

Preprocessing Data dari Twitter

Pemrosesan Teks



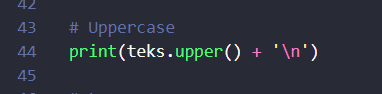
Nim : 5200411488

Nama : Arieska Restu Harpian Dwika

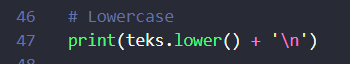
Mengambil Data dari Twitter



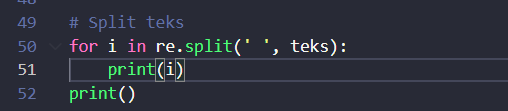
Melakukan Uppercase



Melakukan Lowercase



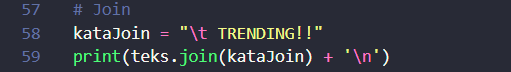
Melakukan Split Text



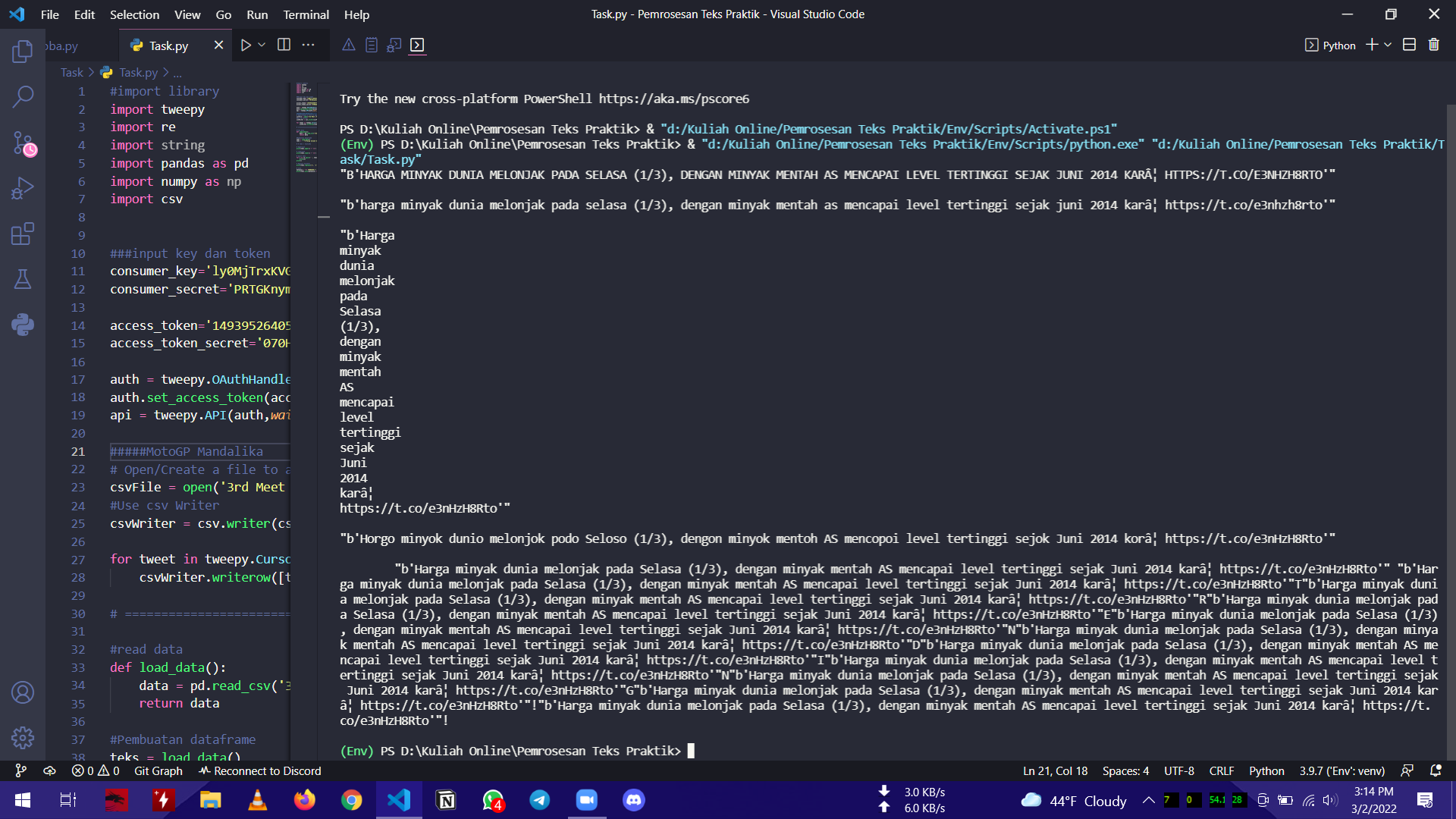
Melakukan Replace Karakter



Melakukan Join



Hasil



Source Code

#import library

import tweepy

import re

import string

import pandas as pd

import numpy as np

import csv

###input key dan token

consumer\_key='ly0MjTrxKVG3CIbYwxSfGnYZq'

consumer\_secret='PRTGKnymI1kKFFgNUS7jYv5p3UpgAh7afHEU7hF6hgwbTuFB57'

access\_token='1493952640577044486-prDDvGVRngcTyJjT6KDMtfUBYODBvg'

access\_token\_secret='070Hm6qiwLtP9ZUUZUOpB2QBxcyRdZrFR8UQKb4FtPvzJ'

auth = tweepy.OAuthHandler(consumer\_key, consumer\_secret)

auth.set\_access\_token(access\_token, access\_token\_secret)

api = tweepy.API(auth,*wait\_on\_rate\_limit*=True)

#####MotoGP Mandalika

# Open/Create a file to append data

csvFile = open('3rd Meet (Thursday, February, 24th 2022)\Practice\Data\Rusia.csv', 'a')

#Use csv Writer

csvWriter = csv.writer(csvFile)

for tweet in tweepy.Cursor(api.search\_tweets,*q*="#rusia",*count*=2,*lang*="id").items():

    csvWriter.writerow([tweet.text.encode('utf8')])

# =========================

#read data

def load\_data():

    data = pd.read\_csv('3rd Meet (Thursday, February, 24th 2022)\Practice\Data\Rusia.csv')

    return data

#Pembuatan dataframe

teks = load\_data()

teks = '''"b'Harga minyak dunia melonjak pada Selasa (1/3), dengan minyak mentah AS mencapai level tertinggi sejak Juni 2014 kar\xe2\x80\xa6 https://t.co/e3nHzH8Rto'"'''

# # teks = texts[0]

# Uppercase

print(teks.upper() + '\n')

# Lowercase

print(teks.lower() + '\n')

# Split teks

for i in re.split(' ', teks):

    print(i)

print()

# Replace karakter

print(teks.replace('a', 'o') + '\n')

# Join

kataJoin = "\t TRENDING!!"

print(teks.join(kataJoin) + '\n')