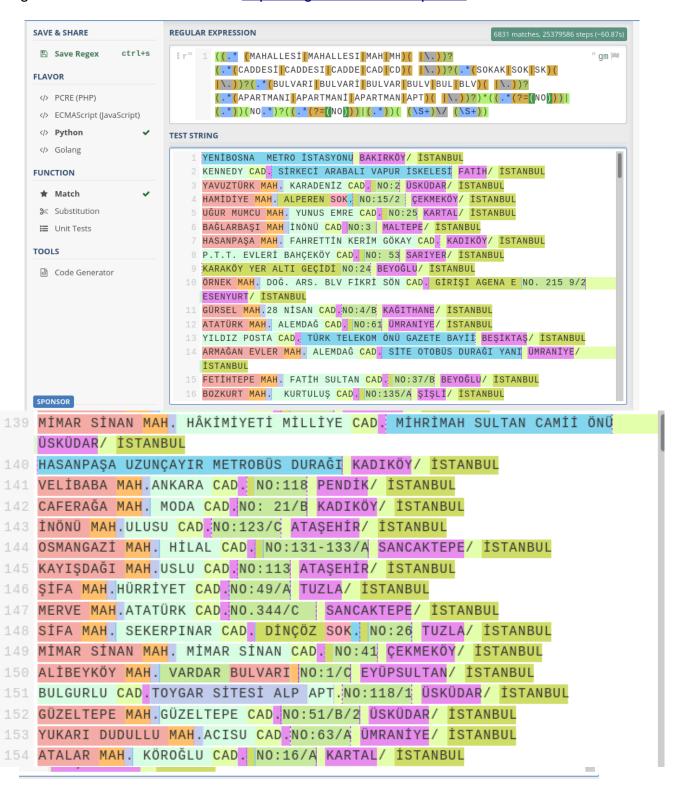


Doğal Dil İşleme 1. Ödev

Ahmet Elgün 16011016

Regex: ((.* (MAHALLESİ|MAHALLESI|MAH|MH)(|\.))?(.*(CADDESİ|CADDESI|CADDE|CAD|CD)(|\.))?(.*(SOKAK|SOK|SK)(|\.))?(.*(BULVARI|BULVARİ|BULVAR|BLV)(|\.))? (.*(APARTMANI|APARTMANİ|APARTMAN|APT)(|\.))?)*((.*(?=(NO)))|(.*))(NO.*)?((.*(?=(NO)))|(.*))((\S+)V (\S+))

Regex101 üzerinde deneme linki: https://regex101.com/r/rXpiZT/1



Eksik bulunan: Veri boyutu çok büyük olduğu için kaç tanesinin eksik olduğu tespit edilememiştir ama büyük çoğunluğunun eksiksiz bulunduğu görülmektedir. Eksik bulunanların bir kısmı içinde cadde, mahalle gibi kelimeler geçen tarifler ve NO kelimesi yazılmadan verilen NO'lardan oluşuyor.

Bazı eksik bulunan adresler:

ÖRNEK MAH. DOĞ. ARS. BLV FİKRİ SÖN CAD. GİRİŞİ AGENA E NO. 215 9/2

MUSTAFA KEMAL PAŞA CAD. AZİMKAR SOK.14/1 FATİH/ İSTANBUL

DOĞAN ARASLI BULV.NO:158 KOYICI MEVKI ESENYURT/ İSTANBUL