

הנחיות כלליות

1. במבחן זה תעבדו עם שתי טבלאות:

○ **imdb** – טבלת סרטים.

○ **disasters** – טבלת אסונות.

2. הגשה: אתם תקבלו קובץ מחולק לפי שאלות, שבכל שאלה יש מקום ל:

○ פקודת **SQL** אחת או יותר שעונה על השאלה (חייבת להיות בפורמט שיאפשר לקחת אותה ולהריץ אותה כמות שהיא - והיא תעבוד)

○ את התוצאה שהתקבלה מהרצת השאילתה (מספר / ערך מתאים).

לדוגמה:

-- Question 1

-- SQL:

<sql here>

-- Answer: <answer here>

ועם ערכים:

-- Question 1

-- SQL:

SELECT COUNT(*) from products;

-- Answer: 232

3. הקפידו על קוד מסודר, קריא ומדויק.

חלק א' – שאלות על טבלת הסרטים (imdb)

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

1. כמה סרטים יש להם דירוג **IMDB** גבוה מ-8?
2. מהו דירוג ה-**IMDB** הגבוה ביותר שמופיע במערכת הנתונים?
3. כמה סרטים שייכים לז'אנר **Drama**?
4. כמה סרטים שייכים לז'אנר **Action**?
5. כמה סרטים יצאו בשנת **2010**?
6. מהו זמן הסרט (**Runtime**) המינימלי שמופיע בטבלה?
7. כמה סרטים יש להם ציון **Meta Score** גבוה מ-85?
8. כמה סרטים יש להם ערך **Gross** גבוה מ-100 מיליון?
9. מהו דירוג ה-**IMDB** הממוצע של כל הסרטים בטבלה?
10. כמה סרטים קיבלו תעודת צפייה (**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 צפייה (**Runtime**) בין 95 ל-140 דקות (כולל)?
14. איזו שנת יציאה (**Release Year**) אחרי 1940 (לא כולל 1940) היא השנה שבה יש הכי מעט סרטים?
15. מהו ערך ה-**Gross** המקסימלי מבין סרטים בז'אנר **Comedy**?
16. כמה סרטים יש להם דירוג **IMDB** בין 7.3 ל-8.1 (כולל)?
17. מהו ציון ה-**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** הממוצע 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 כשזה נדרש.

21. כמה אסונות התרחשו בשנת 2020?
22. מהו מספר הנפגעים (**casualties**) המקסימלי שנרשם באירוע אסון אחד כלשהו בטבלה?
23. כמה אסונות סווגו כסוג "Flood"?
24. כמה אסונות סווגו כסוג "Earthquake" והתרחשו ב-Brazil או India?
25. מהו מספר הנפגעים הממוצע (**average casualties**) באירועים מסוג Wildfire או Hurricane, כאשר זמן התגובה (**response time**) היה פחות מ-10 שעות?
26. כמה אסונות יש להם אפס נפגעים (**casualties 0**)?
27. מהו ערך הנזק הכלכלי (**economic loss**) המינימלי שדווח במאגר?
28. כמה אירועים גרמו לנזק כלכלי הגדול מ-1 מיליארד (10^9) דולר?
29. באיזו שנה התרחש המספר הגדול ביותר של אסונות?
30. מהי השנה המוקדמת ביותר שמופיעה 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 שמופיע הכי 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 שנה היא השנה עם מספר האסונות הכולל הנמוך ביותר?
35. מהו ערך הנזק הכלכלי (**economic loss**) הגבוה ביותר מבין כל האירועים מסוג Flood?
36. כמה אירועי אסון התרחשו בין השנים 2010–2024 (כולל), כאשר מדד החומרה (**severity**) גדול מ-5?
37. מהו מספר הנפגעים הממוצע באירועי Earthquake שבהם: מספר ימי ההחלמה (**recovery days**) גדול מ-40 או זמן התגובה (**response time**) היה בין 5 ל-15 שעות (כולל)?
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 ל**נזק הכספי הכולל הגבוה ביותר**, כאשר מסכמים **economic loss + aid given**?

חלק ג' – שאלת מחקר: קשר בין מספר אסונות לז'אנרים של סרטים

הנחת מחקר

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

ביצוע החלק מחולק לשני שלבים:

1. הכנת הנתונים (עבודה עם טבלאות מסכמות).
2. בדיקת ההנחה בעזרת JOIN ושאלות נוספות.

חלק ג'.1 – הכנת הנתונים

א. יצירת טבלה מסכמת לסרטים לפי שנה

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

- עמודת שנה (**year_value**)
- מספר סרטי **Action** שיצאו באותה שנה (**action_count**)
- מספר סרטי **Comedy** שיצאו 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. הוספת עמודת שנה לטבלת האסונות

1. הוסיפו לטבלת **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 מחרוזת לתאריך (אם צריך)

○ YEAR(...) לשליפת 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 שהוספתם בסעיף הקודם.

- מספר האסונות הכולל בשנה (**disaster_count**).

הטבלה תיווצר מתוך SELECT (ראו נספח בסוף המבחן).

חלק ג' 2. – בדיקת ההנחה בעזרת JOIN

בשלב זה עליכם להשתמש בטבלאות:

- **imdb_yearly** – סיכום סרטים לפי שנה

- **disasters_yearly** – סיכום אסונות לפי שנה

1. בצעו **JOIN** בין שתי הטבלאות כך שלכל שנה יהיו:

- מספר סרטי Action

- מספר סרטי Comedy

- מספר האסונות באותה שנה

2. כתבו שאילתות נוספות לפי בחירתכם, על מנת לבדוק האם הנתונים תומכים בהנחת המחקר.

3. 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 מהשאילתות שביצעתם (מספרים, ממוצעים, השוואות וכו').

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

- יצירת הטבלאות המסכמות (CREATE TABLE ... AS SELECT)

- עדכון עמודת השנה

- כל שאילתה שבה השתמשותם לצורך בדיקת המחקר

נספח – יצירת טבלה מתוך SELECT

לעיתים נרצה לשמור תוצאה של שאילתה כטבלה חדשה, לצורך עבודה נוחה יותר בהמשך.
ניתן לעשות זאת בעזרת התחביר:

```
CREATE TABLE table_name AS
```

```
SELECT ...
```

```
FROM ...
```

```
WHERE ...;
```


משימת בונוס – יצירת טבלת תשובות למבחן

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

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

שם העמודה	סוג נתון	תיאור
full_name	VARCHAR(...)	שם התלמיד
class	VARCHAR(...)	כיתה
id_number	VARCHAR(...)	ת"ז
question_num	INT	מספר השאלה במבחן
sql_command	TEXT	המלאה שנכתבה SQL פקודת ה- לתשובה
answer_value	TEXT	התוצאה שהתקבלה מהרצת השאילתה

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

3. כל רשומה חייבת לכלול:

- את פרטי התלמיד (שם, כיתה, ת"ז)
- מספר השאלה
- פקודת ה-SQL המלאה שאתם הגשתם
- התוצאה שהתקבלה בהרצה

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