metin, ekran görüntüsü içeren bir resim

Açıklama otomatik olarak oluşturuldu

Nunique = number unique = eşsiz değer sayısı

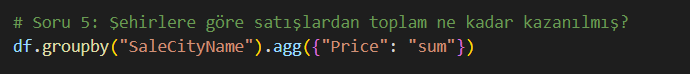
Value counts = değerlerin sayısı

metin, ekran görüntüsü, yazı tipi, çizgi içeren bir resim

Açıklama otomatik olarak oluşturuldu

İki değere birlikte bakmak için bu değişkenleri gruplamak gerekmektedir.

Bu sayede kırılımları inceleyebiliriz



Burada SaleCityName ve Price değişkenleri gruplanmıştır.

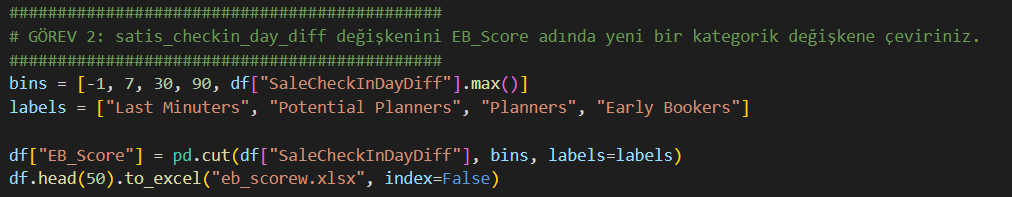
Yapılan gruplamanın ise “sum” yani toplam değerine bakılmaktadır.

Buna diğer bir örnek

metin, yazı tipi, ekran görüntüsü, çizgi içeren bir resim

Açıklama otomatik olarak oluşturuldu

Burada kullanılan by=…. İfadesi ise benzer gruplamaların birbirine karışmaması içindir.



satis\_checkin\_day\_diff değişkeni SAYISAL görünen ama aslında kategorik niteliği ifade eden bir değişkendir.

Bu nedenle aralarında 1 gün – 7 gün – 30 ve 90 gün – en son olarakta en max kaç ise o kadar gün şeklinde gruplama yaparak sayısal değerler kategorize edilmek istenmiştir.

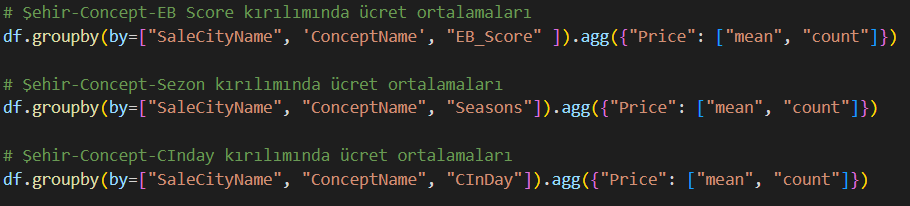
Bu sayede veriler sınıflandırılmış ve daha iyi bir gözlem elde edilebilmiştir (okunaklılık artırılmıştır.)

Diğer satırda ise yapılan bu sınıflandırmaya ise isimler verilmiştir. (labels, etikeleme işlemi)

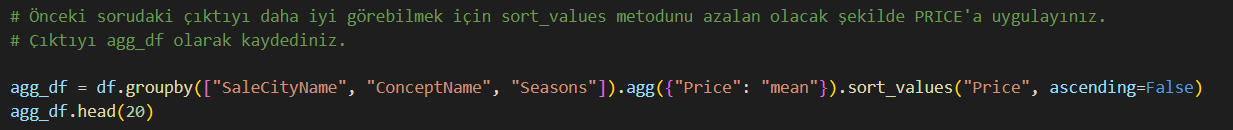
Bu sınıf işlemi EB\_Score adında bir değişkene kesilerek aktarılmış ve bir nevi yeni bir veri seti elde edilmiştir.

To\_excel metoduyla da excel kaydı oluşturulmuştur.

EB\_Score’u diğer değişkenlerle birlikte inceleyerek (gruplandırarak) anlamlandırılıyor.



GRUPLAMA VE SORT\_VALUES METODU KULLANIMI



Çıktının düzgün görünmesi ve okunaklılığın artırılması için groupby işleminden sonra nokta(.) kullanarak **sort\_values** metodu ekleniyor.

Sort\_Values kullanarak yapılan sıralama işlemini PRİCE değişkenine göre ayarlamak içinde parametre olarak giriliyor

Ascendig = artan

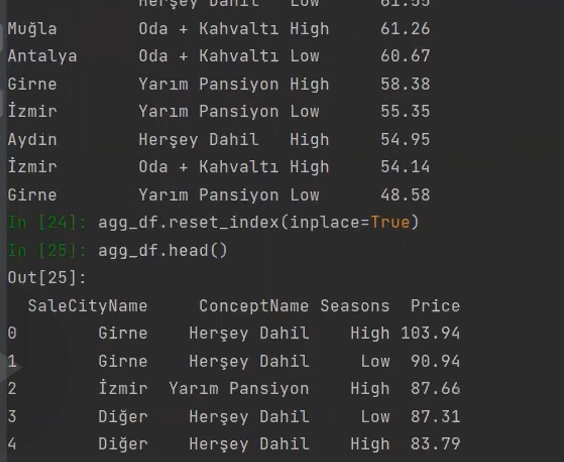
Demektir.

Ascendig = false yapılarak artandan azalana değilde (yüksekten - düşüğe)

Düşükten yükseğe olarak kullanılması isteniyor parametreyle

Eğer ascendig= true olsaydı artandan azalana doğru bir sıralama yapılırdı. Tercihe göre artık

..



İlk çıktıda değişkenler, muğla, Antalya gibi isimler ve bunları istemiyoruz.

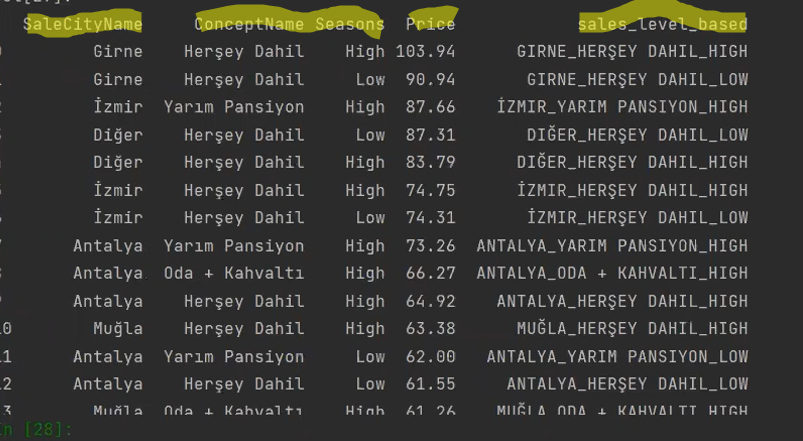
Burada yer alan isimleri index olarak ele almak istiyoruz.

metin, ekran görüntüsü, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu

Bunun için agg\_df’i yani kesilen değeri reset\_index metoduyla sıfırlıyoruz ve eski değerine çeviriyoruz

Fakat inplace’e dikkat etmek gerekiyor çünkü kalıcı bir etki bırakarak veri setini bozmak istemiyoruz



Görselde görüldüğü gibi 4 tane değişken var

Bizden istenen ise bu değişkenleri tek bir çatı altında toplamak ve bu değerleri keserek toplu ve yeni değişken elde etmek

ekran görüntüsü, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu

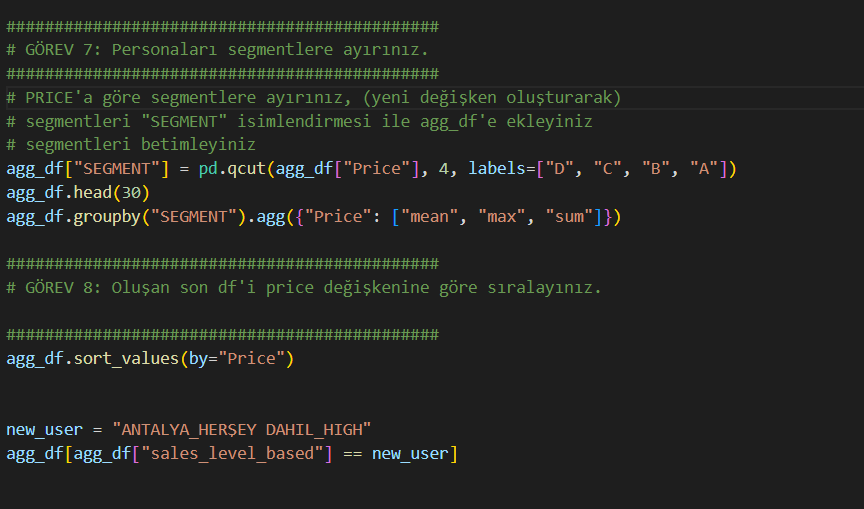
Agg\_df[‘sales\_level\_based’] DEĞŞKENİNİ agg\_df[[“SalesCityName”, “Conceptname”, “Seasons”]] ile aggregate (birleştirme) yap

Bu işlemi de verilen şartlara göre (lambda x: “ \_ ”.join(x).upper(), axis=1) ile yap

Diyor ki lambda da

X’lerden oluşuyor, Buradaki x’Leri alt tre ( \_ ) kullanarak birleştir

Birleştirirken de upper yap ve bunları da son olarak axis = 1 yani satırlara göre yap



Agg\_df[SEGMENT] yaparak yeni değişken oluşturuldu

Bu değişkene pd.qcut ile PRİCE değişkenindeki değerler sınıflandırılarak atandı

Bu sınıfın mean,max ve sum değerleri alındı

Sort\_values ile sıralandı

En son ise NEW\_USER değişkeni oluşturuldu

Yapılan işlemler bu değişkene atandı

Bu sayede persona oluşmuş oldu