

עיבוד נתונים מתקדם – מועד א' 14:00-9:00 7.7.2024

הנחיות כלליות:

שם הקורס: עיבוד נתונים מתקדם.

קוד קורס: 150-330221

1. משך הבחינה: 3 שעות + תוספות זמן לזכאים.
2. שימוש לב: רכיב הבוחן פתוח לפחות זמן יותר מאשר הזמן המותר.
3. תגרור ציון 0 ללא בדיקת הבחינה – יש לדאוג ללחוץ על סיום בחינה והגשה בתום המענה.
4. בתחילת הבחינה יש לדאוג להריץ את קובץ הנתונים Northwind הקיים במודול.
5. הבחינה כוללת 10 שאלות:
 - a. 5 שאלות מעשיות, משקל כל שאלה 15 נקודות, סה"כ 75 נקודות.
 - b. 5 שאלות תיאורטיות, משקל כל שאלה 5 נקודות, סה"כ 25 נקודות.
6. יש לענות על השאלות המעשיות בתוכנת ה-SQL ולהעתיק את התשובה לרכיב הבוחן במודול. יש לשים לב שהעתקה מבוצעת לשאלה הנכונה – ערורים לא יתקבלו בגין טעות במיקום ההדבקה.
7. את השאלות התיאורטיות אפשר לענות ישירות ברכיב הבוחן, או לחילופין באקסל / וורד ולהעתיק את התשובה לרכיב הבוחן תחת השאלה המתאימה.
8. אין להתעסק בעיצוב רכיב הבוחן, אנו יודעים לקרוא את הקוד. חבל לשבץ זמן על כך.
9. בשם פנים ואפוי אין להדק תמונות ברכיב הבוחן!! הרכיב לא שומר תמונות ואנחנו נקבע רכיב ריק, שימוש לב כי לא ניתן לעורר במידה והדבקתם תמונה.
10. חומר פתוח עד 10 דפים דו-צדדיים.
11. יש לשמור את הקובץ עליו עובדים בכונן S כל כמה דקות /או אחרי כל שאלה כדי למצער את הנזק במקרה של הפסקת חשמל או תקלות אחרות.
12. אין צורך לצרף קבצים.
13. שאלות בבחינה לצוות הקורס יבוצעו בנוהל פתקים כמפורט:

כל שאלה של כל סטודנט תכתב על גבי "טופס" המצו依 אצל המশגיחים. צוות הקורס יעבור על הפטקים בכלל סבב.

שעות איסוף פתקים לנבחנים ב-9:00: 9:45, 10:30, 11:15.

12:15 (למי שיש התאמת למידה של הארכת זמן)

שעות איסוף פתקים לנבחנים ב-14:00: 14:45, 15:30, 16:15.

17:15 (למי שיש התאמת למידה של הארכת זמן)

חלק מעשי – 5 שאלות (15 נקודות לשאלת)

שאלה 1

סמכ"ל המכירות של החברה קיבל סכום כסף נכבד מהmarketיעים והחליט לפקח את העובדים הטובים ביותר בבורנוו שמן. עובד טוב ביותר הינו עובד אשר הפדיון מהמכירות שלו ברבעון מסוים בשנה מסוימת היה הגבוה ביותר.

לשם כך, פנה אליך בבקשה להוכיח לו דוח מכירות לכל העובדים עבור כל רביעון בכל שנה. עוד ביקש הסמכ"ל שתדאגו שבדו"ח העמודות יהיו הרבעונים ובשורות יהיה את שם

העובד, השנה וסך הפדיון מהמכירות שלו לאותו רביעון באותה שנה. אנו

במידה וישנם רביעונים שבהם העובד לא ביצע מכירות, יש לדאוג שתופיע הספרה 0. ← העומד

Coalesce
isnull (null, 0)

```
--Q1
--select employeeName, y,
--       coalesce([1],0) '1', isnull([2],0) '2', isnull([3],0) '3', isnull([4],0) '4'
from(
    select FirstName + ' ' + LastName employeeName, YEAR(OrderDate) y,
           DATEPART(QUARTER, OrderDate) qrt, UnitPrice*Quantity sumPidyon
    from Employees e join orders o
        on e.EmployeeID = o.EmployeeID
    join [Order Details] od
        on o.OrderID=od.OrderID) T
pivot(
    sum(sumPidyon) <=
        for qrt in ([1],[2],[3],[4])
)pivotTable
order by employeeName, y
```

employeeName	y	1	2	3	4
Andrew Fuller	1996	0.00	0.00	5940.80	16893.90
Andrew Fuller	1997	7639.30	25667.50	19999.75	21652.05
Andrew Fuller	1998	45101.41	34854.55	0.00	0.00
Anne Dodsworth	1996	0.00	0.00	4955.30	6410.40
Anne Dodsworth	1997	2079.30	4512.30	12368.40	9717.55
Anne Dodsworth	1998	32519.25	9501.50	0.00	0.00

שאלה 2

כתבו שאלתה המציגת את 3 אנשי המכירות (Sales Representative) בעלי ביצועי המכירות הטובים ביותר בתבסס על סך הפדיון מהמכירות שנבעו מההזמנות שהם טיפולו.

ע"י שימוש ב- CTE יש להראות לכל איש מכירות, את שמו, כמות ההזמנות שהם טיפולו. כמות הלקוחות בהםם טיפול, סך הפדיון של העובד מתוך סך הפדיון הכללי בארגון.

תרומות הלקוחות מחושבת כמספר של העובד מתוך סך הפדיון הכללי בארגון. ← group by

יש לדאוג למצב בו איש המכירות לא טיפול באף הזמנה או הזמינות שאין להם פדיון, במקרה,

יש להוציא מהדו"ח אנשי מכירות שלא טיפול באף הזמנה.

יש להציג את 3 העובדים בסדר יורד לפי כמות ההזמנות order by

הערה: על מנת להציג את 3 אנשי המכירות הטובים ביותר יש להשתמש בפונקציה row

.number

```
-- Q2
with CTE2
as(
    select e.EmployeeID, FirstName + ' ' + LastName fullName, Title,
           COUNT( distinct o.OrderID) totalOrders,
           COUNT( distinct CustomerID) distinctCustomers,
           sum(UnitPrice*Quantity) revenue,
           ROW_NUMBER() over( order by COUNT( distinct o.OrderID) desc) ranking
    from Employees e left join Orders o
        on e.EmployeeID=o.EmployeeID
    left join [Order Details] od
        on o.OrderID=od.OrderID
    group by e.EmployeeID ,FirstName + ' ' + LastName, Title
)

select fullName,totalOrders,distinctCustomers,revenue, revenue / (select sum(UnitPrice*Quantity)
                                                               from [Order Details]) contribution
from CTE2
where totalOrders>0 and
      title like 'Sales Representative' and
      ranking <=3
order by totalOrders desc
```

fullName	totalOrders	distinctCustomers	revenue	contribution
Margaret Peacock	156	75	250187.45	0.1847
Janet Leverling	127	63	213051.30	0.1572
Nancy Davolio	123	65	202143.71	0.1492

3 שאלה

לאחר שיבת סעורה בין המנכ"ל לכיספים, החליטו השנאים שצרכ' להגדיל את היקף הקניות של החברה בזרה ממשמעותית ויביעו מחרות בין הליקוחות השונים על היקף הניכוי שרם מבצעים, בסוף התחרות נתנו להבנה משמעותית ל-5 המנכ"לים. התבקשתם ע"י מנהלת מעדן מהלקלת מעדן להכלן לשניי ח"ח שיציג את 5 הליקוחות בעלי סך הדרויים בגובה בשנה האחוריה (גינוי שהתאריך של היום הימנ"ל 1997-07-07). לכל אחד מלהקוחות אלו יש להציג:

- מס' לקוחות.
- ✓ שם לקוחות.
- ✓ סך הדרויים.
- ✓ כמהו ההזמנות של הלקוח.
- ✓ דירוג הלקוח בהתקבוס על סך הדרויים.

בהתבסס על תוצאות אלו סמנכ"ל הכספיים והמנכ"ל יכול לחתה מלהקוחות שצ'ו ב-5 המקומות הראשוניים, וכך עם הלהקוחות יוכל לעתה כהה חסר להם לעמוד הלקוח שזכה במקום אחד מעלייהם.

group by
count distinct

לפ"ט

```
--Q3
select CustomerID, CompanyName, revenue, nuOfOrders, revenueRanking, revenue - LAG(revenue,1) over(order by revenue desc) 'diff From Precedur Customer'
from(
    select c.CustomerID, CompanyName, sum(UnitPrice*Quantity) revenue, count(distinct o.OrderID) nuOfOrders,
    ROW_NUMBER() over(order by sum(UnitPrice*Quantity) desc) revenueRanking
    from Customers c join Orders o
        on c.CustomerID=o.CustomerID
    join [Order Details] od
        on o.OrderID=od.OrderID
    where o.OrderDate between '1996-07-07' and '1997-07-07'
    group by c.CustomerID, CompanyName
) t
where revenueRanking<=5
```

CustomerID	CompanyName	revenue	nuOfOrders	revenueRanking	diff From Precedur Customer
1	QUICK	49342.30	13	1	NULL
2	ERNSH	37749.00	12	2	-11593.30
3	SAVEA	28499.54	7	3	-9249.46
4	RATTC	27502.30	10	4	-997.24
5	MEREP	23281.90	9	5	-4220.40

4 שאלה

צור פונקציה בשם GetTopCustomersByCategory המתקבלת 2 משתנים:

- תאריך תחילת סוג Date
 - תאריך סיום סוג Date
- על הפונקציה להחזיר לכל קטגוריה את 3 הלהקוחות שלם ממוצרי הקטגוריה הנימם מהגובהים ביחס בזמן בתווך התאריכים שיבחר.

לכל לקוחות יש להציג:
order
orderID
customer
1. מס' לקוחות {
2. שם החברה {
3. מספר קטגוריה {
4. שם הקטגוריה {
5. סך הדרויים של הלקוח ממוצרי הקטגוריה

6. את דירוג הדרויים שלו בתוך הקטגוריה לפ"ט לפ"ט
הפרצתוריה צריכה להחזיר את התוצאות לפי הדירוג בתוך כל קטגוריה בסדר עולה.
יש לבדוק את הפונקציה על כל שנת 1997

```
--Q4
create or alter procedure getTopCustomersByCategory @startDate Date, @endDate Date
as
begin
    select*
    from(
        select c.CompanyName, c.CustomerID, cat.CategoryID, cat.CategoryName, sum(od.UnitPrice *Quantity) sumRev,
        dense_rank() over(partition by (cat.CategoryID) order by sum(od.UnitPrice *Quantity) desc) catRank
        from Categories cat join Products p
            on cat.CategoryID=p.CategoryID
        join [Order Details] od
            on p.ProductID = od.ProductID
        join Orders o
            on o.OrderID=od.OrderID
        join Customers c
            on c.CustomerID = o.CustomerID
        where o.OrderDate between @startDate and @endDate
        group by cat.CategoryID, cat.CategoryName, c.CompanyName, c.CustomerID )q
        where catRank<=3
        order by CategoryID, catRank
    end
execute getTopCustomersByCategory '1997-01-01', '1997-12-31'
```

89 %

CompanyName	CustomerID	CategoryID	CategoryName	sumRev	catRank
1 QUICK-Stop	QUICK	1	Beverages	12257.00	1
2 M?re Paillarde	MEREP	1	Beverages	11665.70	2
3 Simons bistro	SIMOB	1	Beverages	10540.00	3
4 Ernst Handel	ERNSH	2	Condiments	8906.70	1
5 QUICK-Stop	QUICK	2	Condiments	7451.70	2
6 Save-a-lot Markets	SAVEA	2	Condiments	4377.00	3
7 QUICK-Stop	QUICK	3	Confections	11671.60	1
8 Ernst Handel	ERNSH	3	Confections	10736.20	2

מנול מוחלת המכירות מעוניין לנתח את נתוני המכירות של לקוחות שתרומות לארגן גובוהה מ-8% (בأfon ליל' לא קשור לתקופות). התנונה לארגון מוגדרת בסכום הפדיון של הלוקו ביחס לסכום הפדיון הכללי לארגון. א. אלו נזק ב Select בהתאם לדרישתך או של המנהל, שיוציאו זו"ח השוואה עברו כל הלוקו, של הפדיון בשני רביעונים ממוצע בהמשך ולחשב את סכום התרומות שהזינו בשני הרביעונים ביחס. יש לך חישובן אך ורק לתקופות קיימות (כלומר, לא יכולה שהנתונים שלהם חסרים באחת התקופאות). איזט נזק

תוצאות הדוח ימינו בהתאם למספר ההזדמנות שביצעו הלוקו, מהכמות הנמוכה ביותר לבוגהו ביחס.

הערה: ברישום 8% יש לרשותם 0.08

```
--Q5
select c.CompanyName, sumRev1_1997, sumRev3_1997, countOrder1_1997-countOrder3_1997 countOrders,
       countTotalOrders, tot.sumRev / (select sum(UnitPrice*Quantity)
                                         from [Order Details]) contribution
from Customers c join (select customerID, sum(UnitPrice*Quantity) sumRev1_1997,
                           count(distinct o.OrderID)/countOrder1_1997
                         from Orders o
                         join [Order Details] od
                           on o.OrderID=od.OrderID
                         where DATEPART(QUARTER, OrderDate) = 1
                               and Year(OrderDate) = 1997
                         group by CustomerID)q1_1997
on c.CustomerID=q1_1997.CustomerID
join (
      select customerID, sum(UnitPrice*Quantity) sumRev3_1997,
             count(distinct o.OrderID)/countOrder3_1997
           from Orders o
           join [Order Details] od
             on o.OrderID=od.OrderID
           where DATEPART(QUARTER, OrderDate) = 3
                 and Year(OrderDate) = 1997
           group by CustomerID)q3_1997
on q1_1997.CustomerID=q3_1997.CustomerID
join (select CustomerID, count(distinct o.OrderID) countTotalOrders,
            sum(UnitPrice*Quantity) sumRev
          from Orders o join [Order Details] od
                        on o.OrderID=od.OrderID
            group by CustomerID) tot
on q3_1997.CustomerID=tot.CustomerID
where tot.sumRev / (select sum(UnitPrice*Quantity)
                     from [Order Details]) >0.08
```

	CompanyName	sumRev1_1997	sumRev3_1997	countOrders	countTotalOrders	contribution
1	Ernst Handel	11307.40	14704.05	7	30	0.0836
2	QUICK-Stop	6092.20	8568.00	4	28	0.0867
3	Save-a-lot Markets	7889.10	25251.00	8	31	0.0854

הדו"ח יכול את הפריטים הבאים:
 א. שם חברת הלוקו - Customer
 ב. סכום הפדיון בربعון 1 של שנת 1997
 ג. סכום הפדיון בربعון 3 של שנת 1997
 ד. מספר ההזדמנות של הלוקו בשני הרביעונים הנ"ל
 ה. מספר ההזדמנות של הלוקו באופן כללי לא קשור לרביעונים
 ו. תרומת הלוקו לארגון

(1) נפח בף שיקויים גודלו נסמן בפונקציית COUNT()

(2) מיפוי וJOIN עם WHERE

(3) WHERE, GROUP BY

(4) CASE WHEN

(5) CASE WHEN Select AS

```
select CustomerID,
       case when sum(UnitPrice*Quantity)>10000 then 'platinum'
             when sum(UnitPrice*Quantity) between 5000 and 10000 then 'gold'
             when sum(UnitPrice*Quantity)between 1000 and 5000 then 'platinum'
             else 'bronze'
       end customerRank
from orders o join [Order Details] od
               on o.OrderID=od.OrderID
where YEAR(OrderDate)=(YEAR(GETDATE()) -1)
group by CustomerID
```

חלק תיאורי – 5 שאלות (5 נקודות לשאלת)

בثور מנהלי מערכות המידע בראשות מסחר אינטרנטית, התבקשנות לסואג ללקוחות לפי מדריך

המדד והקריטריונים כדלקמן:

- פלטינום: סך הכנסות מכירות גבוהה מ-10,000 דולר.
- זהב: סך הכנסות מכירות בין 5,000 דולר ל-10,000 דולר. מושג case
- כסף: סך הכנסות מכירות בין 1,000 דולר ל-5,000 דולר.
- ברונזה: סך הכנסות מכירות נמוך מ-1,000 דולר.

הסבירו שלב אחריו שלב כיצד תבצעו את סיווג הלוקוות לקритריונים השונים באמצעות SQL ללא כתיבת קוד.

בזאת שיטות סטטיסטיות נבדקו סימני תבואה מסוימים, כגון גיל, נסיעה ועוד, על מנת להעניק ללקוחות מרכיב שתקזב אחר הסקרים ששולאו. מנהלת הספריה לעזרה לה

זמןאים להשכלה בהרגע שספר השכר או הוחרה.

הסבירו שלב אחר שבו ציד תטמעו את התוכנית שמקשת מנהלת הספריה באמצעות טריגרים ללא כתיבת קוד.

הערה: מותר להניח הנחות לגבי טבלאות ועמודות, רק יש לציין את ההנחות.

לפחות 2/3 מהלקוחות נסמן כ'כסף' או 'ברונזה'. רגע מילוי הטעינה נסמן כ'כסף' או 'ברונזה'.

רבע מילוי הטעינה נסמן כ'כסף' או 'ברונזה'.

תיאורטית 3

בהתור אנגליסטים תוחחים בחנות מסחר מוביל בשוק, קיבלתם משימת חינכה לעבד חדש שהגיעו ואותם רוצים ל夺取 פוריקט שקיבלו מנהל החנות. החנות מעוניינת להזות את המוצרים הנמכרים בייצור בכל חדש.

הסבירו לעובד הצעיר שאתם חונכים כיצד יש לבצע את השאלתה הנ"ל ע"י שימוש בפונקציות חלון שלב אחר שלב ללא כתיבת קוד.

select * from row where row_number = 1;

```
--Q3
select ProductName, ROW_NUMBER() over(partition by MONTH(o.orderDate) order by sum(quantity*od.unitPrice) desc) ranking
from Products p join [Order Details] od
  on p.ProductID=od.ProductID
  join Orders o
  on o.OrderID=od.OrderID
group by ProductName, MONTH(o.orderDate)
```

הערה: מותר להניח הנחות, לציין כאשר ביצעתם זאת.

תיאורטית 4

מה ההבדל בין CTE לחת שאלתה ב-From ?(Derived Table) יש לנו מינות לפחות 3 הבדלים.

from - a subject like an animal or a person can be from a place with { } . CTE (4) and CTE (3) are used here.

תיאורטית 5

```

select ProductName,
       UnitPrice
  from Products
 where UnitPrice <
        (SELECT UnitPrice
         from Products
        where ProductName like 'Ravioli Angelo')
      and
        UnitsInStock > (SELECT avg(UnitsInStock)
                           from Products) +20
  order by ProductName

```

ඩීප් මාලා ගුරු

මානුව සැකිල්ල නියමිත තේව්‍ය ප්‍රස්ථාන නිරූපණය කළ ඇති ප්‍රාග්ධනය

רשות את השאלה עבורה בזע הפטרואן

הערה: אין לרשום את תיאור השאלה. כלומר, מה כל שורה עשו.

