

# Version 2

## עיבוד נתונים מתקדם – מועד א' 19:00-14:00 7.7.2024

הנחיות כלליות:

שם הקורס: עיבוד נתונים מתקדם.

קוד קורס: 150-330221

1. משך הבדיקה: 3 שעות + תוספות זמן לזכאים.
  2. שימוש לב: רכיב הבוחן פתוח לזמן ארוך יותר מאשר זמן הבדיקה. אולם, הגשה לאחר הזמן המותר **תגרור ציון 0 ללא בדיקת הבדיקה – יש לדאוג ללחוץ על סיום בחינה והגשה בתום המענה.**
  3. בתחילת הבדיקה יש לדאוג להריץ את קובץ הנתונים Northwind השמור בכונן D וקיים במודול.
  4. יש לענות על כל השאלות. אין שאלות בחירה.
  5. הבדיקה 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 המשישות בתוכנת-hSQL ולהעתיק את התשובה לרכיב הבוחן במודול. יש לשים לב שהעתקה מבוצעת לשאלה הנכונה – ערורים לא יתקבלו בגין טעות במיקום הבדיקה.
  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 לbezבז זמן על כך.
  9. בשות פנים ואופן אין להדק תמונות ברכיב הבוחן!! הרכיב לא שומר תמונות ואנוanno נקבע רכיב ריק, שימוש לב כי לא ניתן לעורר במידה והדבקתם תמונה.
  10. חומר פתוח עד 10 דפים דו-צדדיים.
  11. יש לשמר את הקובץ עליו עובדים כל כמה דקות ו/או אחרי כל שאלה כדי למדוע את הנזק במקרה של הפסקת חשמל או תקלות אחרות.
  12. אין צורך לצרף קבצים.
  13. שאלות בבדיקה לצוות הקורס יבוצעו בנהל פתקים כdllkmn:
- 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)

### שאלה 1

```
--Q1
select CompanyName, ContactName, Y,
       ISNULL([1],0)'1', ISNULL([2],0)'2', ISNULL([3],0)'3', ISNULL([4],0)'4'
  from (select CompanyName, ContactName, YEAR(OrderDate) Y,
              DATEPART(QUARTER, OrderDate) q, Quantity*UnitPrice totalSales
            from Customers c join orders o
              on c.CustomerID=o.CustomerID
             join [Order Details] od
                on o.OrderID=od.OrderID)src
 pivot(avg([totalSales])
      for q in ([1],[2],[3],[4]))t
```

סמנכ"ל תחת המכירות של החברה חוששת שישנה מגמה שלילית לחילוק מליקות החברה. מוגמת שלילית הנה ירידת ממוצע המכירות של הלוקוח על פני הרכבעים השונים בשנים בהם היה לוקוח החברה. לשם כך, פנתה אליך בבקשה להזכיר לה זו ד"ח מכירות לכל הלוקוחות עבור כל רביען בכל שנה. עד ביקשה הסמנכ"לית שתתגאו שבד"ח העמודות יהיו הרבעונים ובשורות יהיה את שם החברה, שם איש הקשר, השנה וממוצע המכירות שלו לאוטו רביען באותה שנה. במידה ושניהם רביעונים שבהם הלוקוח לא ביצע מכירות, יש לדאוג שתופיע הספרה 0.

```
--Q2
with Top5ProfitProducts
as
(
    select p.ProductID, ProductName, sum(Quantity) sumQuan,
           count(distinct CustomerID) countCustomers,
           sum(Quantity*od.UnitPrice) sumRev,
           ROW_NUMBER() over(order by sum(Quantity*od.UnitPrice) desc) rowNum
      from orders o left join [Order Details] OD
        on O.OrderID=OD.OrderID
       left join Products p
         on p.ProductID=od.ProductID
     group by p.ProductID, ProductName
)
select *, round(sumRev/(select sum(Quantity*UnitPrice) from [Order Details])*100,2) contribution
from Top5ProfitProducts
where sumQuan>0 and rowNum<=5
order by sumQuan
```

### orders, OD , product

#### שאלה 2

כתבו שאלתת המציג את 5 המוצרים שהניבו את הפדיון*היכי גבוח* בחברה. ע"י שימוש ב- CTE יש להראות את קוד מוצר, את קוד המוצר, שם המוצר, סך הכמות (quantity) שנמכרו ממנו, כמות הלוקוחות שכשו אותו, סך ההכנסות מכירותו ואת התרומה מהכנסות שלו. תרומות הכנסות מחושבת כסך הפדיון של המוצר מטור סך הפדיון הכללי בארגון. יש לדאוג למצב בו מוצר לא נמכר כלל ובוסף, יש להוציא מהדו"ח מוצרים שלא נמכרו כלל. יש להציג את 5 המוצרים בסדר יורד לפ' סך הפדיון מהמכירות. הערה: על מנת להציג את 5 המוצרים הנמכרים ביותר יש להשתמש בפונקציה *row\_number*.

```
--Q3
select*, avgMonthlyRev- LEAD(avgMonthlyRev,1) over(order by rowNum) diff
from(
    select e.EmployeeID, FirstName + ' ' + LastName fullName,
           MONTH(OrderDate) orderMonth,
           avg(Quantity*od.UnitPrice) avgMonthlyRev,
           count(distinct O.OrderID) countOrders,
           ROW_NUMBER() over(order by AVG(Quantity*od.UnitPrice) desc) rowNum
      from Employees e join orders o
        on e.EmployeeID=o.EmployeeID
       join [Order Details] OD
         on O.OrderID=OD.OrderID
     where OrderDate between '1996-07-07' and '1997-07-07'
   group by e.EmployeeID, FirstName + ' ' + LastName, MONTH(OrderDate))a
where rowNum<=5
```

לאחר שיבת שערת בין סמנכ"ל הכספיים למנחתת משבבי אנוש, החליטו השנים שצרכו להגדיל את ריקף המכירות של החברה בזורה ממשמעוויות ויחלטו לבצע תחרות בין העובדים השונים על היקף המכירות שלהם מבצעים, בסוף התחרות נתנו בונוס שמן 5-5 המנצחים.

על מנת לעוזר להם לראות מי 5 העובדים המנצחים בעלי ממוצע הפדיון*היכי גבוח* בשנה[\(נניח שהתאריך של יומם הינו 07-07-1997\)](#), מנהלת מערכות המידע ביקשה

מספר ד"ח שיציג לכל עובד את: 20-07-2006 ✓

- מספרו. א' ל' ↗ employee ↘
  - שמו המלא.
  - ממוצע הפדיון *חדש* של העובד ↗ OD ↘
  - כמות הזמנות שבצעו ע"י העובד. ↗ partition ↘
  - דירוג העובד *חדש* שבסדר החודשי. נ' פאיין → - ↗ ↘
  - ההפרטן שדורג לראשונה בחודש.
- העובד שעבד בתבוסס על ממוצע הפדיון החדש. נ' פאיין → - ↗ ↘
- בהתבסס על תוצאות אלו סמנכ"ל הכספיים ומורת מתן שמאנו יכולים לחתם בונוס לעובדים שזכו ב-5 המקבילות הראשונות, וכך שמי העובדים יוכל לדעת כמה היה חסר להם לעומת העובד שזכה באותו מקום אחד מתחתיים.

```
--Q4
create or alter procedure GetEmployeeOrderSummery @empID int
as
begin
    select o.OrderID, OrderDate, CompanyName,
           count(distinct ProductID) numOfProd,
           sum(Quantity*od.UnitPrice) sumRev,
           Rank() over(order by sum(Quantity*od.UnitPrice) desc) orderRank
      from Customers c join Orders o
        on c.CustomerID = o.CustomerID
       join [Order Details] od
         on o.OrderID=od.OrderID
     where o.EmployeeID = @empID
   group by o.OrderID, OrderDate, CompanyName
end
exec GetEmployeeOrderSummery 2
```

שאליה 4  
צורו פרוצדורה בשם *GetEmployeeOrderSummery* המקבלת משתנה אחד:

- מספר עובד מסווג int

על הפרוצדורה להחזיר את סיכום הזמן של העובד.

כל עובד יש להציג:

1. מס' הזמנה

2. תאריך הזמנה

3. שם הלוקוח

4. מספר המוצרים בהזמנה

5. סך הכנסות מהזמן

6. דירוג הפדיון בהתייחס לכל הזמן בהם טיפול העובד ↗ ↘

Rank ↗ Row Num ↘  
יש לבדוק פרוצדורה זאת על עובד מס' 2

מן-היל מחלוקת המכירות מעוניין לנתח את נתוני המכירות של לקוחות שתורמתם לאגרן **(באזור כליל קשור לתקופות)**. התורמה לארגן מוגדרת כסכום הפדיון נוכחית מ-2%

Select → entire page

**לעומת** של הלקוק ביחס לסטטוס הפלידי הכללי בארץ. בהתחם לדרישה זו של המנהל, יש ליצור דוח השוואה עבור כל הליקוות, של הפלידי בשני רבעונים מכיוון שהמשר ולחשב את מספר ההזמנות שהצמינו בשני הרבעונים בידך. יש לקחת בחשבון אך ורק לקוחות קיימים (כלומר, לא כolumbia שהנתנו להם חסרים באחת הרשוואות).

תוצאות הדוח ימיינו בהתאם למספר ההזדמנויות שביבוצעו הלוקחות, מהכמויות הנמצאה ביחס

**Order by** → לגובהה ביותר.

הערה: ברישום 2% יש לרשום 0.02

הדו"ח יכול את הפרטים הבאים:

**✓ א. שם חברת הלקוח**

✓ ב. סכום הפדיון ברבעון 2 של

✓ ג. סכום הפטורון ברבעון 3 של שנת 1997

✓ ד. מספר ההזדמנויות של הלקוח בשני הרבעים

ה. מספר ההזמנה של הח

```
--Q5
select CompanyName, sumRev_1998_2,sumRev_1997_3, countOrder_1998_2+countOrder_1997_3 totalOrdersInQuarters_2_3,
       totRev/(select count(Quantity*UnitPrice)
                  from Orders o join [Order Details] od
                                on o.OrderID=od.OrderID) teruma
from Customers c join (select o.CustomerID, sum(Quantity*UnitPrice) sumRev_1998_2,
                               count(distinct o.OrderID) countOrder_1998_2
                          from orders o join [Order Details] od
                            on o.OrderID=od.OrderID
                          where year(o.OrderDate)=1998 and DATEPART(QUARTER, o.OrderDate) =2
                          group by o.CustomerID
                        )Q_2_1998
on c.CustomerID=Q_2_1998.CustomerID
join (
      select o.CustomerID, sum(Quantity*UnitPrice) sumRev_1997_3,
             count(distinct o.OrderID) countOrder_1997_3
            from orders o join [Order Details] od
              on o.OrderID=od.OrderID
              where year(o.OrderDate)=1997 and DATEPART(QUARTER, o.OrderDate) =3
              group by o.CustomerID
        ) Q_3_1997
on Q_2_1998.CustomerID=Q_3_1997.CustomerID
join (
      select o.customerID, sum(Quantity*UnitPrice) totRev, count(distinct o.orderID) totalOrders
         from orders o join [Order Details] od
           on o.OrderID=od.OrderID
           group by o.customerID
     ) Q_all
on Q_3_1997.CustomerID=Q_all.CustomerID
where totRev/
      (select sum(Quantity*UnitPrice)
         from Orders o join [Order Details] od
           on o.OrderID=od.OrderID)
<0.02
```

### חלק תאורטי – 5 שאלות (5 נקודות לשאלת)

וְאֵת שֶׁבּוֹנָה כִּי-כַּאֲמָרָה בְּשָׂרְבָּרָה  
בְּעֵדָה וְבְּמִזְבֵּחַ כִּי-כַּאֲמָרָה בְּשָׂרְבָּרָה  
בְּעֵדָה וְבְּמִזְבֵּחַ כִּי-כַּאֲמָרָה בְּשָׂרְבָּרָה

### תיאורית 1

בثور מנהלי מערכות המידע באוניברסיטה מובילה, התקבךשותם לשוג סטודנטים (GPA = Grade Point Average).  
מלגות המבוסס על ממוצע הציונים שלהם.

- המדרג והקריטריונים כדלקמן:
  - פלטינום: GPA שווה או גבוהה מ-4.0
  - זהב: GPA בין 3.5 ל-3.99
  - כסף: GPA בין 3.0 ל-3.49
  - ברונזה: GPA נמוך מ-3.0

הסבירו בשלב אחרי שלב כיצד תבצעו את סיווג הסטודנטים לקריטריונים השונים באמצעות SQL ללא כתיבת קוד.

על מנת להטמייע את המערךת כך שבאופן אוטומטי ישמרו הלוגים של כל השינויים שמבצעים במהלך המוצר ולעתק את מנהל המחלקה כאשר ישינו קייזון במחיר, נשימוש בטראגרים.

1. ניצור טריגר עבור שינוי במחיר:
  - a. ניצור טריגר בשם LogPriceChange שמודיע לאחר פעולה עדכן על `updatedOn` בटבלת `Price` בעמודת `Product`.
  - b. בתוך הטריגר, סופר שנוחה לטבלת `PriceChangeLog` המזקיקה נתונים מוצר כגון: מספר המוצר, מחיר ישן ומחר חדש ותאריך ועתה השינוי.
2. ניצור טריגר עבור עדכן מנהל המחלקה:
  - a. ניצור טריגר בשם NotifyAdminOnPriceChange שמודיע לאחר פעולה `updatedOn` בטבלת `PriceChangeLog`.
  - b. בתוך הטריגר, נבדוק אם השינוי במחיר מרשים (לדוגמא, עלייה של 100% במחיר).
  - c. אם השינוי קייזון, נשלח הודעה למנהל המחלקה (לדוגמא אימייל).

תיאורית 2

בתווך עובי מערוכות מידע בחברת eCommerce, התבקשתם ע"י סמכ"ל השיווק והמכירות לעזרו להם לבנות מערכת שתציג את הלוגים של כל השינויים שבוצעים לפחות מהאזור המוצר ומאת הנהלה הממליצה כאשר יש שינוי במחירים.

הסבירו שלב אחר ציד תטמעו את התוכנית שבבקש סמכ"ל השיווק והמכירות באמצעות ריגרים ללא תכנית גוד.

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

תיאורטיות 6

בBOR אנליסטים תוחמים בחברת פיננסים, קיבילתם משימת חניכה לעובד חדש שהגיעה ואתם רצחים למתן לו פרויקט שקיבילתם ממנהל המஸחר. המנהל רצה לנתח דפומי עסקאות של לקוחות ולראות אם הלקוחות עם התנהגות הוצאות אסורה.

הסבירו לעוד חמוץ שאתם זוכרים כיצד יש לבצע את השאלתה הנ"ל ע"י שימוש בפתרונות חלון שב אחר שלב לא כתיבת קו.

הערה: מותר להניח הנחות, לפחות כאשר ביצעתם זאת.

## תיאורטית 4

מה ההבדל בין CTE לdatable זמניות (temp tables)?

יש למנות לפחות 3 הבדלים.

1. היקף:  
 a. CTEs רק בוטה הביצוע של השאלתה שבה הן מוגדרות. אין מאחסנות למשימות ומוחקות לאחר הרשות ביצוע השאלתית.  
 b. tempdb שטבלאות מנוגות במקודם התאריכים שבעדtempdb ויכולים להמשר מעבר לביצוע של שאלתה בודדת. הן מנוגות ב�רך כל ליל הפעלה או שאלתה שיש להרשאות גישה מתאימות.

- a.** שאלות מושגשיות בעיקר כדי לפחות שאלות מורכבות על ידי מtran אפשרות להגדירה של שאלות מורכבות שונות שיתן להפנות אליהן בתוך השאלות הראשונות.

**b.** טבלאות דינמיות מושגשיות לאחיזון תוצאות בণיאים או פירוק פעולות מורכבות שללבם טקסטים ייחודיים ונומינטיביים. ניתן ליעזר, לאחסן ותעדן זומרם כטבלאות רגולריות.

- a. מוגדרות באמצעות מילת המפתח WITH בתחילת שאלתה, ואחריה שם CTE והגדرتה

b. טבלאות זמניות נוצרות באמצעות משפט CREATE TABLE, עם קידומת אופיינית של # (יחד עם טבלאות זמינות מקוניות או ## עבור טבלאות זמניות גלובליות).

4. אחסון נתונים:  
 a. ain מאחסנת נתונים בעצם; הם רק דרך להציג ערכת נתונים  
 זמינים שיבת הפקות אליה בתוך שאילתאות.  
 b. טבלאות מיוחדות מאחסנת נתונים באופן מתמשך בסיס נתונים של tempdb עד שום נשים טריים במשמעות או שההפעלה מסתיימת.

- a. CTEs. כפופות לאוטום כליל, בקורס ישנה כמו השאלתה הראשית. על המשתמשים להיות בעלי הרשות מתחילה למשתמשים לשלבאות הבסיסיות ולבצע את השאליתה המכליה את ה-CTE.

ב-. לטבלאות דומות יש קבצת הרשות משללים וניתן להעניק או לדוחות אותן במפורש למשתמשים או לתפקידים.

## תיאורטיות 5

```
select ProductName,  
       UnitPrice  
  from Products  
 where UnitPrice <  
           (SELECT UnitPrice  
            from Products  
           where ProductName like 'NuNuCa Nu?-Nougat-Creme')  
      and  
       UnitsInStock > (SELECT avg(UnitsInStock)  
                         from Products) -20  
order by ProductName
```

הוּא 88 מילוי ↓  
הוּא מילוי ↓  
הוּא מילוי ↓  
הוּא מילוי ↓

רשום את השאלה עבורה בוצע הפתרון.

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