

```
import pandas as pd
```

```
#####
```

```
# Kural Tabanlı Sınıflandırma ile Potansiyel Müşterileri Getirisi Hesaplama
```

```
# Abdurrahman Bulut
```

```
#####
```

```
#####
```

```
# Aşılabilir Sorun
```

```
#####
```

```
# Bir oyun şirketi müşterilerinin bazı özelliklerini kullanarak seviye tabanlı (level based) yeni müşterileri tanımlar (persona)
```

```
# oluşturmak ve bu yeni müşterileri tanımlarına göre segmentler oluşturup bu segmentlere göre yeni gelebilecek müşterilerinin şirkete
```

```
# ortalama ne kadar kazandırabileceğini tahmin etmek istemektedir.
```

```
# Arayışın: Türkiye'den iOS kullanıcısı olan 25 yaşındaki bir erkek kullanıcısının ortalama ne kadar kazandırabileceğini belirlemek isteniyor.
```

```
#####
```

```
# Veri Seti Hikayesi
```

```
#####
```

```
# Persona.csv veri seti uluslararası bir oyun şirketinin sattığı ürünlerin fiyatları ve bu ürünleri satın alan kullanıcıların bazı
```

```
# demografik bilgilerini barındırmaktadır. Veri seti her satış işlemiyle oluşan kayıtlardan meydana gelmektedir. Bunun anlamı tablo
```

```
# tekilleştirilmiştir. Diğer bir ifadeyle demografik özelliklere sahip bir kullanıcı birden fazla alışveriş yapabilir.
```

```
# Price: Müşterinin harcama tutarı
```

```
# Source: Müşterinin baktığı cihaz türü
```

```
# Sex: Müşterinin cinsiyeti
```

```
# Country: Müşterinin ülkesi
```

```
# Age: Müşterinin yaşı
```

```
##### Uygulama 1 #####
```

```
# PRICE SOURCE SEX COUNTRY AGE
# 0 39 android male bra 17
# 1 39 android male bra 17
# 2 49 android male bra 17
# 3 29 android male tur 17
# 4 49 android male tur 17
```

```
##### Uygulama Sonrası #####
```

```
# customers_level_based PRICE SEGMENT
# 0 BRA_ANDROID_FEMALE_0_18 1139.800000 A
# 1 BRA_ANDROID_FEMALE_19_23 1070.600000 A
# 2 BRA_ANDROID_FEMALE_24_30 508.142857 A
# 3 BRA_ANDROID_FEMALE_31_40 233.166667 C
# 4 BRA_ANDROID_FEMALE_41_66 236.666667 C
```

```
#####
```

```
# PROJE GİRİŞİ
```

```
#####
```

```
#####
```

```
# GİRİŞ 1: Aşağıdaki soruları yanıtlayınız.
```

```
#####
```

```
# Soru 1: persona.csv dosyasını okutunuz ve veri seti ile ilgili genel bilgileri gösteriniz.
```

```
df = pd.read_csv("modul2/tasks/persona.csv")
```

```
df.head()
```

```
df.describe().T
```

```
df.info()
```



```

# list comp ile customers_level_based deÄYerleri oluÅYturulduktan sonra bu deÄYerlerin
tekillenÅYtirilmesi gerekmektedir.
# ÅrneÄYin birden fazla ÅYu ifadeden olabilir: USA_ANDROID_MALE_0_18
# BunlarÅ± groupby'a alÅ±p price ortalamalarÅ±nÅ± almak gerekmektedir.
# Yeni seviye tabanlÅ± mÅYÅYterileri tanÅ±mlama
agg_df['customers_level_based'] = agg_df['COUNTRY'].str.upper() + '_' +
agg_df['SOURCE'].str.upper() + '_' + agg_df[
    'SEX'].str.upper() + '_' + agg_df['AGE_CAT'].str.upper()

agg_df = agg_df.groupby('customers_level_based')
['PRICE'].mean().reset_index().sort_values(ascending=False, by='PRICE')
agg_df
#####
# GÅ-REV 7: Yeni mÅYÅYterileri (USA_ANDROID_MALE_0_18) segmentlere ayÅ±rÅ±nÅ±z.
#####
# PRICE'a gÅYre segmentlere ayÅ±rÅ±nÅ±z,
# segmentleri "SEGMENT" isimlendirmesi ile agg_df'e ekleyiniz,
# segmentleri betimleyiniz,
agg_df['SEGMENT'] = pd.qcut(agg_df['PRICE'], q=4, labels=['D', 'C', 'B', 'A'])

# default observed=true deprecated oluyormuÅY
segment_summary = agg_df.groupby('SEGMENT', observed=False)['PRICE'].agg(['mean', 'max', 'sum'])
segment_summary

#####
# GÅ-REV 8: Yeni gelen mÅYÅYterileri sÅ±nÅ±flandÅ±rÅ±nÅ±z ne kadar gelir getirebileceÄYini tahmin
ediniz.
#####
# 33 yaÅYÅ±nda ANDROID kullanan bir TÅYrk kadÅ±nÅ± hangi segmente aittir ve ortalama ne kadar
gelir kazandÅ±rmasÅ± beklenir?
new_user = "TUR_ANDROID_FEMALE_31_40"
agg_df[agg_df['customers_level_based'] == new_user].SEGMENT # B segmenti
# 35 yaÅYÅ±nda IOS kullanan bir FransÅ±z kadÅ±nÅ± hangi segmente ve ortalama ne kadar gelir
kazandÅ±rmasÅ± beklenir?
# Yeni gelen mÅYÅYterilerin sÅ±nÅ±flandÅ±rÅ±lmasÅ± ve gelir tahmini
new_user = "FRA_IOS_FEMALE_31_40"
agg_df[agg_df['customers_level_based'] == new_user].SEGMENT # C segmenti

```