**SQL PROJE**

A screenshot of a computer

Description automatically generated

**CASE 1 SİPARİŞ ANALİZİ**

**-SORU 1**

-Aylık olarak order dağılımını inceleyiniz.

**SQL QUERY**:

select

to\_char(order\_approved\_at, 'YYYY-MM' ) AS tarih,

count(distinct order\_id) AS order\_count

from orders

where to\_char(order\_approved\_at, 'YYYY-MM' ) IS NOT NULL

group by 1

ORDER BY 1

|  |  |
| --- | --- |
| **tarih** | **order\_count** |
| 1.09.2016 | 1 |
| 1.10.2016 | 320 |
| 1.12.2016 | 1 |
| 1.01.2017 | 760 |
| 1.02.2017 | 1765 |
| 1.03.2017 | 2689 |
| 1.04.2017 | 2374 |
| 1.05.2017 | 3693 |
| 1.06.2017 | 3252 |
| 1.07.2017 | 3974 |
| 1.08.2017 | 4348 |
| 1.09.2017 | 4301 |
| 1.10.2017 | 4590 |
| 1.11.2017 | 7395 |
| 1.12.2017 | 5832 |
| 1.01.2018 | 7187 |
| 1.02.2018 | 6706 |
| 1.03.2018 | 7288 |
| 1.04.2018 | 6778 |
| 1.05.2018 | 7066 |
| 1.06.2018 | 6164 |
| 1.07.2018 | 6176 |
| 1.08.2018 | 6620 |
| 1.09.2018 | 1 |

Elimizdeki verilere baktığımızda bu alışveriş sitesinin artış gösteren ve sonrasında belli aralıkta devam eden sipariş sayısı olduğunu gözlemleyebiliyoruz.

**-SORU 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?

**SQL QUERY**:

select to\_char(order\_approved\_at, 'YYYY-MM' ) AS tarih,

order\_status,

count(order\_id) AS order\_count

from orders

group by 1,2

order by 1,2 DESC

-DELIVERED STATÜSÜ İÇİN

select to\_char(order\_approved\_at, 'YYYY-MM' ) AS tarih,

count(DISTINCT order\_id) AS order\_count, order\_status

from orders

where order\_status IN ('delivered')

group by 1,3

order by 1,3

-CANCELED VE SHIPPED STATÜSÜ İÇİN

select to\_char(order\_approved\_at, 'YYYY-MM' ) AS tarih,

count(DISTINCT order\_id) AS order\_count, order\_status

from orders

where order\_status IN ('canceled','shipped')

group by 1,3

order by 1,3

Her iki tabloyu beraber incelediğimizde müşterilere ulaşan siparişlerin sayısı düzenli artış göstermiştir, özellikle 2017-11 ayında en yüksek sayıya ulaşılmıştır, o dönem yapılan bir indirim veya kampanya etkili olmuş olabilir.

Gönderimde olan siparişler de genel olarak dengeli gitmektedir. 2018-03 ayı hem gönderimde olan sipariş sayısı hem de ulaşan siparişler bakımından yüksek olduğu görülmektedir.

İptal edilen siparişlerde artışın görüldüğü 2018-02 ayı belirgin olarak öne çıkmaktadır. Verilerde tarihi belli olmayan fakat iptal edilen 14 adet daha sipariş görülmektedir. Bu siparişlerin aylara düzenli mi dağıldığı ya da çoğunun belli bir dönemde mi biriktiği bilinemediği için iptal edilen siparişler de belli bir aralıkta ilerlemiştir.

**-SORU 3**

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

**SQL QUERY**:

SELECT

product\_category\_name,

category\_name\_english,

count(DISTINCT orders.order\_id) as sold\_product\_count

from orders

inner join order\_items on orders.order\_id = order\_items.order\_id

inner join products on products.product\_id = order\_items.product\_id

join translation on products.product\_category\_name=translation.category\_name

group by 1,2

order by 3 desc

|  |  |  |
| --- | --- | --- |
| **product\_category\_name** | **category\_name\_english** | **sold\_product\_count** |
| cama\_mesa\_banho | bed\_bath\_table | 9417 |
| beleza\_saude | health\_beauty | 8836 |
| esporte\_lazer | sports\_leisure | 7720 |
| informatica\_acessorios | computers\_accessories | 6689 |
| moveis\_decoracao | furniture\_decor | 6449 |
| utilidades\_domesticas | housewares | 5884 |
| relogios\_presentes | watches\_gifts | 5624 |
| telefonia | telephony | 4199 |
| automotivo | auto | 3897 |
| brinquedos | toys | 3886 |
| cool\_stuff | cool\_stuff | 3632 |
| ferramentas\_jardim | garden\_tools | 3518 |
| perfumaria | perfumery | 3162 |
| bebes | baby | 2885 |
| eletronicos | electronics | 2550 |

\*Tabloda ilk 15 kategori yer almaktadır.

Toplam sipariş sayısı ve kategorilere bakacak olursak ev tekstil ürünleri, sağlık-güzellik ürünleri ve spor- eğlence ürünlerinde satışların daha fazla olduğunu gözlemliyoruz.

-Özel günler için analizler

1) Noel/Aralık Ayı

SQL QUERY:

with siralama as (SELECT to\_char(order\_approved\_at, 'YYYY-MM' ) AS tarih,

count(DISTINCT orders.order\_id) as sold\_product\_count, product\_category\_name,

row\_number() over(partition by to\_char(order\_approved\_at, 'YYYY-MM' ) order by count(DISTINCT orders.order\_id) desc) as rn

from orders

join order\_items on orders.order\_id = order\_items.order\_id

join products on products.product\_id = order\_items.product\_id

WHERE to\_char(order\_approved\_at, 'YYYY-MM' ) IN ('2016-12','2017-12')

group by 1,3

order by 1,2 desc)

select tarih, max(product\_category\_name) as product\_category\_name, sold\_product\_count

from siralama

where rn IN (1,2,3,4,5)

group by 1,3

order by 1,3 desc

\*Grafikte ilk 5 kategori yer almaktadır.

Noel dönemi için öne çıkan kategoriler; ev tekstil ürünleri, oyuncaklar, sağlık ve güzellik ürünleri, spor-eğlence ürünleri ve saat olarak görülmektedir. Yılbaşı dönemi olduğu için daha çok ev dekorasyon ürünlerinin öne çıkmasını beklerdik fakat daha çok kişisel hediyelerin satın alındığını gözlemliyoruz.

2) 3 Kral Bayramı / 6 Ocak

SQL QUERY:

with siralama as (SELECT to\_char(order\_approved\_at, 'YYYY-MM-DD' ) AS tarih,

count(DISTINCT orders.order\_id) as sold\_product\_count, product\_category\_name

from orders

join order\_items on orders.order\_id = order\_items.order\_id

join products on products.product\_id = order\_items.product\_id

WHERE to\_char(order\_approved\_at, 'YYYY-MM-DD' ) IN ('2017-01-06','2018-01-06')

group by 1,3

order by 1,2 desc)

select tarih, max(product\_category\_name) as product\_category\_name, sold\_product\_count

from siralama

group by 1,3

order by 1,3 desc

\*Grafikte ilk 10 kategori yer almaktadır.

Brezilya’da önemli olarak görülen bu bayramda ise daha çok dekorasyon ürünlerinin ve ev tekstil ürünlerinin öne çıktığı görülmektedir. Noel döneminde beklediğimiz hediye seçeneklerinin bu dönemde daha fazla olduğu görülmektedir.

3)Brezilya Aşıklar Günü / 12 Haziran

SQL QUERY:

with siralama as (SELECT to\_char(order\_approved\_at, 'YYYY-MM-DD' ) AS tarih,

count(DISTINCT orders.order\_id) as sold\_product\_count, product\_category\_name

from orders

join order\_items on orders.order\_id = order\_items.order\_id

join products on products.product\_id = order\_items.product\_id

WHERE to\_char(order\_approved\_at, 'YYYY-MM-DD' ) IN ('2017-06-12','2018-06-12')

group by 1,3

order by 1,2 desc)

select tarih, max(product\_category\_name) as product\_category\_name, sold\_product\_count

from siralama

group by 1,3

order by 1,3 desc

Brezilya’da aşıklar günü olarak kutlanan on iki haziran tarihi için bakacak olursak sağlık güzellik ürünleri, tekstil ürünleri ve saat gibi kişisel hediyelerin öne çıktığını görüyoruz. Kişisel hediyelerin alınması ise aslında beklenen bir sonuç.

-**SORU 4**

-Haftanın günleri (pazartesi, perşembe,..) ve ay günleri (ayın 1’i,2’si gibi) bazında order sayılarını inceleyiniz.

**SQL QUERY**:

Ayın günleri için dağılım

select count(DISTINCT order\_id) AS order\_count,

to\_char(order\_purchase\_timestamp, 'DD' ) AS ayin\_gunu

from orders

group by 2

ORDER BY 1 desc

|  |  |
| --- | --- |
| **order\_count** | **ayin\_gunu** |
| 3877 | 24 |
| 3581 | 16 |
| 3524 | 15 |
| 3483 | 4 |
| 3468 | 6 |
| 3445 | 5 |
| 3430 | 18 |
| 3387 | 14 |
| 3364 | 19 |
| 3363 | 7 |

\*Tablo ve grafikte en çok alışveriş yapılan 10 gün yer almaktadır.

Toplam verilen siparişlere baktığımızda Brezilya’da alışverişlerin ay ortasından sonra yapıldığı görülmektedir. Özellikle ayın 24. Günü verilen sipariş sayılarının diğer günlere fazla olması maaş ödemelerinin bu tarihe yakın yapıldığını varsaymamıza neden olacaktır.

Haftanın günleri için dağılım

select count(DISTINCT order\_id) AS order\_count,

to\_char(order\_purchase\_timestamp, 'DAY' ) AS haftanin\_gunu

from orders

group by 2

ORDER BY 1 DESC

|  |  |
| --- | --- |
| **order\_count** | **haftanin\_gunu** |
| 16196 | MONDAY |
| 15963 | TUESDAY |
| 15552 | WEDNESDAY |
| 14761 | THURSDAY |
| 14122 | FRIDAY |
| 11960 | SUNDAY |
| 10887 | SATURDAY |
|  |  |

Bu grafikte çoğunlukla hafta içi alışveriş yaptığı görülmektedir. İnsanların hafta içinde eksiklerini tamamlamak için da fazla zaman ayırdığını söyleyebiliriz.

**CASE 2 MÜŞTERİ ANALİZİ**

**-SORU 1**

-Hangi şehirlerdeki müşteriler daha çok alışveriş yapıyor? Müşterinin şehrini en çok sipariş verdiği şehir olarak belirleyip analizi ona göre yapınız.

**SQL QUERY**:

select

customer\_city,

sum(order\_count) as total\_order\_count

from(select

customer\_unique\_id,

order\_count,

customer\_city,

row\_number() over (partition by customer\_unique\_id order by order\_count ) as rn

from

(with siralama1 as (SELECT

customer\_unique\_id,

customer\_city,

count(distinct order\_id) as order\_count

from customers as cus

join orders as o on o.customer\_id=cus.customer\_id

group by 1,2)

select customer\_unique\_id,

customer\_city,

sum(order\_count) as order\_count

from siralama1

group by 1,2

order by 3 desc) as siralama2) as siralama3

group by 1

order by 2 desc

|  |  |
| --- | --- |
| **customer\_city** | **total\_order\_count** |
| sao paulo | 15540 |
| rio de janeiro | 6882 |
| belo horizonte | 2773 |
| brasilia | 2131 |
| curitiba | 1521 |
| campinas | 1444 |
| porto alegre | 1379 |
| salvador | 1245 |
| guarulhos | 1189 |
| sao bernardo do campo | 938 |

\*Tablo ve grafikte ilk 10 şehir gösterilmektedir.

Grafiği incelediğimizde geliri daha yüksek olan büyük şehirlerdeki müşterilerin daha çok alışveriş yaptığını görmekteyiz.

**CASE 3 SATICI ANALİZİ**

**-SORU 1**

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

**SQL QUERY**:

with top\_5 as

(select

COUNT(DISTINCT orders.order\_id) AS order\_count,

SELLERS.seller\_id as seller\_id,

AVG(AGE(order\_delivered\_customer\_date,order\_purchase\_timestamp)) as avg\_delivery\_time

from orders

join order\_items as oi on orders.order\_id= oi.order\_id

join sellers on sellers.seller\_id = oi.seller\_id

GROUP BY 2

HAVING COUNT(DISTINCT orders.order\_id) > 35

ORDER BY 3

LIMIT 5)

select

top\_5.seller\_id,

order\_count,

avg\_delivery\_time,

count(review\_comment\_message) as comment\_count,

round(avg(review\_score),2) as avg\_score

from top\_5

join order\_items as oi on oi.seller\_id= top\_5.seller\_id

join reviews on oi.order\_id= reviews.order\_id

group by 1,2,3

order by 3

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **seller\_id** | **order\_count** | **avg\_delivery\_time** | **comment\_count** | **avg\_score** |
| d13e50eaa47b4cbe9eb81465865d8cfc | 67 | 4 days 23:20:31 | 12 | 4,81 |
| 6b90f847357d8981edd79a1eb1bf0acb | 45 | 5 days 11:28:02 | 17 | 4,38 |
| d566c37fa119d5e66c4e9052e83ee4ea | 65 | 5 days 19:52:16 | 23 | 4,72 |
| 138dbe45fc62f1e244378131a6801526 | 49 | 5 days 29:02:19 | 24 | 3,98 |
| 325f3178fb58e2a9778334621eecdbf9 | 36 | 5 days 30:40:02 | 13 | 4,46 |

Ortalama sipariş sayısının otuz iki olduğu hesaplanarak tablo ve grafikte bu sayının üzerinde satış yapan ve ürünlerini en hızlı ulaştıran satıcılar yer almaktadır.

Genel olarak bakıldığında ise bu beş satıcının da birbirine yakın sipariş sayıları ve teslim süreleri bulunmaktadır. Ortalama skorları ve yorum sayılarında da bariz bir farklılık görülmemektedir fakat tablodaki ilk satıcı hem ortalama ürün sayısının iki katını satmış hem de en hızlı ulaştıran olmuştur. Ortalama en yüksek skor da ona aittir.

Tablodaki dördüncü satıcı ise aralarındaki en düşük skora sahip satıcıdır yorum sayısı da diğerlerine göre fazladır. Ürü şikayetinden kaynaklı yorum sayısı fazla olabilir bu yüzden de dört ortalamanın altında bir skora sahip olabilir.

**-SORU 2**

-Hangi satıcılar daha fazla kategoriye ait ürün satışı yapmaktadır?

Fazla kategoriye sahip satıcıların order sayıları da fazla mı?

**SQL QUERY**:

with tablo as (select

count(distinct orders.order\_id) as total\_order\_count,

oi.seller\_id,

rank()over(order by count(orders.order\_id) desc) seller\_rn

from orders

join order\_items as oi on orders.order\_id=oi.order\_id

group by 2

order by 1 desc)

SELECT

order\_items.seller\_id,

count(DISTINCT product\_category\_name) as total\_order\_category\_count,

tablo.total\_order\_count,

tablo.seller\_rn

from order\_items

join products on order\_items.product\_id=products.product\_id

join tablo on order\_items.seller\_id = tablo.seller\_id

group by 1,3,4

order by 2 DESC, 3 desc

|  |  |  |  |
| --- | --- | --- | --- |
| **seller\_id** | **total\_order\_category\_count** | **total\_order\_count** | **seller\_order\_rn** |
| b2ba3715d723d245138f291a6fe42594 | 27 | 337 | 46 |
| 955fee9216a65b617aa5c0531780ce60 | 23 | 1287 | 6 |
| 4e922959ae960d389249c378d1c939f5 | 23 | 420 | 30 |
| 1da3aeb70d7989d1e6d9b0e887f97c23 | 21 | 265 | 57 |
| f8db351d8c4c4c22c6835c19a46f01b0 | 19 | 667 | 18 |
| 18a349e75d307f4b4cc646a691ed4216 | 17 | 121 | 154 |
| 70a12e78e608ac31179aea7f8422044b | 15 | 315 | 57 |
| 6edacfd9f9074789dad6d62ba7950b9c | 15 | 208 | 90 |
| 7178f9f4dd81dcef02f62acdf8151e01 | 14 | 203 | 92 |
| 44073f8b7e41514de3b7815dd0237f4f | 14 | 144 | 150 |

\*Tablo ve grafikte ilk 10 satıcı yer almaktadır.

Tabloda yer alan 2 numaralı satıcı hem sipariş sayısı açısından hem de sipariş sayısına göre yapılan satıcılar arasındaki sıralamada ön plana çıkmaktadır.

Toplam satış yaptığı kategori sayısı da tabloya bakarsak yüksek olarak değerlendirilebilir. Toplamda 3095 satıcı olduğunu düşünürsek tablodaki satıcılar bu sıralama içeresinde ön plana çıkmaktadır. 2 numaralı satıcı dışında diğer satıcılar ise bulundukları kategorilerde sipariş sayılarını arttırarak toplam sıralamada üste çıkabilirler.

**CASE 4 PAYMENT ANALİZİ**

**-SORU 1**

-Ödeme yaparken taksit sayısı fazla olan kullanıcılar en çok hangi bölgede yaşamaktadır? Bu çıktıyı yorumlayınız.

**SQL QUERY**:

select

count(customer\_city)as more\_installment\_count,

customer\_city

from (select payment\_installments,

customer\_city,

payments.order\_id

from payments

join orders on orders.order\_id = payments.order\_id

join customers on customers.customer\_id = orders.customer\_id

where payment\_installments >6

order by 1 desc) as installment\_count

group by 2

order by 1 desc

|  |  |
| --- | --- |
| **more\_installment\_count** | **customer\_city** |
| 1437 | sao paulo |
| 846 | rio de janeiro |
| 341 | belo horizonte |
| 239 | brasilia |
| 198 | porto alegre |
| 190 | curitiba |
| 174 | salvador |
| 160 | campinas |
| 119 | guarulhos |
| 97 | recife |

\*Tablo ve grafikte ilk 10 şehir yer almaktadır.

Altı ay üzerindeki taksitlerin fazla kabul edilerek incelendiği bu grafikte, büyük şehirlerde yaşayan kişilerin gelirinin yüksek olması hem de daha fazla eşyaya sahip olma isteğini göz önüne alırsak şehirler arasında beklenen düzeyde bir sıralama oluşmuştur.

**-SORU 2**

-Ödeme tipine göre başarılı order sayısı ve toplam başarılı ödeme tutarını hesaplayınız. En çok kullanılan ödeme tipinden en az olana göre sıralayınız.

**SQL QUERY**:

select

payment\_type,

order\_status,

count(order\_status) as success\_order\_count,

sum(payment\_value) as total\_payment

from orders

join payments on orders.order\_id= payments.order\_id

where order\_status = 'delivered'

group by 1,2

order by 3 desc

|  |  |  |  |
| --- | --- | --- | --- |
| **payment\_type** | **order\_status** | **success\_order\_count** | **total\_payment** |
| credit\_card | delivered | 74586 | $ 12.101.094.879.999.800 |
| boleto | delivered | 19191 | $ 27.699.325.800.000.100 |
| voucher | delivered | 5493 | $ 3.430.131.900.000.000 |
| debit\_card | delivered | 1486 | $ 2.084.211.200.000.000 |

\*Başarı kriteri olarak teslim edilen siparişler baz alınmıştır.

Grafik ve tabloya baktığımızda en çok kullanılan ödeme yönteminin açık ara kredi kartı olduğunu görüyoruz, buna karşılık boleto ödeme yöntemi ile daha fazla miktarlarda alışveriş yapılmış bu sebeple de kredi kartı ile yapılan alışveriş tutarının iki katından daha fazla tutara sahiptir.

**-SORU 3**

-Tek çekimde ve taksitle ödenen siparişlerin kategori bazlı analizini yapınız. En çok hangi kategorilerde taksitle ödeme kullanılmaktadır?

**SQL QUERY**:

TEK ÇEKİM

select

product\_category\_name,category\_name\_english, count(DISTINCT pay.order\_id) as single\_paying\_count

from products

join order\_items as oi on oi.product\_id =products.product\_id

join payments as pay on oi.order\_id=pay.order\_id

join translation as tr on tr.category\_name = products.product\_category\_name

WHERE payment\_installments = 1

group by 1,2

order by 3 desc

|  |  |  |
| --- | --- | --- |
| **product\_category\_name** | **category\_name\_english** | **single\_paying\_count** |
| esporte\_lazer | sports\_leisure | 4299 |
| informatica\_acessorios | computers\_accessories | 4177 |
| beleza\_saude | health\_beauty | 3880 |
| cama\_mesa\_banho | bed\_bath\_table | 3535 |
| moveis\_decoracao | furniture\_decor | 3152 |
| utilidades\_domesticas | housewares | 2732 |
| telefonia | telephony | 2552 |
| eletronicos | electronics | 2052 |
| automotivo | auto | 2041 |
| brinquedos | toys | 1902 |

\*Tablo ve grafikte ilk 10 kategori yer almaktadır.

Tek çekimde yapılan alışverişlerde spor-eğlence ürünleri, bilgisayar aksesuarları ve sağlık-güzellik ürünleri ön plana çıkmaktadır.

TAKSİTLİ ALIŞVERİŞ

select

product\_category\_name, category\_name\_english, count(DISTINCT pay.order\_id) as installment\_paying\_count

from products

join order\_items as oi on oi.product\_id =products.product\_id

join payments as pay on oi.order\_id=pay.order\_id

join translation as tr on tr.category\_name = products.product\_category\_name

WHERE payment\_installments <>1

group by 1,2

order by 3 desc

|  |  |  |
| --- | --- | --- |
| **product\_category\_name** | **category\_name\_english** | **installment\_paying\_count** |
| cama\_mesa\_banho | bed\_bath\_table | 5965 |
| beleza\_saude | health\_beauty | 5006 |
| relogios\_presentes | watches\_gifts | 3794 |
| esporte\_lazer | sports\_leisure | 3480 |
| moveis\_decoracao | furniture\_decor | 3353 |
| utilidades\_domesticas | housewares | 3198 |
| informatica\_acessorios | computers\_accessories | 2562 |
| cool\_stuff | cool\_stuff | 2217 |
| brinquedos | toys | 2008 |
| perfumaria | perfumery | 1945 |

\*Tablo ve grafikte ilk 10 kategori yer almaktadır.

Taksitli yapılan alışverişlerde ise ev tekstil ürünleri, sağlık- güzellik ürünleri ve saatler ön plana çıkmaktadır.

**CASE 5 RFM ANALİZİ**

**SQL QUERY:**

with recency as ( with t1 as (SELECT customer\_id,

max(invoicedate) ::date as max\_date

from rfm

where quantity>0 and customer\_id IS NOT NULL

group by 1

order by 1)

select t1.customer\_id,

max\_date,

'2011-12-09' - max\_date as recency

from t1

order by 1

),

frequency as (SELECT customer\_id,

count(distinct invoiceno) AS frequency

from rfm

where QUANTITY >0 and customer\_id IS NOT NULL

group by 1

order by 2 desc),

monetary as (SELECT customer\_id,

sum(DISTINCT quantity\*unitprice) as monetary

from rfm

where QUANTITY >0 and customer\_id IS NOT NULL

group by 1

order by 1

), scores as (

select r.customer\_id,

r.recency,

ntile(5) over(order by recency desc, r.customer\_id) as recency\_score,

f.frequency,

case when frequency between 1 and 2 then '1'

when frequency between 3 and 5 then '2'

when frequency between 6 and 10 then '3'

when frequency between 11 and 20 then '4'

else 5 end as frequency\_score,

m.monetary,

ntile(5) over(order by monetary) as monetary\_score

from

recency as r

join frequency as f on r.customer\_id =f.customer\_id

join monetary as m on m.customer\_id=r.customer\_id)

select customer\_id,

recency\_score || '-' ||frequency\_score || '-'|| monetary\_score as rfm\_scores

from scores

order by 1