', 'viewfinder', 'bluffs', 'textbook', 'binging', 'free']

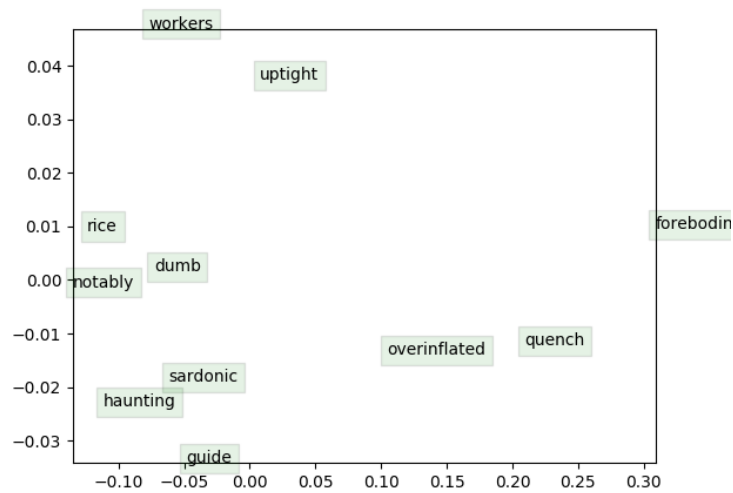


The neighbor of 'snow' contains 'soft'. This word usually used to describe the 'snow'. So, it is reasonable that they appear in the same position in a document.

The target word is :dumb

The top 10 neighbors words are:

['overinflated', 'notably', 'foreboding', 'sardonic', 'quench', 'haunting', 'rice', 'guide', 'uptight', 'workers']



The result of target word 'dumb' is not very reasonable.