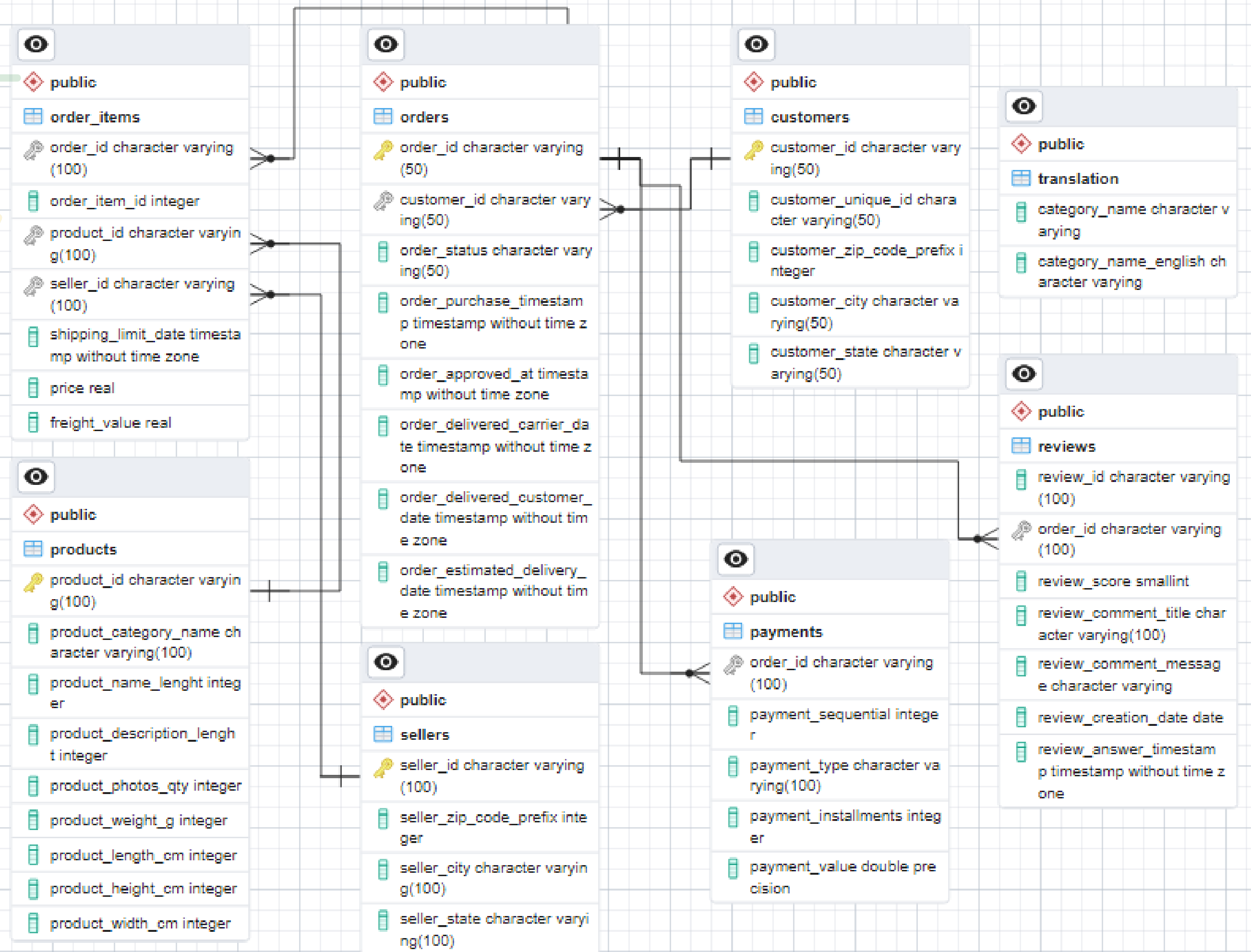


BRAZILIAN E-COMMERCE DATASET



CASE 1 : SİPARİŞ ANALIZI

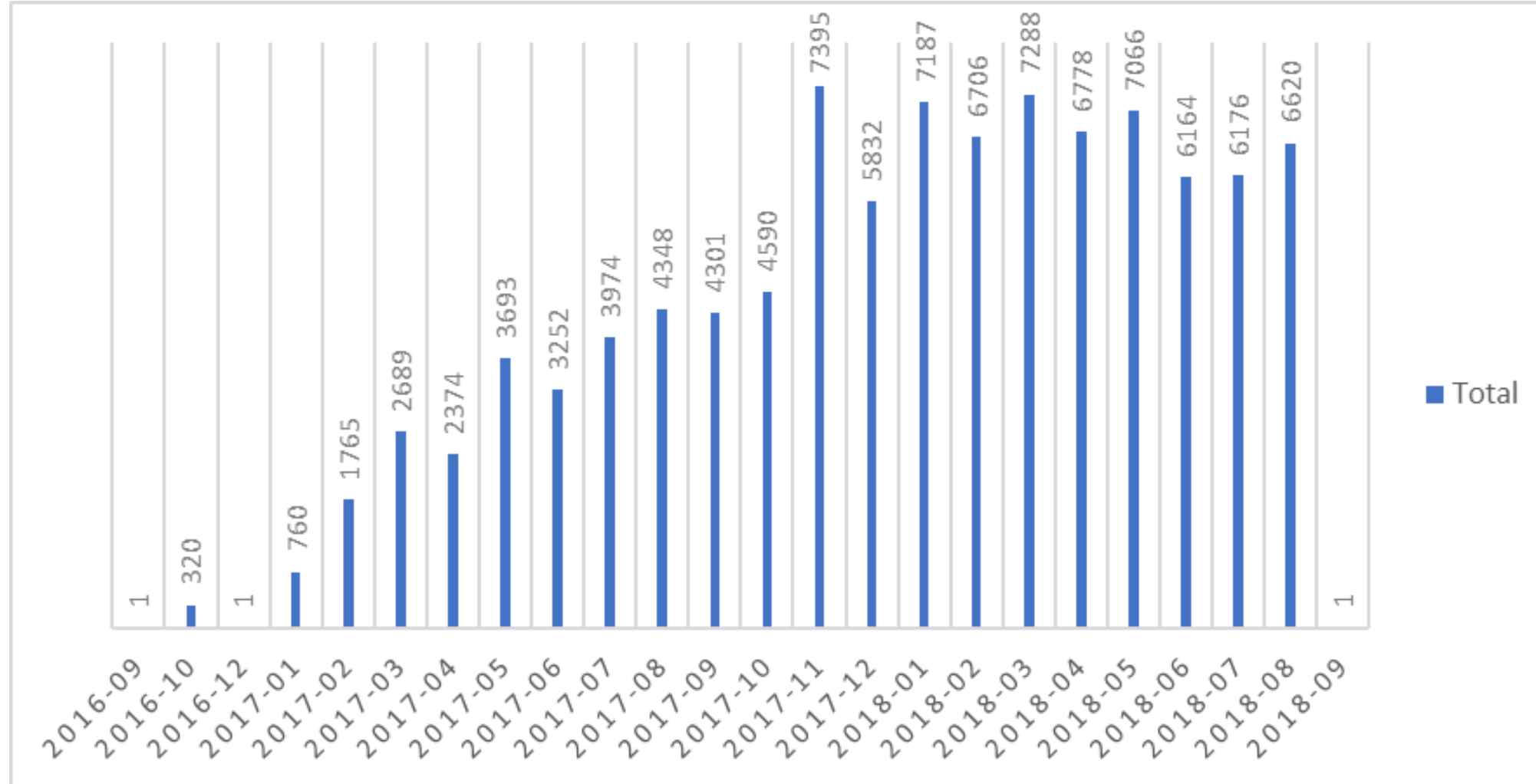
QUESTION 1 :

AYLIK OLARAK ORDER DAĞILIMINI INCELEYİNİZ. TARİH VERİSİ İÇİN ORDER_APPROVED_AT KULLANILMALIDIR.

SELECT TO_CHAR(ORDER_APPROVED_AT,'YYYY-MM') AS ORDER_MONTHLY,

COUNT(ORDER_ID) AS ORDER_COUNT

FROM ORDERS WHERE ORDER_APPROVED_AT IS NOT NULL GROUP BY ORDER_MONTHLY



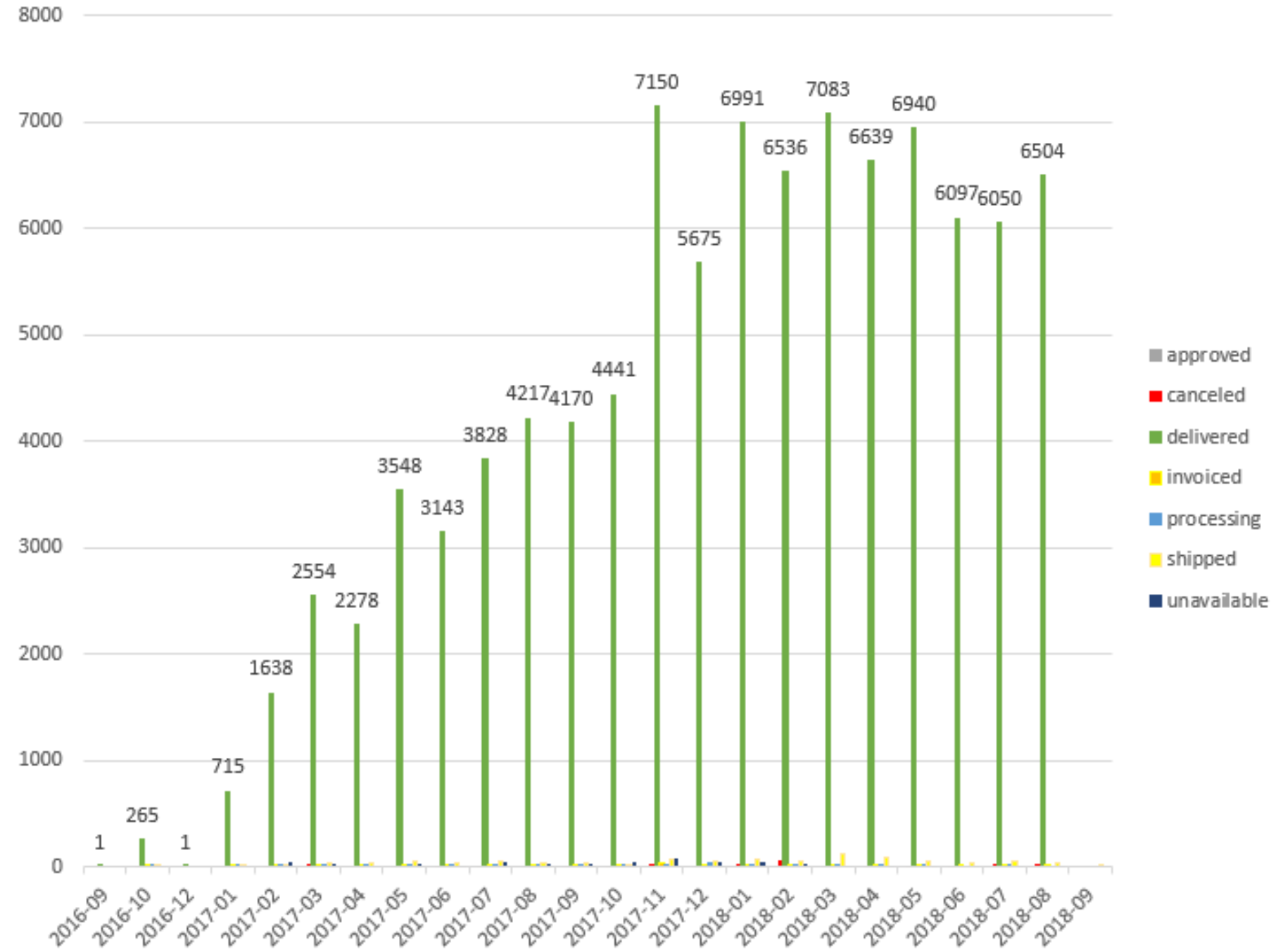
Aylık olarak sipariş dağılımlarını incelediğimde E-ticaret sürecinde 2016 4.çeyrekte sonrasıl olarak 2017 yılından itibaren ivmelendiğini söyleyebilirim ve bu yıllar geçtikçe artmaya devam etmiştir.

En çok siparişin alındığı 2017-11.ay ise o sene 11.aya ait önemli alışveriş günleri için kampanyalar, reklamlar, kuponlar ile satışların çok olmasına etkili çalışmaların bariz şekilde yapıldığı görülmektedir. Bir sonraki ay itibarıyla normal artış seyrinde artarak sipariş adetleri değişmiştir.

QUESTION 2 :

AYLIK OLARAK ORDER STATUS KIRILIMINDA ORDER SAYILARINI INCELEYINIZ. SORGU SONUCUNDA ÇIKAN OUTPUTU EXCEL İLE GÖRSELLEŞTİRİNİZ.DRAMATİK BİR DÜŞÜŞÜN YA DA YÜKSELİŞİN OLDUĞU AYLAR VAR MI? VERİYİ İNCELEYEREK YORUMLAYINIZ.

```
SELECT TO_CHAR(ORDER_APPROVED_AT,'YYYY-MM') AS ORDER_MONTHLY,  
ORDER_STATUS, COUNT(ORDER_ID) AS ORDER_COUNT  
FROM ORDERS WHERE ORDER_APPROVED_AT IS NOT NULL GROUP BY ORDER_MONTHLY, ORDER_STATUS
```



order_status				
approved	canceled	delivered	invoiced	processing
shipped	unavailable			

Sipariş statüsünde tümü seçili olsa dahi en ön plana çıkan hızlı yükseliş gösteren 'Delivery' statüsünü inceleyeceğim.

Teslim edilme durumunda bir önceki aya göre neredeyse 2ye katlandığı görülen 2017 Kasım ayıdır.Bu artışın nedenlerini Black Friday, Cyber Monday ve Cumhuriyetin İlanı gibi günlerin Kasım ayında bulunmasına bağlayabiliriz. Bu günlere ait kampanyaların, indirimlerin ve talebin sayesinde sipariş teslim edilme sayılarında yoğun artış buna bağlanabilir.Bir sonraki ay ani düşüş olarak görünse de normal aylık artış seyrine göre ilerlediğini söyleyebiliriz.

Delivery haricinde de diğer önemli kırılımlar olan 'Canceled' ve 'Unavailable' durumlarını da inceleyelim.

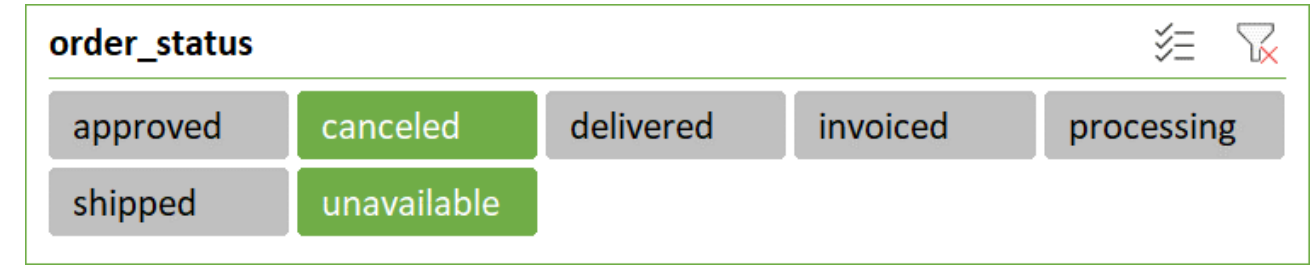
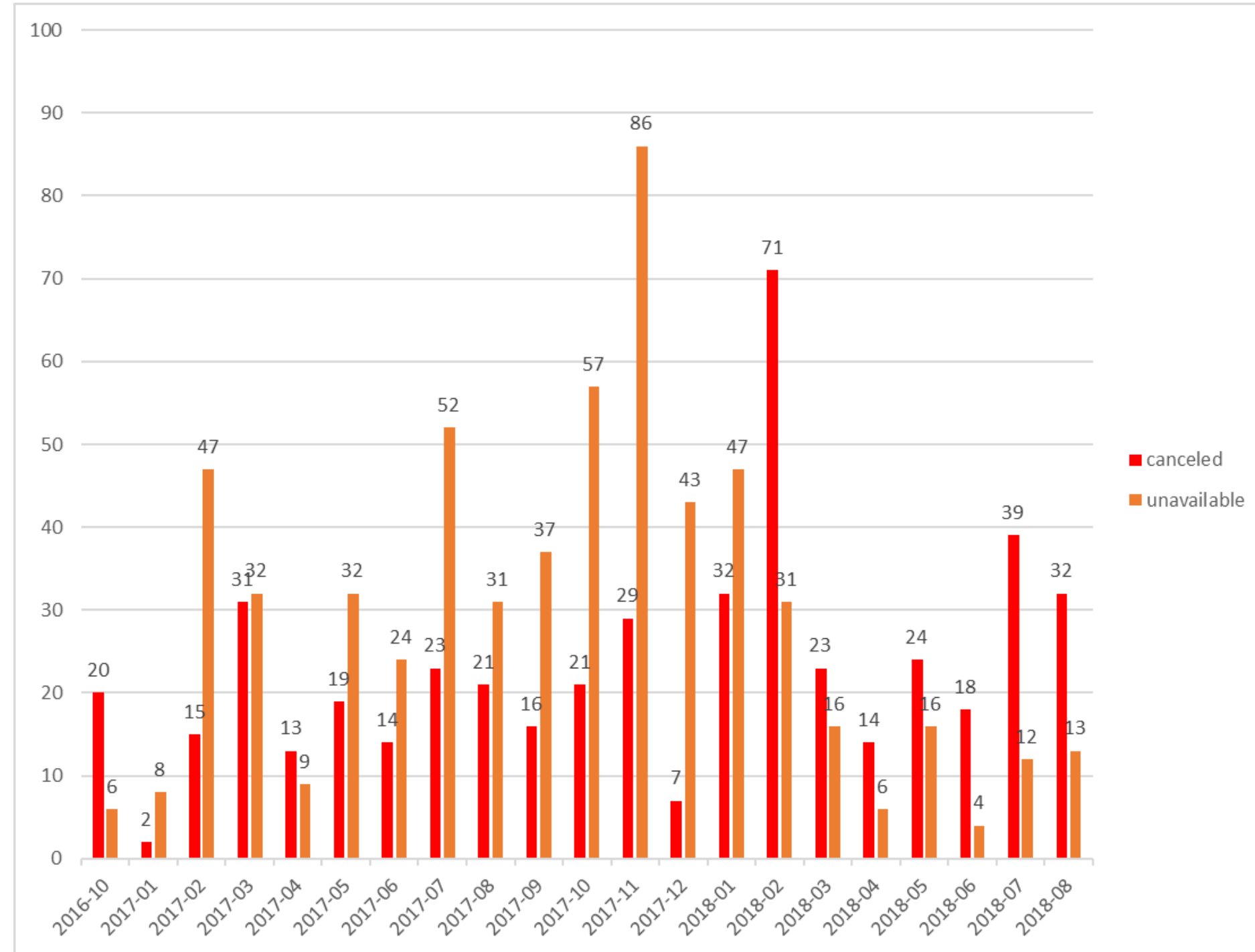
QUESTION 2 :

AYLIK OLARAK ORDER STATUS KIRILIMINDA ORDER SAYILARINI INCELEYINIZ. SORGU SONUCUNDA ÇIKAN OUTPUTU EXCEL İLE GÖRSELLEŞTİRİNİZ.DRAMATİK BİR DÜŞÜŞÜN YA DA YÜKSELİŞİN OLDUĞU AYLAR VAR MI? VERİYİ İNCELEYEREK YORUMLAYINIZ.

SELECT TO_CHAR(ORDER_APPROVED_AT,'YYYY-MM') AS ORDER_MONTHLY,

ORDER_STATUS, COUNT(ORDER_ID) AS ORDER_COUNT

FROM ORDERS WHERE ORDER_APPROVED_AT IS NOT NULL GROUP BY ORDER_MONTHLY, ORDER_STATUS



'Canceled' statüsünde görülen hızlı artış ve azalış durumunu anlatmak için 2018'in 1.çeyrek dönemini ele alacağım.2018-01 ayında söylenen siparişler geciken teslim tarihleri nedeniyle yığılarak Şubat ayına yansımış bu da Şubat ayında bir önceki aya göre 2katına yakın siparişi iptal etme verisini ortaya çıkarmıştır.Bunların dışında müşterinin olağan şekilde siparişi onaylansa dahi kargoya teslim edilmeden önce iptal etme durumları da söz konusudur ama özellikle Ocak sonunda verilen siparişler onaylanıp aynı günü veya ertesi günü kargoya teslim edilse dahi kargocunun teslimatı 1 ayı bulmaktadır.Şubat ayında Karnaval ayın 1 haftasını aşan süreçte ve yoğunlukta olduğuna bu iptal ve aksamalara sebebiyet vermiş olabilir.

'Unavailable' statüsü durumu mevcut değil statüsünde kullanılan bir kırılımdır ve canceled sayısı kadar veri tablosunda yer almaktadır. Statüsü mevcut olmayan durumlar için siparişin o anki statusu seçilmemiş ya da yoğun iş anında statunun güncellenmemiş olduğu yönünde tahmin yürütebilirim.Bunlara bağlı siparişlerin incelenip durumlarının güncellenmesi unavailable kırılımının azalmasını sağlayacaktır.

QUESTION 3 :

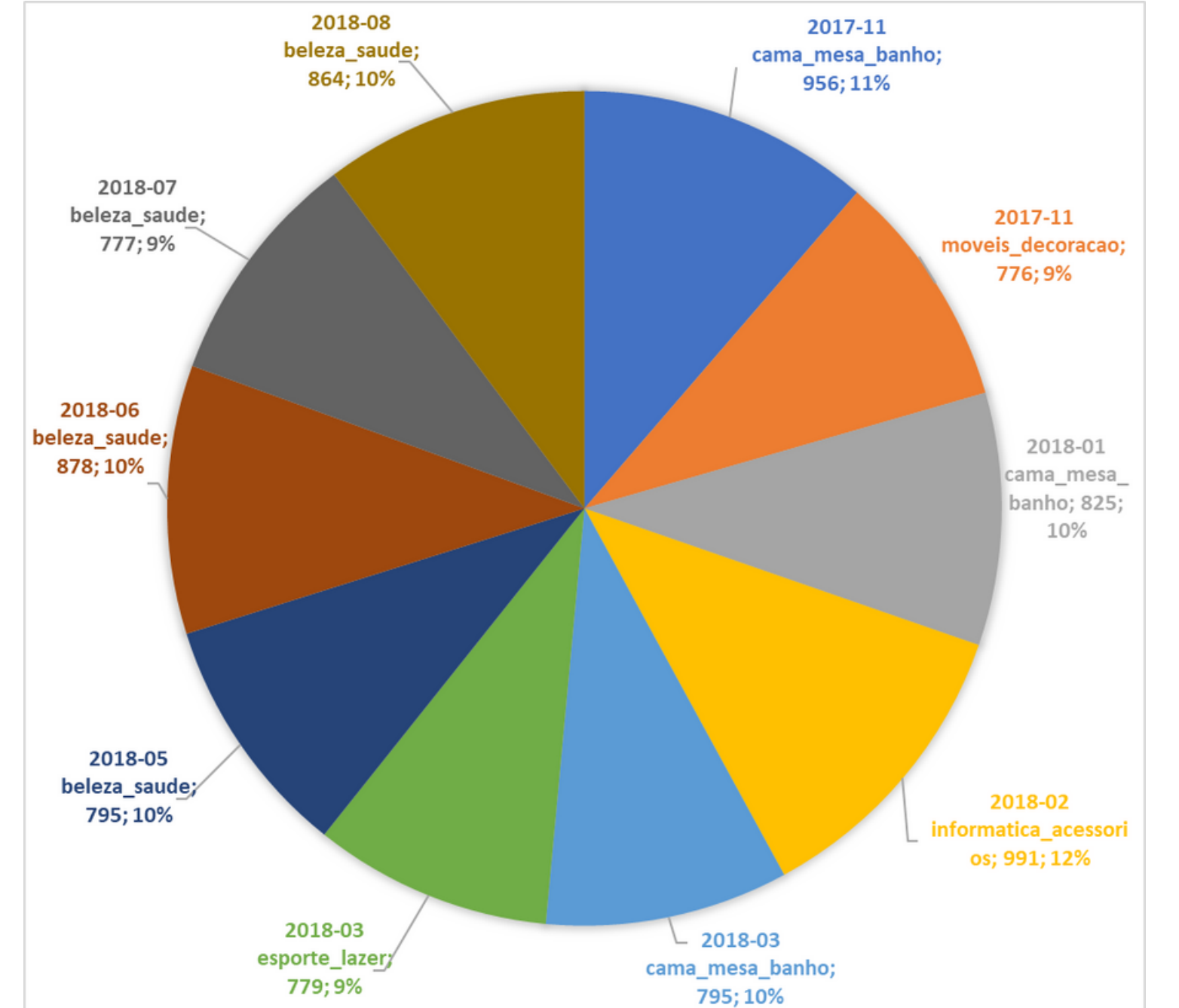
ÜRÜN KATEGORİSİ KIRILIMINDA SİPARİŞ SAYILARINI İNCELEYİNİZ.

ÖZEL GÜNLERDE ÖNE ÇIKAN KATEGORİLER NELERDİR? ÖRNEĞİN YILBAŞI, SEVGİLİLER GÜNÜ...

```
SELECT TO_CHAR(ORDER_APPROVED_AT,'YYYY-MM') AS ORDER_MONTHLY,  
P.PRODUCT_CATEGORY_NAME, COUNT(O.ORDER_ID) AS ORDER_COUNT  
FROM ORDERS O LEFT JOIN ORDER_ITEMS OI ON OI.ORDER_ID=O.ORDER_ID LEFT JOIN PRODUCTS P ON  
P.PRODUCT_ID=OI.PRODUCT_ID  
GROUP BY ORDER_MONTHLY, PRODUCT_CATEGORY_NAME  
ORDER BY ORDER_COUNT DESC;
```

En çok satın alınan kategori miktarlarının toplamını inceledim ve listedeki ilk 10 kategori üzerine yorum yapacağım.991 adet satış miktarıyla 2018-02 ayında en çok satılan kategori informatica_acessorios (bilgisayar aksesuarları) olarak görünmektedir.Şubat ayı özel günler karnaval ve sevgililer günü gibi günler için daha erken sipariş verme durumunu düşünürsek 2018-Ocak ayında satın alınan kategorilere bakmak daha mantıklı olacaktır.Mobilya,ev eşyaları,dekor ürünleri hediyeelik eşyalar ön plana çıkmaktadır.

Çoğunlukla müşterilerin tercih ettiği kategoriler yılbaşı,özel günler için mobilya,güzellik ürünleri,dekorasyon,bilgisayar aksesuarları en çok tercih edilen kategorilerdir.



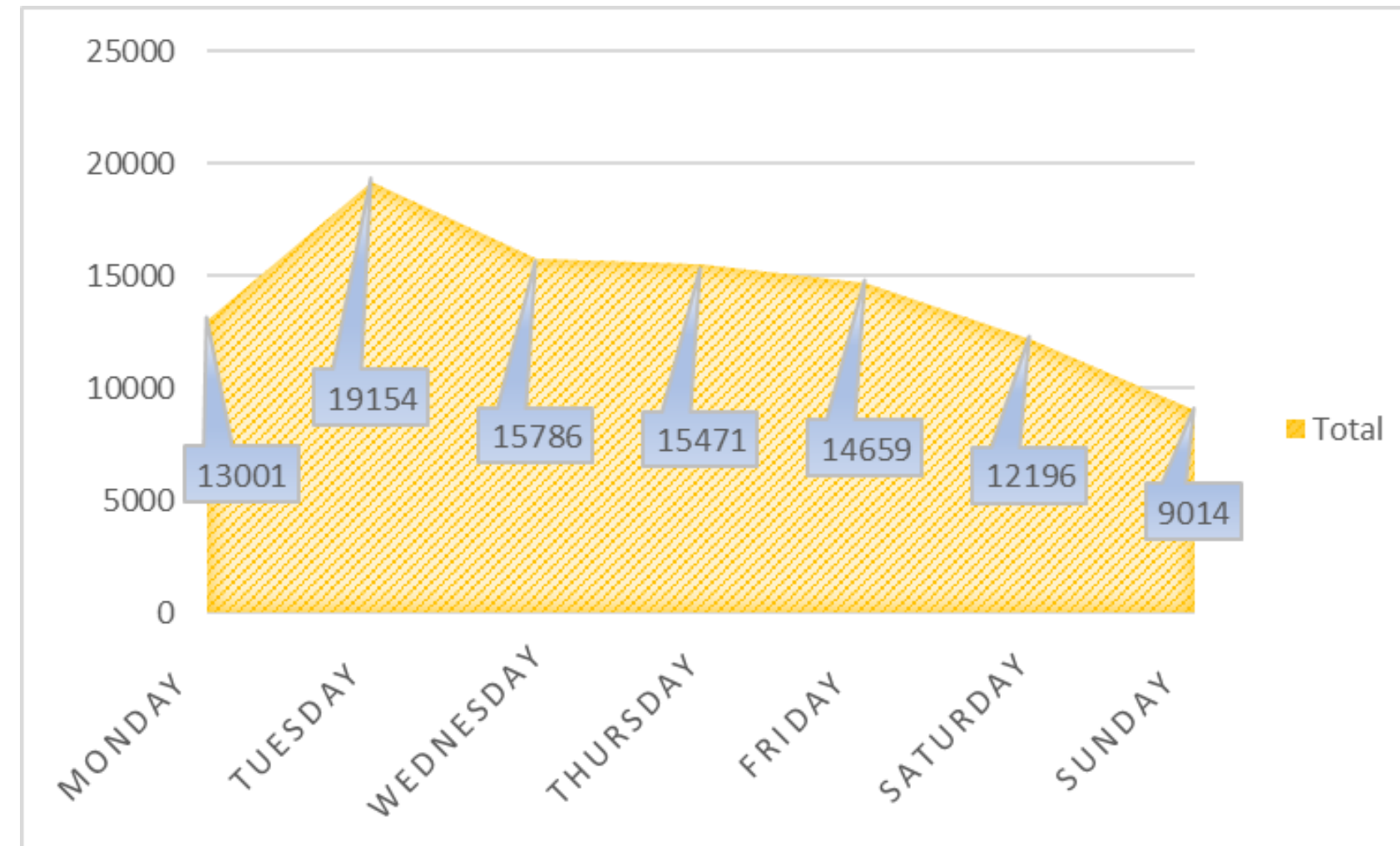
QUESTION 4 :

HAFTANIN GÜNLERİ(PAZARTESİ, PERŞEMBE, ...) VE AY GÜNLERİ (AYIN 1’İ,2’Sİ GIBI) BAZINDA ORDER SAYILARINI INCELEYİNİZ.

YAZDIĞINIZ SORGUNUN OUTPUTU İLE EXCEL’DE BİR GÖRSEL OLUŞTURUP YORUMLAYINIZ.

HAFTANIN GÜNLERİ

```
SELECT TO_CHAR(ORDER_APPROVED_AT,'DAY') AS ORDER_DAILY,  
COUNT(ORDER_ID) AS ORDER_COUNT  
FROM ORDERS  
GROUP BY ORDER_DAILY;
```



Sipariş sayılarının en çok Salı günü yapıldığını görmekteyiz.En az sipariş sayısı da Pazar gününe aittir.Haftasonu alışveriş yapmayıp yoğun olarak hafta içi sipariş alışkanlığı bulunan verilere göre yoğun olarak indirim,reklam çalışmaları da hafta içinde yapılması satışları daha da artıracaktır.

QUESTION 4 :

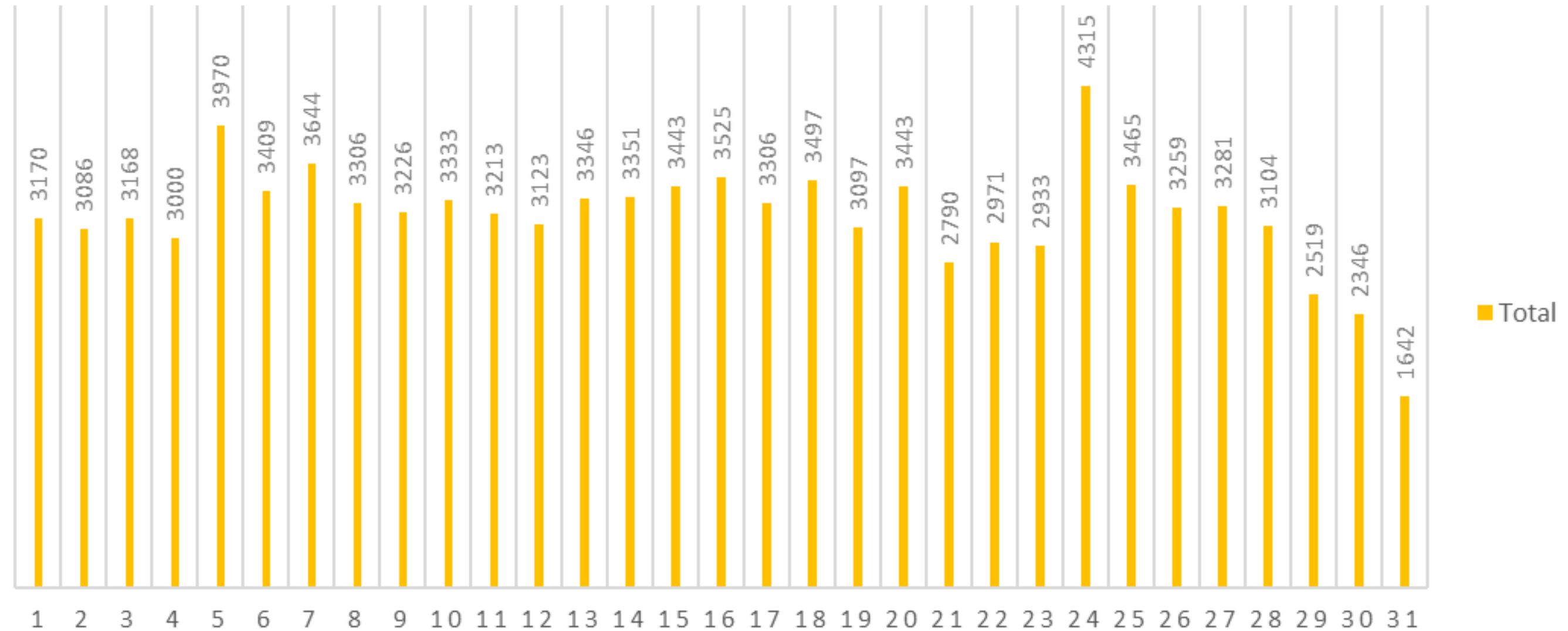
HAFTANIN GÜNLERİ(PAZARTESİ, PERŞEMBE,) VE AY GÜNLERİ (AYIN 1’İ,2’Sİ GIBI) BAZINDA ORDER SAYILARINI INCELEYİNİZ.

YAZDIĞINIZ SORGUNUN OUTPUTU İLE EXCEL’DE BİR GÖRSEL OLUŞTURUP YORUMLAYINIZ.

AYIN GÜNLERİ

```
SELECT EXTRACT(DAY FROM ORDER_APPROVED_AT) AS ORDER_DAILY,  
COUNT(ORDER_ID) AS ORDER_COUNT  
FROM ORDERS  
GROUP BY ORDER_DAILY ORDER BY ORDER_DAILY;
```

Ayin günlerine göre toplam sipariş miktarını incelediğimde ayın 24ü en yüksek 31'i en düşük sipariş toplamları olduğunu söyleyebilirim.Ancak genel olarak gün dağılımda birbirine yakın adetler mevcut olduğundan istikrarlı bir adetsel alışverişin varlığından bahsedebiliriz.Ayın 24ü ay sonlarına gelen özel gün,kampanya ve indirimlerin tarihlerine denk gelmesinden ötürü yüksek miktar,aysonları iş yoğunlukları ya da haftasonlarına denk gelmesi nedeniyle de düşük miktar olarak çıkması yüksek ihtimallerden birkaçı olabilir.

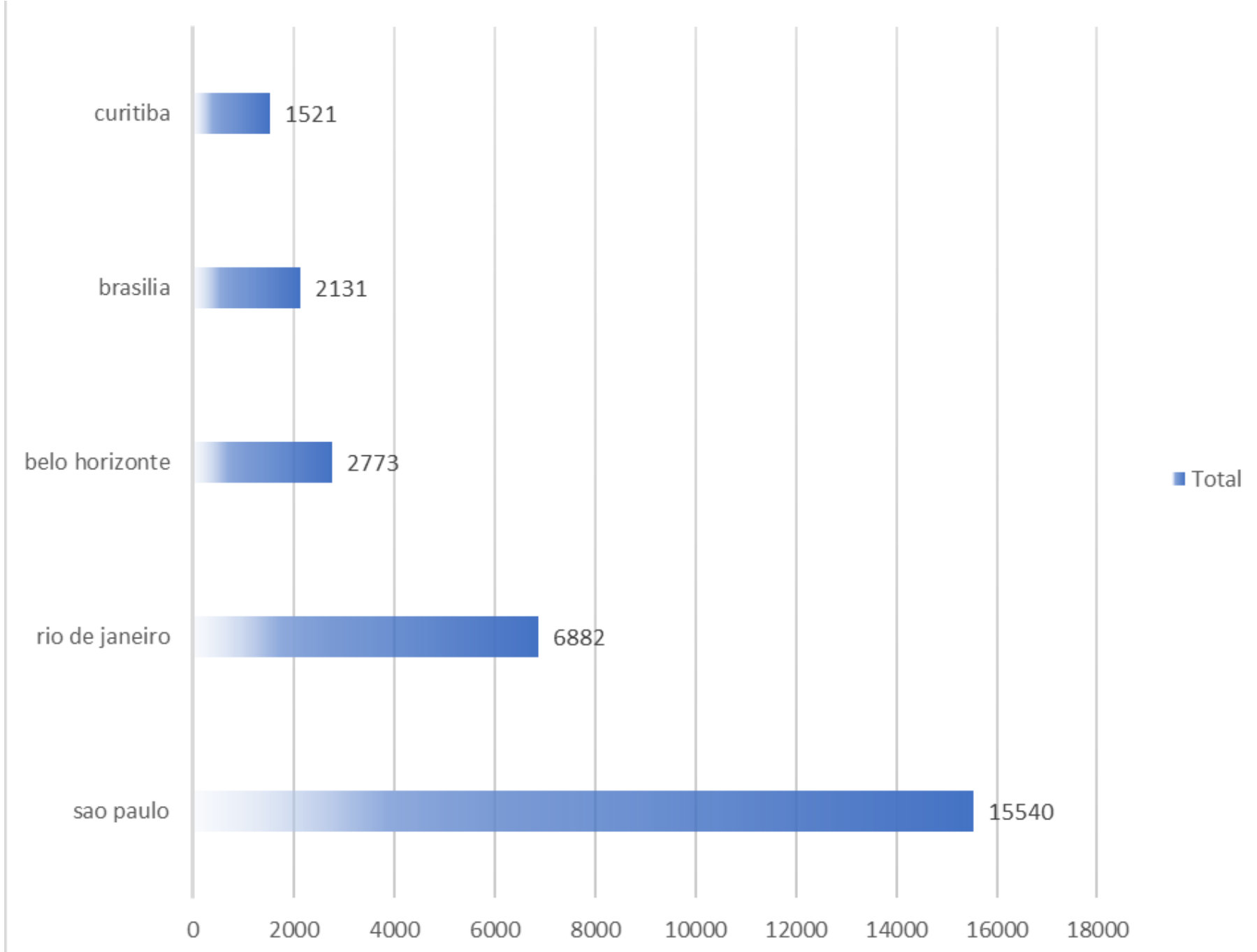


--CASE 2 : MÜŞTERİ ANALIZI

--QUESTION 1 :

--HANGI ŞEHİRLERDEKİ MÜŞTERİLER DAHA ÇOK ALIŞVERİŞ YAPIYOR?

--MÜŞTERİNİN ŞEHRİNİ EN ÇOK SİPARİŞ VERDİĞİ ŞEHİR OLARAK BELİRLEYİP ANALIZI ONA GÖRE YAPINIZ.



```
SELECT C.CUSTOMER_CITY,  
COUNT(O.ORDER_ID) AS ORDER_COUNT  
FROM CUSTOMERS C INNER JOIN ORDERS O ON  
O.CUSTOMER_ID=C.CUSTOMER_ID  
GROUP BY CUSTOMER_CITY  
ORDER BY ORDER_COUNT DESC
```



En çok alışverişin yapıldığı ilk 2 bölge Sao Paulo ve Rio De Janerio.Diğer bölgelere göre daha fazla nüfusa ve alım gücüne sahip olan bu bölgeler de en çok sipariş verilmesi olağan bir süreçtir.Sosyo ekonomik düzey ve nüfusun bu iki bölgede diğer yerlere göre katlayacak düzeyde fazla olması en çok alışveriş yapılan şehirler olarak bulunmalarını açıklayacaktır.

CASE 3: SATICI ANALIZI

QUESTION 1 :

SİPARİŞLERİ EN HIZLI ŞEKİLDE MÜŞTERİLERE ULAŞTIRAN SATICILAR KİMLERDİR? TOP 5 GETİRİNİZ.

```
WITH SELLER_ORDER_DELIVERY AS (  
SELECT OI.SELLER_ID,  
EXTRACT(DAY FROM ORDER_DELIVERED_CUSTOMER_DATE - ORDER_PURCHASE_TIMESTAMP) AS DAY_DIFFERENCE  
FROM ORDERS O  
LEFT JOIN ORDER_ITEMS OI ON O.ORDER_ID=OI.ORDER_ID)  
SELECT SELLER_ID,  
ROUND(AVG(DAY_DIFFERENCE)::NUMERIC) AS AVG_DELIVERY  
FROM SELLER_ORDER_DELIVERY  
GROUP BY SELLER_ID  
ORDER BY AVG_DELIVERY LIMIT 5
```

	seller_id 	avg_delivery 
	character varying (100)	numeric
1	96f7c797de9ca20efbe14545bed63eec	1
2	2c00c85d30361cd2ced2969cffbbffa3	1
3	5e063e85d44b0f5c3e6ec3131103a5...	1
4	751e274377499a8503fd6243ad9c56...	1
5	f3511c85f59f8dec53d140501ee8e7a8	1

Sipariş adetleri en yüksek olan 5 satıcı bilgisine ulaştıktan sonra bunların ortalaması alınarak ilk 5 satıcıya ulaşılmıştır.Müşterilere en hızlı teslimat sağlayan satıcılar listelenmiştir.

BU SATICILARIN ORDER SAYILARI ILE ÜRÜNLERİNDEKİ YORUMLAR VE PUANLAMALARI İNCELEYİNİZ VE YORUMLAYINIZ.

```
WITH TOP_5 AS (  
  WITH SELLER_ORDER_DELIVERY AS (  
    SELECT OI.SELLER_ID,  
    EXTRACT(DAY FROM ORDER_DELIVERED_CUSTOMER_DATE - ORDER_PURCHASE_TIMESTAMP) AS DAY_DIFFERENCE  
  FROM ORDERS O  
  LEFT JOIN ORDER_ITEMS OI  
  ON O.ORDER_ID=OI.ORDER_ID)  
  SELECT SELLER_ID,  
  ROUND(AVG(DAY_DIFFERENCE)::NUMERIC) AS AVG_DELIVERY  
  FROM SELLER_ORDER_DELIVERY  
  GROUP BY SELLER_ID  
  ORDER BY AVG_DELIVERY LIMIT 5)  
  SELECT T.SELLER_ID,  
  COUNT(DISTINCT OI.ORDER_ID) AS ORDER_COUNT,  
  ROUND(AVG(R.REVIEW_SCORE)::NUMERIC) AS REVIEW_SCORE_AVG,  
  COUNT(DISTINCT REVIEW_COMMENT_MESSAGE) AS REVIEW_MESSAGE  
  FROM TOP_5 T  
  LEFT JOIN ORDER_ITEMS OI  
  ON OI.SELLER_ID = T.SELLER_ID  
  LEFT JOIN REVIEWS R  
  ON R.ORDER_ID = OI.ORDER_ID  
  GROUP BY T.SELLER_ID
```

Sipariş sayıları en çok olan top 5 listesindeki satıcıların yorum ve mesaj sayılarının 0 olduğunu görüyoruz.Bununla ilgili tüketicilerden olumlu olumsuz feedback almak üzerine çalışma yapılması satıcı ve tüketi için avantaj sağlayacaktır.Sipariş sayıları ile ön plana çıkan satıcıların yorumlarının olumlu olma ihtimali fazla olacağını düşünürsek diğer müşterilerin yorumları görerek satın alma yüzdesi artacaktır.Aynı şekilde sipariş miktarı fazla olan ama puanı 1 olan satıcı incelenip hangi ürün düşük puana neden oldu ya da ürün kalemlerinde düşük olan siparişi fazla olan ürünler neler onlar incelenip iyileştirme sağlanabilir.

	seller_id character varying (100)	order_count bigint	review_score_avg numeric	review_message bigint
1	2c00c85d30361cd2ced2969cffbbffa3	1	5	0
2	5e063e85d44b0f5c3e6ec3131103a5...	1	5	0
3	751e274377499a8503fd6243ad9c56...	1	5	0
4	96f7c797de9ca20efbe14545bed63eec	1	1	0
5	f3511c85f59f8dec53d140501ee8e7a8	1	5	0

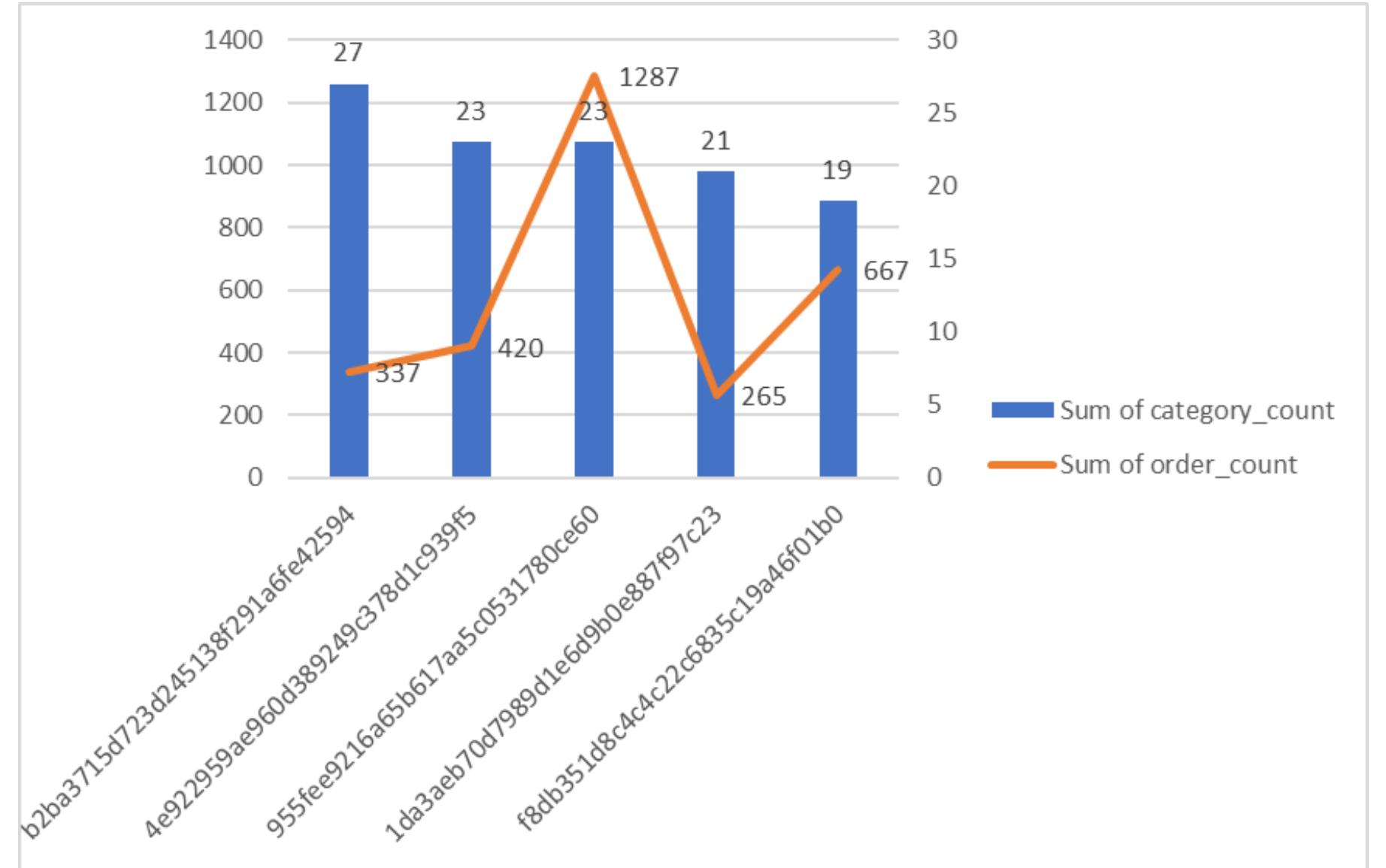
--QUESTION 2 :

HANGI SATICILAR DAHA FAZLA KATEGORIYE AIT ÜRÜN SATIŞI YAPMAKTADIR?

FAZLA KATEGORIYE SAHIP SATICILARIN ORDER SAYILARI DA FAZLA MI?

```
SELECT OI.SELLER_ID,  
COUNT(DISTINCT P.PRODUCT_CATEGORY_NAME) AS CATEGORY_COUNT,  
COUNT(DISTINCT OI.ORDER_ID) AS ORDER_COUNT  
FROM ORDER_ITEMS AS OI  
LEFT JOIN PRODUCTS AS P  
ON OI.PRODUCT_ID = P.PRODUCT_ID  
GROUP BY OI.SELLER_ID  
ORDER BY CATEGORY_COUNT DESC
```

	seller_id character varying (100)	category_count bigint	order_count bigint
1	b2ba3715d723d245138f291a6fe42594	27	337
2	4e922959ae960d389249c378d1c939...	23	420
3	955fee9216a65b617aa5c0531780ce60	23	1287
4	1da3aeb70d7989d1e6d9b0e887f97c...	21	265
5	f8db351d8c4c4c22c6835c19a46f01b0	19	667
6	18a349e75d307f4b4cc646a691ed42...	17	121
7	6edacfd9f9074789dad6d62ba7950b9c	15	208
8	70a12e78e608ac31179aea7f842204...	15	315
9	7178f9f4dd81dcef02f62acdf8151e01	14	203
10	fd386aa7bed2af3c7035c65506c9b4a3	14	69

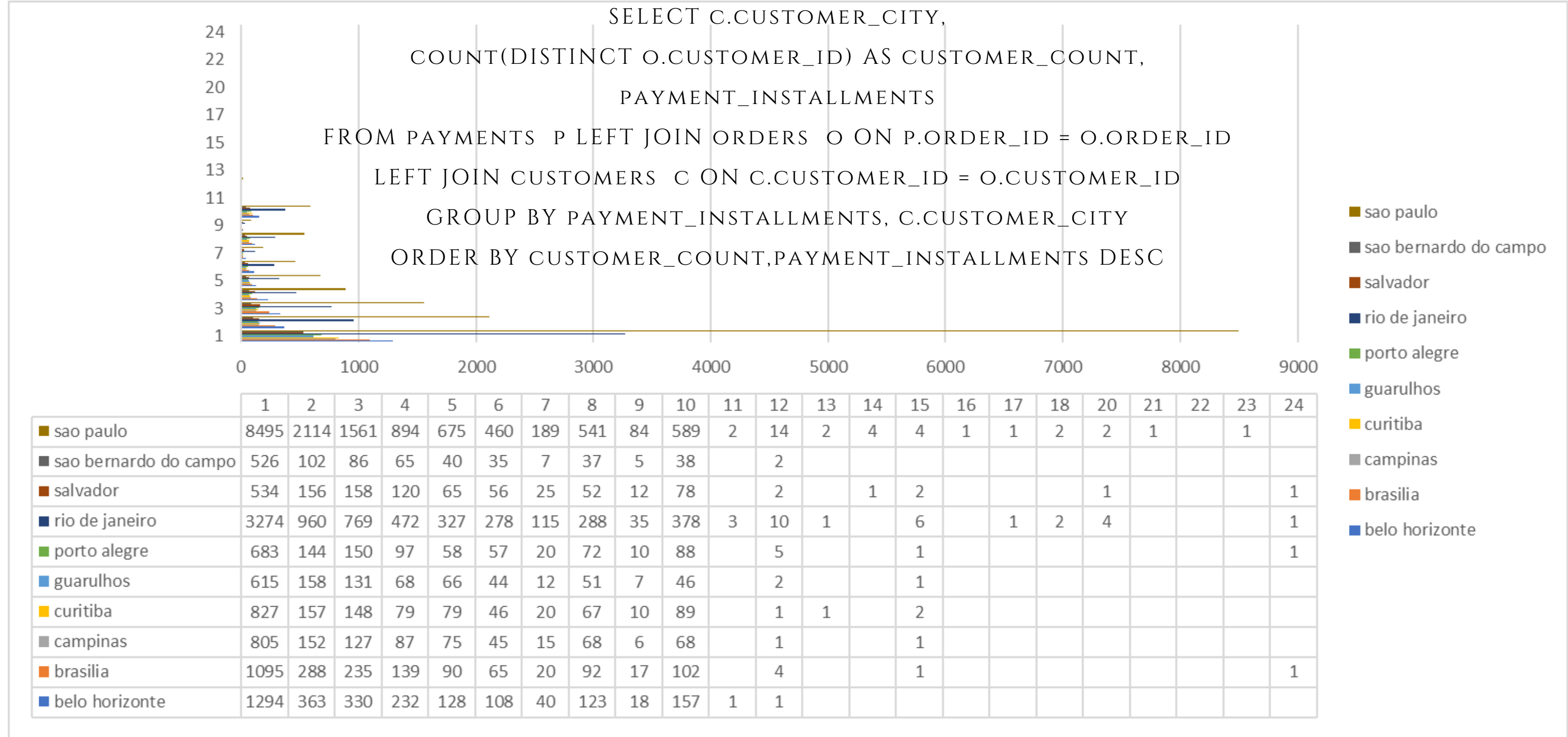


En fazla kategoriye sahip satıcılar listelenmiştir. Solda bulunan grafikte en fazla kategorilere sahip 5 satıcı ve o satıcılara ait sipariş adetlerini görebiliriz. Grafiği incelediğimizde fazla kategori ile fazla sipariş sayısı arasında doğru orantı bulunmamaktadır. Kategori çeşitliliği satış adetlerini olumlu yönde etkileyebilir ama satış adedinde kesin artışı sağlar diyemeyiz. Datamızda da bunu kanıtlar niteliktedir.

CASE 4 : PAYMENT ANALIZI

QUESTION 1 :

ÖDEME YAPARKEN TAKSIT SAYISI FAZLA OLAN KULLANICILAR EN ÇOK HANGİ BÖLGEDE YAŞAMAKTADIR? BU ÇIKTIYI YORUMLAYINIZ.



En fazla taksit yapılan bölgeler farklılık göstermektedir çünkü son 20-24 taksit sayısına baktığımızda metropol büyük şehirlerin de ismi geçiyor daha ufak çaplı gelir düzeyi düşük şehirlerin de ismi geçiyor.Bu aslında kredi kartı kullanımı günümüzde çok yaygın olduğu için maddi yönden çok iyi bir düzeyde olan da taksit yapıyor ya da kartla ödeme yapıyor aynı şekilde durumu iyi olmayıp bol taksitli seçenikle alım gerçekleştiren dar gelirli de bu yolu izliyor.Ancak yorumlama açısından 20-24 taksit aralığında taksit seçeneği kullanan bölgeler ve en az taksit seçeneği kullanan Rio de Janerio ve Sao Paulo gibi Brezilya'nın ilk sıralarda yoğun nüfus yogunluguna sahip,metropol ,gelir düzeyi yüksek bölgelerdir.Bu bölgelerde taksit,alışveriş,sipariş sayıları en fazla olarak bulunmaktadır.

QUESTION 2 :

ÖDEME TIPINE GÖRE BAŞARILI ORDER SAYISI VE TOPLAM BAŞARILI ÖDEME TUTARINI HESAPLAYINIZ. EN ÇOK KULLANILAN ÖDEME TIPINDEN EN AZ OLANA GÖRE SIRALAYINIZ.

```
SELECT PAYMENT_TYPE,  
ROUND(SUM(PAYMENT_VALUE)::NUMERIC,2) AS SUM_PAYMENT_VALUE,  
COUNT(DISTINCT O.ORDER_ID) AS ORDER_COUNT  
FROM PAYMENTS P LEFT JOIN ORDERS O  
ON P.ORDER_ID = O.ORDER_ID  
GROUP BY PAYMENT_TYPE ORDER BY SUM_PAYMENT_VALUE DESC
```

En çok kullanılan ödeme tipi kredi kartıdır. Boleto Brezilya’da yaygın olarak kullanılan 2.ödemeye yönetimidir.

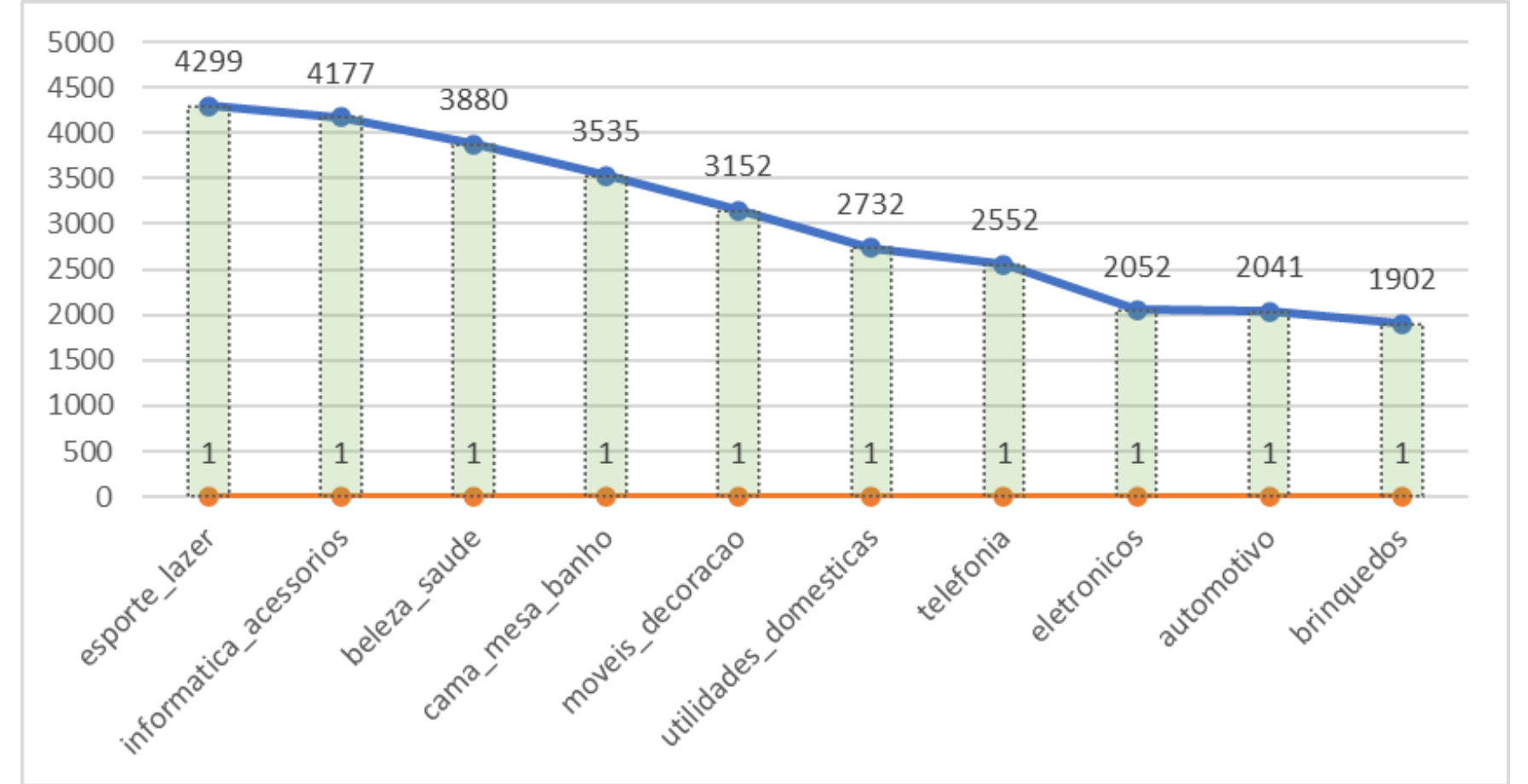
	payment_type character varying (100)	sum_payment_value numeric	order_count bigint
1	credit_card	12542084.19	76505
2	boleto	2869361.27	19784
3	voucher	379436.87	3866
4	debit_card	217989.79	1528
5	not_defined	0.00	3

QUESTION 3 :

TEK ÇEKİMDE VE TAKSİTLE ÖDENEN SİPARİŞLERİN KATEGORİ BAZLI ANALİZİNİ YAPINIZ. EN ÇOK HANGİ KATEGORİLERDE TAKSİTLE ÖDEME KULLANILMAKTADIR?

TEK ÇEKİM

```
SELECT PRODUCT_CATEGORY_NAME,PAYMENT_INSTALLMENTS,
COUNT(DISTINCT O.ORDER_ID) AS ORDER_COUNT
FROM ORDERS AS O
LEFT JOIN PAYMENTS AS P
ON P.ORDER_ID = O.ORDER_ID
LEFT JOIN ORDER_ITEMS AS OI
ON OI.ORDER_ID = O.ORDER_ID
LEFT JOIN PRODUCTS AS PR ON PR.PRODUCT_ID=OI.PRODUCT_ID
WHERE PAYMENT_INSTALLMENTS=1
GROUP BY PAYMENT_INSTALLMENTS,PRODUCT_CATEGORY_NAME
ORDER BY ORDER_COUNT DESC, PAYMENT_INSTALLMENTS DESC
```



Tek çekimle satın alınan ürün kategoriler başında Esporte_lazer, informatica_acessorios, beleza_saude (spor-eğlence, bilgisayar_aksesuarları, güzellik_sağlık)kategorileri yer almaktadır.

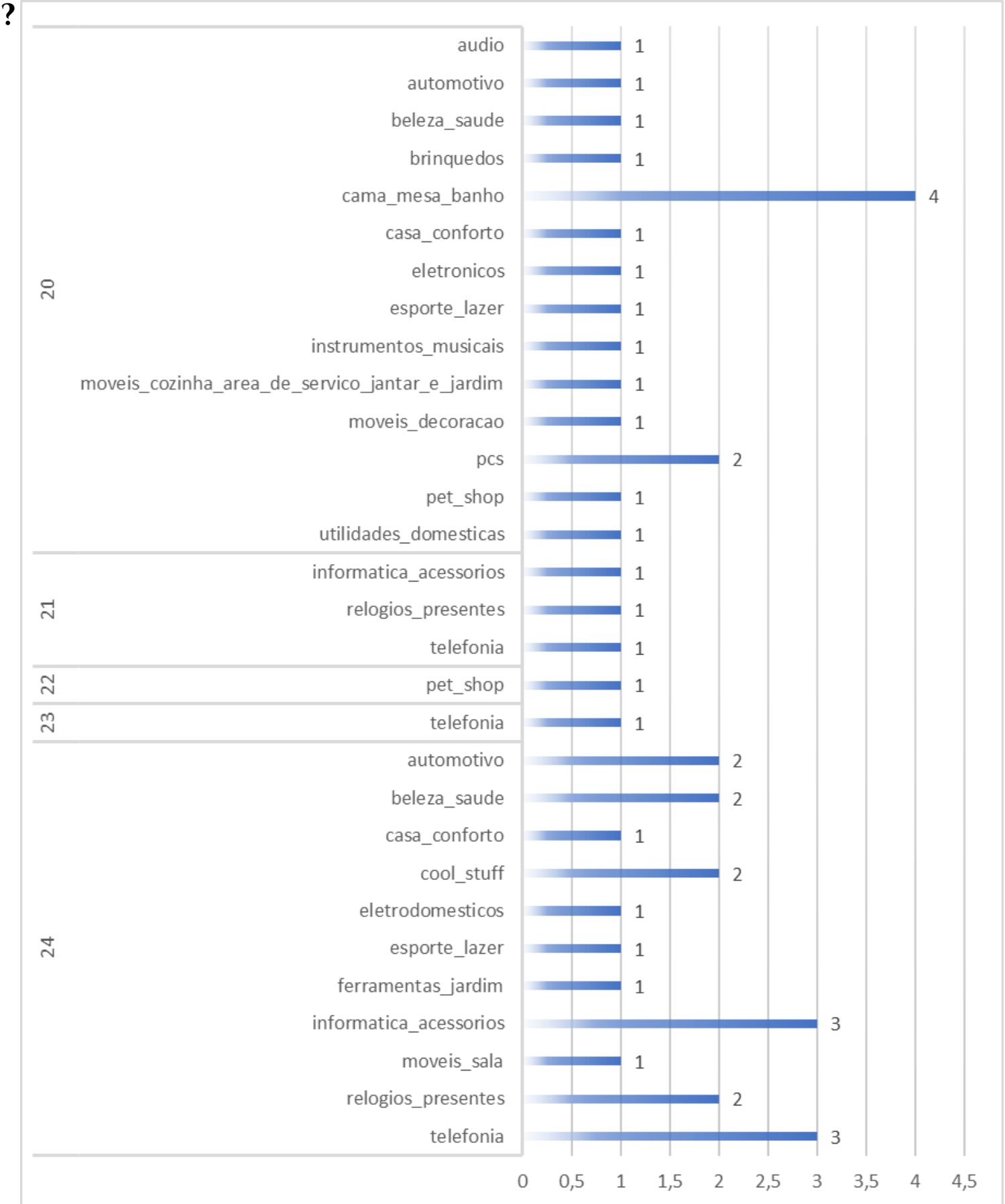
QUESTION 3 :

TEK ÇEKİMDE VE TAKSİTLE ÖDENEN SİPARİŞLERİN KATEGORİ BAZLI ANALİZİNİ YAPINIZ. EN ÇOK HANGİ KATEGORİLERDE TAKSİTLE ÖDEME KULLANILMAKTADIR?

TAKSİTLİ ÇEKİM

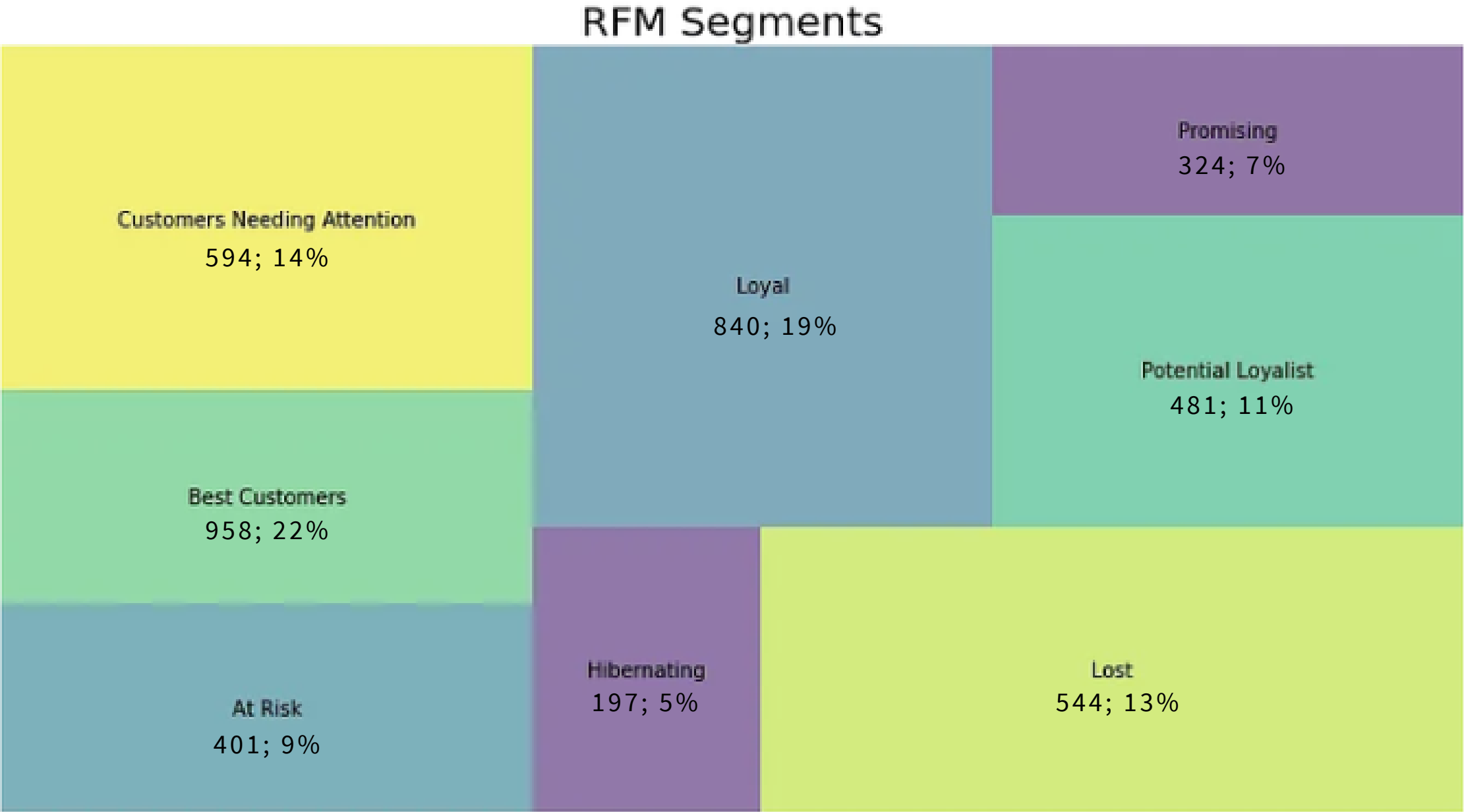
```
SELECT PRODUCT_CATEGORY_NAME,PAYMENT_INSTALLMENTS,
COUNT(DISTINCT O.ORDER_ID) AS ORDER_COUNT
FROM ORDERS AS O
LEFT JOIN PAYMENTS AS P
ON P.ORDER_ID = O.ORDER_ID
LEFT JOIN ORDER_ITEMS AS OI
ON OI.ORDER_ID = O.ORDER_ID
LEFT JOIN PRODUCTS AS PR ON PR.PRODUCT_ID=OI.PRODUCT_ID
WHERE PAYMENT_INSTALLMENTS>1
GROUP BY PAYMENT_INSTALLMENTS,PRODUCT_CATEGORY_NAME
ORDER BY PAYMENT_INSTALLMENTS DESC, ORDER_COUNT DESC
```

Taksitli alışveriş ile satış miktarı en fazla olan ,24 taksit yapılan kategori Telefon ve Bilgisayar aksesuarları kategorileri (telefon, informatica_acessorios) olarak görünmektedir.20-24 taksit yapılan kategori ve satış miktarları grafikte yer almaktadır.Fazla taksitle yapılan alışveriş kategorileri için telefon,bilgisayar,otomotiv gibi yüksek tutarlara sahip bol taksitle ödemesi kolay olunacak kategorilerin fazla olduğunu söyleyebilirim.



Case 5 : RFM Analizi

```
WITH rfm_data AS (  
  SELECT customer_id,  
  invoiceno,  
  quantity,  
  invoicedate,  
  unitprice  
  FROM rfmm  
  WHERE invoiceno NOT LIKE 'C%'  
  AND quantity > 0  
  AND customer_id IS NOT NULL  
  AND invoiceno IS NOT NULL  
),  
date_invoice as(  
  select customer_id,  
  max(invoicedate) as max_invoice_date  
  from rfm_data  
  group by 1  
,  
recency as(  
  select customer_id,  
  max_invoice_date,  
  ('2011-12-09' - max_invoice_date::date) as recency  
  from date_invoice  
,  
frequency as(  
  select customer_id,  
  COUNT(distinct invoiceno) as frequency  
  from rfm_data  
  GROUP BY 1  
,  
monetary as(  
  select customer_id,  
  ROUND(SUM(quantity * unitprice) :: numeric, 2) AS monetary  
  FROM rfm_data  
  GROUP BY 1  
,
```



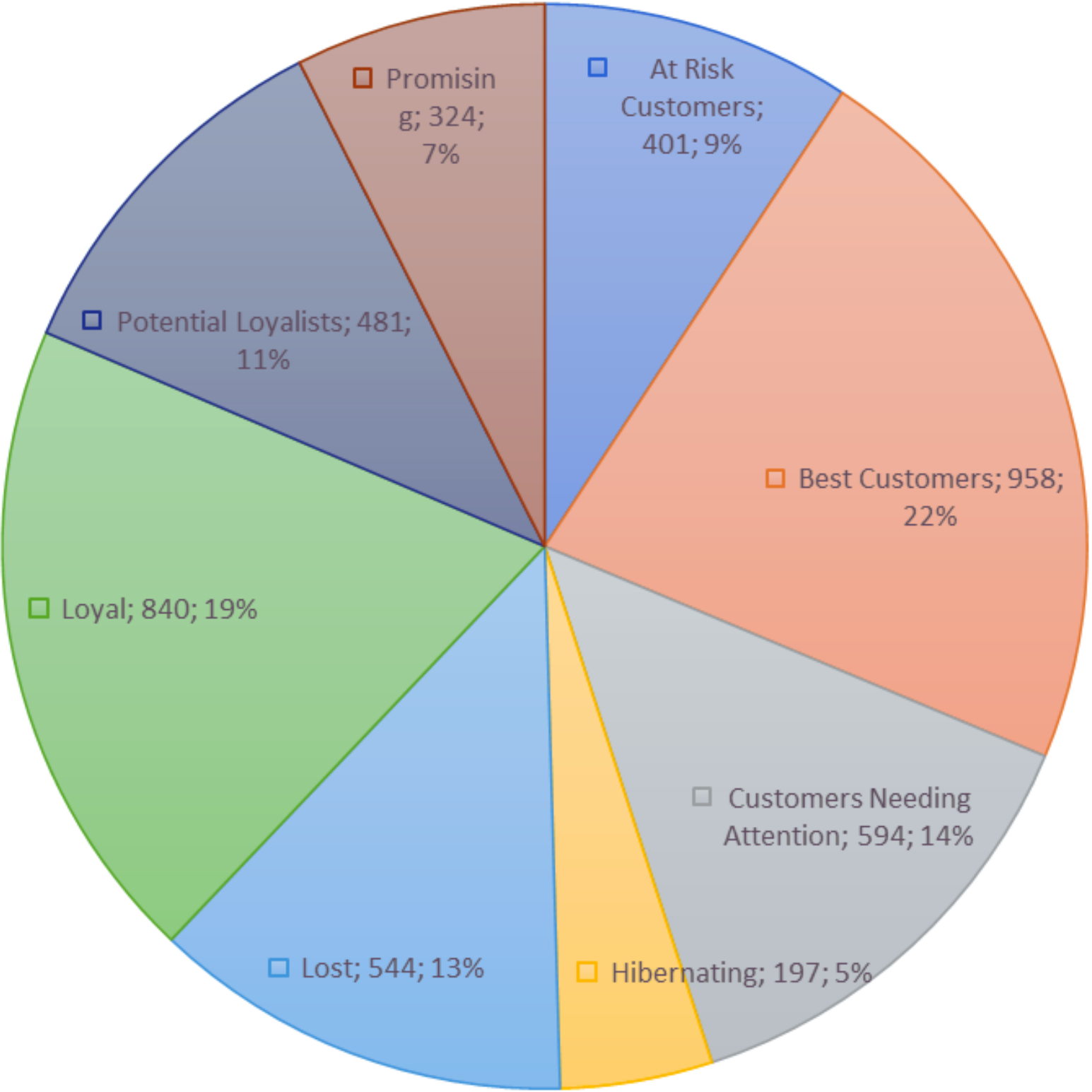
- "Best Customers" En iyi Müşteriler
- "Loyal" Sadık Müşteriler
- "Customers Needing Attention" İlgiye İhtiyacı Olan Müşteriler
- "Lost" Kayıp/Ulaşılamayan Müşteriler
- "Potential Loyalists" Potansiyel Sadıklar
- "At Risk Customers" Risk Altındaki Müşteriler
- "Promising" Gelecek Vaat Eden Müşteriler
- "Hibernating" Uykuda Olan Müşteriler

```
scores AS (  
SELECT r.customer_id,  
r.recency,  
NTILE(5) OVER (ORDER BY recency desc) AS rfm_recency,  
NTILE(5) OVER (ORDER BY frequency) AS rfm_frequency,  
NTILE(5) OVER (ORDER BY monetary) AS rfm_monetary  
FROM recency r  
join frequency f on r.customer_id = f.customer_id  
join monetary m on r.customer_id = m.customer_id  
,
```

```
rfm_segment AS (  
SELECT customer_id,  
rfm_recency,  
rfm_frequency,  
rfm_monetary,  
CASE WHEN rfm_recency >= 4 AND rfm_frequency >= 4 AND rfm_monetary >= 4 THEN 'Best Customers'  
WHEN rfm_recency >= 3 AND rfm_frequency >= 3 AND rfm_monetary >= 3 THEN 'Loyal'  
WHEN rfm_recency >=3 AND rfm_frequency >= 1 AND rfm_monetary >= 2 THEN 'Potential Loyalists'  
WHEN rfm_recency >=3 AND rfm_frequency >= 1 AND rfm_monetary >= 1 THEN 'Promising'  
WHEN rfm_recency >=2 AND rfm_frequency >= 2 AND rfm_monetary >= 2 THEN 'Customers Needing Attention'  
WHEN rfm_recency >=1 AND rfm_frequency >= 2 AND rfm_monetary >= 2 THEN 'At Risk Customers'  
WHEN rfm_recency >=1 AND rfm_frequency >=1 AND rfm_monetary >= 2 THEN 'Hibernating'  
else 'Lost'  
END AS segment  
FROM scores  
)
```

```
--select customer_id,rfm_recency,  
--rfm_frequency,  
--rfm_monetary,  
--segment from rfm_segment
```

```
select segment,  
COUNT(DISTINCT customer_id) AS customer_count  
from rfm_segment  
GROUP BY segment  
ORDER BY 2 DESC;
```



RFM Analizi sonucunda çeşitli müşteri segmentleri oluşturdum. Busegmentleri yüzdelik ve toplam müşteri sayılarına göre grupladım.Bunların sonucunda Promising, Potential Loyalist,Hibernating,Customers Needing Attention segmentlerindeki müşterileri sadık veya en iyi müşteri segmentlerine taşıyabilmek için büyük potansiyel taşımaktalar.Bu potansiyel At Risk ve Lost segmentlerine göre daha yüksektir.Lost ve At Risk segmentleri için daha önceki alışverişlerine istinaden kategorilere göre indirimler ,özel gün ve kampanya zamanları mail bilgilendirme çalışmaları ile alışveriş yapmaya geri döndürmek için çalışmalar yapılabilir.Best Customers ve Loyal segmentleri için en cazip harcamaya teşvik yöntemleri harcadıkça puan kazanımı ya da ürünün o an aldığıda tutara bağlı indirimler gibi daha çok kişinin harcamasına yönelik avantajlar,iyileştirmeler tanımlanarak devamlılık sağlanabilir.