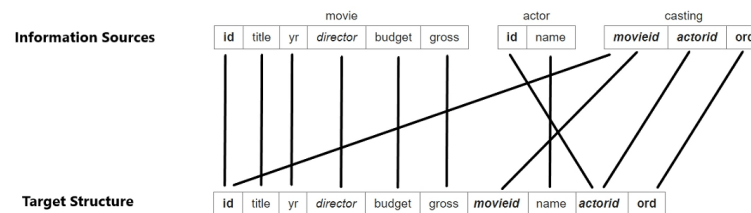


## Homework 3+4 – Business Intelligence

### חלק 1 – STTM

השתמשו ב-DW ותהליך ETL אשר הגדרתם בתרגיל מספר 2 ובצעו את הפעולות הבאות:

1. בדקו שההוראות לתהליך ETL כתובות עפ"י פורמט STTM בקובץ אקסל.
2. צרו סכימה ויזואלית אשר מציגה את ה-STTM של הפרויקט שלכם.  
לצורך המחשה, תשובתכם צריכה להיות בפורמט:



3. הוסיפו להגשה את קובץ ה-STTM שיצרתם + את הסכימה הנ"ל.

### חלק 2 – data mining techniques

עבור ה-DW שיצרתם (לאחר ה-ETL), ענו בקצרה (2-3 משפטים) על השאלות הבאות:

1. מהו תהליך ה-KDD אשר אתם מבצעים? יש לפרט האם בחרתם בשיטת Predictive או Descriptive.
2. מהם הטכניקות אשר תממשו על אוסף הנתונים?  
a. יש להתייחס לסוגי הנתונים.  
b. יש להראות שני תרחישים (use-case) עם דוגמה מספרית (ניתן להמציא ערכים אקראיים לצורך ההמחשה בלבד).
3. הגדירו מדד דמיון עבור ה-DW שבניתם. ניתן להשתמש במדדי דמיון אשר ראינו בכיתה או לחילופין להחליט על מדד דמיון שרלוונטי עבור העבודה שלכם. יש להסביר בקצרה מדוע מדד זה אכן מהווה דמיון בין הנתונים.
4. עבור כל אחת מהשאלות העסקיות, יש לנסח 2 השערות: H0, H1 ולהציע דרך כיצד להחליט איזו השערה לקבל ומדוע.

### חלק 3 – שאילות

השתמשו בשאילות SQL וב Window Functions ובנו 6 שאילות שונות אשר מטרתם לחלץ מידע מה-DW אשר עונה על השאלות העסקיות.

תזכורת: השאילות הנ"ל יהיו התוצר אשר נרצה להציג בבלי ויזואלי לאחר בניית ה-DW. חשבו על שאילות ספציפיות ומדויקות אשר עונות על השאלות העסקיות (ולא רק שליפת נתונים).

### חלק 4 – ניהול גרסאות

השתמשו בקובץ הוראות של תרגול 8 (GitHub) והוסיפו להגשתכם קישור לפרויקט שלכם, כאשר חבר צוות אחד הוא ה-admin של repository וכל שאר חברי הצוות הם collaborators.