

BLM4580 - Doğal Dil İşlemeye Giriş

Ödev 1

Canberk Güllüoğlu

18011602

Regex Komut Satırı:

```
(([A-zĞÇÖŞİÜ"\.,0-9\ - ]+MA?H\.?)?([A-zĞÇÖŞİÜ"\.,0-9\ - ]+CA?D\.?)?([A-zĞÇÖŞİÜ"\.,0-9\ - ]+\sSO?K\.?)?(\s*NO\s*[.:]*[0-9\ - ]*\[A-Z\ -0-9]*)?)(.*\s)([A-zĞÇÖŞİÜ]+)\V([A-zĞÇÖŞİÜ"\.,0-9\ - ]+)
```

Bulunan Adres Sayısı : 6818

Bulunamayan Adres Sayısı: 13

Bulunamayan Adresler:

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

ANADOLU MAH. YILDIRIM BEYAZIT. CAD. NO:6/A Arnavutköy/ İSTANBUL x2

KARTALTEPE MAH. ATATÜRK PARKI NO:4 Küçükçekmece/ İSTANBUL x2

KAYAŞEHİR 3. ETAP 13-14. BÖLGE Başakşehir/ İSTANBUL x2

BAŞAKŞEHİR MAH. GÜLBAHÇE SOK. NO:2 Başakşehir/ İSTANBUL x2

MEHMET AKİF ERSOY MAH. NECİP FAZIL KISAKÜREK CAD.NO:1/1A Çekmeköy/ İSTANBUL x2

YAYLA MAH. YAY MEYDANI CAD. NO:1 Şişli/ İSTANBUL x2

Bulunamayan adreslerde küçük harf kullanıldığı için bulunamadı, ayrıca bu adresler adresler.txt dosyasında iki defa tekrar ediyor.

Bulunamayan adreslerin dışında, gruplandırma yapılırken yanlış gruplanan ya da mah./cad./sok./ formatında olmayan adresler de var, bu adresler eksik/yanlış şekilde ayrılıyor.

Diğer eksik bulunan adreslerin en yaygın sebebi ise “No” kısmında olan farklı yazımlar.

Eksik bulunan veya yanlış gruplanan adreslere bazıları:

TURGUT REİS MAHALLESİ FATİH BULVARI NO:235A 1 SULTANBEYLİ/ İSTANBUL

VELİBABA MAH. HUKUKÇULAR CAD. HUKUÇULAR SİTESİ NO:50-1 PENDİK/ İSTANBUL

ATATÜRK BULVARI 32 C FATİH/ İSTANBUL

HASANPAŞA UZUNÇAYIR METROBÜS DURAĞI KADIKÖY/ İSTANBUL

Ekran Görüntüsü:

REGULAR EXPRESSION v4 6818 matches, 1592200 steps (~877ms)

```
i / ((([A-zĞÇÖŞİÜ"\.,0-9\-\ ]+MA?H\.\.?)?([A-zĞÇÖŞİÜ"\.,0-9\-\ ]+CA?D\.\.?)?([A-zĞÇÖŞİÜ"\.,0-9\-\ ]+SOSOK\.\.?)?(\s*NO\s*[.:]\s*[0-9\-\ ]*\s*[A-Z\-\.\.0-9]*)?)(\s*)([A-zĞÇÖŞİÜ]+)\s*\/([A-zĞÇÖŞİÜ"\.,0-9\-\ ]+)
```

TEST STRING

İNÖNÜ MH KAR SK KÖŞEBAŞI ATAŞEHİR/ İSTANBUL
MECİDİYE MH. FATİH BULVARI NO:470 SULTANBEYLİ/ İSTANBUL
YENİSAHRA MAH. KARAMAN CAD. NO:3/1A ATAŞEHİR/ İSTANBUL
MİMAR SİNAN MAH. SELAMİ ALİ EFENDİ CAD. NO9/1 ÜSKÜDAR/ İSTANBUL
KAYIŞDAĞ CAD. NO:221/A KADIKÖY/ İSTANBUL
ÇAKMAK MAH. TAVUKÇUYOLU CAD. NO:30/3 ÜMRANIYE/ İSTANBUL
ÇINARDERE MAH. AYAZMA CAD. NO:4 PENDİK/ İSTANBUL
AKPINAR KÖYÜ NO:3/A EYÜPSULTAN/ İSTANBUL
ZEYİNLİK MAH. PTT ÖNÜ GAZETE BAYII BAKIRKÖY/ İSTANBUL
MERKEZ MAH. ŞEHİT MİTHAT YILMAZ CAD. KEKİK SOK. NO:1 SARIYER/ İSTANBUL
ABDİ İPEKÇİ MH. İSTASYON CD. 1549. SK. NO:18/C DARICA/ İSTANBUL
AYDINTEPE MAH.İBRAHİM KORAY SOK. NO:3/A TUZLA/ İSTANBUL
DENİZ CD. KÜÇÜKSO SK. NO:15 PENDİK PENDİK/ İSTANBUL
CAFER AĞA MAH. ASIM GÜNDÜZ CAD. NO: 84 KADIKÖY/ İSTANBUL
HURRIYET MAH. ESKİ EDİRNE ASFALTI 184 GAZİOSMANPAŞA/ İSTANBUL
ZÜHTÜPAŞA MAH. BAĞDAT CAD. NO: 39 KADIKÖY/ İSTANBUL
RASİMPAŞA MAH. HALİTAĞA CAD. NO: 39/2/1A KADIKÖY/ İSTANBUL
FINDIKLI MAH. ÇINAR CAD..NO:19 MALTEPE/ İSTANBUL
İSTİNYE MAH. DERE SOK. NO:15 SARIYER/ İSTANBUL
KURUÇEŞME MAH. KIRBAŞ SOK. NO:25 BEŞİKTAŞ/ İSTANBUL
KOCATEPE MAH. TARLABAŞI BULVARI NO:18A BEYOĞLU/ İSTANBUL
MERKEZ EFENDİ G-48. SOK. NO: 9/A ZEYİNBURNU/ İSTANBUL
UGUR MUMCU MAH. L CAD. KÖŞESİ NO:594 SULTANGAZİ/ İSTANBUL
ALTIN ŞEHİR MAH. OSMANLI CAD. NO:7/A BAŞAKŞEHİR/ İSTANBUL
AYDOS CAD. RAKİP SOK. NO:2 VELİBABA. MAH. PENDİK/ İSTANBUL
YUKARI DUDULLU MAH CURDA GAZİ GÜLBABA SOK NO : 3-B ÜMRANIYE/ İSTANBUL
KOCASINAN MERKEZ MAH. MAHMUTBEY YOLU NO:168/D BAĞCELİEVLER/ İSTANBUL
OSMANGAZİ MAH. PAPATYA CAD. NO: 23/2 ESENİYURT/ İSTANBUL
MEŞRUTİYET MAH. HALASKAR GAZİ CAD. ŞAİR NİGAR SOK. NO:94 MAYA APT SİSLİ/ İSTANBUL
ÖZYURT SOK. BAŞKAN APT. NO: 21 ATAŞEHİR/ İSTANBUL
İSTİKLAL MAH. MİTHATPAŞA CAD. NO:10/A ÜMRANIYE/ İSTANBUL
YAMANEVLER MAH. ALEMDAĞ CAD. NO: 155 ÜMRANIYE/ İSTANBUL
SELMANİ PAK CAD. NO: 42 ÜSKÜDAR/ İSTANBUL