

שאלות SQL - מטלת בית 3

1. האם קיים קשר בין אזור בשיקגו לכמות וסוג הפשעים ולאחוז העצורים?

KPI 1 - יותר מ-50% מהפשעים בשיקגו קרו באותו אזור

```
SELECT AreasCodesandNames.Name,  
COUNT(SampleDataAfterSTTM1000Rows.ID) OVER (PARTITION BY  
AreasCodesandNames.ID) AS TotCrimes  
FROM AreasCodesandNames INNER JOIN SampleDataAfterSTTM1000Rows  
ON AreasCodesandNames.ID = SampleDataAfterSTTM1000Rows.CommunityArea  
ORDER BY TotCrimes DESC
```

שאלתה זו תחזיר כפלט את האיזור בעיר וכמות הפשעים שבוצעו באותו האיזור ותאפשר לנו להציג ולענות על ה-KPI.

KPI 2 - יותר מ-50% מהפשעים באותו אזור בשיקגו הם מאותו הסוג

```
SELECT AreasCodesandNames.Name,  
SampleDataAfterSTTM1000Rows.PrimaryType,  
COUNT(SampleDataAfterSTTM1000Rows.ID) OVER (PARTITION BY  
AreasCodesandNames.ID, SampleDataAfterSTTM1000Rows.PrimaryType) AS  
TotCrimes  
FROM AreasCodesandNames INNER JOIN SampleDataAfterSTTM1000Rows  
ON AreasCodesandNames.ID = SampleDataAfterSTTM1000Rows.CommunityArea  
ORDER BY TotCrimes DESC
```

שאלתה זו תחזיר כפלט את האיזור בעיר, לצד סוג הפשע וכמות הפשעים שבוצעו באותו האיזור ותאפשר לנו להציג ולענות על ה-KPI בכך שנוכל לסווג את סוג הפשעים והכמות בכל איזור.

KPI 3 - אחוז העצורים מסך הפשעים באזור הפשיעה הנפוץ במדינה גדול מאחוז העצורים מסך הפשעים בשאר האזורים

```
SELECT S1.Name, COUNT(SampleDataAfterSTTM1000Rows.ID) AS TotArrested ,  
S1.TotCrimes  
FROM SampleDataAfterSTTM1000Rows INNER JOIN  
(SELECT AreasCodesandNames.ID AS ID, AreasCodesandNames.Name AS
```

```

Name, COUNT(SampleDataAfterSTTM1000Rows.ID) OVER (PARTITION BY
AreasCodesandNames.ID) AS TotCrimes
FROM AreasCodesandNames INNER JOIN
SampleDataAfterSTTM1000Rows
ON AreasCodesandNames.ID =
SampleDataAfterSTTM1000Rows.CommunityArea
ORDER BY TotCrimes DESC
LIMIT 1 ) AS S1
ON SampleDataAfterSTTM1000Rows.CommunityArea = S1.ID AND
SampleDataAfterSTTM1000Rows.Arrest = TRUE

```

שאלתה זו היא שאלתה מקוננת. פלט השאלתה הפנימית יהיה שורה אחת בלבד, בגלל פונקצית ה-LIMIT, אשר יציג את שם האיזור, מספר האיזור הייחודי ומס' הפשעים אשר בוצעו באיזור עם הכי הרבה פשעים במדינה. לאחר מכן, בשאלתה החיצונית, נציג כפלט אך ורק את האיזור עם כמות הפשעים הרבה ביותר וליצידו נציג את כמות הפשעים והעצורים באותו האיזור. פלט זה יאפשר לנו לענות על KPI זה.

```

SELECT D1.Name, COUNT(F1.ID) OVER (PARTITION BY F1.CommunityArea) AS
TotArrested
FROM SampleDataAfterSTTM1000Rows AS F1, (SELECT
SampleDataAfterSTTM1000Rows.communityarea AS ID
, COUNT(SampleDataAfterSTTM1000Rows.ID) OVER (PARTITION BY
SampleDataAfterSTTM1000Rows.communityarea) AS TotCrimes
FROM
SampleDataAfterSTTM1000Rows
ORDER BY
TotCrimes DESC
LIMIT 1) AS S1,
AreasCodesandNames AS D1
WHERE F1.communityarea <> S1.ID AND F1.Arrest = TRUE AND D1.ID =
F1.communityarea
ORDER BY TotArrested DESC

```

שאלתה מקוננת זאת, תחזיר את שאר האיזורים ללא האיזור עם מס' הפשעים הגדול ביותר ואת מס' הפשעים והעצורים בכל איזור ובכך נוכל לענות על KPI זה.

2. האם יש קשר בין מגפת הקורונה(שנת 2020- שיא המגפה) לכמות וסוג הפשעים?
הגדרה: תקופת הקורונה – שנת 2020, אינטרוול זמן זהה בתקופה סטנדרטית – שנת 2022

KPI 1 - כמות פשעים ממוצעת בתקופת הקורונה גדולה ביותר מ-30% מכמות הפשעים באינטרוול זמן זהה בתקופה סטנדרטית

```
SELECT F1.Year, COUNT(F1.ID) OVER (PARTITION BY F1.year) AS Number_Of_Crimes  
FROM SampleDataAfterSTTM1000Rows AS F1;
```

שאלתה זו תציג עבור כל שנה את מספר הפשעים שקרו בה, ע"י כך נראה האם ישנה עליה/ירידה בכמות פשעים הממוצעת בתקופת הקורונה אל מול כמות הפשעים באינטרוול זמן זהה בתקופה סטנדרטית.

KPI 2 - יחס גניבות בשווי גבוה מ-\$500 בתקופת הקורונה גדול יותר מאותו היחס באינטרוול זמן זהה בתקופה סטנדרטית

```
SELECT F1.Year, COUNT(F1.ID) OVER (PARTITION BY F1.year) AS Crimes_Over_500$  
FROM SampleDataAfterSTTM1000Rows AS F1  
WHERE F1.Description = "Over $500";
```

בדומה לשאלתה הקודמת, שאלתה זו תציג את מספר הפשעים אשר סווגו כ"מעל \$500" לפי כל שנה. על ידי כך נוכל להבין ולהציג את היחס ביניהם.

KPI 3 - עלייה של 50% במספר פשעים במשפחה בזמן מגפת הקורונה לעומת מספר פשעים במשפחה באינטרוול זמן זהה בתקופה סטנדרטית

```
SELECT F1.Year, COUNT(F1.ID) OVER (PARTITION BY F1.year) AS Domestic_Crimes  
FROM SampleDataAfterSTTM1000Rows AS F1  
WHERE F1.Domestic = "True";
```

שאלתה זו תציג את מספר הפשעים אשר סווגו כפשעים בתוך המשפחה (על ידי עמודת Domestic בה הערכים הם בוליאנים) בחלוקה לפי שנים. על ידי כך נוכל להבין ולסיק האם ישנה עליה/ירידה במספר הפשעים במשפחה בזמן מגפה אל מול אינטרוול זמן זהה בתקופה סטנדרטית.