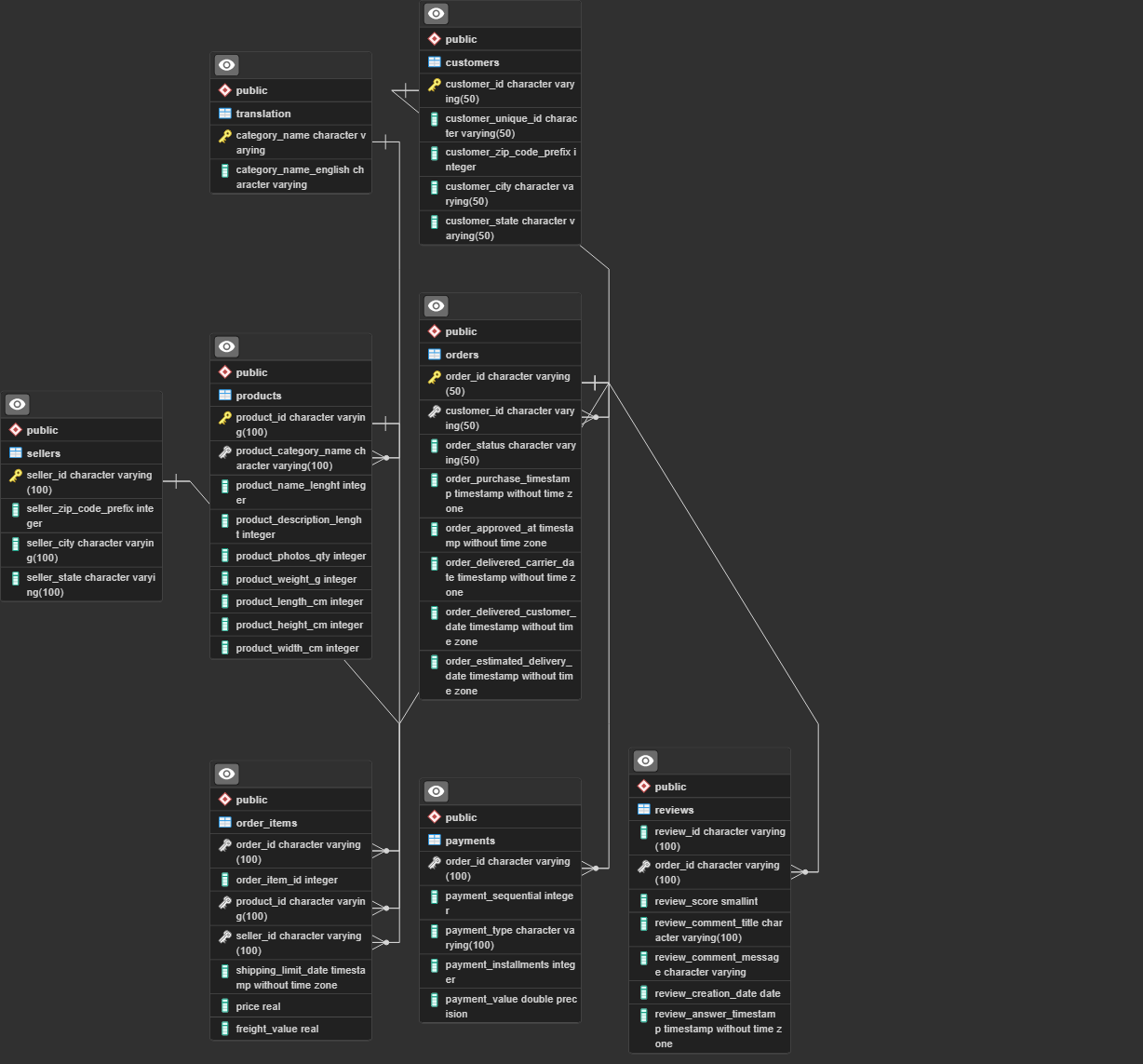
**ERD Schema**



**CASE 1 : Sipariş Analizi**

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

**NOT:**

**İptal edilen ve mevcutta bulunmayan ürünleri filtrelemedim.**

**Order\_approved\_at kolonunda boş olan fakat gönderilmiş veya hazırlanmış siparişler olduğu için null kolonu çıkıyor.**

SELECT

date\_trunc('month',order\_approved\_at)::date AS months,

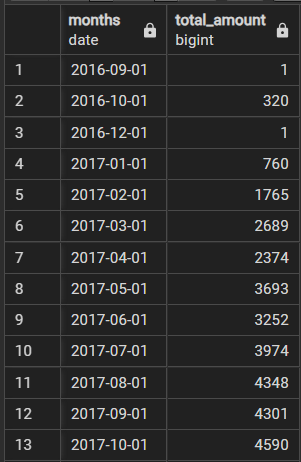
COUNT (\*)total\_amount

FROM

orders

GROUP BY

1;



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.iptal edilen veya yolda olanları ayrı gösterebilirsin

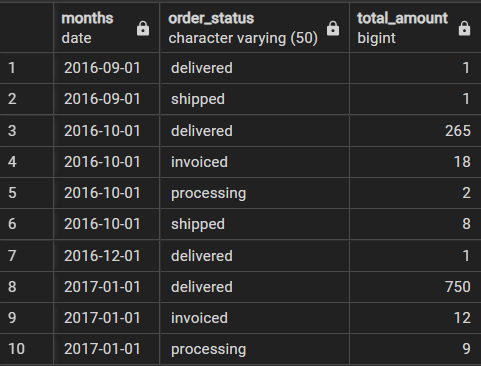
**NOT:**

**Bir önceki soruda belirttiğim gibi , order\_approved\_at kolonu boş olan, fakat gönderilmiş**

**veya hazırlanmış siparişler olduğu için order\_purchase\_timestamp kolonunun daha güvenilir**

**olacağını düşündüm.**

**BAŞARILI SİPARİŞLER İÇİN;**



SELECT

date\_trunc('month',order\_purchase\_timestamp)::date AS months,

order\_status,

COUNT (\*) total\_amount

FROM

orders

WHERE

order\_status NOT IN ('unavailable','canceled')

GROUP BY

1,2

ORDER BY

1;

**Elimizdeki verilere göre en çok 2017 Kasım ayında sipariş verilmiş.Diğer iki yılın Kasım verileri olmadığından bir karşılaştırma yapamıyoruz. Black Friday ve Black Friday’in dijital hali olan ,2012 yılından beri Brezilya’da Black Friday’in hemen ertesindeki Pazartesi kutlanan Cyber Monday gibi büyük indirim günlerinin Kasım ayında olmalarının siparişlerin artmasında etkileri olduğunu düşünüyorum.Bu iki büyük alışveriş festivalinin yanında Brezilya’da 12 kasım çocuk günü olarak kutlanıyor ve aileler bu tarihte çocuklarına hediyeler alıyor.Ayrıca Amerikan bayramı Halloween’in Brezilya versiyonu olan The Day of Saci de 31 kasımda gerçekleşiyor.Bu festival de kostüm,makyaj malzemesi tarzı ürünlerin siparişlerinin artmasına etki edebilir.**

**BAŞARISIZ SİPARİŞLER İÇİN;**

SELECT

date\_trunc('month',order\_purchase\_timestamp)::date as month,

order\_status,

COUNT (\*) total\_amount

FROM

orders

WHERE

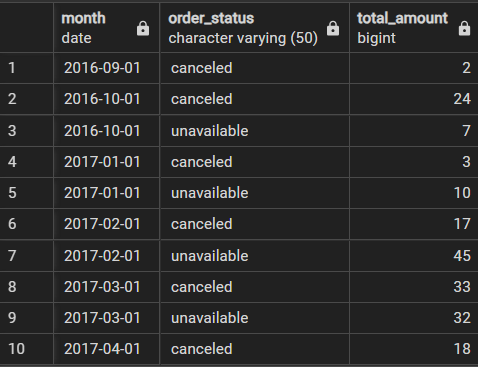
order\_status IN ('unavailable','canceled')

GROUP BY

1,2

ORDER BY

1;



**Elimizdeki verilere göre müşterilere en ‘yok’ denen ay , aynı zamanda en çok sipariş verilen kasım ayı.Buna yoğun siparişlere ürün yetiştirememek neden olmuş olabilir.En yüksek sipariş iptalinin olduğu dönem ise 2018 ağustos ayı.Bu 84 iptalin 53’ünün kategori bilgisi yok ve bu 53 ürünün 28’i de en çok sipariş verilen Sao Paolo bölgesinden.**

3 :Ürün kategorisi kırılımında sipariş sayılarını inceleyiniz. Özel günlerde öne çıkan kategoriler nelerdir?

Örneğin yılbaşı, sevgililer günü…Her kategoriden ne kadar sipariş var mesela yılbaşında

**NOT:**

**-Sipariş verme tarihini Black Friday'e göre aldım.Çünkü önceki soru da gördüğümüz üzere en çok sipariş verilen ay Kasım ve sonraki soruda görebileceğimiz gibi ayın 24’ü en çok sipariş verilen gün.Bu tarihte bize Black Friday’i veriyor.Brezilya’da Black Friday indirimleri yaklaşık 1 hafta sürdüğünü öğrendim.Bu sebeple sorgudaki filtrelememi buna göre yaptım.**

WITH BLACKFRIDAY AS (

SELECT

P.PRODUCT\_CATEGORY\_NAME,

T.CATEGORY\_NAME\_ENGLISH,

COUNT( O.ORDER\_ID)BLACKFRIDAY

FROM ORDERS O

JOIN

ORDER\_ITEMS OI ON O.ORDER\_ID = OI.ORDER\_ID

JOIN

PRODUCTS P ON OI.PRODUCT\_ID = P.PRODUCT\_ID

LEFT JOIN

TRANSLATION T ON P.PRODUCT\_CATEGORY\_NAME = T.CATEGORY\_NAME

WHERE EXTRACT(MONTH FROM ORDER\_PURCHASE\_TIMESTAMP) = 11

AND EXTRACT(DAY FROM ORDER\_PURCHASE\_TIMESTAMP) BETWEEN 20 AND 29

AND ORDER\_STATUS NOT IN ('unavailable','canceled')

GROUP BY

P.PRODUCT\_CATEGORY\_NAME,

T.CATEGORY\_NAME\_ENGLISH

ORDER BY

COUNT( O.ORDER\_ID) DESC

)

,REST AS (

SELECT

P.PRODUCT\_CATEGORY\_NAME,

T.CATEGORY\_NAME\_ENGLISH,

COUNT( O.ORDER\_ID)SALES\_NOT\_BF

FROM ORDERS O

JOIN

ORDER\_ITEMS OI ON O.ORDER\_ID = OI.ORDER\_ID

JOIN

PRODUCTS P ON OI.PRODUCT\_ID = P.PRODUCT\_ID

LEFT JOIN

TRANSLATION T ON P.PRODUCT\_CATEGORY\_NAME = T.CATEGORY\_NAME

WHERE EXTRACT(MONTH FROM ORDER\_PURCHASE\_TIMESTAMP) != 11

AND EXTRACT(DAY FROM ORDER\_PURCHASE\_TIMESTAMP) NOT BETWEEN 20 AND 29

AND ORDER\_STATUS NOT IN ('unavailable','canceled')

GROUP BY

P.PRODUCT\_CATEGORY\_NAME,

T.CATEGORY\_NAME\_ENGLISH

ORDER BY

COUNT( O.ORDER\_ID) DESC

)

,PERCENTAGE AS (

SELECT

BF.PRODUCT\_CATEGORY\_NAME,

BF.CATEGORY\_NAME\_ENGLISH,

ROUND (

(SUM(BF.BLACKFRIDAY) / SUM(BF.BLACKFRIDAY) OVER ()) \* 100,2) AS BLACKFRIDAY\_PERCENTAGE,

ROUND (

(SUM(REST.SALES\_NOT\_BF) / SUM(REST.SALES\_NOT\_BF) OVER ()) \* 100,2) AS SALES\_NOT\_BF\_PERCENTAGE

FROM

BLACKFRIDAY BF

JOIN

REST ON BF.PRODUCT\_CATEGORY\_NAME = REST.PRODUCT\_CATEGORY\_NAME

GROUP BY

BF.PRODUCT\_CATEGORY\_NAME,

BF.CATEGORY\_NAME\_ENGLISH,

BF.BLACKFRIDAY,

REST.SALES\_NOT\_BF

)

SELECT

PRODUCT\_CATEGORY\_NAME,

CATEGORY\_NAME\_ENGLISH,

BLACKFRIDAY\_PERCENTAGE-SALES\_NOT\_BF\_PERCENTAGE INCREASE\_PER

FROM

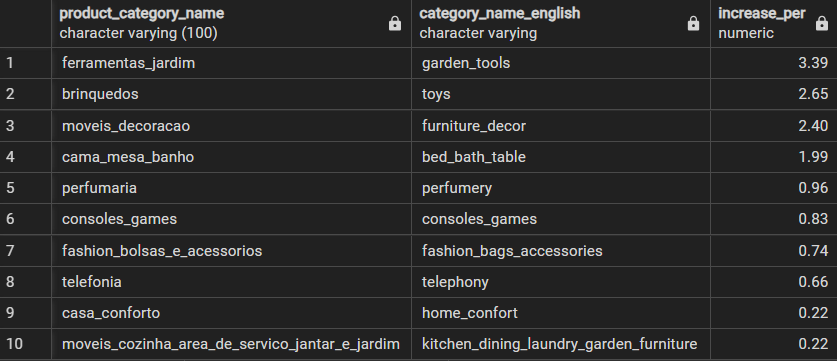
PERCENTAGE

WHERE

BLACKFRIDAY\_PERCENTAGE-SALES\_NOT\_BF\_PERCENTAGE>0

ORDER BY

BLACKFRIDAY\_PERCENTAGE-SALES\_NOT\_BF\_PERCENTAGE DESC;



**Bu sorgumda Black Friday’de satışı diğer günlere göre artan ürünleri ve artış oranlarını listeledim.Normal günlerde de zaten çok sipariş alan ürünlerin dışında daha keyfi ürünler olan toys,perfumery,consoles games,fashion bags and accessories kategorilerindeki artışlar görülüyor..Burada indirimlerin satışlara olumlu etkileri olduğunu söyleyebiliriz.**

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

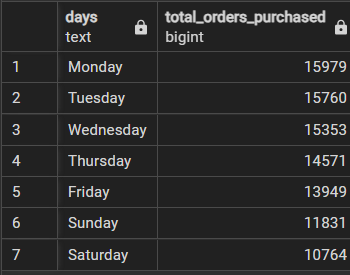
yorumlayınız.

**NOT :**

**-Siparişin verildiği tarihten aldım.**

**-İptal ve mevcutta bulunmayan siparişleri filtreledim.**

**-Soruda vurgulanan order sayısı olduğu için to\_char kullandım.**

SELECT

INITCAP(TO\_CHAR(ORDER\_PURCHASE\_TIMESTAMP,'day'))DAYS,

COUNT(DISTINCT ORDER\_ID) TOTAL\_ORDERS\_PURCHASED

FROM

ORDERS

WHERE

ORDER\_STATUS NOT IN ('canceled','unavailable')

GROUP BY

1

ORDER BY

2 DESC;

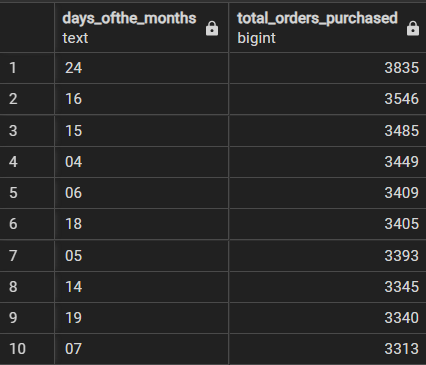
**Bu sorguda başarılı siparişlerin daha çok hafta içi verildiğini görüyoruz.Bunun nedeni insanların teslim süresini de düşünerek yeni aldıkları ürünleri hafta sonları kullanmak istemeleri olabilir veya insanlar hafta içi günlerde daha az sosyalleştikleri için ürün incelemelerini ve siparişlerini bu günler de yapıyor olabilirler.**

SELECT

TO\_CHAR(ORDER\_PURCHASE\_TIMESTAMP,'dd') DAYS\_OFTHE\_MONTHS,

COUNT(DISTINCT ORDER\_ID) TOTAL\_ORDERS\_PURCHASED

FROM

 ORDERS

WHERE

ORDER\_STATUS NOT IN ('canceled','unavailable')

GROUP BY

1

ORDER BY

2 DESC;

**Brezilya da Black Friday 22-29 kasım tarihleri arasında kutlanıyor.Ayrıca yılbaşı da ay sonu.Noel de ayın 24’ünde bu da gün içinde teslim edilebilecek yiyecek ve içeçek tarzı siparişleri etkileyebilir.En çok sipariş verilen bölge olan Sao Paolo’nun kurtuluşu da ayın 25’inde.Tüm bu özel günlere baktığımızda en çok sipariş verilen günün ayın 24'nde diğer günlere göre yüksek olması beklenen bir durum.**

**CASE 2 : Müşteri Analizi**

1 :Hangi şehirlerdeki müşteriler daha çok alışveriş yapıyor? Müşterinin şehrini en çok sipariş

verdiği şehir olarak belirleyip analizi ona göre yapınız.

Örneğin; Sibel Çanakkale’den 3, Muğla’dan 8 ve İstanbul’dan 10 sipariş olmak üzere 3 farklı

şehirden sipariş veriyor.Sibel’in şehrini en çok sipariş verdiği şehir olan İstanbul olarak

seçmelisiniz ve Sibel’in yaptığı siparişleri İstanbul’dan 21 sipariş vermiş şekilde görünmelidir.

**NOTLAR:**

**1.İptal edilen ve mevcutta bulunmayan siparişleri denklemden çıkardım.**

**2.ROW\_NUMBER() yerine RANK() kullandığımda bazı müşteriler farklı şehirlerden eşit sayıda sipariş verdikleri için o iki şehir de yazacaktı.Tek şehir görünmesi için Row kullandım.**

WITH CITY\_RANK AS(

SELECT

C.CUSTOMER\_UNIQUE\_ID,

C.CUSTOMER\_CITY,

COUNT(O.ORDER\_ID)TOTAL\_ORDER,

ROW\_NUMBER() OVER (PARTITION BY C.CUSTOMER\_UNIQUE\_ID ORDER BY COUNT(O.ORDER\_ID) DESC) RANKING

FROM

CUSTOMERS C

JOIN

ORDERS O ON O.CUSTOMER\_ID = C.CUSTOMER\_ID

WHERE

O.ORDER\_STATUS NOT IN ('unavailable','canceled')

GROUP BY

1,2

)

,TOTAL AS(

SELECT

CUSTOMER\_UNIQUE\_ID,

CUSTOMER\_CITY,

SUM(TOTAL\_ORDER)TOPLAM

FROM

CITY\_RANK

GROUP BY

1,2

ORDER BY

SUM(TOTAL\_ORDER) DESC

)

,COMBINE AS (

SELECT

CITY\_RANK.CUSTOMER\_UNIQUE\_ID,

CITY\_RANK.CUSTOMER\_CITY,

TOTAL.TOPLAM

FROM

CITY\_RANK

JOIN

TOTAL ON CITY\_RANK.CUSTOMER\_UNIQUE\_ID = TOTAL.CUSTOMER\_UNIQUE\_ID

WHERE

CITY\_RANK.RANKING = 1

)

SELECT

CUSTOMER\_CITY,

SUM(TOPLAM)HOP

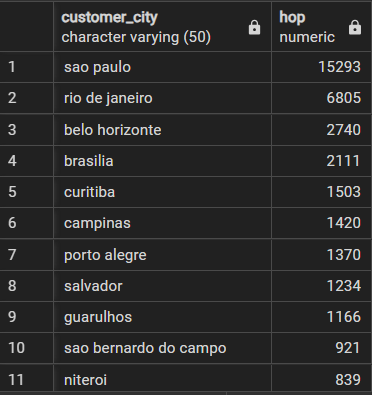
FROM COMBINE

GROUP BY

1

ORDER BY

2 DESC;



**CASE 3: Satıcı Analizi**

1 :Siparişleri en hızlı şekilde müşterilere ulaştıran satıcılar kimlerdir? Top 5 getiriniz.

Bu satıcıların order sayıları ile ürünlerindeki yorumlar ve puanlamaları inceleyiniz ve

yorumlayınız.

**Ortalama satış adetini 36 bulduğum için 36 siparişten fazla gönderim yapanları aldım.**

**Siparişin verildiği ve müşterinin eline ulaştığı zaman aralığını kullandım.**

**Tek order\_id de farklı satıcılar olduğunu bulduğum için order\_items tablosundaki siparişleri saydım.**

WITH DELIVERY\_TIMES AS(

SELECT

OI.SELLER\_ID,

ROUND(AVG(O.ORDER\_DELIVERED\_CUSTOMER\_DATE::date - O.ORDER\_PURCHASE\_TIMESTAMP::date),2)ORDER\_DELIVERY\_TIME

FROM

ORDERS O

JOIN

ORDER\_ITEMS OI ON O.ORDER\_ID = OI.ORDER\_ID

GROUP BY

OI.SELLER\_ID

HAVING

COUNT(DISTINCT OI.ORDER\_ID) > 36

ORDER BY

ORDER\_DELIVERY\_TIME

)

SELECT

DT.SELLER\_ID,

ORDER\_DELIVERY\_TIME,

COUNT(DISTINCT OI.ORDER\_ID)ORDER\_TOTAL,

ROUND(AVG(R.REVIEW\_SCORE),2)REV\_SCORE,

COUNT(R.REVIEW\_COMMENT\_TITLE)COMMENT\_COUNT,

COUNT(R.REVIEW\_COMMENT\_MESSAGE)MESSAGE\_COUNT

FROM

DELIVERY\_TIMES DT

JOIN

ORDER\_ITEMS OI ON DT.SELLER\_ID = OI.SELLER\_ID

JOIN

REVIEWS R ON OI.ORDER\_ID = R.ORDER\_ID

GROUP BY

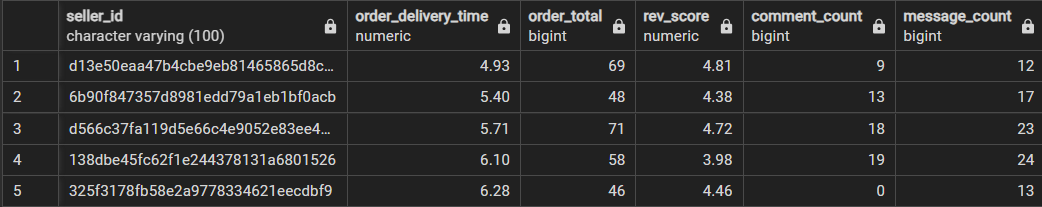
DT.SELLER\_ID,2

ORDER BY

ORDER\_DELIVERY\_TIME

LIMIT

5;



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

Fazla kategoriye sahip satıcıların order sayıları da fazla mı?

**NOT:**

**Kategorisi boş olmayanları ele aldım**

**İptal edilen ve mevcutta bulunmayan siparişleri denklemden çıkardım.**

**1 order\_id de 1 den çok kategori olabiliyor.Örnek order\_id’ler '005d9a5423d47281ac463a968b3936fb',**

**'0c0aeabd6ae9d7a303af615334a220d0','0f20a3721e8997ae2f771ec8cf0c4d3d' gibi....Bu sebeple distinct kullanmadan sayım yapmanın daha doğru olacağını düşündüm.**

WITH CATEGORY\_SALES AS (

SELECT

OI.SELLER\_ID,

P.PRODUCT\_CATEGORY\_NAME,

COUNT(O.ORDER\_ID)ORDER\_TOTAL

FROM

ORDER\_ITEMS OI

JOIN

ORDERS O ON OI.ORDER\_ID = O.ORDER\_ID

JOIN

PRODUCTS P ON P.PRODUCT\_ID = OI.PRODUCT\_ID

WHERE

ORDER\_STATUS NOT IN ('unavailable','canceled')

GROUP BY

OI.SELLER\_ID,

P.PRODUCT\_CATEGORY\_NAME

ORDER BY

OI.SELLER\_ID,

ORDER\_TOTAL DESC

)

SELECT

SELLER\_ID,

COUNT(PRODUCT\_CATEGORY\_NAME)CATEGORY\_COUNT,

SUM(ORDER\_TOTAL)ORDERS\_SUM,

ROUND((SUM(ORDER\_TOTAL)/COUNT(DISTINCT PRODUCT\_CATEGORY\_NAME)),2) as sales\_rate\_per\_category

FROM

CATEGORY\_SALES

WHERE

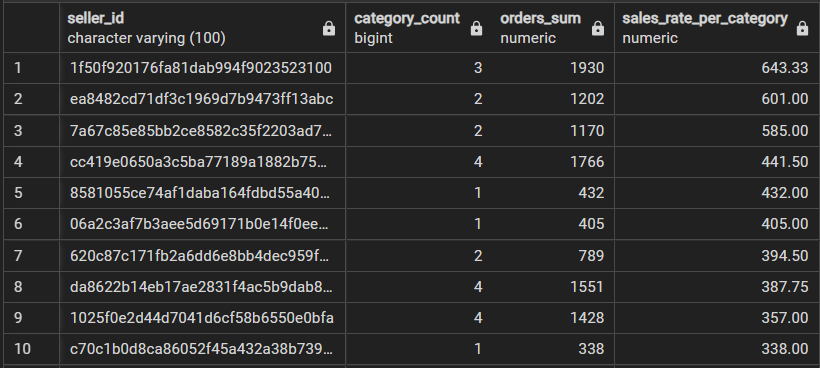
PRODUCT\_CATEGORY\_NAME IS NOT NULL

GROUP BY

SELLER\_ID

ORDER BY

4 DESC;



**Bu sorguyu toplam satışa göre getirdiğimizde en fazla başarılı satış yapanların kategori sayıları , bu sorguda en çok kategoriye sahip olan satıcılardan oldukça az.Burdan anlayabileceğimiz satış adetinin kategori sayısında bağımsız olduğudur.Daha çok kategoriye sahip olanların daha fazla satış yaptığını söyleyemeyiz.Ayrıca kategori başına ürün satışlarında da bu belli oluyor.Listemizin ilk 10 sırasındaki satıcıların kategorileri diğer satıcılara göre az ama kategori başına satış adetleri fazla.**

**CASE 4 : Payment Analizi**

1 :Ödeme yaparken taksit sayısı fazla olan kullanıcılar en çok hangi bölgede yaşamaktadır?

Bu çıktıyı yorumlayınız.

**ORTALAMA TAKSİT SAYISINI 3 DEN FAZLASINI ALDIM.**

SELECT

C.CUSTOMER\_STATE,

COUNT(DISTINCT O.ORDER\_ID)TOTAL\_DISTINCT\_ORDER

FROM

PAYMENTS P

JOIN

ORDERS O ON P.ORDER\_ID = O.ORDER\_ID

JOIN

CUSTOMERS C ON O.CUSTOMER\_ID = C.CUSTOMER\_ID

WHERE

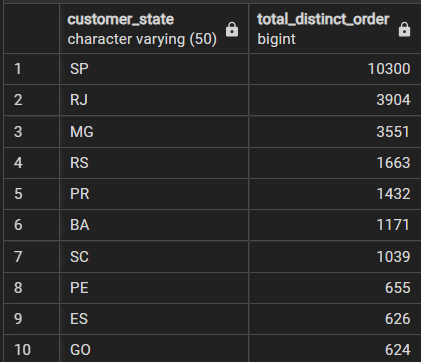
P.PAYMENT\_INSTALLMENTS > 3

GROUP BY

1

ORDER BY

2 DESC;



**Taksitli alışverişi daha çok kullanan bölgeler aynı zamanda ülkenin nüfusunun en yoğun olduğu ve alışverişin daha çok yapıldığı bölgeler,az olan bölgeler ise nüfusun az olduğu bölgeler.Burada taksit ve alışveriş adetlerinin nüfus ile doğru orantılı olduğunu görebiliyoruz.**

**Fakat bence aşağıda yazdığım sorgudan daha sağlıklı bilgi edinebiliriz.Bölgelerin pesin ve taksitli ödemelerine bakıp o bölgenin ekonomik durumu hakkında daha rahat yorum yapabiliriz.Örneğin yukarıda 3 taksitten fazla taksitle ödeme yapanlar SP bölgesinde en çok.Fakat aşağıda gördüğümüz gibi yüzde olarak 3 taksitten daha az kullanan bölge yani taksitsiz veya 3 taksitten az alışverişin daha fazla olduğu bölge.**

WITH TOPLAM AS(

SELECT

C.CUSTOMER\_STATE,

COUNT(DISTINCT O.ORDER\_ID)TOTAL\_DISTINCT\_ORDER

FROM

PAYMENTS P

JOIN

ORDERS O ON P.ORDER\_ID = O.ORDER\_ID

JOIN

CUSTOMERS C ON O.CUSTOMER\_ID = C.CUSTOMER\_ID

GROUP BY

1

ORDER BY

2 DESC

),

PESIN AS(

SELECT

C.CUSTOMER\_STATE,

COUNT(DISTINCT O.ORDER\_ID)PESIN

FROM

PAYMENTS P

JOIN

ORDERS O ON P.ORDER\_ID = O.ORDER\_ID

JOIN

CUSTOMERS C ON O.CUSTOMER\_ID = C.CUSTOMER\_ID

WHERE

P.PAYMENT\_INSTALLMENTS <3

GROUP BY

1

ORDER BY

2 DESC

),

TAKSIT AS(

SELECT

C.CUSTOMER\_STATE,

COUNT(DISTINCT O.ORDER\_ID)TAKSIT

FROM

PAYMENTS P

JOIN

ORDERS O ON P.ORDER\_ID = O.ORDER\_ID

JOIN

CUSTOMERS C ON O.CUSTOMER\_ID = C.CUSTOMER\_ID

WHERE

P.PAYMENT\_INSTALLMENTS > 3

GROUP BY

1

ORDER BY

2 DESC

)

SELECT

T.CUSTOMER\_STATE,

ROUND((PESIN \* 1.0 / TOTAL\_DISTINCT\_ORDER \* 1.0) \* 100,2) LESS\_THAN\_3,

ROUND((TAKSIT \* 1.0 / TOTAL\_DISTINCT\_ORDER \* 1.0) \* 100,2) MORE\_THAN\_3

FROM

TOPLAM T

JOIN

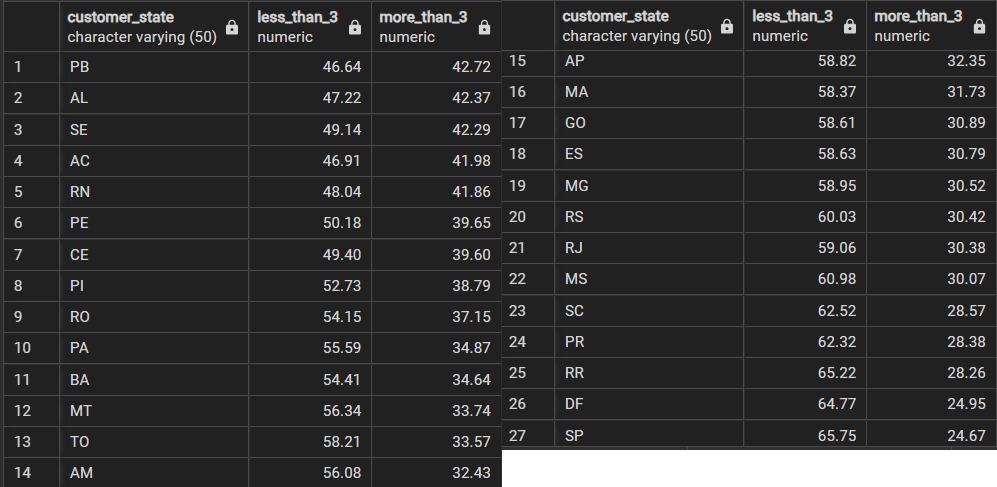
PESIN P ON T.CUSTOMER\_STATE = P.CUSTOMER\_STATE

JOIN

TAKSIT TA ON P.CUSTOMER\_STATE = TA.CUSTOMER\_STATE

ORDER BY

MORE\_THAN\_3 DESC;



2 :Ödeme tipine göre başarılı order sayısı ve toplam başarılı ödeme tutarını hesaplayınız.

En çok kullanılan ödeme tipinden en az olana göre sıralayınız.

**Her bir order tek bir ödeme yöntemiyle ödenmiş**

SELECT

P.PAYMENT\_TYPE,

COUNT(DISTINCT O.ORDER\_ID)ORDER\_COUNT,

ROUND(SUM(P.PAYMENT\_VALUE::decimal),2)TOTAL\_PAYMENT

FROM

PAYMENTS P

JOIN

ORDERS O ON P.ORDER\_ID = O.ORDER\_ID

WHERE

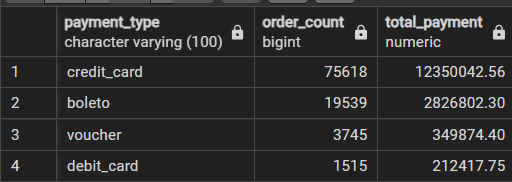
ORDER\_STATUS NOT IN ('unavailable','canceled')

GROUP BY

1

ORDER BY

2 DESC;



3 :Tek çekimde ve taksitle ödenen siparişlerin kategori bazlı analizini yapınız. En çok hangi

kategorilerde taksitle ödeme kullanılmaktadır?

**NOT:**

**Payment sequential kolonu aylık bazda değil ödeme bazlı taksit.Her ay veya belli dönemlerde değil ama ödemeyi bölerek yapılan ödemeler.Payment installments ise her ay düzenli ödeme olduğu için taksit kolonu olarak payment installments’ı aldım.Ayrıca github ve medium tarzı sitelerde de yaptığım araştırmalarda taksit için yine payment installments kullanılmış.**

WITH PAYMENT\_STYLE AS (

SELECT

PR.PRODUCT\_CATEGORY\_NAME,

T.CATEGORY\_NAME\_ENGLISH,

P.PAYMENT\_INSTALLMENTS,

COUNT(DISTINCT O.ORDER\_ID) TOTAL\_ORDERS,

CASE

WHEN PAYMENT\_INSTALLMENTS <= 1 THEN 'Tek Çekim'

ELSE 'Taksitli'

END AS PAY\_TYPE,

DENSE\_RANK() OVER (PARTITION BY CASE WHEN PAYMENT\_INSTALLMENTS <= 1 THEN 'Tek Çekim'

ELSE 'Taksitli'

END ORDER BY COUNT(DISTINCT O.ORDER\_ID) DESC) AS RANKING

FROM

PAYMENTS P

JOIN

ORDERS O ON P.ORDER\_ID = O.ORDER\_ID

JOIN

ORDER\_ITEMS OI ON O.ORDER\_ID = OI.ORDER\_ID

JOIN

PRODUCTS PR ON PR.PRODUCT\_ID = OI.PRODUCT\_ID

LEFT JOIN

TRANSLATION T ON PR.PRODUCT\_CATEGORY\_NAME = T.CATEGORY\_NAME

WHERE

O.ORDER\_STATUS NOT IN ('unavailable','canceled')

GROUP BY

1,2,3

)

SELECT

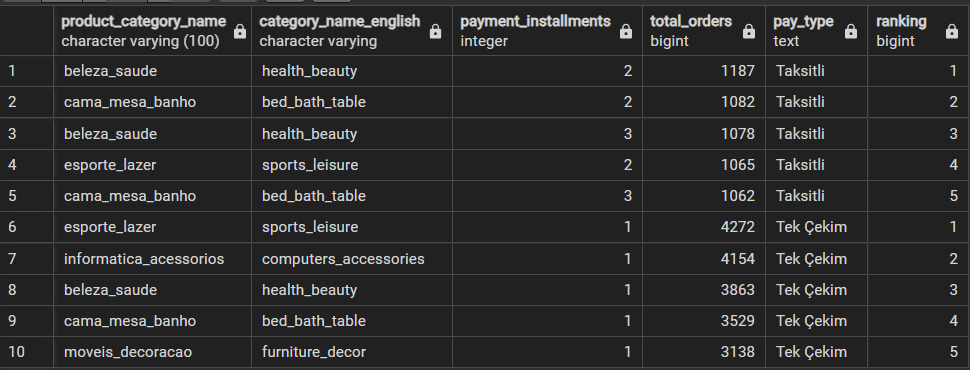
\*

FROM

PAYMENT\_STYLE

WHERE

RANKING BETWEEN 1 AND 5;

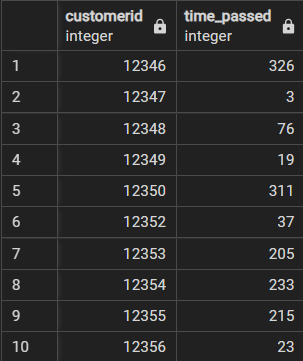


**En çok taksit seçeneği kullanılan kategori health\_beauty , en çok peşin ödenen kategori ise sport\_leisure**

**RFM ANALİZİ**

**Burada amacımız kullanıcıların sipariş alışkanlıklarını anlamak ve değerli kullanıcıları belirlemek olduğu için eksi değerleri ve customerid kolonu boş olan kullanıcıları analize dahil etmiyorum.**

**RECENCY :**



SELECT

CUSTOMERID,

'2011-12-10' - MAX(INVOICEDATE::date)TIME\_PASSED

FROM

RFM

WHERE

CUSTOMERID IS NOT NULL

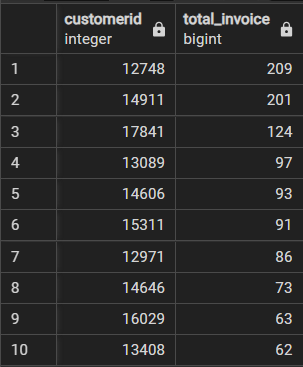
AND UNITPRICE \* QUANTITY > 0

GROUP BY

CUSTOMERID

ORDER BY

CUSTOMERID;

**FREQUENCY :**

SELECT

CUSTOMERID,

COUNT (DISTINCT INVOICENO)TOTAL\_INVOICE

FROM

RFM

WHERE

CUSTOMERID IS NOT NULL

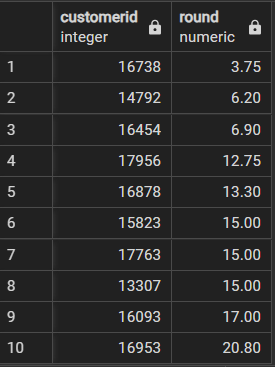
AND UNITPRICE \* QUANTITY > 0

GROUP BY

CUSTOMERID

ORDER BY

2 DESC;

**MONETARY:**

SELECT

CUSTOMERID,

ROUND(SUM(UNITPRICE \* QUANTITY)::decimal,2)

FROM

RFM

WHERE CUSTOMERID IS NOT NULL

AND UNITPRICE \* QUANTITY > 0

GROUP BY

CUSTOMERID

ORDER BY

2;

**RFM :**

**WITH RFM\_VALUES AS (**

**SELECT**

**CUSTOMERID,**

**'2011-12-10' - MAX(INVOICEDATE::date)RECENCY,**

**COUNT (DISTINCT INVOICENO)FREQUENCY,**

**ROUND(SUM(UNITPRICE \* QUANTITY)::decimal,2)MONETARY**

**FROM**

**RFM**

**WHERE CUSTOMERID IS NOT NULL**

**AND UNITPRICE \* QUANTITY > 0**

**GROUP BY**

**CUSTOMERID)**

**,RFM\_RATES AS (**

**SELECT**

**CUSTOMERID,**

**RECENCY,**

**NTILE(5) OVER (ORDER BY RECENCY DESC) RECENCY\_SCORE,**

**FREQUENCY,**

**NTILE(5) OVER (ORDER BY FREQUENCY)FREQUENCY\_SCORE,**

**MONETARY,**

**NTILE(5) OVER (ORDER BY MONETARY)MONETARY\_SCORE**

**FROM**

**RFM\_VALUES)**

**,RFM\_ANALYSIS AS (**

**SELECT**

**CUSTOMERID,**

**RECENCY\_SCORE,**

**FREQUENCY\_SCORE,**

**MONETARY\_SCORE,**

**CONCAT(RECENCY\_SCORE,FREQUENCY\_SCORE,MONETARY\_SCORE) RFM\_SCORE**

**FROM**

**RFM\_RATES**

**ORDER BY**

**CUSTOMERID)**

**SELECT \*,**

**CASE**

**WHEN (RECENCY\_SCORE = 1 OR RECENCY\_SCORE = 2) AND (FREQUENCY\_SCORE = 1 OR FREQUENCY\_SCORE = 2) THEN 'Hibernating'**

**WHEN (RECENCY\_SCORE = 1 OR RECENCY\_SCORE = 2) AND (FREQUENCY\_SCORE = 3 OR FREQUENCY\_SCORE = 4) THEN 'At Risk'**

**WHEN (RECENCY\_SCORE = 1 OR RECENCY\_SCORE = 2) AND FREQUENCY\_SCORE = 5 THEN 'Cant Loose'**

**WHEN RECENCY\_SCORE = 3 AND (FREQUENCY\_SCORE = 1 OR FREQUENCY\_SCORE = 2) THEN 'About\_to\_Sleep'**

**WHEN RECENCY\_SCORE = 3 AND FREQUENCY\_SCORE = 3 THEN 'Need\_Attention'**

**WHEN (RECENCY\_SCORE = 3 OR RECENCY\_SCORE = 4) AND (FREQUENCY\_SCORE = 4 OR FREQUENCY\_SCORE = 5) THEN 'Loyal\_Customers'**

**WHEN RECENCY\_SCORE = 4 AND FREQUENCY\_SCORE = 1 THEN 'Promising'**

**WHEN RECENCY\_SCORE = 5 AND FREQUENCY\_SCORE = 1 THEN 'New\_Customers'**

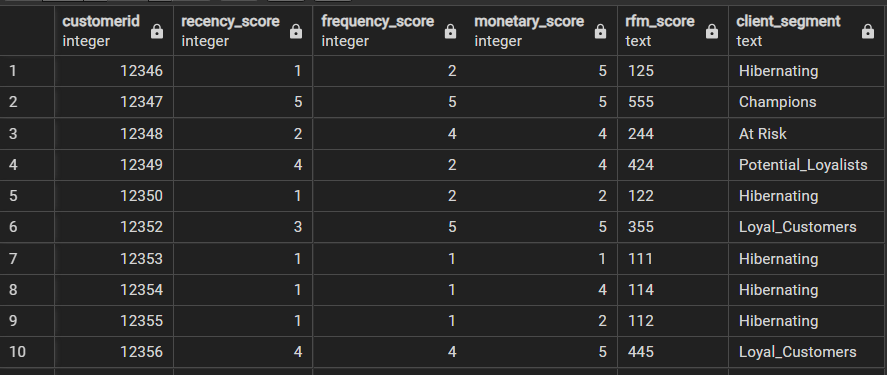
**WHEN (RECENCY\_SCORE = 4 OR RECENCY\_SCORE = 5) AND (FREQUENCY\_SCORE = 2 OR FREQUENCY\_SCORE = 3) THEN 'Potential\_Loyalists'**

**WHEN RECENCY\_SCORE = 5 AND (FREQUENCY\_SCORE = 4 OR FREQUENCY\_SCORE = 5) THEN 'Champions'**

**END AS CLIENT\_SEGMENT**

**FROM**

**RFM\_ANALYSIS;**

****

**NOT:**

**NTILE() eşit böldüğü için frequency değerleri aynı olmasına rağmen bir noktadan itibaren frequency scoreları farklı çıkabiliyor.**