

TBY414 VERİ MADENCİLİĞİ

Ders Çizelgesi: Salı 09:00 – 11:50, LAB C-307

Öğretim Görevlisi:	Bora GÜNGÖREN
Oda No:	Belirlenecek
Telefon:	
Eposta:	bora.gungoren@gmail.com / borag@metu.edu.tr
Haftalık görüşme:	Randevu ile ayarlanacaktır.
Ders web sayfası:	Dönem Github sayfası ağırlıklı olarak kullanılacaktır. https://github.com/boragungoren-portakalteknoloji/BASKENT-TBY414-Bahar2020
Ders Tanımı:	
<p>Çağdaş iş dünyasındaki yöneticilerin işle ilgili problemlerini çözerken artan düzeyde analitik olmaları gerekmektedir. Bu süreç problemlerin ve problem çözümü için gereken bilgilerin tespit edilmesi anlamına gelmektedir. Bu ders iş ve işletme verisi odaklı problemlerin çözülmesinde kullanılacak araçlara odaklanmaktadır. Dersin ana konusu veri işleme olmakla birlikte veri kaynaklarının tespit edilmesi, veri paylaşımı ve bir işletmenin veri akışının verinin daha kullanılabilir hale getirilmesi için değiştirilmesi de ders konuları arasındadır.</p>	
Öğrenme Hedefleri	
<p>Ders öğrenme hedefleri:</p> <p>Derse özel hedefler:</p> <ol style="list-style-type: none">1. İş verilerinin türlerini ve iş verisi işlemenin genel türlerini belirleyebilmek.2. Bir veri analizi problemini analiz edip sınıflandırabilmek.3. İş hayatındaki karar verme problemlerini çözmek için basit Python programları yazabilmek.4. İşletme veri kümelerini paylaşmak ve işlemekte kullanılan çağdaş araçları ve veri biçimlerini öğrenmek. <p>Disipline özel hedefler:</p> <ol style="list-style-type: none">5. Verili bir yönetsel bir problemten yola çıkarak veri yoğunluklu bir karar problemini oluşturabilmek.6. İş kararlarını vermek için veri kalitesi gereksinimlerini anlamak. <p>Kişisel becerileri:</p> <ol style="list-style-type: none">7. Analitik beceriler.8. Yazılı iletişim becerileri.	
Öğrenme Yöntemleri:	
<p>Ders uygulamalı olacaktır. Tüm oturumlar bilgisayar laboratuvarında gerçekleştirilecektir. Hemen her hafta öğrenciler belirli bir problemin çözümü için laboratuvarında program yazıyor olacaktır. Değerlendirme için birden fazla programlama ödevi, laboratuvar ortamında gerçekleştirilecek iki test ve bir adet takım projesi olacaktır. Takım projesi final sınavı yerine geçecektir.</p>	
Ders Kitabı ve Yardımcı Kitaplar Reading:	
<p>Karaçay, T. 2017. "Python 3. 3. Baskı" Seçkin Yayıncılık.</p> <p>Ceder, N. 2018. "The Quick Python Book. 3rd Ed." Manning.</p> <p>Ceder, N. 2018. "Exploring Data with Python." Manning.</p> <p>Pilgrim, M. 2009. "Dive into Python 3." Apress.</p> <p>Björgevinnson, T. 2017. "The Art of Data Usability." Manning.</p> <p>Cielen, D, et a. 2016. "Introducing Data Science." Manning.</p>	

Değerlendirme:				
Değerlendirme Biçimi	% Katkı	Değerlendirme Süresi	Öğrenme Hedeflerine Katkı	Geribildirim Yöntemi
Laboratuvar	10	20-40 dakikalık programlama görevleri	1,2,3,5,7	Sözlü ve yazılı
Ödevler	20	Her ödev için en az 2-4 saatlik bireysel çalışma	Tamamı	Sözlü ve yazılı
Test 1	20	Laboratuvarda yapılacak olan 60-90 dakikalık sınav	1,2,3,4,5,7	Sözlü ve yazılı
Test 2	20	Laboratuvarda yapılacak olan 60-90 dakikalık sınav	1,2,3,4,5,7	Sözlü ve yazılı
Team Project	30	Takım üyesi başına 10-20 saatlik bireysel çalışma	Tamamı	Sözlü ve yazılı

Notlar:

Kopya: Kopya çekilmesi durumunda akademik kurallar katı biçimde uygulanacaktır. Ödevlerin hazırlanmasında bilgi paylaşmak, fikir alışverişi yapmak, İnternet kaynaklarını araştırmak kopya çekmek değildir. Ancak dersin doğası gereği program yazıldığı için programların hiç bir bireysel katkı konmaksızın direkt olarak kopyala-yapıştır olarak aktarılması durumunda kopya çekilmiş olarak görülecektir.

İntihal: Ders ödevleri ve özellikle takım projelerinde raporlarda üçüncü kaynaklardan bilgilere başvurulması gerekecektir. Bu kaynakların düzgün biçimde belirtilmesi gerekmektedir. Aksi takdirde kopya olarak değerlendirileceklerdir.

Ders Gerekleri: Derse düzenli devam edilmesi çok önemlidir. Ders konuları bir önceki konular üzerine inşa edildiği için uzun süre devamsızlık yapılması durumunda dersi takip etmek çok zor gelebilmektedir. BU nedenle derse devam zorunludur ve ders devamı izlenecektir. Her hafta, **bazen bir defadan daha sık** biçimde devam alınacaktır.

Github: Tüm öğrencilerin bir adet Github hesabı açması zorunludur. Tüm ödevler ve projeler Github üzerinden teslim edilecektir. Github, kodların kim tarafından, hangi kaynaktan, hangi gün ve saatte yüklendiği, ne şekilde değişiklikler içerdiğini aktif biçimde izlemeye yardımcı bir araçtır. Bu nedenle özellikle parçalar halinde yapılması ve ilerleme gerçekleşmesi öngörülen ödevler ve projeler yanı sıra takım projesinde takım arkadaşlarının projeye olan katkısı yönünde izlenebilirlik sağlamaktadır.

Takım Projesi: Öğrenciler 3 ila 5 kişiden oluşan takımlar oluşturacak ve **kendi önerecekleri** bir iş problemini çözmek için Python ile bir uygulama geliştireceklerdir. Projeler veri toplama veya üçüncü kaynaklardan veri çekmeyi de içeren uçtan uca bir çaba olarak görülmelidir. Bu nedenle projelerin dönem sonundaki teslim süresine az bir zaman kala başlatılması sıkıntılı olacaktır.

Dönem Planı

Hafta	Topic	
1	Derse Giriş <ul style="list-style-type: none">İş hayatındaki veri yoğun problemler ve bir veri projesinin yapısı.Veri paylaşma ve işleme araçlarındaki alternatiflerimiz.Github.Örnek vazı veri kümeleri.	
2	Python Adım 1 <ul style="list-style-type: none">Girintileme ile kapsam belirlemeDeğişkenler ve deyimlerMetin ve sayı değişkenleri ile çalışmakTemel işlemlerNone değeriKomut satırı girdisi almak	
3	Python Adım 2 <ul style="list-style-type: none">Tuple and List türleriSözlük türüProgram akışını yönetmekİşlev tanımlamak	
4	Python Adım 3 <ul style="list-style-type: none">Modül tanımlamakDosya sistemine erişimDosyaları işlemekPandas kullanımı	
5	Python Adım 4 <ul style="list-style-type: none">İş verisinin türleriVeri kaynaklarıVeri kalitesi problemleriVeri ön işleme kavramlarıVeri ön işleme teknikleri	
6	Python Step 5 <ul style="list-style-type: none">Grafik türleri ve kullanımlarıMatplotlib kitaplığı ile grafik çizdirmekGrafiklerle ilgili veri kümesi problemleri	
7	TEST 1 (Ders saatinde)	
8	Örnek Proje 1 – Zaman Serisi Kestirimi <ul style="list-style-type: none">Dönemselliği anlamakSimple exponential smoothing (SES),Holt ve Holt-Winter teknikleri,	
9	Örnek Proje 1 – Zaman Serisi Kestirimi <ul style="list-style-type: none">Durağanlığın sınanmasıARIMA modellerinin tartışılmasıVeri projelerinin nasıl yürütüleceğinin tartışılması	
10	Örnek Proje 2 – Anomali Tespiti <ul style="list-style-type: none">Anomali türleriAykırlıklar ve Yeniliklerin ayrışmasıYoğunluk teknikleri: K-Nearest Neighbor (KNN) ve Local Outlier Factor (LOF).Kümeleme teknikleri: K-Means.	

	<ul style="list-style-type: none"> • Kaggle'daki kredi kartı sahtekarlık veri kümesi ile bir uygulama • Gerçek bir problemin nasıl ele alınması gerektiğinin tartışılması 	
11	Örnek Proje 3 – İş Kuralları <ul style="list-style-type: none"> • Kural motoru kavramı • Kuralların tespit edilmesi – Çağırışım Kuralı Madenciliği 	
12	TEST 2 (Ders Saatinde)	
13	Örnek Proje 4 – Öneriler <ul style="list-style-type: none"> • Öneri sistemlerinin karşılaştırılması: İşbirlikçi ve içerik merkezli yöntemler • Modeller, yanlılık ve varyans • Öneri sistemi performansının değerlendirileceği • Python ile basit bir öneri sistemi: Surprise 	
14	Veri Kalitesi ve Veri Kullanılabilirliği <ul style="list-style-type: none"> • Temel tanımlar • Bir iş akışının veri kalitesi odaklı gözden geçirilmesi 	