

Yıldız Teknik Üniversitesi Bilgisayar Mühendisliği Bölümü Doğal Dil İşlemeye Giriş Ödev-1: Regex

Recep Furkan Koçyiğit 16011043

Regex komut satırı:

(?<mahalle>[A-Z,0-9,Ü,Ğ,İ,Ş,Ç,Ö,\-,\.]*? *(MA?H\.?|MAHALLES(I|İ)))? *(?<cadde>[A-Z,0-9,Ü,Ğ,İ,Ş,Ç,Ö,\-,\.]*? *(CA?D\.?|CADDES(I|İ)))? *(?<sokak>[A-Z,0-9,Ü,Ğ,İ,Ş,Ç,Ö,\-,\.]*? *(SO?K(AK)?\.))? *(.*)?(?<numara>NO.*)?(+(?<ilce>[A-Z,Ü,Ğ,İ,Ş,Ç,Ö]{4,})\/ *(?<il>[A-Z,Ü,Ğ,İ,Ş,Ç,Ö]{4,}))

Mahalle, cadde, sokak, numara, ilçe ve il bilgileri istenilen veriler olduğundan bu gruplar ? operatörü ile işaretlenmiştir. Mahallesi, Mah. , mh. gibi farklılıklar | (veya) operatörü kullanılarak bulunması sağlanmıştır.

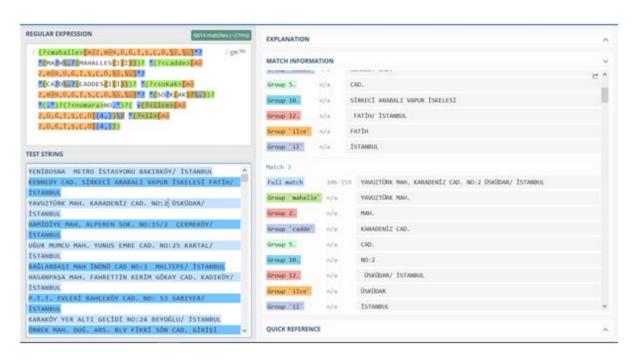
Toplam Adres Sayısı: 6831

Eşleşen Adres Sayısı: 6814

Eşleşmeyen Adres Sayısı: 17

Başarı Oranı: %99.8

Ekran Görüntüsü:



Eşleşmeyen Adresler:

YAYLA MAH. YAY MEYDANI CAD. NO:1 Şişli/ İSTANBUL

FATİH MAH. İMAM AZAM CAD. NO:35/B Pendik/ İSTANBUL

MEHMET AKİF ERSOY MAH. NECİP FAZIL KISAKÜREK CAD.NO:1/1A Çekmeköy/İSTANBUL

YAYLA MAH. YAY MEYDANI CAD. NO:1 Şişli/ İSTANBUL

FATİH MAH. İMAM AZAM CAD. NO:35/B Pendik/ İSTANBUL

MEHMET AKİF ERSOY MAH. NECİP FAZIL KISAKÜREK CAD.NO:1/1A Çekmeköy/ İSTANBUL

ANADOLU MAH. YILDIRIM BEYAZIT. CAD. NO:6/A Arnavutköy/ İSTANBUL

KARTALTEPE MAH. ATATÜRK PARKI NO:4 Küçükçekmece/ İSTANBUL

ISLAM BEY MAH. TAHTA MİNARE CAD. NO:1 EyüpSultan/ İSTANBUL

ANADOLU MAH. YILDIRIM BEYAZIT. CAD. NO:6/A Arnavutköy/ İSTANBUL

KARTALTEPE MAH. ATATÜRK PARKI NO:4 Küçükçekmece/ İSTANBUL

KAYAŞEHİR 3. ETAP 13-14. BÖLGE Başakşehir/ İSTANBUL

BAŞAKŞEHİR MAH. GÜLBAHÇE SOK. NO:2 Başakşehir/ İSTANBUL

KAYAŞEHİR 3. ETAP 13-14. BÖLGE Başakşehir/ İSTANBUL

BAŞAKŞEHİR MAH. GÜLBAHÇE SOK. NO:2 Başakşehir/ İSTANBUL

AŞAĞI DUDULLU MAH. ALEMDAĞ CAD. NO:465/A Ümraniye/ İSTANBUL

AŞAĞI DUDULLU MAH. ALEMDAĞ CAD. NO:465/A Ümraniye/ İSTANBUL

https://regex101.com/r/WPujo5/1