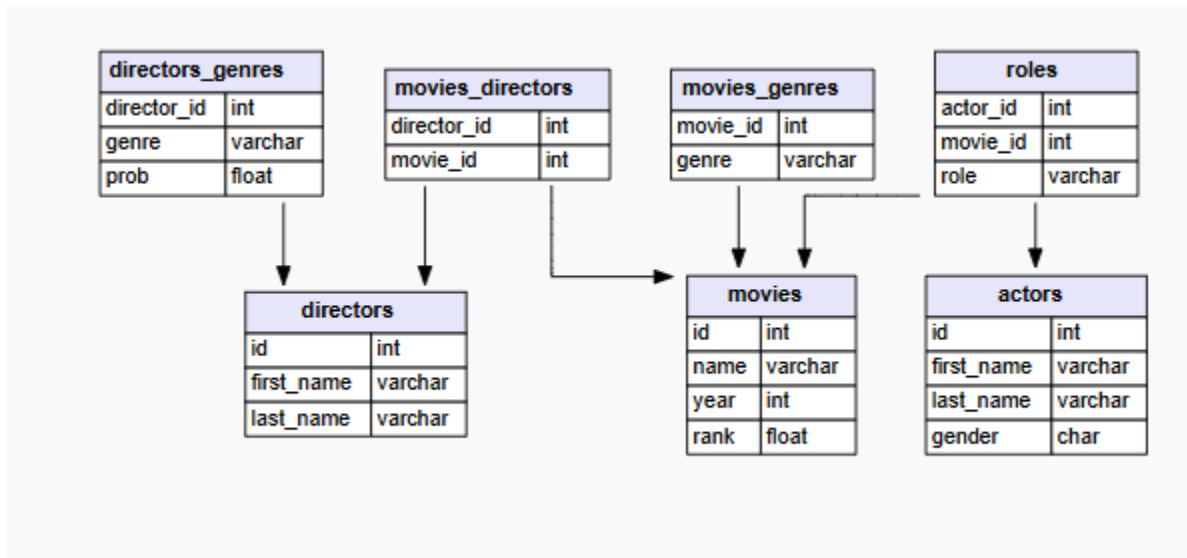


Databases for data analytics

דוגמת בחינה

השאלות בבחינה תתייחסנה לסכמה של [IMDB](#) המובאת בתרשים למטה.



1. שליפות

במהלך הבחינה מותר, ואפילו רצוי, להשתמש בשאלות שכתבתם קודם. כדאי לעשות זאת אנא השתמשו ב views וציינו בשאלתה המשתמשת את המקום בו הגדרתם אותה.

במידה ויש סעיף שהתקשתם בפיתורו, אתם יכולים לציין שם של view שפותר אותו, לכתוב מה השדות בו, ולהשתמש בו בסעיפים אחרים.

- שילפו את שם של movie שמספרו 7
- שילפו את ה rank המקסימלי של movie
- שילפו את ה roles שמעולם לא היו בסרט מסוג comedy
- שילפו את מספר ה genre השונים לכל actor
- עבור movies שיש להם actors, שילפו את ה actors per movie distribution. כלומר, מספר הפעמים שבסרט היה שחקן 1, 2, 3 וכדומה.
- הרחיבו את הסעיף הקודם ל movies ללא שחקנים.
- הציגו את זוגות ה directors שביימו סרט באותו השם
- מה מספר התפקידים המקסימלי לשחקן בסרט
- בנוס: מיצאו את הבמאי ששיחק הכי הרבה בסרטיו

2. תקינות נתונים

- a. האם יתכנו genre שאין להם סרטים? מה היתרונות והחסרונות ביצוג הזה.
- b. שדה genre בטבלת movies_genres הוא מסוג varchar. האם כדאי להחליפו בשדה המפוענח כ foreign key מטבלה חדשה של ז'אנרים? יש נימוקים טובים לתשובה חיובית ולשלילית. הציון לא ינתן על התשובה אלא על הנימוקים. אנא התייחסו ל-
- i. חשיבות השדה
 - ii. סט הערכים אפשרי
 - iii. משמעות הטעות ואופן הטיפול בה
 - iv. האפשרויות אחרות למנוע שגיאות
- c. נחליף את יצוג ה genre לטבלה
- i. כיתבו שאילתה היוצרת טבלת genres. ניתן להשתמש ב
ROW_NUMBER() OVER (ORDER BY name) AS row_num
כדי לקבל מפתח מספרי בשם row_num עבור כל רשומה
 - ii. יצרו טבלת new_movies_genres בה ה genre מיוצג בעזרת טבלת ה genre.
- d. איזה סוג יחס יש בין movies ו director (אחד לאחד, אחד לרבים, רבים לרבים)
- e. איזה סוג יחס יש בין movie ו genre? במה תלויה תשובתכם?
- f. מה הבעיה בטבלת directors_genres. מה התועלת בשימוש בה?

3. עיצוב

עצבו מערכת עבור קופת חולים.

המערכת צריכה לתמוך בשימושים הבאים.

1. הצגת כל הרופאים של מטופל, הטיפולים שקיבל, ותחום הטיפול במטופל. שימו לב שרופא לבצע טיפולים שונים ומטופל יכול להעזר ברופאים שונים.
 2. הצגת כל הבדיקות מסוג מסויים שרופא ביקש.
 3. הצגת כלל הבדיקות של מטופל ותוצאותן.
-
1. כיתבו קוד ליצירת הטבלאות הנדרשות. הקפידו על נרמול.
 2. ציינו מה היחסים בין הישויות (1:1, 1:n, או n:n)
 3. אנא נמקו את החלוקה לטבלאות, טיפוס השדות שבחרת, מפתחות, foreign keys, constraints.
 4. נמקו את היצוג של היחס (רופא, מטופל, תחום). פרטו מה היתרונות והחסרונות של בחירתכם.
 5. נמקו את היצוג של בדיקות ותוצאות. פרטו מה היתרונות והחסרונות של בחירתכם.
 6. ממשו את השאילתה "הצגת כל הבדיקות ללא תוצאה" על בסיס הנתונים שלכם.
 7. בונסו: ממשו את השאילתה "בדיקות שתוצאתן טובה מכל הבדיקות העבר". הניחו שכלל שתוצאה בבדיקה נמוכה יותר היא טובה יותר