דוגמה לדוח אנליטי

* **תקציר מנהלים**
* **הלקוח -**חברת ההפקות Siren Pictures Inc. אשר יצרה את הסדרה המצליחה משחקי הדיונון, גייסה אותנו לעזור עם המלצות על הפקת סדרה או סרט חדש.

Siren Pictures is a production company founded in 2011 by KIM Ji-yeon, who had previously produced "A Million" (2009) and "Second Half" (2010). It found commercial and critical acclaim with its first production, "The Fortress" (2017), a film adaptation of the novel ‘Namhan Mountain Fortress’ written by KIM Hoon, who is none other than KIM Ji-yeon’s father.

* **מסדי נתונים -**עומדים לרשותנו 4 מסדי נתונים מסוג CSV.
	+ האחד Netflix TV Shows and Movies.csv
	כולל מידע כללי על הסרטים והסדרות, דירוג IMDB וכמות המדרגים. [לינק](https://www.kaggle.com/datasets/thedevastator/netflix-imdb-scores)
	+ השני titles.csv
	כולל מידע על הז'אנר וכן דירוגים מאתר TMDB. [לינק](https://www.kaggle.com/datasets/victorsoeiro/netflix-tv-shows-and-movies)
	+ השלישי Credits
	כולל את שמות השחקנים/ות. [לינק](https://www.kaggle.com/datasets/victorsoeiro/netflix-tv-shows-and-movies)
	+ הרביעי Netflix titles
	כולל מידע על שמות הסרטים והסדרות ומאפיינים כמו במאי, סיווג צפייה (מתאים לילדים וכו'), משך וכו'. [לינק](https://www.kaggle.com/datasets/shivamb/netflix-shows)
* **טיוב נתונים -**כיצד ניתן לחבר בין 4 מסדי הנתונים? האם יש שדה משותף?
אם משתמש במפתח מסוג STRING יש לקחת בחשבון CASE SENSITIVITY (למשל סרט שנקרא במסד אחד TAXI DRIVER ובמסד השני Taxi Driver. )
* שדה GENRES כולל רשימה של ז'אנרים, צריך "לפרק" לשדות נפרדים כדי לנתח השפעה של ז'אנר על הציון.
* עבור חלק מהרשומות חסר חלק מהמידע וצריך להחליט מה לעשות עם הרשומות החסרות, לדוגמה:





 **שדות חדשים - לפרט**

* **תרשים של המודל (חיבור בין טבלאות)**

****

 **חיבורים:** A + B + C מתחברים על בסיס שדה משותף ID

D מתחבר ל-A על בסיס שדה TITLE (יש לבצע UPPER CASE כדי שהחיבור יעבוד תקין) ובנוסף עם
 שדה TYPE (סרט או סדרת טלוויזיה) . יש לשים לב ששדה TYPE בבסיס נתונים D מכיל TV SHOW לעומת
 אותו שדה בבסיס נתונים A שבו הערך הוא SHOW (בלי TV)

* **שאלת מחקר**

בהתבסס על הדירוג של IMDB וכמות המדרגים

* באיזה ז'אנר/ שילוב של ז'אנרים כדאי להתמקד?
* האם כדאי להפיק סרט או סדרה?
* עם אילו עם במאים ושחקנים כדאי לעבוד?
* **ניתוח הממצאים (לפחות 10 ממצאים)**דוגמה לגרף שמציג את 10 שילובי הז'אנרים שקיבלו את הציונים הגבוהים ביותר בדירוג IMDB.
ניתן לראות שסרט / סדרה המשלבים מד"ב עם אנימציה, פעולה ופנטזיה, ואשר מתאימים לכל המשפחה, מקבלים ציון ממוצע של 9.3 המתבסס על 297 אלף נתוני הצבעות.


דוגמה לגרף שמציג את הבמאים/ות אשר ביימו לפחות סרט אחד וגם סדרה אחת עם הציונים שלהם,
מגבוה לנמוך:


* **מסקנות**
* **המלצה ללקוח**
* **ביבליוגרפיה**
* **כלים שנעשה בהם שימוש - אקסל + טאבלו**