



gezinomi
her gün tatil olsa...



Kural Tabanlı Sınıflandırma ile
Potansiyel Müşteri Getirisi Hesaplama

miuul

İş Problemi

Gezinomi yaptığı satışların bazı özelliklerini kullanarak seviye tabanlı (level based) yeni satış tanımları oluşturmak ve bu yeni satış tanımlarına göre segmentler oluşturup bu segmentlere göre yeni gelebilecek müşterilerin şirkete ortalama ne kadar kazandırabileceğini tahmin etmek istemektedir.

Örneğin:

Antalya'dan Herşey Dahil bir otele yoğun bir dönemde gitmek isteyen bir müşterinin ortalama ne kadar kazandırabileceği belirlenmek isteniyor.



İŞ PROBLEMİ

VERİ SETİ

DEĞİŞKENLER

GÖREVLER

Veri Seti Hikayesi

gezinomi_miuul.xlsx veri seti Gezinomi şirketinin yaptığı satışların fiyatlarını ve bu satışlara ait bilgiler içermektedir. Veri seti her satış işleminde oluşan kayıtlardan meydana gelmektedir. Bunun anlamı tablo tekilleştirilmemiştir. Diğer bir ifade ile müşteri birden fazla alışveriş yapmış olabilir.



iŞ PROBLEMI

› VERİ SETİ

DEĞİŞKENLER

GÖREVLER

Değişkenler

miuul_gezinomi.xlsx

SaleId: Satış id

SaleDate: Satış Tarihi

CheckInDate: Müşterinin otele giriş tarihi

Price: Satış için ödenen fiyat

ConceptName: Otel konsept bilgisi

SaleCityName: Otelin bulunduğu şehir bilgisi

CInDay: Müşterinin otele giriş günü

SaleCheckInDayDiff: Check in ile giriş tarihi gün farkı

Season: Otele giriş tarihindeki sezon bilgisi

İŞ PROBLEMİ

VERİ SETİ

DEĞİŞKENLER

GÖREVLER



Proje Görevleri

Uygulama Öncesi Veri Seti :

SaleId	SaleDate	CheckInDate	Price	ConceptName	SaleCityName	CInDay	SaleCheckInDayDiff	Seasons
415122	12/3/2022	12/3/2022	79.30403	Herşey Dahil	Antalya	Saturday	0	Low
415103	12/3/2022	12/3/2022	45.9707	Yarım Pansiyon	Antalya	Saturday	0	Low
404034	9/12/2022	9/13/2022	77.83883	Herşey Dahil	Antalya	Tuesday	1	High
415094	12/3/2022	12/10/2022	222.7106	Yarım Pansiyon	İzmir	Saturday	7	Low
414951	12/1/2022	12/3/2022	140.4762	Yarım Pansiyon	İzmir	Saturday	2	Low

Hedeflenen çıktı :

sales_level_based	SEGMENT	Price
GIRNE_HERŞEY DAHİL_HIGH	A	103.9354
GIRNE_HERŞEY DAHİL_LOW	A	90.93594
İZMİR_YARIM PANSİYON_HIGH	A	87.6573
DIĞER_HERŞEY DAHİL_LOW	A	87.31088
DIĞER_HERŞEY DAHİL_HIGH	A	83.78727

Görev 1: Aşağıdaki Soruları Yanıtlayınız

Soru 1: miuul_gezinomi.xlsx dosyasını okutunuz ve veri seti ile ilgili genel bilgileri gösteriniz..

Soru 2: Kaç unique şehir vardır? Frekansları nedir?

Soru 3: Kaç unique Concept vardır?

Soru 4: Hangi Concept'den kaçar tane satış gerçekleşmiş?

Soru 5: Şehirlere göre satışlardan toplam ne kadar kazanılmış?

Soru 6: Concept türlerine göre göre ne kadar kazanılmış?

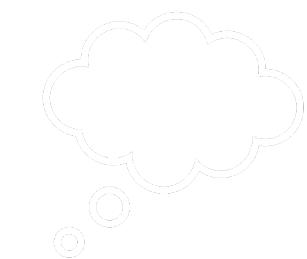
Soru 7: Şehirlere göre PRICE ortalamaları nedir?

Soru 8: Conceptlere göre PRICE ortalamaları nedir?

Soru 9: Şehir-Concept kırılımında PRICE ortalamaları nedir?

Görev 2: SaleCheckInDayDiff değişkenini kategorik bir değişkene çeviriniz.

- SaleCheckInDayDiff değişkeni müşterinin CheckIn tarihinden ne kadar önce satın almını tamamladığını gösterir.
- Aralıkları ikna edici şekilde oluşturunuz.
Örneğin: '0_7', '7_30', '30_90', '90_max' aralıklarını kullanabilirsiniz.
- Bu aralıklar için "Last Minuters", "Potential Planners", "Planners", "Early Bookers" isimlerini kullanabilirsiniz.



Görev 3: COUNTRY, SOURCE, SEX, AGE kırılımında **ortalama kazançlar** nedir?

Şehir-Concept-EB Score, Şehir-Concept- Sezon, Şehir-Concept-CInDay kırılımında ortalama ödenen ücret ve yapılan işlem sayısı cinsinden inceleyiniz ?

Elde edilmesi gereken çıktı:

	SaleCityName	ConceptName	EB_Score	PRICE	
				mean	count
0	Antalya	Hersey Dahil	Last Minuters	62.75	14148
1			Potential Planners	64.90	8874
2			Planners	67.88	4490
3			Early Bookers	66.49	3281
4		Oda + Kahvaltı	Last Minuters	65.35	503

Görev 4: City-Concept-Season kırılımının çıktısını PRICE'a göre sıralayınız.

- Elde ettiğiniz çıktıyı agg_df olarak kaydediniz.

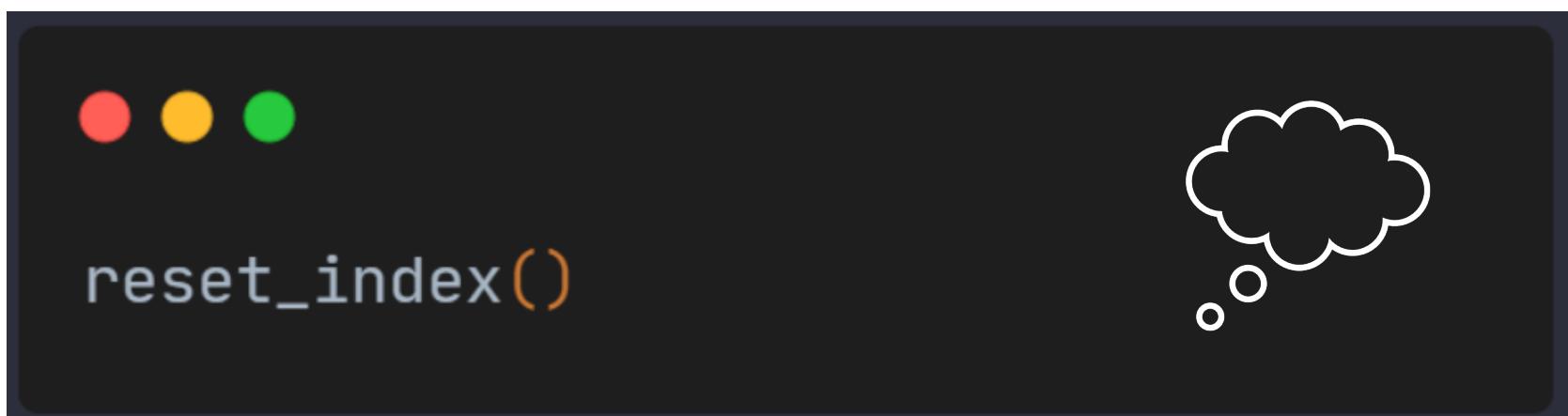
Elde edilmesi gereken çıktı:

	SaleCityName	ConceptName	Seasons	PRICE
0	Girne	Herşey Dahil	High	103.94
1			Low	90.94
2	İzmir	Yarım Pansiyon	High	87.66
3	Diğer	Herşey Dahil	Low	87.31
4			High	83.79

Görev 5: Indekste yer alan isimleri değişken ismine çeviriniz.

- Üçüncü sorunun çıktısında yer alan PRICE dışındaki tüm değişkenler index isimleridir. Bu isimleri değişken isimlerine çeviriniz.

İpucu:



Görev 6: Yeni seviye tabanlı müşterileri (persona) tanımlayınız.

- Yeni **seviye tabanlı satışları** tanımlayınız ve veri setine değişken olarak ekleyiniz.
- Yeni eklenecek değişkenin adı: sales_level_based
- Önceki soruda elde edeceğiniz çıktıdaki gözlemleri bir araya getirerek **sales_level_based** değişkenini oluşturmanız gerekmektedir.

Bu tabloda bulunan gözlemler bir araya gelecek

SaleCityName	ConceptName	Seasons	PRICE
Girne	Hersey Dahil	High	103.9354
Girne	Hersey Dahil	Low	90.93594
İzmir	Yarım Pansiyon	High	87.6573
Diğer	Hersey Dahil	Low	87.31088
Diğer	Hersey Dahil	High	83.78727

Elde edilmesi gereken çıktı

SaleCityName	ConceptName	Seasons	PRICE	sales_level_based
Girne	Hersey Dahil	High	103.9354	GIRNE_HERSEY DAHIL_HIGH
Girne	Hersey Dahil	Low	90.93594	GIRNE_HERSEY DAHIL_LOW
İzmir	Yarım Pansiyon	High	87.6573	İZMİR_YARIM PANSİYON_HIGH
Diğer	Hersey Dahil	Low	87.31088	DIĞER_HERSEY DAHIL_LOW
Diğer	Hersey Dahil	High	83.78727	DIĞER_HERSEY DAHIL_HIGH

Görev 7: Yeni müşterileri (personaları) segmentlere ayırınız.

- Yeni personaları **PRICE**'a göre 4 segmente ayırınız.
- Segmentleri **SEGMENT** isimlendirmesi ile değişken olarak **agg_df**'e ekleyiniz.
- Segmentleri betimleyiniz (Segmentlere göre group by yapıp price mean, max, sum'larını alınız).

İpucu:

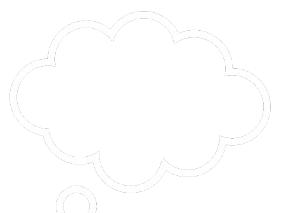


```
● ● ●  
pd.qcut(agg_df["PRICE"], 4, labels=["D", "C", "B", "A"])
```

Görev 8: Yeni gelen müşterileri sınıflandırıp, ne kadar gelir getirebileceklerini tahmin ediniz.

- Antalya'da herşey dahil ve yüksek sezonda tatil yapmak isteyen bir kişinin ortalama ne kadar gelir kazandırması beklenir?
- Girne'de yarım pansiyon bir otele düşük sezonda giden bir tatilci hangi segmentte yer alacaktır?

İpucu:



```
new_user = "ANTALYA_HERSEY DAHİL_HIGH"  
agg_df[agg_df[sales_level_based]== new_user]
```

miuul

miuul.com