

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
```

```
In [2]: df = pd.read_csv('spam.csv',encoding='latin-1')[['v1','v2']]
df.columns = ['label', 'message']
df.head()
```

```
-----
FileNotFoundError                                Traceback (most recent call last)
<ipython-input-2-a92773bb66c4> in <module>
----> 1 df = pd.read_csv('spam.csv',encoding='latin-1')[['v1','v2']]
      2 df.columns = ['label', 'message']
      3 df.head()

/ext/anaconda2020.02/lib/python3.7/site-packages/pandas/io/parsers.py in read_csv(fi
lepath_or_buffer, sep, delimiter, header, names, index_col, usecols, squeeze, prefi
x, mangle_dupe_cols, dtype, engine, converters, true_values, false_values, skipiniti
alspace, skiprows, skipfooter, nrows, na_values, keep_default_na, na_filter, verbos
e, skip_blank_lines, parse_dates, infer_datetime_format, keep_date_col, date_parser,
dayfirst, cache_dates, iterator, chunksize, compression, thousands, decimal, lineter
minator, quotechar, quoting, doublequote, escapechar, comment, encoding, dialect, er
ror_bad_lines, warn_bad_lines, delim_whitespace, low_memory, memory_map, float_preci
sion)
    684     )
    685
--> 686     return _read(filepath_or_buffer, kwds)
    687
    688

/ext/anaconda2020.02/lib/python3.7/site-packages/pandas/io/parsers.py in _read(filep
ath_or_buffer, kwds)
    450
    451     # Create the parser.
--> 452     parser = TextFileReader(fp_or_buf, **kwds)
    453
    454     if chunksize or iterator:

/ext/anaconda2020.02/lib/python3.7/site-packages/pandas/io/parsers.py in __init__(se
lf, f, engine, **kwds)
    934         self.options["has_index_names"] = kwds["has_index_names"]
    935
--> 936         self._make_engine(self.engine)
    937
    938     def close(self):

/ext/anaconda2020.02/lib/python3.7/site-packages/pandas/io/parsers.py in _make_engin
e(self, engine)
   1166     def _make_engine(self, engine="c"):
   1167         if engine == "c":
-> 1168             self._engine = CParserWrapper(self.f, **self.options)
   1169         else:
   1170             if engine == "python":

/ext/anaconda2020.02/lib/python3.7/site-packages/pandas/io/parsers.py in __init__(se
lf, src, **kwds)
   1979         if kwds.get("compression") is None and encoding:
   1980             if isinstance(src, str):
-> 1981                 src = open(src, "rb")
   1982                 self.handles.append(src)
   1983
```

`FileNotFoundError: [Errno 2] No such file or directory: 'spam.csv'`

In [3]:

```
df.groupby('label').describe()
```

```
-----
NameError                                Traceback (most recent call last)
<ipython-input-3-a32778b9f09d> in <module>
----> 1 df.groupby('label').describe()

NameError: name 'df' is not defined
```

In [4]:

```
import string
from nltk.corpus import stopwords
from nltk import PorterStemmer as Stemmer
def process(text):
    text = text.lower()
    text = ''.join([t for t in text if t not in string.punctuation])
    text = [t for t in text.split() if t not in stopwords.words('english')]
    st = Stemmer
    text = [st.stem(t) for t in text]
    return text
```

In [5]:

```
process('it\'s holidays and we are playing cricket. jeff is playing very welllll')
```

```
-----
OSError                                Traceback (most recent call last)
<ipython-input-5-f64581f61b96> in <module>
----> 1 process('it\'s holidays and we are playing cricket. jeff is playing very welllll')

<ipython-input-4-7601d7249212> in process(text)
      5     text = text.lower()
      6     text = ''.join([t for t in text if t not in string.punctuation])
----> 7     text = [t for t in text.split() if t not in stopwords.words('english')]
      8     st = Stemmer
      9     text = [st.stem(t) for t in text]

<ipython-input-4-7601d7249212> in <listcomp>(.0)
      5     text = text.lower()
      6     text = ''.join([t for t in text if t not in string.punctuation])
----> 7     text = [t for t in text.split() if t not in stopwords.words('english')]
      8     st = Stemmer
      9     text = [st.stem(t) for t in text]

/ext/anaconda2020.02/lib/python3.7/site-packages/nltk/corpus/reader/wordlist.py in words(self, fileids, ignore_lines_startswith)
     23         return [
     24             line
----> 25             for line in line_tokenize(self.raw(fileids))
     26             if not line.startswith(ignore_lines_startswith)
     27         ]

/ext/anaconda2020.02/lib/python3.7/site-packages/nltk/corpus/reader/wordlist.py in raw(self, fileids)
     32         elif isinstance(fileids, string_types):
     33             fileids = [fileids]
----> 34         return concat([self.open(f).read() for f in fileids])
     35
     36

/ext/anaconda2020.02/lib/python3.7/site-packages/nltk/corpus/reader/wordlist.py in <listcomp>(.0)
     32         elif isinstance(fileids, string_types):
```

```

33         fileids = [fileids]
--> 34         return concat([self.open(f).read() for f in fileids])
35
36

/ext/anaconda2020.02/lib/python3.7/site-packages/nltk/corpus/reader/api.py in open(self, file)
211         """
212         encoding = self.encoding(file)
--> 213         stream = self._root.join(file).open(encoding)
214         return stream
215

/ext/anaconda2020.02/lib/python3.7/site-packages/nltk/data.py in join(self, fileid)
353     def join(self, fileid):
354         _path = os.path.join(self._path, fileid)
--> 355         return FileSystemPathPointer(_path)
356
357     def __repr__(self):

/ext/anaconda2020.02/lib/python3.7/site-packages/nltk/compat.py in _decorator(*args, **kwargs)
226     def _decorator(*args, **kwargs):
227         args = (args[0], add_py3_data(args[1])) + args[2:]
--> 228         return init_func(*args, **kwargs)
229
230     return wraps(init_func)(_decorator)

/ext/anaconda2020.02/lib/python3.7/site-packages/nltk/data.py in __init__(self, _path)
331         _path = os.path.abspath(_path)
332         if not os.path.exists(_path):
--> 333             raise IOError('No such file or directory: %r' % _path)
334         self._path = _path
335

OSError: No such file or directory: '/ext/data/nltk_data/corpora/stopwords/englis'

```

In [0]: