

BLM4580 DOĞAL DİL İŞLEMEYE GİRİŞ

ÖDEV-1
VERILEN ADRESLERIN REGEX KOMUTLARI İLE
BULUNMASI

SERHAD ÇALIŞKAN 19011615 Bu çalışmada adresler regex komutları ile mahalle, cadde, ilçe, il ve apartman numarası gibi kısımlarına ayrılmıştır.

Regex ifadesi:

```
REGULAR EXPRESSION

6831 matches, 3878290 steps (~8.12s)

"gm (.*\bBU?LVA?R?I?\.?)?(.*\bCA?D?\.?)?(.*\bSO?K\.?)?

(.*\bBU?LVA?R?I?\.?)?(.*\bİŞ ME?RKE?Zİ?.?)?(.*\bASFA?L?T?I?

\.?\b)?(.*\bVAD\.)?(.*\bMEVKİİ?)?(.*\bDURAKĞ?I?)?(\s*N?O?\s*

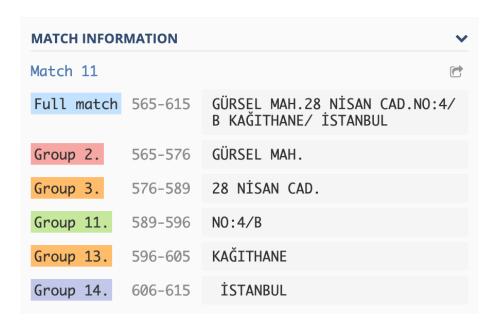
[:;,.]?\s*[0-9]{1,}[A-Z]?\/?[A-Z0-9]{0,2}\s)?(.*\s)?(\b[a-zA-Z0-9ğüşöçİĞÜŞÖÇ]*)?/(.*)?
```

Eşleşen ifadelere örnek:

```
MAKASÇILAR CAD. NO:59/10 FATİH/ İSTANBUL
CAKMAK MAH. OKUL SOK. NO:67/B UMRANIYE/ ISTANBUL
İSTASYON MAH. YENİ DOĞAN MEVKİİ ŞEN DURAK İST. CAD. NO:17 KÜÇÜKÇEKMECE/
İSTANBUL
ATAKENT MAH. TÜRKLER CAD. NO:2/1 ÜMRANİYE/ İSTANBUL
CEBECİ MAH. 1.CEBECİ CAD. NO:6 C SULTANGAZİ/ İSTANBUL
SULTAN MAH. MILLET CAD. NO:23/4 FATIH/ ISTANBUL
SAHİP MOLLA CAD. NO:37 PAŞABAHÇE BEYKOZ/ İSTANBUL
ÜNALAN MAH. ISMETPASA CAD. NO:53/A ÜSKÜDAR/ İSTANBUL
İSTKLAL MAH. MİTHATPAŞA CAD. NO:6/E ÜMRANİYE/ İSTANBUL
19 MAYIS MAH. SEMSETTİN GÜNALTAY CAD. NO:144 KADIKÖY/ İSTANBUL
YUNUS EMRE MAH. VEYSEL KARANİ CAD. NO: 105 SANCAKTEPE/ İSTANBUL
ÇAMLIK MAH. MUHSİN YAZICIOGLU CAD. NO:6/B ÇEKMEKÖY/ İSTANBUL
MECLİS MAH. ESKİ AANKARA ASFALTI NO:74/C SANCAKTEPE/ İSTANBUL
FİKİRTEPE MAH FAHRETTİN KERİM GÖKAY CAD. ABDİBEY SOK. KESİŞİMİ
KADIKÖY/ İSTANBUL
KAVACIK MAH. CUMHURİYET CAD. BEYKOZ/İSTANBUL
AŞAĞI DUDULLU MAH. ALEMDAĞ CAD. NO:465/A ÜMRANİYE/ İSTANBUL
KADIKÖY DENİZ OTOBÜSÜ İSEKELE ÖNÜ RASİMPAŞA KADIKÖY/ İSTANBUL
YENİ MAH. DR. ÖMER BESİMPAŞA CAD. NO:64 BEYKOZ/ İSTANBUL
GÖZTEPE MAH. TÜTÜNCÜ MEHMET EFENDİ CAD.RIDVANPAŞA SOK. KADIKÖY/
İSTANBUL
BAĞDAT CAD. NO:445 KADIKÖY/ İSTANBUL
İNKİLAP MAH. VATAN CAD. NO:44
                                ÜMRANİYE/ İSTANBUL
BAŞIBÜYÜK CAD. ÇİĞDEM SOK.NO:59B MALTEPE/ İSTANBUL
```

Seçilen bir ifadenin nasıl ayrıldığının gösterimi:

GÜRSEL MAH. 28 NİSAN CAD. NO: 4/B KAĞITHANE/ İSTANBUL



Yapılan çalışma sonucu herhangi bir eşleşmeyen ifade bulunmamıştır.

Toplam 6831 eşleşme bulunmuştur. Eşleşme oranı %100'dür.

Fakat bazı ifadeyi kapsamayan -eksik- adreslerde yanlış bir ayrışma söz konusudur.

Örnek:

"CEBECİ MAH. ATATÜRK B.L.V. NO: 97/1" SULTANGAZİ/ İSTANBUL

Görüldüğü gibi B.L.V tanımlı olmadığından dolayı numara ile Atatürk Bulvarı'nı birbirinden ayıramamıştır.