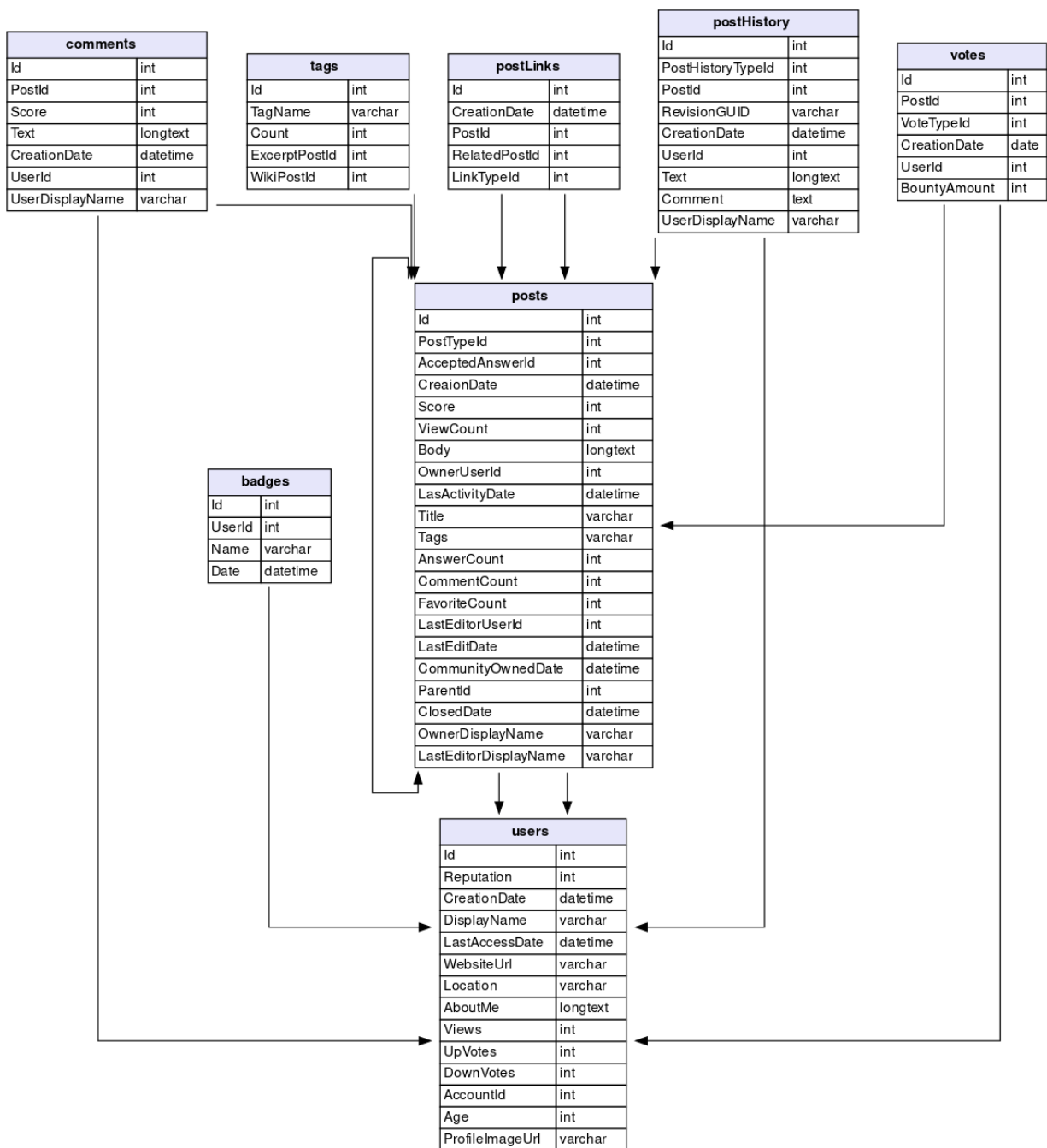


גישה לנתונים | - מסדי נתונים ושפת SQL

Databases for data analytics

מועד א

השאלות בבחינה תתייחסנה לסכמה של [StackOverflow](https://stackoverflow.com) המובאת בתרשים למטה.



1. שליפות (35 נקודות)

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 בפתרונו, אתם יכולים לציין שם של view שפותר אותו, לכתוב מה השדות בו, ולהשתמש בו בסעיפים אחרים.

- a. שילפו את מספר ה tag ששמו 'Sql'
- b. שילפו את מספר ה users
- c. שילפו את ה users שהם לא ה owners של אף post
- d. שילפו את מספר ה comments לכל post, בעזרת טבלת comments
- e. עבור posts שקיבלו הערות, שילפו את ה comments per post distributions.
- f. כלומר, מספר הפעמים ש post קיבל הערה 1, 2, 3 וכדומה.
- g. הרחיבו את הסעיף הקודם ל posts שלא קיבלו הערות.
- g. הציגו את זוגות ה post ids שיותר מ 10 אנשים זהים הצביעו לשניהם (כלומר אותו אדם הצביע לשני ה posts). שילפו באופן סימטרי, כלומר (a,b) וגם (b,a)

2. תקינות נתונים (30 נקודות)

a. ב StackOverflow כותבים שונים יכולים ליצור posts. האם כדאי ש title יהיה מפתח ראשי בטבלת posts? יש נימוקים טובים לתשובה חיובית ולשלילית. הציון לא ינתן על התשובה אלא על הנימוקים. אנא התייחסו ל-

- i. ה type של השדה
- ii. הכרות הגולש עם הערכים
- iii. שימושיות לגולש המחפש תשובות
- iv. אפשרויות אחרות למנוע כפילויות
- b. בטבלת posts יש שדה tags. השדה מכיל רשימה של אינדקסי tags (לדוגמה [4,7], [82])

- 1. איזה סוג יחס יש בין tag ו post (אחד לאחד, אחד לרבים, רבים לרבים)
- 2. כיצד היה צריך ליצג את ה tags של post בעזרת טבלה
- 3. שלפו את כל ה post שיש להם את ה tag מספר 7. שימו לב לא לשלוף את אלו עם 17 או 70.
- 4. מימוש הסעיף הקודם היה עצוב. אנא ציינו יתרון של שימוש בשדה.
- c. בטבלה posts יש שדה בשם CommentsCount. בנוסף ניתן לחשב את השדה בעזרת טבלת comments
 - i. מה היתרונות של הכפילות בהוספת השדה?
 - ii. מה החסרונות?
 - iii. כתבו שאלתה המוצאת אי הסכמות בין השדה לטבלת comments. השתמשו בשאלתה משאלה 1.
 - iv. כיצד ניתן לדאוג שלא יהיו אי הסכמות?

3. עיצוב (35 נקודות)

עצבו בסיס נתונים עבור ספריית ספרים.
המערכת צריכה לתמוך בשימושים הבאים.

1. הצגת שם, סופר, תיאור ותאריך הוצאה לכל ספר.
 2. הצגת מספר הכותרים (עותקים) של כל ספר בספרייה.
 3. הצגת כל הכותרים הנמצאים אצל קורא מסויים.
 4. הצגת כל הקוראים שלא השאילו מעולם ספר.
-
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. האם מצב בו קורא מעולם לא השאיל ספר הגיוני? האם ניתן ורצוי למנוע אותו מראש? תשובות שונות אפשריות, כל עוד הן מנומקות.
 5. ממשו את השאילתה "הצגת כל הקוראים שלא השאילו מעולם ספר." על בסיס הנתונים שלכם.