

## הנחיות כלליות

1. ב厰בחן זה תעבדו עם שתי טבלאות:

o **db\_imdb** – טבלת סרטים.

o **disasters** – טבלת אסונות.

2. הגשה: אתם תקבלו קובץ מוחלך לפי שאלות, שככל שאלה יש מקום ל:

o פקודות **SQL** אחת או יותר שיענו על השאלה (חייבת להיות בפורמט שיאפשר ללקחת אותה ולהריץ אותה כמו שהיא - והיא תעבוד)

o את התוצאה שהתקבל מהרצת השאלתה (מספר / ערך מתאים).

לדוגמה:

-- Question 1

-- SQL:

<sql here>

-- Answer: <answer here>

ועם ערכיהם:

-- Question 1

-- SQL:

SELECT COUNT(\*) from products;

-- Answer: 232

3. הקפידו על קוד מסודר, קריא ומודוק.

## חלק א' – שאלות על טבלת הסרטים (imdb)

ענו על השאלות הבאות על בסיס טבלת **imdb**.

1. כמה סרטים יש להם דירוג **IMDB** גובה מ-8?
2. מהו דירוג ה-IMDB הגבוה ביותר שמופיע במערכת הנתונים?
3. כמה סרטים שייכים לזרן **Drama**?
4. כמה סרטים שייכים לזרן **Action**?
5. כמה סרטים יצאו בשנת **2010**?
6. מהו זמן הסרט (**Runtime**) המינימלי שמופיע בטבלה?
7. כמה סרטים יש להם ציון **Meta Score** גובה מ-85?
8. כמה סרטים יש להם ערך **Gross** גובה מ-100 מיליון?
9. מהו דירוג ה-IMDB הממוצע של כל הסרטים בטבלה?
10. כמה סרטים קיבלו תעודה צפיה (**Certificate** מסוג **'R'**)?
11. איזה ז'אנר (**Genre**) הוא הז'אנר שמופיע הכי הרבה פעמים בטבלה?
12. כמה סרטים יש מהז'אנר של "**Adventure**" ויצאו אחרי **2010** והדירוג שלהם הוא לא **PG**?
13. כמה סרטים יש להם זמן צפיה (**Runtime**) בין **95** ל-**140** דקות (כולל)?
14. איזו שנת יצאה (**Release Year**) אחרי 1940 (לא כולל 1940) היא השנה שבה יש הכי מעט סרטים?
15. מהו ערך ה-**Gross** המקורי, מבין סרטים בז'אנר **Comedy**?
16. כמה סרטים יש להם דירוג **IMDB** בין **7.3** ל-**8.1** (כולל)?
17. מהו ציון **Meta Score** הממוצע של סרטים בז'אנר **Thriller**?
18. כמה סרטים יצאו לפני שנת **1980**?
19. איזה ז'אנר (**Genre**) הוא הז'אנר בעל דירוג ה-IMDB הגבוה הנמוך ביותר?
20. מהו זמן הסרט (**Runtime**) הארוך ביותר מבין סרטים בז'אנר **Horror**?

## חלק ב' – שאלות על טבלת האסונות (disasters)

ענו על השאלות הבאות על בסיס טבלת **disasters**. היעזרו בפונקציה YEAR כדי לחלץ את השנה כshaذا נדרש.

21. כמה אסונות התרחשו בשנת **2020**?

22. מהו מספר הנפגעים (**casualties**) המקסימלי שנרשם באירוע אסון אחד כלשהו בטבלה?

23. כמה אסונות סוגם כסוג "**Flood**"?

24. כמה אסונות סוגם כסוג "**Earthquake**" והתרכשו ב-**India** או **Brazil**?

25. מהו מספר הנפגעים הממוצע (**average casualties**) באירועים מסווג **Hurricane** או **Wildfire** כאשר זמן התגובה (**response time**) היה **פחות מ-10 שעות**?

26. כמה אסונות יש להם **אפס נפגעים** (**0 casualties**)?

27. מהו ערך הנזק הכלכלי (**economic loss**) המינימלי שדווח במאגר?

28. כמה אירועים גרמו לנזק כלכלי גדול מ-**1 מיליארד (10⁹) דולר**?

29. באיזו שנה התרחש **המספר הגדל ביותר של אסונות**?

30. מהי השנה המוקדמת ביותר שמויפה בטבלה שבה: מדד החומרה (**severity index**) גבוה מ-**4** וגם סכום הסיוע שניתן (**aid given**) גדול מ-**50,000**?

31. איזה סוג אסון (**Disaster Type**) הוא הסוג שמויף כי הרבה פעמים בטבלה?

32. כמה אירועי אסון יש שבהם סוג האסון (**Disaster Type**) מכיל את המילה "**Storm**"?

33. כמה אסונות היו עם מספר נפגעים (**casualties**) בין **50 ל-500** (כולל)?

34. איזו היא השנה עם **מספר האסונות הכולל הנמור ביותר**?

35. מהו ערך הנזק הכלכלי (**economic loss**) הגבוה ביותר מבין כל האירועים מסווג **Flood**?

36. כמה אירועי אסון התרחשו בין השנים **2010–2024** (כולל), כאשר מדד החומרה (**severity**) גדול מ-**5**?

37. מהו מספר הנפגעים הממוצע באירוע **Earthquake** שבו: מספר ימי ההחלמה (**recovery**) ובו **レスポンס טיים (response time)** היה בין **5 ל-15 שעות (days)**?

38. מהו ערך הנזק הכלכלי (**economic loss**) הגבוה ביותר מבין האסונות בעלי **מדד החומרה הגבוה ביותר** (**ה-severity index המקסימלי**)?

39. מהו האסון שבו **זמן הטיפול הכולל** (מספר ימי ההחלמה + זמן התגובה בשעות) ביחידות של **ימים** היה **הנמוך ביותר?**

40. באיזו שנה התרחש האסון שגרם **לנזק הכספי הכולל הגבוה ביותר**, כאשר מסכמים **economic** ?**loss + aid given**

## חלק ג' – שאלת מחקר: קשר בין מספר אסונות לדינרים של סרטים

### הנחה מחקר

הנחה המחקר: בשנים שבהן יש מספר גדול של אסונות, יוצאים יותר סרטים **Action** מאשר **Comedy**.

ביצוע החלק מחולק לשני שלבים:

1. **הנתון הנטוניים** (עבודה עם טבלאות מסכמות).
2. **בדיקה ההנחה** בעזרת NOIJ ושאלות נוספות נוספות.

---

### חלק ג'.1 – הנתון הנטוניים

א. ייצרת טבלה מסכמת לסרטים לפי שנה

צרו טבלה חדשה שתסכם את טבלת **imdb** לפי שנה.  
טבלה צריכה לכלול:

- עמודת שנה (**year\_value**)
- מספר סרטים **Action** שייצאו באותה שנה (**action\_count**)
- מספר סרטים **Comedy** שייצאו באותה שנה (**comedy\_count**)

יש לוודא כי לכל שנה קיימת שורה אחת בלבד. הפקודות שמייצרות את הטבלה נמצאות בדף הבא.

השתמשו בסקריפט הבא (התאימו שמות טבלאות/עמודות במידת הצורך):

```
CREATE TABLE imdb_yearly AS  
  
SELECT  
  
    Released_Year AS year_value,  
  
    SUM(  
  
        CASE  
  
            WHEN Genre LIKE '%Action%'  
  
                AND Genre NOT LIKE '%Comedy%'  
  
            THEN 1 ELSE 0  
  
        END  
  
    ) AS action_count,  
  
    SUM(  
  
        CASE  
  
            WHEN Genre LIKE '%Comedy%'  
  
                AND Genre NOT LIKE '%Action%'  
  
            THEN 1 ELSE 0  
  
        END  
  
    ) AS comedy_count  
  
FROM imdb  
  
GROUP BY Released_Year;
```

## ב. הוספה עמודה חדשה לatable האסונות

1. הוסיפו לatable **disasters** עמודה חדשה שתכיל רק את שנת האירוע, למשל בשם **.year\_value**
2. מלאו את העמודה על בסיס עמודת התאריך הקיימת בטבלה (Date), בעזרת:  
**STR\_TO\_DATE ( . . . )** להמרת מחזורת לתאריך (אם צריך) ○

## ◦ YEAR ( . . . ) לשילוף השנה מן התאריך

### ג. יצירת טבלת סיכום אסונות לפי שנה

צרו טבלה חדשה שתסכם את טבלת **disasters** לפי שנה.  
הטבלה צריכה לכלול:

- עמודת שנה (**year\_value**) – מתוך העמודה שהוספتم בסעיף הקודם.
  - מספר האסונות הכלול בשנה (**disaster\_count**).
- הטבלה תיווצר מຕוך SELECT (ראו נספח בסוף המבחן).

### חלק ג' 2 – בדיקת ההנחה בעזרת JOIN

בשלב זה עליכם להשתמש בטבלאות:

- סיכום סרטים לפי שנה (**imdb\_yearly**)
  - סיכום אסונות לפי שנה (**disasters\_yearly**)
1. בצעו JOIN בין שתי הטבלאות כך שלכל שנה יהיה:
    - מספר סרטי Action
    - מספר סרטי Comedy
    - מספר האסונות באותה שנה
  2. כתבו שאלות נוספות לפי בחירתכם, על מנת לבדוק האם הנתונים תומכים בהנחה המחקר.
    - בסיום, ענו **במילים** על השאלה:

האם לדעתכם הנתונים תומכים בהנחה שבשנים עם מספר רב של אסונות יוצאים יותר סרטי Action מאשר סרטי Comedy?

יש לנמק על בסיס התוצאות שהתקבלו מהשאלות שביבעתם (מספרים, ממוצעים, השוואות וכו').

4. יש להציג את כל פקודות ה-SQL שהרצתם בחלק זה:

- ייצרת הטבלאות המוסכמת (CREATE TABLE ... AS SELECT)
- עדכון עמודת השנה
- כל שאלתה שבה השתמשת לצורך בדיקת המחקר

## **נספח – ייצרת טבלה מຕוך SELECT**

לעתים נרצה לשמר תוצאה של שאלתה כטבלה חדשה, לצורך עבודה נוחה יותר בהמשך.  
ניתן לעשות זאת באמצעות התחביר:

```
CREATE TABLE table_name AS  
SELECT ...  
FROM ...  
WHERE ...;
```

## משימת בונוס – יצירת טבלה תשובות לבחן

במשימה זו תיצורו טבלה ייעודית שתשמש כטבלת הגשה של כל תשובות המבחן.

1. צרו טבלה חדשה בשם **answers** שתכלול את העמודות הבאות:

שם העמודה	סוג נתונים	תיאור
full_name	VARCHAR(...)	שם התלמיד
class	VARCHAR(...)	כיתה
id_number	VARCHAR(...)	ת"ז
question_num	INT	מספר השאלה בבחן
sql_command	TEXT	הملאה שנכתבה SQL פקודה ה- לתשובה
answer_value	TEXT	התוצאה שהתקבלה מהרצת השאלה

2. לאחר יצירת הטבלה, עליכם **להכניס לתוכה רשותה עבור כל שאלה בבחן** (כל שאלה בחלק א', חלק ב', וחלק ג' אם יש).

3. כל רשותה חייבת לכלול:

○ את פרטי התלמיד (שם, כיתה, ת"ז)

○ מספר השאלה

○ פקודה ה-SQL המלא שאתם הגשתם

○ התוצאה שהתקבלה בהרצת

4. בסיום המשימה, כל התשובות של כל חלקי המבחן נמצאות בטבלה אחת מסודרת.