



BLM4580-DOĞAL DİL İŞLEMeye GİRİŞ

Kamil Başkut

15011091

ÖDEV-1(regEX)

Açıklama

Hedef, verilen adres veri setine uygun regex komutunu yazarak adresleri Mahalle, Sokak, Cadde, No, İlçe olmak üzere 5 ayrı parçaya ayırmaktır. Doğru adresi bulması için mahalle, Sokak ya da Cadde, No ve İlçe bilgilerini bulundurmak zorundadır. Veri setinde 6831 adet adres bulunmaktadır.

Tüm adresleri okumak için regEX Komutu:

```
(([A-zÇĞİÖŞÜÂçğİöşü0-9"\.\- ]*\sMA?HA?L?L?E?S?[İİ]?\.?\.?s?)?(((A-zÇĞİÖŞÜÂçğİöşü0-9"\.\- ]*CA?DD?E?S?[İİ]?\.?)|([A-zÇĞİÖŞÜÂçğİöşü0-9"\.\- ]*\sSO?KA?K?Ğ?[İİ]?\.?\.?)?)([A-zÇĞİÖŞÜÂçğİöşü0-9"\.\- ]*NO\s?\.?\.?:?s?[0-9]*\V?[A-Z0-9]*)?(.*)(( [A-zÇĞİÖŞÜÂçğİöşü]{4,})(\V )([A-zÇĞİÖŞÜÂçğİöşü]*)
```

Adresleri Mahalle->Cadde ->sokak->No->Semt formatında okumak için regEX komutu:

```
(([A-zÇĞİÖŞÜÂçğİöşü0-9"\.\- ]*\sMA?HA?L?L?E?S?[İİ]?\.?\.?s?)(((A-zÇĞİÖŞÜÂçğİöşü0-9"\.\- ]*CA?DD?E?S?[İİ]?\.?)|([A-zÇĞİÖŞÜÂçğİöşü0-9"\.\- ]*\sSO?KA?K?Ğ?[İİ]?\.?\.?)))([A-zÇĞİÖŞÜÂçğİöşü0-9"\.\- ]*NO\s?\.?\.?:?s?[0-9]*\V?[A-Z0-9]*)?(.*)(( [A-zÇĞİÖŞÜÂçğİöşü]{4,})(\V )([A-zÇĞİÖŞÜÂçğİöşü]*)
```

Mahalle

- Komut: ([A-zÇĞİÖŞÜÂçğİöşü0-9"\.\-]*\sMA?HA?L?L?E?S?[İİ]?\.?\.?s?)
- Mahalle bilgisi MH, MAH veya MAHALLESİ şeklinde okunabilir.
- Veri setinde bu üç formata uygun 6003 adet Mahalle bilgisi bulunmuştur.

Cadde

- Komut: ([A-zÇĞİÖŞÜÂçğİöşü0-9"\.\-]*CA?DD?E?S?[İİ]?\.?)
- Cadde bilgisi CD, CAD veya CADDESİ şeklinde okunabilir.
- Veri setinde bu üç formata uygun 4333 adet Cadde bilgisi bulunmuştur.

Sokak

- Komut: ([A-zÇĞİÖŞÜÂçğİöşü0-9"\.\-]*\sSO?KA?K?Ğ?[İİ]?\.?\.?s?)
- Sokak bilgisi SK, SOK, SOKAK ya da SOKAĞI şeklinde okunabilir.
- Veri setinde bu dört format için 1421 adet Sokak bilgisi bulunmuştur.

No

- Komut: ([A-zÇĞİÖŞÜÂçğİöşü0-9"\.\-]*NO\s?\.?\.?:?s?[0-9]*\V?[A-Z0-9]*)
- No bilgisi No şeklinde okunabilir.
- Veri setinde 5961 adet No bilgisi bulunmuştur

İlçe

- Komut: ([A-zÇĞİÖŞÜÂçğİöşü]{4,})(\V)([A-zÇĞİÖŞÜÂçğİöşü]*)
- Veri setinde 6831 adet ilçe bulunmuştur.

Tüm Adressler:

```
/ / ([A=zCĞİÖŞÜAcğışöü=9"\.-_~]*$MA?HA?L?E?P?[ii]?[s?]?)?([A=zCĞİÖŞÜAcğışöü=9"\.-_~]*$SO?KA?K?Ç?F[ii]?[.?.])?( [A=zCĞİÖŞÜAcğışöü=9"\.-_~]*$NO?s?:\?:\s?:[0-9]*\/?[A-Z0-9]*)(.*)([A=zCĞİÖŞÜAcğışöü]
(4,))(\V )([A=zCĞİÖŞÜAcğışöü]*)
```

TEST STRING

YENİBOSNA METRO İSTASYONU BAKIRKÖY/ İSTANBUL
KENNEDY CAD. SİRKECI ARABALI VAPUR İSKELESİ FATİH/ İSTANBUL
YAVUZTÜRK MAH. KARADENİZ CAD. NO:2 ÜSKÜDAR/ İSTANBUL
HAMİDYE MAH. ALPEREN SOK. NO:15/2 ÇEKMEKÖY/ İSTANBUL
UĞUR MUMCU MAH. YUNUS EMRE CAD. NO:25 KARTAL/ İSTANBUL
BAĞLARBAŞI MAH. İNÖNÜ CAD NO:3 MALTEPE/ İSTANBUL
HASANPAŞA MAH. FAHRETTİN KERİM GÖKAY CAD. KADIKÖY/ İSTANBUL
P.T.T. EVLERİ BAHCİKÖY CAD. NO: 53 SARIYER/ İSTANBUL
KARAKÖY YER ALTİ GEÇİDİ NO:24 BEYOĞLU/ İSTANBUL
ÖRNEK MAH. DOĞ. ARS. BLV FIKRI SÖN CAD. GİRİŞİ AGENA E NO. 215 9/2 ESENLYURT/ İSTANBUL
GÜRSEL MAH.28 NISAN CAD.NO:4/B KAĞITHANE/ İSTANBUL
ATATÜRK MAH. ALEMDAĞ CAD. NO:61 ÜMRANIYE/ İSTANBUL
YILDIZ POSTA CAD. TÜRK TELEKOM ÖNÜ GAZETE BAYII BESİKTAS/ İSTANBUL
ARMAĞAN EVLER MAH. ALEMDAĞ CAD. ŞİTE OTOBÜS DURAĞI YANI ÜMRANIYE/ İSTANBUL
FETİHTEPE MAH. FATİH SULTAN CAD. NO:37/B BEYOĞLU/ İSTANBUL
BOZKURT MAH. KURTULUS CAD. NO:135/A ŞİSLİ/ İSTANBUL
PAŞABAHH MAH. BARBAROS CAD. NO:4/A BEYKOZ/ İSTANBUL
YESİL PINAR MAH. ŞUKRAN SOK. NO:36/B EYÜPSULTAN/ İSTANBUL
MUSTAFA KEMAL PAŞA CAD. AZİMKAR SOK.14/1 FATİH/ İSTANBUL
YENİ DOĞAN MAH. KIŞLA CAD. NUR SANAYİ SİTESİ NO: 69 BAYRAMPAŞA/ İSTANBUL
DENİZ KÖSKLER MAH. ESKİ LONDRA ASCALTı NO: 22/3 AVCILAR/ İSTANBUL

Doğru Adressler:

GÜRSel MAH. 28 NİSAN CAD. NO:4/B KAGİTHANE/ İSTANBUL
ATATÜRK MAH. ALEMDAĞ CAD. NO:61 ÜMRANIYE/ İSTANBUL
YILDIZ POSTA CAD. TÜRK TELEKOM ÖNÜ GAZETE BAYII BEŞİKTAS/ İSTANBUL
ARMİGAN EVLER MAH. ALEMDAĞ CAD. SİTE OTOBÜS DURAĞI YANI ÜMRANIYE/ İSTANBUL
FETİHTEPE MAH. FATİH SULTAN CAD. NO:37/B BEYOĞLU/ İSTANBUL
BOZKURT MAH. KURTULUŞ CAD. NO:135/A ŞİŞLİ/ İSTANBUL
PAŞABAĞCI MAH. BARBAROS CAD. NO:4/A BEYKOZ/ İSTANBUL
YEŞİL PINAR MAH. ŞÜKRAN SOK. NO:36/B EYÜPSULTAN/ İSTANBUL
MUSTAFA KEMAL PAŞA CAD. AZİMKAR SOK.14/1 FATİH/ İSTANBUL
YENİ DOĞAN MAH. KIŞLA CAD. NUR SANAYİ SİTESİ NO: 69 BAYRAMPASA/ İSTANBUL
DENİZ KÖŞKLER MAH. ESKİ LONDRA ASFALTI NO: 22/3 AVCILAR/ İSTANBUL
ORTAÇEŞME SONDURAK BEYKOZ/ İSTANBUL
MEHMET AKİF ERSOY MAH. NATO YOLU BOSNA BULVARI NO: 115 ÜSKÜDAR/ İSTANBUL
DERBENT MAH. DEREİÇİ OTOBÜS YOLU SORAK SARIYER/ İSTANBUL
LEVAZİM MAH. KORU SOK NO:7A BEŞİKTAS/ İSTANBUL
ÇENGELKÖY MAH. ÇENGELKÖY CAD. NO:35/A ÜSKÜDAR/ İSTANBUL
BÜYÜKDERE CAD. NİMET ABLA CAMİ YANI ŞİŞLİ/ İSTANBUL
REŞİTPAŞA MAH. POSTAYOLU CAD. NO: 85 SARIYER/ İSTANBUL
RUMELİ HİSAR ÜSTÜ BEBEK YOLU SOK SARIYER/ İSTANBUL
HASANPAŞA MAH. KURBAĞALIDERE CAD. NO:23/A KADIKÖY/ İSTANBUL
SEYRANTEPE MAH. İ. KARAOĞLANOĞLU CAD. NO: 153/B KAGİTHANE/ İSTANBUL
TAHTAKALE MAH. GAFFAR OKAN CAD. NO:10 AVCILAR/ İSTANBUL
BABA HASAN ALEMİ MAH. ATATÜRK BULVARI FATİH/ İSTANBUL
TOPKAPI TİCARET MERKEZİ KARŞISI ÜSTGEÇİT ALTI ZEYİNBURNU/ İSTANBUL
MİLLET CAD. YUSUF PAŞA ZİRAAT BANKASI ÖNÜ FATİH/ İSTANBUL

Bazı yanlış Adresler:

YILDIZ POSTA CAD. TÜRK TELEKOM ÖNÜ GAZETE BAYİİ BEŞİKTAŞ/ İSTANBUL -> Mahalle bilgisi yok

DERBENT MAH. DEREİCİ OTOBÜS SON DURAK SARIYER/ İSTANBUL -> Sokak bilgisi yok

RUMELİ HİSAR ÜSTÜ BEBEK YOLU SOK SARIYER/ İSTANBUL -> Mahalle, Sokak, No bilgileri yok

Link: <https://regex101.com/r/zD29LZ/2>