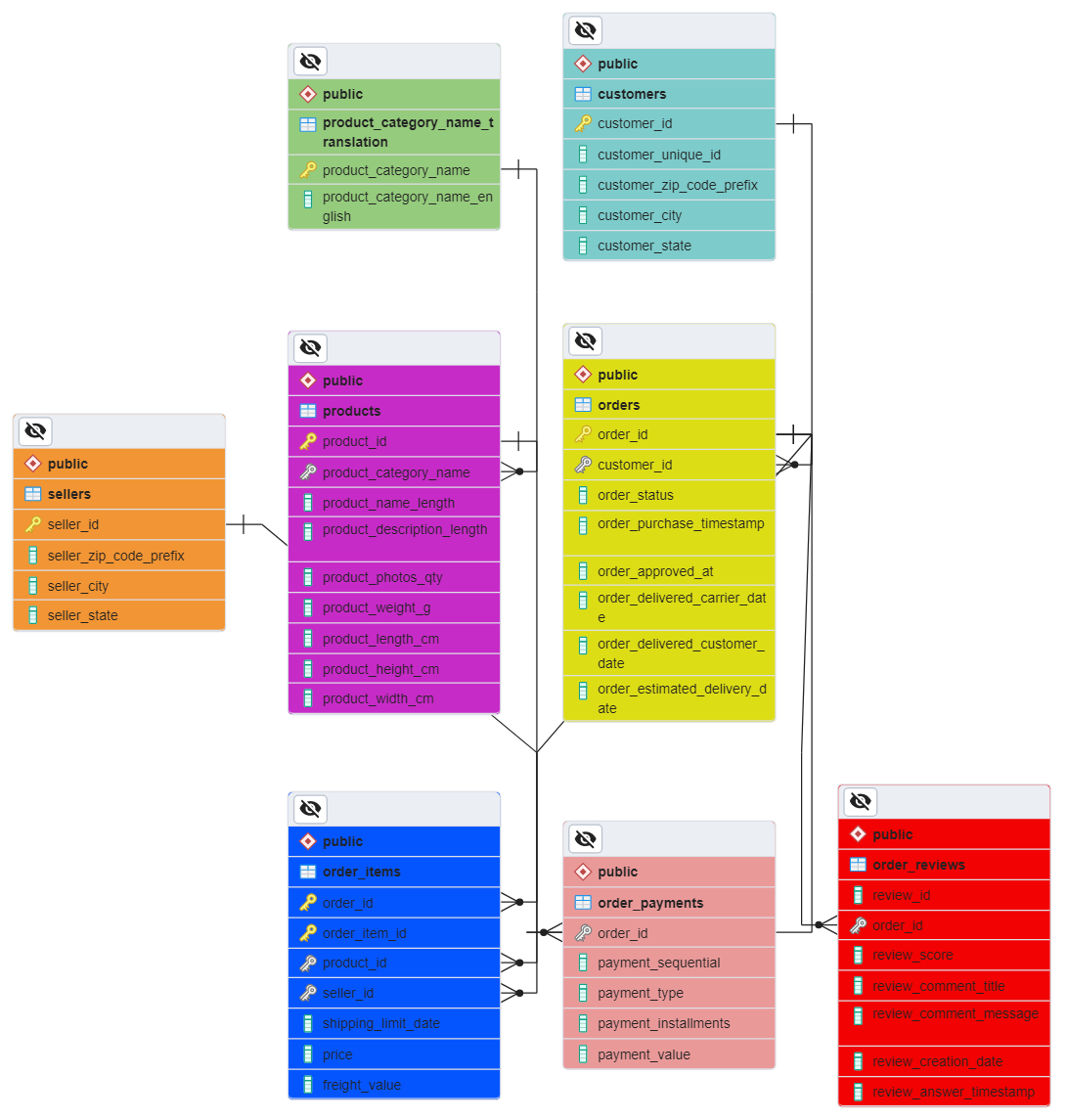
### **SQL PROJE**

Aşağıda veri tabanının ERD görseli eklenmiştir.

****

#### **Case 1 : Sipariş Analizi**

##### **Question 1 :** Aylık olarak order dağılımını inceleyiniz. Tarih verisi için order\_approved\_at kullanılmalıdır.

**SELECT**

**COALESCE(TO\_CHAR(order\_approved\_at, 'YYYY-MM'), 'Bilinmeyen Tarih') AS year\_month,**

**COUNT(\*) AS order\_count**

**FROM**

**orders**

**GROUP BY**

**year\_month**

**ORDER BY**

**year\_month;**

##### **Question 2 :** Aylık olarak order status kırılımında order sayılarını inceleyiniz. Sorgu sonucunda çıkan outputu excel ile görselleştiriniz. Dramatik bir düşüşün ya da yükselişin olduğu aylar var mı? Veriyi inceleyerek yorumlayınız.

**SELECT**

**COALESCE(TO\_CHAR(order\_approved\_at, 'YYYY-MM'), 'Bilinmeyen Tarih') AS year\_month,**

**order\_status,**

**COUNT(\*) AS order\_count**

**FROM**

**orders**

**GROUP BY**

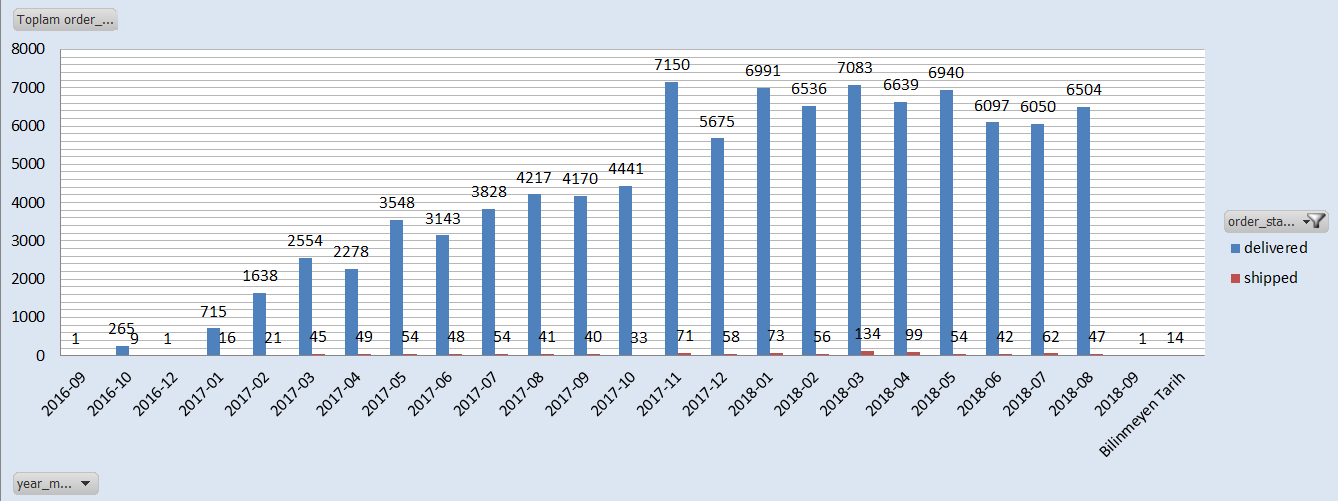
**year\_month, order\_status**

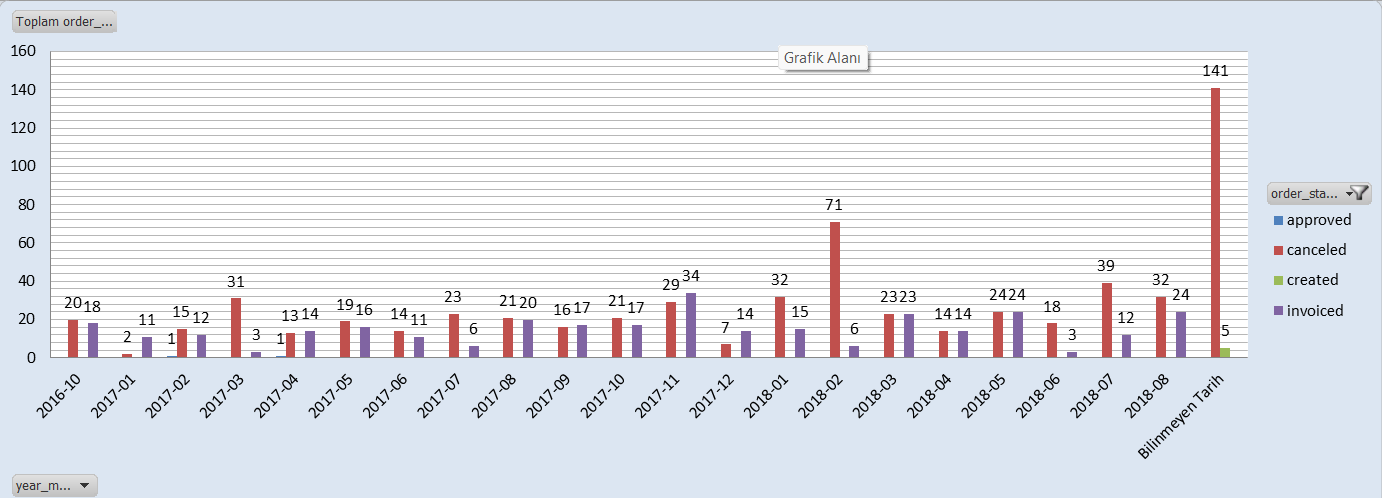
**ORDER BY**

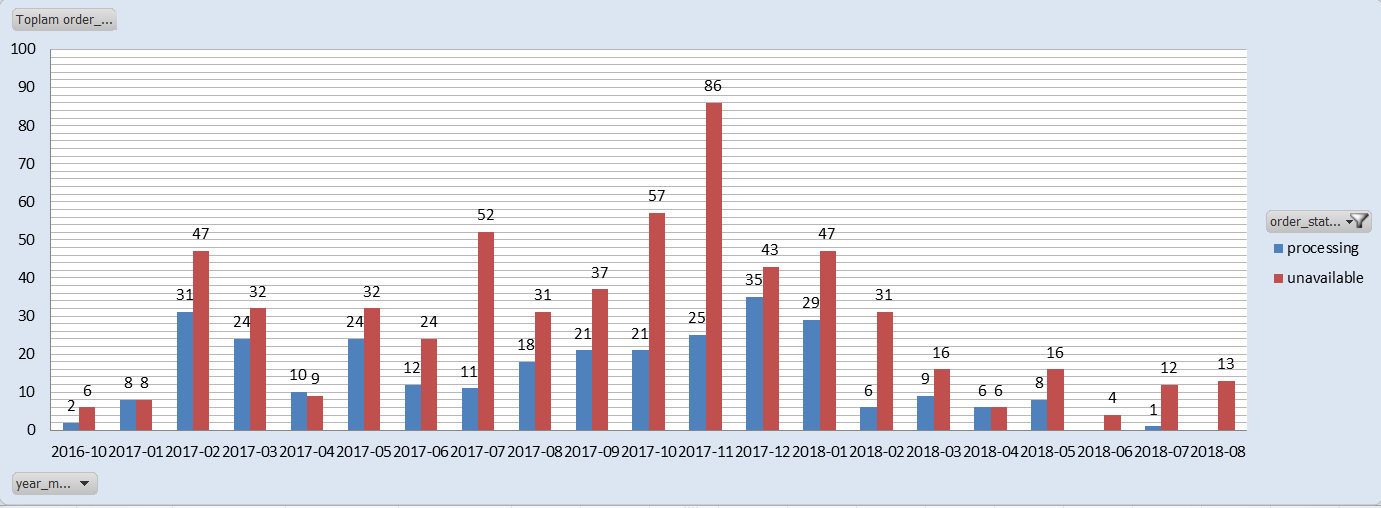
**year\_month, order\_status;**

* Gönderilen ve teslim edilen ürünlere baktığımız da 2017 yılında Kasım ayında daha fazla ürün gönderilip teslim edildiğini görüyoruz. Buradan 2017 yılının Kasım ayında satışın daha fazla olduğu çıkarımına varabiliriz. Bunun sebebi 15 Kasım’ın Bağımsızlık İlan edilme tarihi olabilir.
* Aynı tarihte ‘unavailable’ durumunda da bir artış görmekteyiz. Bunun sebebi stok tükenmiş veya ürün belli bir sınır üzerinde sipariş verilmiş olabilir.
* Grafiklerden çıkardığım izlenimde ‘unavailable’ durumunun genelde diğerlerinden daha fazla olduğudur. Stok durumuna göre veya başka bir sebepten kaynaklı ürünler için sipariş uygunluğu verilmemiş.

**\*\*\* Grafikler aşağıdadır.**

****

****

****

##### **Question 3 :** Ürün kategorisi kırılımında sipariş sayılarını inceleyiniz. Özel günlerde öne çıkan kategoriler nelerdir? Örneğin yılbaşı, sevgililer günü…

**SELECT**

**COALESCE(p.product\_category\_name, 'Unknown') AS product\_category\_name,**

**COUNT(oi.order\_id) AS order\_count**

**FROM**

**order\_items oi**

**LEFT JOIN**

**products p ON oi.product\_id = p.product\_id**

**GROUP BY**

**product\_category\_name**

**ORDER BY**

**order\_count DESC;**

* Bu sorgu sonucunda en çok sipariş verilen kategorinin 'cama\_mesa\_banho' olduğunu görüyoruz.
* Araştırmalarım sonucunda 2. ve 3. ayda karnaval kutlamaları yapılmaktaymış. 2-3 ayların toplam siparişlerine baktığım zaman ise 2017 yılında "moveis\_decoracao" kategorisinde daha çok sipariş alınmış.
* Ayrıca diğer özel günlere baktığım zamanda da "moveis\_decoracao" kategorisinde daha çok sipariş verildiğini görebiliriz.
* 2018 yılında ise "informatica\_acessorios" kategorisinde daha çok sipariş alınmış.

##### 

##### **Question 4 :** Haftanın günleri(pazartesi, perşembe, ….) ve ay günleri (ayın 1’i,2’si gibi) bazında order sayılarını inceleyiniz. Yazdığınız sorgunun outputu ile excel’de bir görsel oluşturup yorumlayınız.

**WITH date\_info AS (**

**SELECT**

**TO\_CHAR(order\_approved\_at, 'DY') AS day\_of\_week,**

**EXTRACT(DAY FROM order\_approved\_at) AS day\_of\_month,**

**COUNT(order\_id) AS order\_count**

**FROM**

**orders**

**WHERE**

**order\_approved\_at IS NOT NULL**

**GROUP BY**

**day\_of\_week, day\_of\_month**

**)**

**SELECT**

**day\_of\_week,**

**day\_of\_month,**

**COALESCE(order\_count, 0) AS order\_count**

**FROM**

**date\_info**

**ORDER BY**

**CASE**

**WHEN day\_of\_week = 'MON' THEN 1**

**WHEN day\_of\_week = 'TUE' THEN 2**

**WHEN day\_of\_week = 'WED' THEN 3**

**WHEN day\_of\_week = 'THU' THEN 4**

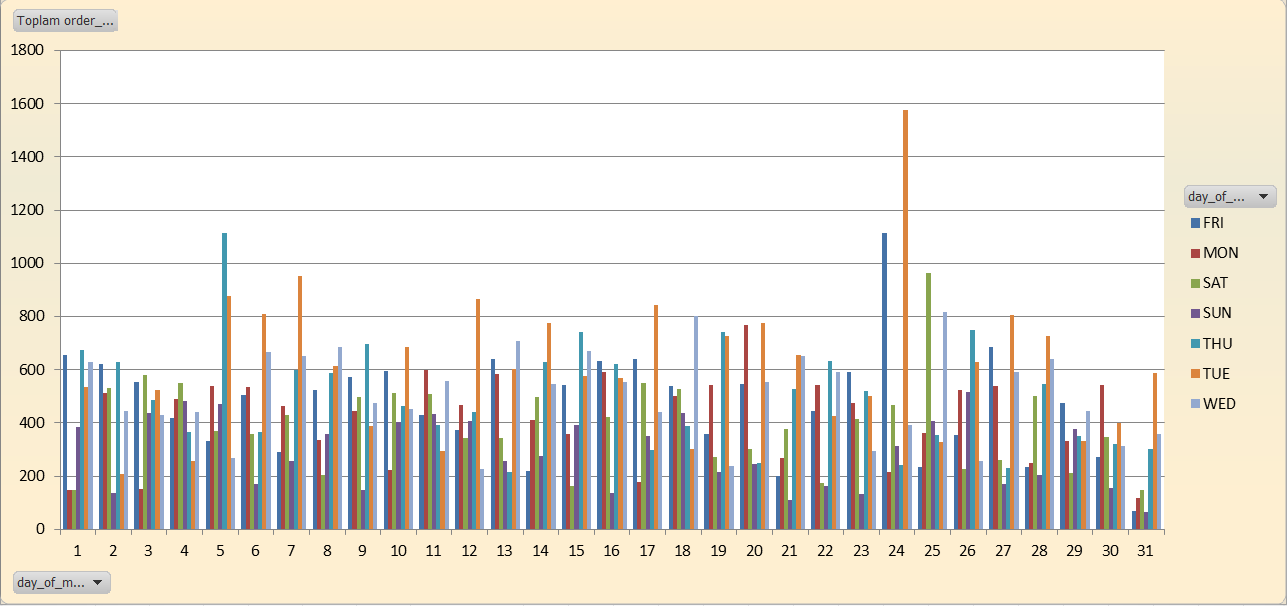
**WHEN day\_of\_week = 'FRI' THEN 5**

**WHEN day\_of\_week = 'SAT' THEN 6**

**WHEN day\_of\_week = 'SUN' THEN 7**

**END,**

**day\_of\_month;**



* Grafiği incelediğim zaman Brezilyalılar Salı günü daha çok sipariş veriyorlar. Ay bazında baktığım zaman da ayın 24 ve 25’inde siparişin arttığını görebiliyoruz. Bunun sebebinin 25 Aralık’ta Noel kutlamaları için verilen siparişler olduğunu düşünüyorum.

#### **Case 2 : Müşteri Analizi**

##### **Question 1 :** Hangi şehirlerdeki müşteriler daha çok alışveriş yapıyor? Müşterinin şehrini en çok sipariş verdiği şehir olarak belirleyip analizi ona göre yapınız.

--En çok sipariş verilen şehirler

**SELECT**

**customer\_city,**

**COUNT(DISTINCT o.order\_id) AS order\_count**

**FROM**

**public.orders o**

**JOIN**

**public.customers c ON o.customer\_id = c.customer\_id**

**GROUP BY**

**customer\_city**

**ORDER BY**

**order\_count DESC;**

---- Müşterinin şehrini en çok sipariş verdiği şehir

**WITH CustomerCityOrderCounts AS (**

**SELECT**

**c.customer\_unique\_id,**

**c.customer\_city,**

**COUNT(o.order\_id) AS order\_count**

**FROM**

**public.orders o**

**JOIN**

**public.customers c ON o.customer\_id = c.customer\_id**

**GROUP BY**

**c.customer\_unique\_id, c.customer\_city**

**)**

**SELECT**

**cco.customer\_unique\_id,**

**(ARRAY\_AGG(cco.customer\_city ORDER BY cco.order\_count DESC))[1] AS most\_order\_city,**

**MAX(cco.order\_count) AS max\_order\_count**

**FROM**

**CustomerCityOrderCounts cco**

**GROUP BY**

**cco.customer\_unique\_id**

**ORDER BY**

**max\_order\_count DESC;**

#### **Case 3: Satıcı Analizi**

##### **Question 1 :** Siparişleri en hızlı şekilde müşterilere ulaştıran satıcılar kimlerdir? Top 5 getiriniz. Bu satıcıların order sayıları ile ürünlerindeki yorumlar ve puanlamaları inceleyiniz ve yorumlayınız.

**SELECT**

**oi.seller\_id,**

**s.seller\_city,**

**AVG(EXTRACT(EPOCH FROM (o.order\_delivered\_customer\_date - o.order\_approved\_at)) / 86400) AS avg\_delivery\_days,**

**COUNT(DISTINCT o.order\_id) AS total\_orders**

**FROM**

**order\_items oi**

**JOIN**

**orders o ON oi.order\_id = o.order\_id**

**JOIN**

**sellers s ON oi.seller\_id = s.seller\_id**

**WHERE**

**o.order\_status = 'delivered'**

**GROUP BY**

**oi.seller\_id, s.seller\_city**

**HAVING**

**COUNT(DISTINCT o.order\_id) > 20**

**ORDER BY**

**avg\_delivery\_days**

**LIMIT 5;**

* Limitsiz olarak diğer satıcılarında ortalama teslimat sürelerini de incelediğim de ortalama olarak 20 ve üzeri ürün satanlar yedi gün içerisinde teslim ederken burada ki satıcılar beş gün içerisinde teslim ediyor.
* Yine aynı şekilde 50 ve üzeri siparişler de diğer satıcılar neredeyse dokuz gün içerisinde teslim ederken burada ki satıcılar daha kısa sürede teslim ediyorlar sattıkları ürün sayısına göre.
* Elimizde yeterince bilgi olmadığı için tam çıkarım yapmak doğru olmasa da araştırmalarım sonucunda bu satıcıların şehirlerinin São Paulo eyaletine bağlı olduğunu gördüm. Metropol olan bu eyalette kargo sisteminin daha çok geliştiğini ve birçok farklı kargo gönderim şirketinin olmasından kaynaklandığını düşünüyorum.

##### **Question 2 :** Hangi satıcılar daha fazla kategoriye ait ürün satışı yapmaktadır?

Fazla kategoriye sahip satıcıların order sayıları da fazla mı?

**SELECT**

**s.seller\_id,**

**s.seller\_city,**

**p.product\_category\_name,**

**COUNT(oi.order\_item\_id) AS total\_sales**

**FROM**

**order\_items oi**

**JOIN**

**products p ON oi.product\_id = p.product\_id**

**JOIN**

**sellers s ON oi.seller\_id = s.seller\_id**

**--WHERE**

**--s.seller\_id = '1f50f920176fa81dab994f9023523100'**

**GROUP BY**

**s.seller\_id, s.seller\_city, p.product\_category\_name**

**ORDER BY**

**total\_sales DESC;**

Yorum satırına aldığım WHERE şartı ile bir kaç satıcıyı kontrol ettiğim de fazla kategoriye sahip satıcıların order sayıları da fazla.

#### 

#### **Case 4 : Payment Analizi**

##### **Question 1 :** Ödeme yaparken taksit sayısı fazla olan kullanıcılar en çok hangi bölgede yaşamaktadır? Bu çıktıyı yorumlayınız.

**SELECT**

**c.customer\_state,**

**c.customer\_city,**

**AVG(op.payment\_installments) AS avg\_installments**

**FROM**

**order\_payments op**

**JOIN**

**orders o ON op.order\_id = o.order\_id**

**JOIN**

**customers c ON o.customer\_id = c.customer\_id**

**GROUP BY**

**c.customer\_state, c.customer\_city**

**ORDER BY**

**avg\_installments DESC**

**LIMIT 100;**

##### **Question 2 :** Ödeme tipine göre başarılı order sayısı ve toplam başarılı ödeme tutarını hesaplayınız. En çok kullanılan ödeme tipinden en az olana göre sıralayınız.

**SELECT**

**op.payment\_type,**

**COUNT(DISTINCT o.order\_id) AS successful\_order\_count,**

**SUM(op.payment\_value) AS total\_successful\_payment\_amount**

**FROM**

**order\_payments op**

**JOIN**

**orders o ON op.order\_id = o.order\_id**

**WHERE**

**o.order\_status = 'delivered'**

**AND op.payment\_type IS NOT NULL**

**GROUP BY**

**op.payment\_type**

**ORDER BY**

**total\_successful\_payment\_amount DESC;**

##### 

##### **Question 3 :** Tek çekimde ve taksitle ödenen siparişlerin kategori bazlı analizini yapınız. En çok hangi kategorilerde taksitle ödeme kullanılmaktadır?

**SELECT**

**COALESCE(pc.product\_category\_name, 'unknown') AS product\_category\_name,**

**op.payment\_type,**

**COUNT(DISTINCT oi.order\_id) AS order\_count,**

**SUM(oi.price) AS total\_order\_amount**

**FROM**

**order\_items oi**

**JOIN**

**order\_payments op ON oi.order\_id = op.order\_id**

**JOIN**

**products p ON oi.product\_id = p.product\_id**

**LEFT JOIN**

**product\_category\_name\_translation pc ON p.product\_category\_name = pc.product\_category\_name**

**WHERE**

**op.payment\_type IN ('credit\_card', 'boleto') --boleto tek çekim kabul edilmiştir.**

**GROUP BY**

**COALESCE(pc.product\_category\_name, 'unknown'), op.payment\_type**

**ORDER BY**

**product\_category\_name, order\_count DESC;**

* En çok taksit kullanılan kategori "cama\_mesa\_banho" kategorisidir.
* "beleza\_saude"
* "esporte\_lazer"

#### **Case 5 : RFM Analizi**

Aşağıdaki e\_commerce\_data\_.csv doyasındaki veri setini kullanarak RFM analizi yapınız.   
Recency hesaplarken bugünün tarihi değil en son sipariş tarihini baz alınız.

**Recency**

**SELECT**

**CustomerID,**

**MAX(InvoiceDate) AS LastPurchaseDate,**

**CURRENT\_TIMESTAMP AS CurrentDate,**

**CURRENT\_TIMESTAMP - MAX(InvoiceDate) AS Recency**

**FROM**

**rfm\_data**

**GROUP BY**

**CustomerID;**

**Frequency**

**SELECT**

**CustomerID,**

**COUNT(DISTINCT InvoiceNo) AS Frequency**

**FROM**

**rfm\_data**

**GROUP BY**

**CustomerID**

**ORDER BY**

**Frequency DESC;**

**Monetary**

**SELECT**

**CustomerID,**

**SUM(Quantity \* UnitPrice) AS MonetaryValue**

**FROM**

**rfm\_data**

**GROUP BY**

**CustomerID**

**ORDER BY**

**MonetaryValue DESC;**

Veri seti bu linkten alınmıştır, veriyi tanımak için linke girip inceleyebilirsiniz.  
  
[E-Commerce Data](https://www.kaggle.com/datasets/carrie1/ecommerce-data)