



YILDIZ TEKNİK ÜNİVERSİTESİ
BİLGİSAYAR MÜHENDİSLİĞİ
DOĞAL DİL İŞLEMESİ 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?İ?))|(.*\sCA?D[\.\s]|.*\sCADDE(S[İİ])?)|(.*\sSO?K[\.\s]|.*\sOKA[K(GI)])|(NO[:;\.\s]?.*)(.*\s\S*\s(İSTANBUL))$
```

RegEX Komutunun kapsadığı tanımlar:

Mahalle:	(.*\sMA?H[\.\s] .*\sMAHALLES?I?İ?)
Cadde:	(.*\sCA?D[\.\s] .*\sCADDE(S[İİ])?)
Sokak:	(.*\sSO?K[\.\s] .*\sOKA[K(GI)])
No:	(NO[:;\.\s]?.*)
Diğer:	(.*\s
Semt:	\S*\s
Şehir:	(İSTANBUL) \$

Eşleşen adreslere örnek:

FIKİRTEPE 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
19 MAYIS MAH. SEMSETTİN GÜNALTAY CAD. NO:144 KADIKÖY/ İSTANBUL
YUNUS EMRE MAH. VEYSEL KARANİ CAD. NO:105 SANCAKTEPE/ İSTANBUL
ÇAMLIK MAH. MUHSİN YAZICIOĞLU CAD. NO:6/B ÇEKMEKÖY/ İSTANBUL
MECLİS MAH. ESKİ AANKARA ASFALTI NO:74/C SANCAKTEPE/ İSTANBUL
İSTKLAL MAH. MİTHATPAŞA CAD. NO:6/E ÜMRANİYE/ İSTANBUL
CEBECİ MAH. 1.CEBECİ CAD. NO:6 C SULTANGAZİ/ İSTANBUL
SULTAN MAH. MİLLET CAD. NO:23/4 FATİH/ İSTANBUL
SAHİP MOLLA CAD. NO:37 PAŞABAHÇE BEYKOZ/ İSTANBUL
ÜNALAN MAH. İSMETPAŞA CAD. NO:53/A ÜSKÜDAR/ İSTANBUL
MAKAŞÇILAR 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/
İSTANBUL
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

Seçilen bir ifadenin ayrıştırılması şu şekildedir:

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

MATCH INFORMATION

Match 1

Full match	0-45	YENİBOSNA METRO İSTASYONU BAKIRKÖY/ İSTANBUL
Group `diger`	0-26	YENİBOSNA METRO İSTASYONU
Group 14.	26-45	BAKIRKÖY/ İSTANBUL
Group `ilce`	27-35	BAKIRKÖY

Sonuç:

Veri dosyasındaki total adres: 6831

Doğru Adresler: 6825

Yanlış adres: 6

Tespit Doğruluk Oranı: (%99.91)

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