

FAZLI BOZATAY

16011081

REGEX

```
(?<Mahalle>.*? *MA?H\.(?<ALLES(i|I)?)?(?<Cadde>.*? *CA?D\.(?<DES(i|I)?)?(?<Sokak>.*? *SO?K(AK)?(\.|))?(.*)?(?<NO>NO.*)?( (?<Ilce>[^0-9\n ]* *)\ / *(?<Il>[^0-9\n ]*))
```

YAZILAN REGEX HAKKINDA:

- Mahalle, Cadde, Sokak ve No gibi bileşenlerin varlığı isteğe bağlı olduğu için o gruplar ? operatörü ile işaretlenmiştir.
- İl ve İlçe ayrımı için / işareti tüm adreslerde ortak olduğu için regex içerisinde de koşul kılınmıştır.
- Boşluklar birden fazla veya hiç olmayabileceği için * operatörü ile uygun yerlere konulmuştur.
- Mah. Mh. Mahallesi gibi çeşitlilikler ? ve | operatörleriyle sağlanmıştır.
- Bir adres satırı ; Mahalle, Cadde, Sokak, No, İl ve İlçe olarak gruplandırılmaya çalışılmıştır. Bu bileşenler arasında olmayan bir bileşen için gruplama yapılmamıştır.

SONUÇLAR:

-Hiç bulunamayan adres satırı yoktur.

-Eksiksiz bileşenlerine ayrılan bir adres örneği :

YAVUZTÜRK MAH. KARADENİZ CAD. NO:2 ÜSKÜDAR/ İSTANBUL

-Eksin bileşenlerine ayrılan bir adres örnekleri

İSTİKLAL MH. MİTHATPAŞA CD. NO:34/B 2. ATAKNET MAH.BURÇ CAD. 3/6 ÜMRANIYE/ İSTANBUL

Bu örnekteki gibi birden fazla mahalle, cadde, sokak veya no gibi bileşenlerin bulunması durumu desteklenmemektedir.

GÖKEVLER MAH.ÖZDEMİR SOK.BELDEKENT SİTESİ A BLOK NO:1 BELDEM MARKET ESENİYURT/ İSTANBUL

Burada NO: dan sonra apartman no, blok no gibi bilgiler beklenirken İlçe ve İl dışında başka bir adres bilgisi ortaya çıkmıştır. Bu bilgi NO ile aynı grup içerisinde bulunarak yanlış bilgi çıkarımı oluşmuştur.

KİREÇ BURNU CAD. HAYDAR ALİYEV CAD. NO 48/B SARIYER/ İSTANBUL

Bu örnekte birden fazla cadde girilmiştir. İlk cadde algılanabilirken ikinci yazılan cadde ekstra adres bilgisi olarak işaretlenmiştir.

SSK OKMEYDANI HASTANE ÖNÜ HALK EKMEK ŞİŞLİ/ İSTANBUL

Bu örnekteki SSK sözcüğü “S” sokağın adı “SK” sokak keywordu olarak regex tarafından yanlış olarak yorumlanmıştır.

-6831 adres satırı içerisinde ortalama yüzde 1 lik bir kısım örneklerdeki gibi yanlış ayrılmıştır. Yani 60 tane eksik bulunan adres vardır.