פרויקט מסכם Power BI

מבוא :

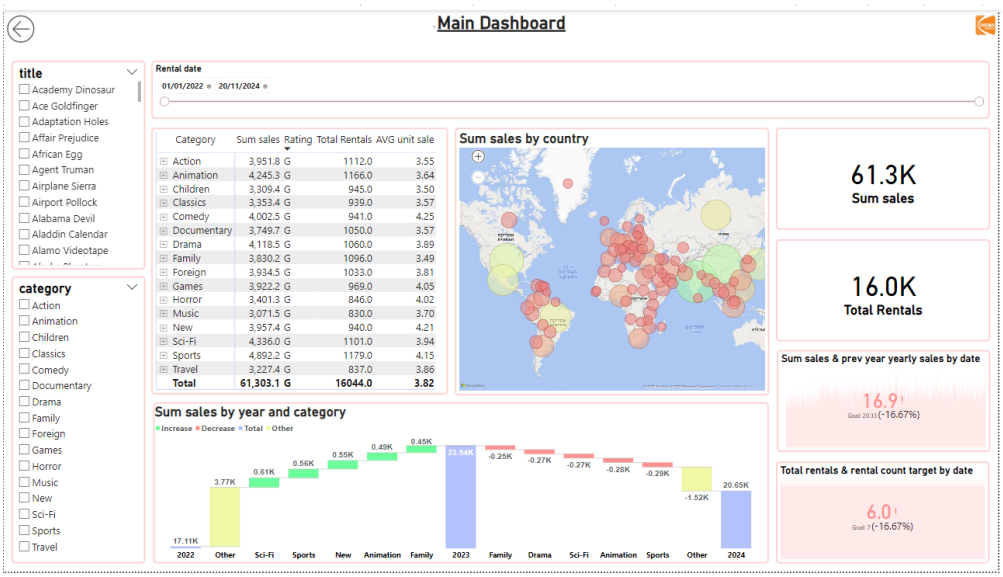
במהלך הפרויקט הנוכחי, ביצענו ניתוח של נתוני השכרות, מכירות ולקוחות בתחום השכרת הסרטים.

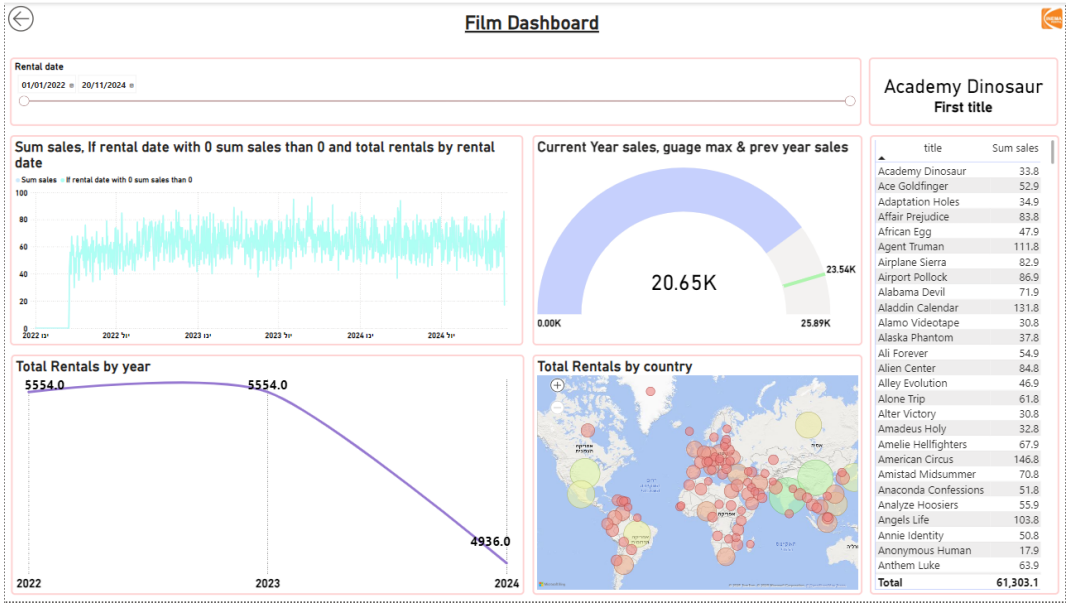
מטרת הפרויקט הייתה לבנות מערך נתונים איכותי , אשר יאפשר הסקת מסקנות והפקת תובנות עסקיות משמעותיות.

באמצעות שימוש ב Power BI יצרנו מודל נתונים מבוסס סכמת כוכב, אשר מחבר בין טבלאות שונות בצורה לוגית ומסודרת.

הפרויקט כלל עיבוד נתונים, יצירת מדדים, עמודות ופיתוח דשבורדים ויזואליים שמציגים תובנות בצורה נגישה וברורה.

המסמך שלפניכם מספק תיעוד מפורט של שלבי העבודה, החל מעיבוד הנתונים ועד להצגת תובנות, תוך מתן דגש על הפתרונות והכלים בהם נעשה שימוש.





ייבוא הנתונים

ייבוא הנתונים החל מקובץ שסופק כחלק מהמטלה.
הנתונים נטענו לתוך Power BI תוך הקפדה על דרישות המטלה.
תהליך זה הבטיח שכל המידע הדרוש לניתוח זמין ומוכן לעיבוד נוסף במסגרת העבודה על המודל .

מארג (Merge) בין טבלאות

במהלך הפרויקט, נעשה שימוש בפונקציית ה Merge כדי לחבר בין טבלאות שונות.
ה Merge משמש לאיחוד נתונים משתי טבלאות או יותר, על סמך עמודה משותפת המשמשת כמפתח. פעולה זו מאפשרת להעשיר טבלה אחת במידע רלוונטי מטבלה אחרת, תוך שמירה על שלמות הנתונים.
בשלב זה נבחרו סוגי החיבור המתאימים, התהליך כלל:

1. זיהוי העמודות המשותפות המשמשות כמפתחות לחיבור.
2. ביצוע ה Merge
3. הרחבת העמודות הנדרשות מתוך הטבלה השנייה, כדי להוסיף נתונים רלוונטיים לטבלת הבסיס.

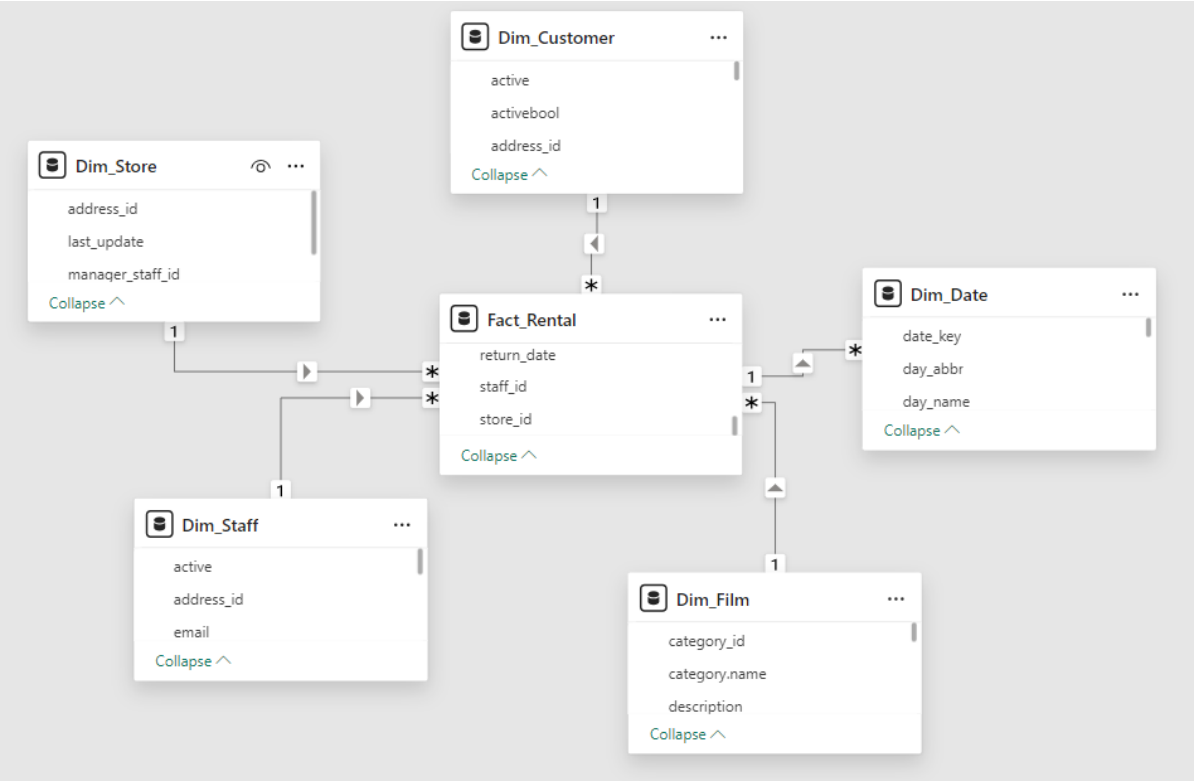
באמצעות פעולה זו נוצר מארג נתונים שלם שמאפשר ניתוח יעיל ומדויק.

הגדרת קשרים בין טבלאות (Relational Model)

כחלק מבניית מודל הנתונים ב Power BI הוגדרו קשרים (Relationships) בין הטבלאות השונות.
קשרים אלו הם הבסיס ליצירת מודל נתונים ומאפשרים ניתוח רוחבי של הנתונים.
השלבים העיקריים בתהליך כללו:

1. זיהוי מפתחות ראשיים לכל טבלה , זוהו העמודות שיכולות לשמש כמפתחות ליצירת קשרים.
   לדוגמה, עמודת מזהה לקוח בטבלת הלקוחות חוברה לעמודה המקבילה בטבלת השכרות.
2. הגדרת סוגי הקשרים :
   קשר של אחד לרבים (One-to-Many) או קשר של אחד לאחד (One-to-One)
3. ניהול כיווניות הקשר חד כיווניים (Single) או דו כיווניים (Both)

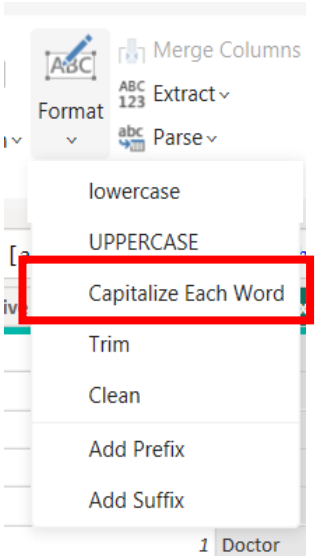
באמצעות הגדרת הקשרים נבנה מודל נתונים מבוסס סכמת כוכב :

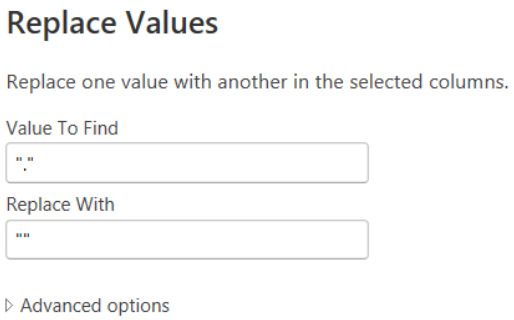


Dim Customer

עמודת הPrefix

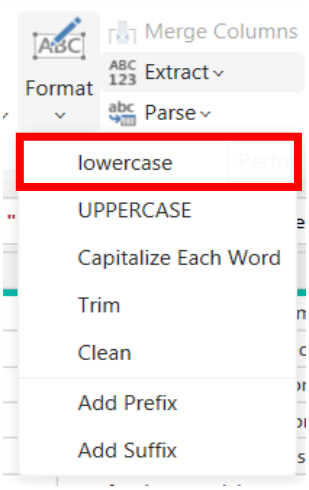
על מנת להפוך את האות הראשונה בכל מילה לגדולה ואת יתר האותיות לקטנות, נעשה שימוש באפשרות
Transform > Format > Capitalize Each Word
לצורך הסרת הנקודה שהופיעה בסוף הערכים, נעשה שימוש בפקודת
Transform > Replace Values
הנקודה הוחלפה בערך ריק, (""), וכך הוסרה לחלוטין.





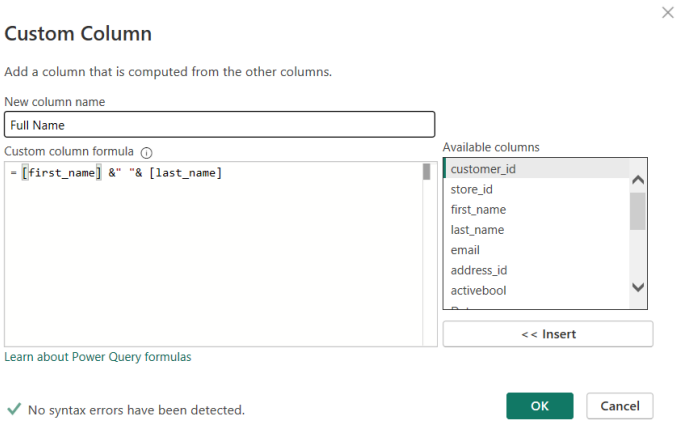
עמודות First_Name & Last_Name

עבור שינוי פורמט האותיות עבור כל עמודה נעשה שימוש באפשרות
Transform > Format > Lowercas
שהמירה את כל האותיות לערכי אותיות קטנות.



עמודת Full_Name

יצירת עמודה חדשה שתכיל את השם המלא של המשתמש, המורכב מהשם הפרטי ומהשם המשפחה.
על מנת להוסיף עמודה חדשה, נבחרה האפשרות Add Column בתפריט העריכה.
לאחר מכן נרשמה הנוסחה שחיברה את עמודת השם הפרטי [FirstName] לעמודת שם המשפחה [LastName] עם רווח ביניהן.

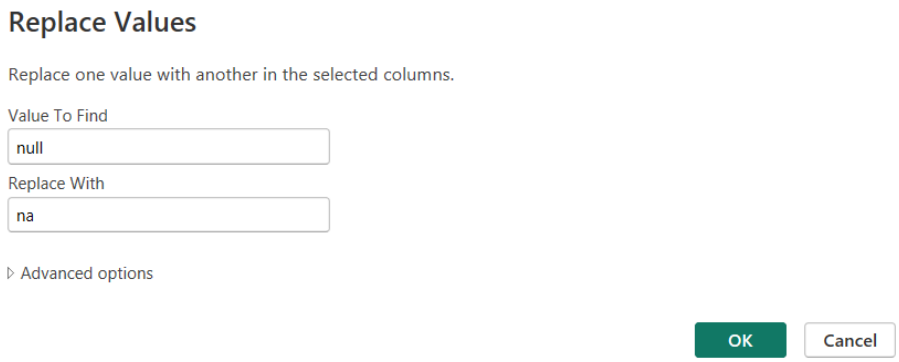


החלפת ערכים ריקים (Null) ב "na"

קודם כל היה צורך בזיהוי העמודות הרלוונטיות (מסוג Text) , לאחר מכן נבחרו כל העמודות מסוג טקסט שבהן עשויים להימצא ערכים ריקים.

לאחר מכן כדי לבצע את הפעולה נעשה שימוש בפקודת

Transform > Replace Values

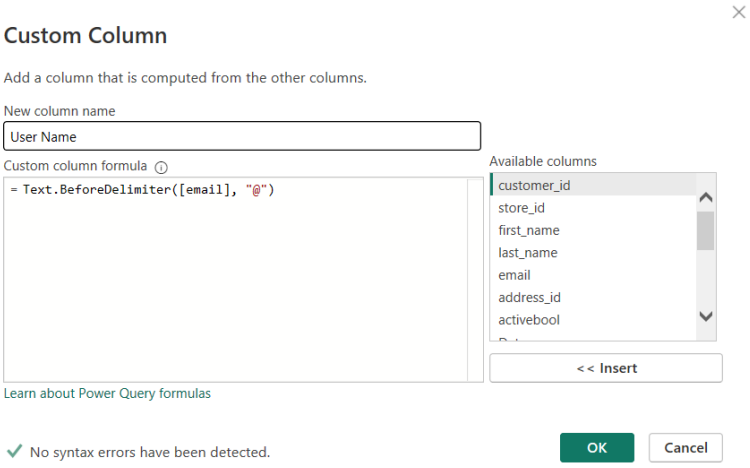


עמודת User_Name

על מנת ליצור עמודה חדשה שתכיל את הטקסט שמופיע לפני התו "@" מתוך הערך בעמודת המייל נבחרה האפשרות Custom Column בתפריט העריכה.

לאחר מכן נרשמה הנוסחה שחילצה את מה שנדרש מהכתובת המלאה.

הפונקציה **Text.BeforeDelimiter** מחזירה את הטקסט שנמצא לפני התו המוגדר (במקרה זה "@").



עמודת Level_Of_Income

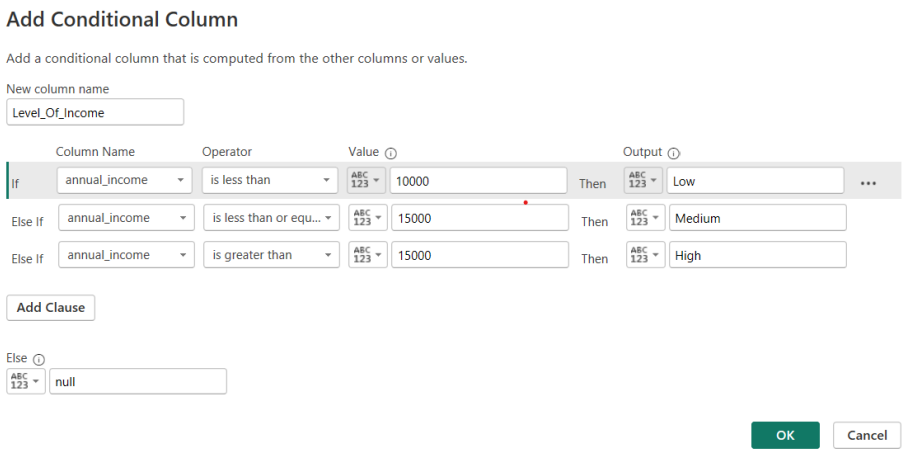
על מנת ליצור עמודה חדשה שתכיל High \ Medium / Low על פי רמת ההכנסה, נבחרה האפשרות Conditional Column בתפריט העריכה.

תנאי 1 אם ההכנסה פחות מ-10,000, אז התוצאה תהיה Low.

תנאי 2 אם ההכנסה פחות מ-15,000 (ולא פחות מ-10,000), אז התוצאה תהיה Medium.

תנאי 3 אם ההכנסה מעל 15,000, התוצאה תהיה High.

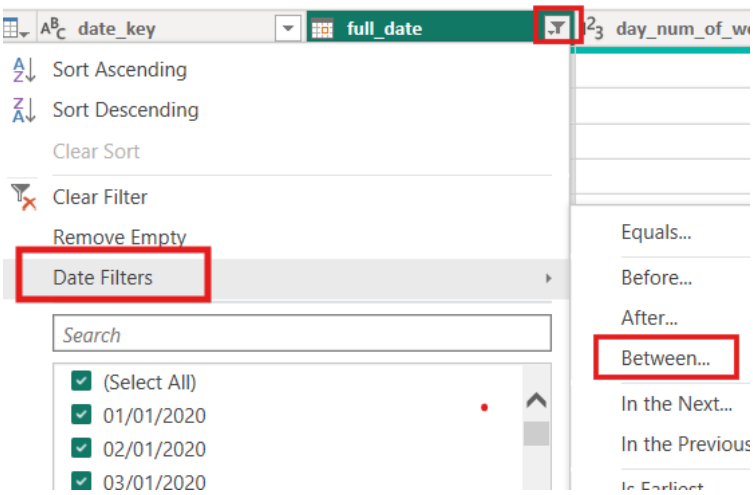
Else התוצאה תהיה Null.

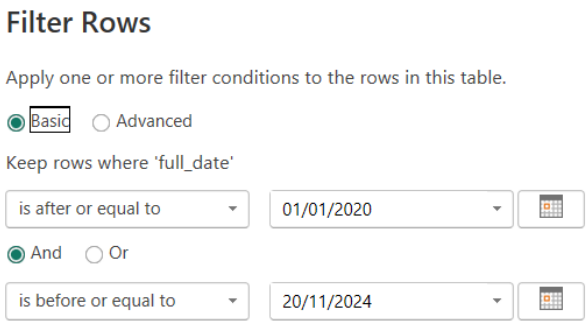


Dim Date

הבאת תאריכים בין תאריך ההשכרה המוקדמת ביותר למאוחרת ביותר

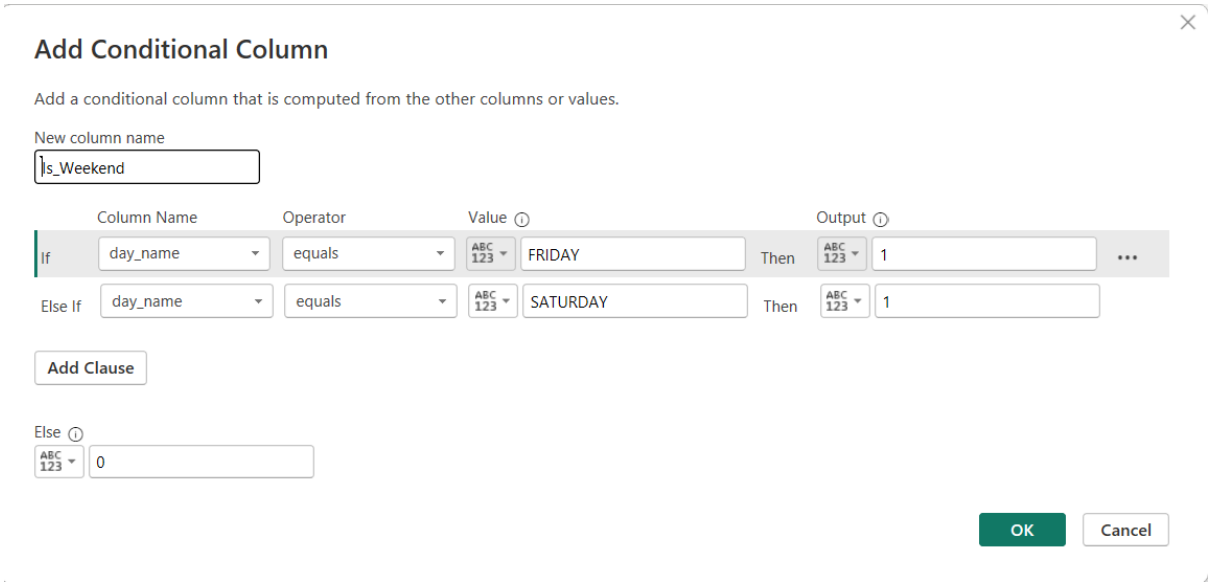
כשלב ראשון יש לזהות את התאריכים של ההשכרה הראשונה והאחרונה בעמודת Rental_Date בטבלת Rental.
לאחר מכן נעשה פילטור של עמודת תאריכים (Full Date) בין התאריך הראשון לתאריך האחרון באמצעות Between.
* באמצעות פעולה זאת נוצרה לנו אפשרות לקשר את Dim Date לטבלת ה Fact.





עמודת Is_Weekend

על מנת ליצור עמודה חדשה שתספק אינדיקציה אם תאריך מסוים הוא חלק מימי סוף השבוע (שישי או שבת), נבחרה האפשרות Conditional Column בתפריט העריכה.
תנאי 1 אם שם היום שווה ל-FRIDAY, אז התוצאה תהיה 1.
תנאי 2 אם שם היום שווה ל-SATURDAY, אז התוצאה תהיה 1.
Else התוצאה תהיה 0.
*כך העמודה החדשה תקבל את הערך 1 אם יום התאריך הוא שישי או שבת וערך 0 אם מדובר ביום אחר בשבוע.

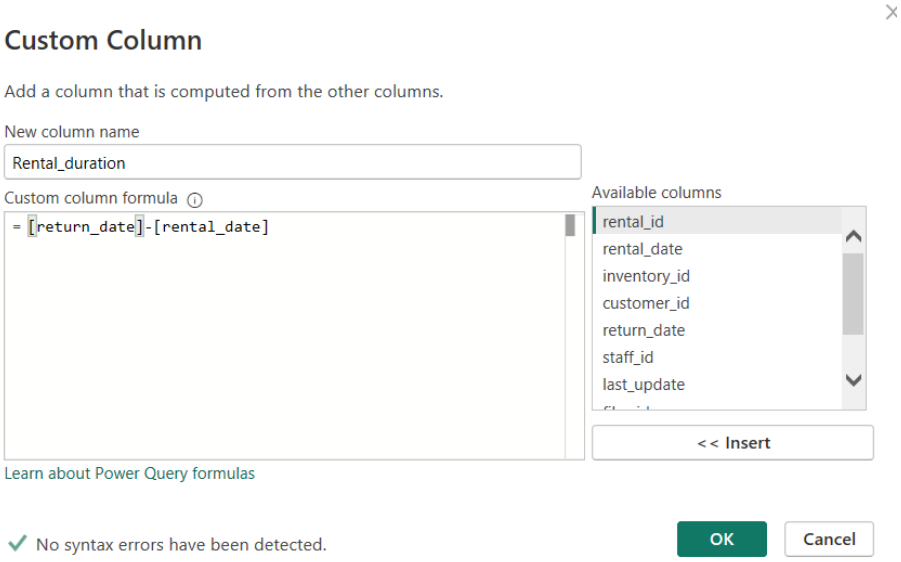


Fact_Table

טבלת ה- Fact שלנו היא טבלת Rental.

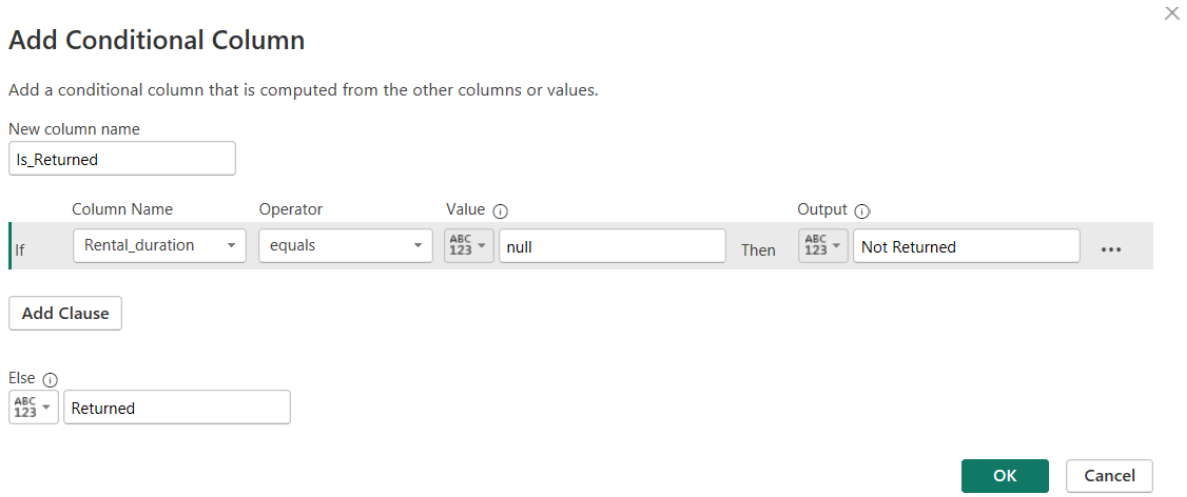
עמודת Rental_duration

עמודה זו מציגה את משך הזמן של כל השכרה בטבלת Rental.
על מנת ליצור עמודה זו נבחרה האפשרות Custom Column בתפריט העריכה.
העמודה מחשבת את ההפרש בין תאריך החזרה (Return Date) לתאריך ההשכרה (Rental Date) ההפרש מוצג במספר ימים.



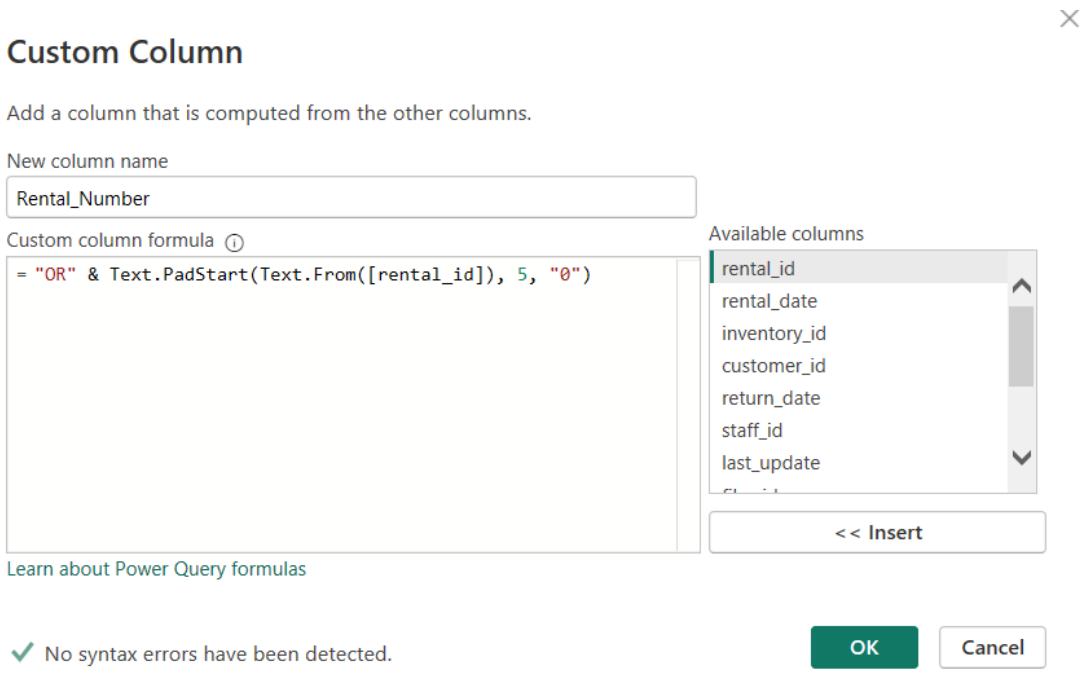
עמודת Is_Returned

עמודה זו מציגה אם סרט הוחזר או לא.
העמודה תבדוק אם קיים ערך בעמודת Rental_duration (משך זמן השכרה) .
במידה וכן המשמעות היא שהסרט הוחזר אחרת, הסרט לא הוחזר.
על מנת ליצור עמודה זו נבחרה האפשרות Conditional Column בתפריט העריכה.
תנאי 1 אם Rental_duration שווה ל- Null , התוצאה תהיה Not Returned.
Else (אם קיים ערך בעמודה), התוצאה תהיה Returned.



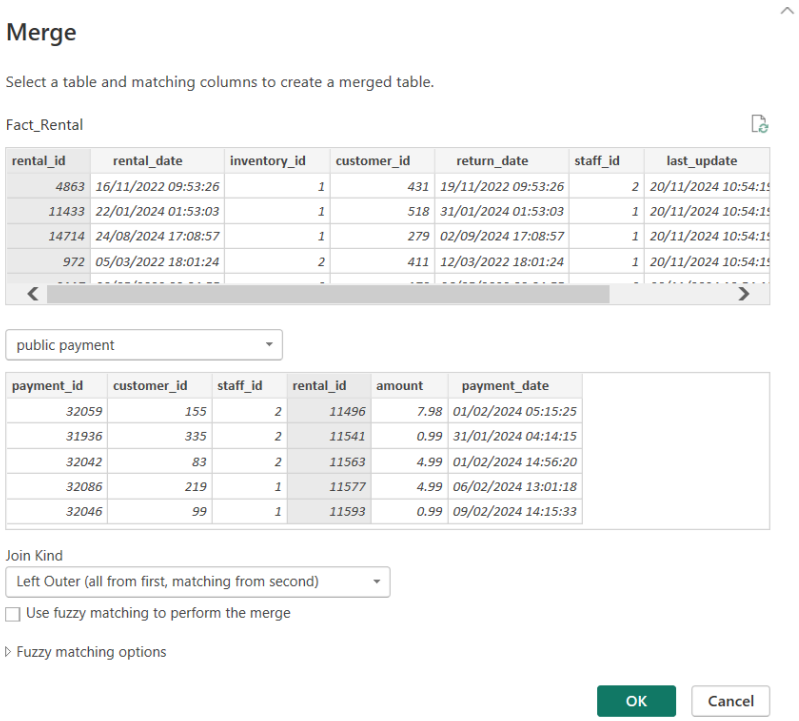
עמודת Rental_Number

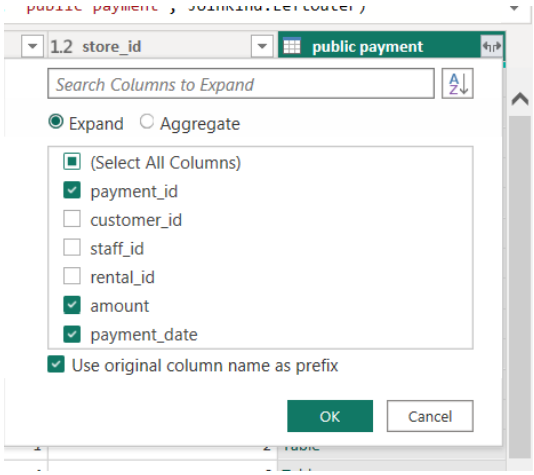
על מנת ליצור עמודת מזהה השכרה חדשה בפורמט " OR" ולאחריו 5 ספרות נבחרה האפשרות Custom Column בתפריט העריכה .
לאחר מכן נרשמה הנוסחה שתיצור את השרשור המבוקש.
הפונקציה Text.PadStart מטרתה להוסיף אפסים (במקרה שלנו) מובילים עד שהאורך של המחרוזת יהיה 5 תווים.
לדוגמה, אם הערך בעמודת rental_id הוא 23, הפונקציה תוסיף אפסים ותחזיר 00023.



טבלת Payment

על מנת להביא את העמודות מטבלת Payment לטבלת הFact_Rental נעשה Merge בעזרת מפתח המקשר ביניהן (Rental_Id).
ולאחר מכן נבחר עמודות חשובות.





Dax

מדד Sum Salse

מדד זה משמש לחישוב סך כל המכירות מתוך טבלת Fact_Rental בהתבסס על עמודת Amount,
בנוסף הוא מאפשר לנו לקבל תמונה ברורה על סך ההכנסות מתוך טבלת ההשכרות ומסייע לנו לבצע ניתוחים עסקיים בצורה חכמה ודינמית .

```
Sum sales = SUM(Fact_Rental[Amount])
```

פירוט הפעולה של המדד:
פונקציית SUM מבצעת חיבור של כל הערכים בעמודה Amount.

מדד Total Rentals

מדד זה משמש לחישוב מספר השכרות הכולל מתוך טבלת Fact_Rental בהתבסס על עמודת Rental_Id.
בנוסף הוא כלי מרכזי בניתוח נתוני השכרות, המאפשר להבין כמה השכרות בוצעו בכלל או לפי פילטרים ספציפיים. הוא מסייע לראות מגמות, כמו גם לבחון את ביצועי החברה על פני תקופות שונות.

```
Total Rentals = Count(Fact_Rental[rental_id])
```

פירוט הפעולה של המדד:
הפונקציה **COUNT()** סופרת את מספר הערכים הלא ריקים בעמודה **rental_id**.
מכיוון שכל השכרה מקבלת מזהה ייחודי, הפונקציה בעצם מחזירה את מספר השורות בטבלה, כלומר מספר העסקאות שבוצעו.

מדד Total Returned Rentals

מדד זה משמש לספירת מספר השכרות שהוחזרו מתוך טבלת Fact_Rental בהתבסס על עמודת Is Returned.
מדד זה חשוב להבנת התנהגות הלקוחות וניהול המלאי.

```
Total Returned Rentals = COUNTROWS(FILTER('Fact_Rental',Fact_Rental[Is Returned] = "Returned"))
```

פירוט הפעולה של המדד:
הפונקציה FILTER() מסננת את הנתונים בטבלה Fact_Rental, היא מחזירה רק את השורות שבהן הערך בעמודה Is Returned הוא "Returned" , כלומר השכרות שהוחזרו.
לאחר הסינון, הפונקציה COUNTROWS() סופרת את מספר השורות שנותרו, מאחר וכל שורה מייצגת השכרה אחת, מתקבלת כמות השכרות שהוחזרו.

מדד Total Not Returned Rentals

מדד זה משמש לספירת מספר השכרות שלא הוחזרו מתוך טבלת Fact_Rental בהתבסס על עמודת Is Returned.
מדד זה מספק אינדיקציה חשובה על מספר השכרות שלא הוחזרו ומאפשר לזהות בעיות ולשפר את תהליך ההשכרה והחזרת הסרטים.

```
Total Not Returned Rentals = COUNTROWS(FILTER('Fact_Rental','Fact_Rental'[Is Returned] = "Not Returned"))
```

פירוט הפעולה של המדד:
הפונקציה FILTER() מסננת את הנתונים בטבלה Fact_Rental, היא מחזירה רק את השורות שבהן הערך בעמודה Is Returned הוא "Not Returned" , כלומר השכרות שלא הוחזרו.
לאחר הסינון, הפונקציה COUNTROWS() סופרת את מספר השורות שנותרו, מאחר וכל שורה מייצגת השכרה אחת, מתקבלת כמות השכרות שלא הוחזרו.

Visualization

Main Dashboard

דשבורד ניתוח השכרות סרטים
הדשבורד מספק תובנות עסקיות מקיפות על ביצועי השכרות של סרטים על פני תקופות זמן שונות ובחלוקה לפי אזורים וקטגוריות.
דשבורד זה מאפשר זיהוי מגמות, ניתוח השפעת קטגוריות הסרטים על המכירות, והשוואה בין ביצועי השנים השונות.
השימוש בויזואליזציות אינטראקטיביות מאפשר קבלת החלטות מבוססות נתונים בצורה מהירה ויעילה.
הדשבורד כולל מספר ויזואליזציות אינטראקטיביות שמאפשרות ניתוח מעמיק של הנתונים:

מפה גיאוגרפית (Visual Map)
המפה מציגה את סך המכירות לפי מדינות וערים, ומאפשרת הבנה מהירה של פיזור המכירות ברחבי העולם.
הצבעים משמשים כמדד לביצועים:
ירוק –אזורים עם מכירות גבוהות.
צהוב –אזורים עם מכירות בינוניות.
אדום –אזורים עם מכירות נמוכות.
נבחרה כדי להציג בצורה ויזואלית וברורה את התפלגות המכירות לפי מיקום גיאוגרפי, כך שניתן יהיה לזהות אזורים חזקים וחלשים במכירות במהירות.
היא מאפשרת ניתוח עומק באמצעות זום וסינון לפי מדינות וערים.

מטריצה(Matrix Table)
מטריצה זו מספקת תובנות מפורטות על סך המכירות וכמות השכרות, תוך חלוקה לפי קטגוריות הסרטים והשמות הספציפיים של הסרטים.
נבחרה כדי להציג פירוט מדויק של המכירות לפי קטגוריות ושמות הסרטים, מה שמאפשר למשתמש לבצע ניתוח השוואתי ולזהות אילו קטגוריות וסרטים הם הרווחיים ביותר.

גרף וואטרפול(Waterfall Chart)
גרף זה מציג את סך המכירות בהתפלגות לפי שנים וקטגוריות סרטים, תוך הדגשת השינויים שחלו במכירות לאורך הזמן
כחול – סך כל המכירות (Total).
ירוק –גידול במכירות (Increase)
אדום –ירידה במכירות (Decrease)
צהוב –נתונים אחרים שאינם משתייכים באופן ישיר לגידול או ירידה (Other)
נבחר כדי להראות את השינויים במכירות לאורך זמן ולהמחיש באופן ברור אילו שנים וקטגוריות תרמו לעלייה או לירידה בהכנסות, דבר שמאפשר זיהוי מגמות והשפעות.

פילטרים(Filters – Date & Category)
פילטר קטגוריות ,**פילטר תאריכים**
נבחרו כדי לאפשר גמישות והתמקדות בפרקי זמן ספציפיים, קטגוריות מסוימות או שמות הסרטים, מה שמאפשר למשתמשים לחקור מגמות שונות בהתאם לצרכים שלהם.

כרטיסי מידע(Cards – Total Sales & Total Rentals)
הכרטיסים מספקים מבט מהיר על שני נתוני מפתח:
סך המכירות – **(Total Sales)** מציג את סך ההכנסות שנוצרו מהשכרות הסרטים.
סך השכרות – **(Total Rentals)** מציג את מספר ההשכרות הכולל שבוצע לאורך כל התקופה.
כרטיסים אלו נבחרו כדי לאפשר למשתמש לקבל מידע מרכזי בצורה ברורה ופשוטה ללא צורך בחישובים נוספים.

מדדי KPI (Key Performance Indicators - KPIs)
KPI סך המכירות– (SUM SALES)
מציג את מגמת המכירות לאורך זמן תוך השוואה לשנה הקודמת.
זהו מדד קריטי להערכת ביצועי החברה ולניתוח צמיחה או ירידה בהכנסות.
KPI סך ההשכרות– (Total Rentals)
מציג את מגמת השכרות הסרטים תוך השוואה לשנה הקודמת.
המדד מסייע להבין האם יש עלייה או ירידה בביקוש לסרטים ולזהות שינויים בהתנהגות הלקוחות לאורך השנים.

Film Dashboard

דשבורד ניתוח לפי סרט יחיד
דשבורד זה מספק תובנות מפורטות על ביצועי סרט ספציפי שנבחר מתוך הדשבורד הראשי.
הוא מאפשר בחינת נתוני השכרות והמכירות של הסרט בחלוקה לפי זמן ומיקום, תוך שימוש בויזואליזציות אינטראקטיביות.
הדשבורד נועד לאפשר ניתוח ממוקד של כל סרט בנפרד, תוך זיהוי מגמות, דפוסי השכרה והשוואה בין שנים ואזורי מכירה.
באמצעות Drillthrough המשתמשים יכולים לעבור בקלות מהדשבורד הראשי לדשבורד מפורט זה ולנתח את ביצועי הסרטים ברזולוציה גבוהה יותר.
הדשבורד כולל מספר ויזואליזציות אינטראקטיביות שמאפשרות ניתוח מעמיק ומדויק של נתוני הסרט:

כפתור Drill Through
כפתור זה מחבר בין הדשבורד הראשי לדשבורד המציג נתונים מפורטים על סרט ספציפי, בהתאם לבחירת המשתמש בפילטר שמות הסרטים.
נבחר כדי לאפשר ניתוח מעמיק ואינטראקטיבי, כך שהמשתמש יוכל להתמקד בסרט מסוים ולקבל עליו מידע מפורט בצורה נוחה וברורה.

כרטיס – (Card) שם הסרט הנבחר
מציג את שם הסרט שנבחר בפילטר של הדשבורד הראשי, ומוודא שהמשתמש תמיד ידע על איזה סרט הדשבורד מציג נתונים.
נבחר כדי לשמור על בהירות ודיוק בעת ניתוח הנתונים של סרט מסוים.

טבלה– (Table)
מציגה את סכום המכירות לכל סרט, מה שמאפשר למשתמש לזהות את הביצועים הפיננסיים של הסרטים השונים.
נבחרה כדי להציג נתונים מספריים ברורים, המאפשרים השוואת רווחיות בין הסרטים.

פילטר תאריך(Slicer - Rental Date)
מאפשר למשתמש לבחור טווח תאריכים ולהתמקד בנתונים של תקופה מסוימת.
נבחר כדי להעניק שליטה טובה יותר בניתוח מגמות ההשכרה והמכירות לאורך זמן.

מפה גיאוגרפית (Visual Map)
מאפשרת ניתוח עומק באמצעות זום וסינון לפי מדינות וערים, מציגה את כמות ההשכרות בכל מדינה באמצעות קידוד צבעים המשמשים כמדד לביצועים:
ירוק –אזורים עם מכירות גבוהות.
צהוב –אזורים עם מכירות בינוניות.
אדום –אזורים עם מכירות נמוכות.
נבחרה כדי להמחיש בצורה ויזואלית וברורה את הפיזור הגיאוגרפי של השכרות הסרטים, ולזהות אזורים עם ביצועים טובים או חלשים.

גרף מד (Gauge Chart)
מציג את סך המכירות של השנה הנוכחית (Current Year Sales) בהשוואה לערך היעד (Prev Year Sales) ולערך המקסימלי שהוגדר (Gauge Max) .
נבחר כדי להמחיש בצורה אינטואיטיבית את הביצועים של השנה הנוכחית ביחס לשנה הקודמת ולמקסימום שהוגדר, מה שעוזר להבין האם היעדים הושגו וכמה עוד נדרש לשיפור.

גרפי קו (Line Chart)
כמות ההשכרות לפי שנים
מציג את מגמת כמות ההשכרות לאורך השנים, ומאפשר לזהות האם חלה עלייה או ירידה בפופולריות הסרטים.
נבחר כדי לעזור בזיהוי מגמות צריכה ולספק תובנות עסקיות על שינויים בהתנהגות הלקוחות לאורך השנים.
סכום ההשכרות לפי תאריכי השכרה
מציג את סכום ההכנסות מהשכרות לאורך ציר הזמן, תוך התחשבות בכך שאם אין השכרות (Rental Date = 0), המכירות יהיו 0.
נבחר כדי להציג בצורה ברורה את הקשר בין תאריכי ההשכרה להכנסות, ולעזור בזיהוי דפוסים עונתיים או מגמות מכירה.