מטלת עבודה POWER BI

מגישים: רן חיים, עילי חביב, גבריאל מזרחי

פירוט תהליך עבודה

## שלב 1: טעינת נתונים, סידור ,עמודות מחושבות ופילטרים

בשלב הראשון של העבודה התחלנו בטעינת הנתונים מהDatabase של Adventure Works מטבלת Sakila.

התחלנו בחיבור כלל הטבלאות באפשרות פונקציית MERGE שנמצאת בסרגל כלים העליון, ונעזרנו בתרשים של סעיף 2. את כלל הטבלאות החדשות סידרנו בסכמת כוכב.

לאחר הטעינה התחלנו בשינוי כלל ההגדרות בטבלאות:

* שינינו הגדרות של כל עמודה לסוג הערך המתאים לה.
* שינוי שמי על מנת שלנו יהיה יותר נוח להשתמש ולהסתכל על הנתונים
* מחיקת עמודות אשר לא הביאו ערך לאנליזה
  + Month abbreviation
  + last\_update שהיה זהה לאורך כל הdatabase
* שינוי תאריכים לפורמטים שבוקשו בסעיף 3

הוספת עמודות חדשות התבצעו בשני דרכים:

* בעזרת הסרגל כלים העליון בשימוש קטגורית Transform עשינו את הדברים הפשוטים כגון:
  + להפוך את כלל העמודה לlowercase/CapitalCase
  + חיפוש והחלפה של NULL לNA
  + שימוש בExtract על מנת להוציא את כל הטקסט שלפני ה"@" בתוך המייל בשביל user\_name
* הדרך השנייה הייתה באמצעות DAX, אותה ביצענו בTable View. שימוש בDAX היה יותר לחיבור בין שני עמודות או לבצע פעולות אשר לא נמצאות בסרגל כלים של Transform.

על מנת להבין יותר את החומר שנלמד במהלך הקורס החלטנו שכדי לעשות את עמודת level\_of\_income אנחנו צריכים לחלק את עמודת annual\_income ל3 BINS שיתנו לנו את האלטרנטיבה לפונקציית ntile בSQL.

ולאחר שקיבלנו את הערכים יכולנו ליצור עמודה מחושבת אשר תבדוק את הערך ותחליט האם הלקוח הוא low/medium/high – income

בסעיף 4 התבקשנו לעשות פילטר על הזמן בין ההתחלה של rental\_date לסוף שלו. את הערכים בדקנו בעזרת measurements של Max Date וMin Date לאחר מכן ביצענו פילטר על העמודה של full\_date. ובעזרת עמודה מחושבת יצרנו את is\_weekend בה בדקנו האם **מספר** היום בשבוע גדול מ5 אם כן שים 1 אחרת לשים 0.

בסעיף 5 בחרנו להוסיף שני עמודות:

1. late\_return העמודה בודקת האם הלקוח החזיר באיחור או לא החזיר בכלל. מחזירה 1 אם הלקוח לא החזיר בכלל או באיחור אחרת תחזיר 0
2. rental\_complete\_payment – שמנו לב שהלקוחות משלמים את אותו המחיר גם כאשר הם לא החזירו את הסרט, כך שרצינו לבדוק כמה כסף לא שולם ולהציג את זה ללקוח בטבלה בסעיפים הבאים.

את עמודת rental number ביצענו בעזרת DAX והוספנו את הrental\_id למחרוזת "OR"

## שלב 2: מדדים וKPI

במהלך השלב הזה התחלנו לחשוב אילו מדדים יהיו רלוונטיים לנו על מנת ליצור דשבורד יעיל שייתן מידע על גבי החברה, ההכנסות שלה, כמות השכרות וKPIS.

כלל המדדים נמצאים תחת טבלת MSR והם מחולקים לתיקיות הבאות:

Dates – כל מה שקשור לסעיף 4 א', פה יש את שני המדדים Min Date ו-Max Date שאיתם בדקנו את תאריכי ההתחלה והסיום של rental\_date.

KPI – כלל המידע הדרוש לוויזואליזציה של KPI, כלל הKPI שלנו מעלים את הסכום של השנה/חודש **האחרונים** ל110% על מנת שנוכל להציב יעדים משמעותיים לחודש/שנה **הנוכחית**.

ארבעת הKPI שיצרנו הינם:

* Last Year Rental
* Last Month Rental
* Revenue Last Month
* Revenue Last Year

**הזוג הראשון** עושים השוואה אל מול ההשכרות של החודש שעבר והשנה שעברה.

**הזוג השני** עושים השוואה אל מול הרווח של החודש שעבר והשנה שעברה.

ישנו עוד מדד אחד Revenue MTD Last Month ששמנו בKPI, מדד זה שייך לוויזואליזציה של Gauge ב[Film Dashboard](#_Gauge) והוא מחשב את כלל ההכנסות מהחודש האחרון בלבד.

Rentals – בתיקייה זו ישנו מדד אחד שמציג כמה השכרות היו לחנות בחודש הנוכחי.

Sales Percentages – שני מדדים שמראים את אחוז המכירות של כל קטגוריה (Sales %) ואחוז מחירות של כל סרט לבדו מתוך הקטגוריה (Film Title Sales %).

Totals – כל מה שקשור לסכמה של השכרות, איחורים של החזרה ותשלומים.

## שלב 3: ויזואליזציה

### Main Dashboard

זהו הדשבורד הראשי בו מוצג כלל הנותנים על העסק כגון השכרות, רווח, השוואות אל מול שנים שעברו.

#### KPIS

בחרנו להשתמש בKPIS של השוואות מול שנה אחת אחורה בגלל ההפסקה של חודשים פברואר 2021 עד יוני 2021. ראינו שבמידע אל מול שנה שעברה ניתן לקבל מסקנות טובות יותר. את הוויזואליזציה של הKPI **לא ניתן** לפלטר באמצעות הקטגוריות.

#### Waterfall

בתרשים זה הראנו את סכום ההכנסות אל מול כל קטגוריה בנפרד. ולאחר ניסוים רבים של סוגי תאריכים החלטנו להישאר על year\_begin\_date, למרות שmonth\_begin\_date מראה יותר נתונים. הסיבה לכך היא שהKPI שלנו הינו על בסיס שנתי וזה אמור להיות עמוד שנותן מידע כללי על העסק לטווח הארוך, הדשבורד השני נותן יותר מקוון על קטגוריות ספציפיות.

#### Matrix

הנתונים שהחלטנו להציג הינם:

* הכנסות של כל סרט בהם שמנו פורמט של "בר" הנותן אינדיקציה לכמה כל סרט רווחי. בנוסף ישנם אייקונים הנותנים אינדיקציה של סרטים שלא הושכרו כלל.
* אחוז מכלל ההכנסות המתחלק לשני חלקים
  + אחוז ההכנסות של כל קטגוריה אל מול כלל הקטגוריות (Sales %)
  + אחוז ההכנסות של כל סרט אל מול הקטגוריה שלו (Film Title Sales %)
* כמות ההשכרות של אותו מוצר
* כמות ההשכרות שהוחזרו **באיחור**, יכול לתת אינדיקציה לבעל העסק האם כדי לו להעלות את זמן ההשכרה של אותו הסרט. על מנת להקל על בעל העסק דאגנו להוסיף אייקונים, האייקונים מתחילים להופיע כאשר יש יותר מ5 איחורים על מנת שישים לב שיש יחסית הרבה איחורים שמנו אייקון של עיגול צהוב ומעל ל10 איחורים יש אייקון של עיגול אדום שנותן הדגשה שישנם הרבה איחורים בהחזרות של סרט זה.
* כמות הכסף שלא שולם במלואו בגלל איבוד סרטים. במהלך ניתוח הנתונים שמנו לב שבטבלת payment, לקוחות שלא החזירו סרטים שילמו רק את מחיר ההשכרה הרגילה ולא את מחיר ההחלפה של הסרט. כדי להראות את זה ללקוח על מנת שיוכל לחקור את הנושא ולדבר עם הלקוחות לשלם את מחיר ההחלפה של הסרט. ישנו פורמט אשר צובע את כל התאים שהערך שלהם גדול מ0 בצבע אדום.

#### Map

במפה ניתן לראות מה סכום ההכנסות מהשכרות בכל מדינה לפי עיגולים, ככל שהעיגול גדול יותר כך יש יותר הכנסות. המפה עובדת באופן יעיל יותר כאשר מפעילים פילטר של קטגוריות.

### Film Dashboard

הדשבורד הזה מיועד בעיקר לDRILL THROUGH מהדשבורד הראשי, הוא מציג את הקטגוריה שלה עשינו drill through והפילטר של הזמן הינו מסוכרן עם הMain

הוספנו פילטר אשר נותן לפצל את התוצאות לכל אחת מהחנויות בנפרד או לראות את שניהם ביחד.

#### Line Charts

הגרפים הקוויים מראי את התשואה החודשית של העסק וכמות ההשכרות החודשית. הם מחושבים על ידי Total Rentals ו-Revenue ופרוסים על ידי month\_begin\_date.

#### Gauge

מראה את הרווח החודשי והמטרה שלו הינו הרווח של חודש שעבר. כאשר נשים את העכבר על השטח של ההכנסה נוכל לראות את 8 הסרטים הרווחים ביותר. בעזרת Tooltip

#### Map

במפה ניתן לראות מה סכום ההכנסות מהשכרות בכל מדינה לפי עיגולים, ככל שהעיגול גדול יותר כך יש יותר הכנסות. המפה בדשבורד הנוכחי מציגה רק לגבי קטגוריה מסוימת.

## הסקת מסקנות ודברים שלמדנו מהפרויקט

אלו כמה מסקנות שהגענו אליהם לאחר שסיימנו את הפרויקט והבנו שכדאי לעשות לפעמים הבאות.

* כדי להשתמש בTransform data בשביל לרכז את כל המידע בטבלאות ראשיות ולנקות/לסדר את המידע שלנו. בTable View עדיף להוסיף את כלל העמודות המחושבות.
  + במהלך העבודה עשינו מיקס בין הוספת עמודות ושינוי המידע כל פעם במקום אחר. בעזרת שאלות לCHAT GPT הבאנו כי זה לא הדבר הנכון לעשות, אלה כדי לתת הגדרה לכל אחת מהאופציות.
* כדי להשתמש במשתנים בתוך המדדים במקום ליצור מדדים חדשים כאשר אין שימוש חוזר בהם על מנת לשמור על סביבה עבודה נקייה ומאורגנת יותר.
* כדי להפריד את הזמנים של הMain Dashboard והSecondary Dashboard מכיוון שכל אחד מהם מיועד לרמות שונות של מנהלים. הMain יותר מיועד למנהלים הכי גבוהים שצריכים מבט גדול יותר על העסק וניתן יהיה להשתמש בSecondary על מנת להיכנס לפרטים קטנים יותר.