

# FILEHANDING\_STRINGMANUPILATION-copy1

July 7, 2022

## 1 IMPORTING PANDAS

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

## 2 READING CONV.TXT FILE

```
[66]: data=pd.read_csv(r"C:\\Users\\NEKEAPOLLO\\scripts\\conv.
↳txt",delimiter='\\t',header=None)
```

```
[67]: data.rename(columns={0:'Dailouges'},inplace=True)
```

```
[68]: data
```

```
[68]:
```

|     | Dailouges   |
|-----|---|
| 0   | WILL: I've never seen wildlings do a thing lik...   |
| 1   | WAYMAR ROYCE: How close did you get?                |
| 2   | WILL: Close as any man would.                       |
| 3   | GARED: We should head back to the wall.             |
| 4   | ROYCE: Do the dead frighten you?                    |
| ... | ...   |
| 106 | NED: Tell me about Jon Arryn.                       |
| 107 | ROBERT: One minute he was fine, and then ... Bur... |
| 108 | NED: We both did.                                   |
| 109 | ROBERT: He never had to teach you much, but me...   |
| 110 | NED: Aye.   |

[111 rows x 1 columns]

## 3 PREPROCESSING DATA

```
[69]: new_data=data['Dailouges'].str.split(':',expand=True,)
```

```
[70]: new_data
```

*#SEPRATING CHARACTERS*  
*↳AND THEIR DAILOUGES*

```
[70]:
```

|     | 0            | 1  |
|-----|--------------|--|
| 0   | WILL         | I've never seen wildlings do a thing like thi...   |
| 1   | WAYMAR ROYCE | How close did you get?                             |
| 2   | WILL         | Close as any man would.                            |
| 3   | GARED        | We should head back to the wall.                   |
| 4   | ROYCE        | Do the dead frighten you?                          |
| ..  | ...          | ...  |
| 106 | NED          | Tell me about Jon Arryn.                           |
| 107 | ROBERT       | One minute he was fine, and then ... Burned rig... |
| 108 | NED          | We both did.                                       |
| 109 | ROBERT       | He never had to teach you much, but me ... You ... |
| 110 | NED          | Aye.   |

[111 rows x 2 columns]

```
[71]: new_data.rename(columns={0:'Characters'},inplace=True)
new_data.rename(columns={1:'Dailouges'},inplace=True)
```

```
[72]: new_data
```

```
[72]:
```

|     | Characters   | Dailouges  |
|-----|--------------|--|
| 0   | WILL         | I've never seen wildlings do a thing like thi...   |
| 1   | WAYMAR ROYCE | How close did you get?                             |
| 2   | WILL         | Close as any man would.                            |
| 3   | GARED        | We should head back to the wall.                   |
| 4   | ROYCE        | Do the dead frighten you?                          |
| ..  | ...          | ...  |
| 106 | NED          | Tell me about Jon Arryn.                           |
| 107 | ROBERT       | One minute he was fine, and then ... Burned rig... |
| 108 | NED          | We both did.                                       |
| 109 | ROBERT       | He never had to teach you much, but me ... You ... |
| 110 | NED          | Aye.   |

[111 rows x 2 columns]

```
[73]: character_count=new_data.groupby('Characters').count()
↳#GROUPING CHARACTERS
character_count
```

```
[73]:
```

|            | Dailouges |
|------------|-----------|
| Characters |           |
| ARYA       | 4         |
| BRAN       | 10        |
| CASSEL     | 3         |
| CATELYN    | 10        |
| CERSEI     | 5         |
| GARED      | 3         |

|               |    |
|---------------|----|
| JAIME         | 3  |
| JON           | 10 |
| NED           | 27 |
| ROBB          | 7  |
| ROBERT        | 9  |
| ROYCE         | 5  |
| SANSA         | 3  |
| SEPTA MORDANE | 2  |
| THEON         | 5  |
| WAYMAR ROYCE  | 1  |
| WILL          | 4  |

```
[74]: character_count.shape #THIS TELLS US THAT THERE ARE 17
      ↪ UNIQUE CHARACTERS
```

```
[74]: (17, 1)
```

```
[75]: character_groups=new_data.groupby("Characters")
      character_groups.first()
```

```
[75]:                                     Dailouges
Characters
ARYA                                     Where's the Imp?
BRAN                                     Jon said he was a deserter.
CASSEL      Lord Stark. My lady. A guardsman just rode in...
CATELYN                                     Do you have to?
CERSEI                                     What if Jon Arryn told someone?
GARED                                     We should head back to the wall.
JAIME                                     But who would he tell?
JON                                     Go on. Father's watching.
NED      And which one of you was a marksman at ten? K...
ROBB                                     Relax your bow arm.
ROBERT                                     You've got fat.
ROYCE                                     Do the dead frighten you?
SANSA                                     Thank you.
SEPTA MORDANE      Fine work, as always. Well done.
THEON                                     Mountain lion?
WAYMAR ROYCE                                     How close did you get?
WILL      I've never seen wildlings do a thing like thi...
```

## 4 ARYA

```
[87]: C1=character_groups.get_group('ARYA')['Dailouges']
      C1 #STORING DAILOUGES
      ↪ IN A NEW DATAFRAME
```

```
[87]: 92                Where's the Imp?
      95                Arya.
      97    That's Jaime Lannister. The queen's twin brot...
      104                Where's the Imp?
      Name: Dailouges, dtype: object
```

```
[88]: C1.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\C1.txt',index=False,header=False)
      ↪#SAVING FILE AS CSV
```

```
[89]: #CLEANING DATA

      text_file=open(r'C:\\Users\\NEKEAPOLLO\\scripts\\C1.txt')
      text=text_file.read()
      text=text.lower()                                #CONVERTING TO
      ↪LOWER CASES
      words=text.split()                                #SPLITTING WORDS
      ↪SEPRATED BY WHITE SPACE
      words=[word.strip('.,!;() []') for word in words]
      words=[word.replace('â€', '') for word in words]    #UTF ENCODING
      unique=[]
      for word in words:
          if word not in unique:
              unique.append(word)
      unique.sort()
      unique
```

```
[89]: ['arya',
      'brother',
      'imp?',
      'jaime',
      'lannister',
      'queens',
      'thats',
      'the',
      'twin',
      'wheres']
```

```
[85]: ARYA1=unique                                #SAVING THE SERIES
      ARYA2=pd.DataFrame(ARYA1)                    #CONVERTING SERIES TO
      ↪DATAFRAME
```

```
[86]: ARYA2.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\ARYA.txt',index=False,
      ↪header=False) #SAVING FINAL FILE
```

## 5 BRAN

```
[91]: C2=character_groups.get_group('BRAN')['Dailouges']  
C2
```

```
[91]: 35                Jon said he was a deserter.  
37                Our way is the old way?  
39                Is it true he saw the White Walkers?  
41                So he was lying?  
52                Where will they go? Their mother's dead.  
56                NO!  
59                Please, father!  
63                What about you?  
82                I saw the king! He's got hundreds of people!  
84                But he's coming right now! Down our road!  
Name: Dailouges, dtype: object
```

```
[98]: C2.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\C2.txt',index=False ,header=False)
```

```
[101]: text_file=open(r'C:\\Users\\NEKEAPOLLO\\scripts\\C2.txt')  
text=text_file.read()  
text=text.lower()  
words=text.split()  
words=[word.strip('.,!;() []') for word in words]  
words=[word.replace('â€', '') for word in words]  
unique=[]  
for word in words:  
    if word not in unique:  
        unique.append(word)  
unique.sort()  
unique
```

```
[101]: [' ',  
'a',  
'about',  
'but',  
'coming',  
'dead',  
'deserter',  
'down',  
'father! "',  
'go?',  
'got',  
'he',  
'hes',  
'hundreds',  
'i',
```

```

'is',
'it',
'jon',
'king',
'lying?',
'mothers',
'no',
'now',
'of',
'old',
'our',
'people',
'please',
'right',
'road',
'said',
'saw',
'so',
'the',
'their',
'they',
'true',
'walkers?',
'was',
'way',
'way?',
'what',
'where',
'white',
'will',
'you?']

```

```

[102]: BRAN=unique
      BRAN1=pd.DataFrame(BRAN)

```

```

[104]: BRAN1.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\BRAN.txt',index=False,header=False)

```

## 6 CASSEL

```

[105]: C3=character_groups.get_group('CASSEL')['Dailouges']
      C3

```

```

[105]: 21      Lord Stark. My lady. A guardsman just rode in...
      25                                The law is law, my lady.
      53                                They don't belong down here.
      Name: Dailouges, dtype: object

```

```
[106]: C3.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\C3.txt',index=False ,header=False)
```

```
[110]: text_file=open(r'C:\\Users\\NEKEAPOLLO\\scripts\\C3.txt')
        text=text_file.read()
        text=text.lower()
        words=text.split()
        words=[word.strip('.,!;() []') for word in words]
        words=[word.replace('â€', '') for word in words]
        unique=[]
        for word in words:
            if word not in unique:
                unique.append(word)
        unique.sort()
        unique
```

```
[110]: ['"',
        'a',
        'belong',
        'captured',
        'deserter',
        'dont',
        'down',
        'from',
        'guardsman',
        'here',
        'hills',
        'in',
        'is',
        'just',
        'lady',
        'lady."',
        'law',
        'lord',
        'my',
        'nights',
        'rode',
        'stark',
        'the',
        'they',
        'theyve',
        'watch']
```

```
[111]: CASSEL=unique
        CASSEL1=pd.DataFrame(CASSEL)
```

```
[115]: CASSEL1.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\CASSEL.
        ↪txt',header=False,index=False)
```

## 7 CATELYN

```
[116]: C4=character_groups.get_group('CATELYN')['Dailouges']  
C4
```

```
[116]: 23                               Do you have to?  
27               Ned. Ten is too young to see such things.  
72      All these years and I still feel like an outs...  
74               I wonder if the old gods agree.  
76               I am so sorry, my love.  
81               Gods, but they grow fast.  
83      How many times have I told you No climbing!  
85      I want you to promise me. No more climbing.  
89               Your Grace.  
100              My queen.  
Name: Dailouges, dtype: object
```

```
[117]: C4.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\C4.txt',index=False ,header=False)
```

```
[118]: text_file=open(r'C:\\Users\\NEKEAPOLLO\\scripts\\C4.txt')  
text=text_file.read()  
text=text.lower()  
words=text.split()  
words=[word.strip('.,!;() []') for word in words]  
words=[word.replace('â€', '') for word in words]  
unique=[]  
for word in words:  
    if word not in unique:  
        unique.append(word)  
unique.sort()  
unique
```

```
[118]: ['',  
'agree',  
'all',  
'am',  
'an',  
'and',  
'but',  
'climbing',  
'come',  
'do',  
'fast.',  
'feel',  
'gods',  
'grace',  
'grow',
```



```
'have',  
'here',  
'how',  
'i',  
'if',  
'is',  
'like',  
'love."',  
'many',  
'me',  
'more',  
'my',  
'ned',  
'no',  
'old',  
'outsider',  
'promise',  
'queen',  
'see',  
'so',  
'sorry',  
'still',  
'such',  
'ten',  
'the',  
'these',  
'they',  
'things',  
'times',  
'to',  
'to?',  
'told',  
'too',  
'want',  
'when',  
'wonder',  
'years',  
'you',  
'young',  
'your']
```

```
[120]: CATELYN=unique  
CATELYN2=pd.DataFrame(CATELYN)
```

```
[121]: CATELYN2.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\CATELYN.  
↪txt',index=False,header=False)
```

## 8 CERSEI

```
[122]: C5=character_groups.get_group('CERSEI')['Dailouges']
C5
```

```
[122]: 66           What if Jon Arryn told someone?
        68           My husband.
        70           You should be the Hand of the king.
        102      We've been riding for a month, my love. Surel...
        105      Where is our brother? Go find the little beast.
Name: Dailouges, dtype: object
```

```
[123]: C5.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\C5.txt',index=False ,header=False)
```

```
[124]: text_file=open(r'C:\\Users\\NEKEAPOLLO\\scripts\\C5.txt')
        text=text_file.read()
        text=text.lower()
        words=text.split()
        words=[word.strip('.,!;() []') for word in words]
        words=[word.replace('â€', '') for word in words]
        unique=[]
        for word in words:
            if word not in unique:
                unique.append(word)
        unique.sort()
        unique
```

```
[124]: ['"',
        'a',
        'arryn',
        'be',
        'beast',
        'been',
        'brother?',
        'can',
        'dead',
        'find',
        'for',
        'go',
        'hand',
        'husband',
        'if',
        'is',
        'jon',
        'king',
        'little',
        'love',
```

```

'month',
'my',
'of',
'our',
'riding',
'should',
'someone?',
'surely',
'the',
'told',
'wait."',
've',
'what',
'where',
'you']

```

```

[125]: CERSEI=unique
      CERSEI2=pd.DataFrame(CERSEI)

```

```

[139]: CERSEI2.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\CERSEI.
      ↪txt',index=False,header=False)

```

## 9 GARED

```

[128]: C6=character_groups.get_group('GARED')['Dailouges']
      C6

```

```

[128]: 3                We should head back to the wall.
      5      Our orders were to track the wildlings. We tr...
      11                See where they went.
      Name: Dailouges, dtype: object

```

```

[130]: C6.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\C6.txt',index=False ,header=False)

```

```

[131]: text_file=open(r'C:\\Users\\NEKEAPOLLO\\scripts\\C6.txt')
      text=text_file.read()
      text=text.lower()
      words=text.split()
      words=[word.strip('.,!;() []') for word in words]
      words=[word.replace('â€', '') for word in words]
      unique=[]
      for word in words:
          if word not in unique:
              unique.append(word)
      unique.sort()
      unique

```

```
[131]: ['back',
        'head',
        'more',
        'no',
        'orders',
        'our',
        'see',
        'should',
        'the',
        'them',
        'they',
        'to',
        'track',
        'tracked',
        'trouble',
        'us',
        'wall',
        'we',
        'went',
        'were',
        'where',
        'wildlings',
        'wont']
```

```
[132]: GARED=unique
        GARED2=pd.DataFrame(GARED)
```

```
[138]: GARED2.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\GARED.
        ↪txt',index=False,header=False)
```

## 10 JAIME

```
[135]: C7=character_groups.get_group('JAIME')['Dailouges']
        C7
```

```
[135]: 67                                     But who would he tell?
        69      If he told the king, both our heads would be ...
        71      That's an honor I can do without. Their days ...
        Name: Dailouges, dtype: object
```

```
[136]: C7.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\C7.txt',index=False ,header=False)
```

```
[137]: text_file=open(r'C:\\Users\\NEKEAPOLLO\\scripts\\C7.txt')
        text=text_file.read()
        text=text.lower()
        words=text.split()
```

```

words=[word.strip('.,!;() []') for word in words]
words=[word.replace('â€', '') for word in words]
unique=[]
for word in words:
    if word not in unique:
        unique.append(word)
unique.sort()
unique

```

```

[137]: ['',
'a',
'an',
'and',
'are',
'around?',
'arryn',
'be',
'boars',
'both',
'but',
'by',
'can',
'choose',
'city',
'days',
'didnt',
'died',
'do',
'fucking',
'gates',
'go',
'hand',
'he',
'heads',
'hes',
'him',
'his',
'honor',
'hunting',
'i',
'if',
'is',
'it',
'job',
'jon',
'king',
'knew',

```

```
'know',
'life',
'lives',
'long',
'new',
'now',
'of',
'off',
'on',
'on."',
'or',
'other',
'our',
'robert',
'short."',
'skewered',
'someone',
'tell?',
'thats',
'the',
'their',
'to',
'told',
'too',
'way',
'whatever',
'while',
'who',
'whores',
'will',
'with',
'without',
'would']
```

```
[140]: JAMIE=unique
      JAMIE2=pd.DataFrame(JAMIE)
```

```
[142]: JAMIE2.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\JAMIE.
      ↪txt',index=False,header=False)
```

## 11 JON

```
[144]: C8=character_groups.get_group('JON')['Dailouges']
      C8
```

```
[144]: 13             Go on. Father's watching.
        14             And your mother.
        19             Don't think too much, Bran.
        31             Father will know if you do.
        33             You did well.
        43             What is it?
        50             Now there are five.
        51             You want to hold it?
        61 Lord Stark? There are five pups. One for each...
        64             I'm not a Stark. Get on.
Name: Dailouges, dtype: object
```

```
[145]: C8.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\C8.txt',index=False,header=False)
```

```
[146]: text_file=open(r'C:\\Users\\NEKEAPOLLO\\scripts\\C8.txt')
        text=text_file.read()
        text=text.lower()
        words=text.split()
        words=[word.strip('.,!;() []') for word in words]
        words=[word.replace('â€', '') for word in words]
        unique=[]
        for word in words:
            if word not in unique:
                unique.append(word)
        unique.sort()
        unique
```

```
[146]: ['"',
        'a',
        'and',
        'are',
        'bran."',
        'children',
        'did',
        'direwolf',
        'do',
        'dont',
        'each',
        'father',
        'fathers',
        'five',
        'for',
        'get',
        'go',
        'have',
        'hold',
        'house',
```

```

'if',
'im',
'is',
'it?',
'know',
'lord',
'meant',
'mother',
'much',
'not',
'now',
'of',
'on',
'one',
'pups',
'sigil',
'stark',
'stark?',
'the',
'them',
'there',
'they',
'think',
'to',
'too',
'want',
'watching',
'well',
'were',
'what',
'will',
'you',
'your']

```

```

[147]: JON=unique
      JON2=pd.DataFrame(JON)

```

```

[149]: JON2.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\JON.txt',index=False,header=False)

```

## 12 NED

```

[150]: C9=character_groups.get_group('NED')['Dailouges']
      C9

```

```

[150]: 18      And which one of you was a marksman at ten? K...
      22      Get the lads to saddle their horses.

```



```

24             He swore an oath, Cat.
26             Tell Bran he's coming, too.
28     He won't be a boy forever. And winter is coming.
30             King of the Andals and the First Men ...
32     Lord of the Seven Kingdoms and protector of t...
34             You understand why I did it?
36             But do you understand why I had to kill him?
38     The man who passes the sentence should swing ...
40     The White Walkers have been gone for thousand...
42             A madman sees what he sees.
45             There are no mountain lions in these woods.
47             It's a direwolf.
48             Tough old beast.
54     Better a quick death. They won't last without...
60             I'm sorry, Bran.
62     You will train them yourselves. You will feed...
73     You have five northern children. You're not a...
75             It's your gods with all the rules.
77             Tell me.
86             Your Grace.
91     Guarding the North for you, Your Grace. Winte...
99             My queen.
106            Tell me about Jon Arryn.
108            We both did.
110            Aye.
Name: Dailouges, dtype: object

```

```
[151]: C9.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\C9.txt',index=False ,header=False)
```

```
[152]: text_file=open(r'C:\\Users\\NEKEAPOLLO\\scripts\\C9.txt')
        text=text_file.read()
        text=text.lower()
        words=text.split()
        words=[word.strip('.,!;() []') for word in words]
        words=[word.replace('â€', '') for word in words]
        unique=[]
        for word in words:
            if word not in unique:
                unique.append(word)
        unique.sort()
        unique
```

```
[152]: ['"',
        'a',
        'about',
        'all',
        'an',
```

'and',  
'andals',  
'are',  
'arryn',  
'at',  
'aye',  
'be',  
'beast',  
'been',  
'better',  
'both',  
'boy',  
'bran',  
'bran."',  
'bury',  
'but',  
'cat."',  
'children',  
'coming',  
'death',  
'did',  
'die',  
'die."',  
'direwolf',  
'do',  
'eddard',  
'feed',  
'first',  
'five',  
'for',  
'forever',  
'get',  
'go',  
'gods',  
'gone',  
'grace',  
'guarding',  
'had',  
'have',  
'he',  
'hes',  
'him?',  
'horses',  
'house',  
'i',  
'if',  
'im',

'in',  
'is',  
'it?',  
'its',  
'jon',  
'keep',  
'kill',  
'king',  
'kingdoms',  
'lads',  
'last',  
'lions',  
'lord',  
'madman',  
'man',  
'marksman',  
'me',  
'men',  
'mother',  
'mountain',  
'my',  
'no',  
'north',  
'northern',  
'not',  
'oath',  
'of',  
'old',  
'on."',  
'one',  
'outsider',  
'passes',  
'practicing',  
'protector',  
'queen',  
'quick',  
'realm',  
'rules',  
'saddle',  
'sees',  
'sentence',  
'seven',  
'should',  
'sorry',  
'stark',  
'swing',  
'sword',

```
'swore',
'tell',
'ten?',
'the',
'their',
'them',
'there',
'these',
'they',
'thousands',
'to',
'too."',
'tough',
'train',
'understand',
'walkers',
'warden',
'was',
'we',
'what',
'which',
'white',
'who',
'why',
'will',
'winter',
'winterfell',
'with',
'without',
'wont',
'woods',
'years',
'you',
'your',
'youre',
'yours."',
'yourselves',
'yourselves."',
'â€!']
```

```
[153]: NED=unique
      NED2=pd.DataFrame(NED)
```

```
[154]: NED2.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\NED.txt',index=False ,header=False)
```

## 13 ROBB

```
[155]: C10=character_groups.get_group('ROBB')['Dailouges']  
C10
```

```
[155]: 20                                Relax your bow arm.  
29                                Lad, go run back and get the rest.  
49                There are no direwolves south of the Wall.  
57                                Put away your blade.  
65                                What is it?  
78                I hear the prince is a right royal prick.  
80    Go on, Tommy, shave him good. He's never met ...  
Name: Dailouges, dtype: object
```

```
[156]: C10.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\C10.txt',index=False ,header=False)
```

```
[157]: text_file=open(r'C:\\Users\\NEKEAPOLLO\\scripts\\C10.txt')  
text=text_file.read()  
text=text.lower()  
words=text.split()  
words=[word.strip('.,!;() []') for word in words]  
words=[word.replace('â€', '') for word in words]  
unique=[]  
for word in words:  
    if word not in unique:  
        unique.append(word)  
unique.sort()  
unique
```

```
[157]: ['"',  
'a',  
'and',  
'are',  
'arm',  
'away',  
'back',  
'better',  
'blade',  
'bow',  
'direwolves',  
'get',  
'girl',  
'go',  
'good',  
'hair."',  
'he',  
'hear',
```

```
'hes',
'him',
'his',
'i',
'is',
'it?',
'lad',
'likes',
'met',
'never',
'no',
'of',
'on',
'own',
'prick',
'prince',
'put',
'relax',
'rest."',
'right',
'royal',
'run',
'shave',
'south',
'than',
'the',
'there',
'tommy',
'wall',
'what',
'your']
```

```
[158]: ROBB=unique
      ROBB2=pd.DataFrame(ROBB)
```

```
[159]: ROBB2.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\ROBB.txt',index=False,header=False)
```

## 14 ROBERT

```
[162]: C11=character_groups.get_group('ROBERT')['Dailouges']
      C11
```

```
[162]: 87                You've got fat.
      88                Cat!
      90      Nine years. Why haven't I seen you? Where the...
      94                Who have we here? You must be Robb.
```

```

96      Ooh. Show us your muscles. You'll be a soldier.
101     Take me to your crypt. I want to pay my respe...
103                                     Ned.
107     One minute he was fine, and then ... Burned rig...
109     He never had to teach you much, but me ... You ...
Name: Dailouges, dtype: object

```

```
[163]: C11.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\C11.txt',index=False ,header=False)
```

```
[164]: text_file=open(r'C:\\Users\\NEKEAPOLLO\\scripts\\C11.txt')
        text=text_file.read()
        text=text.lower()
        words=text.split()
        words=[word.strip('.,!;() []') for word in words]
        words=[word.replace('â€ ', '') for word in words]
        unique=[]
        for word in words:
            if word not in unique:
                unique.append(word)
        unique.sort()
        unique
```

```
[164]: [' ',
        '16?',
        'a',
        'all',
        'and',
        'at',
        'be',
        'been?',
        'burned',
        'but',
        'cat',
        'crack',
        'crypt',
        'do',
        'fat',
        'fine',
        'fuck',
        'girls',
        'got',
        'had',
        'have',
        'havent',
        'he',
        'hell',
        'here?']
```

'him',  
'i',  
'it',  
'loved',  
'man."',  
'me',  
'minute',  
'much',  
'muscles',  
'must',  
'my',  
'ned',  
'never',  
'nine',  
'one',  
'ooh',  
'pay',  
'remember',  
'respects',  
'right',  
'robb',  
'seen',  
'show',  
'showed',  
'skulls',  
'soldier',  
'take',  
'teach',  
'that',  
'the',  
'then',  
'through',  
'to',  
'us',  
'want',  
'wanted',  
'was',  
'we',  
'what',  
'what."',  
'whatever',  
'where',  
'who',  
'why',  
'years',  
'you',  
'you?'



```
'youll',
'your',
'youve',
'â€¦']
```

```
[165]: ROBERT=unique
ROBERT2=pd.DataFrame(ROBERT)
```

```
[166]: ROBERT2.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\ROBERT.txt',index=False,
↪,header=False)
```

## 15 ROYCE

```
[167]: C12=character_groups.get_group('ROYCE')['Dailouges']
C12
```

```
[167]: 4                      Do the dead frighten you?
6      You don't think he'll ask us how they died? G...
8      It's a good thing we're not children. You wan...
9              Your dead men seem to have moved camp.
12                      What is it?
Name: Dailouges, dtype: object
```

```
[168]: C12.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\C12.txt',index=False ,header=False)
```

```
[169]: text_file=open(r'C:\\Users\\NEKEAPOLLO\\scripts\\C12.txt')
text=text_file.read()
text=text.lower()
words=text.split()
words=[word.strip('.,!;() []') for word in words]
words=[word.replace('â€¦','') for word in words]
unique=[]
for word in words:
    if word not in unique:
        unique.append(word)
unique.sort()
unique
```

```
[169]: ['',
'a',
'again."',
'as',
'ask',
'away',
'back',
'behead',
```

'camp',  
'catch',  
'children',  
'course',  
'dead',  
'deserter',  
'died?',  
'do',  
'dont',  
'first',  
'frighten',  
'get',  
'good',  
'have',  
'hell',  
'horse',  
'how',  
'i',  
'if',  
'is',  
'it',  
'it?',  
'its',  
'men',  
'moved',  
'not',  
'of',  
'on',  
'run',  
'say',  
'seem',  
'south',  
'the',  
'they',  
'thing',  
'think',  
'to',  
'us',  
'want',  
'were',  
'what',  
'will',  
'wont',  
'you',  
'you?',  
'your',  
'â€|']

```
[170]: ROYCE=unique
ROYCE2=pd.DataFrame(ROYCE)
```

```
[172]: ROYCE2.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\ROYCE.txt',index=False,
↵,header=False)
```

## 16 SANSa

```
[173]: C13=character_groups.get_group('SANSa')['Dailouges']
C13
```

```
[173]: 16                Thank you.
93                Will you shut up?
98    Would you please shut up.
Name: Dailouges, dtype: object
```

```
[174]: C13.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\C13.txt',index=False ,header=False)
```

```
[175]: text_file=open(r'C:\\Users\\NEKEAPOLLO\\scripts\\C13.txt')
text=text_file.read()
text=text.lower()
words=text.split()
words=[word.strip('.,!;() []') for word in words]
words=[word.replace('â€', '') for word in words]
unique=[]
for word in words:
    if word not in unique:
        unique.append(word)
unique.sort()
unique
```

```
[175]: ['please', 'shut', 'thank', 'up', 'up?', 'will', 'would', 'you']
```

```
[176]: SANSa=unique
SANSa2=pd.DataFrame(SANSa)
```

```
[178]: SANSa2.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\SANSa.txt',index=False,
↵,header=False)
```

## 17 SEPTA MORDANE

```
[179]: C14=character_groups.get_group('SEPTA MORDANE')['Dailouges']
C14
```

```
[179]: 15             Fine work, as always. Well done.  
      17     I love the detail that you've managed to get ...  
      Name: Dailouges, dtype: object
```

```
[186]: C14.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\C14.txt',index=False ,header=False)
```

```
[187]: text_file=open(r'C:\\Users\\NEKEAPOLLO\\scripts\\C14.txt')  
      text=text_file.read()  
      text=text.lower()  
      words=text.split()  
      words=[word.strip('.,!;() []') for word in words]  
      words=[word.replace('â€ ', '') for word in words]  
      unique=[]  
      for word in words:  
          if word not in unique:  
              unique.append(word)  
      unique.sort()  
      unique
```

```
[187]: ['"',  
      'always',  
      'as',  
      'corners',  
      'detail',  
      'done."',  
      'fine',  
      'get',  
      'i',  
      'in',  
      'love',  
      'managed',  
      'that',  
      'the',  
      'this',  
      'to',  
      'well',  
      'work',  
      'youve']
```

```
[188]: SEPTA_MORDANE=unique  
      SEPTA_MORDANE2=pd.DataFrame(SEPTA_MORDANE)
```

```
[189]: SEPTA_MORDANE2.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\SEPTA_MORDANE.  
      ↪txt',index=False ,header=False)
```

## 18 THEON

```
[190]: C15=character_groups.get_group('THEON')['Dailouges']  
C15
```

```
[190]: 44             Mountain lion?  
46             It's a freak.  
55             Right. Give it here.  
58             I take orders from your father, not you.  
79     Think of all those southern girls he gets to ...  
Name: Dailouges, dtype: object
```

```
[191]: C15.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\C15.txt',index=False,header=False)
```

```
[192]: text_file=open(r'C:\\Users\\NEKEAPOLLO\\scripts\\C15.txt')  
text=text_file.read()  
text=text.lower()  
words=text.split()  
words=[word.strip('.,!;() []') for word in words]  
words=[word.replace('â€', '') for word in words]  
unique=[]  
for word in words:  
    if word not in unique:  
        unique.append(word)  
unique.sort()  
unique
```

```
[192]: ['''',  
'a',  
'all',  
'father',  
'freak',  
'from',  
'gets',  
'girls',  
'give',  
'he',  
'here',  
'his',  
'i',  
'it',  
'its',  
'lion?',  
'mountain',  
'not',  
'of',  
'orders',
```

```
'prick',
'right',
'royal',
'southern',
'stab',
'take',
'think',
'those',
'to',
'with',
'you."',
'your']
```

```
[193]: THEON=unique
      THEON2=pd.DataFrame(THEON)
```

```
[194]: THEON2.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\THEON.txt',index=False,
      ↪,header=False)
```

## 19 WAYMAR ROYCE

```
[195]: C16=character_groups.get_group('WAYMAR ROYCE')['Dailouges']
      C16
```

```
[195]: 1      How close did you get?
      Name: Dailouges, dtype: object
```

```
[196]: C16.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\C16.txt',index=False,header=False)
```

```
[197]: text_file=open(r'C:\\Users\\NEKEAPOLLO\\scripts\\C16.txt')
      text=text_file.read()
      text=text.lower()
      words=text.split()
      words=[word.strip('.,!;() []') for word in words]
      words=[word.replace('â€ ','') for word in words]
      unique=[]
      for word in words:
          if word not in unique:
              unique.append(word)
      unique.sort()
      unique
```

```
[197]: ['close', 'did', 'get?', 'how', 'you']
```

```
[198]: WAYMAR_ROYCE=unique
      WAYMAR_ROYCE2=pd.DataFrame(WAYMAR_ROYCE)
```

```
[199]: WAYMAR_ROYCE2.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\WAYMAR_ROYCE.  
      ↪txt',index=False ,header=False)
```

## 20 WILL

```
[200]: C17=character_groups.get_group('WILL')['Dailouges']  
C17
```

```
[200]: 0      I've never seen wildlings do a thing like thi...  
      2                                Close as any man would.  
      7      Whatever did it to them could do it to us. Th...  
      10                                They were here.  
      Name: Dailouges, dtype: object
```

```
[205]: C17.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\C17.txt',index=False ,header=False)
```

```
[210]: text_file=open(r'C:\\Users\\NEKEAPOLLO\\scripts\\C17.txt')  
      text=text_file.read()  
      text=text.lower()  
      words=text.split()  
      words=[word.strip('.,!;() []') for word in words]  
      words=[word.replace('â€', '') for word in words]  
      unique=[]  
      for word in words:  
          if word not in unique:  
              unique.append(word)  
      unique.sort()  
      unique
```

```
[210]: [''  
      'a',  
      'any',  
      'as',  
      'children',  
      'close',  
      'could',  
      'did',  
      'do',  
      'even',  
      'ever',  
      'here',  
      'in',  
      'it',  
      'ive',  
      'killed',  
      'life.'',
```

```
'like',  
'man',  
'my',  
'never',  
'not',  
'seen',  
'the',  
'them',  
'they',  
'thing',  
'this',  
'to',  
'us',  
'were',  
'whatever',  
'wildlings',  
'would']
```

```
[207]: WILL=unique  
      WILL2=pd.DataFrame(WILL)
```

```
[208]: WILL2.to_csv(r'C:\\Users\\NEKEAPOLLO\\scripts\\WILL.txt',index=False ,header=False)
```