In [ ]:
```
import pandas as pd
import numpy as np
```

In [8]:
```
with open("conv.txt") as f:
    stored_conv = f.read()
```

In [9]:
```
lst = list(stored_conv.split('\n'))
```

In [10]:
```
lst
```

```
 ',
 'ROBERT: Ned.',
 '',
 'ARYA: Whereâ€™s the Imp?',
 '',
 'CERSEI: Where is our brother? Go find the little beast.',
 '',
 'NED: Tell me about Jon Arryn.',
 '',
 'ROBERT: One minute he was fine, and then â€¦ Burned right through him, whatever it was. I loved that ma
n.',
 '',
 'NED: We both did.',
 '',
 'ROBERT: He never had to teach you much, but me â€¦ You remember me at 16? All I wanted to do was crack sku
lls and fuck girls. He showed me what was what.',
 '',
 'NED: Aye.',
 '',
 '']
```

In [11]:
```
def remove_punc(string):
    punc = "''.,!?..."
    clean_text = ''.join(ch for ch in string if ch not in punc)
    return (clean_text)
```

In [12]:
```
lst2 = []
for i in range(len(lst)):
    lst2.append(remove_punc(lst[i]))

lst2
```

```
 '',
 'CATELYN: I want you to promise me No more climbing',
 '',
 'NED: Your Grace',
 '',
 'ROBERT: Youâ€™ve got fat',
 '',
 'ROBERT: Cat',
 '',
 'CATELYN: Your Grace',
 '',
 'ROBERT: Nine years Why havenâ€™t I seen you Where the hell have you been',
 '',
 'NED: Guarding the North for you Your Grace Winterfell is yours',
 '',
 'ARYA: Whereâ€™s the Imp',
 '',
 'SANSA: Will you shut up',
 '',
 'ROBERT: Who have we here You must be Robb'.
```

In [16]:
```
lst_names_of_character = []
lst_dialogues_WILL = []
lst_dialogues_WAYMAR = []
lst_dialogues_GARED = []
lst_dialogues_ROYCE = []
lst_dialogues_JON = []
lst_dialogues_SEPTA = []
lst_dialogues_SANSA = []
lst_dialogues_NED = []
lst_dialogues_ROBB = []
lst_dialogues_CASSEL = []
lst_dialogues_CATELYN = []
lst_dialogues_BRAN = []
lst_dialogues_THEON = []
lst_dialogues_CERSEI = []
lst_dialogues_JAIME = []
lst_dialogues_ROBERT = []
lst_dialogues_ARYA = []
```

In [15]:
```
for i  in range(len(lst2)):
    if lst2[i] == '':
        continue
    else:
        lst_names_of_character.append(lst2[i].split(':')[0])
        if lst_names_of_character[-1] == 'WILL':
            for word in lst2[i].split(':')[1].split():
                if word not in lst_dialogues_WILL:
                    lst_dialogues_WILL.append(word)
            with open('WILL' , mode = 'w') as f:
                for word in lst_dialogues_WILL:
                    f.write(word)
                    f.write('\n')

        elif lst_names_of_character[-1] == 'WAYMAR ROYCE':
            for word in lst2[i].split(':')[1].split():
                if word not in lst_dialogues_WAYMAR:
                    lst_dialogues_WAYMAR.append(word)
            with open('WAYMAR ROYCE' , mode = 'w') as f:
                for word in lst_dialogues_WAYMAR:
                    f.write(word)
                    f.write('\n')

        elif lst_names_of_character[-1] == 'GARED':
            for word in lst2[i].split(':')[1].split():
                if word not in lst_dialogues_GARED:
                    lst_dialogues_GARED.append(word)
            with open('GARED' , mode = 'w') as f:
                for word in lst_dialogues_GARED:
                    f.write(word)
                    f.write('\n')

        elif lst_names_of_character[-1] == 'ROYCE':
            for word in lst2[i].split(':')[1].split():
                if word not in lst_dialogues_ROYCE:
                    lst_dialogues_ROYCE.append(word)
            with open('ROYCE' , mode = 'w') as f:
                for word in lst_dialogues_ROYCE:
                    f.write(word)
                    f.write('\n')

        elif lst_names_of_character[-1] == 'JON':
```

```
            for word in lst2[i].split(':')[1].split():
                if word not in lst_dialogues_JON:
                    lst_dialogues_JON.append(word)
            with open('JON' , mode = 'w') as f:
                for word in lst_dialogues_JON:
                    f.write(word)
                    f.write('\n')

        elif lst_names_of_character[-1] == 'SEPTA MORDANE':
            for word in lst2[i].split(':')[1].split():
                if word not in lst_dialogues_SEPTA:
                    lst_dialogues_SEPTA.append(word)
            with open('SEPTA MORDANE' , mode = 'w') as f:
                for word in lst_dialogues_SEPTA:
                    f.write(word)
                    f.write('\n')

        elif lst_names_of_character[-1] == 'SANSA':
            for word in lst2[i].split(':')[1].split():
                if word not in lst_dialogues_SANSA:
                    lst_dialogues_SANSA.append(word)
            with open('SANSA' , mode = 'w') as f:
                for word in lst_dialogues_SANSA:
                    f.write(word)
                    f.write('\n')

        elif lst_names_of_character[-1] == 'NED':
            for word in lst2[i].split(':')[1].split():
                if word not in lst_dialogues_NED:
                    lst_dialogues_NED.append(word)
            with open('NED' , mode = 'w') as f:
                for word in lst_dialogues_NED:
                    f.write(word)
                    f.write('\n')

        elif lst_names_of_character[-1] == 'ROBB':
            for word in lst2[i].split(':')[1].split():
                if word not in lst_dialogues_ROBB:
                    lst_dialogues_ROBB.append(word)
            with open('ROBB' , mode = 'w') as f:
                for word in lst_dialogues_ROBB:
                    f.write(word)
                    f.write('\n')
```

```
        elif lst_names_of_character[-1] == 'CASSEL':
            for word in lst2[i].split(':')[1].split():
                if word not in lst_dialogues_CASSEL:
                    lst_dialogues_CASSEL.append(word)
            with open('CASSEL' , mode = 'w') as f:
                for word in lst_dialogues_CASSEL:
                    f.write(word)
                    f.write('\n')

        elif lst_names_of_character[-1] == 'CATELYN':
            for word in lst2[i].split(':')[1].split():
                if word not in lst_dialogues_CATELYN:
                    lst_dialogues_CATELYN.append(word)
            with open('CATELYN' , mode = 'w') as f:
                for word in lst_dialogues_CATELYN:
                    f.write(word)
                    f.write('\n')

        elif lst_names_of_character[-1] == 'BRAN':
            for word in lst2[i].split(':')[1].split():
                if word not in lst_dialogues_BRAN:
                    lst_dialogues_BRAN.append(word)
            with open('BRAN' , mode = 'w') as f:
                for word in lst_dialogues_BRAN:
                    f.write(word)
                    f.write('\n')

        elif lst_names_of_character[-1] == 'THEON':
            for word in lst2[i].split(':')[1].split():
                if word not in lst_dialogues_THEON:
                    lst_dialogues_THEON.append(word)
            with open('THEON' , mode = 'w') as f:
                for word in lst_dialogues_THEON:
                    f.write(word)
                    f.write('\n')

        elif lst_names_of_character[-1] == 'CERSI':
            for word in lst2[i].split(':')[1].split():
                if word not in lst_dialogues_CERSI:
                    lst_dialogues_CERSI.append(word)
            with open('CERSI' , mode = 'w') as f:
                for word in lst_dialogues_CERSI:
```

```
                f.write(word)
                f.write('\n')

        elif lst_names_of_character[-1] == 'JAIME':
            for word in lst2[i].split(':')[1].split():
                if word not in lst_dialogues_JAIME:
                    lst_dialogues_JAIME.append(word)
            with open('JAIME' , mode = 'w') as f:
                for word in lst_dialogues_JAIME:
                    f.write(word)
                    f.write('\n')

        elif lst_names_of_character[-1] == 'ROBERT':
            for word in lst2[i].split(':')[1].split():
                if word not in lst_dialogues_ROBERT:
                    lst_dialogues_ROBERT.append(word)
            with open('ROBERT' , mode = 'w') as f:
                for word in lst_dialogues_ROBERT:
                    f.write(word)
                    f.write('\n')

        elif lst_names_of_character[-1] == 'ARYA':
            for word in lst2[i].split(':')[1].split():
                if word not in lst_dialogues_ARYA:
                    lst_dialogues_ARYA.append(word)
            with open('ARYA' , mode = 'w') as f:
                for word in lst_dialogues_ARYA:
                    f.write(word)
                    f.write('\n')

        else:
            continue
```

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]: