**Homework 2 – Business Intelligence**

**חלק 2 – הגדרת Data Warehouse**

1. מתוך הסכמות הקיימות ובהתבסס על טבלת הנתונים שבחרנו במטלה הקודמת סיכמת כוכבים תהיה המתאימה ביותר, מכיוון שהיא מספקת דרך פשוטה ואינטואיטיבית לארגן ולנתח את הנתונים שלנו. על ידי ארגון הנתונים שלנו בצורה זו, נוכל לבצע בקלות שאילתות כדי לנתח את הקשרים בין הממדים והמדדים. לדוגמה, נוכל לנתח בקלות כיצד התמחור משפיע על דירוגי משתמשים עבור תעשייה או קטגוריה ספציפית, או כיצד מספר המשתמשים משתנה לפי קנה המידה העסקי וסוג המשתמש.
2. להלן פירוט של האופן שבו ניתן לארגן את טבלת הנתונים לפי סכמת כוכבים:
3. Fact Table:

* product\_id (Primary Key)
* no\_of\_users
* pricing
* rating

1. Dimension Tables:

* Category Dimension Table:
* category\_id (Primary Key)
* category
* Industry Dimension Table:
* industry\_id (Primary Key)
* industry
* Business Scale Dimension Table:
* business\_scale\_id (Primary Key)
* business\_scale
* User Type Dimension Table:
* user\_type\_id (Primary Key)
* user\_type
* Deployment OS Dimension Table:
* deployment\_os\_id (Primary Key)
* deployment\_os
* Mobile Apps Dimension Table:
* mobile\_apps\_id (Primary Key)
* mobile\_apps

עיצוב סכמה זה מאפשר לטבלת העובדות לאחסן את הנתונים הניתנים למדידה כגון מספר המשתמשים, התמחור והדירוג. טבלאות הממדים לוכדות את התכונות התיאוריות המשויכות לכל נקודת נתונים. העמודה product\_id פועלת כמפתח הראשי המקשר את טבלת העובדות עם טבלאות הממדים המתאימות.

1. מקרה שימוש אחד שבו סכימת כוכבים תהיה יעילה עבור אוסף הנתונים שלנו היא באמצעות סכימת כוכבים, טבלת העובדות תכיל נתונים כגון הכנסות, מספר רישיונות שנמכרו ודירוגי שביעות רצון לקוחות. טבלאות הממדים יכולות לכלול מידע על המוצר, התעשייה, קנה המידה העסקי, סוג המשתמש, הפריסה, מערכת ההפעלה והתמחור. טבלת ממדי המוצר יכולה לכלול מאפיינים כגון שם המוצר, התיאור והקטגוריה. טבלת ממדי התעשייה יכולה לכלול תכונות כגון שם ותיאור של התעשייה. לטבלת הממדים בקנה מידה עסקי יכולות להיות תכונות כגון מספר העובדים בעסק. טבלת הממדים של סוג המשתמש יכולה לכלול תכונות כגון תפקידי משתמש ופונקציות עבודה. טבלת ממדי הפריסה יכולה לכלול תכונות כגון מקומית, ענן או היברידי. טבלת הממדים של מערכת ההפעלה יכולה לכלול תכונות כגון Windows, macOS או Linux. טבלת מאפייני התמחור יכולה לכלול מאפיינים כגון טווח מחירים ותוכניות מנוי.

עם מבנה זה, חברת התוכנה יכולה בקלות לאחזר ולנתח נתונים כדי לענות על שאלות כגון:

* מהם המוצרים הנמכרים ביותר לפי ענף וסוג משתמש?
* כיצד משתנה דירוג שביעות רצון הלקוח לפי פריסה ומערכת הפעלה?
* מה ההכנסה שנוצרת מכל קטגוריית מוצר וקנה מידה עסקי?
* מהו המספר הממוצע של רישיונות שנמכרו ללקוח לפי ענף ותוכנית תמחור?

סכמת הכוכבים יעילה במקרה זה מכיוון שהיא מאפשרת שליפה מהירה של מידע, שכן טבלת העובדות היא הטבלה המרכזית ומותאמת לשאילתות. השימוש בטבלאות ממדים גם מקטין את הצורך בצירופים, מה שעלול להאט את השאילתות והניתוח. נתונים אלה יכולים לעזור לחברת התוכנה לקבל החלטות מושכלות לגבי פיתוח מוצר, אסטרטגיות תמחור ומיקוד לקוחות, מה שמוביל לתוצאות עסקיות משופרות.