

YILDIZ TEKNİK ÜNİVERSİTESİ BİLGİSAYAR MÜHENDİSLİĞİ DOĞAL DİL İŞLEMEYE GİRİŞ ÖDEV 1 - RegEX Kullanımı

Maziyar Mazaheri 11011902 Üzerinde çalışma yapılan adres dosyasında 6831 adet adres bilgisi bulunmaktadır.

RegEX Komutu:

Bu çalışmada adresler regex komutları ile mahalle, cadde, ilçe, il ve apartman numarası gibi kısımlarına ayrılmıştır.

((.*\sMA?H[\.\s]|.*\sMAHALLES?!?i?)|(.*\sCA?D[\.\s]|.*\sCADDE(S[ii])?)|(.*\sS O?K[\.\s]|.*SOKA[K(GI)])|(NO[;:\.]\s?.*))*(.*)\s(\S*\/\s(iSTANBUL))\$

RegEX Komutunun kapsadığı tanımlar:

Mahalle: (.*\sMA?H[\.\s]|.*\sMAHALLES?!?i?)

Cadde: (.*\sCA?D[\.\s]|.*\sCADDE(S[Ii])?)

Sokak: (.*\sSO?K[\.\s]|.*SOKA[K(GI)])

No: (NO[;:\.]\s?.*)

Diğer: (.*)\s

Semt: \S*\/\s

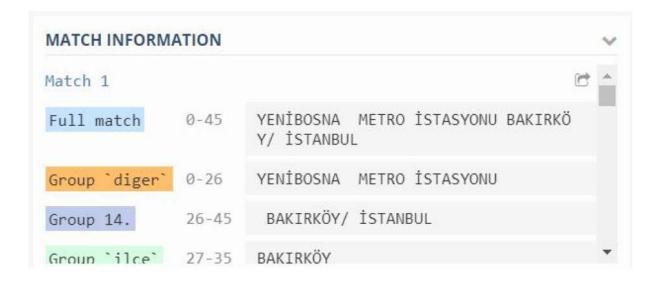
Şehir: (İSTANBUL) \$

Eşleşen adreslere örnek:

```
FİKİRTEPE MAH FAHRETTİN KERİM GÖKAY CAD. ABDİBEY SOK. KESİSİMİ
KADIKÖY/ İSTANBUL
KAVACIK MAH. CUMHURİYET CAD. BEYKOZ/ İSTANBUL
ASAĞI DUDULLU MAH. ALEMDAĞ CAD. NO:465/A ÜMRANİYE/ İSTANBUL
KADIKÖY DENİZ OTOBÜSÜ İSEKELE ÖNÜ RASİMPASA KADIKÖY/ İSTANBUL
19 MAYIS MAH. SEMSETTİN GÜNALTAY CAD. NO:144 KADIKÖY/ İSTANBUL
YUNUS EMRE MAH. VEYSEL KARANÎ CAD. NO: 105 SANCAKTEPE/ ÎSTANBUL
ÇAMLIK MAH. MUHSİN YAZICIOGLU CAD. NO:6/B ÇEKMEKÖY/ İSTANBUL
MECLÍS MAH. ESKÍ AANKARA ASFALTI NO:74/C SANCAKTEPE/ ÍSTANBUL
İSTKLAL MAH. MİTHATPAŞA CAD. NO:6/E ÜMRANİYE/ İSTANBUL
CEBECÍ MAH. 1.CEBECÍ CAD. NO:6 C SULTANGAZÍ/ ÍSTANBUL
SULTAN MAH. MÎLLET CAD. NO:23/4 FATÎH/ ÎSTANBUL
SAHİP MOLLA CAD. NO:37 PAŞABAHÇE BEYKOZ/ İSTANBUL
ÜNALAN MAH. ISMETPAŞA CAD. NO:53/A ÜSKÜDAR/ İSTANBUL
MAKASÇILAR CAD. NO:59/10 FATİH/ İSTANBUL
ÇAKMAK MAH. OKUL SOK. NO:67/B ÜMRANİYE/ İSTANBUL
İSTASYON MAH. YENİ DOĞAN MEVKİİ ŞEN DURAK İST. CAD. NO:17 KÜÇÜKÇEKMECE/
ATAKENT MAH. TÜRKLER CAD. NO:2/1 ÜMRANİYE/ İSTANBUL
YENİ MAH. DR. ÖMER BESİMPAŞA CAD. NO:64 BEYKOZ/ İSTANBUL
GÖZTEPE MAH. TÜTÜNCÜ MEHMET EFENDİ CAD.RIDVANPAŞA SOK. KADIKÖY/
İSTANBUL
 Q
           計
                                                Æ
```

Seçilen bir ifadenin ayrıştırlması şu şekildedir:

19 MAYIS MAH.SEMSETTÍN GÜNALTAY CAD. NO:144 KADIKÖY/ İSTANBUL



Sonuç:

Veri dosyasındaki total adres: 6831

Doğru Adresler: 6825

Yanlış adres: 6

Tespit Doğruluk Oranı: (%99.91)

Testler https://regex101.com/ websitesi kullanılarak yapılmıştır.