**סכמת הכוכב**

**הערה**: בחרנו לייצר ID לכל אחד מהישויות במערכת, זאת הן במטרה לאפשר לנו לפעול ללא תלות ב-yelp, והן בכדי לגרום לשאילתות ה-join לפעול ביעילות על משתנים מסוג int (ולא Varchar, המתקבלים מ-yelp). על מנת לאפשר עדכונים של הערכים מנתונים עתידיים מ-yelp, ובמקביל לחסוך להמנע בטבלאות Cross-Reference, נשמור בטבלאות את המפתח המקורי של yelp.

**טבלת ממד המשתמשים: dim\_users**

|  |  |  |  |
| --- | --- | --- | --- |
| שדה | טיפוס נתונים | מקור נתונים | הערות |
| user\_id | INT | AUTO\_INCREMENT | מפתח ראשי  NOT NULL |
| yelp\_user\_id | Varchar(20) | yelp\_academic\_dataset\_user.json | קישור למפתח של yelp |
| friends\_count | INT | yelp\_academic\_dataset\_user.json | תמונת מצב קיים, יש לעדכן |
| review\_count | INT | yelp\_academic\_dataset\_user.json | תמונת מצב קיים, יש לעדכן |
| fans | INT | yelp\_academic\_dataset\_user.json | תמונת מצב קיים, יש לעדכן |
| is\_elite | Binary: 0/1 | yelp\_academic\_dataset\_user.json | תמונת מצב קיים, יש לעדכן |
| yelping\_since | Date צריך לשאול את דימה אם משתמשים בתאריך או במפתח שלו בטבלת Date. | yelp\_academic\_dataset\_user.json | קבוע |

**טבלת ממד העסקים: dim\_business**

|  |  |  |  |
| --- | --- | --- | --- |
| שדה | טיפוס נתונים | מקור נתונים | הערות |
| business\_id | INT | AUTO\_INCREMENT | מפתח ראשי  NOT NULL |
| Yelp\_business\_id | Varchar(20) | yelp\_academic\_dataset\_business.json | קבוע |
| business\_name | Varchar(20) | yelp\_academic\_dataset\_business.json | קבוע |
| is\_open | Binary: 0/1 | yelp\_academic\_dataset\_business.json | תמונת מצב קיים, יש לעדכן |
| neighborhood | Varchar(20) | yelp\_academic\_dataset\_business.json | קבוע |
| latitude | Float | yelp\_academic\_dataset\_business.json | קבוע |
| longitude | Float | yelp\_academic\_dataset\_business.json | קבוע |
| stars | Float | yelp\_academic\_dataset\_business.json | תמונת מצב קיים, יש לעדכן |
| review\_count | INT | yelp\_academic\_dataset\_business.json | תמונת מצב קיים, יש לעדכן |
| Category | Varchar(40) | yelp\_academic\_dataset\_business.json | קטגוריה אחת מתוך עשר |
| checkin\_sum | INT | yelp\_academic\_dataset\_checkin.json | תמונת מצב קיים, יש לעדכן |

**טבלת ממד התאריכים: dim\_dates**

|  |  |  |  |
| --- | --- | --- | --- |
| שדה | טיפוס נתונים | מקור נתונים | הערות |
| date\_Id | INT | AUTO\_INCREMENT | מפתח ראשי  NOT NULL |
| date | DATE | פונקציות, משתנים ולולאות | קבוע |
| day | INT | פונקציות, משתנים ולולאות | קבוע |
| month | INT | פונקציות, משתנים ולולאות | קבוע |
| year | INT | פונקציות, משתנים ולולאות | קבוע |
| day\_name | Varchar(45) | פונקציות, משתנים ולולאות | קבוע |
| Holiday | Varchar(45) | הכנסת נתונים חיצוניים | קבוע |

**טבלת העובדות- ביקורות: facts\_review**

|  |  |  |  |
| --- | --- | --- | --- |
| שדה | טיפוס נתונים | מקור נתונים | הערות |
| review\_id | INT | AUTO\_INCREMENT | מפתח ראשי  NOT NULL |
| yelp\_review\_id | Varchar(20) | yelp\_academic\_dataset\_review.json | קישור למפתח של yelp |
| user\_id | INT | חיבור בין yelp\_academic\_dataset\_review.json לבין טבלת ממד המשתמשים | מפתח זר לממד המשתמשים |
| business\_id | INT | חיבור בין yelp\_academic\_dataset\_review.json לבין טבלת ממד העסקים | מפתח זר לממד העסקים |
| date\_id | INT | חיבור בין yelp\_academic\_dataset\_review.json לבין טבלת ממד התאריכים | מפתח זר לממד התאריכים |
| stars | Float | yelp\_academic\_dataset\_review.json | תמונת מצב קיים, יש לעדכן |
| votes\_aggregate | INT | yelp\_academic\_dataset\_review.json | תמונת מצב קיים, יש לעדכן. סכימת כלל ההצבעות |