

Help on module Co\_Occurrence:

NAME

Co\_Occurrence

CLASSES

builtins.object

CoOccur

```
class CoOccur(builtins.object)
```

```
| CoOccur(dataframe)
```

```
|
```

```
| This class contains methods for updating the document frequency and
```

```
| tokenizing the article.
```

```
|
```

```
| Methods defined here:
```

```
|
```

```
| __init__(self, dataframe)
```

```
| This is the constructor for this class which initialises the instance
```

```
| variable.
```

```
| :param dataframe: This is a pandas dataframe storing every news article.
```

```
|
```

```
| find(self)
```

```
| This method updates the document frequency by iterating through each
```

```
| document in the dataframe.
```

```
| :return: It returns a list which contains all unique words across all
```

```
| documents(document corpus).
```



```
| df = Empty DataFrame
```

```
| Columns: []
```

```
| Index: []
```

## DATA

```
docFreq = {}
```

```
ps = <PorterStemmer>
```

```
stop_words = {'!', '"', ''', '(', ')', ',', ...}
```

```
stopwords = <WordListCorpusReader in 'C:\\Users\\Smit\\AppData\\Roamin...
```

## FILE

```
c:\\users\\smit\\pycharmprojects\\searchengine\\co_occurrence.py
```