

Muhammed Emin Bal



- <https://www.youtube.com/c/MuhammedEminBal>

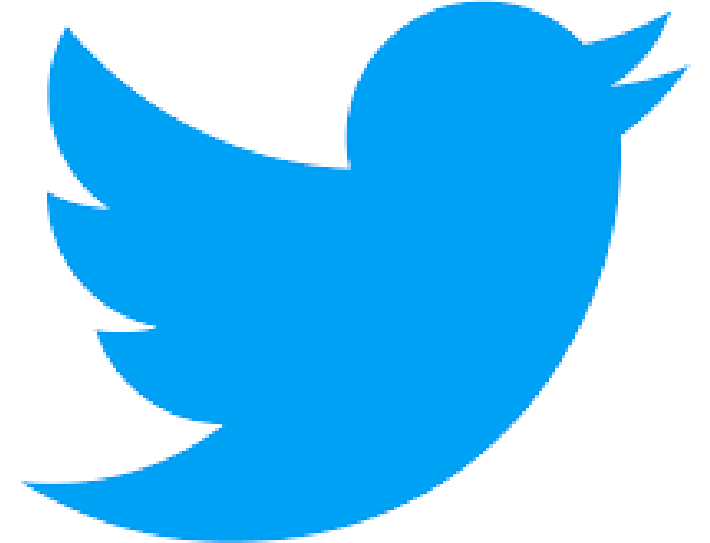
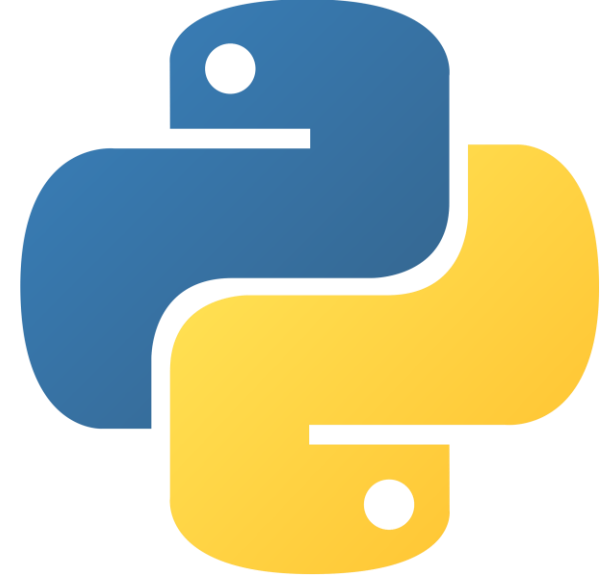


- <https://www.linkedin.com/in/muhammedeminbal/>



- m.balemin@gmail.com
- <https://muhammedbal.com/>

Python İle Sosyal Medya ve E- Ticaret Sitelerinden Veri Çekme Yöntemleri



Program Akışı

- 1- Kullanacağımız modülleri tanımak
- 2- Bazı Http Durumlarını inceleme
- 3- Modüllerin Kurulumu
- 4- Requests ile örnek yapma
- 5- Selenium ile örnek yapma

Kullanacağımız Modüller

- **1- Requests Modülü:**

- Bu modül ile web sitelerine isteklerimizi gönderebiliyoruz.
- Get, post gibi işlemler yapılabilir.
- Çok hızlı ve kullanışlı bir modüldür.
- Login olma işlemi yapılabilir.

- **2- Selenium Modülü:**

- Selenium bilgisayarımıza kuracağımız bir driver ile tarayıcı açarak web sitelerine gitme ve oralarda gezinme imkanı sunar.
- Selenium test işlemleri içinde çok kullanılmaktadır.
- Selenium ile bir siteye giderek nasıl veri alınabileceğini işleyeceğiz.
- Requests modülüne göre oldukça yavaştır. Kullanım yerine göre ikisi arasında tercih yapılır.

- **3- BeautifulSoup:**

- Aldığımız verileri parçalamak için kullanacağız.

- **4- lxml:**

- Html taglerini BeautifulSoup ile parçalarken lxml kullanacağız.

- **5- Time Modülü:**

- Selenium da siteye gittiğimiz zaman sitenin yüklenmesini beklemek yani kodu bekletmek için kullanacağımız zaman modülüdür.

- **6- Pandas:**

- İnternette aldığımız verileri kolayca excel/csv ye çevirmek için kullanacağız.

Sık Kullanılan HTTP Durum Kodları

Http durum kodları siteye istek attığımız zaman başarılı mı başarısız mı olduğunun sonucunu bize veren kodlardır. Http durum kodlarının toplamı 70-80 tane vardır fakat bizim en çok karşılaştığımız durumları maddeler halinde öğrenelim.

- 1- HTTP 200** : Siteye istek attığımız zaman bu kod dönerse başarılı olduğunu anlıyoruz ve başarılı bir şekilde siteye gitmiş oluyoruz.
- 2- HTTP 403** : Siteye gitme işlemi başarılı oluyor. Fakat site size erişim izni vermiyor bir nevi engelliyor.
- 3- HTTP 404** : En çok bilinen hatalardan :). Gittiğiniz sayfanın bulunamadığını söylüyor.
- 4- HTTP 500** : İstek atılıyor fakat isteğin tamamlanamadığını bildiriyor.