

**KARAMANOĞLU MEHMETBEY ÜNİVERSİTESİ**

**MÜHENDİSLİK FAKÜLTESİ**

**BİLGİSAYAR MÜHENDİSLİĞİ**

**ADI:** Alper

**SOYADI:** SARGIN

**Numarası:** 201312030

**Ayşe ELDEM**

**Gül Cihan HABEK**

**2023,Karaman**

import pandas as pd

import nltk

from nltk.corpus import stopwords

metin= """Bugün hava çok güzel, yürüyüşe çıkabilirim. Yarın, 3 sınavım var. Şimdi çıkmazsam; bir sonraki treni kaçıracağım. Marketten alınacaklar: çay, kahve, kalem. . #."""

print(metin)

list\_metin=metin.split(".")

seri\_metin=pd.Series(list\_metin)

print(seri\_metin)

seri\_metin = seri\_metin[:-3]

print(seri\_metin)

#normalization

seri\_metin=seri\_metin.apply(lambda x:" ".join(kelime.lower() for kelime in x.split()))

print(seri\_metin)

#noktalama işaretleri ve sayıların silinmesi /işaretler silindi

seri\_metin=seri\_metin.str.replace("[^\w\s]", "",regex=True)

print(seri\_metin)

#sayılar silindi

seri\_metin=seri\_metin.str.replace("[\d]","",regex=True)

print(seri\_metin)

stopwordsList=stopwords.words("english")

seri\_metin=seri\_metin.apply(

    lambda x:" " .join(kelime for kelime in x.split() if kelime not in stopwordsList)

    )

print(seri\_metin)

#az geçen 3 kelimeyi çıkartıyoruz

words=pd.Series(" ".join(seri\_metin).split())

words\_count=words.value\_counts()

print(words)

print(words\_count)

print("\*\*\*\*\*\*\*\*\*")

enSonUcKelime=words\_count[-3:]

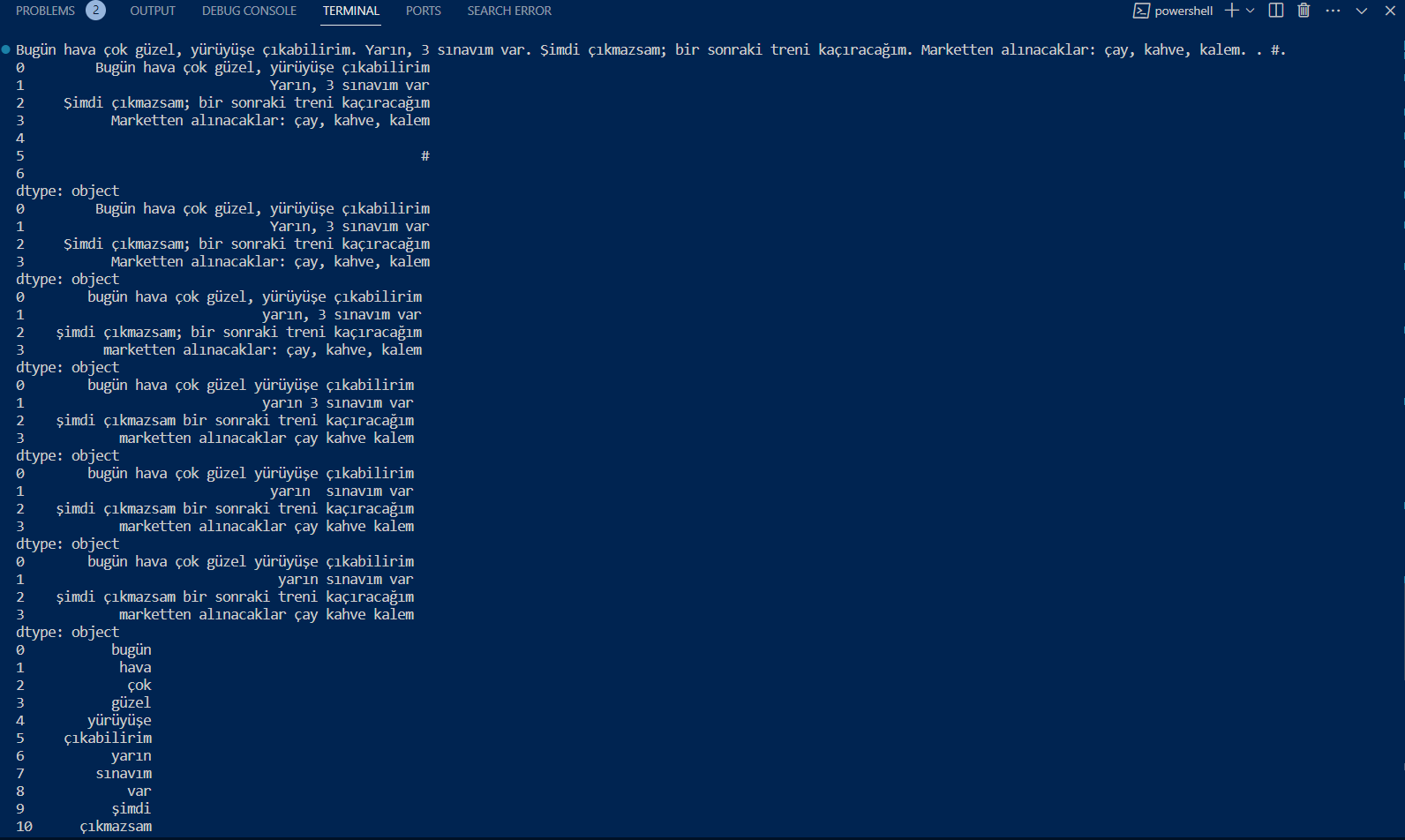
print(enSonUcKelime)

seri\_metin=seri\_metin.apply(

    lambda x:" " .join(kelime for kelime in x.split() if kelime not in enSonUcKelime)

    )

print(seri\_metin)

**Çıktılar**

