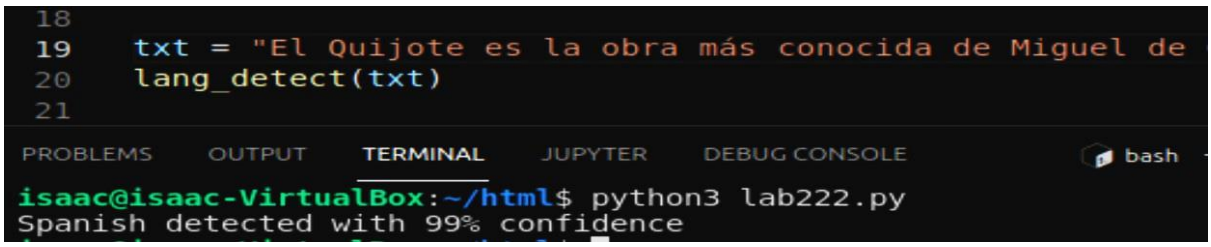
#------------------
Lab 9
Isaac Huang
23019722
#------------------

1. DOMINANT LANGUAGE DETECTION

```
import boto3
client = boto3.client('comprehend')

my_dict = {'en': 'English', 'es': 'Spanish', 'fr': 'French', 'it':
'Italian'}

def lang_detect(txt):
    response = client.detect_dominant_language(Text=txt)
    result = str(response['Languages'])
    y = result.split("""': '""")[1].split("""','""")[0]
    x = result.index('.')
    conf = result[x+1:x+3]

    print(my_dict.get(y)+" detected with "+conf+"% confidence" )

txt = "L'amor che move il sole e l'altre stelle."
lang_detect(txt)
```

1.1 Results



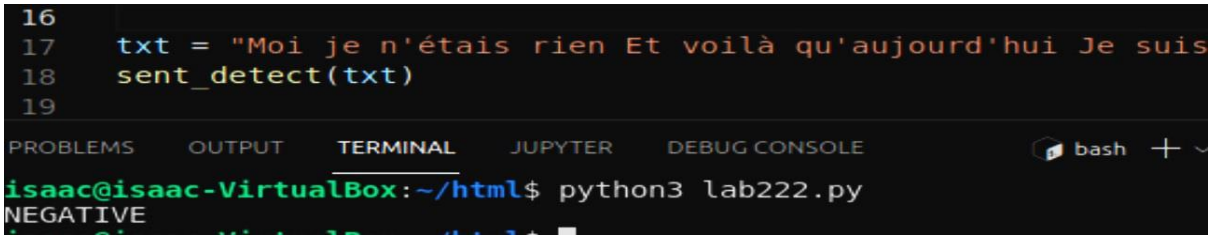
```
18
19    txt = "L'amor che move il sole e l'altre stelle."
20    lang_detect(txt)
21
PROBLEMS  OUTPUT  TERMINAL  JUPYTER  DEBUG CONSOLE
isaac@isaac-VirtualBox:~/html$ python3 lab222.py
Italian detected with 99% confidence
```

2. SENTIMENT DETECTION

```
import boto3
client = boto3.client('comprehend')

def sent_detect(txt):
    response = client.detect_sentiment(
    Text=txt,
    LanguageCode=code_detect(txt)
    )
    print(response['Sentiment'])

def code_detect(txt):
    response = client.detect_dominant_language(Text=txt)
    result = str(response['Languages'])
    y = result.split("""': '""")[1].split("""',""")[0]
    return y

txt = "L'amor che move il sole e l'altre stelle."
sent_detect(txt)
```

2.1 Results

```
16
17    txt = "El Quijote es la obra más conocida de Miguel de C
18    sent_detect(txt)
19
PROBLEMS  OUTPUT  TERMINAL  JUPYTER  DEBUG CONSOLE  bash
isaac@isaac-VirtualBox:~/html$ python3 lab222.py
NEUTRAL
```



3. ENTITIES, KEY PHRASES & SYNTAX?

3.1 Entities

Entities are the key elements in a text , like names of people, places, organisations, brands, dates, quantities, etc. They are usually something with distinct and independent existence. Generally speaking, if they have a specific page on Wikipedia, it usually means they are entities. Extracting entities helps us sort unstructured data and detect important information.

3.2 Key Phrases

Key phrases literally mean key phrases. They might not be entities, but they are closely connected to that context or occur repeatedly in that context. They are usually extracted via statistic methods or algorithm.

3.3 Syntax

Syntax is the grammatical arrangement of words in the text. AWS Syntax Detection shows us what words are verbs, nouns, adjectives, etc. and conjunctives, punctuations, and pronouns, etc which are all essential to construct grammatically correct sentences.

4. ENTITIES DETECTION

```python
import boto3
from pprint import pprint
client = boto3.client('comprehend')

def ent_detect(txt):
    response = client.detect_entities(
    Text=txt,
    LanguageCode=code_detect(txt)
    )
    pprint(response['Entities'])

def code_detect(txt):
    response = client.detect_dominant_language(Text=txt)
    result = str(response['Languages'])
    y = result.split("""': '""")[1].split("""',""")[0]
    return y

txt = "L'amor che move il sole e l'altre stelle."
ent_detect(txt)
```

4.1 Results

```
17
18    txt = "El Quijote es la obra más conocida de Miguel de
19    ent_detect(txt)
20
PROBLEMS   OUTPUT   TERMINAL   JUPYTER   DEBUG CONSOLE   bash

  'Type': 'QUANTITY'},
 {'BeginOffset': 336,
  'EndOffset': 356,
  'Score': 0.9015272259712219,
  'Text': 'Quijote de Cervantes',
  'Type': 'TITLE'},
 {'BeginOffset': 374,
  'EndOffset': 421,
  'Score': 0.9891144633293152,
  'Text': 'El ingenioso caballero don Quijote de la Mancha',
  'Type': 'TITLE'}]
isaac@isaac-VirtualBox:~/html$
```

```
17
18    txt = "Moi je n'étais rien Et voilà qu'aujourd'hui Je
19    ent_detect(txt)
20
```

PROBLEMS OUTPUT TERMINAL JUPYTER DEBUG CONSOLE bash

```
isaac@isaac-VirtualBox:~/html$ python3 lab222.py
[{'BeginOffset': 32,
  'EndOffset': 43,
  'Score': 0.9656052589416504,
  'Text': "aujourd'hui",
  'Type': 'DATE'},
 {'BeginOffset': 127,
  'EndOffset': 138,
  'Score': 0.6522696018218994,
  'Text': "Tout ce qu'",
  'Type': 'QUANTITY'}]
isaac@isaac-VirtualBox:~/html$
```

```
17
18    txt = "L'amor che move il sole e l'altre stelle."
19    ent_detect(txt)
```

PROBLEMS OUTPUT TERMINAL JUPYTER DEBUG CONSOLE bash

```
isaac@isaac-VirtualBox:~/html$ python3 lab222.py
[]
```

5. KEY PHRASE DETECTION

```python
import boto3
from pprint import pprint
client = boto3.client('comprehend')

def phr_detect(txt):
    response = client.detect_key_phrases(
    Text=txt,
    LanguageCode=code_detect(txt)
    )
    pprint(response['KeyPhrases'])

def code_detect(txt):
    response = client.detect_dominant_language(Text=txt)
    result = str(response['Languages'])
    y = result.split("""': '""")[1].split("""',""")[0]
    return y

txt = "L'amor che move il sole e l'altre stelle."
phr_detect(txt)
```

5.1 Results

```
17
18    txt = "El Quijote es la obra más conocida de Miguel de
19    phr_detect(txt)
```

PROBLEMS OUTPUT TERMINAL DEBUG CONSOLE bash

```
{'BeginOffset': 315,
 'EndOffset': 331,
 'Score': 0.9999786019325256,
 'Text': 'la segunda parte'},
{'BeginOffset': 336,
 'EndOffset': 356,
 'Score': 0.9999293088912964,
 'Text': 'Quijote de Cervantes'},
{'BeginOffset': 361,
 'EndOffset': 370,
 'Score': 0.9999479055404663,
 'Text': 'el título'},
{'BeginOffset': 377,
 'EndOffset': 421,
 'Score': 0.9364849328994751,
 'Text': 'ingenioso caballero don Quijote de la Mancha'}]
saac@isaac-VirtualBox:~/html$
```

```
17
18    txt = "Moi je n'étais rien Et voilà qu'aujourd'hui
19    phr_detect(txt)
```

PROBLEMS OUTPUT TERMINAL DEBUG CONSOLE

```
{'BeginOffset': 200,
 'EndOffset': 204,
 'Score': 0.8932915329933167,
 'Text': 'tout'},
{'BeginOffset': 223,
 'EndOffset': 227,
 'Score': 0.9624954462051392,
 'Text': 'tout'},
{'BeginOffset': 241,
 'EndOffset': 243,
 'Score': 0.9998551607131958,
 'Text': 'Je'},
{'BeginOffset': 244,
 'EndOffset': 246,
 'Score': 0.9999161958694458,
 'Text': "l'"}]
saac@isaac-VirtualBox:~/html$
```

```
17
18    txt = "L'amor che move il sole e l'altre stelle."
19    phr_detect(txt)
```

PROBLEMS OUTPUT TERMINAL DEBUG CONSOLE

```
[{'BeginOffset': 0,
  'EndOffset': 6,
  'Score': 0.9999063611030579,
  'Text': "L'amor"},
 {'BeginOffset': 7,
  'EndOffset': 10,
  'Score': 0.9997649788856506,
  'Text': 'che'},
 {'BeginOffset': 16,
  'EndOffset': 23,
  'Score': 0.999977171421051,
  'Text': 'il sole'},
 {'BeginOffset': 26,
  'EndOffset': 40,
  'Score': 0.9999290108680725,
  'Text': "l'altre stelle"}]
isaac@isaac-VirtualBox:~/html$
```

6. SYNTAX DETECTION

```
import boto3
from pprint import pprint
client = boto3.client('comprehend')

def syn_detect(txt):
    response = client.detect_syntax(
    Text=txt,
    LanguageCode=code_detect(txt)
    )
    pprint(response['SyntaxTokens'])

def code_detect(txt):
    response = client.detect_dominant_language(Text=txt)
    result = str(response['Languages'])
    y = result.split("""': '""")[1].split("""',""")[0]
    return y

txt = "L'amor che move il sole e l'altre stelle."
syn_detect(txt)
```

6.1 Results

```
17
18    txt = "El Quijote es la obra más conocida de Miguel de
19    syn_detect(txt)
```

PROBLEMS OUTPUT TERMINAL DEBUG CONSOLE bash

```
{'BeginOffset': 412,
 'EndOffset': 414,
 'PartOfSpeech': {'Score': 0.999546229839325, 'Tag': 'DET'},
 'Text': 'la',
 'TokenId': 78},
{'BeginOffset': 415,
 'EndOffset': 421,
 'PartOfSpeech': {'Score': 0.9804901480674744, 'Tag': 'PROPN'},
 'Text': 'Mancha',
 'TokenId': 79},
{'BeginOffset': 421,
 'EndOffset': 422,
 'PartOfSpeech': {'Score': 0.9999538660049438, 'Tag': 'PUNCT'},
 'Text': '.',
 'TokenId': 80}]
isaac@isaac-VirtualBox:~/html$
```

```
17
18    txt = "Moi je n'étais rien Et voilà qu'aujourd'hui Je
19    syn_detect(txt)
```

PROBLEMS OUTPUT TERMINAL DEBUG CONSOLE bash

```
{'BeginOffset': 246,
 'EndOffset': 250,
 'PartOfSpeech': {'Score': 0.9997336268424988, 'Tag': 'VERB'},
 'Text': 'aime',
 'TokenId': 52},
{'BeginOffset': 251,
 'EndOffset': 252,
 'PartOfSpeech': {'Score': 0.9999866485595703, 'Tag': 'ADP'},
 'Text': 'à',
 'TokenId': 53},
{'BeginOffset': 253,
 'EndOffset': 259,
 'PartOfSpeech': {'Score': 0.9999510049819946, 'Tag': 'VERB'},
 'Text': 'mourir',
 'TokenId': 54}]
isaac@isaac-VirtualBox:~/html$
```

```
17
18    txt = "L'amor che move il sole e l'altre stelle."
19    syn_detect(txt)
```

```
PROBLEMS    OUTPUT    TERMINAL    DEBUG CONSOLE                bash
{'BeginOffset': 28,
 'EndOffset': 33,
 'PartOfSpeech': {'Score': 0.682819664478302, 'Tag': 'ADJ'},
 'Text': 'altre',
 'TokenId': 9},
{'BeginOffset': 34,
 'EndOffset': 40,
 'PartOfSpeech': {'Score': 0.9943856000900269, 'Tag': 'NOUN'},
 'Text': 'stelle',
 'TokenId': 10},
{'BeginOffset': 40,
 'EndOffset': 41,
 'PartOfSpeech': {'Score': 0.9999809265136719, 'Tag': 'PUNCT'},
 'Text': '.',
 'TokenId': 11}]
isaac@isaac-VirtualBox:~/html$
```

7. AWS REKOGNITION

7.1 Script

```python
import boto3
from pprint import pprint

client = boto3.client('rekognition')
bucket_name = '23019722-cloudstorage'

def photo_rek(photo):
    response = client.detect_labels(
        Image={
            'S3Object':{
                'Bucket': bucket_name,
                'Name': photo
            }
        }
    )
    pprint(response['Labels'][0])

def mod_rek(photo):
    response = client.detect_moderation_labels(
        Image={
            'S3Object':{
```

```
                'Bucket': bucket_name,
                'Name': photo
            }
        }
    )
    pprint(response['ModerationLabels'][:2])

def face_rek(photo):
    response = client.detect_faces(
        Image={
            'S3Object':{
                'Bucket': bucket_name,
                'Name': photo
            }
        }
    )
    pprint(response['FaceDetails'][0])

def text_rek(photo):
    response = client.detect_text(
        Image={
            'S3Object':{
                'Bucket': bucket_name,
                'Name': photo
            }
        }
    )
    for word in response['TextDetections']:
        print(word['DetectedText'])

photo1 = 'urban.jpg'
photo2 = 'bikini.jpg'
photo3 = 'faces.jpg'
photo4 = 'quote.jpg'
photo_rek(photo1)
mod_rek(photo2)
face_rek(photo3)
text_rek(photo4)
```

7.2 Results

```
16      pprint(response['Labels'][0])
17
18  photo1 = 'urban.jpg'
19  photo_rek(photo1)
```

PROBLEMS OUTPUT TERMINAL DEBUG CONSOLE

```
isaac@isaac-VirtualBox:~/html$ python3 lab222.py
{'Confidence': 99.93859100341797,
 'Instances': [],
 'Name': 'Landscape',
 'Parents': [{'Name': 'Outdoors'}, {'Name': 'Nature'}]}
isaac@isaac-VirtualBox:~/html$
```

```
19    photo2 = 'bikini.jpg'
20    photo3 = 'faces.jpg'
21    photo4 = 'quote.jpg'
22    mod_rek(photo2)
```

PROBLEMS OUTPUT TERMINAL DEBUG CONSOLE

```
isaac@isaac-VirtualBox:~/html$ python3 lab222.py
[{'Confidence': 95.46479797363281, 'Name': 'Suggestive',
 {'Confidence': 95.46479797363281,
  'Name': 'Revealing Clothes',
  'ParentName': 'Suggestive'},
 {'Confidence': 93.16560363769531,
  'Name': 'Female Swimwear Or Underwear',
  'ParentName': 'Suggestive'}]
isaac@isaac-VirtualBox:~/html$
```

```
20    photo3 = 'faces.jpg'
21    photo4 = 'quote.jpg'
22    face_rek(photo3)
```

PROBLEMS OUTPUT TERMINAL DEBUG CONSOLE

```
isaac@isaac-VirtualBox:~/html$ python3 lab222.py
{'BoundingBox': {'Height': 0.26907387375831604,
                 'Left': 0.8263019323348999,
                 'Top': 0.7124719023704529,
                 'Width': 0.10592147707939148},
 'Confidence': 99.9985580444336,
 'Landmarks': [{'Type': 'eyeLeft',
                'X': 0.8528038859367371,
                'Y': 0.8239902853965759},
               {'Type': 'eyeRight',
                'X': 0.8996672034263611,
                'Y': 0.8248480558395386},
               {'Type': 'mouthLeft',
                'X': 0.8570965528488159,
                'Y': 0.9150832295417786},
```


```
22    photo4 = 'quote.jpg'
23    text_rek(photo4)
```

PROBLEMS OUTPUT TERMINAL DEBUG CONSOLE

```
isaac@isaac-VirtualBox:~/html$ python3 lab222.py
42
"You are enough
just as you
are."
-MEGHAN MARKLE
RS
```