

---

# Data Engineering Masterclass

Homework

Çarşamba, 24.02.2021

---

## Konu

Bu ödev, IMDB verisetleri üzerinde çeşitli soruların kolayca cevaplanabilmesi için bir veritabanı ve bir arayüz geliştirmeyi amaçlar.

## İçerik

1. [IMDB veri setlerini](#) indirin.
  2. Bu veri setlerini tutabileceğiniz bir MySQL instance'ı kurun. Bunun için tercihinize göre Virtualbox, Vagrant, Docker kullanabilirsiniz. (Hint: Docker kullanıyorsanız, persistent volume kullanmak isteyebilirsiniz.)
  3. IMDB veri setlerini tutabilmek için MySQL üzerinde gerek gördüğünüz tabloları oluşturun. (Hint: tercihinize göre her veri seti için ayrı bir tablo kullanabilirsiniz veya optimize bir yapı kurabilirsiniz.)
  4. İndirdiğiniz veri setlerini MySQL'e kaydedecek bir uygulama yazın. Bu uygulama sadece bir SQL query'si çalıştıran bir Bash script'i olabileceği gibi, tercihinize göre Python ya da Java gibi bir programlama dili de kullanabilirsiniz. Bu yazdığınız uygulamayı ödevle birlikte submit edeceksiniz, o nedenle nasıl çalıştırılacağı bilgilerini dokumante etmelisiniz.
-

5. Aşağıdaki TV dizilerinden birini seçin.
  - a. Game of Thrones
  - b. Breaking Bad
  - c. The Mandalorian
  - d. Rick and Morty
  
6. Seçtiğiniz dizi ile ilgili olarak aşağıdaki sorulara cevap veren SQL query'lerini (ya da bir cevabı üreten bir uygulama) yazın. Bu query'leri bulduğunuz cevaplarla birlikte submit edeceksiniz, o nedenle her query'yi ve query'ye karşılık bulduğunuz cevabı dokümente etmelisiniz. (Bonus: Query'lerin çalışma sürelerini eklerseniz, bonus kazanacaksınız.)
  - a. Seçtiğiniz dizinin IMDB average rating'i en yüksek olan 5 bölümü
  - b. Seçtiğiniz dizinin IMDB average rating'i en düşük olan 5 bölümü
  - c. Seçtiğiniz dizinin IMDB'de en çok kişi tarafından oylanmış 5 bölümü
  - d. Seçtiğiniz dizinin IMDB'de en az kişi tarafından oylanmış 5 bölümü
  
7. Aşağıdaki sorulara cevap veren SQL query'lerini yazın.
  - a. 2000-2010 arasında oyuncularından biri tarafından yönetilmiş filmlerin listesi
  - b. 2000 yılından günümüze bir filmde müzik departmanında da yer almış aktrislerin listesi
  - c. 4 ya da daha fazla aktris bulunduran filmlerin sayısı
  
8. Minik bir dinamik web sayfası hazırlayın. Hazırladığınız web sayfası MySQL veritabanınıza bağlanıp seçtiğiniz dizinin bölümlerini IMDB rating'ine göre azalan biçimde ekranda listelemelidir. (Hint: en aşina olduğunuz programlama dilinde hazırlayabilirsiniz. Hiçbir dile aşina değilseniz, Python, PHP, NodeJS gibi dillerden biriyle kolayca hazırlayabilirsiniz.)

## Deliverables

**2 ve 3. Madde:** Hazırladığınız MySQL veritabanının dump'ını Google Drive, Dropbox, Mega gibi bir dosya paylaşım servisine yüklemeli ve linkini ödevle birlikte teslim etmelisiniz.

**4. Madde:** Verisetlerini MySQL'e kayıt eden uygulamanızı bir Github repository'si haline getirmeli ve linkini ödevle birlikte paylaşmalısınız. Hazırladığınız repo bir README dosyası içermelidir ve herhangi bir IDE'ye ihtiyaç duymadan kodun nasıl çalıştırılabileceğini detaylıca anlatmalıdır. Lütfen olabildiğiniz kadar ayrıntılı şekilde (gerektiğini düşündüğünüz yerlerde fazladan ayrıntı vererek) detaylı bir README hazırlamaya çalışın.

**5. Madde:** Seçtiğiniz dizinin adını ödevle birlikte teslim edin.

**6. Madde:** Her sorunun cevabı ve bu cevabı üreten uygulamayı Github reponuzda ayrı bir dizine koyun ve nasıl çalıştırılacağını detaylıca yazın.

**7. Madde:** Her sorunun cevabı ve bu cevabı üreten uygulamayı Github reponuzda ayrı bir dizine koyun ve nasıl çalıştırılacağını detaylıca yazın.

**8. Madde:** Hazırladığınız web sayfasını Heroku (ya da dışarıdan erişime izin veren herhangi bir servise) deploy edin. Bu sayfanın linkini ödevle birlikte teslim edin.

## Rapor

Ödevde nasıl bir yol izlediğinizi anlatan bir rapor teslim etmenizi bekliyoruz. Burada amaç rapor yazmış olmanızı sağlamak değil, izlediğiniz yolu takip edebilmemiz. Bu yüzden detaylı ve özenli bir rapor hazırlamanız, sınavda başarılı olmanız için çok önemli. da rapor içerisine ekleyin. Nihai olarak içerisinde tüm linkleri ve ödevle ait cevapları bir rapor-PDF teslim edeceksiniz.

## Teslim tarihi

Ödevi 7 Mart 2021 23:59'a kadar teslim etmenizi bekliyoruz. Ödevi teslim edebileceğiniz linki bu belgenin içinde bulunduğu epostada bulabilirsiniz.

## Sorular

Olabildiğinizce sadece yukarıdaki bilgiler ışığında ödevi yapmanızı bekliyoruz. İçinden çıkamadığınız durumlarda Telegram grubundan -publicly- sınıf arkadaşlarınıza soru sorabilirsiniz.

## Etik

Ödevi yalnızca kendi çabanızla yapmanızı ve kopya çekmemenizi bekliyoruz. Lütfen bunun tamamen öğrenmeye ve teknik kişisel gelişime yönelik bir çaba olduğunu aklınızdan çıkarmayın. Ödevi teslim ederken, lütfen direkt olarak faydalandığınız kaynakları belirtin.

Ödevi teslim eden her katılımcı, tüm ödevi kendi çabasıyla ve yalnızca kendi çabasıyla yaptığını taahhüt etmiş sayılır.