**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.

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

WITH order\_counts AS (

SELECT cc.customer\_id,

customer\_city,

COUNT(DISTINCT o.order\_id) ct\_order,

FROM customers c

LEFT JOIN orders o ON cc.customer\_id = o.customer\_id

GROUP BY 1,2

),

customer\_city\_rn AS(

SELECT

ROW () NUMBER OVER (PARTITION BY customer\_id ORDER BY ct\_order DESC) as rn,

customer\_id,

customer\_city,

FROM order\_counts

) ,

customer\_city AS(

SELECT cc.customer\_id,

customer\_city,

FROM customer\_city\_rn

WHERE rn=1

)

SELECT cc.customer\_city,

COUNT(o.order\_id)

FROM orders as o

LEFT JOIN customer\_city as cc

ON o.customer\_id =cc.customer\_id

GROUP BY 1

ORDER BY 2 DESC

**Case 3: Satıcı Analizi**

Question 1 :

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

WITH delivery AS (

SELECT

s.seller\_id,

COUNT(DISTINCT o.order\_id)::integer AS ct\_orders,

AVG(order\_delivered\_customer\_date - order\_approved\_at) AS avg\_delivery\_time\_interval,

CASE WHEN AVG(order\_delivered\_customer\_date - order\_approved\_at) <= '11 days' THEN 'quick\_delivery'

ELSE 'quick\_sellers1' END AS quick\_sellers,

CASE WHEN COUNT(DISTINCT o.order\_id)::integer >= 26200 THEN 'quick\_orders'

ELSE 'slow\_orders' END AS quick\_orders

FROM sellers s

LEFT JOIN order\_items oi ON oi.seller\_id = s.seller\_id

LEFT JOIN orders o ON o.order\_id = oi.order\_id

WHERE order\_approved\_at IS NOT NULL AND order\_status != 'canceled'

GROUP BY 1

LIMIT 5

)

SELECT s.seller\_id,

avg\_delivery\_time\_interval,

quick\_sellers,

quick\_orders,

SUM(ct\_orders::integer) AS total\_orders

FROM delivery d

LEFT JOIN sellers s ON d.seller\_id = s.seller\_id

LEFT JOIN order\_items oi ON s.seller\_id = oi.seller\_id

LEFT JOIN orders o ON o.order\_id = oi.order\_id

LEFT JOIN reviews r ON r.order\_id = o.order\_id

WHERE order\_approved\_at IS NOT NULL AND order\_status != 'canceled'

GROUP BY 1, 2, 3, 4

ORDER BY avg\_delivery\_time\_interval,quick\_sellers,quick\_orders

LIMIT 5;

OUTPUT: Tek satıra sığmadı çıktı.

"003554e2dce176b5555353e4f3555ac8" "4 days 15:12:46"

"0015a82c2db000af6aaaf3ae2ecb0532" "9 days 29:42:12.666667"

"001cca7ae9ae17fb1caed9dfb1094831" "12 days 15:25:58.363248"

"004c9cd9d87a3c30c522c48c4fc07416" "13 days 26:21:58.095809"

"002100f778ceb8431b7a1020ff7ab48f" "14 days 29:52:36.166667"

"quick\_delivery" "slow\_orders" 1

"quick\_delivery" "slow\_orders" 9

"quick\_sellers1" "slow\_orders" 47800

"quick\_sellers1" "slow\_orders" 27475

"quick\_sellers1" "slow\_orders" 2856

-----------------------------------------------------------------------------------------------------------

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

**WITH order\_delivery\_date AS(**

**SELECT oi.seller\_id,**

**extract(day from order\_delivered\_customer\_date -order\_purchase\_timestamp) as order\_delivery\_time**

**FROM orders as o**

**LEFT JOIN order\_items as oi**

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

**),**

**SELECT seller\_id,**

**avg (order\_delivery\_time) as avg\_delivery**

**FROM order\_delivery\_date**

**GROUP BY 1**

**ORDER BY 2**

**LIMIT 5**

**;**

---------------------------------------------------------------------------------------------------------------

Bu satıcıların order sayıları ile ürünlerindeki yorumlar ve puanlamaları inceleyiniz ve yorumlayınız.

WITH top\_5\_sellers AS(

WITH order\_delivery\_date AS(

SELECT oi.seller\_id,

extract(day from order\_delivered\_customer\_date -order\_purchase\_timestamp) as order\_delivery\_time,

FROM orders as o

LEFT JOIN order\_items as oi

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

)

SELECT seller\_id,

avg (order\_delivery\_time) as avg\_delivery

FROM order\_delivery\_date

GROUP BY 1

ORDER BY 2

LIMIT 5

)

SELECT t5.seller\_id,

COUNT ( DISTINCT oi.order\_id ) as order\_count,

AVG (r.review\_score) as avg\_score,

COUNT (DISTINCT review\_count\_message ) as total\_review\_message,

FROM **top\_5\_sellers** as t5

LEFT JOIN order\_items as oi

ON t5.seller\_id=oi.seller\_id

LEFT JOIN reviews as rr

ON r.order\_id=oi.order\_id

GROUP BY 1;

WITH delivery AS (

SELECT

s.seller\_id,

COUNT(DISTINCT o.order\_id)::integer AS ct\_orders,

AVG(order\_delivered\_customer\_date - order\_approved\_at) AS avg\_delivery\_time\_interval,

CASE WHEN AVG(order\_delivered\_customer\_date - order\_approved\_at) <= '11 days' THEN 'quick\_delivery'

ELSE 'quick\_sellers1' END AS quick\_sellers,

CASE WHEN COUNT(DISTINCT o.order\_id)::integer >= 26200 THEN 'quick\_orders'

ELSE 'slow\_orders' END AS quick\_orders

FROM sellers s

LEFT JOIN order\_items oi ON oi.seller\_id = s.seller\_id

LEFT JOIN orders o ON o.order\_id = oi.order\_id

WHERE order\_approved\_at IS NOT NULL AND order\_status != 'canceled'

GROUP BY 1

LIMIT 5

)

SELECT

s.seller\_id,

avg\_delivery\_time\_interval,

quick\_sellers,

quick\_orders,

SUM(ct\_orders::integer) AS total\_orders,

ROUND(AVG(review\_score), 2) AS avg\_score,

COUNT(DISTINCT review\_comment\_title)::integer AS comment\_count

FROM delivery d

LEFT JOIN sellers s ON d.seller\_id = s.seller\_id

LEFT JOIN order\_items oi ON s.seller\_id = oi.seller\_id

LEFT JOIN orders o ON o.order\_id = oi.order\_id

LEFT JOIN reviews r ON r.order\_id = o.order\_id

WHERE order\_approved\_at IS NOT NULL AND order\_status != 'canceled'

GROUP BY 1, 2, 3, 4

ORDER BY avg\_delivery\_time\_interval,quick\_sellers,quick\_orders

LIMIT 5;

-------------------------------------------------------------------------------------

OUTPUT: (tüm çıktı bu belgede tek bir satıra sığmadı maalesef!)

"003554e2dce176b5555353e4f3555ac8" "4 days 15:12:46"

"0015a82c2db000af6aaaf3ae2ecb0532" "9 days 29:42:12.666667"

"001cca7ae9ae17fb1caed9dfb1094831" "12 days 15:25:58.363248"

"004c9cd9d87a3c30c522c48c4fc07416" "13 days 26:21:58.095809"

"002100f778ceb8431b7a1020ff7ab48f" "14 days 29:52:36.166667"

OUTPUT DEVAMI:

"quick\_delivery" "slow\_orders" 1 5.00 0

"quick\_delivery" "slow\_orders" 9 3.67 0

"quick\_sellers1" "slow\_orders" 47800 3.90 3

"quick\_sellers1" "slow\_orders" 27475 4.14 3

"quick\_sellers1" "slow\_orders" 2856 3.98 0

NOT: Bu 3.kod diğer 2 koda göre daha doğru bir sonuç verdi. Ancak burda çıktıda 1.ve 2. sıradaki 1 ve 9 adet siparişi olan satıcılar tabii ki doğru değil. Ancak onları bu listeden yok edemedim!:(

---------------------------------------------------------------------------------------------------------------

YORUM:

Bu sorguyu yazmadan önce toplam sipariş adetini, ortalama sipariş teslim gününü ve 2.bir koşul olarak ortalama sipariş adetini median olarak hesapladım.

En hızlı şekilde sipariş ulaştıran müşterilerden ilk 5 ini getirdikten sonra bir sorun oluştu. Örneğin 1 adet siparişi var ancak ilk sütundaki satıcı, hızlı satıcı olarak çıktı. Bunun nedeni bu satıcı sitede yeni bir satıcı olup daha ilk ürününü satmış olabilir.

Buna ek olarak sadece kargo teslimat gününü hesaplamak tek başına bu sorgu için yetersiz kalıyor benim fikrime göre. Çünkü çok daha fazla sipariş teslimi başarılı olan satıcılarda mevcut çıktıda. Örneğin üstteki outputta fazla sipariş teslim eden ve ortalama sipariş teslim gün sayısını +1 ve +2 gün süreyle geçen satıcılar, çok daha fazla teslimat yapmış ve daha başarılı diğer satıcılara göre. Tabiiki bu satıcılar sitede ilk iki hızlı teslimat yapan satıcıya göre çok daha eski satıcı profiline sahip olabilir.

Ancak hızlı kargo teslim=diğer satıcılara göre çok daha az sürede/gün +saatte kargoyu müşteriye teslim eden satıcılar olarakta anlaşılma ihtimali oldukça yüksektir. Tam net bir cevap çıkmadı maalesef 2 koşul koyduğumda!

Şimdi çıktıda ilk satıcı 1 adet sipariş teslimi 4 günde yapmış. 5 puan almış. Müşteri ürün tesliminin hızlı olmasından ve üründen memnun kalmış. Ancak yorum yapmamış.

2.satıcımız 9 günde 9 ürün teslim etmiş. Nerdeyse her gün 1 adet ürün teslimi yapmış olabilir. 5 üzerinden 4 e yakın bir puan almış. Müşteri çocunlukla memnun kalmış.Teslimat açısından başarılı ancak ürün sayısı düşük! Hiç yorum almamış!

3.satıcımız ise 47.800 adetle en yüksek sipariş rakamını yakalamış. Ortalama 12 gün süreyle, 47.800 ürünü teslim etmiş. Nerdeyse 4 puan almış. Müşteri çoğunlukla

memnun kalmış. İlk 5 satıcının içinde bu satıcı en başarılısıdır. 3 tane yorumda almış. Bu tabiiki teslim edilen ürün adetine göre oldukça düşük bir rakam!

Hem teslimatı hemde ürün satış rakamı başarılıdır. Çoğunlukla müşteride memnun kalmıştır.

4. satıcımız 27.475 adet sipariş ile bu ilk 5'te bana göre 2. en başarılı satıcıdır. Oldukça fazla ürün siparişi almış. Yorumları da 4 ün üzerinde. Müşteri hem ürünlerden

hem de hızlı teslim sürecinden memnun kalmış belli ki. Ancak o da çok az yorum almış.

5. satıcımızda 2.865 adet ürün siparişi ile ilk 5 in içerisinde en başarılı ilk 3. üncü satıcı benim fikrime göre. 4 e yakın bir puan almış. Müşteri yorumu hiç yok!

Ortalama 14 günde ürünleri teslim etmiş. Burada kargo gün sayısı arttıkça çocğunlukla puanları olumsuz yönde etkilemiş.

Bunlara ilaveten aslında teslim edilen ürünlerin Brezilaya'nın hangi bölgedeki şehrine teslim ediliyor yani distance/mesafeleri de bu teslimatları etkileyen önemli bir KPI dır. Ancak elimizde bu veri detay olmadığı için bir yorum ya da hesaplama yapamadım.

Question 2 :

-Hangi satıcılar daha fazla kategoriye ait ürün satışı yapmaktadır?

SELECT s.seller\_id,

COUNT (DISTINCT p.product\_category\_name) total\_category,

COUNT(DISTINCT p.product\_id) ct\_products

FROM sellers s

LEFT JOIN order\_items oi ON s.seller\_id =oi.seller\_id

LEFT JOIN products p ON p.product\_id = oi.product\_id

GROUP BY 1

ORDER BY 2 DESC

;

OUTPUT:

"b2ba3715d723d245138f291a6fe42594" 27 183

"4e922959ae960d389249c378d1c939f5" 23 136

"955fee9216a65b617aa5c0531780ce60" 23 104

"1da3aeb70d7989d1e6d9b0e887f97c23" 21 155

"f8db351d8c4c4c22c6835c19a46f01b0" 19 146

"18a349e75d307f4b4cc646a691ed4216" 17 42

-------------------------------------------------------------------------------------------------

--Fazla kategoriye sahip satıcıların, order sayıları da fazla mı?

BAŞKA ÇÖZÜM:

**WITH category AS(**

**SELECT s.seller\_id,**

**COUNT (DISTINCT p.product\_category\_name) total\_category,**

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

**FROM order\_items as oi**

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

**GROUP BY 1**

**ORDER BY 2 DESC**

**;**

BENİM ÇÖZÜMÜM: UZUN!!!!

WITH category AS(

SELECT s.seller\_id,

COUNT (DISTINCT p.product\_category\_name) total\_category,

COUNT(DISTINCT p.product\_id) product\_count

FROM sellers s

LEFT JOIN order\_items oi ON s.seller\_id =oi.seller\_id

LEFT JOIN products p ON p.product\_id = oi.product\_id

GROUP BY 1

ORDER BY 2 DESC

)

SELECT s.seller\_id,

total\_category,

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

FROM category c

LEFT JOIN sellers s ON c.seller\_id=s.seller\_id

LEFT JOIN order\_items oi ON s.seller\_id =oi.seller\_id

LEFT JOIN orders o ON o.order\_id=oi.order\_id

LEFT JOIN products p ON p.product\_id = oi.product\_id

WHERE order\_approved\_at IS NOT NULL and order\_status != 'canceled'

GROUP BY 1,2

ORDER BY 2 DESC

;

OUTPUT:

"b2ba3715d723d245138f291a6fe42594" 27 333

"4e922959ae960d389249c378d1c939f5" 23 420

"955fee9216a65b617aa5c0531780ce60" 23 1286

"1da3aeb70d7989d1e6d9b0e887f97c23" 21 264

"f8db351d8c4c4c22c6835c19a46f01b0" 19 666

"18a349e75d307f4b4cc646a691ed4216" 17 119

YORUM:

Fazla kategoriye sahip olan satıcıların sipariş sayıları diğer satıcılara göre çoğunlukla daha fazladır. Bunun nedeni ürün çeşitliliğine sahip olmalarından kaynaklı olabilir. Örneğin; 23 kategoride ürün satışına sahip olan satıcıdan tam 1.286 adet ürün satın alınmıştır. Ancak çıktıda bazı istisnalar mevcuttur.

Örneğin aşağıdaki satıcı 7 kategoride ürün satmış . Ancak sipariş adeti 27 ya da 23 ya da 21 kategoriye sahip olan satıcıdan çok daha fazladır.

"4a3ca9315b744ce9f8e9374361493884" 7 1804

Bunun nedeni satıcının çoğunlukla insanların sık kullandığı ve hızlı tükettiği ya da yıprattığı ürünlerin kategorilerine ağırlık vermesi olabilir. Ya da sattığı kategorilerdeki

ürünler oldukça kaliteli olduğundan müşteri memnun kalıp tekrar sipariş etmesi dahilinde sipariş çıktı rakamı yüksek olabilir. Tüketiciye hitap eden, çoğunlukla ihtiyaç dahilinde olan ve günlük kullanımda hızlı yıpranan/tüketilen ürün siparişleri çok daha sık ve hızlı satılır. Bu satıcı kategorilerini oldukça başarılı seçmiştir.

Yine aşağıdaki veri çıktısında satıcı yanlızca 3 adet kategoride ürün satmış. Ancak daha fazla kategoriye sahip satıcılardan yine sipariş adeti bu satıcıdan oldukça daha azdır.

Son bir olasılık olarak daha az kategoriye sahip ancak daha fazla kategoriye sahip olan satıcılardan daha fazla satış yapan istisna olan satıcılar, diğer satıcılara kıyasla sitede oldukça eski kayıtlı satıcılar olabilir.

"1f50f920176fa81dab994f9023523100" 3 1403

DİP NOT:

Satıcıların fazla kategoride ürün satması bir avantaj ancak doğru kategori yani hızlı tüketilen, gündelik hayatta her gün kullanılan ürünler olmadığı zaman ya da ürün kalitesi düşük olunca sipariş adeti de düşebilir ya da daha az olabilir. Örneğin banyo masası,cool stuff malzemeler,pet shop ürünleri vb. sık kullanılıp hızlı yıpranan ya da hızlı tüketilen ürünler olduğu için satış hızı ve adeti daha fazla ve yüksektir. Ama bir air conditionerı, müşteri evinde bir oda yada 2-3 oda için 1 ya da 2-3 kez alabilir. Bozuk çıkarsa iade eder. Kaliteli ise yakınlarına tavsiye eder. Onlarda alabilir gibi... Satıcılar bu ürün kategorilerine dikkat etmelidir! Ancak doğrudan belli kategorilerdeki ürünleri üretiyorsa fabrikasında satıcı, tabii ki bu zor bir ihtimaldir!

**Case 4 : Payment Analizi**

Question 1 :

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

**DOĞRU ÇÖZÜM:**

SELECT payment\_installments,

c.customer\_city,

COUNT (DISTINCT o.customer\_id ) as customer\_count,

FROM payments as p

LEFT JOIN orders as o ON p.order\_id=o.order\_id

LEFT JOIN customers c ON c.customer\_id=o.customer\_id

GROUP BY 1,2

ORDER BY 1,3 DESC

--ORDER BY 3,1 DESC

;

Question 2 : \*\*\* her şirkette olan KPI DIR BU!!!!!

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

SELECT p.payment\_type,

COUNT(DISTINCT o.order\_id)::integer total\_orders,

ROUND (SUM (payment\_value::integer),2) total\_payments

FROM payments p

LEFT JOIN orders o ON o.order\_id = p.order\_id

WHERE o.order\_delivered\_customer\_date IS NOT NULL OR order\_status != 'canceled' AND p.payment\_value IS NOT NULL

GROUP BY 1

ORDER BY 1 DESC

;

SQL QUERY 2:

SELECT p.payment\_type,

SUM (CASE WHEN o.order\_status NOT IN ('canceled' , 'unavailable') THEN p.payment\_value END) as succ\_payment\_value,

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

FROM payments p

LEFT JOIN orders o ON o.order\_id = p.order\_id

GROUP BY 1

;

OUTPUT:

"voucher" 3772 353728.00

"debit\_card" 1521 215291.00

"credit\_card" 76067 12445771.00

"boleto" 19689 2851980.00

---------------------------------------------------------------------------------------------------------------

-En çok kullanılan ödeme tipinden en az olana göre