



Customer Segmentation using **RFM**

İş Problemi

> İŞ PROBLEMİ

VERİ SETİ

DEĞİŞKENLER

GÖREVLER

Bir e-ticaret şirketi müşterilerini segmentlere ayırıp bu segmentlere göre pazarlama stratejileri belirlemek istemektedir.

Şirket, ortak davranışlar sergileyen müşteri segmentleri özelinde pazarlama çalışmaları yapmanın gelir artışı sağlayacağını düşünmektedir.

Örneğin şirket için çok kazançlı olan müşterileri elde tutmak için farklı kampanyalar, yeni müşteriler için farklı kampanyalar düzenlenmek istenmektedir.

Müşterileri **RFM yöntemi** ile segmentlere ayırınız.



Veri Seti Hikayesi

Online Retail II isimli veri seti İngiltere merkezli online bir satış mağazasının 01/12/2009 - 09/12/2011 tarihleri arasındaki satışlarını içermektedir.

Bu şirketin ürün kataloğunda hediyeelik eşyalar yer almaktadır.

Şirketin müşterilerinin büyük çoğunluğu kurumsal müşterilerdir.



Değişkenler

InvoiceNo – Fatura Numarası

Eğer bu kod C ile başlıyorsa işlemin iptal edildiğini ifade eder.

StockCode – Ürün kodu

Her bir ürün için eşsiz numara.

Description – Ürün ismi

Quantity – Ürün adedi

Faturalardaki ürünlerden kaçar tane satıldığını ifade etmektedir.

InvoiceDate – Fatura tarihi

UnitPrice – Fatura fiyatı (Sterlin)

CustomerID – Eşsiz müşteri numarası

Country – Ülke ismi

İŞ PROBLEMİ

VERİ SETİ

> DEĞİŞKENLER

GÖREVLER



Proje Görevleri

DSMLBC | VERİ BİLİMİ OKULU

Veri Setine Eriřim

Ařağıdaki buton yardımı ile verisetini indirebilirsiniz.



Çalışmanız için:

[online_retail_II.xlsx](#) veri setinin “**Year2010-2011**” isimli sheet’ine RFM analizi uygulayınız.



Görev 1:

Veriyi Anlama ve Hazırlama

1. Online Retail II excelindeki 2010-2011 verisini okuyunuz. Oluşturduğunuz dataframe'in kopyasını oluşturunuz.
2. Veri setinin betimsel istatistiklerini inceleyiniz.
3. Veri setinde eksik gözlem var mı? Varsa hangi değişkende kaç tane eksik gözlem vardır?
4. Eksik gözlemleri veri setinden çıkartınız. Çıkarma işleminde 'inplace=True' parametresini kullanınız.
5. Eşsiz ürün sayısı kaçtır?
6. Hangi üründen kaçar tane vardır?
7. En çok sipariş edilen 5 ürünü çoktan aza doğru sıralayınız.
8. Faturalardaki 'C' iptal edilen işlemleri göstermektedir. İptal edilen işlemleri veri setinden çıkartınız.
9. Fatura başına elde edilen toplam kazancı ifade eden 'TotalPrice' adında bir değişken oluşturunuz.



Görev 2:

RFM metriklerinin hesaplanması

- Recency, Frequency ve Monetary tanımlarını yapınız.
- Müşteri özelinde Recency, Frequency ve Monetary metriklerini **groupby**, **agg** ve **lambda** ile hesaplayınız.
- Hesapladığınız metrikleri **rfm** isimli bir değişkene atayınız.
- Oluşturduğunuz metriklerin isimlerini **recency**, **frequency** ve **monetary** olarak değiştiriniz.



Not 1: recency değeri için bugünün tarihini (2011, 12, 11) olarak kabul ediniz.

Not 2: rfm dataframe'ini oluşturduktan sonra veri setini “monetary>0” olacak şekilde filtreleyiniz.

İpucu:

```
rfm = df.groupby('Customer ID').agg({'InvoiceDate': lambda date: (--, --).days,  
                                     'Invoice': lambda num: --,  
                                     'TotalPrice': lambda TotalPrice: --})
```



Görev 3:

RFM skorlarının oluşturulması ve tek bir değişkene çevrilmesi

- Recency, Frequency ve Monetary metriklerini qcut yardımı ile 1-5 arasında skora çeviriniz.
- Bu skorları recency_score, frequency_score ve monetary_score olarak kaydediniz.
- Oluşan 2 farklı değişkenin değerini tek bir değişken olarak ifade ediniz ve RFM_SCORE olarak kaydediniz.

Örneğin;

Ayrı ayrı değişkenlerde sırasıyla 5, 2 olan recency_score, frequency_score skorlarını RFM_SCORE değişkeni isimlendirmesi ile oluşturunuz.



DİKKAT! Monetary skoru dahil etmiyoruz.



Görev 4:

RFM skorlarının segment olarak tanımlanması

- Oluşturulan RFM skorların daha açıklanabilir olması için segment tanımlamaları yapınız.
- Aşağıdaki seg_map yardımı ile skorları segmentlere çeviriniz.

İpucu:

```
# RFM isimlendirmesi
seg_map = {
    r'[1-2][1-2]': 'hibernating',
    r'[1-2][3-4]': 'at_Risk',
    r'[1-2]5': 'cant_loose',
    r'3[1-2]': 'about_to_sleep',
    r'33': 'need_attention',
    r'[3-4][4-5]': 'loyal_customers',
    r'41': 'promising',
    r'51': 'new_customers',
    r'[4-5][2-3]': 'potential_loyalists',
    r'5[4-5]': 'champions'
}
```



Görev 5:

Aksiyon zamanı!

- Önemli bulduğunuz 3 segmenti seçiniz. Bu üç segmenti;
 - Hem aksiyon kararları açısından,
 - Hem de segmentlerin yapısı açısından (ortalama RFM değerleri) yorumlayınız.
- “**Loyal Customers**” sınıfına ait customer ID’leri seçerek excel çıktısını alınız.





Data Science and Machine Learning Bootcamp



İletişim: veribilimi@veribilimiokulu.com